



Internationale Standardlehrbücher der Wirtschafts- und Sozialwissenschaften

Herausgegeben von Universitätsprofessor Dr. Lutz Kruschwitz

Bisher erschienene Werke:

Bagozzi u. a., Marketing Management
Bergstrom · Varian, Trainingsbuch zu
Varian, Grundzüge der
Mikroökonomik, 3. A.
Büning · Naeve · Trenkler · Waldmann,
Mathematik für Ökonomen im
Hauptstudium
Dixit · Norman, Außenhandelstheorie,
4. A.
Dornbusch · Fischer, Makroökonomik,
6. A.
Ethier, Moderne Außenwirtschafts-
theorie, 4. A.
Gordon, Makroökonomik, 4. A.
Granvogl · Perridon, Sozioökonomie
Heike · Târcolea, Grundlagen der
Statistik und
Wahrscheinlichkeitsrechnung
Hillier · Lieberman, Einführung in
Operations Research, 5. A.
Kneis, Mathematik für Wirtschafts-
wissenschaftler
Kruschwitz, Finanzierung und
Investition, 2. A.
Kruschwitz, Investitionsrechnung, 8. A.

Mehler-Bicher, Mathematik für
Wirtschaftswissenschaftler
Meissner, Strategisches Internationales
Marketing, 2. A.
Pindyck · Rubinfeld, Mikroökonomie,
4. A.
Sargent, Mikroökonomik
Schäfer · Kruschwitz · Schwake,
Studienbuch Finanzierung und
Investition, 2. A.
Sloman, Mikroökonomie, 3. A.
Smith, Einführung in die
Volkswirtschaftslehre, 2. A.
Stiglitz, Volkswirtschaftslehre, 2. A.
Stiglitz · Schönfelder,
Finanzwissenschaft, 2. A.
Varian, Grundzüge der Mikroökonomik,
4. A.
Zäpfel, Strategisches Produktions-
Management, 2. A.
Zäpfel, Taktisches Produktions-Mana-
gement, 2. A.
Zwer, Internationale Wirtschafts- und
Sozialstatistik, 2. A.

Mathematik für Ökonomen im Hauptstudium

Von

Universitätsprofessor Dr. Herbert Büning

Universitätsprofessor Dr. Peter Naeve

Universitätsprofessor Dr. Götz Trenkler

Universitätsprofessor Dr. Karl-Heinz Waldmann

R. Oldenbourg Verlag München Wien

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Mathematik für Ökonomen im Hauptstudium / von Herbert Büning -
München ; Wien : Oldenbourg, 2000
(Internationale Standardlehrbücher der Wirtschafts- und
Sozialwissenschaften)
ISBN 3-486-20986-8

© 2000 Oldenbourg Wissenschaftsverlag GmbH
Rosenheimer Straße 145, D-81671 München
Telefon: (089) 45051-0
www.oldenbourg-verlag.de

Das Werk einschließlich aller Abbildungen ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Bearbeitung in elektronischen Systemen.

Gedruckt auf säure- und chlorfreiem Papier
Gesamtherstellung: Druckhaus „Thomas Müntzer“ GmbH, Bad Langensalza

ISBN 3-486-20986-8

Vorwort

Mathematik löst bei Studienanfängern der Wirtschaftswissenschaft erfahrungsgemäß keine Begeisterungstürme aus. Das mag zwei Gründe haben. Zum einen, weil diejenigen, die sich für das Studienfach Wirtschaftswissenschaft entschieden haben, in der Tendenz große Lücken schon in der Elementarmathematik aufweisen (was Eingangstests immer wieder belegen) und zum anderen, weil für viele der Studienanfänger nicht einsichtig wird, warum sie sich durch die Mathematik „quälen“ sollen, wird doch — was sich schnell herumspricht — im Verlauf des Studiums bei nicht wenigen Dozenten die Mathematik mehr oder weniger ausgeklammert. So dient die Auseinandersetzung mit der „Linearen Algebra“ und der „Analysis“ als mittlerweile obligatorische Lehrveranstaltungen einer jeden wirtschaftswissenschaftlichen Grundausbildung oft nur zum notwendigen Scheinerwerb.

Das mathematische Pflichtprogramm im wirtschaftswissenschaftlichen Grundstudium hat zu einer kaum noch überschaubaren Menge von Lehrbüchern geführt, die an den verschiedenen wirtschaftswissenschaftlichen Fakultäten eingesetzt werden. Dabei ist der Spielraum der zu behandelnden mathematischen Themen im Grundstudium nicht so groß, um die Fülle solcher Lehrbücher zu rechtfertigen. Der zuständige Dozent für die Mathematik im Grundstudium bedient sich aus nachvollziehbaren Gründen meist seines eigenen Buches oder Skriptes.

Die im Grundstudium erworbenen mathematischen Kenntnisse reichen im wirtschaftswissenschaftlichen Hauptstudium der Betriebs- und Volkswirtschaftslehre mit einer stärkeren quantitativen Ausrichtung bei weitem nicht aus, so z.B. nicht in der Statistik und Wahrscheinlichkeitsrechnung, in der Ökonometrie, im Operations Research und insbesondere bei der Optimierung oder bei der Analyse dynamischer Systeme in der Mikro- und Makroökonomie. Hier ist ein nennenswertes „Mehr“ an Mathematik gefordert. Von Studierenden der Wirtschaftswissenschaft wird in steigendem Maße eine bessere mathematische Ausbildung verlangt, um die komplexen Strukturen ökonomischer Prozesse adäquat beschreiben und analysieren zu können. Doch für ein Hauptstudium mit den oben beispielhaft genannten Schwerpunkten ist im deutschsprachigen Raum das Angebot entsprechender mathematisch-ökonomischer Literatur sehr gering, um nicht zu sagen eine „Fehlanzeige“.

Dieses „Vakuum“ soll das vorliegende Buch helfen zu beseitigen. Gewisse Grundkenntnisse auf dem Gebiet der Differential- und Integralrechnung sowie der linearen Algebra werden vorausgesetzt, eine Reihe diesbezüglicher Begriffe und ihrer Eigenschaften sind zur Erinnerung im 1. Kapitel zusammengestellt. Wir sind uns wohl bewußt, daß wir mit diesem Buch nicht alle Studierenden der Wirtschaftswissenschaft ansprechen werden (und auch nicht wollen), sondern nur diejenigen, die bereit sind, sich die zum Verständnis ökonomischer Theorien notwendigen mathematischen Methoden in einem quantitativ orientierten Hauptstudium der Betriebs- und Volkswirtschaftslehre anzueignen. Dazu zählen insbesondere Studierende mit den Vertiefungsfächern Statistik, Ökonometrie, Operations Research oder mathematische Wirtschaftstheorie im Bereich der Wirtschaftswissenschaft oder Studierende im Hauptfach Statistik. Darüberhinaus sei das Buch auch denjenigen empfohlen, die sich in ihrem ökonomischen Arbeitsfeld außerhalb der Universität mit mathematischen Methoden auseinandersetzen haben.

Natürlich kann ein solches Lehrbuch nicht alle Gebiete einer mathematisch orientierten Ökonomie gleichermaßen abdecken. Und das nicht nur wegen des limitierten Umfangs, sondern auch, weil die Wirtschaftswissenschaft wie jede andere Wissenschaft „ständig im Fluß“ ist. So werden immer wieder neue wichtige Gebiete hinzukommen, ohne daß dabei klassische Gebiete an Bedeutung verlieren. Als ein solches neues Gebiet sei hier nur die Chaostheorie genannt, die in den letzten Jahren auch in die Ökonomie verstärkt Einzug gehalten hat und zu der mittlerweile eine umfangreiche eigenständige Literatur vorliegt. Auf die Chaostheorie werden wir deshalb hier nicht näher eingehen, sondern verweisen bei der Behandlung nichtlinearer dynamischer Systeme nur auf entsprechende Literaturstellen. Einzelne Kapitel dieses Buches basieren auf Manuskripten der Autoren zu Lehrveranstaltungen im wirtschaftswissenschaftlichen Hauptstudium an ihren Universitäten. Unser generelles Bemühen war es, die mathematischen Themenkreise problem- und anwendungsorientiert darzustellen, was in zahlreichen Beispielen aus den verschiedensten Anwendungsbereichen der Betriebs- und Volkswirtschaftslehre zum Ausdruck kommt. Auf Beweise von Sätzen kann ein Mathematiker nicht gänzlich verzichten, dafür beleuchtet ein Beweis auf zu eindrucksvolle Weise den Inhalt einer mathematischen Aussage mit allen ihren Prämissen. Aber es bestand für uns keine Verpflichtung, „alles“ zu beweisen, um somit mehr Raum für ökonomische Modelle, ihrer Beschreibung und Analyse, zu haben.

Ein gutes Lehrbuch sollte auch auf Aufgaben zu den behandelten Themen nicht verzichten, um dem Lernenden die Überprüfung seines Kenntnisstandes zu ermöglichen. So finden sich am Ende eines jeden Kapitels eine Reihe von Aufgaben, zu solchen mit ungerader Nummer sind Lösungen im Anhang des Buches zusammengestellt. Definitionen, Sätze, Beispiele, Abbildungen und Tabellen sind kapitelweise durchnummeriert, wobei die erste Ziffer die Kapitelnummer angibt. Das Ende eines Beweises ist mit \square gekennzeichnet, das Ende eines Beispiels mit \circ .

Selbst bei einer Beteiligung von vier Autoren hängt das Gelingen eines Buches von der Hilfe weiterer Personen ab, sei es bei Fragen des auszuwählenden Stoffes und seiner Darstellung, sei es bei der Erstellung der druckfertigen Vorlage und beim Korrekturlesen. Hier sind wir allen voran Herrn Priv. Doz. Dr. Dietrich Trenkler zu ganz besonderem Dank verpflichtet. Herr Trenkler hat nicht nur die druckfertige Vorlage mit allen Grafiken und Tabellen erstellt, sondern weit darüberhinaus in Formulierung und Darstellung „eingegriffen“, was der Lesbarkeit des Buches zweifelsohne zugute gekommen ist. Unser Dank gilt auch Frau Karin Wüstenbecker, einer unermüdlichen Korrekturleserin, sowie Herrn Martin Weigert vom Oldenbourg Verlag für die gute Zusammenarbeit.

Berlin, Bielefeld,
Dortmund und Karlsruhe

H. Büning, P. Naeve,
G. Trenkler und K.-H. Waldmann

Inhaltsverzeichnis

Einleitung	1
1 Mathematische Grundlagen	5
1.1 Vorbemerkung	5
1.2 Komplexe Zahlen	5
1.3 Matrizen	9
1.3.1 Definitionen und Rechenregeln	9
1.3.2 Rang einer Matrix	12
1.3.3 Transponierte und Inverse einer Matrix	13
1.3.4 Spur einer Matrix	14
1.3.5 Spezielle Matrizen	15
1.3.6 Partitionierte Matrizen	16
1.3.7 Kroneckerprodukt	19
1.4 Lineare Transformationen	20
1.5 Determinanten	21
1.6 Eigenwerte und Eigenvektoren	25
1.7 Diagonalisierung von Matrizen	27
1.8 Quadratische Formen und Definitheit	30
1.9 Vektorräume	34
1.9.1 Definition und Beispiele	34
1.9.2 Unterräume	36
1.9.3 Lineare Unabhängigkeit und Basis	37

1.9.4	Konvexe Teilmengen eines Vektorraumes	39
1.10	Metrische Räume	40
1.11	Geordnete Mengen und Ungleichungen	45
1.12	Funktionen und ihre Eigenschaften	51
1.13	Aufgaben	55
2	Differential- und Integralrechnung	59
2.1	Vorbemerkungen zur Gliederung	59
2.2	Differentialrechnung	60
2.2.1	Die Grenzen des Grundstudiums	60
2.2.2	Vektorwertige Funktionen	61
2.2.3	Stetigkeit vektorwertiger Funktionen	65
2.2.4	Differenzierbarkeit vektorwertiger Funktionen	68
2.2.5	Invertierbare Funktionen	81
2.2.6	Implizite Funktionen	86
2.2.7	Extrema von Funktionen ohne Nebenbedingungen	93
2.2.8	Extrema von Funktionen mit Nebenbedingungen	96
2.3	Riemann-Integral	105
2.3.1	Einleitung	105
2.3.2	Das Riemann-Integral im \mathbb{R}	106
2.3.3	Ein Blick voraus	107
2.3.4	Das Riemann-Integral im \mathbb{R}^2	108
2.3.5	Sätze und Regeln zum Doppelintegral	111
2.3.6	Vom Doppelintegral zum zweifach iterierten Integral	113
2.3.7	Iteriertes Integral über allgemeinen Bereichen	117
2.3.8	Das Riemann-Integral im \mathbb{R}^n	120
2.3.9	Der Satz von Fubini	122
2.3.10	Variablentransformation in n -dimensionalen Integralen	123
2.3.11	Uneigentliche Integrale	132

2.4	Riemann-Stieltjes-Integral	133
2.4.1	Der Fall $\mathbb{R} \rightarrow \mathbb{R}$	133
2.4.2	Anwendungen in der Statistik	136
2.4.3	Eigenschaften des Riemann-Stieltjes-Integrals	138
2.4.4	Rechenregeln für Riemann-Stieltjes-Integrale	139
2.4.5	Das n -dimensionale Riemann-Stieltjes-Integral	140
2.5	Matrizen, Vektoren und Differentiation	140
2.5.1	Differentiation nach einem Skalar	140
2.5.2	Differentiation nach einem Vektor	143
2.5.3	Differentiation nach einer Matrix	147
2.6	Aufgaben	148
3	Verallgemeinerte inverse Matrizen	151
3.1	Vorbemerkungen	151
3.2	Motivierende Beispiele	153
3.2.1	Lineares Regressionsmodell	153
3.2.2	Verzerrte Schätzung multivariater Parameter	156
3.3	Verallgemeinerte Inverse	157
3.4	Lösung von linearen Gleichungssystemen	161
3.5	Umgang mit g -Inversen	166
3.6	Invarianz und nichtnegative Definitheit	167
3.7	Moore-Penrose-Inverse	170
3.8	Approximative Lösung von Gleichungssystemen	175
3.9	Projektionen und Projektoren	179
3.10	Numerische Methoden	187
3.10.1	Bestimmung einer g -Inversen	188
3.10.2	Bestimmung der MP-Inversen nach Dwivedi	194
3.11	Aufgaben	196

4 Nichtnegative Matrizen	201
4.1 Vorbemerkungen	201
4.2 Motivierende Beispiele	201
4.2.1 Statisches Leontief-Modell	202
4.2.2 Erweitertes Leontief-Modell	203
4.2.3 Einkommensverteilungsmodell	204
4.2.4 Dynamisches Leontief-Modell	204
4.3 Grundlegende Eigenschaften	205
4.3.1 Nichtnegativität	205
4.3.2 Matrixnormen	210
4.3.3 Konvergenz von Matrizen	213
4.3.4 Diagonaldominanz	216
4.3.5 Zerlegbarkeit	221
4.4 Frobenius-Wurzel	225
4.5 Stochastische Matrizen	233
4.6 Aufgaben	240
5 Differenzen- und Differentialgleichungen in der dynamischen Analyse	243
5.1 Vorbemerkungen	243
5.2 Einführende Beispiele	244
5.2.1 Cobweb-Modell (Spinnweb-Modell)	244
5.2.2 Multiplikator-Akzelerator-Modell von Samuelson	245
5.2.3 Angebot-Nachfrage-Modell von Evans	246
5.2.4 Cobweb-Modell mit Lagerbeständen	247
5.2.5 Harrod-Domar-Modell	247
5.3 Grundlegende Begriffe und Sätze	248
5.3.1 Definition von Differenzen- und Differentialgleichungen	248
5.3.2 Existenz und Eindeutigkeit von Lösungen	250
5.3.3 Struktur des Lösungsraumes	252
5.3.4 Stabilität und Gleichgewicht	259

5.4	Lineare Differenzgleichungen mit konstanten Koeffizienten	261
5.4.1	Lineare Differenzgleichungen 1. Ordnung	261
5.4.2	Lineare Differenzgleichungen 2. Ordnung	268
5.4.3	Lineare Differenzgleichungen n -ter Ordnung	276
5.4.4	Systeme von linearen Differenzgleichungen	281
5.4.5	Darstellung der Lösung mit Hilfe von Eigenwerten	283
5.5	Lineare Differentialgleichungen mit konstanten Koeffizienten	289
5.5.1	Lineare Differentialgleichungen 1. Ordnung	289
5.5.2	Lineare Differentialgleichungen 2. Ordnung	293
5.5.3	Lineare Differentialgleichungen n -ter Ordnung	299
5.5.4	Systeme von linearen Differentialgleichungen	303
5.6	Operatordarstellung linearer Differenzen- und Differentialgleichungen	309
5.7	Nichtlineare Differenzen- und Differentialgleichungen	313
5.7.1	Vorbemerkungen	313
5.7.2	Nichtlineare Differentialgleichungen 1. Ordnung	317
5.7.3	Stabilität und Gleichgewicht	320
5.8	Aufgaben	323
6	Lineare Optimierung	327
6.1	Einführung	327
6.2	Graphische Lösung	333
6.3	Die Standardform	334
6.4	Der Simplex Algorithmus	341
6.5	Phase I und Phase II des Simplex Algorithmus	352
6.6	Der duale Simplex Algorithmus	357
6.7	Dualität	359
6.8	Zweipersonen-Nullsummenspiele	367
6.9	Das Transportproblem	375
6.10	Parametrische Optimierung	383
6.10.1	Variation der rechten Seite	383
6.10.2	Variation der Zielfunktion	386
6.11	Aufgaben	389

7	Ganzzahlige Optimierung	397
7.1	Einführende Beispiele	397
7.2	Lösungsansätze	400
7.3	Das Branch-and-Bound-Verfahren	403
7.4	Das Schnittebenenverfahren	410
7.5	Simulated Annealing	412
7.6	Tabu Search	417
7.7	Genetische Algorithmen	420
7.8	Aufgaben	422
8	Nichtlineare Optimierung	427
8.1	Einführende Beispiele	427
8.2	Optimalitätsbedingungen	431
8.3	Konvexe Optimierungsprobleme	437
8.4	Quadratische Optimierungsprobleme	442
8.5	Gradientenverfahren	443
8.5.1	Methode des steilsten Abstiegs	444
8.5.2	Verfahren konjugierter Richtungen	445
8.5.3	Newton-Verfahren	447
8.5.4	Quasi-Newton-Verfahren	447
8.6	Verfahren zulässiger Richtungen	448
8.7	Methode der äußeren Approximation	452
8.8	Methode der Straffunktionen	453
8.8.1	Methode der äußeren Straffunktionen	453
8.8.2	Methode der inneren Straffunktionen	455
8.9	Die Komplementaritätsmethode	457
8.10	Aufgaben	461

9 Dynamische Optimierung	465
9.1 Einführung	465
9.2 Deterministische dynamische Optimierung	471
9.2.1 Das Basismodell	471
9.2.2 Das Optimalitätskriterium	473
9.2.3 Die Optimalitätsgleichung	474
9.2.4 Wertiteration	476
9.2.5 Anwendungsbereiche	478
9.3 Stochastische dynamische Optimierung	481
9.3.1 Das Basismodell	481
9.3.2 Das Optimalitätskriterium	483
9.3.3 Die Optimalitätsgleichung	486
9.3.4 Wertiteration	488
9.3.5 Ein Kontrollmodell	489
9.3.6 Optimalität strukturierter Strategien	490
9.4 Verallgemeinerungen	506
9.5 Aufgaben	506
Lösungen ausgewählter Aufgaben	511
Literatur	553
Index	559

Einleitung

Der Einsatz der Mathematik in den Wirtschaftswissenschaften hat in den vergangenen Jahrzehnten in zunehmendem Maße an Bedeutung gewonnen. Wohl kaum ein Gebiet der Ökonomie ist im Laufe der Zeit von der Mathematik „verschont“ geblieben. Es würde allerdings den Rahmen dieses Buches sprengen, wollten wir auch nur alle wesentlichen Gebiete ansprechen und das dafür notwendige mathematische Instrumentarium bereitstellen. So bleibt die Auswahl der zu behandelnden mathematischen Themen in einem solchen Lehrbuch notgedrungen stets mit einer gewissen Willkür verbunden bzw. von der Sichtweise der Autoren abhängig, welche mathematischen Methoden sie als wichtig erachten. Es gibt nicht wenige Ökonomen, die der fortschreitenden Mathematisierung der Ökonomie mit großer Skepsis begegnen. Ein Grund mag sein, daß vielen von ihnen, die nicht über das notwendige mathematische „Know-how“ verfügen, die mathematisch-ökonomische Literatur unverständlich bleibt, wie beispielsweise die Monographie des Nobelpreisträgers für Ökonomie, G. Debreu, *Theory of Value — An Axiomatic Analysis of Economic Equilibrium* (1959), um nur eines der älteren, schon stärker mathematisch ausgerichteten ökonomischen Bücher zu nennen. Zum Verständnis dieser Monographie werden fundierte Kenntnisse geordneter topologischer Räume vorausgesetzt. Der im Vorwort geäußerten Einschätzung des Autors „... *small amount of mathematics necessary for a full understanding* ...“ wird ein Mathematiker zustimmen, weniger jedoch ein anwendungsorientierter Ökonom. Unstrittig sollte jedoch sein, daß die Operationalisierung und Quantifizierung wirtschaftlicher Phänomene beim Ineinanderspiel ökonomischer Variablen zum besseren Verständnis der ökonomischen Theorie beiträgt, ermöglicht diese Operationalisierung mit Hilfe mathematischer Methoden doch eine präzisere Darstellung der wirtschaftlichen Realität. Sicher, über die Adäquatheit von Modellen zur Beschreibung empirischer Phänomene läßt sich wahrlich oft trefflich streiten; das gilt für die Wirtschaftswissenschaften wie für jede andere empirische Wissenschaft. Aber ohne den Versuch zu wagen, das zu beobachtende Phänomen durch ein Modell abzubilden, bleibt die Theorie meist vage und Mißinterpretationen werden Tür und Tor geöffnet. Dies möge der Leser als ein Plädoyer für den Erwerb eines gewissen mathematischen Rüstzeuges im Hinblick auf ein erfolgreiches Ökonomie-Studium,

das internationalen Standards gerecht wird, auffassen.

Bei der Frage nach der Auswahl mathematischer Themen haben wir uns an Fächern und Inhalten eines quantitativ ausgerichteten Hauptstudiums der Betriebs- und Volkswirtschaftslehre orientiert, als da sind: Statistik, Operations Research, Ökonometrie, Mikro- und Makroökonomie u.a. Dementsprechend sind auch die Themen der einzelnen Kapitel gewählt. Dabei geht es um Operationen mit Vektoren, Vektorräumen, Matrizen, Determinanten, das Lösen von Gleichungs- und Ungleichungssystemen mit und ohne Nebenbedingungen, die Untersuchung von Funktionen und ihren Eigenschaften und vieles mehr. Das Buch umfaßt insgesamt 9 Kapitel und einen Anhang mit Lösungen von ausgewählten Aufgaben. Die einzelnen Kapitel bauen nicht generell aufeinander auf, sondern können teilweise auch unabhängig voneinander gelesen werden, so z.B. das 3. und 4. Kapitel zur Matrizenrechnung. Zum besseren Verständnis der Kapitel 7, 8 und 9 empfiehlt sich allerdings erst das Studium des 6. Kapitels.

Zu den Kapiteln im einzelnen:

Das 1. Kapitel dient der Wiederholung mathematischer Begriffe und einer Zusammenstellung der wichtigsten Sätze aus dem Grundstudium der Mathematik für Ökonomen, wie sie in dem Buch von Wetzel u.a. (1981) zu finden sind; dazu einige Ergänzungen zu quadratischen Formen, Eigenwerten, metrischen Räumen, geordneten Mengen und Ungleichungen.

Im 2. Kapitel werden im ersten Teil Fragen der Differenzierbarkeit reell- und vektorwertiger Funktionen sowie von impliziten und inversen Funktionen mit Anwendungen bei der Untersuchung relativer Extrema behandelt. Im zweiten Teil geht es um das Riemannsche Integral von Funktionen einer und mehrerer Variablen, um uneigentliche Integrale, um das Riemann-Stieltjes-Integral, um Variablentransformation in n -dimensionalen Integralen mit Anwendungen in der Statistik (Satz von Jacobi) und im letzten Teil um die Differentiation in Matrixsymbolik.

Das 3. Kapitel beschäftigt sich mit verallgemeinerten Inversen einer beliebigen Matrix. Das Konzept der verallgemeinerten Inversen (auch g -Inverse genannt, „ g “ steht für generalized) ist ein mächtiges Werkzeug zur Lösung linearer Gleichungssysteme und findet insbesondere in der Inferenzstatistik Anwendung beim Schätzen der Parameter des linearen Regressionsmodells. Ein Spezialfall einer g -Inversen ist die Moore-Penrose-Inverse, die für jede Matrix existiert und eindeutig ist. Numerische Methoden zur Bestimmung der verallgemeinerten Inversen werden in einem gesonderten Abschnitt zur Verfügung gestellt. Das Kapitel schließt mit der Darstellung von Projektionen und Projektoren und stellt dabei ihren Zusammenhang mit den vorab betrachteten verallgemeinerten Inversen her. Schätzer und ihre Eigenschaften im linearen Regressionsmodell lassen sich mathematisch elegant mit Hilfe solcher Projektoren beschreiben.

Das 4. Kapitel ist nichtnegativen Matrizen gewidmet, die bei der Untersuchung linearer ökonomischer Modelle eine wichtige Rolle spielen. Dazu werden zu Beginn einige motivierende Beispiele sogenannter Leontief-Modelle vorgestellt. Bei der Suche nach Lösungen der entsprechenden linearen Gleichungs- und Ungleichungssysteme müssen die Komponenten des Lösungsvektors nichtnegativ sein, damit die Lösung ökonomisch sinnvoll ist. Es folgt die Untersuchung von Eigenschaften solcher nichtnegativer Matrizen in Verbindung mit Begriffen wie Matrixnorm, Diagonaldominanz und Zerlegbarkeit von Matrizen. Nichtnegative Matrizen haben stets einen reellen, nichtnegativen Eigenwert, die sogenannte Frobenius-Wurzel, der betragsmäßig mindestens so groß wie alle anderen Eigenwerte der Matrix. Das wird im vorletzten Abschnitt gezeigt. Der letzte Abschnitt handelt von speziellen nichtnegativen Matrizen, sogenannten stochastischen Matrizen, wie sie bei Einkommensverteilungsmodellen oder in der stochastischen Optimierung (siehe 9. Kapitel) Anwendung finden.

Im 5. Kapitel werden Differenzen- und Differentialgleichungen, ihre Lösungen und das Verhalten der Lösungen unter Zugrundelegung gewisser Stabilitätskriterien behandelt. Motiviert wird die Untersuchung solcher Differenzgleichungen (bei diskreter Zeit) und Differentialgleichungen (bei stetiger Zeit) mit Anwendungen in der dynamischen ökonomischen Analyse, so z.B. für Angebots- und Nachfrage-Modelle sowie für Volkseinkommen- und Konsum-Modelle. Dabei beschränken wir uns vorrangig auf lineare Differenzen- und Differentialgleichungen und speziell auf solche mit konstanten Koeffizienten, wie sie in den Anwendungen oft zu finden sind. Für lineare homogene und inhomogene Differenzen- und Differentialgleichungen ist die Struktur der Lösung recht einfach. Nach einführenden Beispielen werden grundlegende Sätze über die Existenz und Eindeutigkeit der Lösungen und ihrer Struktur hergeleitet sowie die Begriffe Stabilität und Gleichgewicht eingeführt. Die folgenden Abschnitte sind der Untersuchung linearer Differenzgleichungen bzw. linearer Differentialgleichungen mit konstanten Koeffizienten gewidmet sowie Systemen solcher Gleichungen. Der nächste Abschnitt dient einer einheitlichen Darstellung linearer Differenzen- und Differentialgleichungen mit Hilfe linearer Operatoren. Am Ende des Kapitels wird noch ein Ausblick auf nichtlineare Differenzen- und Differentialgleichungen gegeben.

Das 6. Kapitel beschäftigt sich mit der linearen Optimierung. Ziel ist es, eine lineare Funktion unter Nebenbedingungen, die in Form von linearen Ungleichungen vorliegen, zu maximieren oder minimieren. Ausgangspunkt der Untersuchungen bildet eine Reihe von Beispielen, die das breite Spektrum potentieller Anwendungen aufzeigen. Diese führen auf eine Standardform, die die Grundlage der algorithmischen (Simplex-Algorithmus) und theoretischen Behandlung (Dualitätssatz, Satz vom komplementären Schlupf) bildet. Neben allgemeinen linearen Optimierungsproblemen werden strukturierte lineare Optimierungsprobleme wie Zweipersonen-Nullsummenspiele oder Transportprobleme ausführlich behandelt. Eine ökonomische Interpretation

der Ergebnisse sowie eine Sensitivitätsanalyse (Parametrische Optimierung) schließt das Kapitel ab.

Das 7. Kapitel hat die ganzzahlige Optimierung zum Gegenstand, bei der die Entscheidungsvariablen der linearen Optimierung auf ganzzahlige Werte oder sogar nur auf die Werte 1 oder 0 (im Falle von ja-nein Entscheidungen) eingeschränkt werden. Verbunden mit der Einschränkung ist ein erheblicher zusätzlicher Rechenaufwand, dem durch moderne Heuristiken (Simulated Annealing, Tabu Search, Genetische Algorithmen), die die klassischen exakten Verfahren (Branch and Bound Verfahren, Schnittebenenverfahren) vervollständigend, Rechnung getragen wird.

Im 8. Kapitel wird die nichtlineare Optimierung behandelt. Hierzu wird die Annahme einer linearen Zielfunktion und/oder linearer Nebenbedingungen fallengelassen. Dadurch wird eine größere Realitätsnähe erreicht, die jedoch durch mathematisch anspruchsvollere und numerisch aufwendigere Lösungsverfahren erkaufte wird. Einen wichtigen Spezialfall bilden konvexe Optimierungsprobleme, bei denen eine konvexe Funktion über einer konvexen Menge minimiert wird. Die theoretische (Karush-Kuhn-Tucker Theorem) und praktische Lösung konvexer Optimierungsprobleme bildet einen Schwerpunkt des Kapitels. Darüber hinaus werden quadratische Optimierungsprobleme behandelt (Komplementaritätsmethode) und ausgewählte Lösungsverfahren (Gradientenverfahren, Verfahren zulässiger Richtungen, Methoden der äußeren Approximation, Strafmethode) vorgestellt.

Das 9. und letzte Kapitel beschäftigt sich mit der dynamischen Optimierung, einer rekursiven Lösungstechnik, die hauptsächlich bei mehrstufigen Entscheidungsproblemen Anwendung findet. Einführende Beispiele veranschaulichen den Lösungsansatz und zeigen das breite Spektrum der möglichen Anwendungen auf. Sie führen auf das Basismodell der deterministischen und stochastischen dynamischen Optimierung und deren Analyse mittels Optimalitätsgleichung (Bellman'sche Funktionalgleichung). Die Optimalität einfach strukturierter Entscheidungsregeln ist ein weiterer Schwerpunkt des Kapitels. Diesbezügliche Strukturaussagen werden unter geeigneten Monotonieannahmen für Stopp-, Ersetzungs-, Lagerhaltungs- und Kassenhaltungsprobleme hergeleitet.

Der Anhang bringt Lösungen zu denen am Ende eines jeden Kapitels gestellten Aufgaben, und zwar jeweils für alle Aufgaben mit ungerader Nummer.

Kapitel 1

Mathematische Grundlagen

1.1 Vorbemerkung

In diesem einführenden Kapitel wollen wir eine Reihe von Begriffen aus der „Linearen Algebra“ und „Analysis“ zusammenstellen, deren Kenntnis wichtig für das Verständnis der folgenden Kapitel ist. Einige dieser Begriffe sind bereits aus dem mathematischen Grundstudium für Wirtschaftswissenschaftler bekannt, siehe z.B. Wetzel u.a. (1981); der Geschlossenheit der Darstellung wegen werden sie hier noch einmal aufgegriffen. Die Begriffe der linearen Unabhängigkeit von Vektoren und der Basis von Unterräumen des \mathbb{R}^n werden hier als bekannt vorausgesetzt; sie sind ein Spezialfall der im Zusammenhang mit der Theorie allgemeiner Vektorräume diskutierten Begriffe in Abschnitt 1.9.

Um dieses Kapitel nicht zu stark auszudehnen, ist die Behandlung der einzelnen Themen bewusst kurz gehalten; auf die Beweise der Sätze bzw. Regeln wird nahezu gänzlich verzichtet. Die meisten, z.T. schon aus dem Grundstudium bekannten Beweise sind recht einfach und mögen dem Leser als Übung dienen. Ansonsten wird auf das Literaturverzeichnis im Anhang verwiesen.

1.2 Komplexe Zahlen

Eine komplexe Zahl z hat die Form $z = a + bi$, wobei $i := \sqrt{-1}$ die **imaginäre Einheit** ist; die Zahl $a \in \mathbb{R}$ heißt **Realteil** und $b \in \mathbb{R}$ ist der **Imaginärteil** von z . Die Menge aller komplexen Zahlen bezeichnen wir mit \mathbb{C} . Offensichtlich gilt $\mathbb{R} \subset \mathbb{C}$.

Zwei komplexe Zahlen $z_1 = a_1 + b_1i$ und $z_2 = a_2 + b_2i$ sind genau dann gleich, wenn $a_1 = a_2$ und $b_1 = b_2$ gilt. Einige grundlegende Rechenoperationen zwischen komplexen Zahlen werden eingeführt in der

Definition 1.1

Gegeben seien die Zahlen $a_1, b_1, a_2, b_2 \in \mathbb{R}$.

(i) **Addition/Subtraktion:**

$$(a_1 + b_1 i) \pm (a_2 + b_2 i) := (a_1 + a_2) \pm (b_1 + b_2) i \quad (1.1)$$

(ii) **Multiplikation:**

$$(a_1 + b_1 i) \cdot (a_2 + b_2 i) := (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1) i \quad (1.2)$$

(iii) **Division:** Für $a_2 \neq 0$ oder $b_2 \neq 0$ ist

$$\frac{a_1 + b_1 i}{a_2 + b_2 i} := \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + \frac{a_2 b_1 - a_1 b_2}{a_2^2 + b_2^2} i. \quad (1.3)$$

Komplexe Zahlen lassen sich als Zahlenpaare in der sogenannten **Gaußschen Zahlenebene** darstellen:

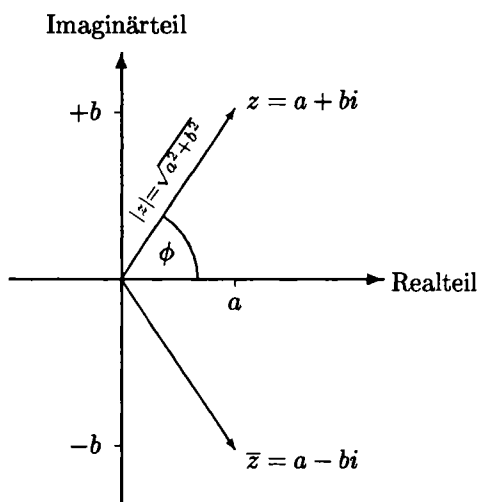


Abbildung 1.1: Darstellung von z , \bar{z} und $|z|$ in der Gaußschen Zahlenebene

Die durch Spiegelung an der Realteil-Achse erhaltene komplexe Zahl $\bar{z} := a - bi$ heißt die zu $z = a + bi$ **konjugiert komplexe Zahl**.

Hier einige Rechenregeln für konjugiert komplexe Zahlen:

Satz 1.1 *Gegeben seien die Zahlen $z, z_1, z_2 \in \mathbb{C}$. Dann gilt*

$$\overline{\overline{z}} = z \quad (1.4)$$

$$\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2} \quad (1.5)$$

$$\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2} \quad (1.6)$$

$$\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}. \quad (1.7)$$

Der **Absolutbetrag** oder **Norm** $|z|$ der komplexen Zahl $z = a + bi$ ist definiert durch $|z| := \sqrt{a^2 + b^2}$; er gibt den Abstand von z zum Nullpunkt an (siehe Abbildung 1.1).

Rechenregeln für Absolutbeträge komplexer Zahlen sind formuliert im

Satz 1.2

Gegeben seien die Zahlen $z, z_1, z_2 \in \mathbb{C}$. Dann gilt

$$|z| \geq 0 \quad (1.8)$$

$$|z|^2 = z \cdot \overline{z} \quad (1.9)$$

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2| \quad (1.10)$$

$$|z_1 + z_2| \leq |z_1| + |z_2| \quad (\text{Dreiecksungleichung}). \quad (1.11)$$

Jede komplexe Zahl $z = a + bi$ weist eine zusätzliche Darstellung auf. Danach gilt $a = |z| \cos \phi$ und $b = |z| \sin \phi$ (siehe Abbildung 1.1). Mithin gilt für z die Darstellung

$$z = |z|(\cos \phi + i \sin \phi). \quad (1.12)$$

Die Zahlen $r := |z|$ und ϕ heißen **Polarkoordinaten** von z . Das Produkt zweier komplexer Zahlen $z_1 = a_1 + b_1 i = r_1(\cos \phi_1 + i \sin \phi_1)$ und $z_2 = a_2 + b_2 i = r_2(\cos \phi_2 + i \sin \phi_2)$ ist dann in Polarkoordinaten wie folgt gegeben:

$$z_1 \cdot z_2 = r_1 r_2 (\cos(\phi_1 + \phi_2) + i \sin(\phi_1 + \phi_2)), \quad (1.13)$$

d.h., die Absolutbeträge werden multipliziert und die Winkel werden addiert.

Aus der Formel für das Produkt ergibt sich die sogenannte **Moivresche Formel** für die n -te Potenz einer komplexen Zahl:

$$z^n = r^n(\cos n\phi + i \sin n\phi). \quad (1.14)$$

Speziell für $r = 1$ ergibt sich:

$$(\cos \phi + i \sin \phi)^n = \cos n\phi + i \sin n\phi. \quad (1.15)$$

Weiterhin gilt:

$$\sqrt[n]{z} = \sqrt[n]{r} \left(\cos \frac{\phi + 2k\pi}{n} + i \sin \frac{\phi + 2k\pi}{n} \right) \quad (1.16)$$

mit $k = 0, 1, \dots, n - 1$. Der Winkel ϕ (siehe Abbildung 1.1) ist dabei im Bogenmaß ausgedrückt.

Aus den Reihendarstellungen der Funktionen e^x , $\sin x$ und $\cos x$ erhält man die sogenannte **Eulersche Formel**:

$$e^{ix} = \cos x + i \sin x. \quad (1.17)$$

Damit läßt sich jede komplexe Zahl $z = r(\cos \phi + i \sin \phi)$ in der Form $z = re^{i\phi}$ schreiben, und es ergibt sich für das Produkt zweier komplexer Zahlen $z_1 = r_1 e^{i\phi_1}$ und $z_2 = r_2 e^{i\phi_2}$ die Darstellung:

$$z_1 \cdot z_2 = r_1 r_2 e^{i(\phi_1 + \phi_2)}. \quad (1.18)$$

Ersetzen wir in $e^{ix} = \cos x + i \sin x$ das Argument x durch $-x$, so erhalten wir durch Addition bzw. Subtraktion der beiden Formeln die Beziehungen

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{und} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}, \quad (1.19)$$

d.h., die trigonometrischen Funktionen $\sin x$ und $\cos x$ mit reellem Argument x werden durch die Exponentialfunktion mit „rein imaginärem“ Argument ausgedrückt.

1.3 Matrizen

1.3.1 Definitionen und Rechenregeln

Definition 1.2

Ein geordnetes rechteckiges Schema von Zahlen der Form

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

heißt eine **Matrix der Ordnung** (m, n) , kurz (m, n) -**Matrix**. Für die Gesamtheit aller (m, n) -Matrizen $\mathbf{A} = (a_{ij})$ mit reellen oder komplexen Zahlen a_{ij} findet man gelegentlich die Schreibweise $\mathbb{R}^{m \times n}$ bzw. $\mathbb{C}^{m \times n}$.

Ist $m = n$, so heißt \mathbf{A} **quadratisch**. Zwei Matrizen $\mathbf{A} = (a_{ij})$ und $\mathbf{B} = (b_{ij})$ der Ordnung (m, n) heißen **gleich**, in Zeichen $\mathbf{A} = \mathbf{B}$, wenn $a_{ij} = b_{ij}$ für alle $i = 1, \dots, m$ und $j = 1, \dots, n$ gilt.

Spezielle Matrizen sind die **Nullmatrix** $\mathbf{0}$, deren Elemente alle gleich 0 sind, sowie die (n, n) -**Einheitsmatrix** $\mathbf{I} = (\delta_{ij})$, wobei δ_{ij} das sogenannte **Kronecker-Symbol** ist, welches durch

$$\delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad (1.20)$$

definiert ist.

Jedes Element $\mathbf{x} = (x_i)$ des \mathbb{R}^n kann als eine $(n, 1)$ -Matrix der Form

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

aufgefaßt werden. Wir sprechen dann von einem **Spaltenvektor**. In analoger Weise beschreibt die $(1, n)$ -Matrix $\mathbf{x}^\top := (x_1, x_2, \dots, x_n)$ einen **Zeilenvektor**.

Im folgenden werden wir des öfteren den j -ten Einheitsvektor $\mathbf{e}_j \in \mathbb{R}^n$ für $j = 1, 2, \dots, n$ betrachten, der durch

$$\mathbf{e}_j := \begin{pmatrix} \delta_{1j} \\ \delta_{2j} \\ \vdots \\ \delta_{nj} \end{pmatrix}, \text{ d.h. } \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

definiert ist. Es handelt sich also um einen Spaltenvektor, der aus $n - 1$ Nullen besteht, und nur an der j -ten Stelle befindet sich eine Eins.

Jede (m, n) -Matrix $\mathbf{A} = (a_{ij})$ ist so aufgebaut, daß n Spaltenvektoren $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ nebeneinander geschrieben werden, wobei die j -te Spalte gegeben ist durch

$$\mathbf{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Analog entsteht \mathbf{A} , indem m Zeilenvektoren $\mathbf{a}_{(1)}^\top, \mathbf{a}_{(2)}^\top, \dots, \mathbf{a}_{(m)}^\top$ untereinander geschrieben werden, wobei für die i -te Zeile gilt

$$\mathbf{a}_{(i)}^\top = (a_{i1}, a_{i2}, \dots, a_{in}).$$

Fassen wir diese Ergebnisse zusammen, so erhalten wir die Spalten- bzw. Zeilenschreibweise einer (m, n) -Matrix

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \begin{pmatrix} \mathbf{a}_{(1)}^\top \\ \mathbf{a}_{(2)}^\top \\ \vdots \\ \mathbf{a}_{(m)}^\top \end{pmatrix}. \quad (1.21)$$

Ein Spezialfall ist die (n, n) -Einheitsmatrix \mathbf{I} , die in der Spaltenschreibweise die Darstellung $\mathbf{I} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ besitzt, wobei \mathbf{e}_j der j -te Einheitsvektor des \mathbb{R}^n ist.

Zwischen Matrizen \mathbf{A} und \mathbf{B} können nun wie folgt Rechenoperationen definiert werden:

Definition 1.3

Es seien $\mathbf{A} = (a_{ij})$ und $\mathbf{B} = (b_{ij})$ zwei (m, n) -Matrizen. Dann ist $\mathbf{A} + \mathbf{B} := (a_{ij} + b_{ij})$.

Für die Matrizenaddition gelten folgende Regeln:

Satz 1.3

- (i) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (*Kommutatives Gesetz*)
- (ii) $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$ (*Assoziatives Gesetz*)
- (iii) $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$
- (iv) Für $\mathbf{A} = (a_{ij})$ sei $-\mathbf{A} := (-a_{ij})$, $i = 1, \dots, m$, $j = 1, \dots, n$. Dann ist

$$\mathbf{A} - \mathbf{A} = \mathbf{A} + (-\mathbf{A}) = \mathbf{0}.$$

Definition 1.4

Für eine reelle Zahl $k \in \mathbb{R}$ und eine Matrix $\mathbf{A} = (a_{ij})$ ist $k \cdot \mathbf{A} := (ka_{ij})$.

Mit Skalaren $k, k_1, k_2 \in \mathbb{R}$ und geeigneten Matrizen \mathbf{A} und \mathbf{B} gelten die folgenden Regeln für die Skalarmultiplikation:

Satz 1.4

- (i) $(k_1 + k_2) \cdot \mathbf{A} = k_1 \cdot \mathbf{A} + k_2 \cdot \mathbf{A}$
- (ii) $k \cdot (\mathbf{A} + \mathbf{B}) = k \cdot \mathbf{A} + k \cdot \mathbf{B}$
- (iii) $k_1 \cdot (k_2 \cdot \mathbf{A}) = (k_1 k_2) \cdot \mathbf{A}$.

Definition 1.5

Für eine (m, n) -Matrix \mathbf{A} und eine (n, p) -Matrix \mathbf{B} heißt die (m, p) -Matrix $\mathbf{A} \cdot \mathbf{B} := (c_{ij})$ mit

$$c_{ij} = \sum_{h=1}^n a_{ih} b_{hj}$$

für alle $i = 1, \dots, m$ und alle $j = 1, \dots, p$, das Produkt von \mathbf{A} und \mathbf{B} .

Falls Addition und Multiplikation definiert sind, gelten die folgenden Rechenregeln für die Matrizenmultiplikation:

Satz 1.5

- (i) $\mathbf{A} \cdot \mathbf{0} = \mathbf{0} \cdot \mathbf{A} = \mathbf{0}$
- (ii) $\mathbf{A} \cdot \mathbf{I} = \mathbf{I} \cdot \mathbf{A} = \mathbf{A}$
- (iii) $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C}$ (*Assoziatives Gesetz*)
- (iv) $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$ (*Distributives Gesetz*).

Hinweise:

- (a) Es gilt *nicht* generell $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$.
- (b) Aus $\mathbf{A} \cdot \mathbf{B} = \mathbf{0}$ folgt nicht $\mathbf{A} = \mathbf{0}$ oder $\mathbf{B} = \mathbf{0}$.
- (c) Statt $\underbrace{\mathbf{A} \cdot \mathbf{A} \cdot \dots \cdot \mathbf{A}}_{n \text{ Faktoren}}$ schreiben wir \mathbf{A}^n .

Wenn keine Verwechslung zu befürchten ist, werden wir im folgenden das Zeichen \cdot im Matrizenprodukt $\mathbf{A} \cdot \mathbf{B}$ fortlassen.

1.3.2 Rang einer Matrix**Definition 1.6**

Die Maximalzahl linear unabhängiger Zeilenvektoren (= Maximalzahl linear unabhängiger Spaltenvektoren) einer (m, n) -Matrix \mathbf{A} heißt **Rang** von \mathbf{A} , in Zeichen $\text{rg}(\mathbf{A})$ oder kurz $\text{rg}(\mathbf{A})$.

Offensichtlich gilt: $\text{rg}(\mathbf{A}) \leq \min\{m, n\}$. Ist $\text{rg}(\mathbf{A}) = m$ bzw. $\text{rg}(\mathbf{A}) = n$, so heißt \mathbf{A} **zeilen-** bzw. **spaltenregulär**. Ist $\text{rg}(\mathbf{A}) = n$ für eine quadratische (n, n) -Matrix \mathbf{A} , so heißt \mathbf{A} **regulär**, anderenfalls **singulär**.

Es gelten folgende Rangaussagen:

Satz 1.6

- (i) $\text{rg}(\mathbf{A} + \mathbf{B}) \leq \text{rg}(\mathbf{A}) + \text{rg}(\mathbf{B})$
- (ii) $\text{rg}(\mathbf{A}\mathbf{B}) \leq \min\{\text{rg}(\mathbf{A}), \text{rg}(\mathbf{B})\}$
- (iii) Für reguläre Matrizen \mathbf{B} und \mathbf{C} ist $\text{rg}(\mathbf{A}) = \text{rg}(\mathbf{A}\mathbf{B}) = \text{rg}(\mathbf{C}\mathbf{A})$.

1.3.3 Transponierte und Inverse einer Matrix

Definition 1.7

Es sei $A = (a_{ij})$ eine (m, n) -Matrix. Dann heißt die (n, m) -Matrix $A^\top := (a_{ji})$ die zu A **transponierte Matrix**.

Für die transponierte Matrix gelten die folgenden Regeln:

Satz 1.7

Gegeben seien die (m, n) -Matrizen A und B sowie die (n, p) -Matrix C . Dann ist

$$(i) \quad (A^\top)^\top = A$$

$$(ii) \quad (A + B)^\top = A^\top + B^\top$$

$$(iii) \quad (AC)^\top = C^\top A^\top$$

$$(iv) \quad \text{rg}(A^\top) = \text{rg}(A)$$

$$(v) \quad \text{rg}(A^\top A) = \text{rg}(AA^\top) = \text{rg}(A).$$

Sind $\mathbf{a} = (a_1, \dots, a_m)^\top$ und $\mathbf{b} = (b_1, \dots, b_n)^\top$ Vektoren, so heißt die (m, n) -Matrix $\mathbf{a}\mathbf{b}^\top$ **dyadisches Produkt** von \mathbf{a} und \mathbf{b} . Ist speziell $m = n$, so heißt die reelle Zahl $\mathbf{a}^\top \mathbf{b} = \sum_{i=1}^n a_i b_i$ **inneres Produkt** von \mathbf{a} und \mathbf{b} . Insbesondere heißen zwei Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ **orthogonal**, in Zeichen $\mathbf{a} \perp \mathbf{b}$, wenn für ihr inneres Produkt gilt $\mathbf{a}^\top \mathbf{b} = 0$. Jedem Vektor $\mathbf{a} \in \mathbb{R}^n$ ist seine **euklidische Norm**

$$\|\mathbf{a}\| := \sqrt{\mathbf{a}^\top \mathbf{a}} = \sqrt{\sum_{i=1}^n a_i^2} \quad (1.22)$$

zugeordnet. Er heißt **normiert**, wenn $\|\mathbf{a}\| = 1$ gilt. Schließlich heißt eine Menge $\{\mathbf{a}_1, \dots, \mathbf{a}_k\} \subset \mathbb{R}^n$ **orthonormiert**, wenn jedes \mathbf{a}_i normiert ist und die Vektoren \mathbf{a}_i und \mathbf{a}_j orthogonal sind für jedes Paar i, j mit $i \neq j$. Orthonormiertheit läßt sich kompakter in der Form $\mathbf{a}_i^\top \mathbf{a}_j = \delta_{ij}$ beschreiben, wobei δ_{ij} das Kronecker-Symbol aus (1.20) ist. Offenbar sind die Einheitsvektoren $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ orthonormiert.

Ist A eine (m, n) - und B eine (n, p) -Matrix, so sind diese in der Spalten- oder Zeilenschreibweise gegeben durch (1.21) bzw.

$$\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p) = \begin{pmatrix} \mathbf{b}_{(1)}^\top \\ \mathbf{b}_{(2)}^\top \\ \vdots \\ \mathbf{b}_{(n)}^\top \end{pmatrix}.$$

Das Matrizenprodukt kann dann mit Hilfe dyadischer bzw. innerer Produkte wie folgt geschrieben werden:

$$\mathbf{AB} = \sum_{i=1}^n \mathbf{a}_i \mathbf{b}_{(i)}^\top = \begin{pmatrix} \mathbf{a}_{(1)}^\top \mathbf{b}_1 & \dots & \mathbf{a}_{(1)}^\top \mathbf{b}_p \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{(m)}^\top \mathbf{b}_1 & \dots & \mathbf{a}_{(m)}^\top \mathbf{b}_p \end{pmatrix} \quad (1.23)$$

Definition 1.8

Eine Matrix \mathbf{X} mit der Eigenschaft $\mathbf{XA} = \mathbf{AX} = \mathbf{I}$ heißt die zu \mathbf{A} **inverse Matrix**, in Zeichen $\mathbf{X} =: \mathbf{A}^{-1}$.

Eine Matrix \mathbf{A} hat genau dann eine inverse Matrix \mathbf{A}^{-1} , wenn \mathbf{A} regulär ist. Man sagt dann auch, daß \mathbf{A} **invertierbar** ist.

Unter geeigneten Bedingungen gelten für die inverse Matrix die folgenden Regeln:

Satz 1.8

- (i) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- (ii) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$
- (iii) $(\mathbf{A}_1 \mathbf{A}_2 \cdots \mathbf{A}_{n-1} \mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1} \mathbf{A}_{n-1}^{-1} \cdots \mathbf{A}_2^{-1} \mathbf{A}_1^{-1}$
- (iv) $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$.

1.3.4 Spur einer Matrix

Definition 1.9

Für eine (n, n) -Matrix $\mathbf{A} = (a_{ij})$ heißt $\text{tr}(\mathbf{A}) := \sum_{i=1}^n a_{ii}$ **Spur** (englisch: *trace*) von \mathbf{A} .

Sind \mathbf{A} , \mathbf{B} und \mathbf{C} geeignet gewählte Matrizen, so gelten für die Spur der Matrizen die folgenden Regeln:

Satz 1.9

(i) $\operatorname{tr}(\mathbf{A}^\top) = \operatorname{tr}(\mathbf{A})$

(ii) $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B})$

(iii) $\operatorname{tr}(k\mathbf{A}) = k \operatorname{tr}(\mathbf{A}), \quad k \in \mathbb{R}$

(iv) $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$

(v) $\operatorname{tr}(\mathbf{B}^{-1}\mathbf{AB}) = \operatorname{tr}(\mathbf{A})$

(vi) $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{BCA}) = \operatorname{tr}(\mathbf{CAB})$

Die Regel (vi) gilt nur für zyklische Vertauschungen von \mathbf{A} , \mathbf{B} und \mathbf{C} ; im allgemeinen ist jedoch beispielsweise $\operatorname{tr}(\mathbf{ABC}) \neq \operatorname{tr}(\mathbf{ACB})$.

1.3.5 Spezielle Matrizen**Definition 1.10**

Eine quadratische Matrix \mathbf{A} heißt

- (i) **symmetrisch**, falls gilt $\mathbf{A}^\top = \mathbf{A}$.
- (ii) **schiefssymmetrisch**, falls gilt $\mathbf{A}^\top = -\mathbf{A}$.

Ist \mathbf{A} regulär und symmetrisch, so ist wegen $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$ auch \mathbf{A}^{-1} symmetrisch.

Definition 1.11

Eine quadratische Matrix $\mathbf{A} = (a_{ij})$ heißt

- (i) **diagonal**, falls $a_{ij} = 0$ für alle $i \neq j$ ist.
- (ii) **dreieckig**, falls entweder

$$a_{ij} = 0 \quad \text{für } i > j \text{ (obere Dreiecksmatrix) oder}$$

$$a_{ij} = 0 \quad \text{für } i < j \text{ (untere Dreiecksmatrix).}$$

Für eine Diagonalmatrix mit den Elementen $d_1 := a_{11}, \dots, d_n := a_{nn}$ schreiben wir

$$\text{diag}(d_1, \dots, d_n) := \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}.$$

So gilt beispielsweise $\mathbf{I} = \text{diag}(1, \dots, 1)$.

Definition 1.12

Eine quadratische Matrix \mathbf{A} heißt

- (i) **orthogonal**, falls gilt $\mathbf{A}^\top \mathbf{A} = \mathbf{I} = \mathbf{A} \mathbf{A}^\top$ (d.h. $\mathbf{A}^{-1} = \mathbf{A}^\top$).
- (ii) **idempotent**, falls gilt $\mathbf{A}^2 = \mathbf{A}$.

Offenbar ist eine Matrix $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ genau dann orthogonal, wenn alle Spaltenvektoren \mathbf{a}_j orthonormiert sind.

Für orthogonale und idempotente Matrizen gelten folgende Regeln:

Satz 1.10

- (i) Sind \mathbf{A} und \mathbf{B} orthogonal, so auch \mathbf{AB} und \mathbf{BA} .
- (ii) Ist \mathbf{A} orthogonal, so auch \mathbf{A}^{-1} und \mathbf{A}^\top .
- (iii) Ist \mathbf{A} idempotent, so auch \mathbf{A}^\top und $\mathbf{I} - \mathbf{A}$ (generell jedoch nicht $\mathbf{A} - \mathbf{I}$).
- (iv) Ist $\mathbf{AB} = \mathbf{A}$ und $\mathbf{BA} = \mathbf{B}$, dann sind \mathbf{A} und \mathbf{B} idempotent.
- (v) Ist \mathbf{A} idempotent, so gilt $\text{rg}(\mathbf{A}) = \text{tr}(\mathbf{A})$.

1.3.6 Partitionierte Matrizen

In vielen Anwendungen ist es zweckmäßig, eine Matrix \mathbf{A} in Untermatrizen zu zerlegen (wir sagen auch: zu partitionieren), um z.B. eine bestimmte „Blockstruktur“ von \mathbf{A} deutlich zu machen.

Ist \mathbf{A} eine (m, n) -Matrix, so ist folgende Zerlegung möglich:

$$\mathbf{A} = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1s} & a_{1,s+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rs} & a_{r,s+1} & \cdots & a_{rn} \\ \hline a_{r+1,1} & \cdots & a_{r+1,s} & a_{r+1,s+1} & \cdots & a_{r+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{ms} & a_{m,s+1} & \cdots & a_{mn} \end{array} \right) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad (1.24)$$

worin \mathbf{A}_{11} eine (r, s) -, \mathbf{A}_{12} eine $(r, n - s)$ -, \mathbf{A}_{21} eine $(m - r, s)$ - und \mathbf{A}_{22} eine $(m - r, n - s)$ -Matrix ist.

Natürlich können die Untermatrizen \mathbf{A}_{11} , \mathbf{A}_{12} , \mathbf{A}_{21} und \mathbf{A}_{22} weiter in Untermatrizen zerlegt werden. Allgemein ergibt sich die Darstellung:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1q} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{p1} & \mathbf{A}_{p2} & \cdots & \mathbf{A}_{pq} \end{pmatrix}.$$

Sind \mathbf{A} , \mathbf{A}_{11} , \mathbf{A}_{22} , \dots , \mathbf{A}_{qq} quadratisch ($p = q$), und gilt $\mathbf{A}_{ij} = \mathbf{0}$ für alle $i \neq j$, so erhalten wir eine sogenannte **Blockdiagonalmatrix**:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{qq} \end{pmatrix}.$$

Für Matrizen der Form (1.24) gelten folgende Regeln, die sich unmittelbar auch auf Matrizen mit mehr als vier Untermatrizen erweitern lassen.

Satz 1.11

Gegeben seien zwei Matrizen

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \quad \text{und} \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}.$$

Dann gilt:

(i)

$$\mathbf{A}^\top = \begin{pmatrix} \mathbf{A}_{11}^\top & \mathbf{A}_{21}^\top \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22}^\top \end{pmatrix}$$

(ii)

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} \mathbf{A}_{11} + \mathbf{B}_{11} & \mathbf{A}_{12} + \mathbf{B}_{12} \\ \mathbf{A}_{21} + \mathbf{B}_{21} & \mathbf{A}_{22} + \mathbf{B}_{22} \end{pmatrix},$$

falls die Ordnungen der entsprechenden Summanden gleich sind.

(iii)

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix},$$

falls die einzelnen Produkte erklärt sind.

(iv) Falls alle auftretenden Inversen existieren, so ist

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}$$

mit

$$\begin{aligned} \mathbf{C}_{11} &:= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} \\ \mathbf{C}_{12} &:= -\mathbf{C}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{C}_{21} &:= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{C}_{11} \\ \mathbf{C}_{22} &:= \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{C}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{aligned} \quad (1.25)$$

(v)

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{qq} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{qq}^{-1} \end{pmatrix},$$

falls die jeweiligen Matrizen \mathbf{A}_{ii} regulär sind.

1.3.7 Kroneckerprodukt

Das Matrizenprodukt \mathbf{AB} setzt voraus, daß die Spaltenzahl von \mathbf{A} mit der Zeilenzahl von \mathbf{B} übereinstimmt. Das in der folgenden Definition angegebene Kroneckerprodukt kann für beliebige Matrizen \mathbf{A} und \mathbf{B} gebildet werden.

Definition 1.13

Gegeben sei die (m, n) -Matrix $\mathbf{A} = (a_{ij})$ und die (p, q) -Matrix $\mathbf{B} = (b_{kl})$. Dann bezeichnet die (mp, nq) -Blockmatrix

$$\mathbf{A} \otimes \mathbf{B} := \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix}$$

das **Kroneckerprodukt** oder **direktes Produkt** von \mathbf{A} und \mathbf{B} .

Für dieses Produkt gelten folgende Regeln, einige im Unterschied zum Matrizenprodukt \mathbf{AB} .

Satz 1.12

- (i) $k(\mathbf{A} \otimes \mathbf{B}) = (k\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (k\mathbf{B}), \quad k \in \mathbb{R}$
- (ii) $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = (\mathbf{A} \otimes \mathbf{C}) + (\mathbf{B} \otimes \mathbf{C})$
- (iii) $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$
- (iv) $(\mathbf{A} \otimes \mathbf{B}) \otimes (\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$
- (v) $(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top$
- (vi) *Sind \mathbf{A} und \mathbf{B} regulär, so ist auch $\mathbf{A} \otimes \mathbf{B}$ regulär. Die Inverse ist gegeben durch $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$*
- (vii) $\text{rg}(\mathbf{A} \otimes \mathbf{B}) = \text{rg}(\mathbf{A}) \text{rg}(\mathbf{B})$
- (viii) $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A}) \text{tr}(\mathbf{B})$
- (ix) $\mathbf{A} \otimes \mathbf{B}$ ist orthogonal, falls \mathbf{A} und \mathbf{B} orthogonal sind.

1.4 Lineare Transformationen

Die im folgenden zu diskutierenden linearen Transformationen und ihre Verknüpfungen stehen im engen Zusammenhang mit Matrizen und ihren Rechenoperationen.

Definition 1.14

Eine Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt **lineare Transformation**, wenn für alle $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ und $k_1, k_2 \in \mathbb{R}$ gilt

$$T(k_1\mathbf{x}_1 + k_2\mathbf{x}_2) = k_1T(\mathbf{x}_1) + k_2T(\mathbf{x}_2).$$

Offensichtlich ist die Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$, linear, wobei \mathbf{A} eine (m, n) -Matrix ist. Es läßt sich zeigen, daß auch umgekehrt jede beliebige lineare Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch eine (m, n) -Matrix \mathbf{A} beschrieben werden kann, so daß für jedes $\mathbf{x} \in \mathbb{R}^n$ gilt: $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Eine **affine Transformation**, also eine Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$ mit $\mathbf{b} \neq \mathbf{0}$, hingegen ist *keine* lineare Abbildung im Sinne obiger Definition, was leicht nachzuprüfen ist.

Addition und Multiplikation linearer Transformationen T_1 und T_2 können durch die entsprechenden Matrizenoperationen beschrieben werden.

Addition: Gegeben seien zwei lineare Transformationen $T_1, T_2 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $T_1(\mathbf{x}) = \mathbf{A}\mathbf{x}$ und $T_2(\mathbf{x}) = \mathbf{B}\mathbf{x}$, worin \mathbf{A} und \mathbf{B} (m, n) -Matrizen sind. Dann ist

$$\begin{aligned} T_1 + T_2 : \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ \mathbf{x} &\mapsto (\mathbf{A} + \mathbf{B})\mathbf{x} \end{aligned}$$

eine lineare Transformation.

Multiplikation: Gegeben seien lineare Transformationen $T_1 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $T_2 : \mathbb{R}^m \rightarrow \mathbb{R}^k$ mit $T_1(\mathbf{x}) = \mathbf{A}\mathbf{x}$ und $T_2(\mathbf{y}) = \mathbf{B}\mathbf{y}$, worin \mathbf{A} bzw. \mathbf{B} eine (m, n) - bzw. eine (k, m) -Matrix ist. Dann ist

$$\begin{aligned} T_2 \circ T_1 : \mathbb{R}^n &\rightarrow \mathbb{R}^k \\ \mathbf{x} &\mapsto T_2(T_1(\mathbf{x})) = \mathbf{B}\mathbf{A}\mathbf{x} \end{aligned}$$

eine lineare Transformation. Dabei ist $\mathbf{B}\mathbf{A}$ eine (k, n) -Matrix.

Ist die lineare Transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ bijektiv, so ist auch die Umkehrabbildung $T^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Transformation, der die Matrix \mathbf{A}^{-1} zugeordnet ist, so daß gilt $T^{-1}(\mathbf{y}) = \mathbf{A}^{-1}\mathbf{y}$. Dies setzt natürlich auch voraus, daß \mathbf{A} regulär ist.

Jeder linearen Transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{Ax}$, sind zwei wichtige Untermengen des \mathbb{R}^n bzw. des \mathbb{R}^m zugeordnet. Zum einen ist der **Kern** von T die Menge $\{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, T(\mathbf{x}) = \mathbf{0}\} \subset \mathbb{R}^n$. Weiter ist der **Bildraum** von T die Menge $T[\mathbb{R}^n] := \{\mathbf{y} \mid T(\mathbf{x}) = \mathbf{y}, \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m$, siehe auch Definition 1.44. Wegen der engen Beziehung zwischen der linearen Transformation T und der (m, n) -Matrix \mathbf{A} kann der Kern bzw. der Bildraum von T äquivalent in der Form

$$\mathcal{N}(\mathbf{A}) := \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} = \mathbf{0}\} \quad \text{bzw.} \quad \mathcal{M}(\mathbf{A}) := \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\} \quad (1.26)$$

geschrieben werden. $\mathcal{N}(\mathbf{A})$ beschreibt somit die Lösungsmenge des homogenen Gleichungssystems $\mathbf{Ax} = \mathbf{0}$. Die Menge $\mathcal{M}(\mathbf{A})$ wird auch der von \mathbf{A} **erzeugte lineare Unterraum** oder kurz als **Spaltenraum** von \mathbf{A} bezeichnet.

1.5 Determinanten

Jeder quadratischen Matrix ist in eindeutiger Weise eine reelle Zahl zugeordnet.

Definition 1.15

Die **Determinante** $|\mathbf{A}|$ einer quadratischen (n, n) -Matrix $\mathbf{A} = (a_{ij})$ ist wie folgt definiert: Für $n = 1$ ist $\mathbf{A} = (a_{11})$ und $|\mathbf{A}| := a_{11}$. Für $n \geq 2$ sei j mit $1 \leq j \leq n$ fest vorgegeben.

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

worin \mathbf{A}_{ij} diejenige $(n-1, n-1)$ -Untermatrix von \mathbf{A} ist, die man durch Streichen der i -ten Zeile und der j -ten Spalte von \mathbf{A} erhält. Statt $|\mathbf{A}|$ findet man auch die Schreibweise $\det(\mathbf{A})$.

Man kann zeigen, daß diese Entwicklung von der Wahl der Spalte \mathbf{a}_j unabhängig und somit eindeutig ist. Ferner können wir auch eine Darstellung von $|\mathbf{A}|$ als Entwicklung nach einer beliebigen Zeile $\mathbf{a}_{(i)}^T$ von \mathbf{A} angeben, welche zum selben Wert führt.

Beispiel 1.1

Wir führen die Entwicklung nach der dritten Zeile $\mathbf{a}_{(3)}^\top$ für eine $(4, 4)$ -Matrix \mathbf{A} vor.

$$\begin{aligned}
 |\mathbf{A}| &= \begin{vmatrix} 1 & 4 & 2 & -1 \\ 3 & 0 & 5 & 2 \\ -2 & 1 & 3 & 6 \\ 5 & -4 & 3 & 0 \end{vmatrix} \\
 &= (-1)^{3+1} \cdot (-2) \cdot \begin{vmatrix} 4 & 2 & -1 \\ 0 & 5 & 2 \\ -4 & 3 & 0 \end{vmatrix} + (-1)^{3+2} \cdot 1 \cdot \begin{vmatrix} 1 & 2 & -1 \\ 3 & 5 & 2 \\ 5 & 3 & 0 \end{vmatrix} \\
 &\quad + (-1)^{3+3} \cdot 3 \cdot \begin{vmatrix} 1 & 4 & -1 \\ 3 & 0 & 2 \\ 5 & -4 & 0 \end{vmatrix} + (-1)^{3+4} \cdot 6 \cdot \begin{vmatrix} 1 & 4 & 2 \\ 3 & 0 & 5 \\ 5 & -4 & 3 \end{vmatrix}.
 \end{aligned}$$

Jede der 4 dreireihigen Unterdeterminanten kann nun wieder über eine Entwicklung nach einer Zeile oder Spalte als Summe von 3 zweireihigen Unterdeterminanten dargestellt werden; letztlich sind hier $4 \cdot 3 = 12$ zweireihige Unterdeterminanten zu bestimmen.

Eine zweireihige Determinante ist aber nach Definition 1.15 einfach zu berechnen, denn es gilt

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}, \quad (1.27)$$

was leicht nachzuprüfen ist. ○

Eine solche wiederholte Entwicklung nach einer Zeile oder Spalte ist aber zumeist sehr aufwendig. Die Aussage (i) des folgenden Satzes beinhaltet ein Kriterium, wie $|\mathbf{A}|$ mit Hilfe einfacher Rechenoperationen leichter bestimmt werden kann.

Satz 1.13

Es sei \mathbf{A} eine (n, n) -Matrix.

- (i) *Die Determinante von \mathbf{A} ändert ihren Wert nicht, wenn man das c -fache, $c \in \mathbb{R}$, einer Zeile bzw. Spalte zu einer anderen Zeile bzw. Spalte addiert.*

- (ii) *Verwandelt man eine Matrix A durch Vertauschen zweier Zeilen bzw. Spalten in eine Matrix B , so gilt $|B| = -|A|$.*
- (iii) *Verwandelt man eine Matrix A durch Multiplikation einer Zeile oder Spalte mit einer reellen Zahl $k \in \mathbb{R}$ in eine Matrix B , so gilt $|B| = k|A|$.*
- (iv) *Für $k \in \mathbb{R}$ gilt $|kA| = k^n|A|$ wegen (iii).*
- (v) $|A^T| = |A|$.
- (vi) *Für die Determinanten des Produkts zweier quadratischen Matrizen A und B gilt*

$$|AB| = |A||B|.$$

Zur Bestimmung der Determinante einer beliebigen Matrix liegt es nahe, diese mit Hilfe der gemäß Satz 1.13(i) „zulässigen“ Operationen in eine Dreiecksmatrix überzuführen, deren Determinante offensichtlich das Produkt der Diagonalelemente ist und mit der Determinante der Ausgangsmatrix übereinstimmt.

Für einige spezielle Matrizen können folgende Aussagen über ihre Determinanten gemacht werden:

Satz 1.14

- (i) *A ist genau dann regulär, wenn $|A| \neq 0$ gilt. Mithin ist A genau dann singular, wenn $|A| = 0$ ist.*
- (ii) *Ist A regulär, so ist $|A^{-1}| = |A|^{-1} = 1/|A|$.*
- (iii) *Ist A orthogonal, so ist $|A| = \pm 1$.*
- (iv) *Ist A idempotent, so ist $|A| = 0$ oder 1 .*
- (v) *Ist in der partitionierten Matrix*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

die Matrix A_{11} regulär und A_{22} quadratisch, so gilt

$$|A| = |A_{11}| |A_{22} - A_{21} A_{11}^{-1} A_{12}|.$$

Ist analog A_{11} quadratisch und A_{22} regulär, so gilt

$$|A| = |A_{22}| |A_{11} - A_{12} A_{22}^{-1} A_{21}|.$$

(vi) Für eine Blockdiagonalmatrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{qq} \end{pmatrix}$$

gilt $|\mathbf{A}| = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22}| \cdot \cdots \cdot |\mathbf{A}_{qq}|$.

(vii) Für die Inverse einer regulären Matrix \mathbf{A} gilt:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix}.$$

Die aus den algebraischen Komplementen $\Delta_{ij} = (-1)^{i+j} |\mathbf{A}_{ij}|$ gebildete Matrix $\text{adj } \mathbf{A} := (\Delta_{ij})$ heißt **Adjungierte von \mathbf{A}** , worin \mathbf{A}_{ij} wie in Definition 1.15 festgelegt ist. Es ist also $|\mathbf{A}| \mathbf{A}^{-1} = (\Delta_{ij})$.

(viii) Eine quadratische Matrix \mathbf{A} besitzt genau dann den Rang r , wenn die Determinanten aller (k, k) -Untermatrizen für $k > r$ gleich 0 sind und wenn die Determinante mindestens einer (r, r) -Untermatrix ungleich 0 ist.

Im Zusammenhang mit der Einteilung quadratischer Formen in Abschnitt 1.8 und dem Nachweis relativer Extrema von Funktionen mehrerer Veränderlicher spielen die sogenannten Hauptabschnittsdeterminanten eine wichtige Rolle:

Definition 1.16

Sei $\mathbf{A} = (a_{ij})$ eine (n, n) -Matrix. Dann heißen die n Unterdeterminanten

$$D_i = \begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix}, \quad 1 \leq i \leq n,$$

Hauptabschnittsdeterminanten von \mathbf{A} . Insbesondere gilt $D_1 = a_{11}$, $D_2 = a_{11}a_{22} - a_{21}a_{12}$ und $D_n = |\mathbf{A}|$.

1.6 Eigenwerte und Eigenvektoren

Definition 1.17

Gegeben sei eine (n, n) -Matrix A . Erfüllen $\lambda \in \mathbb{C}$ und $\mathbf{x} \in \mathbb{R}^n$ mit $\mathbf{x} \neq \mathbf{0}$ das linear homogene Gleichungssystem $A\mathbf{x} = \lambda\mathbf{x}$, so heißt λ **Eigenwert** von A und \mathbf{x} heißt **zugehöriger Eigenvektor** von A . Die Gesamtheit aller Eigenwerte von A ist das **Spektrum** von A .

Die Gleichung $A\mathbf{x} = \lambda\mathbf{x}$ ist äquivalent mit $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Dieses homogene Gleichungssystem hat genau dann eine nichttriviale Lösung $\mathbf{x} \neq \mathbf{0}$, wenn der Eigenwert λ Nullstelle des sogenannten **charakteristischen Polynoms** $p(\lambda) = |A - \lambda I|$ ist. Es ist leicht einzusehen, daß p ein Polynom in λ vom Grade n ist, das die Darstellung

$$p(\lambda) = (-\lambda)^n + c_{n-1}(-\lambda)^{n-1} + \cdots + c_1(-\lambda) + c_0$$

aufweist. Da ein Polynom n -ten Grades nach dem Hauptsatz der Algebra genau n Nullstellen hat, gibt es also zu A genau n , nicht notwendig verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, die auch komplex sein können. Die zu λ_i gehörenden Eigenvektoren \mathbf{x}_i von A sind dann Lösungen von $(A - \lambda_i I)\mathbf{x}_i = \mathbf{0}$.

Ist λ eine k -fache Nullstelle des charakteristischen Polynoms, so heißt λ **k -facher Eigenwert** von A .

Beispiel 1.2

Das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 0 & 2 & -1 \\ 1 & 0 & 3 \end{pmatrix}$$

ist gegeben durch

$$p(\lambda) = \begin{vmatrix} 1 - \lambda & -2 & 2 \\ 0 & 2 - \lambda & -1 \\ 1 & 0 & 3 - \lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - 9\lambda + 4 = (1 - \lambda)^2(4 - \lambda).$$

An dieser Darstellung erkennt man sofort, daß die Nullstellen des charakteristischen Polynoms gegeben sind durch $\lambda_1 = \lambda_2 = 1$ und $\lambda_3 = 4$, d.h., 1 ist zweifacher und 4 ist einfacher Eigenwert von A .

Ein zu $\lambda_1 = \lambda_2 = 1$ gehörender Eigenvektor \mathbf{x}_1 von \mathbf{A} ergibt sich als Lösung der Gleichung $(\mathbf{A} - 1\mathbf{I})\mathbf{x} = \mathbf{0}$. Sie wird vom Eigenvektor $\mathbf{x}_1 = (2, -1, -1)^\top$ erfüllt. Analog ist $\mathbf{x}_2 = (2, -1, 2)^\top$ ein zu $\lambda_3 = 4$ gehörender Eigenvektor, da für ihn gilt $(\mathbf{A} - 4\mathbf{I})\mathbf{x}_2 = \mathbf{0}$. \circ

Im obigen Beispiel ist $\mathbf{x}_1 = (2, -1, -1)^\top$ ein Eigenvektor. Offenbar ist dann auch $-3\mathbf{x}_1 = -3(2, -1, -1)^\top = (-6, 3, 3)^\top$ ein Eigenvektor. Ferner ist $\lambda_1 = 1$ ein $k = 2$ -facher Eigenwert. Dazu gibt es einen linear unabhängigen Eigenvektor $\mathbf{x}_2 = (2, -1, 2)^\top$. Diese Eigenschaften sind Spezialfälle von Aussagen, die im folgenden Satz formuliert sind:

Satz 1.15

Gegeben sei die (n, n) -Matrix \mathbf{A} mit den Eigenwerten $\lambda_1, \dots, \lambda_n$.

- (i) Ist \mathbf{x} Eigenvektor von \mathbf{A} , so ist auch $c\mathbf{x}$ für jedes $c \in \mathbb{R}$ Eigenvektor von \mathbf{A} .
- (ii) Zu einem einfachen Eigenwert von \mathbf{A} gibt es genau einen linear unabhängigen Eigenvektor.
- (iii) Zu einem k -fachen Eigenwert von \mathbf{A} gibt es wenigstens einen linear unabhängigen, aber höchstens k linear unabhängige Eigenvektoren. Die zu einem Eigenwert λ von \mathbf{A} gehörenden linear unabhängigen Eigenvektoren spannen einen Unterraum des \mathbb{R}^n auf, den sogenannten Eigenraum von λ . (Zum Begriff des Unterraumes siehe Abschnitt 1.9.2.)
- (iv) Zu einem k -fachen Eigenwert einer symmetrischen Matrix \mathbf{A} gibt es genau k linear unabhängige Eigenvektoren.
- (v) Sind $\mathbf{x}_1, \dots, \mathbf{x}_k$ Eigenvektoren zum Eigenwert λ , so ist auch $\sum_{i=1}^k l_i \mathbf{x}_i$ mit $l_1, \dots, l_k \in \mathbb{R}$ ein Eigenvektor von \mathbf{A} .
- (vi) Eigenvektoren, die zu verschiedenen Eigenwerten von \mathbf{A} gehören, sind linear unabhängig.
- (vii) Eigenvektoren einer symmetrischen Matrix \mathbf{A} , die zu verschiedenen Eigenwerten von \mathbf{A} gehören, sind orthogonal.
- (viii) $|\mathbf{A}| = \lambda_1 \lambda_2 \cdots \lambda_n$. Da für eine reguläre Matrix $|\mathbf{A}| \neq 0$ ist, folgt insbesondere, daß alle Eigenwerte $\lambda_i \neq 0$ sind. Umgekehrt besitzt eine singuläre Matrix mindestens einen Eigenwert mit $\lambda_i = 0$.
- (ix) $\text{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
- (x) Für die Anzahl k der von Null verschiedenen Eigenwerte einer Matrix \mathbf{A} gilt: $k \leq \text{rg}(\mathbf{A})$.

Für einige spezielle Matrizen gelten folgende Aussagen über ihre Eigenwerte und Eigenvektoren.

Satz 1.16

- (i) Die Eigenwerte einer symmetrischen Matrix sind reelle Zahlen.
- (ii) Die Eigenwerte einer Diagonalmatrix $\text{diag}(d_1, \dots, d_n)$ sind gerade die Diagonalelemente d_1, \dots, d_n ; speziell für die Einheitsmatrix \mathbf{I} ist $\lambda = \lambda_i = 1$ für $i = 1, \dots, n$ ein n -facher Eigenwert.
- (iii) Für jeden Eigenwert λ einer orthogonalen Matrix gilt $|\lambda| = 1$.
- (iv) Eine idempotente Matrix besitzt nur die Eigenwerte 0 oder 1.
- (v) Ist λ Eigenwert der regulären Matrix \mathbf{A} mit Eigenvektor \mathbf{x} , so ist $1/\lambda$ Eigenwert von \mathbf{A}^{-1} mit Eigenvektor \mathbf{x} .
- (vi) Für jeden Eigenwert λ einer Matrix \mathbf{A} ist $\alpha\lambda$ Eigenwert von $\alpha\mathbf{A}$.
- (vii) Für jeden Eigenwert λ einer Matrix \mathbf{A} ist λ^m Eigenwert von \mathbf{A}^m .

1.7 Diagonalisierung von Matrizen

Die Bestimmung der Eigenwerte einer Matrix \mathbf{A} gelingt oft leichter über eine sogenannte „Diagonalisierung“ von \mathbf{A} .

Definition 1.18

Zwei (n, n) -Matrizen \mathbf{A} und \mathbf{B} heißen **ähnlich**, wenn eine reguläre Matrix \mathbf{S} existiert, so daß $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ ist.

Ähnliche Matrizen besitzen dasselbe charakteristische Polynom und damit gleiche Eigenwerte, denn gilt $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, so ist

$$\begin{aligned}
 |\mathbf{B} - \lambda\mathbf{I}| &= |\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda\mathbf{I}| \\
 &= |\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda\mathbf{S}^{-1}\mathbf{S}| \\
 &= |\mathbf{S}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{S}| \\
 &= |\mathbf{S}^{-1}||\mathbf{A} - \lambda\mathbf{I}||\mathbf{S}| \\
 &= |\mathbf{A} - \lambda\mathbf{I}|.
 \end{aligned}$$

Definition 1.19

Eine (n, n) -Matrix \mathbf{A} heißt **diagonalisierbar**, wenn \mathbf{A} einer Diagonalmatrix $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ ähnlich ist, d.h., es existiert eine reguläre Matrix \mathbf{S} mit $\mathbf{D} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$.

Für diagonalisierbare Matrizen gelten folgende Regeln:

Satz 1.17

- (i) Eine diagonalisierbare Matrix hat die Eigenwerte d_1, \dots, d_n .
- (ii) A ist genau dann diagonalisierbar, wenn A insgesamt n linear unabhängige Eigenvektoren s_1, \dots, s_n besitzt. In diesem Fall erfüllt $S = (s_1, \dots, s_n)$ die Diagonalisierbarkeitsbedingung.
- (iii) Hinreichend (aber nicht notwendig) für die Diagonalisierbarkeit von A ist, daß alle n Eigenwerte von A verschieden sind.
- (iv) Ist A symmetrisch, so ist A diagonalisierbar über eine orthogonale Matrix, d.h., es gibt eine orthogonale Matrix P , so daß gilt $P^{-1}AP = P^TAP = D = \text{diag}(d_1, \dots, d_n)$. Die Spalten von P bestehen aus paarweise orthogonalen Eigenvektoren von A . Es folgt die Spektraldarstellung von A :

$$A = PDP^T. \quad (1.28)$$

- (v) Sei A regulär und symmetrisch und $m \in \mathbb{N}$. Dann ist

$$A^m = PD^mP^T \quad \text{mit} \quad D^m = \text{diag}(d_1^m, \dots, d_n^m). \quad (1.29)$$

Sind alle Eigenwerte von A positiv, so kann Gleichung (1.29) für $r \in \mathbb{Z}$ und $s \in \mathbb{N}$ verallgemeinert werden zu

$$A^{r/s} := PD^{r/s}P^T \quad \text{mit} \quad D^{r/s} := \text{diag}(d_1^{r/s}, \dots, d_n^{r/s}). \quad (1.30)$$

Sind einige Eigenwerte von A gleich 0, so gelten die Darstellungen in (1.29) und (1.30) auch dann, wenn die Exponenten negativ sind. Ist beispielsweise $d_i \geq 0$ für alle $i = 1, \dots, n$, so ist

$$A^{1/2} = PD^{1/2}P^T \quad \text{mit} \quad D^{1/2} = \text{diag}(\sqrt{d_1}, \dots, \sqrt{d_n}), \quad (1.31)$$

und ist $d_i > 0$ für alle $i = 1, \dots, n$, so ist

$$A^{-1/2} = PD^{-1/2}P^T \quad \text{mit} \quad D^{-1/2} = \text{diag}(1/\sqrt{d_1}, \dots, 1/\sqrt{d_n}). \quad (1.32)$$

Besitzt die (n, n) -Matrix A mehrfache Eigenwerte, so ist es möglich, daß A nicht diagonalisierbar ist. Nach Satz 1.17(ii) muß A dann weniger als n linear unabhängige Eigenvektoren besitzen. Doch auch in diesem Fall existiert eine einfach aufgebaute Matrix, die zu A ähnlich ist, denn es gilt der

Satz 1.18

Sei A eine (n, n) -Matrix mit den t verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_t$, wobei $\lambda_i, i = 1, \dots, t$, in der Vielfachheit n_i auftritt, so daß $n_1 + \dots + n_t = n$ ist. Dann existiert eine reguläre (n, n) -Matrix T , so daß gilt $J = T^{-1}AT$. Die Matrix J ist eine Blockdiagonalmatrix der Form:

$$J = \begin{pmatrix} J_{11} & 0 & \cdots & 0 \\ 0 & J_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{tt} \end{pmatrix},$$

wobei die (n_i, n_i) -Matrix J_{ii} gegeben ist durch

$$J_{ii} = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

J heißt **Jordan-Matrix** und T ist die **Jordan-Transformationsmatrix** von A . Jede (n, n) -Matrix A ist also einer Jordan-Matrix ähnlich. Wir sagen auch: A kann auf **Jordan-Normalform** transformiert werden.

Die Bestimmung von T und J für vorgegebenes A erfordert meist einen erheblichen numerischen Aufwand, siehe z.B. Murata (1977). Für den Fall, daß alle n Eigenwerte verschieden sind, geht die Jordan-Matrix in eine „reine“ Diagonalmatrix über.

In den folgenden Kapiteln wird mehrfach von der Singulärwertdarstellung einer Matrix Gebrauch gemacht, die im folgenden Satz formuliert wird.

Satz 1.19

Ist A eine beliebige (m, n) -Matrix mit $\text{rg}(A) = r$, so gibt es in Analogie zur Spektraldarstellung einer symmetrischen Matrix A (siehe Satz 1.17(iv)) eine

verallgemeinerte Diagonalisierung, die durch zwei orthogonale Matrizen \mathbf{U} und \mathbf{V} ermöglicht wird. Danach gilt die Singulärwertdarstellung von \mathbf{A} mit

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T. \quad (1.33)$$

Dabei ist Δ^2 die (r, r) -Diagonalmatrix der positiven Eigenwerte von $\mathbf{A}^T \mathbf{A}$ (und damit auch der von $\mathbf{A} \mathbf{A}^T$).

Insbesondere folgt so die „Diagonalisierung von $\mathbf{A} \mathbf{A}^T$ gemäß Satz 1.17(iv)“:

$$\mathbf{U}^T \mathbf{A} \mathbf{A}^T \mathbf{U} = \begin{pmatrix} \Delta^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

sowie die „Diagonalisierung von $\mathbf{A}^T \mathbf{A}$ gemäß Satz 1.17(iv)“:

$$\mathbf{V}^T \mathbf{A}^T \mathbf{A} \mathbf{V} = \begin{pmatrix} \Delta^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Die Diagonalelemente von Δ heißen **Singulärwerte** von \mathbf{A} . Sie sind die positiven Wurzeln der positiven Eigenwerte von $\mathbf{A} \mathbf{A}^T$. Zur Konstruktion der orthogonalen Matrizen \mathbf{U} und \mathbf{V} , die im allgemeinen nicht eindeutig sind, siehe z.B. Bunse und Bunse-Gerstner (1985).

1.8 Quadratische Formen und Definitheit

Definition 1.20

Gegeben sei die symmetrische (n, n) -Matrix $\mathbf{A} = (a_{ij})$. Dann heißt $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$, die zu \mathbf{A} gehörende **quadratische Form** in \mathbf{x} .

Durch $Q(\mathbf{x})$ wird also eine Abbildung $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ induziert, für die insbesondere gilt $Q(\mathbf{0}) = 0$. Weiter ist auf Grund der Symmetrie von \mathbf{A} für $\mathbf{x} = (\mathbf{x}_i)$:

$$\begin{aligned}
 Q(\mathbf{x}) &= \sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} x_j \\
 &= \sum_{i=1}^n a_{ii} x_i^2 + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n x_i a_{ij} x_j \\
 &= \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j.
 \end{aligned}$$

Die Symmetrie von \mathbf{A} kann o.B.d.A. angenommen werden. Ist \mathbf{A} nämlich nicht symmetrisch, so ersetze man \mathbf{A} durch die symmetrische Matrix $(\mathbf{A} + \mathbf{A}^\top)/2$. Dann ist wegen $\mathbf{x}^\top \mathbf{A} \mathbf{x} \in \mathbb{R}$:

$$\begin{aligned}
 \mathbf{x}^\top \mathbf{A} \mathbf{x} &= \mathbf{x}^\top \left(\frac{\mathbf{A} + \mathbf{A}^\top}{2} \right) \mathbf{x} \\
 &= \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}}{2} + \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}}{2} \\
 &= \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}}{2} + \frac{(\mathbf{x}^\top \mathbf{A}^\top \mathbf{x})^\top}{2} \\
 &= \mathbf{x}^\top \left(\frac{\mathbf{A} + \mathbf{A}^\top}{2} \right) \mathbf{x}.
 \end{aligned}$$

Definition 1.21

Eine quadratische Form $Q(\mathbf{x})$ heißt

- (i) **positiv definit**, falls $Q(\mathbf{x}) > 0$ für alle $\mathbf{x} \neq \mathbf{0}$ gilt.
- (ii) **positiv semidefinit**, falls $Q(\mathbf{x}) \geq 0$ für alle $\mathbf{x} \neq \mathbf{0}$ und $Q(\mathbf{x}_0) = 0$ für mindestens ein $\mathbf{x}_0 \neq \mathbf{0}$ gilt.
- (iii) **negativ definit**, falls $Q(\mathbf{x}) < 0$ für alle $\mathbf{x} \neq \mathbf{0}$ gilt.
- (iv) **negativ semidefinit**, falls $Q(\mathbf{x}) \leq 0$ für alle $\mathbf{x} \neq \mathbf{0}$ und $Q(\mathbf{x}_0) = 0$ für mindestens ein $\mathbf{x}_0 \neq \mathbf{0}$ gilt.

In allen anderen Fällen heißt $Q(\mathbf{x})$ **indefinit**. $Q(\mathbf{x})$ wird **nichtnegativ definit** genannt, falls $Q(\mathbf{x})$ entweder positiv definit oder positiv semidefinit ist, Kurzschreibweise: **n.n.d.** Entsprechend wird $Q(\mathbf{x})$ **nichtpositiv definit** genannt, falls $Q(\mathbf{x})$ entweder negativ definit oder negativ semidefinit ist, Kurzschreibweise: **n.p.d.**

Diese Definitheitsbegriffe von $Q(\mathbf{x})$ werden auch auf die zugehörige Matrix \mathbf{A} angewandt. Die Matrix \mathbf{A} heißt also positiv definit, wenn $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ positiv definit ist u.s.w.

Beispiel 1.3

Wir illustrieren die Definitheitsbegriffe für den Fall $n = 2$:

(i) Mit

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

ist $Q(\mathbf{x}) = x_1^2 + x_2^2 > 0$ für alle $\mathbf{x} \neq \mathbf{0}$. Damit ist $Q(\mathbf{x})$ positiv definit.

(ii) Mit

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

ist $Q(\mathbf{x}) = (x_1 + x_2)^2 \geq 0$ für alle $\mathbf{x} \neq \mathbf{0}$. Für $\mathbf{x}_0 = (1, -1)^\top \neq \mathbf{0}$ gilt $Q(\mathbf{x}_0) = 0$. Damit ist $Q(\mathbf{x})$ positiv semidefinit.

(iii) Mit

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

ist $Q(\mathbf{x}) = -(x_1^2 + x_2^2) < 0$ für alle $\mathbf{x} \neq \mathbf{0}$. Damit ist $Q(\mathbf{x})$ negativ definit.

(iv) Mit

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

ist $Q(\mathbf{x}) = -(x_1 + x_2)^2 \leq 0$ für alle $\mathbf{x} \neq \mathbf{0}$. Für $\mathbf{x}_0 = (1, -1)^\top \neq \mathbf{0}$ gilt $Q(\mathbf{x}_0) = 0$. Damit ist $Q(\mathbf{x})$ negativ semidefinit.

(v) Die Matrix

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

mit $Q(\mathbf{x}) = x_1^2 - x_2^2$ ist indefinit. ○

Für eine beliebige symmetrische (n, n) -Matrix \mathbf{A} läßt sich meist nicht so leicht wie in den obigen Beispielen die Frage nach der Definitheit beantworten. Im folgenden wird ein einfaches Kriterium über die Hauptabschnittsdeterminanten D_i von \mathbf{A} angegeben (siehe Definition 1.16):

Satz 1.20

Die quadratische Form $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ ist genau dann

- (i) positiv definit, falls $D_i > 0$ für alle $i = 1, \dots, n$ gilt.*
- (ii) negativ definit, falls gilt $D_1 < 0, D_2 > 0, D_3 < 0, D_4 > 0, \dots$*

Inbesondere ist jede positiv definite oder negativ definite Matrix regulär.

Eine Entscheidung über die Art der Definitheit von \mathbf{A} gelingt auch über die Kenntnis der Eigenwerte $\lambda_1, \dots, \lambda_n$ von \mathbf{A} .

Satz 1.21

Die quadratische Form $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ ist genau dann

- (i) positiv definit, falls $\lambda_i > 0$ für alle $i = 1, \dots, n$ ist.*
- (ii) positiv semidefinit, falls $\lambda_i \geq 0$ für alle $i = 1, \dots, n$ gilt und es einen Eigenwert mit $\lambda_i = 0$ gibt.*
- (iii) negativ definit, falls $\lambda_i < 0$ für alle $i = 1, \dots, n$ ist.*
- (iv) negativ semidefinit, falls $\lambda_i \leq 0$ für alle $i = 1, \dots, n$ gilt und es einen Eigenwert mit $\lambda_i = 0$ gibt.*

Besitzt \mathbf{A} einen positiven und einen negativen Eigenwert, so ist \mathbf{A} indefinit.

Satz 1.21 ist eine unmittelbare Folgerung von

Satz 1.22

Ist A eine symmetrische (n, n) -Matrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n$, so existiert eine orthogonale Transformation $\mathbf{y} = (y_i) = P^T \mathbf{x}$, so daß gilt

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \sum_{i=1}^n \lambda_i y_i^2.$$

Das ergibt sich aus der Spektraldarstellung $A = PDP^T$ von A , siehe Satz 1.17(iv).

Weitere Aussagen über quadratische Formen bzw. ihre zugehörigen Matrizen sind formuliert im

Satz 1.23

- (i) Für jede (m, n) -Matrix A sind $A^T A$ und AA^T nichtnegativ definit.
- (ii) A ist genau dann positiv definit, wenn A^{-1} positiv definit ist.
- (iii) Ist A positiv definit und B regulär, so ist $B^T A B$ positiv definit.

1.9 Vektorräume

Der in der Regel bereits im Grundstudium für Wirtschaftswissenschaftler eingeführte Begriff des Vektorraumes steht meist im engen Zusammenhang mit der Lösbarkeit linear homogener Gleichungssysteme, deren Lösungsraum einen Vektorraum bildet, siehe Wetzell u.a. (1981). Die Elemente dieses Vektorraumes sind also n -Tupel reeller Zahlen, d.h., die dort betrachteten Vektorräume sind ausgezeichnete Teilmengen des \mathbb{R}^n oder der \mathbb{R}^n selbst. Der Begriff des Vektorraumes kann aber viel allgemeiner gefaßt werden, was im folgenden geschehen soll.

1.9.1 Definition und Beispiele

Definition 1.22

Auf einer Menge \mathcal{V} sei eine Addition der Elemente $x, y, z \in \mathcal{V}$ mit folgenden Eigenschaften definiert:

- (i) $x + y \in \mathcal{V}$
- (ii) $x + y = y + x$ (Kommutativität)
- (iii) $(x + y) + z = x + (y + z)$ (Assoziativität)

- (iv) Es gibt ein **Nullelement** $o \in \mathcal{V}$, so daß für jedes $x \in \mathcal{V}$ gilt $x + o = x$
 (v) Für jedes $x \in \mathcal{V}$ gibt es ein **inverses Element** $a \in \mathcal{V}$ mit $x + a = o$.

Ferner sei eine sogenannte **Skalarmultiplikation** von Elementen c, c_1, c_2 aus \mathbb{R} mit Elementen aus \mathcal{V} erklärt, die folgende Eigenschaften hat:

- (vi) $c \cdot x \in \mathcal{V}$
 (vii) $c \cdot (x + y) = c \cdot x + c \cdot y$ (Distributivgesetz)
 (viii) $(c_1 + c_2) \cdot x = c_1 \cdot x + c_2 \cdot x$ (Distributivgesetz)
 (ix) $(c_1 c_2) \cdot x = c_1 \cdot (c_2 \cdot x)$
 (x) $1 \cdot x = x$.

Dann heißt \mathcal{V} **Vektorraum** oder **linearer Raum** über \mathbb{R} , und die Elemente von \mathcal{V} werden ebenfalls als **Vektoren** bezeichnet.

Ein Vektorraum ist also wegen (i) abgeschlossen bezüglich seiner „inneren“ Verknüpfung (Addition) und ebenso wegen (vi) abgeschlossen bezüglich seiner „äußeren“ Verknüpfung (Skalarmultiplikation) mit \mathbb{R} . Statt der Menge \mathbb{R} der reellen Zahlen kann auch die Menge \mathbb{C} der komplexen Zahlen als Skalarmenge gewählt werden.

Für den Vektorraum \mathcal{V} mit der zugehörigen Skalarmenge \mathbb{R} bzw. \mathbb{C} und den Verknüpfungen $+$ und \cdot schreiben wir kurz $(\mathcal{V}, \mathbb{R}, +, \cdot)$ bzw. $(\mathcal{V}, \mathbb{C}, +, \cdot)$.

Beispiel 1.4

- $(\mathbb{R}, \mathbb{R}, +, \cdot)$, $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$ $(\mathbb{C}, \mathbb{R}, +, \cdot)$, $(\mathbb{C}, \mathbb{C}, +, \cdot)$; aber nicht $(\mathbb{R}, \mathbb{C}, +, \cdot)$.
- $(\mathcal{Z}, \mathbb{R}, +, \cdot)$, wobei \mathcal{Z} die Menge aller reellen Zahlenfolgen $[x_n]$ ist mit $[x_n] + [y_n] := [x_n + y_n]$ und $c \cdot [x_n] := [cx_n]$.
- $(\mathbb{R}^{m \times n}, \mathbb{R}, +, \cdot)$, wobei $\mathbb{R}^{m \times n}$ die Menge aller (m, n) -Matrizen ist mit der in Abschnitt 1.3 erklärten Addition und Skalarmultiplikation, siehe Definition 1.2.
- $(\mathcal{F}, \mathbb{R}, +, \cdot)$, wobei \mathcal{F} die Menge aller Funktionen $f: [a, b] \rightarrow \mathbb{R}$ ist mit $(f + g)(x) := f(x) + g(x)$ und $(c \cdot f)(x) := cf(x)$ für alle $x \in [a, b]$ und alle $c \in \mathbb{R}$. \mathcal{F} heißt auch **reeller Funktionenraum** über $[a, b]$.
- $(\mathcal{P}, \mathbb{R}, +, \cdot)$, wobei \mathcal{P} die Menge aller Polynome $p: \mathbb{R} \rightarrow \mathbb{R}$ ist mit $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ für ein festes $n \in \mathbb{N}$. Addition und Skalarmultiplikation sind wie in 4. definiert.

1.9.2 Unterräume

Aus den Eigenschaften (i) bis (x) eines Vektorraumes \mathcal{V} lassen sich nun folgende Regeln ableiten:

Satz 1.24

- (i) *Es existiert ein $z \in \mathcal{V}$ mit $x + z = y$ für alle $x, y \in \mathcal{V}$*
- (ii) *$0 \cdot x = o$ mit $0 \in \mathbb{R}$ und $o \in \mathcal{V}$*
- (iii) *$c \cdot o = o$ für alle $c \in \mathbb{R}$*
- (iv) *$(-c) \cdot x = -(c \cdot x)$ für alle $c \in \mathbb{R}$ und alle $x \in \mathcal{V}$.*

Definition 1.23

Eine Teilmenge \mathcal{U} eines Vektorraumes \mathcal{V} , die selbst wieder einen Vektorraum bildet, heißt **Unterraum** von \mathcal{V} .

Um zu zeigen, daß die Teilmenge \mathcal{U} Unterraum des Vektorraumes \mathcal{V} ist, sind nur die Bedingungen $x + y \in \mathcal{U}$ und $cx \in \mathcal{U}$ für alle $x, y \in \mathcal{U}$ und alle $c \in \mathbb{R}$ zu prüfen, weil alle anderen Bedingungen für \mathcal{U} als Teilmenge von \mathcal{V} „automatisch“ erfüllt sind.

Beispiel 1.5

Gegeben sei eine (endliche oder unendliche) Teilmenge \mathcal{M} des \mathbb{R}^n . Wir zeigen, daß das sogenannte **orthogonale Komplement**

$$\mathcal{M}^\perp := \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^\top \mathbf{m} = 0 \text{ für alle } \mathbf{m} \in \mathcal{M}\}$$

ein linearer Unterraum des \mathbb{R}^n ist. Seien $\mathbf{x}, \mathbf{y} \in \mathcal{M}^\perp$. Dann gilt $\mathbf{x}^\top \mathbf{m} = 0$ und $\mathbf{y}^\top \mathbf{m} = 0$ für alle $\mathbf{m} \in \mathcal{M}$. Mithin gilt auch $(\mathbf{x} + \mathbf{y})^\top \mathbf{m} = 0$ für alle $\mathbf{m} \in \mathcal{M}$, woraus $(\mathbf{x} + \mathbf{y}) \in \mathcal{M}^\perp$ folgt. Weiterhin ist offensichtlich $k\mathbf{x} \in \mathcal{M}^\perp$ für alle Skalare $k \in \mathbb{R}$. Damit ist \mathcal{M}^\perp ein Unterraum des \mathbb{R}^n .

Der Leser beachte, daß \mathcal{M}^\perp auch dann ein Unterraum ist, wenn \mathcal{M} kein Unterraum des \mathbb{R}^n ist. ○

Über die folgenden beiden Definitionen lassen sich nun leicht Unterräume eines Vektorraumes \mathcal{V} konstruieren.

Definition 1.24

Für die Elemente $x_1, \dots, x_n \in \mathcal{V}$ heißt $\sum_{i=1}^n c_i x_i \in \mathcal{V}$ mit $c_i \in \mathbb{R}$ oder $c_i \in \mathbb{C}$ eine **Linearkombination** von x_1, \dots, x_n .

Definition 1.25

Es sei \mathcal{T} eine beliebige Teilmenge eines Vektorraumes \mathcal{V} . Dann heißt die Menge aller Linearkombinationen von Elementen aus \mathcal{T} die **lineare Hülle** von \mathcal{T} , in Zeichen $\mathcal{L}(\mathcal{T})$.

Man kann leicht zeigen, daß die lineare Hülle einer Teilmenge \mathcal{T} eines Vektorraumes \mathcal{V} ein Unterraum von \mathcal{V} ist.

Der Begriff der Linearkombination von Elementen aus \mathcal{V} spielt eine wichtige Rolle in dem nun folgenden Abschnitt über lineare Unabhängigkeit bzw. Abhängigkeit von Elementen eines Vektorraumes.

1.9.3 Lineare Unabhängigkeit und Basis**Definition 1.26**

Die Elemente x_1, \dots, x_n des Vektorraumes \mathcal{V} heißen **linear unabhängig**, falls die Gleichung $\sum_{i=1}^n c_i x_i = 0$ nur für $c_1 = \dots = c_n = 0$ erfüllt ist; anderenfalls heißen sie **linear abhängig**.

Sind also die Elemente $x_1, \dots, x_n \in \mathcal{V}$ linear abhängig, so läßt sich mindestens eines der x_i als Linearkombination der anderen Elemente darstellen.

Beispiel 1.6

1. Die Einheitsvektoren $e_i = (\delta_{ij}) \in \mathbb{R}^n$, $i = 1, \dots, n$ sind linear unabhängig. Dabei ist δ_{ij} das Kronecker-Symbol (1.20).
2. Die Vektoren $(a, a)^\top, (b, b)^\top \in \mathbb{R}^2$ mit $a \neq b$ sind linear abhängig.
3. Gegeben seien n Funktionen $f_i: [a, b] \rightarrow \mathbb{R}$. Diese sind linear unabhängig, wenn aus $\sum_{i=1}^n c_i f_i = 0$ (Nullfunktion) folgt $c_1 = \dots = c_n = 0$. Die Gleichung $\sum_{i=1}^n c_i f_i = 0$ ist äquivalent damit, daß $\sum_{i=1}^n c_i f_i(x) = 0$ für alle $x \in [a, b]$ gilt. Wesentlich ist dabei die Vorgabe des Intervalls $[a, b]$. So sind die Funktionen $f_1(x) = x$ und $f_2(x) = |x|$ auf $[-1, +1]$ linear unabhängig, auf $[-1, 0]$ oder auf $[0, 1]$ aber linear abhängig.

Weitere Beispiele für linear unabhängige Funktionen sind:

- (a) $f_1(x) = e^x$ und $f_2(x) = e^{-x}$ für alle $x \in \mathbb{R}$.
- (b) $f_0(x) = 1$, $f_1(x) = x$, $f_2(x) = x^2$, \dots , $f_n(x) = x^n$ für alle $x \in \mathbb{R}$.
- (c) $f_1(x) = x$ und $f_2(x) = x - 1$ für alle $x \in \mathbb{R}$.
- (d) $f_1(x) = \sin x$, $f_2(x) = \sin 2x$, \dots , $f_n(x) = \sin nx$ für alle $x \in \mathbb{R}$.

Definition 1.27

Eine Teilmenge \mathcal{B} des Vektorraumes \mathcal{V} heißt **Basis** von \mathcal{V} , wenn

- (i) alle Elemente von \mathcal{B} linear unabhängig sind,
- (ii) jedes Element von \mathcal{V} als Linearkombination von Elementen aus \mathcal{B} dargestellt werden kann.

Es gilt der

Satz 1.25

- (i) *Jeder Vektorraum \mathcal{V} besitzt eine Basis \mathcal{B} .*
- (ii) *Die Darstellung eines jeden Elementes $x \in \mathcal{V}$ als Linearkombination von Elementen einer Basis \mathcal{B} ist eindeutig.*
- (iii) *Verschiedene Basen \mathcal{B}_1 und \mathcal{B}_2 eines Vektorraumes \mathcal{V} haben stets dieselbe Mächtigkeit. Dabei heißen zwei Mengen **gleichmächtig**, wenn zwischen ihnen eine bijektive Abbildung existiert.*

Definition 1.28

Die Mächtigkeit einer Basis des Vektorraumes \mathcal{V} heißt **Dimension** von \mathcal{V} , in Zeichen $\dim(\mathcal{V})$.

Ist die Basis \mathcal{B} endlich — nur solche Basen werden in den folgenden Kapiteln betrachtet — so ist $\dim(\mathcal{V})$ gerade gleich der Maximalzahl linear unabhängiger Elemente von \mathcal{V} . Ist etwa $\dim(\mathcal{V}) = r$, so bestehen alle Basen von \mathcal{V} wegen Satz 1.7(iii) aus genau r linear unabhängigen Elementen b_1, \dots, b_r . Alle anderen Elemente von \mathcal{V} lassen sich dann als Linearkombinationen von b_1, \dots, b_r darstellen.

Beispiel 1.7

1. Die n Einheitsvektoren $\mathbf{e}_i \in \mathbb{R}^n$ bilden eine Basis des \mathbb{R}^n , woraus $\dim(\mathbb{R}^n) = n$ folgt.
2. Die Vektoren $\mathbf{e}_1 = (1, 0, 0)^\top$ und $\mathbf{e}_2 = (0, 1, 0)^\top$ sind zwar unabhängig, bilden jedoch keine Basis des \mathbb{R}^3 .
3. Die Funktionen $f_i: \mathbb{R} \rightarrow \mathbb{R}$ mit $f_i(x) = x^i$, $i = 0, 1, 2, \dots, n$ bilden eine Basis des Vektorraumes \mathcal{P} aller reellen Polynome vom Grade n . Mithin ist $\dim(\mathcal{P}) = n + 1$, siehe 5. in Beispiel 1.4 und (3b) in Beispiel 1.6.

1.9.4 Konvexe Teilmengen eines Vektorraumes

Der im folgenden erklärte Begriff einer konvexen Menge spielt in der linearen und nichtlinearen Optimierung, die in den Kapiteln 6 und 8 behandelt werden, eine wichtige Rolle.

Definition 1.29

Eine Teilmenge \mathcal{K} eines Vektorraumes \mathcal{V} über \mathbb{R} heißt **konvex**, wenn gilt $cx + (1 - c)y \in \mathcal{K}$ für alle $x, y \in \mathcal{K}$ und alle c mit $0 \leq c \leq 1$. Dabei heißt $cx + (1 - c)y$ als spezielle Linearkombination auch **konvexe Linearkombination** von x und y .

Für $\mathcal{V} = \mathbb{R}^n$ stellt $\{cx + (1 - c)y \mid 0 \leq c \leq 1\}$ für gegebene Punkte $x, y \in \mathbb{R}^n$ die Menge aller Punkte im \mathbb{R}^n dar, die auf der Verbindungsstrecke zwischen x und y liegen. Konvexität einer Menge $\mathcal{K} \subset \mathbb{R}^n$ bedeutet also, daß mit je zwei Punkten aus \mathcal{K} auch die Verbindungsstrecke der beiden Punkte in \mathcal{K} enthalten ist.

Beispiel 1.8

Rechtecke, Kreise, Dreiecke und Polyeder sind konvexe Teilmengen des \mathbb{R}^2 . Intervalle sind die einzigen konvexen Teilmengen des \mathbb{R}^1 . Die in der Abbildung 1.2 gezeichnete Teilmenge des \mathbb{R}^2 ist nicht konvex.

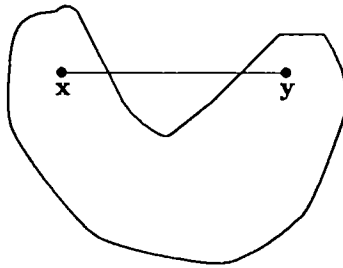


Abbildung 1.2: Nichtkonvexe Menge

○

Man kann leicht zeigen:

Satz 1.26

Der Durchschnitt beliebig vieler konvexer Teilmengen von \mathcal{V} ist konvex.

Ist \mathcal{M} eine beliebige Teilmenge eines Vektorraumes \mathcal{V} , so können wir die kleinste konvexe Teilmenge von \mathcal{V} auszeichnen, die \mathcal{M} umfaßt:

Definition 1.30

Gegeben sei eine Teilmenge \mathcal{M} des Vektorraumes \mathcal{V} . Der Durchschnitt aller konvexen Teilmengen \mathcal{N} von \mathcal{V} mit $\mathcal{M} \subset \mathcal{N}$ heißt **konvexe Hülle** $\mathcal{K}(\mathcal{M})$ von \mathcal{M} .

Explizit ist $\mathcal{K}(\mathcal{M})$ gegeben durch

$$\mathcal{K}(\mathcal{M}) = \left\{ \sum_{i=1}^n c_i x_i \mid x_i \in \mathcal{M}, c_i \in \mathbb{R}, c_i \geq 0, n \in \mathbb{N}, i = 1, \dots, n, \sum_{i=1}^n c_i = 1 \right\}.$$

1.10 Metrische Räume

In diesem Abschnitt werden wir uns mit Mengen beschäftigen, für die eine Distanzfunktion (Metrik) definiert ist. Diese sogenannten metrischen Räume sind spezielle topologische Räume, zu denen ein System von Umgebungen (Topologien) mit gewissen Eigenschaften gehört. Wir wollen hier nicht das allgemeine Konzept topologischer Räume verfolgen, sondern uns auf metrische Räume beschränken. Durch den Begriff der Metrik wird dann ein spezieller Umgebungsbegriff (offene Kugel) induziert.

Definition 1.31

Es sei \mathcal{M} eine nichtleere Menge. Eine Abbildung $d: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ mit den Eigenschaften

- (i) $d(x, y) \geq 0$ für alle $x, y \in \mathcal{M}$
- (ii) $d(x, y) = 0$ gilt genau dann, wenn $x = y$ ist
- (iii) $d(x, y) = d(y, x)$ (Symmetrie)
- (iv) $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in \mathcal{M}$ (**Dreiecksungleichung**)

heißt **Metrik** auf \mathcal{M} . Die Menge \mathcal{M} ist dann ein **metrischer Raum**, und $d(x, y)$ bezeichnet den **Abstand** zwischen x und y .

Beispiel 1.9

1. Betrachte $\mathcal{M} = \mathbb{R}^n$. Für $\mathbf{x} = (x_i), \mathbf{y} = (y_i) \in \mathbb{R}^n$ und $r \geq 1$ ist die **Minkowski-Metrik** festgelegt durch

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_r := (\sum_{i=1}^n |x_i - y_i|^r)^{1/r}.$$

Speziell ist für

(a) $r = 1$: $d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n |x_i - y_i|$,

(b) $r = 2$: $d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2} = \sqrt{(\mathbf{x} - \mathbf{y})^\top (\mathbf{x} - \mathbf{y})}$ (**euklidischer Abstand**). Mit dieser Metrik wird \mathbb{R}^n als **euklidischer Raum** bezeichnet.

2. Für $\mathcal{M} = \mathbb{R}^n$ ist $d(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq n} |x_i - y_i|$ eine Metrik.
3. Es bezeichne \mathcal{H} den sogenannten **Hilbertraum** aller reellen Zahlenfolgen $[x_n]$ mit konvergenter Quadratsumme $\sum_{i=1}^{\infty} x_i^2$. Dann ist durch $d([x_n], [y_n]) := \sum_{i=1}^{\infty} (x_i - y_i)^2$ eine Metrik auf \mathcal{H} festgelegt. Mit dieser Metrik kann \mathcal{H} als eine Verallgemeinerung des euklidischen Raumes \mathbb{R}^n angesehen werden.
4. Für Zahlen $a, b \in \mathbb{R}$ mit $a < b$ sei $\mathcal{C}[a, b]$ die Menge aller stetigen Funktionen $f: [a, b] \rightarrow \mathbb{R}$ auf dem Intervall $[a, b]$. Für $f, g \in \mathcal{C}[a, b]$ wird durch $d(f, g) := \max\{|f(x) - g(x)| \mid a \leq x \leq b\}$ eine Metrik auf $\mathcal{C}[a, b]$ erklärt.

○

Definition 1.32

Gegeben sei der metrische Raum \mathcal{M} mit Metrik d . Für $x_0 \in \mathcal{M}$ und $\varepsilon > 0$ heißt die Menge $\mathcal{U}_\varepsilon(x_0) := \{x \mid x \in \mathcal{M}, d(x, x_0) < \varepsilon\}$ eine **Kugelumgebung** von x_0 mit Radius ε oder kürzer: **ε -Umgebung** von x_0 .

Beispiel 1.10

1. Wir betrachten den euklidischen Raum \mathbb{R}^2 mit der Metrik $d(\mathbf{x}, \mathbf{x}_0) = \|\mathbf{x} - \mathbf{x}_0\|_2$. Die ε -Umgebung des Punktes $\mathbf{x}_0 \in \mathbb{R}^2$ ist das Innere der Kreisfläche mit \mathbf{x}_0 als Mittelpunkt und ε als Radius.
2. Wir betrachten den metrischen Raum \mathbb{R}^2 mit der Metrik $d(\mathbf{x}, \mathbf{x}_0) = \max_{i=1,2} |x_i - x_{i0}|$. Die ε -Umgebung des Punktes \mathbf{x}_0 ist ein Quadrat mit der Kantenlänge 2ε um \mathbf{x}_0 als Mittelpunkt.

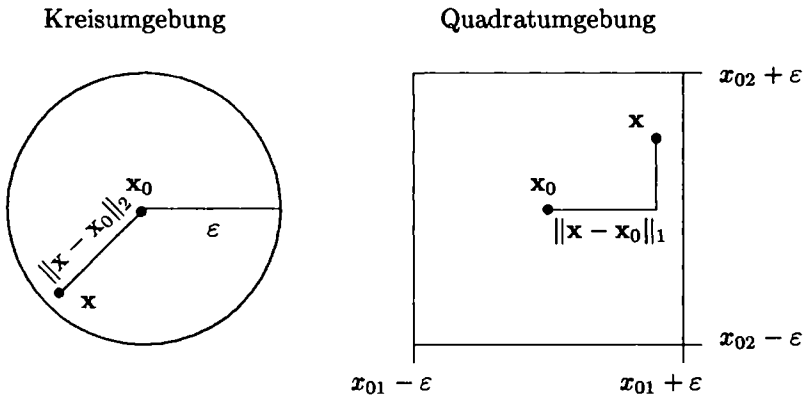


Abbildung 1.3: Kreis- und Quadratumgebung

○

Wir führen nun eine Reihe wichtiger topologischer Begriffe für einen metrischen Raum \mathcal{M} mit Metrik d ein.

Definition 1.33

Sei \mathcal{N} Teilmenge von \mathcal{M} . Ein Element $x_0 \in \mathcal{M}$ heißt

- (i) **innerer Punkt** von \mathcal{N} , wenn ein $\varepsilon > 0$ existiert mit $\mathcal{U}_\varepsilon(x_0) \subset \mathcal{N}$.
- (ii) **äußerer Punkt** von \mathcal{N} , wenn ein $\varepsilon > 0$ existiert mit $\mathcal{U}_\varepsilon(x_0) \cap \mathcal{N} = \emptyset$.
- (iii) **Randpunkt** von \mathcal{N} , wenn er weder innerer noch äußerer Punkt von \mathcal{N} ist.

Die Menge aller Randpunkte von \mathcal{N} heißt **Rand** $\mathcal{R}(\mathcal{N})$ von \mathcal{N} .

Beispiel 1.11

Für $a, b \in \mathbb{R}$ mit $a < b$ sei $(a, b) = \{x \mid x \in \mathbb{R}, a < x < b\}$ ein „offenes“ Intervall. Dann sind alle $x_0 \in (a, b)$ innere Punkte und a und b sind Randpunkte. Jede reelle Zahl x_0 mit $x_0 < a$ oder $b < x_0$ ist äußerer Punkt. ○

Definition 1.34

Eine Teilmenge \mathcal{N} von \mathcal{M} heißt

- (i) **offen**, wenn \mathcal{N} nur innere Punkte enthält.

- (ii) **abgeschlossen**, wenn $\mathcal{M} \setminus \mathcal{N}$ offen ist. Dabei bezeichnet $\mathcal{M} \setminus \mathcal{N}$ das **Komplement** von \mathcal{N} in \mathcal{M} , d.h. die Menge aller Elemente von \mathcal{M} , die nicht zu \mathcal{N} gehören.

Der \mathbb{R}^n und die leere Menge \emptyset sind sowohl offen als auch abgeschlossen. Weiterhin gilt:

Satz 1.27

- (i) *Die Vereinigung von beliebig vielen offenen Mengen ist offen.*
- (ii) *Der Durchschnitt von endlich vielen offenen Mengen ist offen.*
- (iii) *Die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen.*
- (iv) *Der Durchschnitt von beliebig vielen abgeschlossenen Mengen ist abgeschlossen.*
- (v) *Ist die Menge \mathcal{N} abgeschlossen, so ist $\mathcal{R}(\mathcal{N}) \subset \mathcal{N}$. Ist die Menge \mathcal{N} offen, so ist $\mathcal{N} \cap \mathcal{R}(\mathcal{N}) = \emptyset$.*

Definition 1.35

Für eine Menge $\mathcal{N} \subset \mathcal{M}$ heißt $x_0 \in \mathcal{M}$

- (i) **Berührungspunkt** von \mathcal{N} , wenn jede ε -Umgebung von x_0 mindestens ein Element von \mathcal{N} enthält.
- (ii) **Häufungspunkt** von \mathcal{N} , wenn jede ε -Umgebung von x_0 mindestens ein von x_0 verschiedenes Element von \mathcal{N} enthält, d.h. wenn x_0 Berührungspunkt von $\mathcal{N} \setminus \{x_0\}$ ist.

Ein Berührungspunkt von \mathcal{N} ist also genau dann *kein* Häufungspunkt von \mathcal{N} , wenn er Element von \mathcal{N} ist und es eine ε -Umgebung gibt, in der kein weiteres Element von \mathcal{N} liegt. Ein solcher Punkt heißt **isolierter Punkt**.

Es gilt: Eine Menge $\mathcal{N} \subset \mathcal{M}$ ist genau dann abgeschlossen, wenn jeder Häufungspunkt von \mathcal{N} auch Element von \mathcal{N} ist.

Definition 1.36

Gegeben sei eine Menge $\mathcal{N} \subset \mathcal{M}$. Die Menge $\overline{\mathcal{N}}$ aller Berührungspunkte von \mathcal{N} heißt **abgeschlossene Hülle** von \mathcal{N} .

Für eine Teilmenge \mathcal{N} des metrischen Raumes \mathcal{M} gilt: