



Lineare Algebra 1/2

Hans Joachim Pflug
Rechen- und Kommunikationszentrum
der RWTH Aachen

MATSE-Ausbildung

Studiengang „Scientific Programming“
FH Aachen

Überarbeitung:
Benno Willemsen
willemsen@rz.rwth-aachen.de

20. September 2013

Inhaltsverzeichnis

1	Vorbereitung	9
1.1	Matrizen	9
1.1.1	Grundlegende Begriffe	9
1.1.1.1	Definition einer Matrix	9
1.1.1.2	Gleichheit zweier Matrizen	9
1.1.1.3	Nullmatrix	9
1.1.1.4	Transponierte Matrix	10
1.1.2	Rechenoperationen für Matrizen	10
1.1.3	Quadratische Matrizen	11
1.1.4	Spezielle quadratische Matrizen	11
1.1.4.1	Dreiecksmatrix	11
1.1.4.2	Diagonalmatrix	12
1.1.4.3	Symmetrische Matrix	12
1.1.5	Das Matrixprodukt	12
1.1.6	Rechenregeln für das Matrixprodukt	14
1.1.7	Matrizen in MATLAB	15
1.2	Lösung eines linearen Gleichungssystems	16
1.2.1	Lineare Gleichungssysteme	16
1.2.2	Übersicht über die Lösungsverfahren	16
1.2.3	Lineare Gleichungssysteme in Matrixschreibweise	17
1.2.4	Lösungsansatz mit Gauß-Algorithmus und Gauß-Jordan-Algorithmus	18
1.2.5	Lösung mit MATLAB	20
1.2.6	Unter- und überbestimmte Gleichungssysteme	21
1.2.6.1	Genau eine Lösung	22
1.2.6.2	Keine Lösung	23
1.2.6.3	Mehrere Lösungen	23
2	Analytische Geometrie	27
2.1	Vektoren	27
2.1.1	Definition von Vektoren	27
2.1.1.1	Einführung	27
2.1.1.2	Kurze Einführung in Vektorräume	28

2.1.1.3	Der Vektorraum \mathbb{R}^n	29
2.1.2	Der \mathbb{R}^n in der analytischen Geometrie	30
2.1.2.1	Gebundene Vektoren und Ortsvektoren	31
2.1.2.2	Geometrische Interpretation von Addition und Multiplikation im \mathbb{R}^n	31
2.1.2.3	Differenzvektoren	33
2.1.3	Parallele Vektoren	34
2.1.4	Das Skalarprodukt	34
2.1.5	Die Norm eines Vektors	36
2.1.5.1	Einheitsvektoren	37
2.1.5.2	Normieren von Vektoren	38
2.1.5.3	Orthogonale (senkrechte) Vektoren	38
2.1.5.4	Satz des Pythagoras	39
2.1.6	Der Winkel zwischen zwei Vektoren	40
2.1.6.1	Die senkrechte Projektion eines Vektors auf einen anderen	40
2.1.6.2	Der Winkel zwischen zwei Vektoren	41
2.1.7	Weitere Ungleichungen	44
2.1.7.1	Die Schwarzsche Ungleichung	44
2.1.7.2	Die Dreiecksungleichung	45
2.1.8	Das Vektorprodukt (Kreuzprodukt)	46
2.2	Geraden und Ebenen	49
2.2.1	Geradengleichung im \mathbb{R}^2	49
2.2.1.1	Parametergleichungen	49
2.2.1.2	Parameterlose Gleichungen	50
2.2.2	Hyperebenen im \mathbb{R}^n	52
2.2.3	Geraden im \mathbb{R}^3	53
2.2.4	Zusammenfassung	53
2.2.5	Umrechnen zwischen Parameterform und parameter- loser Form	54
2.2.5.1	Bilden eines senkrechten Vektors	54
2.2.5.2	Geraden im \mathbb{R}^2	54
2.2.5.3	Ebene im \mathbb{R}^3	55
2.2.5.4	Hyperebenen im \mathbb{R}^n	56
2.2.6	Parallelität und Orthogonalität von Geraden und Hy- perebenen	56
2.2.7	Schnittmengen zwischen Geraden und Ebenen im \mathbb{R}^2 und \mathbb{R}^3	57
2.2.7.1	Schnittmenge einer Gerade oder Ebene mit einer Hyperebene	57
2.2.7.2	Bestimmung des Schnittpunkts zweier Gera- den im \mathbb{R}^3	59
2.2.8	Abstandsbestimmung im \mathbb{R}^2 und \mathbb{R}^3	60
2.2.8.1	Einführung	60

2.2.8.2	Der Abstand eines Punktes und einer Geraden im \mathbb{R}^3	63
2.3	Die Determinante im \mathbb{R}^2 und \mathbb{R}^3	64
2.3.1	Die Determinante von Matrizen bis 3×3	64
2.3.1.1	Geometrische Bedeutung der Determinante im \mathbb{R}^3	65
2.3.1.2	Rechenregeln	66
2.3.2	Lineare 3×3 -Gleichungssysteme	67
2.4	Kegelschnitte	69
2.4.1	Kurven und Gleichungen	69
2.4.2	Kegelschnitte	70
2.4.3	Lage von Kegelschnitten	70
2.4.4	Übersicht über die Kegelschnitte	72
2.4.4.1	Übersicht über die Standard-Kegelschnitte	72
2.4.4.2	Übersicht über die entarteten Kegelschnitte	73
2.4.5	Kegelschnittgleichungen	74
2.4.6	Bestimmung der Kegelschnitte aus der Kegelschnittgleichung	75
3	Vektorräume	79
3.1	Gruppen, Körper, Vektorräume	79
3.1.1	Äquivalenzrelationen	79
3.1.2	Gruppen	80
3.1.2.1	Einführung	80
3.1.2.2	Restklassen	82
3.1.3	Körper	83
3.1.4	Kartesisches Produkt	85
3.1.5	Vektorräume (lineare Räume)	85
3.1.6	Untervektorräume (Unterräume)	88
3.2	Lineare Unabhängigkeit, Basis, Dimension	89
3.2.1	Lineare Unabhängigkeit	89
3.2.1.1	Nachweis der linearen Unabhängigkeit im Vektorraum \mathbb{R}^n	92
3.2.1.2	Lineare Unabhängigkeit von n Vektoren des \mathbb{R}^n	93
3.2.1.3	Nachweis der Linearen Unabhängigkeit im Vektorraum $C[a, b]$	94
3.2.2	Basis und Dimension	97
3.2.3	Unendliche Mengen linear unabhängiger Vektoren	100
3.3	Polynome	100
3.3.1	Polynome und Lineare Unabhängigkeit	100
3.3.2	Test auf lineare Unabhängigkeit von Polynomen	104
3.4	Skalarprodukt, euklidische und unitäre Räume	107
3.4.1	Das Skalarprodukt	107

3.4.2	Die Norm eines Skalarprodukts	112
3.4.3	Metriken	113
3.5	Orthogonalität	114
3.5.1	Eigenschaften orthogonaler Vektoren	115
3.5.2	Das Orthonormalisierungsverfahren nach Gram-Schmidt	118
3.5.2.1	„Sinn“ des Verfahrens	119
3.5.2.2	Funktionsweise des Verfahrens	119
3.5.2.3	Bemerkungen zur Orthonormalisierung	122
3.5.2.4	Orthonormalisierung in anderen euklidischen Vektorräumen als \mathbb{R}^n	122
4	Lineare Abbildungen	125
4.1	Abbildungen	125
4.1.1	Einführung	125
4.1.2	Injektivität und Surjektivität	126
4.2	Lineare Abbildungen	128
4.2.1	Bild und Kern	129
4.2.2	Lineare Abbildungen und Matrizen	131
4.2.2.1	Aussehen von linearen Abbildungen	131
4.2.2.2	Matrizen als lineare Abbildungen	133
4.2.3	Berechnung linearer Abbildungen	134
4.2.4	Die Dimensionsformel für lineare Abbildungen	139
4.2.5	Beispielaufgaben zur Berechnung von Kern und Bild .	140
4.2.5.1	Weg über Berechnung der Abbildungsmatrix	140
4.2.5.2	Andere Rechenwege	142
4.2.6	Isomorphismus und Umkehrabbildung	143
4.3	Matrizen	147
4.3.1	Allgemeines	147
4.3.2	Die Inverse einer Matrix	148
4.3.3	Wie invertiert man eine Matrix?	150
4.3.3.1	Erläuterung anhand eines Beispiels	150
4.3.3.2	Allgemeine Erläuterung: Elementarmatrizen	151
4.3.4	Matrizen als Vektorräume	153
4.3.5	Der Rang einer Matrix	154
4.4	Determinanten	157
4.4.1	Einführung	157
4.4.2	Eigenschaften der Determinante	158
4.4.3	Geometrische Interpretation	160
4.4.4	Schematische Darstellung der Determinante	161
4.4.5	Rechenverfahren	162
4.4.5.1	Leibnizsche Formel	162
4.4.5.2	Entwicklungssatz nach Laplace	164
4.4.5.3	Gauß-Algorithmus	166
4.4.5.4	Sarrussche Regel	167

4.4.5.5	Kombination der Regeln	168
4.4.6	Rechenaufwand zur Berechnung einer Determinanten	168
4.4.7	Berechnung der Inversen mit der Methode der Unterdeterminanten	169
4.4.8	Unterdeterminanten und Rang einer Matrix	171
4.5	Transformationsmatrizen	172
4.5.1	Basiswechsel	172
4.5.2	Basistransformation bei Abbildungsmatrizen	174
4.5.3	Koordinatentransformationen in der Ebene	175
5	Lineare Gleichungssysteme	179
5.1	Lösbarkeit eines linearen Gleichungssystems	179
5.1.1	Lösbarkeit von quadratischen Systemen mit $\det=0$	183
5.1.2	Äquivalente Bedingungen zur Lösbarkeit eines Gleichungssystems	186
5.2	Die Cramersche Regel	187
5.3	Überbestimmte lineare Gleichungssysteme	189
5.3.1	Orthogonalprojektion auf Unterräume	189
5.3.2	Methode der kleinsten Quadrate	190
5.3.3	Ausgleichsrechnung mit der Methode der kleinsten Quadrate	192
5.4	Unterbestimmte lineare Gleichungssysteme	196
6	Spezielle Matrizen	199
6.1	Orthogonale Matrizen	199
6.1.1	Eigenschaften orthogonaler Matrizen	199
6.2	Die QR-Zerlegung	202
6.3	Eigenwerte von Matrizen	205
6.3.1	Berechnung der Eigenwerte und Eigenvektoren	207
6.3.2	Eigenwerte spezieller Matrizen	210
6.3.3	Diagonalisierbarkeit linearer Abbildungen	211
6.4	Symmetrische Matrizen	213
6.5	Definitheit von Matrizen	218
6.5.1	Definitheit	218
6.5.2	Die Matrix $A^T \cdot A$	221
6.6	Hauptachsentransformationen	222
6.6.1	Einführung	222
6.6.2	Hauptachsentransformation von Kegelschnitten	223
6.6.2.1	Hauptachsentransformation von quadratischen Formen	223
6.6.2.2	Hauptachsentransformation von Kegelschnitten	223
6.7	Hadamardsche Ungleichung	227

Kapitel 1

Vorbereitung

1.1 Matrizen

1.1.1 Grundlegende Begriffe

1.1.1.1 Definition einer Matrix

Definition 1.1: Eine **Matrix** A vom Typ (m, n) ist ein aus $m \cdot n$ Werten bestehendes rechteckiges Schema mit m Zeilen und n Spalten. Diese Werte werden **Elemente** genannt.

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

Die Matrixelemente werden mit a_{ij} bezeichnet, wobei i der Zeilenindex und j der Spaltenindex ist. Die Matrix selbst wird manchmal anstatt mit einem Großbuchstaben auch mit der Form (a_{ij}) bzw. $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ bezeichnet.

1.1.1.2 Gleichheit zweier Matrizen

Definition 1.2: Zwei Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ vom gleichen Typ heißen gleich, wenn alle Matrixelemente übereinstimmen d.h. wenn $\forall i \quad \forall j \quad : \quad a_{ij} = b_{ij}$. Man schreibt dann $A = B$.

1.1.1.3 Nullmatrix

Definition 1.3: Die Elemente einer **Nullmatrix** sind Null, d.h. A ist genau dann eine Nullmatrix, wenn $\forall i \quad \forall j \quad : \quad a_{ij} = 0$. Man verwendet dann auch als Zeichen für die Matrix die 0.

Beispiel 1.4:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

1.1.1.4 Transponierte Matrix

Definition 1.5: Die **transponierte Matrix** A^T entsteht aus einer Matrix, indem man Zeilen und Spalten miteinander vertauscht.

B ist also genau dann gleich A^T , wenn $\forall i \quad \forall j \quad : \quad b_{ij} = a_{ji}$.

Beispiel 1.6:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad A^T = (1 \ 2 \ 3)$$

1.1.2 Rechenoperationen für Matrizen

Die Matrizen vom Typ (m, n) bilden einen sogenannten Vektorraum. Die Definition eines Vektorraums folgt später. Hier beschränken wir uns darauf, die dazu nötigen Operationen auf Matrizen zu definieren:

Bemerkung 1.7: Einfache Operationen sind:

- ⊕: **Addition** und **Subtraktion** von Matrizen: Zwei Matrizen vom gleichen Typ werden addiert bzw. subtrahiert, indem man ihre entsprechenden Elemente addiert bzw. subtrahiert.
- ⊙: **Multiplikation** einer Matrix mit einem Skalar: Die Multiplikation einer Matrix mit einem Skalar erfolgt, indem man jedes Matrixelement mit dem Skalar multipliziert.

Beispiel 1.8: Es sei

$$A = \begin{pmatrix} 3 & -6 & 5 \\ 4 & 0 & 12 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} -6 & 12 & -10 \\ -8 & 0 & -24 \end{pmatrix}$$

Dann gilt:

$$2A = \begin{pmatrix} 6 & -12 & 10 \\ 8 & 0 & 24 \end{pmatrix} \quad \text{und} \quad 2A + B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

1.1.3 Quadratische Matrizen

Definition 1.9: Eine **quadratische Matrix** ist eine Matrix mit gleich vielen Zeilen und Spalten ($m = n$). Die allgemeine Gestalt einer quadratischen Matrix ist:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Dabei heißt die Diagonale von links oben nach rechts unten die **Hauptdiagonale**:

$$A = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}.$$

Der Ausdruck **Nebendiagonale** hat verschiedene Bedeutungen. Einerseits heißt die Diagonale von rechts oben nach links unten Nebendiagonale. Andererseits werden auch die beiden Diagonalen unmittelbar über und unter der Hauptdiagonalen Nebendiagonalen genannt.

Beispiel 1.10: (4×4 -Matrizen): Die Hauptdiagonale wird mit „H“, die Nebendiagonalen mit „n“ bezeichnet.

$$A = \begin{pmatrix} H & n & . & . \\ n & H & n & . \\ . & n & H & n \\ . & . & n & H \end{pmatrix} \quad A = \begin{pmatrix} . & . & . & n \\ . & . & n & . \\ . & n & . & . \\ n & . & . & . \end{pmatrix}$$

1.1.4 Spezielle quadratische Matrizen

1.1.4.1 Dreiecksmatrix

Definition 1.11: Eine Matrix, bei der alle Elemente oberhalb der Hauptdiagonalen gleich 0 sind, heißt **untere Dreiecksmatrix**. Beispiel:

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 2 & 9 & 3 \end{pmatrix}$$

Eine Matrix, bei der alle Elemente unterhalb der Hauptdiagonalen gleich 0 sind, heißt **obere Dreiecksmatrix**. Beispiel:

$$A = \begin{pmatrix} 2 & 9 & 0 \\ 0 & -3 & 7 \\ 0 & 0 & 1 \end{pmatrix}$$

1.1.4.2 Diagonalmatrix

Definition 1.12: Eine Matrix, bei der alle Elemente außerhalb der Hauptdiagonalen gleich 0 sind, heißt **Diagonalmatrix**. Beispiel:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Definition 1.13: Eine Diagonalmatrix, die auf der Hauptdiagonalen ausschließlich den Wert 1 hat, heißt **Einheitsmatrix**. Eine Einheitsmatrix wird gewöhnlich mit dem Buchstaben E oder I bezeichnet und durch einen Index, der die Dimension angibt ergänzt. Beispiel:

$$E = I_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

1.1.4.3 Symmetrische Matrix

Definition 1.14: Sind die Elemente der Matrix spiegelsymmetrisch zur Hauptdiagonalen angeordnet, nennt man die Matrix **symmetrisch**. Es gilt dann $A = A^T$. Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

1.1.5 Das Matrixprodukt

Definition 1.15: Es sei A eine $m \times n$ -Matrix und sei B eine $n \times r$ -Matrix. Dann heißt die Matrix $C = AB$ mit:

$$c_{ik} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^n a_{ij}b_{jk} \quad \forall i = 1, \dots, m; k = 1, \dots, r$$

das **Produkt** der Matrizen A und B .

Diese Regel lässt sich schematisch darstellen:

$$\begin{array}{c}
 \begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1r} \\ b_{21} & \cdots & b_{2k} & \cdots & b_{2r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{nr} \end{pmatrix} = B \\
 \downarrow \\
 \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \xrightarrow{\quad} \begin{pmatrix} \quad & \quad & \quad & \quad \\ \quad & \quad & \quad & \quad \\ \quad & \quad & c_{ik} & \quad \\ \quad & \quad & \quad & \quad \\ \quad & \quad & \quad & \quad \end{pmatrix} = C
 \end{array}$$

Man sieht leicht, dass die Matrix C eine $m \times r$ -Matrix ist.

Beispiel 1.16: Gegeben sind die Matrizen

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad B = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 0 & -1 & -2 & -3 \\ 2 & 5 & 4 & 3 \end{pmatrix}$$

A ist eine 2×3 -Matrix und B ist eine 3×4 -Matrix, also ist das Produkt $C = AB$ eine 2×4 -Matrix. c_{23} ist das Produkt aus der 2. Zeile von A und der 3. Spalte von B .

$$c_{23} = (4 \cdot 2) + (5 \cdot (-2)) + (6 \cdot 4) = 22$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 4 & 3 & 2 & 1 \\ 0 & -1 & -2 & -3 \\ 2 & 5 & 4 & 3 \end{pmatrix} = \begin{pmatrix} \square & \square & \square & \square \\ \square & \square & 22 & \square \end{pmatrix}$$

Das Element c_{14} ist das Produkt aus der 1. Zeile von A und der 4. Spalte von B :

$$c_{14} = (1 \cdot 1) + (2 \cdot (-3)) + (3 \cdot 3) = 4$$

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

Die restlichen Elemente werden analog berechnet.

Bemerkung 1.17: Die Produktbildung ist nur dann möglich, wenn die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt.

Dies lässt sich mit folgendem Schema überprüfen, in dem die Größe der Matrizen nebeneinandergestellt wird:

$$\begin{array}{ccc}
 A & B & AB \\
 m \times n & n \times r & m \times r
 \end{array}
 =$$

AB ist definiert, wenn die beiden inneren Zahlen übereinstimmen. Die beiden äußeren Zahlen geben dann die Größe (den Typ) von AB an.

Beispiel 1.18: Gegeben seien Matrizen A, B und C der folgenden Größen:

$$\begin{array}{ccc}
 A & B & C \\
 2 \times 4 & 4 \times 5 & 5 \times 2
 \end{array}$$

Dann ist AB eine 2×5 -Matrix, CA eine 5×4 -Matrix und BC eine 4×2 -Matrix. Die Produkte AC, BA und CB sind nicht definiert.

1.1.6 Rechenregeln für das Matrixprodukt

Im Folgenden sind die drei wichtigsten Rechenregeln aufgeführt. Wenn die entsprechenden Produkte und Summen definiert sind, dann gilt:

- Assoziativgesetz: $A(BC) = (AB)C$
- Distributivgesetz: $A(B + C) = AB + AC$ und $(A + B)C = AC + BC$
- Es gibt (quadratische) Matrizen A, B mit $AB \neq BA$;
das Matrixprodukt ist also im Allgemeinen **nicht** kommutativ.
Beispiel:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 11 & 11 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 8 & 12 \end{pmatrix}$$

- Es gibt Matrizen $A \neq 0$ und $B \neq 0$ mit $AB = 0$.
Beispiel:

$$\begin{pmatrix} 4 & 12 \\ -2 & -6 \end{pmatrix} \begin{pmatrix} 3 & 6 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

- $(AB)^T = B^T A^T$
- $EA = AE = A$, E ist die (vom Typ her jeweils passende) Einheitsmatrix.
- Es sei 0 eine Nullmatrix. Falls sich die Produkte $0 \cdot A$ und $A \cdot 0$ bilden lassen, sind sie ebenfalls eine Nullmatrix.

1.1.7 Matrizen in MATLAB

Viele der Verfahren in dieser Vorlesung werden gewöhnlich nicht von Hand durchgeführt, sondern vom Computer berechnet. Dazu wird oft das Programmpaket MATLAB verwendet, das primär auf Berechnungen mit Hilfe von Matrizen ausgelegt ist. Alternativen zu MATLAB sind z.B. Scilab oder Octave.

Zunächst einige grundlegende Eigenheiten vorweg. In MATLAB kann man interaktiv Anweisungen ausführen lassen. Dazu schreibt man hinter das Prompt-Zeichen `>>` die Eingabezeile. MATLAB gibt dann darunter das Ergebnis der Eingabezeile aus. Die Ausgabe des Ergebnisses kann man unterdrücken, wenn man die Eingabezeile mit einem Semikolon abschließt.

Eine Matrix wird in MATLAB wie folgt angelegt:

```
>> a=[1 2 3;4 5 6;7 8 9]
```

```
a =
```

```
     1     2     3
     4     5     6
     7     8     9
```

Die Eingabezeile ist die oberste Zeile, alle unteren Zeilen sind die Ergebnisausgabe von MATLAB. Zeilen werden durch ein Semikolon getrennt. Matrizen können addiert und multipliziert werden.

```
>> a=[1 2 3;4 5 6;7 8 9];
```

```
>> b=2*a
```

```
b =
```

```
     2     4     6
     8    10    12
    14    16    18
```

```
>> c=b+a
```

```
c =
```

```
     3     6     9
    12    15    18
    21    24    27
```

```
>> d=b*a
```

d =

$$\begin{array}{ccc} 60 & 72 & 84 \\ 132 & 162 & 192 \\ 204 & 252 & 300 \end{array}$$

Vektoren des \mathbb{R}^n lassen sich als einspaltige Matrizen darstellen.

1.2 Lösung eines linearen Gleichungssystems

1.2.1 Lineare Gleichungssysteme

Wir betrachten ein System von m linearen Gleichungen und n Unbekannten x_1, x_2, \dots, x_n :

$$\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots + \dots + \dots + \dots = \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array}$$

Zum Beispiel:

$$\begin{array}{r} 2x + 3y - 4z = 8 \\ -4x + 2y + 3z = -5 \\ 3x + y + 2z = 13 \end{array}$$

Sind die Werte auf der rechten Seite b_1, \dots, b_n alle gleich 0, dann spricht man von einem **homogenen Gleichungssystem**, ansonsten von einem **inhomogenen Gleichungssystem**.

1.2.2 Übersicht über die Lösungsverfahren

Im Blockkurs Mathematik haben wir solche linearen Gleichungssysteme bereits gelöst. An dieser Stelle geben wir einen Überblick über gängige Lösungsverfahren:

- Einfache Lösungsverfahren sind das **Gleichsetzungsverfahren**, das **Einsetzungsverfahren** und das **Additionsverfahren**. Ab drei Variablen sind diese nicht mehr empfehlenswert.
- Das Standardverfahren ist der **Gauß-Algorithmus** (Gaußsches Eliminationsverfahren), das in diesem Kapitel vorgestellt wird.
- Für spezielle Gleichungssysteme ist die Lösung nach der **Cramerschen Regel** eine Alternative. Wir werden sie in LA2 kennenlernen.
- Eine weitere Methode ist die **Cholesky-Zerlegung**, die in Numerik behandelt wird.

- Für sehr große Gleichungssysteme sind direkte Lösungen zu aufwendig. Hier werden Näherungslösungen mit **iterativen Lösern** berechnet. In der Numerik-Vorlesung werden das **Jacobi-Verfahren** und das **Gauß-Seidel-Verfahren** behandelt.
- Gibt es keine eindeutige Lösung, kann man mit der **Methode der kleinsten Quadrate** die bestmögliche Näherung berechnen. Diese Methode wird sowohl in Analysis als auch in LA2 vorgestellt.

1.2.3 Lineare Gleichungssysteme in Matrixschreibweise

Ein Gleichungssystem kann wie eine Matrix geschrieben werden, wenn man die Unbekannten weglässt. Dadurch wird die Schreibweise kompakter und die Rechnung übersichtlicher. Dabei gibt es, wie im folgenden Beispiel zu sehen ist, zwei unterschiedliche Matrizen.

Zum folgenden Gleichungssystem

$$\begin{aligned} 2x + 3y - 4z &= 8 \\ -4x + 2y + 3z &= -5 \\ 3x + y + 2z &= 13 \end{aligned}$$

gibt es zwei Matrizen: Die **Koeffizientenmatrix** besteht aus den Koeffizienten der linken Seite.

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

Die **erweiterte Matrix** besteht aus der Koeffizientenmatrix und einer zusätzlichen Spalte, in der die Werte der rechten Seite stehen. Im Englischen heißt diese Form **augmented matrix**.

$$\begin{pmatrix} 2 & 3 & -4 & 8 \\ -4 & 2 & 3 & -5 \\ 3 & 1 & 2 & 13 \end{pmatrix}$$

Oft findet man auch eine Art der Matrixschreibweise, bei der linke und rechte Seite durch einen senkrechten Strich getrennt sind.

$$\begin{array}{ccc|c} 2 & 3 & -4 & 8 \\ -4 & 2 & 3 & -5 \\ 3 & 1 & 2 & 13 \end{array}$$

1.2.4 Lösungsansatz mit Gauß-Algorithmus und Gauß-Jordan-Algorithmus

Definition 1.19: Wir nennen zwei Gleichungssysteme **äquivalent**, falls sie die gleiche Lösung haben.¹ Wir nennen zwei Matrizen A und B äquivalent ($A \sim B$), falls die entsprechenden Gleichungssysteme die gleiche Lösung besitzen.

Beispiel 1.20: Das Gleichungssystem

$$\begin{aligned} 2x + 3y - 4z &= 8 \\ -4x + 2y + 3z &= -5 \\ 3x + y + 2z &= 13 \end{aligned} \quad (a)$$

ist äquivalent zu den Gleichungssystemen

$$\begin{aligned} 2x + 3y - 4z &= 8 \\ 8y - 5z &= 11 \\ z &= 1 \end{aligned} \quad (b)$$

und

$$\begin{aligned} x &= 3 \\ y &= 2 \\ z &= 1 \end{aligned} \quad (c)$$

Auch die drei zugehörigen erweiterten Matrizen A, B, C sind dann äquivalent:

$$A = \begin{pmatrix} 2 & 3 & -4 & 8 \\ 3 & 1 & 2 & 13 \\ -4 & 2 & 3 & -5 \end{pmatrix}, B = \begin{pmatrix} 2 & 3 & -4 & 8 \\ 0 & 8 & -5 & 11 \\ 0 & 0 & 1 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Gleichungssystem (b) heißt **Stufenform**. Gleichungssystem (c) heißt **reduzierte Stufenform**. In diesem einfachen Fall² hat die zugehörige Matrix B unterhalb der Hauptdiagonalen nur Nullen (da sie nicht quadratisch ist, ist sie keine Dreiecksmatrix). Die Matrix C hat Einsen auf der Hauptdiagonalen und Nullen darunter und darüber. Nur in der letzten Spalte stehen beliebige Werte.

Es ist leicht zu sehen, dass die Stufenform (Ergebnis des Gauß-Algorithmus) und insbesondere die reduzierte Stufenform (Ergebnis des Gauß-Jordan-Algorithmus) übersichtlicher und die Bestimmung der Lösung deutlich einfacher sind. Bei der erweiterten Stufenform kann man die Lösung sogar einfach ablesen. Darum werden wir versuchen, durch Umformungen ein äquivalentes Gleichungssystem zu bestimmen, das die reduzierte Stufenform besitzt. Solche Umformungen nennt man **Äquivalenzumformungen** der Zeilen. Es gibt davon drei Stück:

¹Dies ist eine Äquivalenzrelation. Die Bedingungen für eine solche Relationen werden später noch genauer erklärt.

²Komplizierte Fälle betrachten wir später.

Z1. Addition eines Vielfachen einer Zeile zu einer anderen.

Z2. Vertauschen zweier Zeilen.

Z3. Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0, \lambda \in \mathbb{R}$.³

Um auf die reduzierte Stufenform zu gelangen, muss man zunächst die Stufenform errechnen. Dies erreicht man, indem man (hauptsächlich durch die Operation Z1) die Einträge unterhalb der Hauptdiagonalen auf 0 bringt. Man beginnt in der ersten Spalte und arbeitet sich spaltenweise vor:

$$\begin{array}{ccc|c}
 2 & 3 & -4 & 8 \\
 -4 & 2 & 3 & -5 & | + 2 \cdot (I) \\
 3 & 1 & 2 & 13 & | - \frac{3}{2} \cdot (I) \\
 \hline
 2 & 3 & -4 & 8 \\
 0 & 8 & -5 & 11 \\
 0 & -\frac{7}{2} & 8 & 2 & | + \frac{7}{16} \cdot (II) \\
 \hline
 2 & 3 & -4 & 8 \\
 0 & 8 & -5 & 11 \\
 0 & 0 & \frac{93}{16} & \frac{93}{16}
 \end{array}$$

Die Stufenform ist nicht eindeutig. Es gibt mehrere mögliche (und äquivalente) Stufenformen. Hat man eine Stufenform erreicht, geht man zeilenweise von unten nach oben vor. Zunächst bringt man das Element auf der Hauptdiagonalen⁴ auf 1, anschließend alle Elemente darüber auf 0. In unserem Beispiel sieht das so aus:

³Die dritte Umformung ist für den Gauß-Algorithmus nicht nötig, aber insbesondere für Handrechnungen vorteilhaft.

⁴ Später werden wir stattdessen das sogenannte **Pivot-Element** nehmen.

$$\begin{array}{r|l}
 2 & 3 & -4 & 8 \\
 0 & 8 & -5 & 11 \\
 0 & 0 & \frac{93}{16} & \frac{93}{16} & | \div \frac{93}{16} \\
 \hline
 2 & -4 & 8 & | + 4 \cdot (III) \\
 0 & 8 & -5 & 11 & | + 5 \cdot (III) \\
 0 & 0 & 1 & 1 \\
 \hline
 2 & 3 & 0 & 12 \\
 0 & 8 & 0 & 16 & | \div 8 \\
 0 & 0 & 1 & 1 \\
 \hline
 2 & 3 & 0 & 12 & | - 3 \cdot (II) \\
 0 & 1 & 0 & 2 \\
 0 & 0 & 1 & 1 \\
 \hline
 2 & 0 & 0 & 6 & | \div 2 \\
 0 & 1 & 0 & 2 \\
 0 & 0 & 1 & 1 \\
 \hline
 1 & 0 & 0 & 3 \\
 0 & 1 & 0 & 2 \\
 0 & 0 & 1 & 1
 \end{array}$$

Transferiert man das Schema wieder in das eigentliche Gleichungssystem zurück, erhält man die Lösung:

$$\begin{aligned}
 x &= 3 \\
 y &= 2 \\
 z &= 1
 \end{aligned}$$

1.2.5 Lösung mit MATLAB

Man sieht schon im letzten Beispiel, dass die Zahlen schnell unhandlich werden und die Rechnung von Hand langsam und fehleranfällig wird, deshalb unterscheiden wir zwei Fälle:

- Einfache Berechnungen von Hand.
- Berechnungen mit dem Computer.

Zu letzterem Punkt benutzen wir wieder das Programmpaket MATLAB. Es gibt in MATLAB mehrere Arten, lineare Gleichungssysteme zu lösen. Universell und praktisch ist der Befehl `rref`. `rref` ist die Abkürzung von *reduced row echelon form*, der englischen Übersetzung von **reduzierte Stufenform**. Wir sehen uns die Wirkungsweise einfach einmal an. Die erste Beispiel-Zeile

`format rat` veranlasst MATLAB, sein Ergebnis in Form von Brüchen auszugeben, was bei unseren kleinen Beispielaufgaben gut klappt. MATLAB rechnet intern aber immer mit Fließkommazahlen. In der zweiten Zeile geben wir die erweiterte Matrix zum obigen Beispiel ein, die anschließend in die reduzierte Stufenform umgerechnet wird.

```
>> format rat
>> a=[2 3 -4 8;-4 2 3 -5;3 1 2 13]

a =

     2         3     -4         8
    -4         2         3     -5
     3         1         2     13

>> rref(a)

ans =

     1         0         0         3
     0         1         0         2
     0         0         1         1
```

Das Resultat des `rref`-Kommandos bedeutet rückübersetzt:

$$\begin{aligned}x &= 3 \\y &= 2 \\z &= 1\end{aligned}$$

was die Lösung des Gleichungssystems ist.

1.2.6 Unter- und überbestimmte Gleichungssysteme

Nicht ganz so leicht zu interpretieren ist die reduzierte Stufenform bei unter- und überbestimmten Gleichungssystemen. **Unterbestimmte Gleichungssysteme** sind Gleichungssysteme mit weniger Gleichungen als Unbekannten, während **überbestimmte Gleichungssysteme** mehr Gleichungen als Unbekannte besitzen. Es gilt, dass ein Gleichungssystem nicht immer (eindeutig) lösbar ist. Vielmehr gibt es insgesamt drei Möglichkeiten:

1. Es existiert eine eindeutige Lösung.
2. Es existiert keine Lösung.

3. Es existieren unendlich viele Lösungen.⁵

Welche dieser drei Fälle eintritt, kann man an der reduzierten Stufenform (wie auch an der einfachen Stufenform) der erweiterten Matrix erkennen. Dazu müssen wir diese Form analysieren. Der erste Nicht-Null-Eintrag jeder Zeile heißt **Pivot-Element**. In der reduzierten Stufenform ist das Pivot-Element immer eine 1. Bezeichnet man schematisch ein Pivot-Element mit einem P und ein beliebiges Element mit einem x, so sieht eine reduzierte Stufenform beispielsweise wie folgt aus:

$$\begin{pmatrix} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \end{pmatrix}$$

Eine Spalte, in der ein Pivot-Element vorkommt, heißt **Pivot-Spalte**. In unserem Beispiel sind alle Spalten, mit Ausnahme der letzten, Pivot-Spalten. Pivot-Spalten enthalten in der reduzierten Stufenform außer dem Pivot-Element selbst nur Nullen. Nicht-Pivot-Spalten können auch andere Elemente enthalten. Wichtig ist außerdem:

- Es dürfen Null-Zeilen vorkommen, also Zeilen, in denen jeder Eintrag 0 ist. Diese müssen die untersten Zeilen der Matrix sein. Null-Zeilen stehen für die Gleichung $0 = 0$ und können ignoriert werden.
- Die Pivot-Spalte der Zeile $n + 1$ muss rechts der Pivot-Spalte der Zeile n liegen.

Reduzierte Stufenformen sind also auch folgende Matrizen:

$$\begin{pmatrix} P & 0 & x & 0 & x \\ 0 & P & x & 0 & x \\ 0 & 0 & 0 & P & x \\ 0 & 0 & 0 & 0 & P \end{pmatrix} \begin{pmatrix} P & 0 & x & x & 0 & x & x & x \\ 0 & P & x & x & 0 & x & x & x \\ 0 & 0 & 0 & 0 & P & x & x & x \end{pmatrix} \begin{pmatrix} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Wir nennen im Folgenden die Spalten 1 bis $n - 1$ der erweiterten Matrix **Koeffizientenmatrix** und die n -te Spalte **letzte Spalte**.

1.2.6.1 Genau eine Lösung

Hat das Gleichungssystem eine eindeutige Lösung, so erhält man immer die Stufenform bzw. die reduzierte Stufenform

$$\begin{array}{c} P \ x \ x | x \\ 0 \ P \ x | x \\ 0 \ 0 \ P | x \end{array} \quad \text{bzw.} \quad \begin{array}{c} P \ 0 \ 0 | x \\ 0 \ P \ 0 | x \\ 0 \ 0 \ P | x \end{array} .$$

⁵Oft steht an dieser Stelle „Es existieren mehrere Lösungen.“ Wenn mehr als eine Lösung existiert, sind es aber automatisch unendlich viele.

1.2.6.2 Keine Lösung

Falls die letzte Spalte der Stufenform eine Pivot-Spalte ist, ist das Gleichungssystem **nicht lösbar**. Man erhält also z.B. die folgende Stufenform bzw. reduzierte Stufenform:

$$\begin{array}{c|c} P & x & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & P \end{array} \quad \text{bzw.} \quad \begin{array}{c|c} P & 0 & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & P \end{array} .$$

Beispiel 1.21:

$$\begin{array}{c|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{array}$$

Die letzte Zeile besteht links aus lauter Nullen; rechts steht eine Zahl ungleich Null. Ausgeschrieben bedeutet das:

$$0 \cdot x + 0 \cdot y + 0 \cdot z = 1$$

Diese Gleichung stimmt nicht, gleichgültig, welche Werte x, y und z annehmen. Daraus folgt, dass das Gleichungssystem nicht lösbar ist.

1.2.6.3 Mehrere Lösungen

Erhält man in der (reduzierten) Stufenform weniger Pivotspalten als das Gleichungssystem Unbekannte hat, so hat das Gleichungssystem mehrere Lösungen. Ein Beispiel für diesen Fall ist das folgende Schema:

$$\begin{array}{c|c} P & x & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & 0 \end{array} \quad \text{bzw.} \quad \begin{array}{c|c} P & 0 & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & 0 \end{array} .$$

Es gibt (mindestens) einen **freien Parameter**, d.h. eine der Variablen kann einen beliebigen Wert annehmen. Befolgt man eine einfache Lösungsregel, dann gehört zu jeder Nicht-Pivot-Spalte der Koeffizientenmatrix eine freie Variable.

Beispiel: Zum Gleichungssystem

$$\begin{array}{l} 2 \cdot x - 4 \cdot y + 2 \cdot z = 8 \\ 1 \cdot x - 1 \cdot y - 7 \cdot z = 6 \end{array}$$

berechnen wir die äquivalente reduzierte Stufenform. Der Bequemlichkeit halber benutzen wir MATLAB:

```
>> format rat
>> a=[2 -4 2 8;1 -1 -7 6];
>> rref(a)
```

ans =

$$\begin{array}{cccc} 1 & 0 & -15 & 8 \\ 0 & 1 & -8 & 2 \end{array}$$

Das Ergebnis bedeutet als Gleichungssystem:

$$\begin{array}{l} 1 \cdot x - 0 \cdot y - 15 \cdot z = 8 \\ 0 \cdot x + 1 \cdot y - 8 \cdot z = 2 \end{array}$$

oder

$$\begin{array}{l} x = 8 + 15 \cdot z \\ y = 2 + 8 \cdot z \end{array}$$

oder in Matrix-Schreibweise (später werden wir **Vektor-Schreibweise** dazu sagen):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} + z \begin{pmatrix} 15 \\ 8 \end{pmatrix}$$

Diese Lösung ist noch nicht vollständig, denn die vollständige Lösung besteht aus dem Tripel (x, y, z) . Den Wert von z können wir aber sofort angeben, denn $z = 0 + 1 \cdot z$. Hängen wir die entsprechende Zeile an, so ergibt sich als Endergebnis:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 15 \\ 8 \\ 1 \end{pmatrix}$$

Zusammengefasst ergibt sich bei nicht eindeutigen Gleichungssystemen folgende Vorgehensweise:

- Die Nicht-Pivot-Spalten werden mit der zugehörigen Variablen versehen auf die rechte Seite gebracht, wobei sich alle Vorzeichen umdrehen.
- Zeilen für fehlende Variablen werden in den Lösungsvektor nach dem Schema $x = x$ eingefügt.

Beispiel: Berechne die Lösung des Gleichungssystems:

$$\begin{array}{rcl} 2a & +b & -2c & = & -2 \\ -4a & -2b & +c & +2d & = & 2 \\ -2a & -b & & +d & = & 1 \end{array}$$

Lösung (hier sparen wir uns wieder die Rechnung von Hand):

```
>> a=[2 1 -2 0 -2;-4 -2 1 2 2;-2 -1 0 1 1];  
>> rref(a)
```

```
ans =
```

```
      1      1/2      0      0      -1  
      0      0      1      0      0  
      0      0      0      1     -1
```

Die zweite Spalte der Matrix ist keine Pivot-Spalte. Das ergibt als Zwischenergebnis:

$$\begin{pmatrix} a \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix}$$

und als Endergebnis:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Kapitel 2

Analytische Geometrie

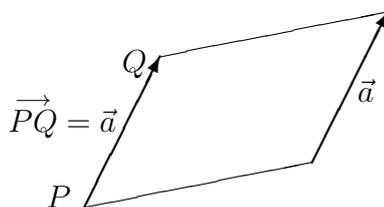
2.1 Vektoren

2.1.1 Definition von Vektoren

2.1.1.1 Einführung

Für den Begriff „Vektor“ wird man, je nach Zusammenhang, eine spezielle oder eine allgemeine Definition finden, wobei auf den ersten Blick nicht zu erkennen ist, wie die Definitionen miteinander vereinbar sind. Die spezielle Bedeutung stammt aus der analytischen Geometrie und lautet etwa wie folgt:

Ein **Vektor** ist eine gerichtete Strecke. Vektoren sind gleich, wenn sie durch Parallelverschiebung ineinander überführt werden können, d.h. gleiche Vektoren sind parallel, gleich lang und gleich gerichtet. Ein Vektor hat also eine bestimmte Richtung und Länge, aber keine bestimmte Lage.



In der Linearen Algebra gibt es eine allgemeinere und umfassendere Definition. Sie heißt kurz:

Ein **Vektor** ist Element eines Vektorraums.

Da der Vektorraum erst in einem späteren Kapitel eingeführt wird, begnügen wir uns hier mit einer knappen Begriffsbestimmung.

2.1.1.2 Kurze Einführung in Vektorräume

Ein Vektorraum besteht aus einer Menge von Elementen (den Vektoren) und zwei Rechenoperationen \odot und \oplus . Mit \oplus können zwei Vektoren addiert, mit \odot kann ein Vektor mit einem Skalar (je nach Vektorraum eine reelle oder komplexe Zahl) multipliziert werden. Die beiden Operatoren \odot und \oplus schreibt man, wenn es sich um abstrakte Operatoren handelt. Konkrete Rechenvorschriften, vor allem aber die kanonischen, werden dann mit „ \cdot “ und „ $+$ “ bezeichnet. Außerdem müssen für einen Vektorraum verschiedene Bedingungen erfüllt sein.

Definition 2.1: Eine Menge V , auf der die Operationen \oplus und \odot definiert sind, heißt **Vektorraum**, wenn für beliebige Vektoren $a, b, c \in V$ und beliebige Skalare α, β gilt:

1. Assoziativität:

$$(a \oplus b) \oplus c = a \oplus (b \oplus c).$$

2. Kommutivität:

$$a \oplus b = b \oplus a.$$

3. Distributivität:

$$(a) \quad \alpha \odot (a \oplus b) = \alpha \odot a \oplus \alpha \odot b.$$

$$(b) \quad (\alpha \oplus \beta) \odot a = \alpha \odot a \oplus \beta \odot a \text{ und } (\alpha\beta) \odot a = \alpha(\beta \odot a).$$

4. Es gibt einen Nullvektor 0 , für den gilt:

$$0 \oplus a = a \oplus 0 = a \text{ für alle } a.$$

Bemerkung: Verwechseln Sie nicht die Zahl 0 mit dem Nullvektor, der ebenfalls mit dem Symbol 0 bezeichnet wird. Zur Unterscheidung wird der Nullvektor aber teilweise auch mit $\vec{0}$ bezeichnet.

5. $1 \odot a = a$, und wenn wir mit $-a$ den Punkt $(-1) \odot a$ bezeichnen, dann gilt:

$$a \oplus (-a) = 0.$$

[Anstelle von $a \oplus (-b)$ schreiben wir $a \ominus b$.]

Mit dieser Art der Definition erhält man Vektoren mit sehr unerwartetem Aussehen. Zum Beispiel gibt es:

- den Vektorraum der auf dem Intervall $[1, 2]$ stetigen Funktionen. Ein Vektor daraus ist z.B. $\sin(x)$
- den Vektorraum der Polynome vom Grad ≤ 3 . Ein Vektor daraus ist z.B. $5x^2 + 7x - 8$.

- den Vektorraum der linearen Abbildungen $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ (siehe LA 2, Kapitel 4). Ein Vektor daraus ist z.B. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Wesentlich ist, dass die Rechenoperationen, die für diese grundsätzlich verschiedenen Objekte definiert werden können, einander entsprechen und den gleichen Gesetzen gehorchen. Wenn beim Beweis einer zusätzlichen Eigenschaft nur die Punkte 1-6 aus der Definition des Vektorraums benutzt werden, gilt die Eigenschaft für alle Vektorräume auf einmal. Dies trifft für einen großen Teil der Sätze in diesem Kapitel zu.

2.1.1.3 Der Vektorraum \mathbb{R}^n

Der \mathbb{R}^n ist ein n -dimensionaler Raum. Die Elemente dieses Vektorraums sind die n -**Tupel**, d.h. n -fache reelle Zahlen

$$(x_1, x_2, \dots, x_n),$$

wenn n eine natürliche Zahl ist. Die Zahlen x_1, \dots, x_n bezeichnen wir als die **Koordinaten**. Die Skalare, die für die Vektorraumaxiome benötigt werden, sind die reellen Zahlen. Die beiden Rechenoperationen sind mit $\vec{a}, \vec{b} \in \mathbb{R}^n$, $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$, $\alpha \in \mathbb{R}$:

Definition 2.2:

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

$$\alpha \cdot \vec{a} = \alpha \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \vdots \\ \alpha a_n \end{pmatrix}$$

Da die Bedingungen des Vektorraums auf diesen Raum zutreffen, ist er ein Vektorraum. Wir werden das später noch eingehend überprüfen.

Dies hat noch nichts mit einer geometrischen Darstellung zu tun, wie das folgende Beispiel zeigt.

Beispiel 2.3: In den Wirtschaftswissenschaften definiert man Räume mit Koordinaten, die die Gesamtproduktion von Waren oder Dienstleistungen wiedergeben, wobei als Maßstab der Preis (z.B. in 1.000.000 Dollar) verwendet wird.¹ Wir würden uns z. B. mit einem 7-dimensionalen Raum befassen, dessen Koordinaten sich auf die folgenden Industriezweige beziehen:

¹ Das entsprechende Verfahren heißt *Input-Output-Analyse* oder *Leontief-Modell*.

1. Stahl 2. Automobil 3. Landwirtschaft 4. Fisch
5. Chemie 6. Kleidung 7. Transport

Dann würde ein Punkt

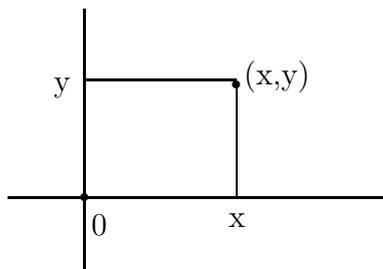
(1000, 800, 550, 300, 700, 200, 900)

in diesem 7-Raum bedeuten, dass die Stahlindustrie Waren in einem Wert von einer Milliarde Dollar und die Chemieindustrie von 700 Millionen Dollar produziert hat.

2.1.2 Der \mathbb{R}^n in der analytischen Geometrie

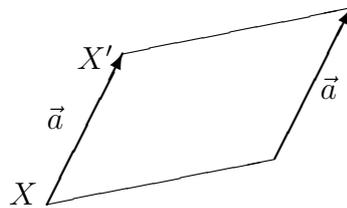
Das klassische Beispiel für einen dreidimensionalen Raum ist natürlich der Raum, in dem wir leben. Die meisten der Beispiele behandeln die Fälle $n = 2$ oder $n = 3$, die sich geometrisch veranschaulichen lassen. Ein n -Tupel kann auf die folgenden Arten geometrisch interpretiert werden:

- Ein n -Tupel kann mit Hilfe einer Basis (siehe Kapitel 3.2.2) als **Punkt** im n -dimensionalen Raum betrachtet werden. In diesem Kapitel wählen wir die kanonische Basis, d.h. die Koordinaten des Tupels geben die intuitiven Koordinaten eines Punktes wieder. Ein Beispiel im \mathbb{R}^2 lautet:



Punkte werden allgemein mit Großbuchstaben bezeichnet.

- Ein n -Tupel kann auch als **Verschiebung** betrachtet werden. Die Koordinaten des Tupels geben an, wie man von irgendeinem Punkt X zu seinem Bildpunkt X' kommt. Verschiebungen werden als **Verschiebungsvektor** oder allgemein als **Vektor** bezeichnet. Dies ist die eingangs erwähnte spezielle Bedeutung des Begriffs Vektor. Ein Verschiebungsvektor hat also eine bestimmte Richtung und Länge, aber keine bestimmte Lage, benötigt also auch keine Basis.



Vektoren aus dem \mathbb{R}^n werden gewöhnlich mit einem Kleinbuchstaben bezeichnet, über den ein Pfeil gesetzt ist (allgemeine Vektoren haben keinen Pfeil).

Ein Punkt (inkl. Basis) und ein Verschiebungsvektor sind formal gesehen die gleichen Objekte; man verbindet nur jeweils eine andere anschauliche Vorstellung mit ihnen.

2.1.2.1 Gebundene Vektoren und Ortsvektoren

Einen Vektor mit festem Anfangs- und Endpunkt nennt man **gebundenen** oder **lokalisierten Vektor** und bezeichnet ihn mit einem Buchstabenpaar, das den Anfangspunkt und den Endpunkt beschreibt, und einem Pfeil darüber, also z.B. \overrightarrow{PQ} . Ein spezieller gebundener Vektor ist der Vektor, der im **Ursprung lokalisiert** ist. Ihn nennt man auch **Ortsvektor** und schreibt $\overrightarrow{0A}$ oder einfach (wie bei „freien“ Verschiebungsvektoren) \vec{a} . Der Endpunkt A ist dabei das gleiche n -Tupel wie \vec{a} .

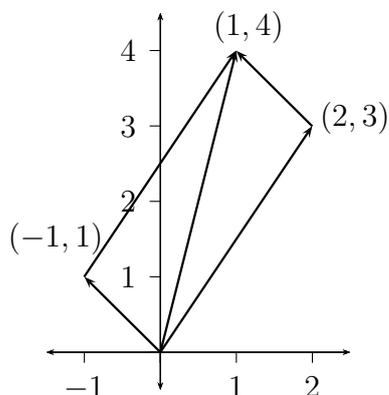
2.1.2.2 Geometrische Interpretation von Addition und Multiplikation im \mathbb{R}^n

Dazu betrachten wir uns einige Beispiele im \mathbb{R}^2 (gleichzeitig können Sie sich veranschaulichen, was im 3-dim. Raum geschieht). In den Grafiken sind jeweils die Ortsvektoren abgebildet.

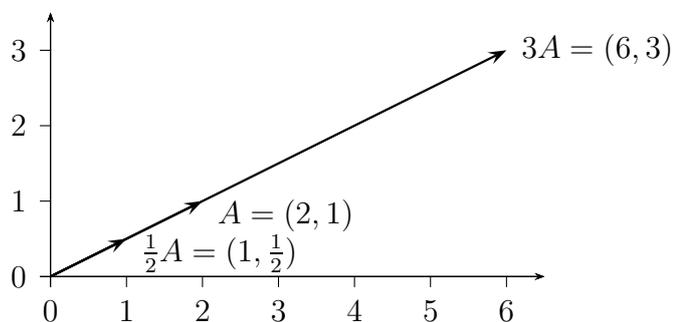
Beispiel 2.4: Gegeben sei $\vec{a} = (2, 3)$ und $\vec{b} = (-1, 1)$. Dann ist

$$\vec{a} + \vec{b} = (1, 4).$$

Grafisch ergibt die Addition das folgende Parallelogramm.



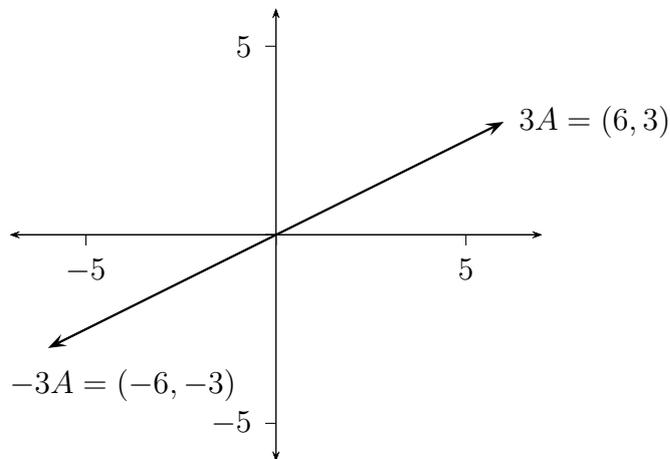
Welche Darstellung ergibt sich nach Multiplikation mit einer Zahl? Wenn $\vec{a} = (2, 1)$ und $\alpha = 3$ ist, dann ist $\alpha\vec{a} = (6, 3)$.



Die Multiplikation mit 3 führt dazu, dass \vec{a} um 3 gedehnt wird. Gemäß diesem Fall führt $\frac{1}{2}\vec{a}$ dazu, dass \vec{a} um die Hälfte "gedehnt" wird, d. h. \vec{a} verkürzt sich auf die Hälfte.

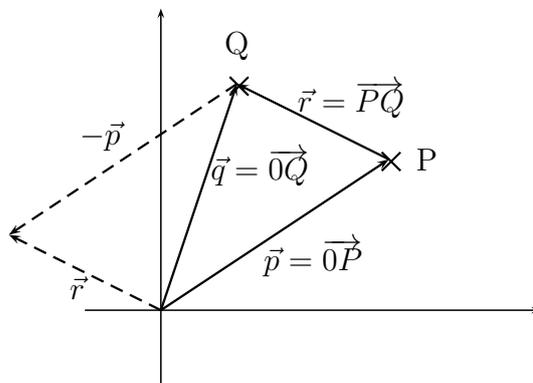
Im Allgemeinen interpretieren wir $t\vec{a}$ für eine positive Zahl $t > 0$ als einen Punkt mit derselben Richtung wie \vec{a} zum Ursprung, jedoch mit einem t -fachen Abstand.

Die Multiplikation mit einer negativen Zahl führt zu einer umgekehrten Richtung. Somit würde $-3\vec{a}$ wie in der folgenden Abbildung dargestellt.



2.1.2.3 Differenzvektoren

Bisher haben wir die Vektoraddition und -multiplikation behandelt. Über diese beiden Operationen lässt sich mit $\vec{a} - \vec{b} = \vec{a} + (-1) \cdot \vec{b}$ auch eine Subtraktion in bekannter Art und Weise definieren. Praktisch gesehen spielt die Subtraktion eine entscheidende Rolle. Folgende Grafik verdeutlicht dies:



Man erkennt, dass folgender Zusammenhang gilt:

$$\begin{aligned} \vec{q} = \vec{OQ} &= \vec{OP} + \vec{PQ} = \vec{p} + \vec{PQ} \\ \Leftrightarrow \vec{PQ} &= \vec{q} - \vec{p} \end{aligned}$$

Interpretiert man die Vektoren \vec{p} und \vec{q} als Punkte, so gibt die Differenz der beiden Vektoren den Verbindungsvektor der Punkte an. Da die Subtraktion bekanntlich nicht kommutativ ist, gibt $\vec{q} - \vec{p}$ den Verbindungsvektor von P nach Q an, während $\vec{p} - \vec{q}$ den Verbindungsvektor von Q nach P angibt und somit genau der um 180° gedrehte Vektor ist.

2.1.3 Parallele Vektoren

Definition 2.5: Zwei Vektoren \vec{a} und \vec{b} werden **parallel** genannt (Schreibweise: $\vec{a} \parallel \vec{b}$), wenn es eine Zahl $\alpha \neq 0$ gibt, so dass $\vec{a} = \alpha\vec{b}$ gilt.

Sie sollen die **gleiche Richtung** haben, wenn es eine Zahl $\alpha > 0$ gibt, so dass $\vec{a} = \alpha\vec{b}$, und **gegensätzliche Richtung** haben, wenn es eine Zahl $\alpha < 0$ gibt, so dass $\vec{a} = \alpha\vec{b}$ gilt.

Der Nullvektor ist zu jedem Vektor parallel.

Beispiel 2.6: Es seien $P = (3, 7)$ und $Q = (-4, 2)$ sowie $A = (5, 1)$ und $B = (-16, -14)$. Dann ist $Q - P = (-7, -5)$ und $B - A = (-21, -15)$. Somit ist \vec{PQ} parallel zu \vec{AB} , da $B - A = 3(Q - P)$. Da $3 > 0$ ist, sehen wir ebenfalls, dass \vec{PQ} und \vec{AB} dieselbe Richtung haben.

2.1.4 Das Skalarprodukt

Es ist selbstverständlich, dass wir während einer Diskussion Vektoren immer aus demselben n -dimensionalen Raum auswählen.

Definition 2.7: Eine Verknüpfung zwischen zwei Vektoren desselben n -dimensionalen Raums \mathbb{R}^n heißt **Skalarprodukt**, wenn folgende Bedingungen erfüllt sind:

SP0 : $\langle \vec{a}, \vec{b} \rangle \in \mathbb{R}$.

SP1 : Es gilt $\langle \vec{a}, \vec{b} \rangle = \langle \vec{b}, \vec{a} \rangle$.

SP2 : Wenn $\vec{a}, \vec{b}, \vec{c}$ drei Vektoren sind, dann ist

$$\langle \vec{a}, (\vec{b} + \vec{c}) \rangle = \langle \vec{a}, \vec{b} \rangle + \langle \vec{a}, \vec{c} \rangle.$$

SP3 : Wenn α eine reelle Zahl ist, dann ist

$$\langle \alpha\vec{a}, \vec{b} \rangle = \alpha\langle \vec{a}, \vec{b} \rangle = \langle \vec{a}, \alpha\vec{b} \rangle.$$

SP4 : Wenn $\vec{a} = \vec{0}$ der Nullvektor ist, dann ist $\langle \vec{a}, \vec{a} \rangle = 0$; sonst ist $\langle \vec{a}, \vec{a} \rangle > 0$.

Es gibt viele unterschiedliche Skalarprodukte im \mathbb{R}^n . Das bei weitem wichtigste ist das **Standardskalarprodukt** oder **euklidische Skalarprodukt**. Im Folgenden werden wir nur noch das Standardskalarprodukt betrachten und es einfach als „Skalarprodukt“ bezeichnen.

Definition 2.8: Gegeben seien zwei Vektoren $\vec{a} = (a_1, \dots, a_n)$ und $\vec{b} = (b_1, \dots, b_n)$. Wir definieren ihr **(Standard-)Skalar-** bzw. **Punktprodukt** $\langle \vec{a}, \vec{b} \rangle$ als

$$\langle \vec{a}, \vec{b} \rangle = a_1 b_1 + \dots + a_n b_n$$

Schreibweisen: Oft wird das Skalarprodukt $\vec{a} \cdot \vec{b}$ geschrieben. In der Mathematik wird jedoch am häufigsten $\langle \vec{a}, \vec{b} \rangle$ verwendet.

Eine spezielle Schreibweise ist $\vec{a}^2 = \langle \vec{a}, \vec{a} \rangle$. Höhere Potenzen, wie \vec{a}^3 haben keine Bedeutung.

Beispiel 2.9: Für

$$\vec{a} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \text{ und } \vec{b} = \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix}$$

gilt:

$$\langle \vec{a}, \vec{b} \rangle = -1 + 12 + 6 = 17.$$

Für diesen Moment geben wir noch keine geometrische Interpretation für dieses Skalarprodukt. Dazu kommen wir später. Zunächst zeigen wir, dass das Standardskalarprodukt wirklich ein Skalarprodukt ist.

BEWEIS Da $a_1 b_1 + \dots + a_n b_n \in \mathbb{R}$, ist **SP0** erfüllt.

In Bezug auf **SP1** stellen wir fest, dass

$$a_1 b_1 + \dots + a_n b_n = b_1 a_1 + \dots + b_n a_n,$$

da für alle Zahlen a, b gilt $ab = ba$.

Für **SP2** soll $\vec{c} = (c_1, \dots, c_n)$ gelten. Dann ist

$$\vec{b} + \vec{c} = (b_1 + c_1, \dots, b_n + c_n)$$

und

$$\langle \vec{a}, \vec{b} + \vec{c} \rangle = a_1(b_1 + c_1) + \dots + a_n(b_n + c_n) = a_1 b_1 + a_1 c_1 + \dots + a_n b_n + a_n c_n.$$

Eine Neuordnung der Terme führt zu

$$a_1 b_1 + \dots + a_n b_n + a_1 c_1 + \dots + a_n c_n,$$

was nichts anderes bedeutet als $\langle \vec{a}, \vec{b} \rangle + \langle \vec{a}, \vec{c} \rangle$. Dies beweist **SP2**.

Die Gesetzmäßigkeit zu **SP3** überlassen wir als Übung.

Zuletzt prüfen wir **SP4** und beobachten, dass wenn eine Koordinate a_i von \vec{a} ungleich 0 ist, dann ist der Term $a_i^2 \neq 0$ und $a_i^2 > 0$ im Skalarprodukt

$$\langle \vec{a}, \vec{a} \rangle = a_1^2 + \cdots + a_n^2.$$

Da jeder Ausdruck ≥ 0 ist, folgern wir, dass die Summe > 0 ist, wie gezeigt werden sollte. ■

2.1.5 Die Norm eines Vektors

Auch hier gibt es eine allgemeine Definition, die auch in der Analysis behandelt wird.

Definition 2.10: Im \mathbb{R}^n müssen für die **Norm** eines Vektors folgende Bedingungen erfüllt sein:

N0 : $\|\vec{x}\| \in \mathbb{R}$.

N1 : $\|\vec{x}\| \geq 0$ für alle $\vec{x} \in \mathbb{R}^n$.

N2 : $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = 0$.

N3 : $\|\lambda\vec{x}\| = |\lambda|\|\vec{x}\|$ für alle $\vec{x} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

N4 : (Dreiecksungleichung) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$

Die im Folgenden definierte Standardnorm hat in der Analytischen Geometrie eine herausragende Bedeutung:

Definition 2.11: Ist ein geeignetes Skalarprodukt definiert, so erhält man die **Standardnorm**

$$\|\vec{a}\| = \sqrt{\langle \vec{a}, \vec{a} \rangle}$$

eines Vektors als die Wurzel des Skalarproduktes des Vektors mit sich selbst. Da $\langle \vec{a}, \vec{a} \rangle \geq 0$ ist, können wir die Wurzel ziehen. Bei einer Definition durch Koordinaten sehen wir, dass

$$\|\vec{a}\| = \sqrt{a_1^2 + \cdots + a_n^2}.$$

Diese Norm stimmt, wenn $n = 2$ oder $n = 3$ ist, mit unserer intuitiven Bezeichnung (hergeleitet durch den Satz des Pythagoras) der Länge eines Vektors überein.

BEWEIS Diese Definition ist tatsächlich eine Norm. **N0** folgt direkt aus **SP0**, ebenso folgen **N1** und **N2** direkt aus **SP4**. **N3** folgt aus **SP3** mit

$$\|\lambda\vec{x}\| = \sqrt{\langle \lambda\vec{x}, \lambda\vec{x} \rangle} = \sqrt{\lambda^2 \langle \vec{x}, \vec{x} \rangle} = \sqrt{\lambda^2} \sqrt{\langle \vec{x}, \vec{x} \rangle} = |\lambda| \|\vec{x}\|$$

N4 werden wir später beweisen. ■

Auch hier werden wir in den folgenden Kapiteln mit „der Norm“ immer die Standardnorm meinen, mit „einer Norm“ jedoch eine unbestimmte Norm. Zunächst werfen wir einen kurzen Überblick auf unterschiedliche Normen im \mathbb{R}^n und ihre Schreibweisen.

- Die Standardnorm heißt auch **euklidische Norm**, **Zweiernorm** oder **Betrag** eines Vektors. Schreibweisen sind $\|\vec{x}\|$, $\|\vec{x}\|_2$ oder $|\vec{x}|$.
- Die **Betragssummennorm** oder **Einernorm** ist definiert als

$$\|\vec{x}\|_1 = |a_1| + \dots + |a_n|.$$

- Die **Maximumnorm** ist definiert als

$$\|\vec{x}\|_\infty = \max\{|a_1|, \dots, |a_n|\}.$$

Beispiel 2.12: Es sei $\vec{b} = (-1, 2, 3)$. Dann ist

$$\begin{aligned} \|\vec{b}\| &= \sqrt{1 + 4 + 9} = \sqrt{14} \\ \|\vec{b}\|_1 &= 1 + 2 + 3 = 6 \\ \|\vec{b}\|_\infty &= \max\{1, 2, 3\} = 3 \end{aligned}$$

2.1.5.1 Einheitsvektoren

Definition 2.13: Wir sagen, dass ein Vektor \vec{e} ein **Einheitsvektor** ist, wenn $\|\vec{e}\| = 1$ ist.

Wichtige Einheitsvektoren sind die **kanonischen Einheitsvektoren**, also Einheitsvektoren in Richtung der Koordinatenachsen:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

bzw im \mathbb{R}^3 :

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

2.1.5.2 Normieren von Vektoren

Gegeben ist irgendein Vektor \vec{a} und sei $a = \|\vec{a}\|$. Ist $a \neq 0$, dann ist der **normierte** Vektor

$$\vec{e}_a = \frac{1}{a} \vec{a}$$

ein Einheitsvektor, da

$$\left\| \frac{1}{a} \vec{a} \right\| = \frac{1}{a} a = 1.$$

Beispiel 2.14: Es soll Vektor $\vec{a} = (1, 2, -3)$ normiert werden. Da $\|\vec{a}\| = \sqrt{14}$, ist $\vec{e}_a = \frac{1}{\sqrt{14}}(1, 2, -3)$.

Bemerkung 2.15: Die Vektoren v und $\lambda \cdot v$ ($\lambda > 0$) zeigen in die gleiche Richtung. Normiert sind beide Vektoren gleich. Man kann beim Normieren also als ersten Schritt einen positiven Vorfaktor des Vektors streichen.

Beispiel 2.16: Normiert man den Vektor

$$\vec{b} = \frac{3}{\sqrt{8}} \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix},$$

kann man statt dessen den Vektor

$$\vec{b}' = \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix}$$

normieren und erhält in beiden Fällen das Ergebnis

$$\frac{\vec{b}}{\|\vec{b}\|} = \frac{\vec{b}'}{\|\vec{b}'\|} = \frac{1}{\sqrt{53}} \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix}.$$

2.1.5.3 Orthogonale (senkrechte) Vektoren

Bemerkung 2.17: Ein Skalarprodukt $\langle \vec{a}, \vec{b} \rangle$ kann sehr wohl gleich 0 sein, ohne dass \vec{a} oder \vec{b} der Nullvektor ist. Wenn z. B.

$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \text{ und } \vec{b} = \begin{pmatrix} 2 \\ 1 \\ -4/3 \end{pmatrix}$$

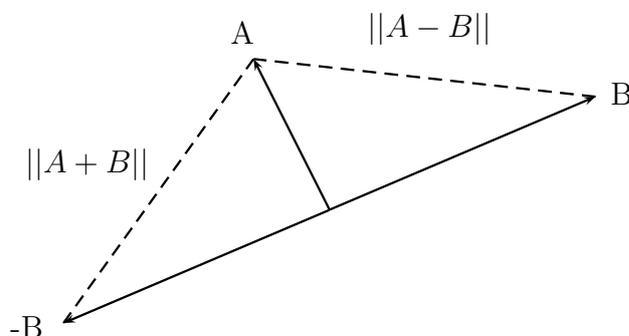
ist, dann ist $\langle \vec{a}, \vec{b} \rangle = 0$.

Definition 2.18: Wir definieren zwei Vektoren als **senkrecht** (oder auch **orthogonal**, Schreibweise $\vec{a} \perp \vec{b}$), wenn $\langle \vec{a}, \vec{b} \rangle = 0$ ist.

Wir müssen allerdings überprüfen, ob diese Definition in der Ebene mit unserer intuitiven geometrischen Beschreibung von „senkrecht“ übereinstimmt, d.h. ob der eingeschlossene Winkel 90° beträgt. Gegeben sind \vec{a}, \vec{b} in der Ebene; die Eigenschaft

$$\|\vec{a} + \vec{b}\| = \|\vec{a} - \vec{b}\|$$

beruht auf dem geometrischen Gesetz, dass \vec{a} senkrecht zu \vec{b} ist.



Satz 2.19: Für alle $\vec{a}, \vec{b} \in \mathbb{R}^n$ gilt

$$\|\vec{a} + \vec{b}\| = \|\vec{a} - \vec{b}\| \Leftrightarrow \langle \vec{a}, \vec{b} \rangle = 0.$$

BEWEIS

$$\begin{aligned} & \|\vec{a} + \vec{b}\| = \|\vec{a} - \vec{b}\| \\ \Leftrightarrow & \|\vec{a} + \vec{b}\|^2 = \|\vec{a} - \vec{b}\|^2 \\ \Leftrightarrow & \vec{a}^2 + 2\langle \vec{a}, \vec{b} \rangle + \vec{b}^2 = \vec{a}^2 - 2\langle \vec{a}, \vec{b} \rangle + \vec{b}^2 \\ \Leftrightarrow & 4\langle \vec{a}, \vec{b} \rangle = 0 \\ \Leftrightarrow & \langle \vec{a}, \vec{b} \rangle = 0. \end{aligned}$$

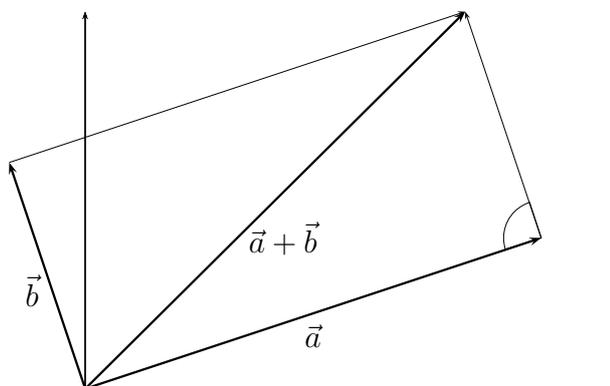
■

2.1.5.4 Satz des Pythagoras

Satz 2.20: Wenn \vec{a} und \vec{b} aus \mathbb{R}^n senkrecht (orthogonal) zueinander sind ($\vec{a} \perp \vec{b}$), dann gilt:

$$\|\vec{a} + \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2.$$

Der Satz wird in der folgenden Abbildung dargestellt.



BEWEIS

$$\begin{aligned} \|\vec{a} + \vec{b}\|^2 &= \|\vec{a}\|^2 + 2\langle \vec{a}, \vec{b} \rangle + \|\vec{b}\|^2 \\ &= \|\vec{a}\|^2 + \|\vec{b}\|^2 \\ \text{da } \langle \vec{a}, \vec{b} \rangle &= 0 \text{ und } \|\vec{a}\|^2 = \langle \vec{a}, \vec{a} \rangle \text{ definiert ist.} \end{aligned}$$

■

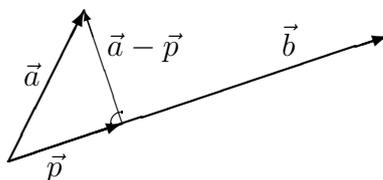
Bemerkung 2.21: Wenn \vec{a} senkrecht zu \vec{b} ist, dann sind auch $\alpha\vec{a}$ und $\beta\vec{b}$ für beliebige reelle Zahlen $\alpha, \beta \in \mathbb{R}$ orthogonal zueinander, weil

$$\langle \alpha\vec{a}, \beta\vec{b} \rangle = \alpha\beta\langle \vec{a}, \vec{b} \rangle = 0.$$

2.1.6 Der Winkel zwischen zwei Vektoren

2.1.6.1 Die senkrechte Projektion eines Vektors auf einen anderen

Wir benutzen jetzt die Definition von “senkrecht”, um die Definition der (senkrechten) “Projektion” abzuleiten. \vec{a}, \vec{b} seien zwei Vektoren und $\vec{b} \neq \vec{0}$. Wir möchten die Projektion \vec{p} von \vec{a} in Richtung von \vec{b} definieren (s. Abb.)



Wir bestimmen einen Vektor \vec{p} , so dass gilt:

1. $\vec{p} \parallel \vec{b} \quad \Rightarrow \quad \vec{p} = \alpha\vec{b}$
2. $(\vec{a} - \vec{p}) \perp \vec{b} \quad \Rightarrow \quad \langle \vec{a} - \vec{p}, \vec{b} \rangle = 0$

Daraus folgt:

$$\begin{aligned} \langle \vec{a} - \alpha \vec{b}, \vec{b} \rangle &= 0 \\ \Leftrightarrow \langle \vec{a}, \vec{b} \rangle - \alpha \langle \vec{b}, \vec{b} \rangle &= 0 \\ \Leftrightarrow \alpha &= \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \\ \Rightarrow \vec{p} &= \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \cdot \vec{b} \end{aligned}$$

\vec{p} nennt man die „**Projektion** von \vec{a} in Richtung \vec{b} “ oder die „**Komponente** von \vec{a} entlang \vec{b} “.

Die Norm der Projektion ist:

$$\|\vec{p}\| = \left\| \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \cdot \vec{b} \right\| = \left| \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \right| \cdot \|\vec{b}\| = \frac{|\langle \vec{a}, \vec{b} \rangle|}{\|\vec{b}\|^2} \cdot \|\vec{b}\| = \frac{|\langle \vec{a}, \vec{b} \rangle|}{\|\vec{b}\|}$$

Bemerkung 2.22: Ist \vec{b} ein Einheitsvektor, so gilt:

$$\vec{p} = \langle \vec{a}, \vec{b} \rangle \cdot \vec{b}.$$

Beispiel 2.23: Es sei

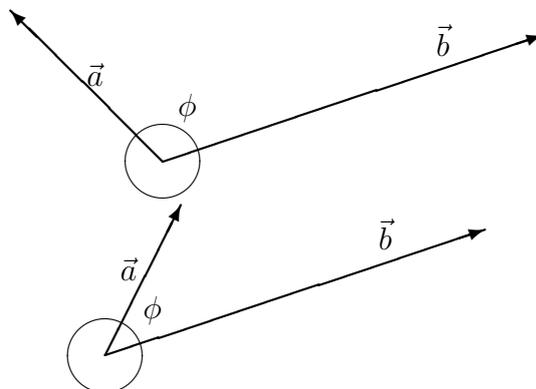
$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Dann ist die Komponente von \vec{a} entlang \vec{b} :

$$\vec{p} = \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \cdot \vec{b} = \frac{1 + 2 - 6}{1 + 1 + 4} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ -1 \end{pmatrix}.$$

2.1.6.2 Der Winkel zwischen zwei Vektoren

Definition 2.24: Unter dem Winkel ϕ zwischen zwei Vektoren \vec{a} und \vec{b} versteht man den kleineren der Winkel zwischen den entsprechenden Ortsvektoren.



Es gilt:

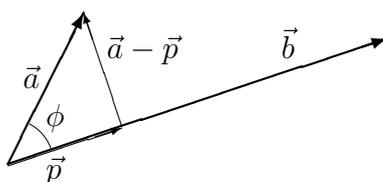
Satz 2.25:

$$\cos \phi = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{b}\|^2} \cdot \frac{\|\vec{b}\|}{\|\vec{a}\|} = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|}$$

$$\langle \vec{a}, \vec{b} \rangle = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos \phi$$

BEWEIS Es sei $\vec{p} = \alpha \vec{b} = \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \vec{b}$ die Projektion von \vec{a} in Richtung \vec{b} . Wir treffen die Fallunterscheidung, ob der Winkel zwischen \vec{a} und \vec{b} spitz ($0 \leq \phi \leq \frac{\pi}{2}$) oder stumpf ($\frac{\pi}{2} < \phi \leq \pi$) ist.

Spitzer Winkel: \vec{b} und \vec{p} zeigen in die gleiche Richtung. Aus $\vec{p} = \alpha \vec{b}$ folgt $\alpha > 0$.

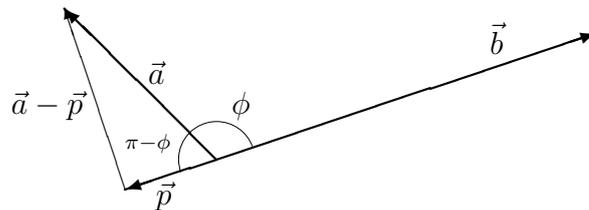


Aus der Ebenegeometrie folgt:

$$\cos \phi = \frac{\|\vec{p}\|}{\|\vec{a}\|} = \frac{|\alpha| \|\vec{b}\|}{\|\vec{a}\|} = \frac{\alpha \|\vec{b}\|}{\|\vec{a}\|}$$

wobei der letzte Schritt gilt, da $\alpha \geq 0$

Stumpfer Winkel: \vec{b} und \vec{p} zeigen in entgegengesetzte Richtungen. Daraus folgt $\alpha < 0$.



Aus der Ebenegeometrie folgt:

$$\cos(\pi - \phi) = \frac{\|\vec{p}\|}{\|\vec{a}\|} = \frac{|\alpha| \|\vec{b}\|}{\|\vec{a}\|} = -\frac{\alpha \|\vec{b}\|}{\|\vec{a}\|}$$

$$\cos \phi = -\cos(\pi - \phi) = \frac{\alpha \|\vec{b}\|}{\|\vec{a}\|}$$

Damit gilt für beide Fälle:

$$\cos \phi = \frac{\alpha \|\vec{b}\|}{\|\vec{a}\|} = \frac{\langle \vec{a}, \vec{b} \rangle \|\vec{b}\|}{\langle \vec{b}, \vec{b} \rangle \|\vec{a}\|} = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|}$$

■

Beispiel 2.26: Es seien $A = (1, 2, -3)$ und $B = (2, 1, 5)$. Bestimmen Sie den Cosinus des Winkels θ zwischen \vec{a} und \vec{b} .

Anhand der Definition muss gelten:

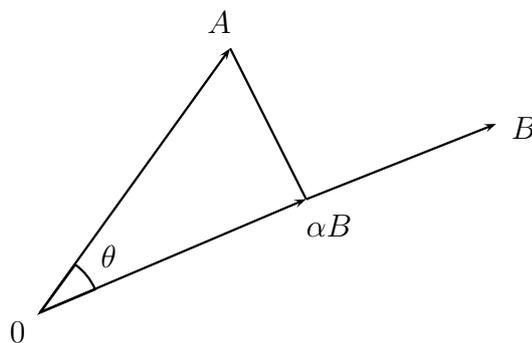
$$\cos \theta = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|} = \frac{2 + 2 - 15}{\sqrt{14} \sqrt{30}} = \frac{-11}{\sqrt{420}} = \frac{-11\sqrt{105}}{210}$$

$$\theta \approx 122,4^\circ$$

Beispiel 2.27: Bestimmen Sie den Cosinus des Winkels zwischen den zwei lokalisierten Vektoren \vec{PQ} und \vec{PR} , wo

$$P = (1, 2, -3), Q = (-2, 1, 5), R = (1, 1, -4).$$

Die Grafik sieht wie folgt aus:



Es gilt

$$A = Q - P = (-3, -1, 8) \text{ und } B = R - P = (0, -1, -1).$$

Dann ist der Winkel zwischen \overrightarrow{PQ} und \overrightarrow{PR} der gleiche wie der zwischen $\vec{a} = \overrightarrow{0A}$ und $\vec{b} = \overrightarrow{0B}$. Somit ist sein Cosinus äquivalent zu

$$\cos \theta = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|} = \frac{0 + 1 - 8}{\sqrt{74} \sqrt{2}} = \frac{-7}{\sqrt{74} \sqrt{2}}.$$

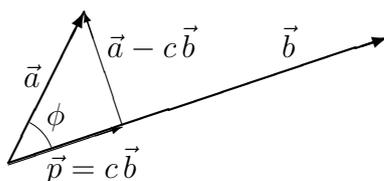
2.1.7 Weitere Ungleichungen

2.1.7.1 Die Schwarzsche Ungleichung

Satz 2.28: (Schwarzsche Ungleichung): \vec{a}, \vec{b} seien zwei Vektoren im \mathbb{R}^n . Dann gilt

$$|\langle \vec{a}, \vec{b} \rangle| \leq \|\vec{a}\| \cdot \|\vec{b}\|.$$

BEWEIS



Wenn $\vec{b} = 0$ ist, dann sind beide Seiten der Ungleichung gleich 0, und somit ist unsere Behauptung offensichtlich gültig. Es sei jetzt $\vec{b} \neq 0$. Dann folgt aus der Zeichnung und dem Satz des Pythagoras:

$$\begin{aligned} \|\vec{a}\| &\geq \|\vec{p}\| = \left\| \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \vec{b} \right\| = \frac{|\langle \vec{a}, \vec{b} \rangle|}{\|\vec{b}\|} \\ \Leftrightarrow |\langle \vec{a}, \vec{b} \rangle| &\leq \|\vec{a}\| \cdot \|\vec{b}\| \end{aligned}$$

■

Da gilt

$$\langle \vec{a}, \vec{b} \rangle = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos \theta,$$

bedeutet die Schwarzsche Ungleichung nicht anderes als:

$$|\cos \theta| \leq 1.$$

2.1.7.2 Die Dreiecksungleichung

Eine Folgerung aus der Schwarzschen Ungleichung ist die Dreiecksungleichung:

Satz 2.29: (Dreiecksungleichung): \vec{a} und \vec{b} seien Vektoren. Dann gilt

$$\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|.$$

BEWEIS Beide Seiten dieser Ungleichung sind positiv oder 0. Daher genügt es, zu beweisen, dass ihre Quadrate die gewünschte Ungleichung erfüllen d.h.

$$\langle \vec{a} + \vec{b}, \vec{a} + \vec{b} \rangle \leq (\|\vec{a}\| + \|\vec{b}\|)^2.$$

Wir sehen, dass

$$\langle \vec{a} + \vec{b}, \vec{a} + \vec{b} \rangle = \langle \vec{a}, \vec{a} \rangle + 2\langle \vec{a}, \vec{b} \rangle + \langle \vec{b}, \vec{b} \rangle.$$

Wegen unseres vorherigen Ergebnisses erfüllt die rechte Seite die Ungleichung

$$\leq \|\vec{a}\|^2 + 2\|\vec{a}\| \|\vec{b}\| + \|\vec{b}\|^2$$

und dies ist genau

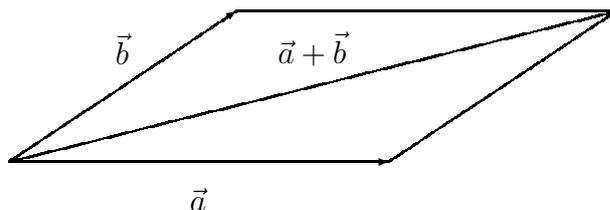
$$(\|\vec{a}\| + \|\vec{b}\|)^2.$$

Also gilt

$$(\|\vec{a} + \vec{b}\|)^2 \leq (\|\vec{a}\| + \|\vec{b}\|)^2$$

und damit ist unser Satz ist bewiesen. ■

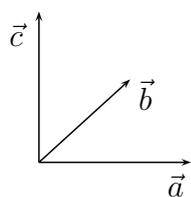
Der Grund für den Namen **Dreiecksungleichung** ist der folgende: Wenn wir ein Dreieck wie in der folgenden Abbildung zeichnen, dann sagt der Satz, dass die Länge einer Seite kleiner gleich der Summe aus den Längen der beiden anderen Seiten ist.



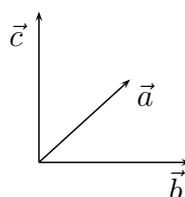
2.1.8 Das Vektorprodukt (Kreuzprodukt)

Die Ergebnisse dieses Abschnitts gelten im Gegensatz zu den bisherigen Ergebnissen nur im \mathbb{R}^3 .

Definition 2.30: Gegeben seien 3 Vektoren \vec{a} , \vec{b} und \vec{c} . \vec{a} und \vec{b} liegen in einer Ebene und umschließen einen Winkel. Wenn man \vec{a} auf kürzestem Wege auf \vec{b} dreht, geschieht das entweder im oder gegen den Uhrzeigersinn. Ist der kürzeste Weg im Uhrzeigersinn und zeigt \vec{c} von der Ebene aus gesehen nach unten, so bilden \vec{a} , \vec{b} und \vec{c} ein **Rechtssystem**.



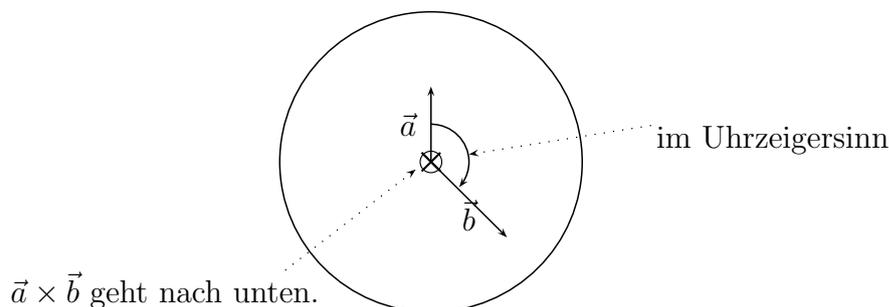
Rechtssystem



Linkssystem

Dazu kann man sich folgende Eselsbrücken merken:

1. Rechte-Hand-Regel: Man zeigt mit dem Daumen der rechten Hand in Richtung des ersten Vektors \vec{a} , während man versucht mit dem Zeigefinger in Richtung des zweiten Vektors \vec{b} zu zeigen. Dann zeigt der Mittelfinger in die Richtung von \vec{c} , wenn er senkrecht zu Daumen und Zeigefinger gehalten wird. Bei stumpfen Winkeln zwischen \vec{a} und \vec{b} widerspricht diese Haltung zwar der menschlichen Anatomie, aber wichtig ist hier nur, dass man mit Daumen und Zeigefinger den kleineren (inneren) der beiden Winkel zwischen \vec{a} und \vec{b} darzustellen versucht.
2. Für Heimwerker: Wenn man die Schraube (mit gewöhnlichem Rechtsgewinde) im Uhrzeigersinn dreht, geht sie nach unten.

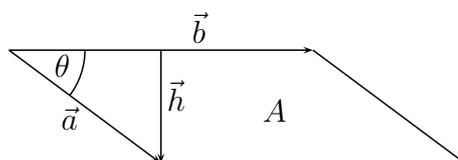


Definition 2.31: Das **Vektorprodukt**, oder auch **Kreuzprodukt**, $\vec{a} \times \vec{b}$ zweier Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^3$ ist für den Fall $\vec{a} = \lambda \vec{b}$ als der Nullvektor definiert, anderenfalls über folgende Eigenschaften:

1. $\vec{a} \times \vec{b}$ steht sowohl senkrecht auf \vec{a} als auch auf \vec{b} .
2. \vec{a}, \vec{b} und $\vec{a} \times \vec{b}$ bilden ein Rechtssystem.
3. Die Länge von $\vec{a} \times \vec{b}$ ist gleich dem Flächeninhalt des von \vec{a} und \vec{b} aufgespannten Parallelogramms, d.h.

$$\|\vec{a} \times \vec{b}\| = \|\vec{a}\| \|\vec{b}\| \sin \theta$$

Für den Fall $\vec{a} = \lambda \vec{b}$ gilt $\vec{a} \times \vec{b} = 0$.



$$A = \|\vec{b}\| \cdot \|\vec{h}\| = \|\vec{b}\| \cdot \|\vec{a}\| \cdot \sin \theta$$

Satz 2.32: Aus der Definition des Kreuzprodukts leiten sich folgende Eigenschaften her:

1. Antikommutativität: $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$.
2. Distributivität:
 - $\vec{a} \times (\vec{b} + \vec{c}) = (\vec{a} \times \vec{b}) + (\vec{a} \times \vec{c})$
 - $(\vec{b} + \vec{c}) \times \vec{a} = (\vec{b} \times \vec{a}) + (\vec{c} \times \vec{a})$
3. Für jede reelle Zahl α gilt: $(\alpha \vec{a}) \times \vec{b} = \alpha(\vec{a} \times \vec{b}) = \vec{a} \times (\alpha \vec{b})$.
4. $\vec{a} \times (\vec{b} \times \vec{c}) \neq (\vec{a} \times \vec{b}) \times \vec{c}$ (Keine Assoziativität)
5. $(\vec{a} \times \vec{b})^2 = \langle \vec{a}, \vec{a} \rangle \langle \vec{b}, \vec{b} \rangle - \langle \vec{a}, \vec{b} \rangle^2$

Aus diesen Eigenschaften lässt sich eine einfache Darstellung ableiten:

Satz 2.33: Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$, dann gilt:

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

BEWEIS Da unser kanonisches Koordinatensystem ein Rechtssystem ist, gilt:

$$\begin{aligned}
 \vec{a} \times \vec{b} &= (a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3) \times (b_1 \vec{e}_1 + b_2 \vec{e}_2 + b_3 \vec{e}_3) \\
 &= a_1 b_1 \underbrace{\vec{e}_1 \times \vec{e}_1}_0 + a_1 b_2 \underbrace{\vec{e}_1 \times \vec{e}_2}_{\vec{e}_3} + a_1 b_3 \underbrace{\vec{e}_1 \times \vec{e}_3}_{-\vec{e}_2} \\
 &\quad + a_2 b_1 \underbrace{\vec{e}_2 \times \vec{e}_1}_{-\vec{e}_3} + a_2 b_2 \underbrace{\vec{e}_2 \times \vec{e}_2}_0 + a_2 b_3 \underbrace{\vec{e}_2 \times \vec{e}_3}_{\vec{e}_1} \\
 &\quad + a_3 b_1 \underbrace{\vec{e}_3 \times \vec{e}_1}_{\vec{e}_2} + a_3 b_2 \underbrace{\vec{e}_3 \times \vec{e}_2}_{-\vec{e}_1} + a_3 b_3 \underbrace{\vec{e}_3 \times \vec{e}_3}_0 \\
 &= \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}
 \end{aligned}$$

■

Beispiel 2.34: $\vec{a} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$ dann ist

$$\vec{a} \times \vec{b} = \begin{pmatrix} 16 \\ -9 \\ 5 \end{pmatrix}.$$

Es gibt 2 Schemata, mit denen sich die Berechnung des Kreuzproduktes leicht merken lässt:

Erste Möglichkeit:

1. Komponente	2. Komponente	3. Komponente
$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} + & - \\ \swarrow & \searrow \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} - & + \\ \swarrow & \searrow \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} + & - \\ \swarrow & \searrow \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$
$15 - (-1) = 16$	$-10 + 1 = -9$	$2 - (-3) = 5$

Entlang der Pfeile wird überkreuz multipliziert und entsprechend der Vorzeichen addiert bzw. subtrahiert.

Zweite Möglichkeit:

$$\ominus \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} \swarrow & \searrow \\ \swarrow & \searrow \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix} \oplus \begin{matrix} -1 & \leftarrow & -1 & \rightarrow & 15 \\ 10 & \leftarrow & 2 & \rightarrow & 1 \\ -3 & \leftarrow & 3 & \rightarrow & 2 \end{matrix} \Rightarrow \begin{pmatrix} 16 \\ -9 \\ 5 \end{pmatrix}$$

Unter beide Vektoren werden die ersten beiden Komponenten des jeweiligen Vektors geschrieben. Dann werden die Komponenten entlang der schrägen Pfeile multipliziert und die Produkte am rechten und linken Rand notiert. Schließlich werden die Produkte am linken Rand von denen am rechten Rand abgezogen.

Bemerkung 2.35: Auch wenn das Vektorprodukt nur für Vektoren aus dem \mathbb{R}^3 definiert ist, so lässt es sich dennoch auch im \mathbb{R}^n für $n > 3$ nutzen, wenn man alle bis auf drei Komponenten der Vektoren wie 0 behandelt.

Beispiel 2.36: Mit Hilfe des Kreuzproduktes wird zu den Vektoren

$$\vec{a} = \begin{pmatrix} 2 \\ 4 \\ 3 \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 5 \end{pmatrix}$$

mit Hilfe der ersten, dritten und vierten Komponente ein orthogonaler Vektor berechnet, wenn die zweite Komponente auf 0 gesetzt wird:

$$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \\ -9 \\ 5 \end{pmatrix}, \quad \text{also ist} \quad \vec{c} = \begin{pmatrix} 16 \\ 0 \\ -9 \\ 5 \end{pmatrix}$$

senkrecht zu \vec{a} und \vec{b} .

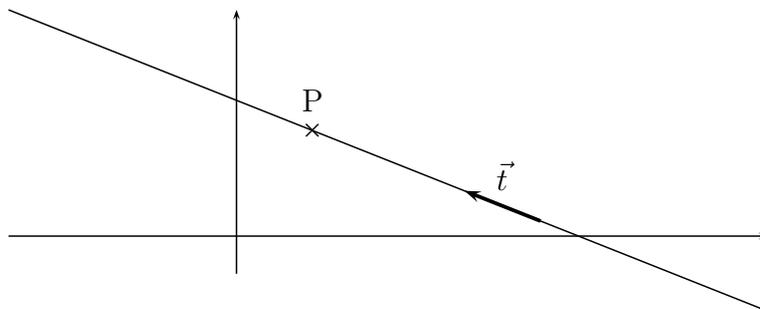
2.2 Geraden und Ebenen

2.2.1 Geradengleichung im \mathbb{R}^2

Eine Geradengleichung im \mathbb{R}^2 kann unterschiedliche Formen annehmen. Zunächst unterscheidet man zwischen Parametergleichungen und parameterlosen Gleichungen.

2.2.1.1 Parametergleichungen

- Ein Punkt und der Richtungsvektor der Geraden sind bekannt \Rightarrow **Punkt-Richtungsgleichung**

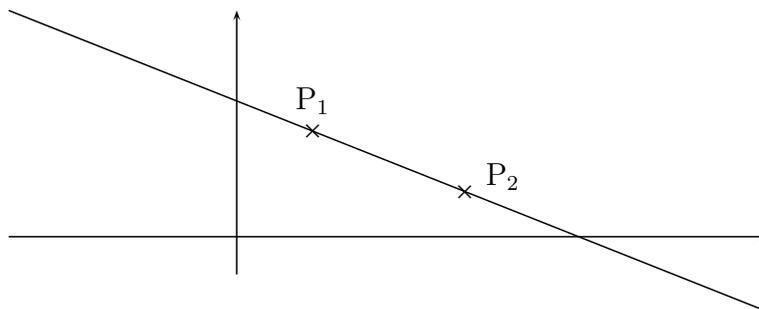


Jeder Punkt auf der Geraden lässt sich beschreiben als

$$\vec{x} = \vec{p} + \gamma \vec{t}$$

γ ist dabei der Parameter der Parametergleichung.

- Zwei Punkte auf der Geraden sind bekannt \Rightarrow **Zweipunktgleichung**



Da sich ein Richtungsvektor aus $\vec{t} = \vec{p}_2 - \vec{p}_1$ gewinnen lässt, erhält man die Geradengleichung mit

$$\vec{x} = \vec{p}_1 + \gamma \cdot (\vec{p}_2 - \vec{p}_1)$$

2.2.1.2 Parameterlose Gleichungen

Aus dem Mathematik-Blockkurs sind schon einige parameterlose Gleichungen bekannt:

- Die Geradengleichung durch zwei Punkte,
- die Gleichung einer Geraden mit vorgegebener Steigung,
- die Achsenabschnittsform einer Geraden
- und die allgemeine Geradengleichung.

Uns interessiert von diesen Gleichungen nur die allgemeine Geradengleichung, die wir im Folgenden Normalform nennen:

Definition 2.37: Eine Gerade im \mathbb{R}^2 ist eindeutig durch ihre **Normalform**

$$ax + by = c$$

beschrieben. Ist \vec{p} ein bekannter Punkt auf der Gerade, so kann die Normalform auch in der Form

$$\langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle$$

dargestellt werden, wobei \vec{n} ein senkrecht zur Gerade stehender **Normalenvektor** und \vec{x} ein beliebiger (unbekannter) Punkt auf der Geraden ist.

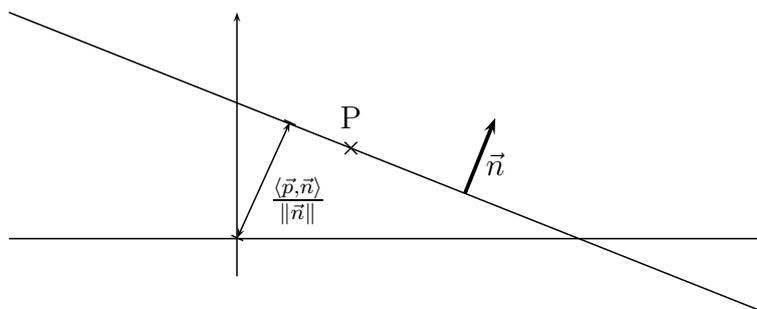
Bemerkung 2.38: Die beiden Darstellungen der Normalform stellen die selbe Gerade dar, d.h. $\langle \vec{x}, \vec{n} \rangle$ ist für alle Punkte \vec{x} auf der Geraden konstant.

BEWEIS Wähle $a = n_1, b = n_2$ und $c = \langle \vec{p}, \vec{n} \rangle$. Bleibt nur noch zu zeigen, dass $\langle \vec{x}, \vec{n} \rangle$ wirklich für alle Punkte auf der Geraden konstant ist:

Für die Gerade gilt: $\vec{x} = \vec{p} + \gamma \cdot \vec{t}$. Multiplikation mit einem Normalenvektor \vec{n} ergibt:

$$\begin{aligned} \langle \vec{x}, \vec{n} \rangle &= \langle \vec{p} + \gamma \vec{t}, \vec{n} \rangle \\ &= \langle \vec{p}, \vec{n} \rangle + \gamma \underbrace{\langle \vec{t}, \vec{n} \rangle}_{=0} \\ &= \langle \vec{p}, \vec{n} \rangle \end{aligned}$$

■



Beispiel 2.39: Es sei $\vec{p} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $\vec{n} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$. Dann gilt für $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$\begin{aligned} \langle \vec{x}, \vec{n} \rangle &= \langle \vec{p}, \vec{n} \rangle \\ \left\langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \end{pmatrix} \right\rangle \\ 5x + 1y &= 11 \end{aligned}$$

Dies ist die bekannte Normalform der Geradengleichung. Man beachte, dass die Vorfaktoren von x und y genau den Komponenten des Normalenvektors entsprechen.

Die Normalform ist nur bis auf einen Faktor genau bestimmt.

Definition 2.40: Eine spezielle Normalform ist die **Hessesche Normalform**, die man erhält, wenn man die Normalform durch $\|\vec{n}\|$ teilt:

$$\frac{\langle \vec{x}, \vec{n} \rangle}{\|\vec{n}\|} = \frac{\langle \vec{p}, \vec{n} \rangle}{\|\vec{n}\|}.$$

Die Besonderheit der Hesseschen Normalform ist, dass der Betrag der rechten Seite der kürzeste Abstand der Geraden zum Nullpunkt ist (Beweis später).

Beispiel 2.41: Die Gerade

$$5x + y = 11$$

hat den Normalenvektor $\vec{n} = (5, 1)$, was sich direkt aus den Vorfaktoren der linken Seite ablesen lässt. Damit ist

$$\|\vec{n}\| = \sqrt{25 + 1} = \sqrt{26}$$

und die Hessesche Normalform lautet:

$$\frac{1}{\sqrt{26}}(5x + y) = \frac{11}{\sqrt{26}}.$$

Der kürzeste Abstand der Geraden zum Nullpunkt ist $l = \frac{11}{\sqrt{26}}$.

2.2.2 Hyperebenen im \mathbb{R}^n

Eine Hyperebene im Raum \mathbb{R}^n ist definiert durch die Gleichung

$$\vec{x} = \vec{p} + \gamma_1 \vec{t}_1 + \gamma_2 \vec{t}_2 + \dots + \gamma_{n-1} \vec{t}_{n-1}$$

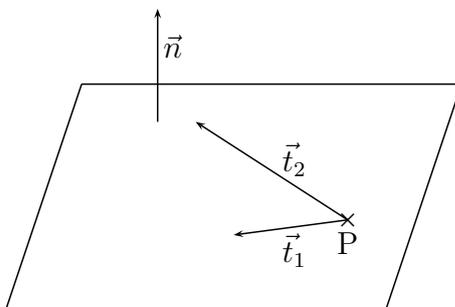
Hierbei müssen $\vec{t}_1, \dots, \vec{t}_{n-1}$ linear unabhängig sein (Erklärung folgt in Kapitel 3.2.1). Für den bisherigen Spezialfall $n = 2$ ergibt sich als Hyperebene

$$\vec{x} = \vec{p} + \gamma_1 \vec{t}_1$$

und damit eine Gerade. Eine Hyperebene des \mathbb{R}^3 ist gegeben durch

$$\vec{x} = \vec{p} + \gamma_1 \vec{t}_1 + \gamma_2 \vec{t}_2.$$

Diese Gleichung ist im \mathbb{R}^3 eine Ebenengleichung. Alle Punkte, die durch Variation von γ_1 und γ_2 erreicht werden können, liegen in einer Ebene im \mathbb{R}^3 (siehe Zeichnung). Voraussetzung ist, dass \vec{t}_1 und \vec{t}_2 linear unabhängig (d.h. in diesem Spezialfall: nicht parallel) sind.



Eine Besonderheit der Hyperebenen ist, dass der Normalenvektor bis auf einen Faktor eindeutig bestimmt ist und eine Normalform

$$\langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle$$

existiert.

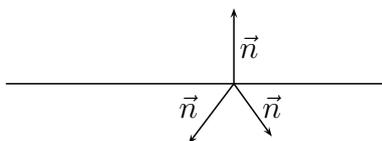
Beispiel 2.42: Die Gleichung

$$ax + by + cz + d = 0$$

beschreibt im \mathbb{R}^3 eine Ebene. Der Normalenvektor der Ebene ist $\vec{n} = (a, b, c)$.

2.2.3 Geraden im \mathbb{R}^3

Bemerkung 2.43: Geraden im \mathbb{R}^3 sind keine Hyperebenen. Es gibt daher für Geraden im \mathbb{R}^3 **keine** parameterlose Form, nur Parameterformen. Der Normalenvektor \vec{n} ist nicht eindeutig.



Die Punkt-Richtungsgleichung einer Geraden im \mathbb{R}^3 lautet

$$\vec{x} = \vec{p} + \gamma \vec{t}$$

und gleicht der entsprechenden Gleichung im \mathbb{R}^2 , nur das \vec{x} , \vec{p} und \vec{t} jetzt Vektoren des \mathbb{R}^3 sind.

2.2.4 Zusammenfassung

Raum	Objekt	Dim.	Punkt-Richtungs-Gl.	Normalgleichung
\mathbb{R}^2	Gerade	1	$\vec{x} = \vec{p} + \gamma \vec{t}$	$\langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle$ $ax + by + c = 0$
\mathbb{R}^3	Gerade	1	$\vec{x} = \vec{p} + \gamma \vec{t}$	Gibt es nicht!
\mathbb{R}^3	Ebene	2	$\vec{x} = \vec{p} + \gamma \vec{t} + \delta \vec{u}$	$\langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle$ $ax + by + cz + d = 0$
\mathbb{R}^n	Hyperebene	n-1	$\vec{x} = \vec{p} + \gamma_1 \vec{t}_1 + \dots + \gamma_{n-1} \vec{t}_{n-1}$	$\langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle$ $a_1 x_1 + \dots + a_n x_n + b = 0$

2.2.5 Umrechnen zwischen Parameterform und parameterloser Form

2.2.5.1 Bilden eines senkrechten Vektors

Bei der Umrechnung einer parameterlosen Form in eine Parameterform stellt sich das Problem, zu einem Vektor mehrere linear unabhängige senkrechte Vektoren zu bilden. Dazu gibt das im Folgenden beschriebene Rezept.

1. Der Ausgangsvektor sei $\vec{n} = (n_1, n_2, \dots, n_k)$. Wir brauchen dazu senkrechte linear unabhängige Vektoren.
2. Man nehme eine Komponente $n_i \neq 0$.
3. Für jeden senkrechten Vektor mache man folgende Schritte:
 - (a) Man vertausche die Komponente mit einer anderen Komponente n_j ; $j \neq i$.
 - (b) Man multipliziere eine der vertauschten Komponenten mit -1.
 - (c) Man setze die anderen Komponenten außer den vertauschten auf 0.

Beispiel 2.44: Im \mathbb{R}^3 sei ein Normalvektor (n_1, n_2, n_3) . n_1 sei ungleich 0. Vertauschen von n_1 mit n_3 führt zu (n_3, n_2, n_1) . Invertieren von n_1 führt zu $(n_3, n_2, -n_1)$. Anschließend wird n_2 auf 0 gesetzt. Das Ergebnis ist $(n_3, 0, -n_1)$. Die Multiplikation

$$\left\langle \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \begin{pmatrix} n_3 \\ 0 \\ -n_1 \end{pmatrix} \right\rangle = n_1 n_3 - n_3 n_1 = 0$$

zeigt, dass die Vektoren wirklich senkrecht aufeinander stehen. Wie wir später sehen, sind mehrere mit diesem Rezept erstellte senkrechte Vektoren auch automatisch linear unabhängig.

2.2.5.2 Geraden im \mathbb{R}^2

Parameterform \rightarrow Normalform

Die Parameterform einer Geraden im \mathbb{R}^2 ist: $\vec{x} = \vec{p} + \gamma \vec{t} = \begin{pmatrix} p_x \\ p_y \end{pmatrix} + \gamma \begin{pmatrix} t_x \\ t_y \end{pmatrix}$.

Der Normalenvektor auf der Gerade steht senkrecht auf \vec{t} und ist somit

$$\vec{n} = \begin{pmatrix} t_y \\ -t_x \end{pmatrix}.$$

Damit können die Werte \vec{p} und \vec{n} in $\langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle$ eingesetzt werden.

Beispiel 2.45: Die Parameterform lautet

$$\vec{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \gamma \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Damit erhält man $\vec{p} = (1, 2)$ und $\vec{n} = (3, 1)$ und die Normalform lautet:

$$\begin{aligned} \langle \vec{x}, \vec{n} \rangle &= \langle \vec{p}, \vec{n} \rangle \\ 3x + y &= 5 \end{aligned}$$

Normalform \rightarrow Parameterform

Hat man die parameterlose Form $ax + by = c$, so braucht man einen Richtungsvektor, der senkrecht zu $\vec{n} = (a, b)$ steht. Dieser errechnet sich zu $\vec{t} = (b, -a)$. Den Punkt \vec{p} erhält man, in dem man x oder y gleich 0 wählt und aus der parameterlosen Form die andere Komponente errechnet.

Beispiel 2.46: Die Normalform einer Geraden im \mathbb{R}^2 ist $3x + y = 5$. Damit ist $\vec{n} = (3, 1)$ und $\vec{t} = (1, -3)$. Setzt man $x = 0$ ergibt sich aus der Normalform $y = 5$ und damit $\vec{p} = (0, 5)$. Zusammengefasst erhält man:

$$\vec{x} = \begin{pmatrix} 0 \\ 5 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

2.2.5.3 Ebene im \mathbb{R}^3

Parameterform \rightarrow Normalform

Zur Umrechnung der Parameter- in die Normalform benötigt man einen Vektor, der senkrecht auf beiden Richtungsvektoren steht. Im \mathbb{R}^3 kann man dazu das Kreuzprodukt verwenden.

Beispiel 2.47: Eine Ebene im \mathbb{R}^3 ist definiert durch

$$\vec{x} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix}$$

Es gilt:

$$\vec{n} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$$

und schließlich

$$-x + y + 4z = 1$$

Normalform \rightarrow Parameterform

Die Umrechnung der Normal- in die Parameterform verläuft ähnlich wie bei Geraden im \mathbb{R}^2 . Aus der Normalgleichung

$$-x + y + 4z = 1$$

ergibt sich der Normalenvektor $\vec{n} = (-1, 1, 4)$ und damit als Richtungsvektoren $\vec{t}_1 = (1, 1, 0)$ und $\vec{t}_2 = (4, 0, 1)$. Setzt man x und z gleich 0, erhält man $y = 1$. Das führt zu $\vec{p} = (0, 1, 0)$ und

$$\vec{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$$

2.2.5.4 Hyperebenen im \mathbb{R}^n

Da das Kreuzprodukt nur im \mathbb{R}^3 definiert ist, benötigt man für die Umrechnung von Parameter- in Normalform in höheren Dimensionen nochmals einen längeren Weg. Man bildet aus der Parameterform ein Gleichungssystem und eliminiert darin die Parameter.

Beispiel 2.48: Da der Rechenweg in beliebigen \mathbb{R}^n funktioniert, hier ein Beispiel im \mathbb{R}^3 . Die Parameterform

$$\vec{x} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix}$$

ergibt das Gleichungssystem

$$\begin{array}{r} x = 3 + \gamma - \delta \\ y = \gamma + 3\delta \\ z = 1 - \delta \\ \hline x - y = 3 - 4\delta \\ z = 1 - \delta \\ \hline x - y - 4z = -1 \end{array}$$

Demgegenüber bleibt der Weg zur Umrechnung von Normal- in Parameterform immer gleich.

2.2.6 Parallelität und Orthogonalität von Geraden und Hyperebenen

- Zwei Geraden heißen parallel, wenn ihre Richtungsvektoren parallel sind. Sie heißen orthogonal, wenn ihre Richtungsvektoren orthogonal sind.

- Zwei Hyperebenen heißen parallel, wenn ihre Normalenvektoren parallel sind. Sie heißen orthogonal, wenn ihre Normalenvektoren orthogonal sind. Der Winkel zwischen zwei Hyperebenen ist der Winkel zwischen ihren Normalenvektoren.

Beispiel 2.49: Ermitteln Sie den Cosinus des Winkels zwischen den Ebenen

$$2x - y + z = 0 \quad \text{und} \quad x + 2y - z = 1.$$

Dieser Cosinus ist der Cosinus des Winkels zwischen den Vektoren

$$\vec{a} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

Dadurch ergibt sich

$$\cos \theta = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|} = -\frac{1}{6} \text{ d.h. } \theta \approx 100^\circ.$$

2.2.7 Schnittmengen zwischen Geraden und Ebenen im \mathbb{R}^2 und \mathbb{R}^3

In diesem Unterabschnitt wird ein zweckmäßiger Weg gezeigt, die Schnittmenge zwischen Geraden und Ebenen zu bestimmen. Es gibt darüber hinaus auch andere, gleichwertige Lösungswege.

2.2.7.1 Schnittmenge einer Gerade oder Ebene mit einer Hyperebene

Wir starten mit zwei Gleichungen, die eine Gerade oder Ebene beschreiben und parameterbehaftet oder parameterlos sein können.

- Der erste Schritt ist, eine Gleichung in eine Parametergleichung und die andere Gleichung in eine parameterlose Gleichung umzuformen. Das geht immer, wenn eine Hyperebene im Spiel ist.² Nur bei zwei Geraden im \mathbb{R}^3 muss man einen anderen Lösungsweg wählen.
- Im zweiten Schritt setzt man die Parametergleichung in die parameterlose ein und erhält den Wert eines Parameters.
- Diesen Wert setzt man in die Parametergleichung ein und erhält das Ergebnis.

²also eine Gerade im \mathbb{R}^2 oder eine Ebene im \mathbb{R}^3

Beispiel 2.50: Es sei

$$Q = (1, 1, 1) \quad \text{und} \quad P = (1, -1, 2).$$

Weiterhin sei

$$\vec{n} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Ermitteln Sie den Schnittpunkt der Geraden durch P in Richtung von \vec{n} mit der Ebene durch Q senkrecht zu \vec{n} .

Naheliegenderweise verwendet man bei der Gleichsetzung für die Ebene die parameterlose und für die Gerade die Parameterform. Die parametrische Gleichung der Geraden durch P in Richtung von \vec{n} ist

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Also:

$$x = 1 + \gamma; \quad y = -1 + 2\gamma; \quad z = 2 + 3\gamma$$

Die Normalengleichung der Ebene ist:

$$\begin{aligned} \langle \vec{x}, \vec{n} \rangle &= \langle \vec{q}, \vec{n} \rangle \\ x + 2y + 3z &= 6. \end{aligned}$$

Einsetzen der parameterlosen Gleichung:

$$\begin{aligned} 1 + \gamma + 2(-1 + 2\gamma) + 3(2 + 3\gamma) &= 6 \\ 5 + 14\gamma &= 6 \\ \gamma &= \frac{1}{14} \end{aligned}$$

Das Ergebnis erhält man durch Einsetzen von γ in die Parametergleichung:

$$\vec{p} + \gamma \vec{n} = (1, -1, 2) + \frac{1}{14} (1, 2, 3) = \left(\frac{15}{14}, -\frac{12}{14}, \frac{31}{14} \right).$$

2.2.7.2 Bestimmung des Schnittpunkts zweier Geraden im \mathbb{R}^3

Dieser Sonderfall ist etwas schwieriger. Die beiden Geradengleichungen sind:

$$G1: \vec{x} = \vec{p}_1 + \lambda \cdot \vec{r}_1; \quad G2: \vec{y} = \vec{p}_2 + \mu \cdot \vec{r}_2$$

Der Weg besteht darin, die beiden Gleichung gleichzusetzen und aufzulösen:

$$\vec{p}_1 + \lambda \cdot \vec{r}_1 = \vec{p}_2 + \mu \cdot \vec{r}_2$$

Jetzt können generell drei Fälle auftreten:

1. Die Geraden haben einen Schnittpunkt. Das Gleichungssystem hat genau eine Lösung.
2. Die Geraden sind parallel. Man erhält entweder mehrere Lösungen (wenn die Geraden genau aufeinander liegen) oder keine Lösung.
3. Die Geraden haben weder einen Schnittpunkt, noch sind sie parallel. In diesem Fall spricht man von **windschiefen** Geraden. Das Gleichungssystem hat keine Lösung.

Beispiel 2.51: Finden Sie den Schnittpunkt der Geraden

$$G1: \vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad G2: \vec{y} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \beta \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix},$$

falls ein solcher existiert.

Ansatz:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \beta \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

$$\alpha \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + \beta \cdot \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -2 \end{pmatrix}$$

Lösung des Gleichungssystems:

$$\begin{array}{ccc|c} -1 & -2 & -1 & \\ 1 & 1 & 0 & +\textcircled{1} \\ 1 & -1 & -2 & +\textcircled{1} \\ \hline -1 & -2 & -1 & \cdot(-1) \\ 0 & -1 & -1 & \cdot(-1) \\ 0 & -3 & -3 & -3 \cdot \textcircled{2} \\ \hline 1 & 2 & 1 & -2 \cdot \textcircled{2} \\ 0 & 1 & 1 & \\ 0 & 0 & 0 & \\ \hline 1 & 0 & -1 & \\ 0 & 1 & 1 & \end{array}$$

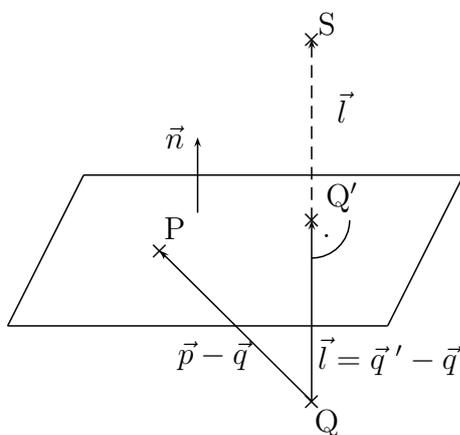
Es wird nur eine der beiden Variablen α oder β gebraucht. Mit $\beta = 1$ ergibt sich als Lösung

$$\vec{s} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$$

2.2.8 Abstandsbestimmung im \mathbb{R}^2 und \mathbb{R}^3

2.2.8.1 Einführung

Satz 2.52: Der kürzeste Verbindungsvektor zwischen einem Punkt, einer Geraden oder einer Ebene stößt immer senkrecht auf die Gerade oder die Ebene.



BEWEIS Der Satz wird für die Verbindung zwischen Punkt und Ebene gezeigt. Für andere Kombinationen gilt er analog. Ein Verbindungsvektor \vec{v} , der nicht senkrecht auf die Ebene stößt, kann zusammengesetzt werden aus einer Komponente parallel zur Ebene \vec{v}_p und einer Komponente \vec{v}_s senkrecht zur Ebene. Da nach Pythagoras

$$\|\vec{v}\|^2 = \|\vec{v}_p\|^2 + \|\vec{v}_s\|^2$$

gilt, kann \vec{v}_s nicht länger als \vec{v} sein. ■

Definition 2.53: Der Punkt Q' ist derjenige Punkt in der Ebene, der den kürzesten Abstand zu Punkt Q besitzt. Er wird **Lotfußpunkt** genannt. Der Punkt S wird **Spiegelpunkt** genannt, da man sich in Q stehend dort sehen würde, falls die Ebene ein Spiegel ist. Für Geraden anstatt Ebenen sind die Definitionen äquivalent.

Aus Satz 2.52 kann man ein Verfahren zur Abstandsbestimmung herleiten, das für fast alle Kombinationen aus Punkten, Geraden und Ebenen im \mathbb{R}^2 und \mathbb{R}^3 funktioniert. Die Ausnahmen sind:

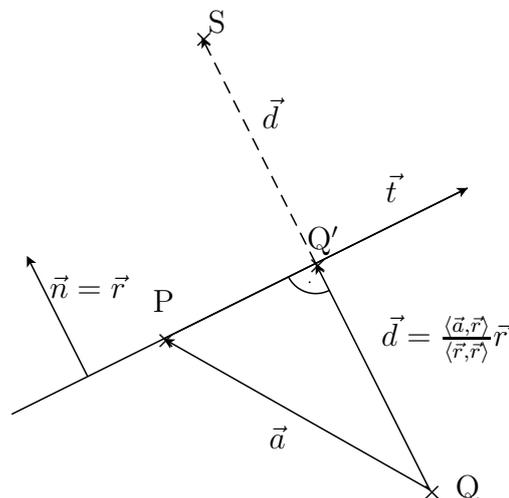
- Der Abstand zwischen zwei Punkten. Hier ergibt sich die Berechnung trivialerweise durch $\|\vec{a} - \vec{b}\|$.
- Der Abstand zwischen einer Geraden und einem Punkt im \mathbb{R}^3 . Hierzu muss man einen zweiten Rechenweg kennen, der im Anschluss erläutert wird.
- Der Abstand zwischen zwei parallelen Geraden im \mathbb{R}^3 lässt sich auf den Abstand zwischen einem Punkt und einer Geraden zurückführen.

Der Rechenweg:

- Bestimme die Richtung \vec{r} des kürzesten Abstandsvektors.
- Bestimme jeweils einen Punkt auf den beiden Objekten und (durch Differenzbildung) den Abstandsvektor \vec{a} zwischen den beiden Punkten.
- Der kürzeste Abstandsvektor \vec{d} ist die Projektion von \vec{a} auf \vec{r} . Der kürzeste Abstand ist $\|\vec{d}\|$.

$$\vec{d} = \frac{\langle \vec{a}, \vec{r} \rangle}{\langle \vec{r}, \vec{r} \rangle} \vec{r}; \quad \|\vec{d}\| = \frac{|\langle \vec{a}, \vec{r} \rangle|}{\langle \vec{r}, \vec{r} \rangle} \|\vec{r}\| = \frac{|\langle \vec{a}, \vec{r} \rangle|}{\|\vec{r}\|}$$

Dieser Zusammenhang ist anhand des Abstandes zwischen einem Punkt und einer Geraden im \mathbb{R}^2 an der folgenden Grafik erläutert:



Die einzelnen Fälle unterscheiden sich nur, wie man \vec{r} erhält.

- Ist eine Hyperebene im Spiel (Gerade im \mathbb{R}^2 oder Ebene im \mathbb{R}^3), dann ist \vec{r} gleich dem Normalenvektor \vec{n} der Hyperebene.

- Bei zwei Geraden im \mathbb{R}^3 muss \vec{r} senkrecht auf beiden Geraden stehen. Daraus ergibt sich \vec{r} als Vektorprodukt der beiden Richtungsvektoren. Bei zwei parallelen Geraden funktioniert das nicht, weil das Vektorprodukt 0 ergibt. Hier muss man den zweiten Lösungsweg wählen.

Beispiel 2.54: Gegeben ist im \mathbb{R}^2 die Gerade g und der Punkt $P = (2, 0)$ mit

$$g : \vec{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Wie groß ist der minimale Abstand zwischen g und P ?

Aus dem Richtungsvektor der Geraden $\vec{t} = (2, 1)$ folgt der Normalenvektor $\vec{n} = (1, -2) = \vec{r}$. Ein Abstandsvektor ergibt sich aus

$$\vec{a} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Der minimale Abstand ist also:

$$d = \frac{|\langle \vec{a}, \vec{r} \rangle|}{\|\vec{r}\|} = \frac{1}{\sqrt{5}}$$

Beispiel 2.55: Gegeben ist im \mathbb{R}^3 der Punkt $Q = (1, 3, 5)$ und die Ebene

$$-x + y - z = -5.$$

Wie groß ist der minimale Abstand zwischen Punkt und Ebene?

Ein Punkt P auf der Ebene ist $P = (5, 0, 0)$. Damit ist $\vec{a} = Q - P = (-4, 3, 5)$. Die Richtung des kürzesten Abstands ist gleich dem Normalenvektor der Ebene $\vec{r} = \vec{n} = (-1, 1, -1)$. Damit ist

$$d = \frac{|\langle \vec{a}, \vec{r} \rangle|}{\|\vec{r}\|} = \frac{|4 + 3 - 5|}{\sqrt{3}} = \frac{2}{\sqrt{3}}$$

Wenn nur der Abstand einer Ebene vom Nullpunkt gesucht ist, dann erhält man die einfachere Formel

$$d = \frac{|\langle \vec{p}, \vec{n} \rangle|}{\|\vec{n}\|},$$

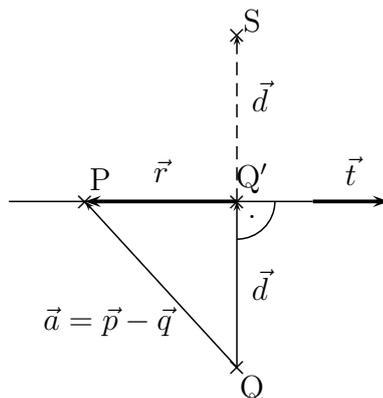
also die Länge der Projektion des Ortsvektors \vec{p} zum Punkt P auf der Ebene mit dem Normalenvektor \vec{n} . Der Vergleich mit der Hesseschen Normalform

$$\frac{\langle \vec{x}, \vec{n} \rangle}{\|\vec{n}\|} = \frac{\langle \vec{p}, \vec{n} \rangle}{\|\vec{n}\|}$$

zeigt, dass der Betrag der rechten Seite der Hesseschen Normalform gerade der Abstand der Ebene zum Nullpunkt ist.

2.2.8.2 Der Abstand eines Punktes und einer Geraden im \mathbb{R}^3

Die Methode aus dem vorigen Abschnitt funktioniert **nicht**, da die Gerade im \mathbb{R}^3 keine Hyperebene, bzw. \vec{n} nicht eindeutig ist. Folgender Ansatz ist recht einfach zu merken:



Man verfolgt nun folgenden Ansatz zur Bestimmung des Abstandes:

1. $\vec{a} = \vec{d} + \vec{r}$, also $\vec{d} = \vec{a} - \vec{r}$.
2. \vec{r} ist die Projektion von \vec{a} auf den Richtungsvektor der Geraden \vec{t} , also

$$\vec{r} = \frac{\langle \vec{a}, \vec{t} \rangle}{\langle \vec{t}, \vec{t} \rangle} \vec{t}.$$

3. Zusammen ergibt sich:

$$\vec{d} = \vec{a} - \frac{\langle \vec{a}, \vec{t} \rangle}{\langle \vec{t}, \vec{t} \rangle} \vec{t}.$$

Beispiel 2.56: Gegeben sind:

$$g : \vec{x} = \begin{pmatrix} 4 \\ 3 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

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

Damit ist $\vec{t} = (2, 2, 1)$ und $\vec{a} = (3, 5, 2)$. \vec{d} und $\|\vec{d}\|$ errechnen sich zu

$$\begin{aligned} \vec{d} &= \vec{a} - \frac{\langle \vec{a}, \vec{t} \rangle}{\langle \vec{t}, \vec{t} \rangle} \vec{t} \\ &= \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} - \frac{6 + 10 + 2}{4 + 4 + 1} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\ \|\vec{d}\| &= \sqrt{2} \end{aligned}$$

Damit ergibt sich für den Lotfußpunkt:

$$\vec{q}' = \vec{q} + \vec{d} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}.$$

2.3 Die Determinante im \mathbb{R}^2 und \mathbb{R}^3

2.3.1 Die Determinante von Matrizen bis 3×3

Determinanten werden allgemein erst in LA2 behandelt. In diesem Kapitel werden die Spezialfälle der Determinanten von (1×1) , (2×2) - und (3×3) -Matrizen vorgestellt. Sätze, die allgemein auch für $(n \times n)$ -Determinanten gelten, werden gekennzeichnet.

Eine Determinante kann nur von einer quadratischen $(n \times n)$ -Matrix gebildet werden, d.h. von einem aus $n \cdot n$ Werten bestehenden rechteckigen Schema mit n Zeilen und n Spalten. Die Matrix kann man dabei auch als Zusammenstellung von n n -dimensionalen Spaltenvektoren verstehen.

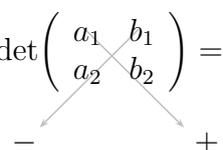
$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = (\vec{a}_1, \dots, \vec{a}_n)$$

Das Ergebnis ist eine Zahl. Allgemein werden Determinanten nach dem Laplaceschen Entwicklungssatz oder nach Umformung in eine Dreiecksmatrix berechnet. Bis $n = 3$ gibt es aber spezielle Gleichungen:

n=1:

$$\det(a_1) = a_1$$

n=2:

$$\det(\vec{a}, \vec{b}) = \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = a_1 b_2 - b_1 a_2$$


Entlang der Diagonalen wird das Produkt gebildet und entsprechend dem angezeigten Vorzeichen aufsummiert.

n=3:

$$\begin{aligned} \det(\vec{a}, \vec{b}, \vec{c}) &= \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \\ &= a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - a_1 c_2 b_3 - b_1 a_2 c_3 - c_1 b_2 a_3 \end{aligned}$$

Diese Formel kann durch die Berechnungsmethode nach **Sarrus** veranschaulicht werden:

$$\begin{array}{ccccccc}
 & a_1 & b_1 & c_1 & | & a_1 & b_1 \\
 & a_2 & b_2 & c_2 & | & a_2 & b_2 \\
 & a_3 & b_3 & c_3 & | & a_3 & b_3 \\
 & \swarrow & \swarrow & \swarrow & & \searrow & \searrow \\
 & - & - & - & & + & + & +
 \end{array}$$

Entlang der Diagonalen werden Produkte gebildet und diese mit den abgebildeten Vorzeichen versehen aufsummiert.

2.3.1.1 Geometrische Bedeutung der Determinante im \mathbb{R}^3

Definition 2.57: Für drei Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ nennt man

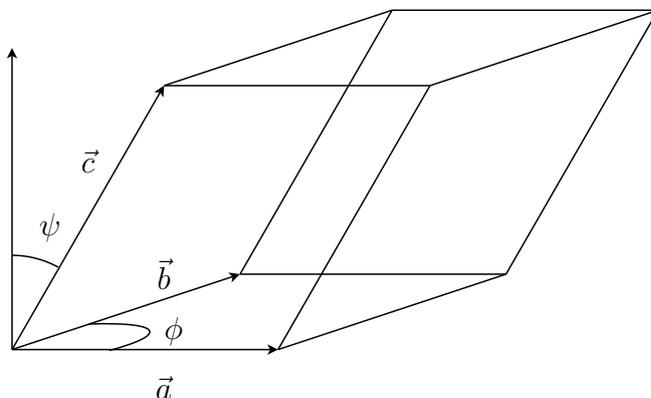
$$\langle \vec{a}, \vec{b} \times \vec{c} \rangle$$

das **Spatprodukt** der drei Vektoren $\vec{a}, \vec{b}, \vec{c}$.

Bemerkung 2.58: Es gilt im \mathbb{R}^3 (Beweis durch Ausmultiplizieren):

$$\det(\vec{a}, \vec{b}, \vec{c}) = \langle \vec{a} \times \vec{b}, \vec{c} \rangle = \|\vec{a} \times \vec{b}\| \cdot \|\vec{c}\| \cos \psi = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \|\vec{c}\| \sin \phi \cos \psi$$

Die Vektoren $\vec{a}, \vec{b}, \vec{c}$ bilden die Kanten eines Körpers im dreidimensionalen Raum, eines **Parallelepiped**s oder **Spats** (s. Abb.).



2.3.1.2 Rechenregeln

Die nachfolgenden Rechenregeln für Determinanten gelten übertragen auch für den \mathbb{R}^n . Im \mathbb{R}^3 können sie durch Beweis der Regeln für das Spatprodukt gezeigt werden.

Satz 2.59: Die Determinante hat folgende Eigenschaften.

$$\mathbf{D1} : \det(\vec{a}, \vec{b}, \vec{c}) = \det(\vec{c}, \vec{a}, \vec{b}) = \det(\vec{b}, \vec{c}, \vec{a}).$$

$$\mathbf{D2} : \det(\vec{a}, \vec{b}, \vec{c}) = -\det(\vec{b}, \vec{a}, \vec{c}).$$

$$\mathbf{D3} : \det(\vec{a}, \vec{a}, \vec{c}) = 0.$$

$$\mathbf{D4} : \text{Für } \alpha \in \mathbb{R} \text{ gilt } \det(\alpha \cdot \vec{a}, \vec{b}, \vec{c}) = \alpha \cdot \det(\vec{a}, \vec{b}, \vec{c}).$$

$$\mathbf{D5} : \det(\vec{a}, \vec{b}, \vec{c} + \vec{d}) = \det(\vec{a}, \vec{b}, \vec{c}) + \det(\vec{a}, \vec{b}, \vec{d}).$$

$$\mathbf{D6} : \det(A) = \det(A^T)$$

Erläuterungen, Beispiele und Folgerungen:

- Beispiel zu [D4](#):

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

- Folgerung aus [D4](#):

$$\det(0, \vec{b}, \vec{c}) = 0 \cdot \det(0, \vec{b}, \vec{c}) = 0$$

Ist in einer Matrix eine Spalte (oder eine Zeile, siehe [D6](#)) gleich 0, dann ist auch deren Determinante gleich 0.

- Beispiel zu [D5](#):

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

- Beispiel zu der geometrisch verblüffenden Regel [D6](#) für transponierte Matrizen, die wir in [Definition 1.5](#) beschrieben hatten:

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

Bemerkung 2.60: Es gilt:

$$\begin{aligned} \det(\vec{a}, \vec{b}, \vec{c} + \alpha\vec{a}) &\stackrel{D5}{=} \det(\vec{a}, \vec{b}, \vec{c}) + \det(\vec{a}, \vec{b}, \alpha\vec{a}) \\ &\stackrel{D4}{=} \det(\vec{a}, \vec{b}, \vec{c}) + \alpha \cdot \det(\vec{a}, \vec{b}, \vec{a}) \\ &\stackrel{D1}{=} \det(\vec{a}, \vec{b}, \vec{c}) + \alpha \cdot \det(\vec{a}, \vec{a}, \vec{b}) \\ &\stackrel{D3}{=} \det(\vec{a}, \vec{b}, \vec{c}) \end{aligned}$$

2.3.2 Lineare 3×3 -Gleichungssysteme

Eine Anwendung von Determinanten ist die Untersuchung von linearen Gleichungssystemen auf eindeutige Lösbarkeit. Durch

$$(LG) \begin{cases} (1) a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 = d_1 \\ (2) b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 = d_2 \\ (3) c_1 \cdot x_1 + c_2 \cdot x_2 + c_3 \cdot x_3 = d_3 \end{cases}$$

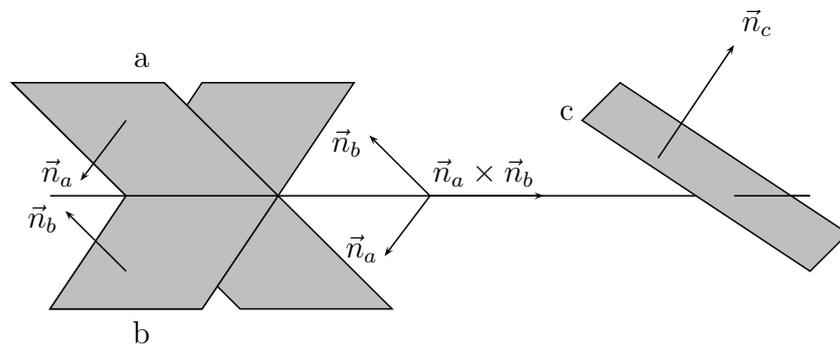
ist ein **lineares Gleichungssystem** mit den Unbekannten (x_1, x_2, x_3) gegeben. Jede Gleichung ist bekanntlich eine Ebene im dreidimensionalen Raum. Eindeutige Lösbarkeit bedeutet also die Eindeutigkeit des Schnittpunkts der drei Ebenen. Der Vektor

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

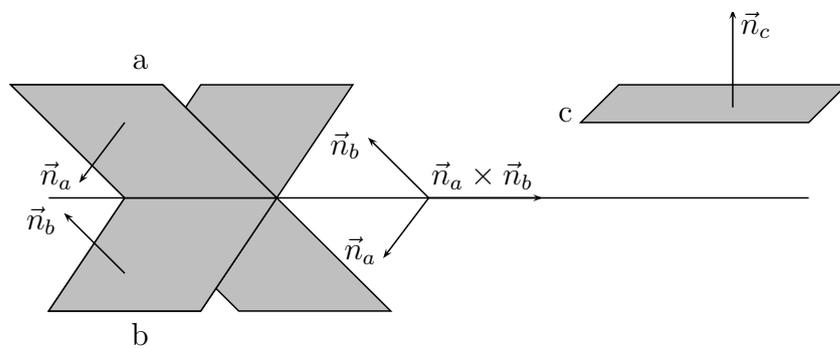
ist der Normalenvektor der Ebene (1). Entsprechend hat man die Normalenvektoren \vec{b} und \vec{c} der Ebenen (2) und (3). Bei einem eindeutigen Schnittpunkt dürfen die Ebenen (1) und (2) nicht parallel sein, d.h. es muss gelten $\vec{a} \times \vec{b} \neq \vec{0}$. Der Vektor $\vec{a} \times \vec{b}$ zeigt in Richtung der Schnittgeraden der beiden Ebenen (1) und (2). Die Schnittgerade darf nicht parallel zur dritten Ebene, d.h. nicht senkrecht zum Normalenvektor \vec{c} , sein. Es muss also gelten

$$\langle (\vec{a} \times \vec{b}), \vec{c} \rangle = \det(\vec{a}, \vec{b}, \vec{c}) \neq 0.$$

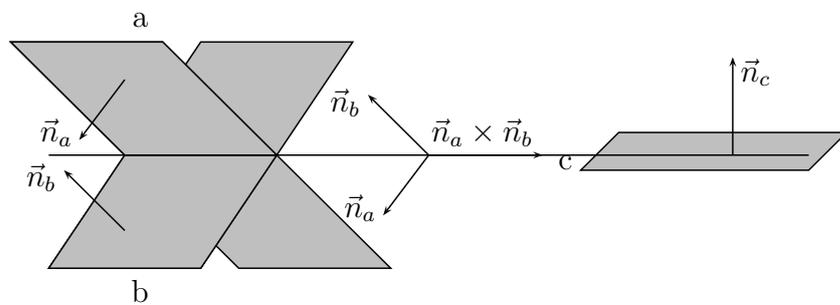
Damit ist die Bedingung für die eindeutige Lösbarkeit des obigen Gleichungssystems gegeben.



Eindeutiger Schnittpunkt: $\langle \vec{n}_a \times \vec{n}_b, \vec{n}_c \rangle \neq 0$



Kein Schnittpunkt: $\langle \vec{n}_a \times \vec{n}_b, \vec{n}_c \rangle = 0$



Schnittgerade, kein eindeutiger Schnittpunkt: $\langle \vec{n}_a \times \vec{n}_b, \vec{n}_c \rangle = 0$

Satz 2.61: Notwendig und hinreichend dafür, dass das lineare Gleichungssystem LG eine eindeutige Lösung besitzt, ist die Bedingung

$$\det(\vec{a}, \vec{b}, \vec{c}) \neq 0.$$

2.4 Kegelschnitte

2.4.1 Kurven und Gleichungen

Es sei $F(x, y)$ eine Funktion, die von $(x, y) \in \mathbb{R}^2$ abhängt. Für eine Zahl $c \in \mathbb{R}$ betrachten wir die Gleichung

$$F(x, y) = c.$$

Zum Beispiel erhält man für $F(x, y) = x + y$ und $c = 2$ die Gleichung

$$x + y = 2.$$

Die Menge der Punkte, die diese Gleichung erfüllt, heißt **Graph der Gleichung**. Im Beispiel ist der Graph eine Gerade.

Auch Funktionen haben Graphen. Zum Beispiel hat die Funktion

$$f(x) = 5 - x$$

ebenfalls eine Gerade als Graph. Es gilt: Wenn $f(x)$ eine reelle Funktion ist, dann entspricht der Graph der Gleichung

$$F(x, y) = y - f(x) = 0$$

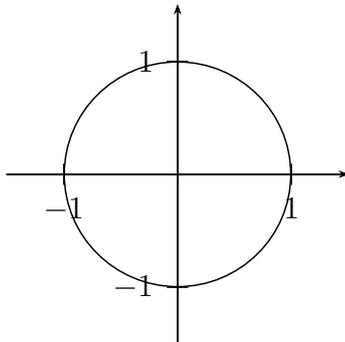
dem Funktionsgraphen von $f(x)$. Das wäre in diesem Beispiel

$$y - 5 + x = 0.$$

Somit lässt sich jeder Funktionsgraph als Graph einer Gleichung darstellen. Umgedreht ist aber nicht jeder Graph einer Gleichung auch ein Funktionsgraph. Zum Beispiel ist

$$x^2 + y^2 = 1$$

kein Funktionsgraph (siehe Abbildung).

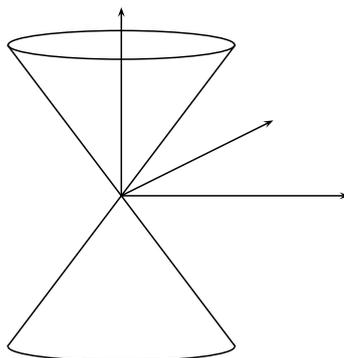


2.4.2 Kegelschnitte

Wichtige Beispiele von Kurven sind Kegelschnitte. Kegelschnitte entstehen, wenn ein Kegel mit einer Ebene im \mathbb{R}^3 geschnitten wird.

Die einfachste Gleichung eines Kegels ist die Menge aller Punkte $(x, y, z) \in \mathbb{R}^3$, für die gilt

$$z^2 = x^2 + y^2.$$



Eine Ebene im \mathbb{R}^3 ist bekanntlich durch die Gleichung

$$ax + by + cz = d$$

gegeben. Löst man diese Gleichung nach einer der Variablen x , y oder z auf - z.B. nach z , wenn $c \neq 0$ ist - dann erhält man die Gleichung

$$z = a'x + b'y + d'.$$

Diese Gleichung kann man nun in die Gleichung des Kegels einsetzen und erhält

$$(a'x + b'y + d')^2 = x^2 + y^2.$$

Durch Ausmultiplizieren erhält man die Gleichung

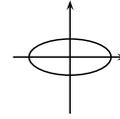
$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0.$$

Dies ist die Gleichung eines Kegelschnitts in allgemeiner Form, wobei mindestens eine der drei Konstanten $\{A, B, C\}$ ungleich 0 sein muss.

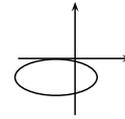
2.4.3 Lage von Kegelschnitten

Bevor wir die verschiedenen Arten von Kegelschnitten untersuchen, werfen wir einen Blick auf die Lage eines Kegelschnitts im Koordinatensystem. Wir nehmen dazu als Beispiel eine Ellipse.

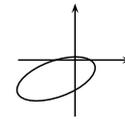
- Der Kegelschnitt kann sich in seiner **Standardlage** befinden. Das heißt, dass der Mittelpunkt des Kegelschnitts der Koordinatenursprung ist und dass sich die Symmetrieachsen des Kegelschnitts mit den Koordinatenachsen decken. Diese Form des Kegelschnitts heißt **Mittelpunktsform**.



- Der Kegelschnitt kann **verschoben** sein, d.h. der Mittelpunkt ist nicht mehr der Koordinatenursprung, aber die Symmetrieachsen laufen noch parallel zu den Koordinatenachsen. Diese Form des Kegelschnitts heißt **Verschiebungsform**. Die Verschiebungsform lässt sich aus der Standardform ableiten, indem man x durch $x - x_0$ und y durch $y - y_0$ ersetzt, wobei (x_0, y_0) der Mittelpunkt des Kegelschnitts ist.



- Der Kegelschnitt kann gedreht bzw. gedreht und verschoben sein. Diese Kegelschnittformen behandeln wir allgemein erst in LA 2.

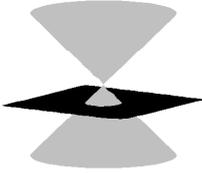
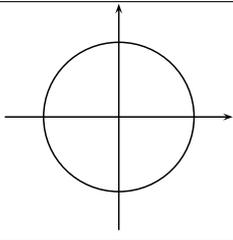
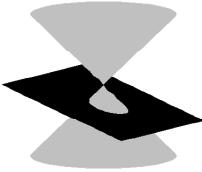
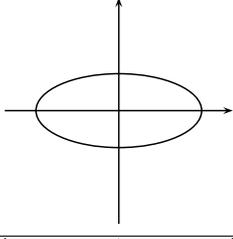
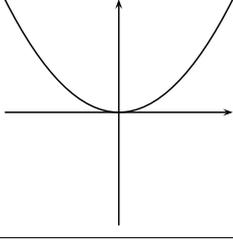
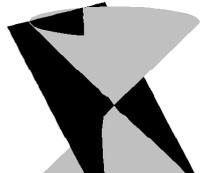
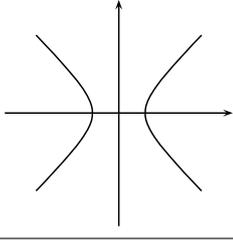


2.4.4 Übersicht über die Kegelschnitte

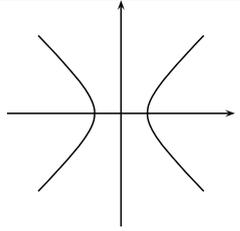
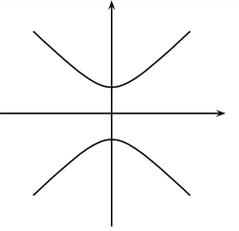
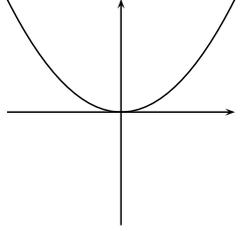
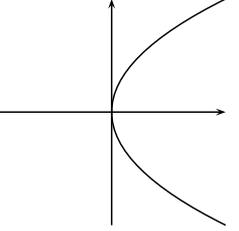
Das Aussehen des Schnitts wird wesentlich von der Neigung n der Ebene bestimmt. Diese kann zwischen 0 (horizontal) und ∞ (vertikal) liegen. Die Neigung 1 entspricht gerade der Neigung der Kegeloberfläche und ist darum ausgezeichnet.

2.4.4.1 Übersicht über die Standard-Kegelschnitte

Folgende Fälle werden unterschieden:

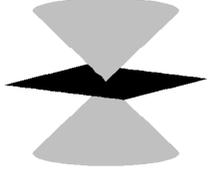
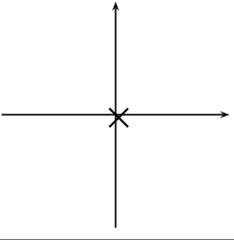
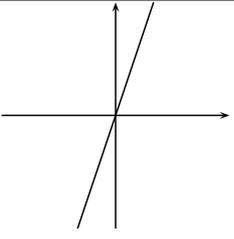
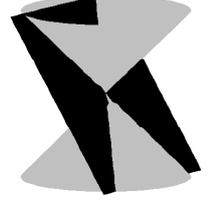
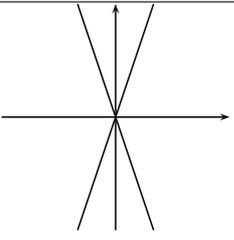
n	Kegelschnitt		
0	Kreis		
(0,1)	Ellipse		
1	Parabel		
> 1	Hyperbel		

Parabeln und Hyperbeln können sowohl stehend als auch liegend sein:

	stehend	liegend
Hyperbel		
Parabel		

2.4.4.2 Übersicht über die entarteten Kegelschnitte

Bei den entarteten Kegelschnitten gibt es zum einen die, die durch den Koordinatenursprung gehen:

n	Kegelschnitt		
$[0, 1)$	Punkt		
1	Gerade		
> 1	Zwei sich kreuzende Geraden		

Andere entartete Kegelschnitte lassen sich nicht durch Schnitt eines Kegels mit einer Ebene darstellen, wohl aber durch die allgemeine Kegelschnittgleichung

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$$

Dies sind zwei parallele Geraden sowie mehrere Sorten von imaginären Kegelschnitten, also Gleichungen, die nur im komplexen Raum Lösungen haben.

2.4.5 Kegelschnittgleichungen

Wichtig sind die nicht entarteten Kegelschnitte Kreis, Ellipse, Parabel und Hyperbel. Die anderen Gleichungen sind weniger relevant.

Kegelschnitt	Mittelpunktsform	Besonderheiten
Ellipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	a, b : Halbachsen
Kreis	$\frac{x^2}{r^2} + \frac{y^2}{r^2} = 1$ oder $x^2 + y^2 = r^2$	r : Radius
Punkt	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$	
Imaginärer Kegelschnitt	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$	
Hyperbel (stehend)	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	$\frac{b}{a}x, -\frac{b}{a}x$: Asymptoten
Hyperbel (liegend)	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$	$\frac{b}{a}x, -\frac{b}{a}x$: Asymptoten
Zwei kreuzende Geraden	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$	
Parabel (stehend)	$ax^2 - y = 0$	(x_0, y_0) : Scheitelpunkt
2 par. Geraden (senkrecht)	$a^2 \cdot x^2 = 1$	
Gerade (senkrecht)	$a^2 \cdot x^2 = 0$	
Imaginärer Kegelschnitt	$a^2 \cdot x^2 = -1$	
Parabel (liegend)	$ay^2 - x = 0$	(x_0, y_0) : Scheitelpunkt
2 par. Geraden (waager.)	$a^2 \cdot y^2 = 1$	
Gerade (waagrecht)	$a^2 \cdot y^2 = 0$	
Imaginärer Kegelschnitt	$a^2 \cdot y^2 = -1$	

Die Verschiebungsformen entstehen, indem x durch $x - x_0$ und y durch $y - y_0$ ersetzt wird.

Kegelschnitt	Mittelpunktsform	Verschiebungsform
Ellipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1$
Kreis	$\frac{x^2}{r^2} + \frac{y^2}{r^2} = 1$	$\frac{(x - x_0)^2}{r^2} + \frac{(y - y_0)^2}{r^2} = 1$
Hyperbel(stehend)	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	$\frac{(x - x_0)^2}{a^2} - \frac{(y - y_0)^2}{b^2} = 1$
Hyperbel(liegend)	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$	$\frac{(x - x_0)^2}{a^2} - \frac{(y - y_0)^2}{b^2} = -1$
Parabel(stehend)	$ax^2 - y = 0$	$a(x - x_0)^2 - (y - y_0) = 0$
Parabel(liegend)	$ay^2 - x = 0$	$a(y - y_0)^2 - (x - x_0) = 0$

2.4.6 Bestimmung der Kegelschnitte aus der Kegelschnittgleichung

Die allgemeine Kegelschnittgleichung lautet:

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$$

wobei A , B und C nicht gleichzeitig 0 werden dürfen.

- Ist ein xy -Term vorhanden ($C \neq 0$), dann ist der Kegelschnitt gedreht. Da wir gedrehte Kegelschnitte erst im nächsten Semester behandeln, nehmen wir hier $C = 0$ an und gehen von der Gleichung

$$Ax^2 + By^2 + Dx + Ey + F = 0$$

aus.

- Nun bestimmen wir die Art des Kegelschnitts anhand folgender Tabelle:

	Kegelschnitt
$A = B$	Kreis
$A \cdot B > 0$	Ellipse
$A \cdot B < 0$	Hyperbel
$A = 0, B \neq 0$	Parabel (liegend)
$B = 0, A \neq 0$	Parabel (stehend)

Dabei können auch die entsprechenden entarteten Varianten auftreten.

- Falls die Gleichung nicht bereits die entsprechende Mittelpunktsform ist (je nach Gleichung D , E oder $F \neq 0$), so ist die Gleichung verschoben. Man erhält die Verschiebungsform, indem man D bzw. E durch quadratische Ergänzung beseitigt.

Beispiel 2.62: Eine typische Aufgabe heißt:

Man bestimme die Schnittkurve des Kegels $z^2 = x^2 + y^2$ mit den jeweiligen Ebenen:

$$(a) \quad z = \frac{2}{3}x + 5$$

$$(b) \quad z = y + 2$$

Es ist nur die Projektion der Schnittkurve in die (x,y) -Ebene zu bestimmen.

Der letzte Satz ist eine Vereinfachung. Eigentlich liegt der Kegelschnitt in der Schnittebene, also einer Ebene im \mathbb{R}^3 . Wenn wir die Projektion des Kegelschnitts auf die (x,y) -Ebene betrachten, können wir die z -Komponente des Kegelschnitts gleich 0 setzen.

(a)

Die Gleichung der Projektion erhält man durch Gleichsetzen der Ebenengleichung $z = \frac{2}{3}x + 5$ und der Kegelgleichung $z^2 = x^2 + y^2$:

$$\begin{aligned} \left(\frac{2}{3}x + 5\right)^2 &= x^2 + y^2 \\ \frac{4}{9}x^2 + \frac{20}{3}x + 25 &= x^2 + y^2 \\ \frac{5}{9}x^2 + y^2 - \frac{20}{3}x - 25 &= 0 \end{aligned}$$

Da kein xy -Term vorhanden ist, ist der Kegelschnitt nicht gedreht. Es gilt: $A \cdot B = \frac{5}{9} \cdot 1 = \frac{5}{9} > 0$. Damit handelt es sich um eine Ellipse. Da $A \neq B$, ist es kein Kreis. Da $D \neq 0$, ist die Ellipse verschoben und die Gleichung muss mit quadratischer Ergänzung in die Verschiebungsform überführt werden. Dazu betrachten wir einzeln die Terme mit x und die mit y .

$$\begin{aligned} \frac{5}{9}x^2 - \frac{20}{3}x &= \frac{5}{9}(x^2 - 12x) \\ &= \frac{5}{9}(x^2 - 12x + 36) - 20 \\ &= \frac{5}{9}(x - 6)^2 - 20 \end{aligned}$$

Es ergibt sich:

$$\begin{aligned}\frac{5}{9}(x-6)^2 + y^2 &= 45 \\ \frac{(x-6)^2}{81} + \frac{y^2}{45} &= 1\end{aligned}$$

Es handelt sich also tatsächlich um eine Ellipse. Sie hat den Mittelpunkt $(6,0)$ und Halbachsen der Länge 9 und $\sqrt{45}$.

(b)

Die Ebenengleichung $z = y + 2$ führt zur Schnittgleichung

$$x^2 - 4y - 4 = 0$$

Da $C = 0$, ist der Kegelschnitt nicht gedreht. Da auch $B = 0$, handelt es sich um eine stehende Parabel. Da $F \neq 0$, ist die Parabel verschoben. Es ergibt sich:

$$\frac{1}{4}x^2 - (y+1) = 0$$

und damit eine Parabel mit dem Scheitelpunkt $(0, -1)$.

Kapitel 3

Vektorräume

3.1 Gruppen, Körper, Vektorräume

3.1.1 Äquivalenzrelationen

Folgende Definition kennen wir schon aus dem Mathematik-Blockkurs:

Definition 3.1: Es sei M eine Menge. Unter einer **Äquivalenzrelation** auf M versteht man eine Relation $x \sim y$, welche die folgenden drei Axiome erfüllt:

1. Reflexivität: $x \sim x \forall x \in M$
2. Symmetrie: $x \sim y \Leftrightarrow y \sim x \forall x, y \in M$
3. Transitivität: $\forall x, y, z \in M$ gilt: $x \sim y \wedge y \sim z \Rightarrow x \sim z$.

Beispiel 3.2: 1. Bei Vektoren \vec{AB} und \vec{CD} im \mathbb{R}^n ist die Äquivalenzrelation $\vec{AB} \sim \vec{CD}$ definiert durch $B - A = D - C$. Dass hier die Axiome (1) - (3) erfüllt sind, ist leicht nachzuweisen.

2. Auf \mathbb{R} ist die Gleichheit $=$ eine Äquivalenzrelation und ebenso auf \mathbb{C} , der Menge der komplexen Zahlen.
3. Die "Kleiner-Relation" $<$ ist auf \mathbb{R} keine Äquivalenzrelation, da Axiom (2) ($a < b \Leftrightarrow b < a$) nicht erfüllt ist.

Beispiel 3.3: Wir betrachten die erste Äquivalenzrelation aus dem vorigen Beispiel. Es sei $P = (1, -1, 3)$, $Q = (2, 4, 1)$ und $C = (1, 5, -2)$. Dann ist \vec{PQ} äquivalent zu $\vec{0C}$, denn $C = Q - P$.

Wenn $A = (4, -2, 5)$ und $B = (5, 3, 3)$, dann ist \vec{PQ} äquivalent zu \vec{AB} , da $Q - P = B - A = (1, 5, -2)$.

Definition 3.4: Es sei \sim eine Äquivalenzrelation auf einer Menge M . Für ein festes $x \in M$ heißt die Teilmenge $[x] = \{y; y \sim x\} \subset M$ die **Äquivalenzklasse** von x bezüglich \sim .

Erläuterung: Die Menge aller zu einem festen Element x äquivalenten Elemente nennt man eine **Äquivalenzklasse**.

Beispiel 3.5: Aus den vorigen beiden Beispielen folgt, dass eine Klasse von gerichteten Strecken, die durch einen Verschiebungsvektor dargestellt wird, eine Äquivalenzklasse ist.

3.1.2 Gruppen

3.1.2.1 Einführung

Die Definition einer Gruppe ist ebenfalls aus dem Mathematik-Blockkurs bekannt:

Definition 3.6: Eine nichtleere Menge A , zwischen deren Elementen eine Verknüpfung „ \circ “ definiert ist, heißt eine **Gruppe**, wenn für die Verknüpfung „ \circ “ folgende Eigenschaften gelten:

1. Abgeschlossenheit: Für alle $x, y \in A$ ist $x \circ y$ eindeutig definiert und ist ein Element von A .

$$\forall x, y \in A \quad x \circ y \in A$$

2. Assoziativität: Die Verknüpfung ist assoziativ, d.h. es gilt :

$$\forall x, y, z \in A \Rightarrow (x \circ y) \circ z = x \circ (y \circ z)$$

3. Neutralement: Es existiert ein eindeutiges neutrales Element $n \in A$ mit der Eigenschaft:

$$\forall x \in A \Rightarrow x \circ n = n \circ x = x$$

4. Inverses Element: Zu jedem $x \in A$ existiert ein eindeutiges inverses Element $x^{-1} \in A$ mit

$$x \circ x^{-1} = x^{-1} \circ x = n.$$

Definition 3.7: Eine Gruppe, für die zusätzlich zu den Gruppeneigenschaften noch die Kommutativität

$$\forall x, y \in A \Rightarrow x \circ y = y \circ x$$

erfüllt ist, heißt **kommutative oder Abelsche Gruppe**.

Es gibt auch wichtige **nichtkommutative** Gruppen. Ein Beispiel ist die Matrizenmultiplikation.

Bemerkung 3.8: Man erkennt an der Definition einer Gruppe, dass die Kommutativität für das inverse und das neutrale Element auch für nichtkommutative Gruppen gilt. Diese Eigenschaften gehören eigentlich nicht zur Definition einer Gruppe, sie gelten aber dennoch für jede Gruppe.

BEWEIS Es gilt: $x \circ x^{-1} = n$, d.h. x^{-1} ist das inverse Element von x . Die Frage ist, ob x auch das inverse Element von x^{-1} ist, d.h. ob $x^{-1} \circ x = n$ gilt. Es sei z das inverse Element von x^{-1} , also $x^{-1} \circ z = n$. Dann ist

$$\begin{aligned} n &= x^{-1} \circ z \\ &= (x^{-1} \circ n) \circ z \\ &= x^{-1} \circ (n \circ z) \\ &= x^{-1} \circ ((x \circ x^{-1}) \circ z) \\ &= x^{-1} \circ (x \circ (x^{-1} \circ z)) \\ &= x^{-1} \circ (x \circ n) \\ &= x^{-1} \circ x \end{aligned}$$

Damit lässt sich auch die Frage beantworten, ob $n \circ x = x$ ist. Es gilt:

$$\begin{aligned} n \circ x &= (x \circ x^{-1}) \circ x \\ &= x \circ (x^{-1} \circ x) \\ &= x \circ n \\ &= x \end{aligned}$$

■

Bemerkung 3.9: Auch die Eindeutigkeit des neutralen und inversen Elementes gehört nicht nativ zur Gruppdefinition, gilt jedoch auch grundsätzlich.

BEWEIS Für ein weiteres neutrales Element n' folgt:

$$n' = n' \circ n = n$$

Es seien y und y' zwei inverse Elemente zu x . Dann folgt:

$$y' = y' \circ n = y' \circ (x \circ y) = (y' \circ x) \circ y = n \circ y = y$$

■

Beispiel 3.10: Zu untersuchen ist, ob A bzgl. \circ eine Gruppe bildet:

1. $A = \mathbb{R}$, \circ ist die gewöhnliche Addition. Das neutrale Element ist die Null, das inverse Element von a ist $-a$.
2. $A = \mathbb{R}^n$, \circ ist die Vektoraddition. Das neutrale Element ist der Nullvektor $\vec{0}$, der zu \vec{a} inverse Vektor ist $-\vec{a}$.
3. $A = \mathbb{R}$, \circ ist die gewöhnliche Multiplikation. A bildet keine Gruppe. Test:

Abgeschlossenheit: $a \cdot b$ ist eindeutig und $\in A$

Assoziativität: Folgt aus der Assoziativität der Multiplikation.

Neutrales Element: $n = 1$. Damit ist $n \in A$

Inverses Element: Existiert nicht für $a = 0$

Damit ist A keine Gruppe.

4. $A = \mathbb{R}_{\neq 0}$, \circ ist die gewöhnliche Multiplikation. Das neutrale Element ist die Eins, das inverse Element von a ist $a^{-1} = 1/a$, also ist A eine Gruppe.
5. Vektoren im \mathbb{R}^3 bilden bzgl. des Vektorproduktes **keine** Gruppe. Test:

Abgeschlossenheit: $\vec{a} \times \vec{b}$ ist eindeutig und $\in A$

Assoziativität: $\vec{a} \times (\vec{b} \times \vec{c}) \neq (\vec{a} \times \vec{b}) \times \vec{c}$

Damit ist A keine Gruppe.

3.1.2.2 Restklassen

Definition 3.11: Ist m eine feste natürliche Zahl, so heißen zwei ganze Zahlen a, b **nach dem Modul m kongruent** oder **kongruent modulo m** wenn die Differenz $a - b$ durch m teilbar ist. Man schreibt dann

$$a \equiv b \pmod{m} \quad \text{oder nur} \quad a \equiv b.$$

Beispiel 3.12: Es gilt:

$$-11 \equiv -7 \equiv -3 \equiv 1 \equiv 5 \equiv 9 \equiv 13 \pmod{4}$$

Satz 3.13: Ist m fest gewählt, so ist die Relation $\equiv \pmod{m}$ eine Äquivalenzrelation in \mathbb{Z} (Beweis als Übungsaufgabe).

Damit ist die Menge aller Zahlen, die bezüglich eines Moduls kongruent sind, eine Äquivalenzklasse. Die Zahlen innerhalb dieser Klasse ergeben bei der Division durch m den gleichen Rest. Diese Äquivalenzklassen heißen **Restklassen**. Durch jedes m wird die Menge der ganzen Zahlen in genau m Restklassen geteilt. Die Restklassen selbst bezeichnen wir mit

$$a + m\mathbb{Z} = \{b; b = a + km; k \in \mathbb{Z}\}.$$

Bei fest vorgegebenem m kann man jeder Restklasse eindeutig einen Divisionsrest $r \in \{0, 1, \dots, m-1\}$ zuordnen. Wir betrachten im Folgenden den Divisionsrest als „Repräsentanten“ der Klasse.

Bemerkung 3.14: Die Menge der möglichen Restklassen mod m bilden eine Gruppe bzgl. der Verknüpfung $a \circ b = (a + b) \pmod{m}$, wobei a und b die Repräsentanten der Restklassen sind (Beweis als Übungsaufgabe).

Beispiel 3.15: Für $m = 5$ ist

$$3 \circ 4 = (3 + 4) \pmod{5} = 2$$

Mit Klassen statt Repräsentanten geschrieben:

$$(3 + 5\mathbb{Z}) \circ (4 + 5\mathbb{Z}) = (2 + 5\mathbb{Z})$$

3.1.3 Körper

In vielen Fällen gibt es zwei Operationen wie z.B. Addition und Multiplikation bei reellen oder bei komplexen Zahlen, wobei die beiden Operationen durch Interaktion zusammenhängen. Dies führt zu der ebenfalls aus dem Blockkurs Mathematik bereits bekannten Definition eines Körpers.

Definition 3.16: Eine nichtleere Menge A heißt **Körper**, wenn zwischen den Elementen von A zwei Verknüpfungen „ \oplus “ und „ \odot “ mit folgenden Eigenschaften definiert sind :

1. A bildet eine kommutative Gruppe bezüglich \oplus .
2. Nach Ausschluss des neutralen Elements der Verknüpfung \oplus bilden die restlichen Elemente von A eine kommutative Gruppe bezüglich \odot , d.h.

$$\forall x, y \in A \Rightarrow x \odot y = y \odot x$$

3. Für die beiden Verknüpfungen \oplus und \odot gelten die Distributivgesetze :

$$\text{a) } \forall x, y, z \in A \Rightarrow (x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$$

$$\text{b) } \forall x, y, z \in A \Rightarrow z \odot (x \oplus y) = (z \odot x) \oplus (z \odot y)$$

Das neutrale Element der Addition \oplus wird meist als „0“ bezeichnet (Null-Element), das neutrale Element der Multiplikation \odot meist als „1“ (Eins-Element). Im folgenden Satz werden die Elemente als $\mathbb{0}$ und $\mathbb{1}$ geschrieben, um eine Verwechslung mit den Zahlen 0 und 1 zu vermeiden.

Die Voraussetzung, dass die $\mathbb{0}$ - das neutrale Element bzgl. der Addition - in der Eigenschaft 2 ausgeschlossen wird, ist notwendig, weil aus den anderen Körperaxiomen folgt, dass $\mathbb{0}^{-1}$ nicht existieren kann, es sei denn, es gilt $\mathbb{1}=\mathbb{0}$. Dies folgt aus:

Satz 3.17: Es ist $a \odot \mathbb{0} = \mathbb{0} \quad \forall a \in K$.

BEWEIS Es ist

$$a \odot (\mathbb{0} + \mathbb{0}) = a \odot \mathbb{0} + a \odot \mathbb{0} \quad \text{und} \quad a \odot (\mathbb{0} + \mathbb{0}) = a \odot \mathbb{0},$$

also

$$a \odot \mathbb{0} + a \odot \mathbb{0} = a \odot \mathbb{0} \quad \implies a \odot \mathbb{0} = \mathbb{0}.$$

■

Wenn also $x = \mathbb{0}^{-1}$ existiert, dann muss gelten $x \odot \mathbb{0} = \mathbb{1}$. Es gilt aber $x \odot \mathbb{0} = \mathbb{0}$, also $\mathbb{0} = \mathbb{1}$.

Prinzipiell ist das möglich. Für einen solchen Körper müsste allerdings gelten: $a = \mathbb{1} \odot a = \mathbb{0} \odot a = \mathbb{0}$, d.h. ein solcher Körper hätte nur ein einziges Element 0 und die Operationen:

$$0 \oplus 0 = 0$$

$$0 \odot 0 = 0$$

Allerdings fordert Eigenschaft 2 eines Körpers, dass die neutralen Elemente bezüglich der Addition und der Multiplikation unterschiedlich sind. Somit ist $\mathbb{0}=\mathbb{1}$ ausgeschlossen.

3.1.4 Kartesisches Produkt

Wir benötigen noch eine Definition, um gewisse mathematische Formulierungen zu vereinfachen.

Definition 3.18: Unter dem **kartesischen Produkt** $X \times Y$ zweier Mengen X und Y versteht man die Menge

$$X \times Y = \{(x, y); x \in X, y \in Y\}.$$

Beispiel 3.19: Falls $x \in \mathbb{R}$ und $y \in \mathbb{R}$, dann ist das Wertepaar $(x, y) \in \mathbb{R}^2$.

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y); x \in \mathbb{R}, y \in \mathbb{R}\}.$$

Beispiel 3.20: Falls $x = (x_1, x_2) \in \mathbb{R}^2$ und $y = (y_1, y_2) \in \mathbb{R}^2$, dann ist das Wertepaar $(x, y) = (x_1, x_2, y_1, y_2) \in \mathbb{R}^4$. Man differenziert nicht zwischen $((x_1, x_2), (y_1, y_2))$ und (x_1, x_2, y_1, y_2) . Deshalb gilt:

$$\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^3 \times \mathbb{R}$$

3.1.5 Vektorräume (lineare Räume)

Wir kommen jetzt zur Definition der Vektorräume, wobei diese Definition die in Kap. 2.1.1.2 besprochenen Eigenschaften des \mathbb{R}^n verallgemeinert.

Definition 3.21: Es sei K ein beliebiger Körper. Eine nicht-leere Menge V , auf der die beiden Abbildungen

$$\begin{aligned} \oplus : V \times V &\rightarrow V, & (x, y) &\mapsto x \oplus y \in V \\ \text{und } \odot : K \times V &\rightarrow V, & (\lambda, x) &\mapsto \lambda \odot x \in V \end{aligned}$$

definiert sind, heißt **Vektorraum über K** , wenn folgende Axiome gelten:

1. V ist eine kommutative Gruppe bzgl. \oplus .
2. $\forall \lambda, \mu \in K$ und $x \in V$ gilt $\lambda \odot (\mu \odot x) = (\lambda\mu) \odot x$, wobei mit $\lambda\mu$ die Multiplikation aus K gemeint ist.
3. $\forall x \in V$ gilt $\mathbb{1} \odot x = x$ ($\mathbb{1}$ ist das neutrale Element der Multiplikation aus K).
4. $\forall \lambda \in K, x, y \in V$ gilt $\lambda \odot (x \oplus y) = (\lambda \odot x) \oplus (\lambda \odot y)$.
5. $\forall \lambda, \mu \in K, x \in V$ gilt $(\lambda + \mu) \odot x = (\lambda \odot x) \oplus (\mu \odot x)$.

Erläuterung: Für die Definition eines Vektorraums benötigt man einen Körper K . Üblicherweise ist $K = \mathbb{R}$ oder $K = \mathbb{C}$. Wenn nichts anderes gesagt wird, wird immer $K = \mathbb{R}$ angenommen. Dann wird V als **reeller Vektorraum** bezeichnet. Im Falle $K = \mathbb{C}$ heißt V **komplexer Vektorraum**. Allgemein spricht man von einem **Vektorraum über dem Körper K** .

Definition 3.22: Die Elemente der Menge V eines Vektorraums heißen **Vektoren**.

Der Vektorraum setzt sich somit zusammen aus:

- einer Menge V (den Vektoren),
- der Abbildung \oplus , die zwei Vektoren addiert,
- und der Abbildung \odot , die einen Vektor mit einem Skalar (einem Element des Körpers) multipliziert.

Schließlich müssen die fünf oben genannten Axiome gelten. Die Multiplikation zweier Vektoren muss nicht definiert sein.

Beispiel 3.23: \mathbb{R}^n ist ein reeller Vektorraum. Es gilt:

$$V : \mathbb{R}^n, \quad K : \mathbb{R}$$

und mit $\vec{x}, \vec{y} \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$:

$$\begin{aligned}\vec{x} + \vec{y} &= (x_1 + y_1, \dots, x_n + y_n) \\ \alpha \cdot \vec{x} &= (\alpha x_1, \dots, \alpha x_n)\end{aligned}$$

Aber mit dieser Definition eines Vektorraumes ergeben sich auch die auf Seite 28 angedeuteten unerwarteten Strukturen.

Beispiel 3.24: $V = C[a, b]$ sei die Menge der auf dem Intervall $[a, b]$ definierten reellen stetigen Funktionen. Die Vektoren dieses Raums sind Funktionen; sie werden also nicht \vec{x}, \vec{y}, \dots geschrieben, sondern $f(x), g(x), \dots$

Der zugehörige Körper K ist wieder \mathbb{R} .

Für $f(x), g(x) \in V$ und $\alpha \in \mathbb{R}$ gilt:

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

Beispiel:

$$\begin{aligned}f(x) &= x^2 - 5 \\ g(x) &= x + 2 \\ (f + g)(x) &= f(x) + g(x) \\ &= x^2 - 5 + x + 2 \\ &= x^2 + x - 3\end{aligned}$$

$$\begin{aligned}f(x) &= \sin(x) \\ (5f)(x) &= 5f(x) = 5 \sin(x)\end{aligned}$$

Bemerkung 3.25: Die Menge der $m \times n$ -Matrizen über dem Körper K ist ein Vektorraum und wird mit $M(m \times n, K)$ bezeichnet. Die Vektoren in diesem Vektorraum sind also Matrizen. Als \oplus wählt man die Matrizenaddition und als \odot die Multiplikation der Matrizen mit einem Skalar. Der Vektorraum unterscheidet sich nur durch die Schreibweise seiner Elemente im Rechteck vom Vektorraum $K^{m \cdot n}$, und er hat die Dimension $m \cdot n$. In Kapitel 4.3 werden wir uns genauer mit diesem Vektorraum beschäftigen.

Beweis: Überprüfung der Vektorraumaxiome.

3.1.6 Untervektorräume (Unterräume)

Gewisse Teilmengen eines Vektorraums sind wieder Vektorräume.

Definition 3.26: Es sei V ein Vektorraum über einem Körper K . Eine Teilmenge $U \subset V$ heißt **Untervektorraum** oder **Unterraum** von V , wenn $U \neq \emptyset$ (leere Menge) und $\forall x, y \in U$ und alle $\lambda \in K$ gilt: $x + y \in U$, $\lambda x \in U$.

Bemerkung 3.27: Erläuterung: Es sei V ein Vektorraum. Ist $U \subset V$ ebenfalls ein Vektorraum, dann nennt man U Untervektorraum von V . Aus $U \subset V$ leiten sich mit einer Ausnahme alle Vektorraumaxiome für U ab. Nur die Abgeschlossenheit muss separat überprüft werden.

Beispiel 3.28: $V = \mathbb{R}^2$ ist ein Vektorraum.

U ist gegeben durch $U = \{(x, 0); x \in \mathbb{R}\} \Rightarrow U \subset V$.

Ist U ein Vektorraum (und damit Untervektorraum von V)?

Es sei $u_1 = (x_1, 0)$ und $u_2 = (x_2, 0)$. Dann gilt:

$$\begin{aligned} u_1 + u_2 &= (x_1 + x_2, 0) \in U \\ \lambda u_1 &= (\lambda x_1, 0) \in U \end{aligned}$$

$\Rightarrow U$ ist ein Untervektorraum von V .

Bemerkung 3.29: Ist U ein Untervektorraum von V , dann ist der Nullvektor 0 von V und zu jedem $x \in U$ auch $-x$ in U enthalten. Aus der Definition eines Vektorraums folgt nämlich, dass $0 \cdot x = 0$ und $(-1) \cdot x = -x$ ist.

Bemerkung 3.30: Es sind $\{0\}$ und V ebenfalls Unterräume von V .

Bemerkung 3.31: Sind U_1 und U_2 Unterräume von V , dann auch $U_1 \cap U_2$.

Frage: Ist $U_1 \cup U_2$ ebenfalls Unterraum?

Gegenbeispiel:

$$\begin{aligned} U_1 &= \{(x, 0); x \in \mathbb{R}\} \\ U_2 &= \{(0, y); y \in \mathbb{R}\} \end{aligned}$$

sind Unterräume des \mathbb{R}^2 .

$$U_v = U_1 \cup U_2$$

ist die Teilmenge des \mathbb{R}^2 bei der mindestens eine der beiden Koordinaten gleich 0 ist. Es sei

$$\begin{aligned} x_1 &= (1, 0) \in U_v \\ x_2 &= (0, 1) \in U_v \\ \Rightarrow x_1 + x_2 &= (1, 1) \notin U_v \end{aligned}$$

Damit ist die Abgeschlossenheit verletzt.

Definition 3.32: Es seien U_1 und U_2 zwei disjunkte Unterräume von V , d.h. $U_1 \cap U_2 = \{0\}$. Die **Summe oder direkte Summe** von U_1 und U_2 ist dann definiert durch

$$U_1 + U_2 = \{x + y; x \in U_1, y \in U_2\}$$

Satz 3.33: Sind U_1 und U_2 zwei Unterräume von V , dann ist die Summe $U_1 + U_2$ wieder Unterraum von V .

Beispiel 3.34:

$$U_1 = \{(x, 0); x \in \mathbb{R}\}$$

$$U_2 = \{(0, y); y \in \mathbb{R}\}$$

sind Unterräume des \mathbb{R}^2 .

$$U_p = U_1 + U_2 = \{(x + 0, 0 + y); x, y \in \mathbb{R}\} = \mathbb{R}^2$$

3.2 Lineare Unabhängigkeit, Basis, Dimension

3.2.1 Lineare Unabhängigkeit

Bekanntlich sagt man, \mathbb{R}^2 habe die Dimension 2, \mathbb{R}^3 die Dimension 3 usw. Außerdem kann jeder Vektor

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

in der Form

$$\vec{x} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

d.h. \vec{x} als Linearkombination der drei Einheitsvektoren dargestellt werden.

Definition 3.35: Seien $v_1, v_2, \dots, v_r \in V$ Vektoren eines Vektorraums V :

- (a) Die Addition solcher Vektoren mit Vorfaktoren

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r \quad \text{mit} \quad \lambda_i \in K$$

nennt man **Linearkombinationen** von v_1, v_2, \dots, v_r .

- (b) Die Menge

$$L = \{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r; \lambda_i \in K\} \subset V$$

aller Linearkombinationen heißt **Lineare Hülle** der r Vektoren v_1, \dots, v_r .

Aus der Definition ergibt sich, dass L ein Unterraum von V ist.

Beispiel 3.36: Gegeben seien die Vektoren

$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Dann ist

$$\vec{c} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} = 1 \cdot \vec{a} + (-1) \cdot \vec{b} \in L(\vec{a}, \vec{b})$$

eine Linearkombination von \vec{a} und \vec{b} und somit auch in der Linearen Hülle

$$L(\vec{a}, \vec{b}) = \{\lambda_1 \vec{a} + \lambda_2 \vec{b}; \lambda_1, \lambda_2 \in \mathbb{R}\}$$

enthalten.

Beispiel 3.37: Beispiele im \mathbb{R}^3 :

$$L(\vec{e}_1, \vec{e}_2, \vec{e}_3) = \{(\lambda_1, \lambda_2, \lambda_3); \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}\} = \mathbb{R}^3$$

$$L(\vec{e}_1, \vec{e}_2) = \{(\lambda_1, \lambda_2, 0); \lambda_1, \lambda_2 \in \mathbb{R}\} \quad (\text{x-y-Ebene})$$

$$L(\vec{e}_1, \vec{e}_2, \vec{e}_1 + \vec{e}_2) = \{(\lambda_1 + \lambda_3, \lambda_2 + \lambda_3, 0); \lambda_1, \lambda_2 \in \mathbb{R}\} \quad (\text{x-y-Ebene})$$

Der englische Ausdruck für „Lineare Hülle“ ist **span**. Auch im Deutschen schreibt man oft $\text{span}(\vec{e}_1, \vec{e}_2)$ statt $L(\vec{e}_1, \vec{e}_2)$. Es gibt auch den deutschen Ausdruck **Spann** für die lineare Hülle, der jedoch eher ungewöhnlich ist. Gängig ist aber ein Satz wie: Die Vektoren \vec{e}_1 und \vec{e}_2 *spannen die x-y-Ebene auf*.

Definition 3.38: Spannen die Vektoren den gesamten Raum auf, d.h. gilt $L(v_1, \dots, v_r) = V$, so nennt man $\{v_1, \dots, v_r\}$ ein **Erzeugendensystem**.

Definition 3.39: Es sei V ein Vektorraum über K . Ein r -Tupel $\{v_1, \dots, v_r\} \subset V$ heißt **linear unabhängig**, wenn aus $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r = 0$ stets folgt, dass $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$ ist.

Erläuterung: Lineare Unabhängigkeit heißt:

- Keiner der Vektoren ist eine Linearkombination der übrigen.
- Keiner der Vektoren ist der Nullvektor.

Beispiel 3.40: (\mathbb{R}^n)

Zwei Vektoren \vec{v}_1 und \vec{v}_2 sind linear abhängig, wenn die Gleichung

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = 0$$

mit λ_1 oder $\lambda_2 \neq 0$ gilt. Sagen wir, die Gleichung gelte für $\lambda_1 \neq 0$ (für $\lambda_2 \neq 0$ ist die Rechnung äquivalent). Dann ist

$$\vec{v}_1 = -\frac{\lambda_2}{\lambda_1} \vec{v}_2.$$

Diese Gleichung ist erfüllt, wenn entweder $\vec{v}_1 = 0$ und $\lambda_2 = 0$ (aber $\lambda_1 \neq 0$) ist oder \vec{v}_1 und \vec{v}_2 parallel sind. Zwei Vektoren des \mathbb{R}^n sind also genau dann linear abhängig, wenn sie entweder parallel sind oder einer der beiden Vektoren der Nullvektor ist.

Für n Vektoren $\vec{v}_1, \dots, \vec{v}_n$ gilt, dass Sie genau dann linear abhängig sind, wenn

$$\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = 0$$

gilt, wobei mindestens einer der Vorfaktoren $\lambda_1, \dots, \lambda_n \neq 0$ ist. Sagen wir $\lambda_1 \neq 0$ (für die anderen λ_i ist die Rechnung äquivalent). Dann gilt die Gleichung in zwei Fällen:

1. $\vec{v}_1 = 0$. Dann können alle anderen $\lambda_i = 0$ sein.
2. $\vec{v}_1 = -\frac{1}{\lambda_1}(\lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n)$. Dann ist \vec{v}_1 eine Linearkombination der anderen \vec{v}_i .

Bemerkung 3.41: Vektoren v aus der linearen Hülle der $\{v_k\}$ sind eindeutig darstellbar, wenn die $\{v_k\}$ linear unabhängig sind.

BEWEIS Es seien

$$\begin{aligned} v &= \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n \\ &= \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n \end{aligned}$$

zwei verschiedene Darstellungen von v . Dann folgt daraus:

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \dots + (\alpha_n - \beta_n)v_n = 0$$

Sind v_1, \dots, v_n linear unabhängig, sind alle Differenzen $\alpha_n - \beta_n$ gleich 0. ■

3.2.1.1 Nachweis der linearen Unabhängigkeit im Vektorraum \mathbb{R}^n

Die Gleichung

$$\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = 0$$

ist ein lineares Gleichungssystem, das sich für beliebige n lösen lässt.

Beispiel 3.42:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}$$

führt zum Gleichungssystem

$$\begin{pmatrix} \lambda_1 + 2\lambda_2 - \lambda_3 = 0 \\ 2\lambda_1 + 3\lambda_2 + 2\lambda_3 = 0 \\ \lambda_1 + \lambda_2 + 5\lambda_3 = 0 \end{pmatrix}$$

Die Lösung nach dem Gauß-Schema ergibt:

$$\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 2 & 3 & 2 & 0 & | & -2 \cdot I \\ 1 & 1 & 5 & 0 & | & -I \\ \hline 1 & 2 & -1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & -1 & 6 & 0 & | & -II \\ \hline 1 & 2 & -1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & 0 & 2 & 0 \end{array}$$

Daraus folgt $\lambda_3 = \lambda_2 = \lambda_1 = 0$.

Beispiel 3.43: Was würde passieren, wenn die Vektoren linear abhängig wären?

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

ergibt

$$\begin{array}{ccc|c}
 1 & 2 & -1 & 0 \\
 2 & 3 & 2 & 0 & | & -2 \cdot I \\
 1 & 1 & 3 & 0 & | & -I \\
 \hline
 1 & 2 & -1 & 0 \\
 0 & -1 & 4 & 0 \\
 0 & -1 & 4 & 0 & | & -II \\
 \hline
 1 & 2 & -1 & 0 \\
 0 & -1 & 4 & 0 \\
 0 & 0 & 0 & 0
 \end{array}$$

Es gibt nur noch zwei Gleichungen mit drei Unbekannten, so dass zur Lösung $\lambda_1=\lambda_2=\lambda_3=0$ (Trivallösung) noch andere Lösungen hinzukommen. In diesem Fall sind dies:

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \alpha \cdot \begin{pmatrix} -7 \\ 4 \\ 1 \end{pmatrix}.$$

3.2.1.2 Lineare Unabhängigkeit von n Vektoren des \mathbb{R}^n

Dieser Fall führt auf ein $n \times n$ -Gleichungssystem. Wir haben gesehen, dass sich ein 3×3 -Gleichungssystem genau dann eindeutig lösen lässt, wenn die Determinante der Koeffizientenmatrix A ungleich 0 ist. In Kapitel 4.4.2 werden wir sehen, dass dies nicht nur für 3×3 -, sondern allgemein für quadratische Gleichungssysteme gilt. Da $\lambda_1 = \dots = \lambda_n = 0$ auf jeden Fall Lösung ist, folgt:

$$\begin{aligned}
 \det(A) \neq 0 &\Leftrightarrow \text{Lösung des Gleichungssystems ist eindeutig} \\
 &\Leftrightarrow \text{Nur } \lambda_1 = \dots = \lambda_n = 0 \text{ ist Lösung} \\
 &\Leftrightarrow \text{Die Vektoren sind linear unabhängig.}
 \end{aligned}$$

Es würde in den vorigen Beispielen also schon ausreichen, die Determinanten der Koeffizientenmatrizen zu prüfen. Für das erste, linear unabhängige Beispiel gilt:

$$\det \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 2 \\ 1 & 1 & 5 \end{pmatrix} = -2$$

Für das zweite, linear abhängige Beispiel ist:

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

3.2.1.3 Nachweis der Linearen Unabhängigkeit im Vektorraum $C[a, b]$

Formal ist hier die identische Eigenschaft wie im Vektorraum \mathbb{R}^n zu prüfen, es ergeben sich jedoch unterschiedliche Ausprägungen:

$$\begin{array}{l}
 \text{statt: } \lambda_1 \vec{v}_1 + \dots + \lambda_r \vec{v}_r = \boxed{0} \Rightarrow \lambda_1 = \dots = \lambda_r = 0 \\
 \text{jetzt: } \lambda_1 f_1(x) + \dots + \lambda_r f_r(x) = \boxed{0} \Rightarrow \lambda_1 = \dots = \lambda_r = 0
 \end{array}$$

Nullfunktion $0(x)$
Ergibt 0 für alle x .

Beispiel 3.44: Beispiele von linear abhängigen Funktionen:

- Die Funktionen

$$f_1(x) = x; \quad f_2(x) = x^2; \quad f_3(x) = x^2 - 2x$$

sind auf einem beliebigen Intervall $[a, b]$ linear abhängig, denn

$$2f_1(x) - f_2(x) + f_3(x) = 2x - x^2 + x^2 - 2x = 0 \quad \forall x \in \mathbb{R}$$

- Die Funktionen

$$f_1(x) = \sin^2(x); \quad f_2(x) = \cos^2(x); \quad f_3(x) = 1$$

sind auf einem beliebigen Intervall $[a, b]$ linear abhängig, denn

$$f_1(x) + f_2(x) - f_3(x) = \cos^2(x) + \sin^2(x) - 1 = 0 \quad \forall x \in \mathbb{R}$$

Beispiel 3.45: Beispiele von linear unabhängigen Funktionen:

Die Funktionen

$$f_1(x) = x^2; \quad f_2(x) = x$$

sind auf einem beliebigen Intervall $[a, b]$ linear unabhängig. Beweis: Zur Lösung der Gleichung

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) = \lambda_1 x^2 + \lambda_2 x = 0$$

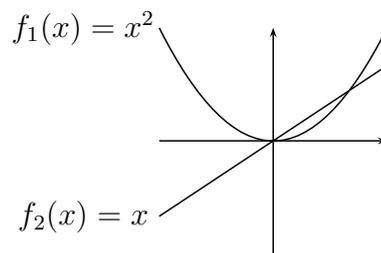
ist eine Fallunterscheidung nötig.

$$(1) \lambda_1 = 0 \Rightarrow \lambda_2 x = 0$$

Dies ist für alle x nur der Fall, wenn $\lambda_2 = 0$.

$$(2) \lambda_1 \neq 0 \Rightarrow x^2 + \frac{\lambda_2}{\lambda_1} x = 0$$

Diese Gleichung ist erfüllt für $x = 0$ und $x = -\frac{\lambda_2}{\lambda_1}$, aber nicht für alle x . Somit bleibt für alle x als einzige Lösung $\lambda_1 = \lambda_2 = 0$. $f_1(x)$ und $f_2(x)$ sind also linear unabhängig.



Allgemeines Testverfahren auf Lineare Unabhängigkeit: Wie überprüft man, ob eine Menge von Funktionen linear unabhängig auf einem Intervall $[a, b]$ ist? Es ist zu ermitteln, ob die Gleichung

$$l(x) = \lambda_1 f_1(x) + \dots + \lambda_n f_n(x) = 0 \quad \forall x \in [a, b] \quad (1)$$

andere Lösungen außer $\lambda_1 = \dots = \lambda_n = 0$ besitzt.

Ansatz: Man wählt willkürlich n x -Werte aus dem Intervall aus und ersetzt Gleichung (1) durch:

$$l(x) = \lambda_1 f_1(x) + \dots + \lambda_n f_n(x) = 0 \quad \text{für } x = x_1, \dots, x_n \quad (2)$$

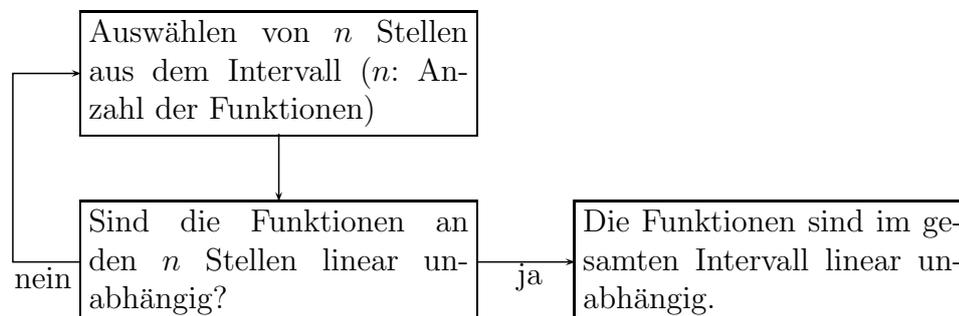
(2) unterscheidet sich von (1) dadurch, dass $\lambda_1 f_1(x) + \dots + \lambda_n f_n(x)$ nicht mehr auf dem ganzen Intervall gleich 0 sein muss, sondern nur an den ausgewählten Punkten. Dies ist wesentlich leichter nachzuprüfen. Man kann jetzt Folgendes sagen:

- $\lambda_1 = \dots = \lambda_n = 0$ ist auf jeden Fall Lösung von (1) und (2).
- Andere Lösungen von (1) müssen auch Lösungen von (2) sein.
- Andere Lösungen von (2) müssen nicht Lösungen von (1) sein, denn $l(x)$ kann auf anderen Punkten als den ausgewählten immer noch $\neq 0$ sein.

Das bedeutet:

- Wenn (2) nur die Lösung $\lambda_1 = \dots = \lambda_n = 0$ hat, dann hat auch (1) nur diese Lösung und die Funktionen sind linear unabhängig.
- Wenn (2) andere Lösungen als $\lambda_1 = \dots = \lambda_n = 0$ hat, kann man gar nichts sagen, denn diese Lösungen müssen nicht Lösung von (1) sein. Man kann mit dieser Methode keine lineare Abhängigkeit beweisen, sondern es nur mit anderen Stellen noch einmal probieren.

Darauf baut sich das folgende Lösungsschema auf:



Wie man leicht sieht, führt das Verfahren in eine Endlosschleife, falls die Funktionen linear abhängig sind. Hierzu benötigt man andere Verfahren.

Beispiel 3.46:

$$f_1(x) = 1; \quad f_2(x) = x^2; \quad f_3(x) = x^4 \quad \text{im Intervall } [-3, 3]$$

Daraus folgt die Gleichung $\lambda_1 + \lambda_2 x^2 + \lambda_3 x^4 = 0$. Wähle 3 x -Werte, z.B. -1, 1 und 2. Daraus ergibt sich das Gleichungssystem

$$\begin{aligned} x = -1 &\Rightarrow \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ x = 1 &\Rightarrow \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ x = 2 &\Rightarrow \lambda_1 + 4\lambda_2 + 16\lambda_3 = 0 \end{aligned}$$

Die Determinante ist

$$\det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 4 & 16 \end{pmatrix} = 0$$

Es ist noch nicht möglich, eine Aussage zu machen. Wähle neue x -Werte, z.B. 0, 1 und 2.

$$\begin{aligned} x = 0 &\Rightarrow \lambda_1 + 0\lambda_2 + 0\lambda_3 = 0 \\ x = 1 &\Rightarrow \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ x = 2 &\Rightarrow \lambda_1 + 4\lambda_2 + 16\lambda_3 = 0 \end{aligned}$$

Die Determinante ist

$$\det \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 4 & 16 \end{pmatrix} = 12$$

Also sind 1, x^2 und x^4 linear unabhängig. Im Verlauf der Vorlesung werden wir noch eine einfachere Methode kennenlernen, die lineare Unabhängigkeit von Polynomen zu überprüfen.

3.2.2 Basis und Dimension

Definition 3.47: Es sei V ein Vektorraum über K . $\{v_1, \dots, v_n\} \subset V$ heißt **Basis** oder **minimales Erzeugendensystem** von V , wenn $\{v_1, \dots, v_n\}$ linear unabhängig sind und V deren Lineare Hülle ist.

Definition 3.48: Sei $\{v_1, \dots, v_n\}$ eine Basis von V und $v \in V$ ein beliebiger Vektor, dann heißen die Vorfaktoren $(\lambda_1, \dots, \lambda_n) \in K^n$ der Linearkombination

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

Koordinaten des Vektors v bzgl. der Basis $\{v_1, \dots, v_n\}$.

Bemerkung 3.49: Die Koordinaten eines Vektors bzgl. einer gegebenen Basis sind eindeutig.

BEWEIS Da V die lineare Hülle von $\{v_1, \dots, v_n\}$ ist, existiert zu jedem $v \in V$ mindestens ein $(\lambda_1, \dots, \lambda_n) \in K^n$, so dass gilt:

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

Da $\{v_1, \dots, v_n\}$ linear unabhängig sind, ist diese Darstellung eindeutig. ■

Definition 3.50: Besitzt der Vektorraum V eine Basis $\{v_1, \dots, v_n\}$, so heißt n die **Dimension** von V .
Bezeichnung : $\dim(V) = n$.

Frage: Ist die Dimension n eindeutig, wenn sie (in \mathbb{N}) existiert? Zur Klärung dieser Frage benötigen wir 3 Sätze:

- Basisergänzungssatz
- Austauschlemma¹
- Eindeutigkeit der Dimension

Satz 3.51: (Basisergänzungssatz) Sei V ein Vektorraum über K und seien

$$v_1, \dots, v_r, w_1, \dots, w_s$$

Vektoren in V . Ist $\{v_1, \dots, v_r\}$ linear unabhängig und ist

$$L(v_1, \dots, v_r, w_1, \dots, w_s) = V,$$

dann kann man $\{v_1, \dots, v_r\}$ durch evtl. Hinzunahme geeigneter Vektoren aus $\{w_1, \dots, w_s\}$ zu einer Basis von V ergänzen.

¹Ein Lemma ist ein Satz, der seine Bedeutung vor allem als Hilfssatz im Beweis eines wichtigeren Satzes hat.

BEWEIS Falls schon $L(v_1, \dots, v_r) = V$, dann ist nichts mehr zu beweisen.

Wir betrachten also den Fall

$$L(v_1, \dots, v_r) \neq V.$$

Dann existiert mindestens ein w_i mit $w_i \notin L(v_1, \dots, v_r)$; denn wären alle $w_i \in L(v_1, \dots, v_r)$, dann müsste $L(v_1, \dots, v_r) = L(v_1, \dots, v_r, w_1, \dots, w_s) = V$ gelten, was der Annahme $L(v_1, \dots, v_r) \neq V$ widerspricht.

Die Menge $\{w_i, v_1, \dots, v_r\}$ ist linear unabhängig, denn aus $\sum_{j=1}^r \lambda_j v_j + \lambda w_i = 0$ folgt $\lambda = 0$, da $w_i \notin L(v_1, \dots, v_r)$, und weiter folgt dann $\lambda_j = 0 \forall j$, da alle v_j linear unabhängig sind.

Möglicherweise ist die Menge $\{w_i, v_1, \dots, v_r\}$ aber noch keine Basis. Dann wird der vorige Schritt wiederholt und es werden solange weitere w_i dazugenommen, bis sie $\{v_1, \dots, v_r\}$ zu einer Basis von V ergänzen. Dies ist nach endlich vielen Schritten möglich, da sich ergibt: $L(v_1, \dots, v_r, w_1, \dots, w_s) = V$. ■

Beispiel 3.52: Es sei V der \mathbb{R}^3 . Weiterhin sei

$$\vec{v}_1 = (1, 0, 0); \quad \vec{v}_2 = (0, 1, 0).$$

\vec{v}_1 und \vec{v}_2 sind linear unabhängig. Außerdem sei

$$\vec{w}_1 = (1, 1, 0); \quad \vec{w}_2 = (0, 0, 1).$$

Da $L(\vec{v}_1, \vec{v}_2, \vec{w}_1, \vec{w}_2) = \mathbb{R}^3$, sind die Voraussetzungen für den Basisergänzungssatz erfüllt.

Der Satz sagt: Da $\{\vec{v}_1, \vec{v}_2\}$ keine Basis des \mathbb{R}^3 ist, ist mindestens eine der Kombinationen $\{\vec{v}_1, \vec{v}_2, \vec{w}_1\}$, $\{\vec{v}_1, \vec{v}_2, \vec{w}_2\}$ oder $\{\vec{v}_1, \vec{v}_2, \vec{w}_1, \vec{w}_2\}$ eine Basis des \mathbb{R}^3 , d.h. linear unabhängig und mit \mathbb{R}^3 als linearer Hülle. In diesem Beispiel sind das die Vektoren $\{\vec{v}_1, \vec{v}_2, \vec{w}_2\}$. Alle vier Vektoren sind keine Basis des \mathbb{R}^3 , da sie linear abhängig sind.

Satz 3.53: (Austauschlemma): Sind $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_m\}$ Basen eines Vektorraums V über K , dann gibt es zu jedem v_i ein w_j , so dass aus $\{v_1, \dots, v_n\}$ wieder eine Basis entsteht, wenn man in ihr v_i durch w_j ersetzt.

Beispiel 3.54: Ein einfaches Beispiel im \mathbb{R}^3 :

$$\vec{v}_1 = (1, 0, 0); \quad \vec{v}_2 = (0, 1, 0); \quad \vec{v}_3 = (0, 0, 1)$$

$$\vec{w}_1 = (2, 0, 0); \quad \vec{w}_2 = (0, 2, 0); \quad \vec{w}_3 = (0, 0, 3)$$

$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ und $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ sind Basen des \mathbb{R}^3 . Aus der ersten Basis wird \vec{v}_3 entfernt. Der Satz sagt jetzt: Ich kann dafür einen der drei \vec{w}_i -Vektoren einsetzen, so dass wieder eine Basis entsteht.

Im Beispiel ist dies \vec{w}_3 . $\{\vec{v}_1, \vec{v}_2, \vec{w}_3\}$ ist eine Basis des \mathbb{R}^3 .

BEWEIS Es seien $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_m\}$ zwei Basen von V . Aus der ersten Basis wird der Vektor v_i gestrichen. Für die restliche Menge $\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\}$ gilt

$$L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) \neq V,$$

denn wäre $L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) = V$, dann wäre auch $v_i \in L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$, ließe sich also durch eine Linearkombination der restlichen Vektoren darstellen. Damit wären $\{v_1, \dots, v_n\}$ linear abhängig und keine Basis, was in Widerspruch zur Voraussetzung steht.

Damit existiert nach dem Basisergänzungssatz ein w_j mit

$$w_j \notin L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$$

und die Vektoren

$$\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n, w_j\}$$

sind linear unabhängig.

Sind die Vektoren auch eine Basis? Wären sie keine Basis, könnten Sie durch einen der Vektoren

$$\{v_1, \dots, v_n\}$$

zu einer Basis ergänzt werden. Alle Vektoren außer v_i kommen dazu trivialerweise nicht in Frage. Aber auch v_i ist nicht möglich. In diesem Fall wäre die Basis nämlich $\{v_1, \dots, v_n, w_j\}$. Da aber bereits $\{v_1, \dots, v_n\}$ eine Basis ist, ist w_j linear abhängig von diesen Vektoren und $\{v_1, \dots, v_n, w_j\}$ ist keine Basis. Damit muss

$$\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n, w_j\}$$

bereits eine Basis von V sein. ■

Damit kommt man jetzt zur Eindeutigkeit der Dimension eines endlichdimensionalen Raumes.

Satz 3.55: (Eindeutigkeit der Dimension): Sind $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_m\}$ Basen eines Vektorraums V über K , dann ist $n = m$.

BEWEIS Nimmt man $n < m$ an (der Fall $n > m$ ist äquivalent), dann kann man mit Hilfe des Austauschlemmas aus der Basis $\{w_1, \dots, w_m\}$ eine Basis $\{v_1, \dots, v_n, w_{n+1}, \dots, w_m\}$ herstellen. Nun war aber $\{v_1, \dots, v_n\}$ bereits eine Basis, so dass $\{v_1, \dots, v_n, w_{n+1}, \dots, w_m\}$ linear abhängig sein müssen. Dies ist ein Widerspruch zur linearen Unabhängigkeit der Basiselemente. ■

Damit können wir folgende Bemerkung aufstellen, die uns bei der Suche nach einer geeigneten Basis für einen vorgegeben Vektorraum einen entscheidenden Hinweis liefert:

Bemerkung 3.56: Ist $\dim(V) = n$ und sind $\{\vec{v}_1, \dots, \vec{v}_n\}$ n linear unabhängige Vektoren in V , so ist $\{\vec{v}_1, \dots, \vec{v}_n\}$ eine Basis von V .

BEWEIS Wäre $\{\vec{v}_1, \dots, \vec{v}_n\}$ keine Basis von V , so könnte man $\{\vec{v}_1, \dots, \vec{v}_n\}$ nach dem Basisergänzungssatz um $k \geq 1$ Vektoren aus einer Basis $\{\vec{w}_1, \dots, \vec{w}_n\}$ zu einer Basis ergänzen. Dann jedoch gäbe es eine Basis der Dimension $n+k$, was ein Widerspruch zur Eindeutigkeit der Dimension ist. ■

Beispiel 3.57: Zeigen Sie, dass $\vec{v}_1 = (1, 0)$ und $\vec{v}_2 = (0, 1)$ eine Basis des \mathbb{R}^2 sind.

$$\lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0$$

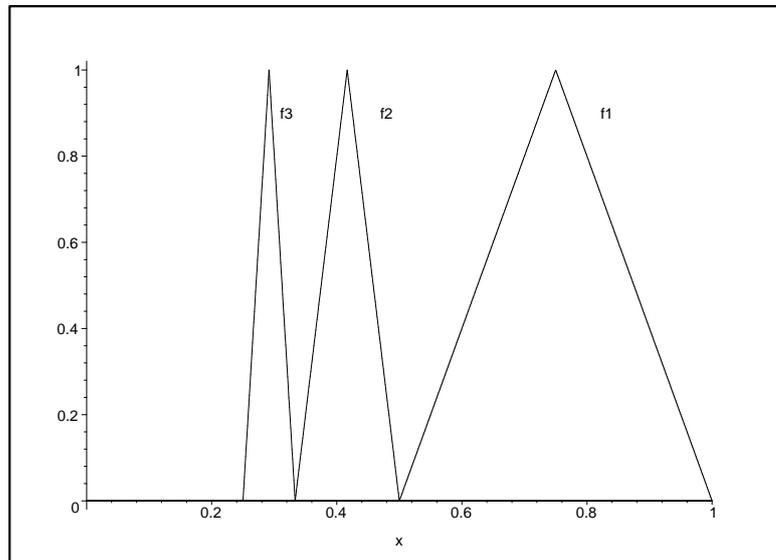
$\Rightarrow \vec{v}_1$ und \vec{v}_2 sind linear unabhängig. Da die Dimension des \mathbb{R}^2 gleich 2 ist, sind die beiden Vektoren auch eine Basis des \mathbb{R}^2 .

3.2.3 Unendliche Mengen linear unabhängiger Vektoren

Beispiel 3.58: Beispiel einer unendlichen Menge linear unabhängiger Elemente in $C[0, 1]$. Es seien stetige Funktionen f_n definiert durch

$$f_n(x) = \begin{cases} 0, & x \notin [\frac{1}{n+1}, \frac{1}{n}] \\ 1, & x = (\frac{1}{n+1} + \frac{1}{n})/2 \\ \text{linear auf den Zwischenintervallen} \end{cases}, n \in \mathbb{N}$$

Die Menge $\{f_1, \dots, f_n\}$ ist linear unabhängig für jedes $n \in \mathbb{N}$. Die folgende Abbildung zeigt die Funktionen f_1, f_2, f_3 , wobei die "Dreiecke" von rechts nach links jeweils die Graphen der Funktionen f_1, f_2, f_3 darstellen. Außerhalb sind die Funktionswerte jeweils gleich Null.



3.3 Polynome

3.3.1 Polynome und Lineare Unabhängigkeit

Die einfachste Anwendung der Theorie der Vektorräume im Bereich der Funktionen ist der Vektorraum der Polynome. Polynome sind einfache Funk-

tionen mit sehr vielen weiteren speziellen Eigenschaften, die mit algebraischen Mitteln behandelt werden können.

Definition 3.59: Ein Polynom (in x) (oder ganzrationale Funktion) ist ein Ausdruck der Form

$$\begin{aligned} p(x) \equiv p_n(x) &= a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = \sum_{k=0}^n a_kx^k \\ &= a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = \sum_{k=0}^n a_{n-k}x^{n-k} \end{aligned}$$

Die Zahlen $a_k \in \mathbb{R}$, $k = 0, 1, 2, \dots, n$, heißen **Koeffizienten**, a_n heißt **Leitkoeffizient**. Ist $a_n \neq 0$, so heißt n **Grad** des Polynoms. Im Fall $a_n = 1$ heißt das Polynom normiert.

Eine wichtige Eigenschaft ist das Abdividieren von Nullstellen. In den folgenden Sätzen soll gezeigt werden, dass ein Polynom vom Grad n höchstens n reelle Nullstellen hat. Dazu gibt es zunächst ein einfaches Ergebnis:

Satz 3.60: Ist x_0 eine beliebige reelle Zahl und $n \in \mathbb{N}$, dann gilt für alle $x \in \mathbb{R}$

$$p_n(x) = (x - x_0)p_{n-1}(x) + r,$$

wobei p_{n-1} ein Polynom vom Grade $n - 1$ ist und r nicht von x abhängt.

BEWEIS Vollständige Induktion:

- Induktionsanfang: $n = 1$

$$\begin{aligned} p_1(x) &= a_1x + a_0 \\ &= a_1(x - x_0) + a_1x_0 + a_0 \\ &= a_1(x - x_0) + r \end{aligned}$$

- Induktionsannahme: Der Satz gilt für $p_n(x)$.
- Induktionsbehauptung: Der Satz gilt auch für $p_{n+1}(x)$
- Beweis:

$$\begin{aligned} p_{n+1}(x) &= p_n(x)x + a_0 \\ &= ((x - x_0)p_{n-1}(x) + r) \cdot x + a_0 \\ &= x(x - x_0)p_{n-1}(x) + rx + a_0 + \overbrace{rx_0 - rx_0}^{=0} \\ &= x(x - x_0)p_{n-1}(x) + r(x - x_0) + rx_0 + a_0 \\ &= (x - x_0) \overbrace{(xp_{n-1} + r)}^{\tilde{p}_n(x)} + \overbrace{rx_0 + a_0}^s \\ &= (x - x_0)\tilde{p}_n(x) + s \end{aligned}$$

■

Bemerkung 3.61: Es gilt $r = p_n(x_0)$.

BEWEIS

$$\begin{aligned} p_n(x_0) &= (x_0 - x_0)p_{n-1}(x_0) + r \\ &= r \end{aligned}$$

■

Bemerkung 3.62: Ist x_0 Nullstelle von p_n , dann gilt

$$p_n(x) = (x - x_0)p_{n-1}(x).$$

BEWEIS Ist x_0 eine Nullstelle, dann gilt $p_n(x_0) = 0$. Da aber nach der vorigen Bemerkung auch gilt $r = p_n(x_0)$, ist $r = 0$ und

$$p_n(x) = (x - x_0)p_{n-1}(x).$$

■

Die letzte Folgerung kann man jetzt auf p_{n-1} anwenden und dies solange fortsetzen, bis man Folgendes erhält:

Satz 3.63: Hat das Polynom p_n n Nullstellen, dann gilt die **Faktorzerlegung**

$$p_n(x) = (x - x_1)(x - x_2) \cdots (x - x_n)a_n.$$

Daraus folgt direkt:

Satz 3.64: Ein Polynom vom Grade n hat höchstens n Nullstellen.

Dies gilt auch im Komplexen. Außerdem können einige der Nullstellen x_k gleich sein.

Satz 3.65: Hat ein Polynom vom Grade $\leq n$ mehr als n Nullstellen, dann ist es das Nullpolynom d.h. es gilt $a_0 = a_1 = \dots = a_n = 0$. Das Nullpolynom ist das Polynom vom Grade Null, das überall den (konstanten) Funktionswert Null hat.

Daraus ergibt sich

Satz 3.66: Die Funktionen $\{1, x, x^2, \dots, x^n\}$ sind auf jedem Intervall $[a, b] \in \mathbb{R}$ linear unabhängig. Außerdem bilden diese Funktionen im Vektorraum P_n der Polynome vom Grade $\leq n$ eine Basis.

BEWEIS Lineare Unabhängigkeit bedeutet:

$$\lambda_0 + \lambda_1 x + \dots + \lambda_n x^n = 0 \quad \forall x \in [a, b] \quad \Rightarrow \quad \lambda_0 = \lambda_1 = \dots = \lambda_n = 0$$

$\lambda_0 + \lambda_1 x + \dots + \lambda_n x^n = 0 \quad \forall x \in [a, b]$ bedeutet, dass das Polynom auf dem ganzen Intervall $[a, b]$ gleich 0 sein muss, für $a \neq b$ sind das unendlich viele Nullstellen. Das sind mehr als n Nullstellen, weswegen die einzige Lösung das Nullpolynom ist, also $\lambda_0 = \lambda_1 = \dots = \lambda_n = 0$. Darum sind $\{1, x, \dots, x^n\}$ auf jedem Intervall $[a, b]$ linear unabhängig. Da $L(\{1, x, \dots, x^n\}) = P_n$ ist, bilden die Polynome auch eine Basis. ■

Der Vektorraum der Polynome P_n ist von der Dimension $n + 1$ und die Funktionen $\{1, x, x^2, \dots, x^n\}$ sind eine Basis des P_n .

Bemerkung 3.67: Zwei Polynome haben also genau dann für alle $x \in \mathbb{R}$ den gleichen Funktionswert, falls alle Koeffizienten gleich sind.

Bemerkung 3.68: Stimmen zwei Polynome vom Grade $\leq n$ in mindestens $(n + 1)$ Stellen überein, dann sind sie gleich, d.h. für

$$p_n(x) = \sum_{k=0}^n a_k x^k \quad \text{und} \quad q_n(x) = \sum_{k=0}^n b_k x^k$$

gilt dann

$$a_k = b_k, \quad k = 0, 1, \dots, n.$$

BEWEIS Das Differenzpolynom $d_n(x)$ ist definiert als

$$d_n(x) = p_n(x) - q_n(x) = \sum_{k=0}^n (a_k - b_k) x^k$$

Da $d_n(x)$ nach Voraussetzung mindestens $n + 1$ Nullstellen hat, ist $d_n(x) = 0$. Also gilt: $p_n(x) = q_n(x)$. ■

Satz 3.69: Ein Polynom p_n von Grade $\leq n$ ist durch die Angabe von $n + 1$ Punkten (x_k, y_k) mit **paarweise verschiedenen** x_k durch $y_k = p_n(x_k)$ eindeutig bestimmt. Dies ist das sogenannte **Interpolationspolynom**.

Beispiel 3.70: Nach dem obigen Satz ist die Parabel (Polynom vom Grad 2) durch die drei Punkte $(-2/1)$, $(-1/-1)$ und $(1/1)$ eindeutig. Die allgemeine Form des Polynoms ist $f(x) = ax^2 + bx + c$. Einsetzen der drei Punkte ergibt die Gleichungen:

$$\begin{aligned}(1/1) : & \quad 1 = a + b + c \\ (-1/-1) : & \quad -1 = a - b + c \\ (-2/1) : & \quad 1 = 4a - 2b + c\end{aligned}$$

Zu lösen ist also das Gleichungssystem:

$$\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 4 & -2 & 1 & 1 \end{array}$$

Die Berechnung mit Matlab ergibt:

```
>> a=[1 1 1 1;1 -1 1 -1;4 -2 1 1];
>> rref(a)
```

ans =

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

Das Ergebnis ist also $a = 1; b = 1; c = -1$. Somit ist das Ergebnispolynom:

$$f(x) = x^2 + x - 1$$

3.3.2 Test auf lineare Unabhängigkeit von Polynomen

Es seien $p_1(x), \dots, p_n(x)$ Polynome vom Grad $\leq m$. Lineare Unabhängigkeit der Polynome bedeutet, dass

$$\lambda_1 p_1(x) + \lambda_2 p_2(x) + \dots + \lambda_n p_n(x) = 0$$

keine andere Lösungen außer $\lambda_1 = \lambda_2 = \dots = 0$ besitzt. Jedes Polynom $p_i(x)$ lässt sich schreiben als

$$p_i(x) = p_{i0} + p_{i1}x + \dots + p_{im}x^m = \left\langle \begin{pmatrix} p_{i0} \\ p_{i1} \\ \vdots \\ p_{im} \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ \vdots \\ x^m \end{pmatrix} \right\rangle$$

Damit gilt:

$$\begin{aligned}
 0 &= \lambda_1 p_1(x) + \dots + \lambda_n p_n(x) \\
 &= \lambda_1 \left\langle \begin{pmatrix} p_{10} \\ p_{11} \\ \vdots \\ p_{1m} \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ \vdots \\ x^m \end{pmatrix} \right\rangle + \dots + \lambda_n \left\langle \begin{pmatrix} p_{n0} \\ p_{n1} \\ \vdots \\ p_{nm} \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ \vdots \\ x^m \end{pmatrix} \right\rangle \\
 &= \left\langle \lambda_1 \begin{pmatrix} p_{10} \\ p_{11} \\ \vdots \\ p_{1m} \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} p_{n0} \\ p_{n1} \\ \vdots \\ p_{nm} \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ \vdots \\ x^m \end{pmatrix} \right\rangle
 \end{aligned}$$

Da $\{1, x, \dots, x^m\}$ linear unabhängig ist, folgt daraus:

$$\lambda_1 \begin{pmatrix} p_{10} \\ p_{11} \\ \vdots \\ p_{1m} \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} p_{n0} \\ p_{n1} \\ \vdots \\ p_{nm} \end{pmatrix} = 0$$

Dies ist ein lineares Gleichungssystem für die Variablen $\{\lambda_1, \dots, \lambda_n\}$. Gibt es andere Lösungen, als $\lambda_1 = \dots = \lambda_n = 0$, sind die Polynome linear abhängig, sonst sind die Polynome linear unabhängig.

Beispiel 3.71: Sind die Polynome

$$p_1(x) = (1-x)^2; \quad p_2(x) = (1-x)x; \quad p_3(x) = x^2$$

linear unabhängig? Es gilt:

$$\begin{aligned}
 p_1(x) &= (1-x)^2 = x^2 - 2x + 1 = \left\langle \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \right\rangle \\
 p_2(x) &= (1-x)x = -x^2 + x = \left\langle \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \right\rangle \\
 p_3(x) &= x^2 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \right\rangle
 \end{aligned}$$

Das zu lösende Gleichungssystem ist also:

$$\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \quad | + 2 \cdot I \\ 1 & -1 & 1 & 0 \quad | - I \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \quad + II \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}$$

Damit ergibt sich als einzige Lösung $\lambda_1 = \lambda_2 = \lambda_3 = 0$ und die Polynome sind linear unabhängig.

3.4 Skalarprodukt, euklidische und unitäre Räume

3.4.1 Das Skalarprodukt

Erinnern wir uns: In Kapitel 2.1.4 hatten wir das Skalarprodukt bereits für den Vektorraum \mathbb{R}^n definiert. Diese Definition kann erweitert werden auf zwei Vektoren eines beliebigen Vektorraums.

Satz 3.72: Sind a, b und c Elemente eines Vektorraums V über einem Körper $K \in \{\mathbb{R}, \mathbb{C}\}$, dann definiert man ein **Skalarprodukt**

$$\langle a, b \rangle : V \times V \rightarrow K$$

durch folgende Eigenschaften:

SP1: Es gilt

$$\langle a, b \rangle = \begin{cases} \langle b, a \rangle & \text{für beliebige } K \text{ außer } \mathbb{C} \\ \overline{\langle b, a \rangle} & \text{für } K = \mathbb{C}. \end{cases}$$

SP2: Wenn a, b, c drei Vektoren sind, dann ist

$$\begin{aligned} \langle a, (b + c) \rangle &= \langle a, b \rangle + \langle a, c \rangle \\ \langle (a + b), c \rangle &= \langle a, c \rangle + \langle b, c \rangle \end{aligned}$$

SP3: Sei $\lambda \in K$, dann gilt:

$$\langle \lambda a, b \rangle = \lambda \langle a, b \rangle = \begin{cases} \langle a, \lambda b \rangle & \text{für beliebige } K \text{ außer } \mathbb{C} \\ \langle a, \overline{\lambda} b \rangle & \text{für } K = \mathbb{C}. \end{cases}$$

SP4: Wenn $a = 0$ der Nullvektor ist, dann ist $\langle a, a \rangle = 0$; sonst ist $\langle a, a \rangle > 0$.

Bemerkung 3.73: $\langle a, a \rangle \in \mathbb{R}$ gilt auch für $K = \mathbb{C}$, so dass das „ $>$ “ in der letzten Bedingung auch wirklich definiert ist.

BEWEIS Nach Bedingung **SP1** muss gelten: $\langle a, a \rangle = \overline{\langle a, a \rangle}$, so dass der Imaginäranteil 0 sein muss. ■

Nicht alle Vektorräume müssen auch ein Skalarprodukt besitzen. Bisher haben wir nur ein Skalarprodukt benutzt, nämlich das Standardskalarprodukt im \mathbb{R}^n . Wir werden jetzt auch

- für andere Vektorräume Skalarprodukte definieren, beispielsweise für $C[a, b]$.
- weitere Skalarprodukte für den \mathbb{R}^n definieren.

Beispiel 3.74: Es sei $C[a, b]$ der lineare Raum der auf dem abgeschlossenen Intervall $[a, b] \subset \mathbb{R}$ stetigen reellen Funktionen. Dann ist

$$\langle f, g \rangle = \int_a^b f(x) \cdot g(x) dx; \quad f, g \in C[a, b]$$

ein Skalarprodukt auf $C[a, b]$.

BEWEIS Für alle $f, g \in C[a, b]$ gilt

SP1:

$$\begin{aligned} \langle f, g \rangle &= \int_a^b f(x) \cdot g(x) dx \\ &= \int_a^b g(x) \cdot f(x) dx \\ &= \langle g, f \rangle \end{aligned}$$

SP2:

$$\begin{aligned} \langle f, g + h \rangle &= \int_a^b f(x) \cdot (g(x) + h(x)) dx \\ &= \int_a^b (f(x) \cdot g(x)) + (f(x) \cdot h(x)) dx \\ &= \int_a^b (f(x) \cdot g(x)) dx + \int_a^b (f(x) \cdot h(x)) dx \\ &= \langle f, g \rangle + \langle f, h \rangle \end{aligned}$$

SP3:

$$\begin{aligned} \langle \lambda f, g \rangle &= \int_a^b \lambda f(x) \cdot g(x) dx \\ &= \lambda \cdot \int_a^b f(x) \cdot g(x) dx \\ &= \lambda \cdot \langle f, g \rangle \end{aligned}$$

SP4:

$$\langle f, f \rangle = \int_a^b f(x)^2 dx$$

Da $f(x)^2 \geq 0 \forall x \in [a, b]$ ist auch $\int_a^b f(x)^2 dx \geq 0$ und es gilt

$$\int_a^b f(x)^2 dx = 0 \quad \Leftrightarrow \quad f(x) = 0 \quad \forall x \in [a, b].$$

■

Beispiel 3.75: Im \mathbb{R}^2 ist durch

$$\langle x, y \rangle = a \cdot x_1 y_1 + b \cdot (x_1 y_2 + x_2 y_1) + c \cdot x_2 y_2$$

ein Skalarprodukt definiert, falls $a > 0$ ist und $ac - b^2 > 0$ ist.

BEWEIS

SP1:

$$\begin{aligned} \langle x, y \rangle &= a \cdot x_1 y_1 + b \cdot (x_1 y_2 + x_2 y_1) + c \cdot x_2 y_2 \\ &= a \cdot y_1 x_1 + b \cdot (y_1 x_2 + y_2 x_1) + c \cdot y_2 x_2 \\ &= \langle y, x \rangle \end{aligned}$$

SP2:

$$\begin{aligned} \langle x, y + z \rangle &= a \cdot x_1 (y_1 + z_1) + b \cdot (x_1 (y_2 + z_2) + x_2 (y_1 + z_1)) + c \cdot x_2 (y_2 + z_2) \\ &= a \cdot x_1 y_1 + b \cdot (x_1 y_2 + x_2 y_1) + c \cdot x_2 y_2 + a \cdot x_1 z_1 + b \cdot (x_1 z_2 + x_2 z_1) + c \cdot x_2 z_2 \\ &= \langle x, y \rangle + \langle x, z \rangle \end{aligned}$$

SP3:

$$\begin{aligned} \langle \lambda x, y \rangle &= a \cdot \lambda x_1 y_1 + b \cdot (\lambda x_1 y_2 + \lambda x_2 y_1) + c \cdot \lambda x_2 y_2 \\ &= \lambda (a \cdot x_1 y_1 + b \cdot (x_1 y_2 + x_2 y_1) + c \cdot x_2 y_2) \\ &= \lambda \langle x, y \rangle \end{aligned}$$

SP4:

$$\begin{aligned} \langle x, x \rangle &= a \cdot x_1^2 + 2b \cdot (x_1 x_2) + c \cdot x_2^2 \\ &= a \left(x_1^2 + \frac{2b}{a} x_1 x_2 \right) + c \cdot x_2^2 \\ &= a \left(x_1^2 + \frac{2b}{a} x_1 x_2 + \overbrace{\frac{b^2}{a^2} x_2^2 - \frac{b^2}{a^2} x_2^2}^{=0} \right) + c \cdot x_2^2 \\ &= a \left(x_1 + \frac{b}{a} x_2 \right)^2 - \frac{b^2}{a} x_2^2 + c \cdot x_2^2 \\ &= \underbrace{\frac{>0}{a}}_{>0} \underbrace{\left(x_1 + \frac{b}{a} x_2 \right)^2}_{\geq 0} + \frac{\overbrace{(ca - b^2)}^{>0}}{a} \underbrace{x_2^2}_{\geq 0} \\ &\geq 0 \end{aligned}$$

$\langle x, x \rangle = 0$ gilt genau dann, wenn $(x_1 + \frac{b}{a} x_2)^2 = 0$ und $x_2^2 = 0$. Aus beiden Bedingungen zusammen folgt $x_1 = x_2 = 0$. Das Standardskalarprodukt im \mathbb{R}^2 ergibt sich aus dem Spezialfall $a = c = 1; b = 0$. Die Skalarprodukte, die durch eine andere Wahl von a, b und c entstehen, haben keiner anschauliche Bedeutung. ■

Im nachfolgenden Beispiel wird der Unterschied der Skalarprodukte noch einmal deutlich.

Beispiel 3.76: Für das Skalarprodukt aus dem vorigen Beispiel sei $a = c = 2$ und $b = 1$. Die beiden Bedingungen $a > 0$ und $ac - b^2 > 0$ sind erfüllt. Es sei

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Damit gilt für das oben vorgestellte Skalarprodukt:

$$\langle x, y \rangle = 2 \cdot 1 \cdot 0 + 1 \cdot (1 \cdot 1 + 0 \cdot 0) + 2 \cdot 0 \cdot 1 = 1$$

Das Standardskalarprodukt ($a = c = 1; b = 0$) ist

$$\langle x, y \rangle = 0$$

Beispiel 3.77: Auf dem linearen Raum \mathbb{C} selbst ist durch

$$\langle z, w \rangle = z \cdot \bar{w} \quad \forall z, w \in \mathbb{C}$$

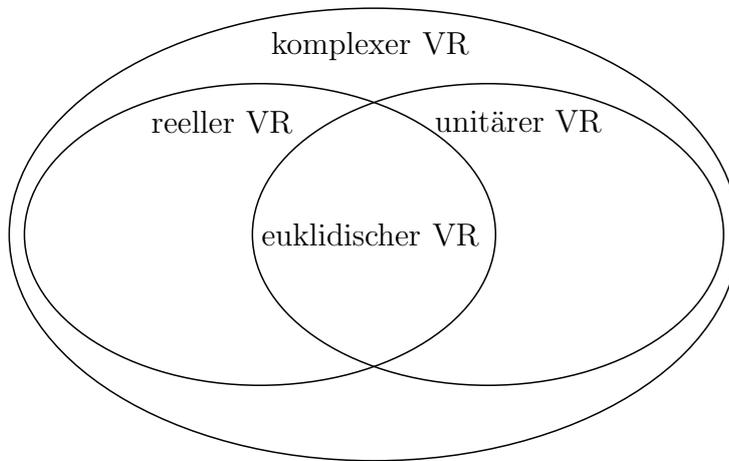
ein komplexes Skalarprodukt definiert. Es gilt dann z.B.

$$\langle z, z \rangle = z \cdot \bar{z} = |z|^2$$

Definition 3.78: Des Weiteren ist folgende Terminologie bzgl. Vektorräumen üblich:

- **Linearer Raum, linearer Vektorraum:** Gleichbedeutend mit Vektorraum.
- **Linearer Unterraum:** Gleichbedeutend mit Untervektorraum.
- **Reeller Vektorraum:** Vektorraum über dem Körper \mathbb{R} .
- **Komplexer Vektorraum:** Vektorraum über dem Körper \mathbb{C} .
- **Euklidischer Vektorraum:** Reeller Vektorraum mit definiertem Skalarprodukt (Bedingungen [SP1-4](#)).
- **Unitärer Vektorraum:** Komplexer Vektorraum mit definiertem Skalarprodukt. Euklidische Räume sind reelle unitäre Räume.

Grafisch kann man die Zusammenhänge als Mengendiagramm wie folgt darstellen:



Definition 3.79: Das Skalarprodukt ist nur ein Spezialfall einer sogenannten Linearform:

- Für eine **Linearform** in der ersten Variablen gilt:

$$\begin{aligned}\langle a + b, c \rangle &= \langle a, c \rangle + \langle b, c \rangle \\ \langle \lambda a, c \rangle &= \lambda \langle a, c \rangle\end{aligned}$$

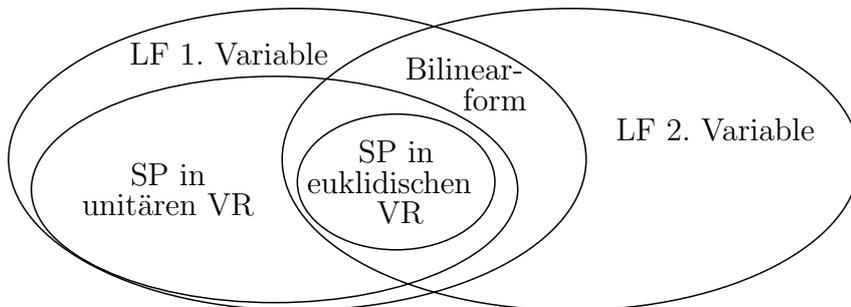
- Für eine **Linearform** in der zweiten Variablen gilt:

$$\begin{aligned}\langle a, b + c \rangle &= \langle a, b \rangle + \langle a, c \rangle \\ \langle a, \lambda c \rangle &= \lambda \langle a, c \rangle\end{aligned}$$

- **Bilinearform:** Linearform in der ersten und in der zweiten Variablen.

Bemerkung 3.80: Skalarprodukte in euklidischen Vektorräumen sind wegen [SP1-3](#) Bilinearformen. Skalarprodukte in unitären Vektorräumen sind Linearformen in der ersten Variablen, müssen aber keine Bilinearformen sein.

Somit ergibt sich die folgende Mengendarstellung:



3.4.2 Die Norm eines Skalarprodukts

Definition 3.81: In einem unitären (bzw. euklidischen) Raum kann man mit Hilfe des Skalarprodukts eine (Standard-)Norm durch

$$\|x\| = \sqrt{\langle x, x \rangle}$$

definieren. Die Bedingungen (N0) bis (N4) aus Kapitel 2.1.5 werden dadurch erfüllt.

Zur Erinnerung: Die Bedingungen einer Norm sind:

$$\text{N0: } \|x\| \in \mathbb{R}.$$

$$\text{N1: } \|x\| \geq 0 \quad \forall x \in V$$

$$\text{N2: } \|x\| = 0 \quad \Leftrightarrow \quad x = 0$$

$$\text{N3: } \|\lambda x\| = |\lambda| \|x\| \quad \forall x \in V, \quad \lambda \in \mathbb{C}$$

$$\text{N4: } \|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V. \text{ (Dreiecksungleichung)}$$

N0 ergibt sich aus **SP1C**: $\langle a, a \rangle = \overline{\langle a, a \rangle} \Rightarrow \text{Im}(\langle a, a \rangle) = 0$. Die Eigenschaften N1-4 wurden bereits in Definition 2.10 für den euklidischen Raum \mathbb{R}^n hergeleitet. Für die Eigenschaften N1-3 haben wir dazu nur die Gleichung $\|x\| = \sqrt{\langle x, x \rangle}$ verwendet und keine der speziellen Eigenschaften des \mathbb{R}^n benötigt. N1-3 brauchen deshalb nicht noch einmal neu bewiesen zu werden. Für den Beweis von N4 haben wir jedoch die Eigenschaft $\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + \dots + x_n y_n$ verwendet und müssen den Satz noch einmal allgemein für unitäre Räume beweisen.

BEWEIS (Beweis der Schwarzschen Ungleichung): Für alle $x, y \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\langle x + \lambda y, x + \lambda y \rangle \geq 0.$$

Es wird jetzt ein spezielles λ eingesetzt, so dass sich aus dieser Gleichung die Schwarzsche Ungleichung ergibt. Da wir in komplexen Vektorräumen sind, wenden wir die Regeln des unitären Skalarprodukts an und erhalten zunächst:

$$\langle x, x \rangle + \bar{\lambda} \langle x, y \rangle + \lambda \langle y, x \rangle + |\lambda|^2 \langle y, y \rangle \geq 0.$$

Für $y = 0$ wird die Schwarzsche Ungleichung zu $\langle x, x \rangle \geq 0$ und ist damit erfüllt. Wir beschränken uns daher auf den Fall $y \neq 0$ und wählen

$$\lambda = -\frac{\langle x, y \rangle}{\langle y, y \rangle}$$

Daraus ergibt sich

$$\bar{\lambda} = -\frac{\langle y, x \rangle}{\langle y, y \rangle}$$

und man erhält

$$\begin{aligned} \langle x, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle} \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, x \rangle + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle^2} \langle y, y \rangle &\geq 0 \\ \langle x, x \rangle + \frac{1}{\langle y, y \rangle} (-2\langle x, y \rangle \langle y, x \rangle + |\langle x, y \rangle|^2) &\geq 0 \\ \langle x, x \rangle + \frac{1}{\langle y, y \rangle} (-2\langle x, y \rangle \overline{\langle x, y \rangle} + |\langle x, y \rangle|^2) &\geq 0 \\ \langle x, x \rangle + \frac{1}{\langle y, y \rangle} (-2|\langle x, y \rangle|^2 + |\langle x, y \rangle|^2) &\geq 0 \\ \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} &\geq 0 \\ \langle x, x \rangle \langle y, y \rangle &\geq |\langle x, y \rangle|^2 \\ \|x\| \|y\| &\geq |\langle x, y \rangle|, \end{aligned}$$

wobei im letzten Schritt auf beiden Seiten die positive Wurzel gezogen wurde. Der Beweis der Dreiecksungleichung aus der Schwarzschen Ungleichung verläuft nun ebenso wie für den Spezialfall aus Satz 2.29. ■

Beispiel 3.82: Es sei $V = C[-1, 1]$ mit dem Skalarprodukt

$$\langle f, g \rangle = \int_{-1}^1 f(x) \cdot g(x) dx.$$

Ferner sei $f(x) = x^2 \in V$. Dann ist

$$\begin{aligned} \|f(x)\|^2 = \|x^2\|^2 &= \int_{-1}^1 x^2 \cdot x^2 dx \\ &= \left[\frac{1}{5} x^5 \right]_{-1}^1 \\ &= \frac{2}{5} \end{aligned}$$

3.4.3 Metriken

Der Abstand zweier Punkte wird mit einer sogenannten Metrik berechnet.

Definition 3.83: Eine **Metrik** ist eine Abbildung $d : V \times V \rightarrow \mathbb{R}$ zweier (Orts-)Vektoren $x \in V$ und $y \in V$ nach \mathbb{R} mit den Eigenschaften:

1. Symmetrie: $d(x, y) = d(y, x)$
2. Dreiecksungleichung: $d(x, y) \leq d(x, z) + d(z, y)$
3. Positive Definitheit: $d(x, y) \begin{cases} = 0 & \text{für } x = y \\ > 0 & \text{sonst} \end{cases}$

Bemerkung 3.84: Durch $d(x, y) = \|x - y\|$ ist die durch eine Norm induzierte **Standardmetrik** definiert.

3.5 Orthogonalität

Definition 3.85: Äquivalent zu der Definition im Vektorraum \mathbb{R}^n aus Kapitel 2.1.5.3 heißen zwei Elemente x, y eines euklidischen oder unitären Raums **orthogonal** zueinander, falls gilt $\langle x, y \rangle = 0$; Schreibweise: $x \perp y$.

Definition 3.86: Die Menge $B = \{x_1, \dots, x_n\}$ heißt **Orthonormalbasis** eines euklidischen bzw. unitären Raums V , falls gilt

- i) B ist eine Basis von V ;
- ii) sämtliche x_k sind paarweise orthogonal;
- iii) für sämtliche x_k gilt $\|x_k\| = 1$.

Ferner gibt es die Begriffe **Orthonormalsystem**, **Orthogonalsystem** und **Orthogonalbasis**, die wie folgt zusammenhängen:

	orthogonal	normiert	Basis
Orthogonalsystem	X	-	-
Orthonormalsystem	X	X	-
Orthogonalbasis	X	-	X
Orthonormalbasis	X	X	X

Für ein Orthogonalsystem B muss noch gelten: $0 \notin B$

Beispiel 3.87: Die Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

sind zwar eine Basis, aber nicht orthogonal, da $\langle \vec{a}_1, \vec{a}_3 \rangle = 1 \neq 0$.

Beispiel 3.88: Die Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

sind eine Basis und orthogonal, aber \vec{a}_3 ist nicht normiert. Also bilden die drei Vektoren eine Orthogonalbasis.

Beispiel 3.89: Die Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

sind orthogonal und normiert, bilden aber keine Basis. Es handelt sich also um ein Orthonormalsystem.

Beispiel 3.90: Die Vektoren

$$\vec{a}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind eine Basis, orthogonal und normiert. Sie bilden also eine Orthonormalbasis.

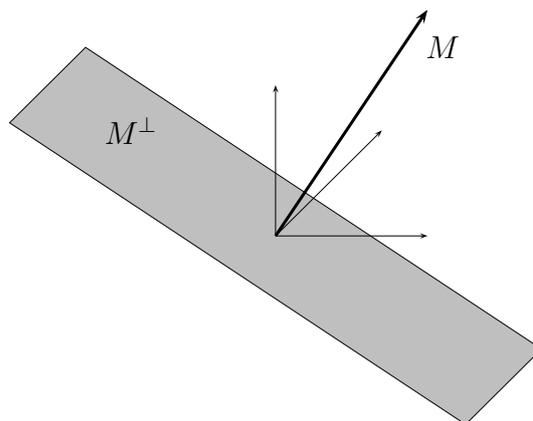
3.5.1 Eigenschaften orthogonaler Vektoren

Definition 3.91: Ist M eine Teilmenge eines euklidischen Vektorraums V , dann heißt

$$M^\perp = \{v \in V; v \perp u \quad \forall u \in M\}$$

das **orthogonale Komplement** von M .

Das orthogonale Komplement M^\perp ist also die Menge aller Vektoren, die senkrecht auf jedem Vektor aus M stehen.



Beispiel 3.92: Es sei $V = \mathbb{R}^3$ und $M = \{(2, 0, 0)\}$, also eine Menge mit nur einem Vektor. Dann ist M^\perp die Menge aller Vektoren \vec{x} , die skalar mit $(2, 0, 0)$ multipliziert das Ergebnis 0 ergeben. Dies ist

$$M^\perp = \left\{ \begin{pmatrix} 0 \\ a \\ b \end{pmatrix}; \quad a, b \in \mathbb{R} \right\}$$

Bemerkung 3.93: M^\perp ist ein Untervektorraum von V .

BEWEIS Es ist die Abgeschlossenheit zu prüfen. Es gilt für $u \in M$; $x, y \in M^\perp$ und $\lambda \in \mathbb{R}$:

$$\langle x + y, u \rangle = \langle x, u \rangle + \langle y, u \rangle = 0$$

$$\langle \lambda x, u \rangle = \lambda \langle x, u \rangle = 0$$

Damit ist M^\perp abgeschlossen und somit ein Untervektorraum von V . ■

Satz 3.94: Ein Orthonormalsystem ist linear unabhängig.

BEWEIS Der Beweis beruht darauf, dass in einem Orthonormalsystem $\langle a_i, a_j \rangle = \delta_{i,j}$ ist. δ ist hierbei das sogenannte Kronecker-Symbol. Es gilt:

$$\delta_{k,l} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases}$$

Der Beweis lautet:

$$\begin{aligned} \lambda_1 a_1 + \dots + \lambda_n a_n &= 0 & | \langle a_i, \dots \rangle \\ \langle a_i, \lambda_1 a_1 + \dots + \lambda_n a_n \rangle &= 0 \\ \lambda_i \langle a_i, a_i \rangle &= 0 \\ \lambda_i &= 0 \end{aligned}$$

Da i alle Werte von 1 bis n annehmen kann, ist $\lambda_1 = \dots = \lambda_n = 0$ und $a_1 \dots a_n$ sind linear unabhängig. ■

Das Skalarprodukt eines Vektors mit einem Einheitsvektor ist der Betrag der Länge der Projektion des Vektors in Richtung des Einheitsvektors. Dieser Satz wurde in Kapitel 2.1.6.1 bereits gezeigt und es wurden dabei keine speziellen Eigenschaften des \mathbb{R}^n gebraucht. Beispiel:

$$\left\langle \begin{pmatrix} 5 \\ 3 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = 5$$

Das führt zu folgendem Satz:

Satz 3.95: Ist $B = \{v_1, \dots, v_n\}$ eine Orthonormalbasis von V , dann gilt für jedes $v \in V$:

$$v = \sum_{k=1}^n \langle v, v_k \rangle v_k,$$

d.h. v hat bzgl. B die Koordinaten $\{\langle v, v_k \rangle; 1 \leq k \leq n\}$.

BEWEIS Da B eine Basis ist, gilt immer die Darstellung

$$v = \sum_{k=1}^n \lambda_k v_k.$$

Bildet man auf beiden Seiten das Skalarprodukt mit v_l , dann erhält man wegen $\langle v_k, v_l \rangle = \delta_{k,l}$:

$$\langle v, v_l \rangle = \lambda_l \langle v_l, v_l \rangle = \lambda_l \quad \forall 1 \leq l \leq n.$$

■

Beispiel 3.96: Die Vektoren

$$\vec{a}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind eine Orthonormalbasis des \mathbb{R}^3 . Der Vektor $\vec{v} = (5, 3, 7)$ lässt sich also als $\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \lambda_3 \vec{a}_3$ schreiben. Die Vorfaktoren λ_1 - λ_3 ergeben sich aus:

$$\begin{aligned} \lambda_1 &= \langle \vec{v}, \vec{a}_1 \rangle = \frac{5}{\sqrt{2}} + \frac{3}{\sqrt{2}} = 4 \cdot \sqrt{2} \\ \lambda_2 &= \langle \vec{v}, \vec{a}_2 \rangle = \frac{5}{\sqrt{2}} - \frac{3}{\sqrt{2}} = \sqrt{2} \\ \lambda_3 &= \langle \vec{v}, \vec{a}_3 \rangle = 7 \end{aligned}$$

Also ist $\vec{v} = 4\sqrt{2} \cdot \vec{a}_1 + \sqrt{2} \cdot \vec{a}_2 + 7 \cdot \vec{a}_3$.

Satz 3.97: Ist $B = \{v_1, \dots, v_r\}$ ein Orthogonalsystem in V und ist $U = L(B)$ der von B aufgespannte Unterraum, dann lässt sich jedes $v \in V$ eindeutig als Summe $v = u + w$ mit $u \in U$ und $w \in U^\perp$ schreiben. Es ist dann

$$u = \sum_{k=1}^r \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k$$

und daher

$$w = v - \sum_{k=1}^r \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k.$$

Die Projektion u von v auf U ist also die Summe der Einzelprojektionen von v auf die Basisvektoren von U .

BEWEIS Zu zeigen ist, dass tatsächlich $w \in U^\perp$ gilt.

$$\begin{aligned} w &= v - \sum_{k=1}^r \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k \\ \langle w, v_i \rangle &= \langle v, v_i \rangle - \sum_{k=1}^r \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} \overbrace{\langle v_k, v_i \rangle}^{\delta_{i,k}} \\ &= \langle v, v_i \rangle - \langle v, v_i \rangle \\ &= 0 \end{aligned}$$

Damit ist $w \in U^\perp$. ■

Beispiel 3.98:

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

ist sogar ein Orthonormalsystem. Seine lineare Hülle ist

$$U = L(B) = \left\{ \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}; a, b \in \mathbb{R} \right\}.$$

Das Orthogonale Komplement $U^\perp = \{v \in V; v \perp u \forall u \in U\}$ ist

$$U^\perp = \left\{ \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}; c \in \mathbb{R} \right\}.$$

Der Satz sagt: Jedes $v \in V$, also jedes $v \in \mathbb{R}^3$ lässt sich zusammensetzen als Summe aus zwei Vektoren $u \in U$ und $w \in U^\perp$, also

$$v = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}; v \in \mathbb{R}^3, a, b, c \in \mathbb{R}.$$

Bemerkung 3.99: Mit Hilfe von Satz 3.97 können wir nun orthogonale Projektionen auf Unterräume, also z.B. auch Lotfußpunkte, berechnen. Diese Eigenschaft werden wir in Kapitel 5.3 auch zur Bestimmung von Näherungslösungen überbestimmter Gleichungssysteme nutzen.

Bemerkung 3.100: Gemäß der Definition 3.32 der Direkten Summe kann man V als direkte Summe von U und U^\perp schreiben, d.h. $V = U + U^\perp$.

3.5.2 Das Orthonormalisierungsverfahren nach Gram-Schmidt

Das Orthonormalisierungsverfahren nach Gram-Schmidt wurde unabhängig voneinander von Erhard Schmidt (1907) und von Jørgen Pederson Gram (1879) veröffentlicht. Beide gelten als „Entdecker“ des Verfahrens, wobei im deutschen Sprachraum manchmal nur vom Orthonormalisierungsverfahren nach Erhard Schmidt die Rede ist. Allerdings wurde das Verfahren schon 1836 von Cauchy benutzt. Eine vereinfachte Variante heißt das Orthogonalisierungsverfahren nach Gram-Schmidt. Hier werden die Vektoren nicht normalisiert.

Was macht das Gram-Schmidtsche Orthonormalisierungsverfahren?

Gegeben seien m linear unabhängige Vektoren aus einem Vektorraum der

Dimension n , wobei $m \leq n$. Das Verfahren erzeugt daraus einen Satz von m orthonormalen Vektoren mit der gleichen linearen Hülle.

Beispiel: Die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}; \quad \vec{v}_2 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}; \quad \vec{v}_3 = \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix}$$

ergeben nach dem Gram-Schmidtschen Orthonormalisierungsverfahren die Vektoren

$$\vec{v}_1' = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}; \quad \vec{v}_2' = \begin{pmatrix} 2/3 \\ -2/3 \\ 1/3 \end{pmatrix}; \quad \vec{v}_3' = \begin{pmatrix} 2/3 \\ 1/3 \\ -2/3 \end{pmatrix}.$$

3.5.2.1 „Sinn“ des Verfahrens

Könnte man im obigen Beispiel nicht gleich die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

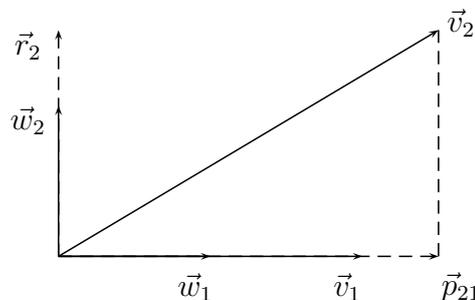
nehmen? Schließlich sind sie auch orthonormal und haben die gleiche lineare Hülle.

Das Verfahren hat aber folgende Vorteile:

- Zunächst einmal kann man den Ausgangsvektoren nicht immer direkt die Lösung ansehen, insbesondere, wenn die v_i keine Basis oder nicht aus dem \mathbb{R}^n sind.
- Es ist ein Beweis dafür, dass n linear unabhängige Vektoren in n orthogonale Vektoren umgewandelt werden können, die die gleiche lineare Hülle haben. Das heißt auch, dass das Verfahren den Beweis erbringt, dass die Ausgangsvektoren linear unabhängig waren.

3.5.2.2 Funktionsweise des Verfahrens

Einfaches Beispiel im \mathbb{R}^2 :



Schritt 1: \vec{v}_1 normieren: $\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$.

Schritt 2a: Projektion von \vec{v}_2 auf \vec{w}_1 bestimmen: $\vec{p}_{21} = \langle \vec{v}_2, \vec{w}_1 \rangle \vec{w}_1$.

Schritt 2b: Diese Projektion von \vec{v}_2 abziehen. Der Rest steht senkrecht auf \vec{v}_1 :

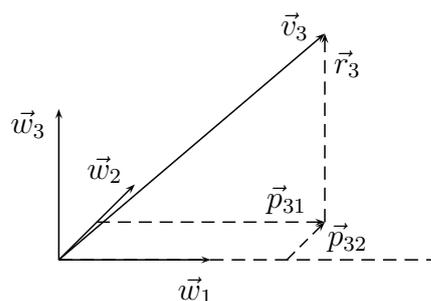
$$\vec{r}_2 = \vec{v}_2 - \langle \vec{v}_2, \vec{w}_1 \rangle \vec{w}_1.$$

Schritt 3: Diesen Vektor normieren. $\vec{w}_2 = \frac{\vec{r}_2}{\|\vec{r}_2\|}$.

$\Rightarrow \vec{w}_1, \vec{w}_2$ bilden ein Orthonormalsystem. Zusammengefasst:

$$\begin{aligned}\vec{w}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} \\ \vec{r}_2 &= \vec{v}_2 - \langle \vec{v}_2, \vec{w}_1 \rangle \vec{w}_1 \\ \vec{w}_2 &= \frac{\vec{r}_2}{\|\vec{r}_2\|}\end{aligned}$$

Drei linear unabhängige Vektoren im \mathbb{R}^3 : Zunächst werden 2 der 3 Vektoren wie im letzten Beispiel orthonormalisiert.



Von \vec{v}_3 werden die beiden Projektionen auf \vec{w}_1 und \vec{w}_2 abgezogen. Der Rest steht senkrecht auf \vec{w}_1 und \vec{w}_2 .

$$\vec{r}_3 = \vec{v}_3 - \overbrace{\langle \vec{v}_3, \vec{w}_1 \rangle \vec{w}_1}^{\text{Proj. auf } \vec{w}_1} - \overbrace{\langle \vec{v}_3, \vec{w}_2 \rangle \vec{w}_2}^{\text{Proj. auf } \vec{w}_2}$$

Normieren des senkrechten Vektors:

$$\vec{w}_3 = \frac{\vec{r}_3}{\|\vec{r}_3\|}$$

Allgemein:

$$\begin{aligned}w_1 &= \frac{v_1}{\|v_1\|} \\ r_{k+1} &= v_{k+1} - \sum_{i=1}^k \langle v_{k+1}, w_i \rangle w_i \\ w_{k+1} &= \frac{r_{k+1}}{\|r_{k+1}\|}\end{aligned}$$

Beispiel 3.101:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix}$$

Berechnung von \vec{w}_1 :

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Berechnung von \vec{w}_2 :

$$\begin{aligned} \vec{r}_2 &= \vec{v}_2 - \langle \vec{v}_2, \vec{w}_1 \rangle \vec{w}_1 \\ &= \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} - \frac{1}{3}(3 + 8 + 10) \cdot \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} - \frac{7}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \\ &= \frac{1}{3} \left[\begin{pmatrix} 9 \\ 12 \\ 15 \end{pmatrix} - \begin{pmatrix} 7 \\ 14 \\ 14 \end{pmatrix} \right] = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \\ \vec{w}_2 &= \frac{\vec{r}_2}{\|\vec{r}_2\|} = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \end{aligned}$$

Berechnung von \vec{w}_3 :

$$\begin{aligned} \vec{r}_3 &= \vec{v}_3 - \langle \vec{v}_3, \vec{w}_1 \rangle \vec{w}_1 - \langle \vec{v}_3, \vec{w}_2 \rangle \vec{w}_2 \\ &= \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{3} \cdot 11 \cdot \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{1}{3} \cdot 13 \cdot \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \\ &= \frac{1}{9} \left[\begin{pmatrix} 63 \\ 9 \\ 9 \end{pmatrix} - \begin{pmatrix} 11 \\ 22 \\ 22 \end{pmatrix} - \begin{pmatrix} 26 \\ -26 \\ 13 \end{pmatrix} \right] \\ &= \frac{1}{9} \begin{pmatrix} 26 \\ 13 \\ -26 \end{pmatrix} = \frac{13}{9} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \\ \vec{w}_3 &= \frac{\vec{r}_3}{\|\vec{r}_3\|} = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \end{aligned}$$

$$\Rightarrow \vec{w}_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \vec{w}_2 = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \quad \vec{w}_3 = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

3.5.2.3 Bemerkungen zur Orthonormalisierung

Bemerkung 3.102: Es gibt mehrere mögliche Ergebnisse. Diese hängen davon ab, in welcher Reihenfolge man die Vektoren orthonormalisiert. Eines ist aber allen Ergebnissen gemeinsam: Sie bilden alle ein Orthonormalsystem mit der gleichen linearen Hülle wie die Ausgangsvektoren.

Bemerkung 3.103: Das Orthonormalisierungsverfahren funktioniert nur für linear unabhängige Vektoren. Im Verlauf des Verfahrens erkennt man, ob die Vektoren linear unabhängig sind oder nicht. Folgende Aussagen sind nämlich äquivalent:

- Die Ausgangsvektoren sind linear abhängig.
- Man erhält im Laufe des Verfahrens einen Nullvektor als Ergebnis

Das heißt aber auch, dass man die lineare Unabhängigkeit der Ausgangsvektoren gezeigt hat, wenn keiner der Ergebnisvektoren der Nullvektor ist.

3.5.2.4 Orthonormalisierung in anderen euklidischen Vektorräumen als \mathbb{R}^n

Beispiel 3.104: Im Vektorraum der Polynome vom Grad ≤ 2 sei ein Skalarprodukt definiert durch

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx$$

Gegeben sind:

$$v_1(x) = 1; \quad v_2(x) = x; \quad v_3(x) = x^2$$

Auch hier kann man die drei Vektoren mit dem bekannten Schema orthonormalisieren:

Bestimmung von $w_1(x)$:

$$\begin{aligned} w_1(x) &= \frac{v_1(x)}{\|v_1(x)\|} = \frac{1}{\|1\|} \\ \|1\| &= (\langle 1, 1 \rangle)^{\frac{1}{2}} \\ &= \left(\int_{-1}^1 1 dx \right)^{\frac{1}{2}} \\ &= ([x]_{-1}^1)^{\frac{1}{2}} = 2^{\frac{1}{2}} = \sqrt{2} \\ \Rightarrow w_1(x) &= \frac{1}{\sqrt{2}} \end{aligned}$$

Bestimmung von $w_2(x)$:

$$\begin{aligned}
 r_2(x) &= v_2(x) - \langle v_2(x), w_1(x) \rangle w_1(x) \\
 &= x - \int_{-1}^1 x \cdot \frac{1}{\sqrt{2}} dx \cdot \frac{1}{\sqrt{2}} \\
 &= x - \frac{1}{2} \left[\frac{1}{2} x^2 \right]_{-1}^1 = x \\
 w_2(x) &= \frac{r_2(x)}{\|r_2(x)\|} \\
 \|x\| &= \left(\int_{-1}^1 x \cdot x dx \right)^{\frac{1}{2}} \\
 &= \left(\left[\frac{1}{3} x^3 \right]_{-1}^1 \right)^{\frac{1}{2}} = \sqrt{\frac{2}{3}} \\
 \Rightarrow w_2(x) &= \sqrt{\frac{3}{2}} \cdot x
 \end{aligned}$$

Bestimmung von $w_3(x)$:

$$\begin{aligned}
 r_3(x) &= v_3(x) - \langle v_3(x), w_1(x) \rangle w_1(x) - \langle v_3(x), w_2(x) \rangle w_2(x) \\
 &= x^2 - \int_{-1}^1 x^2 \frac{1}{\sqrt{2}} dx \cdot \frac{1}{\sqrt{2}} - \int_{-1}^1 x^2 \sqrt{\frac{3}{2}} x dx \cdot \sqrt{\frac{3}{2}} x \\
 &= x^2 - \frac{1}{2} \left[\frac{1}{3} x^3 \right]_{-1}^1 - \frac{3}{2} x \left[\frac{1}{4} x^4 \right]_{-1}^1 \\
 &= x^2 - \frac{2}{6} = x^2 - \frac{1}{3} \\
 w_3(x) &= \frac{r_3(x)}{\|r_3(x)\|}
 \end{aligned}$$

$$\begin{aligned}\|r_3(x)\| &= \left(\int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 dx \right)^{\frac{1}{2}} \\ &= \left(\int_{-1}^1 \left(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\right) dx \right)^{\frac{1}{2}} \\ &= \left(\left[\frac{1}{5}x^5 - \frac{2}{9}x^3 + \frac{1}{9}x \right]_{-1}^1 \right)^{\frac{1}{2}} \\ &= \left(\frac{2}{5} - \frac{4}{9} + \frac{2}{9} \right)^{\frac{1}{2}} \\ &= \left(\frac{2}{5} - \frac{2}{9} \right)^{\frac{1}{2}} = \left(\frac{18}{45} - \frac{10}{45} \right)^{\frac{1}{2}} \\ &= \left(\frac{8}{45} \right)^{\frac{1}{2}} = \left(\frac{16}{9 \cdot 10} \right)^{\frac{1}{2}} = \frac{4}{3 \cdot \sqrt{10}} \\ \Rightarrow w_3(x) &= \frac{3 \cdot \sqrt{10}}{4} \left(x^2 - \frac{1}{3}\right) = \frac{\sqrt{10}}{4} (3x^2 - 1)\end{aligned}$$

$$\Rightarrow w_1(x) = \frac{1}{\sqrt{2}}, \quad w_2(x) = \sqrt{\frac{3}{2}} \cdot x, \quad w_3(x) = \frac{\sqrt{10}}{4} (3x^2 - 1).$$

Kapitel 4

Lineare Abbildungen

4.1 Abbildungen

4.1.1 Einführung

Definition 4.1: A und B sind beliebige Mengen. Jedem $a \in A$ ist ein *eindeutiges* Element $f(a) \in B$ zugeordnet. Die Sammlung von solchen Zuordnungen heißt **Funktion** oder **Abbildung** von A in B . Schreibweise:

$$f : A \rightarrow B \quad \text{oder} \quad f : A \ni a \rightarrow f(a) \in B$$

Die Abbildung ordnet dem Element $a \in A$ das Element $f(a) \in B$ zu. Man nennt $f(a)$

- den Wert von f bezüglich a .
- das Bild von a bezüglich f .

Definition 4.2: Folgende Bezeichnungen werden üblicherweise verwendet:

A : **Definitionsbereich, Definitionsmenge**

B : **Zielmenge, Wertevorrat**

$f(D) := \{b \in B; \exists d \in D : f(d) = b\}$: Menge aller Bilder von $D \subset A$

$f(A) \subset B$: **Bildmenge, Bild** von A bezüglich f , **Wertebereich, Wertemenge**¹

$f^{-1}(\{b\}) := \{a \in A; f(a) = b\} \subset A$: **Urbild** von $b \in B$

$f^{-1}(E) := \{a \in A; f(a) \in E\} \subset A$: **Urbild** der Menge $E \subset B$

¹Wertebereich und Wertemenge werden teilweise auch als Synonyme für Zielmenge verwendet

4.1.2 Injektivität und Surjektivität

Die Begriffe Injektivität, Surjektivität und Bijektivität wurden zwar bereits im Blockkurs Mathematik eingeführt, sollen an dieser Stelle aber noch einmal wiederholt werden.

Definition 4.3: Eine Abbildung $f : V \rightarrow W$ heißt **injektiv** (eindeutig), falls:

$$f(u) = f(v) \Rightarrow u = v$$

bzw.

$$u \neq v \Rightarrow f(u) \neq f(v)$$

Erläuterung in Worten: Unterschiedliche Funktionsargumente liefern immer auch unterschiedliche Ergebnisse.

Definition 4.4: Eine Abbildung $f : V \rightarrow W$ heißt **surjektiv**, falls

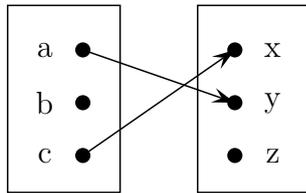
$$f(V) = W$$

bzw. für jedes $y \in W$ gibt es ein $x \in V$ mit $f(x) = y$.

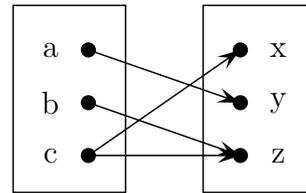
Erläuterung in Worten: Die Surjektivität hängt von der Wahl der Zielmenge ab. Eine Abbildung ist surjektiv, wenn die Zielmenge der Wertemenge entspricht, d.h. jedes Element der Zielmenge auch tatsächlich ein Funktionswert von f ist.

Definition 4.5: Eine Abbildung $f : V \rightarrow W$ heißt genau dann **bijektiv**, falls sie injektiv und surjektiv ist.

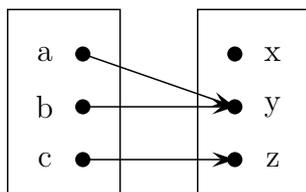
Erläuterung in Worten: Die Bijektivität erlaubt somit für alle Elemente der Zielmenge eine eindeutige Zuordnung eines Elementes des Definitionsbereiches, das dorthin abgebildet wird.



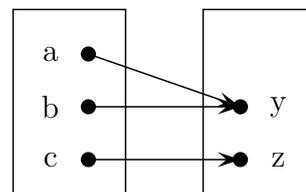
Keine Abbildung, kein Element ist $b \in A$ zugeordnet.



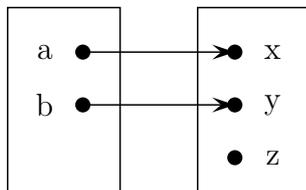
Keine Abbildung, zwei Elemente sind $c \in A$ zugeordnet.



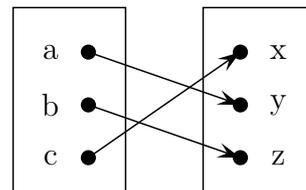
Weder injektiv noch surjektiv.



Surjektiv, aber nicht injektiv.



Injektiv, aber nicht surjektiv.



Injektiv + surjektiv = bijektiv.

Beispiel 4.6: Ist $f(x) = x^2$ surjektiv?

Antwort: Die Frage kann nicht beantwortet werden, da die Angabe der Zielmenge fehlt.

Verbesserte Frage: Ist $f(x) = x^2; f: \mathbb{R} \rightarrow \mathbb{R}$ surjektiv?

Antwort: Nein, denn negative Werte sind keine Funktionswerte. $-1 \neq f(x)$.

Oder auch: Ist $f(x) = x^2; f: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ surjektiv?

Antwort: Ja, denn für alle $y \in \mathbb{R}_{\geq 0}$ existiert ein Urbild $x \in \mathbb{R}$, so dass $f(x) = y$.

Beispiel 4.7: Die Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}^2$, definiert durch $f(t) = t \cdot \vec{a}$, $\vec{a} \in \mathbb{R}^2 \setminus \{(0, 0)\}$, ist injektiv, aber nicht surjektiv.

4.2 Lineare Abbildungen

Die wichtigsten Abbildungen von einem Vektorraum V in einen Vektorraum W sind die linearen Abbildungen.

Definition 4.8: Es seien V und W zwei Vektorräume über einem Körper K . Eine Abbildung $f : V \rightarrow W$ heißt **linear** oder ein **Homomorphismus**, falls gilt:

$$\begin{aligned} f(x+y) &= f(x) + f(y) && \text{(Additivität)} \\ f(\lambda x) &= \lambda f(x) && \text{(Homogenität)} \end{aligned} \quad \forall x, y \in V, \forall \lambda \in K$$

Die Menge der linearen Abbildungen von V nach W wird auch mit $\mathbf{Hom}(V, W)$ bezeichnet.

Beispiel 4.9: Eine Abbildung $f : V = \mathbb{R} \rightarrow W = \mathbb{R}$ ist gegeben durch $f(x) = 5x$. Dann ist die Abbildung f eine lineare Abbildung.

Beweis:

$$\begin{aligned} \text{Additivität: } f(x+y) &= 5(x+y) \\ &= 5x + 5y \\ &= f(x) + f(y) && \forall x, y \in V \\ \text{Homogenität: } f(\lambda x) &= 5(\lambda x) \\ &= \lambda 5x \\ &= \lambda f(x) && \forall x \in V, \forall \lambda \in \mathbb{R} \end{aligned}$$

Beispiel 4.10: Eine Abbildung $f : V = \mathbb{R} \rightarrow W = \mathbb{R}$ ist gegeben durch $f(x) = x^2$. Dann ist die Abbildung f **keine** lineare Abbildung. Sie erfüllt weder die Additivität noch die Homogenität.

Beweis:

$$\begin{aligned} \text{Additivität: } f(x+y) &= (x+y)^2 \\ &= x^2 + 2xy + y^2 \\ &= f(x) + 2xy + f(y) \\ &\neq f(x) + f(y) \\ \text{Homogenität: } f(\lambda x) &= (\lambda x)^2 \\ &= \lambda^2 x^2 \\ &= \lambda^2 f(x) \\ &\neq \lambda f(x) \end{aligned}$$

Beispiel 4.11: Eine Abbildung $f : V = \mathbb{R} \rightarrow W = \mathbb{R}$ ist gegeben durch $f(x) = x + 3$. Dann ist die Abbildung f **keine** lineare Abbildung. Sie erfüllt

weder die Additivität noch die Homogenität.

Beweis:

$$\begin{aligned}
 \text{Additivität:} \quad & f(x+y) = (x+y) + 3 \\
 & f(x) + f(y) = (x+3) + (y+3) = (x+y) + 6 \\
 \Rightarrow \quad & f(x+y) \neq f(x) + f(y) \\
 \text{Homogenität:} \quad & f(\lambda x) = (\lambda x) + 3 \\
 & \lambda f(x) = \lambda(x+3) \\
 \Rightarrow \quad & f(\lambda x) \neq \lambda f(x)
 \end{aligned}$$

Beispiel 4.12: Eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist gegeben durch $f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle$, wobei $\vec{a} \in \mathbb{R}^n$ ein konstanter Vektor ist. Wegen der bekannten Eigenschaften des Skalarprodukts ist f linear. (Ohne Beweis)

Beispiel 4.13: Eine Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ist gegeben durch $f(\vec{x}) = \vec{a} \times \vec{x}$, wobei $\vec{a} \in \mathbb{R}^3$ ein konstanter Vektor ist und $\vec{a} \times \vec{x}$ das bekannte Vektorprodukt zweier Vektoren ist. Wegen der bekannten Eigenschaften des Vektorprodukts ist f linear. (Ohne Beweis)

Beispiel 4.14: In der Praxis begegnet man häufig linearen Abbildungen. So macht man sich z.B. mit Hilfe des Superpositionsprinzips die Eigenschaften physikalischer Größen linearer Probleme, wie z.B. der Überlagerung von Schwingungen, zunutze.

4.2.1 Bild und Kern

Wir hatten bereits das Bild einer Abbildung $f : V \rightarrow W$ definiert als die Menge aller Funktionswerte:

$$f(V) = \{f(v); v \in V\}$$

Aus der Definition der linearen Abbildung folgt direkt, dass $f(V)$ ein Untervektorraum von W ist. Eine weitere wichtige Definition ist der Kern einer linearen Abbildung:

Definition 4.15: Ist V ein Vektorraum und eine f eine (lineare) Abbildung, dann ist der **Kern von f** oder der **Nullraum von f** der Untervektorraum

$$\text{Ker}(f) = \{v \in V; f(v) = 0\}.$$

Der Kern einer linearen Abbildung f kann auch als die Menge der Nullstellen der Abbildung f verstanden werden.

Bemerkung 4.16: Der Kern einer linearen Abbildung hat mindestens das Element 0, d.h. $f(0) = 0$. Dies folgt direkt aus der Additivität oder der Homogenität:

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

wähle $\lambda = 0 \Rightarrow f(0) = 0$

Bemerkung 4.17: Der Kern einer linearen Abbildung $f : V \rightarrow W$ ist ein Untervektorraum von V . Dies folgt mit $x, y \in \text{Ker}(f)$ aus:

$$f(x + y) = f(x) + f(y) = 0 \Rightarrow x + y \in \text{Ker}(f)$$

$$f(\lambda x) = \lambda \cdot f(x) = \lambda \cdot 0 = 0 \Rightarrow \lambda x \in \text{Ker}(f)$$

Der Kern einer linearen Abbildung f enthält also entweder nur ein Element oder unendlich viele Elemente.

Definition 4.18: Die Bezeichnungen *singulär* und *regulär* stehen für:

$$\begin{array}{ll} \text{Ker}(f) = \{0\} & \text{Reguläre Abbildung} \\ \text{Ker}(f) = \{0, x_1, \dots\}; x_1, \dots \neq 0 & \text{Singuläre Abbildung.} \end{array}$$

Die Bezeichnungen sind nicht verwechselt. Es ist genau andersherum, als man intuitiv erwartet.

Bemerkung 4.19: Die lineare Abbildung f ist genau dann injektiv, wenn $\text{Ker}(f) = \{0\}$ ist, das bedeutet: f regulär $\Leftrightarrow f$ injektiv.

BEWEIS Es sei $f(u) = f(v)$, Wegen der Linearität ist das gleichbedeutend mit $f(u - v) = 0$ und $u - v \in \text{Ker}(f)$. Das heißt, es gilt generell:

$$f(u) = f(v) \Leftrightarrow u - v \in \text{Ker}(f).$$

Aus der Regularität $\text{Ker}(f) = \{0\}$ folgt

$$f(u) = f(v) \Leftrightarrow u = v,$$

also die Injektivität von f .

Andersherum folgt aus der Injektivität $f(u) = f(v) \Leftrightarrow u = v$ mit

$$u \in \text{Ker}(f) \Leftrightarrow f(u) = f(0)$$

$$\stackrel{\text{Inj.}}{\Leftrightarrow} u = 0$$

die Regularität. ■

Definition 4.20: Eine lineare bijektive Abbildung heißt auch **Isomorphismus**.

4.2.2 Lineare Abbildungen und Matrizen

4.2.2.1 Aussehen von linearen Abbildungen

Wie sehen lineare Abbildungen aus?

- Sei $V = W = \mathbb{R}$, also $f : \mathbb{R} \rightarrow \mathbb{R}$, dann folgt aus der Homogenität, dass die lineare Abbildung f folgende Gestalt hat:

$$f(x) = f(x \cdot 1) = x \cdot f(1) = x \cdot a, \quad \text{wobei } f(1) = a \text{ ist}$$

$$\Rightarrow f : \mathbb{R} \rightarrow \mathbb{R}, \quad \text{mit } f(x) = a \cdot x, \quad a \in \mathbb{R}.$$

- Seien $V = \mathbb{R}^n$ und $W = \mathbb{R}$, also $f : \mathbb{R}^n \rightarrow \mathbb{R}$, dann hat f folgende Gestalt:

$$\begin{aligned} \vec{x} &= (x_1, \dots, x_n) \\ &= (x_1, 0, \dots, 0) + (0, x_2, 0, \dots, 0) + \dots + (0, \dots, 0, x_n) \end{aligned}$$

Für $f(\vec{x})$ folgt aus der Additivität:

$$f(\vec{x}) = f(x_1, 0, \dots, 0) + f(0, x_2, 0, \dots, 0) + \dots + f(0, \dots, 0, x_n)$$

und aufgrund der Homogenität:

$$\begin{aligned} &= x_1 \cdot \underbrace{f(1, 0, \dots, 0)}_{=: a_1} + x_2 \cdot \underbrace{f(0, 1, 0, \dots, 0)}_{=: a_2} \\ &\quad + \dots + x_n \cdot \underbrace{f(0, \dots, 0, 1)}_{=: a_n} \\ &= x_1 \cdot a_1 + x_2 \cdot a_2 + \dots + x_n \cdot a_n \\ &= \langle \vec{x}, \vec{a} \rangle \end{aligned}$$

$$\Rightarrow f : \mathbb{R} \rightarrow \mathbb{R}, \quad \text{mit } f(\vec{x}) = \langle \vec{x}, \vec{a} \rangle, \quad \vec{a} \in \mathbb{R}^n$$

- Seien $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$, also $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, dann hat f folgende Gestalt:

$$\vec{y} = f(\vec{x}) = f\left(\sum_{i=1}^n x_i \cdot \vec{e}_i\right) = \sum_{i=1}^n x_i \cdot \underbrace{f(\vec{e}_i)}_{\vec{a}_i}$$

$$y_j = \sum_{i=1}^n x_i \cdot a_{ji}$$

$$\begin{aligned} \vec{y} &= (y_1, \dots, y_m) \\ y_i &= \sum_{j=1}^n a_{ij} x_j \quad i = 1, \dots, m \end{aligned}$$

Diese Schreibweise kann man vereinfachen, indem man die aus Kapitel 1.1 bekannten Matrizen verwendet.

Bemerkung 4.21: Eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kann man mit einer $m \times n$ -Matrix schreiben als:

$$\vec{y} = A \cdot \vec{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Vergleicht man die Darstellung mit der Abbildungsgleichung $y_i = \sum_{j=1}^n a_{ij}x_j$, dann erkennt man, dass das i -te Element von \vec{y} sich daraus ergibt, dass die i -te Zeile der Matrix A skalar mit \vec{x} multipliziert wird.

Auch die beiden anderen Fälle kann man auf den letzten Fall zurückführen, indem man

- einen Vektor als eine $1 \times n$ -Matrix sieht,
- einen Skalar als eine 1×1 -Matrix sieht.

Bisher wissen wir zwar schon, wie man eine Matrixmultiplikation durchführt, wissen aber noch nicht, welche Bedeutung eine solche hat. Dabei hilft die folgende Bemerkung:

Bemerkung 4.22: Möchte man k lineare Abbildungen

$$\begin{aligned} f_1 &: \mathbb{R}^{n_0} \rightarrow \mathbb{R}^{n_1} \\ f_2 &: \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2} \\ &\vdots \\ f_k &: \mathbb{R}^{n_{k-1}} \rightarrow \mathbb{R}^{n_k} \end{aligned}$$

mit den dazugehörigen Abbildungsmatrizen

$$\begin{aligned} A_1 &\in \mathbb{R}^{n_1 \times n_0} \\ A_2 &\in \mathbb{R}^{n_2 \times n_1} \\ &\vdots \\ A_k &\in \mathbb{R}^{n_k \times n_{k-1}} \end{aligned}$$

hintereinander ausführen, so ergibt sich die Gesamtabbildung als Multiplikation

$$\begin{aligned} f(\vec{x}) &= f_k(f_{k-1}(\dots f_2(f_1(\vec{x})) \dots)) \\ &= (f_k \circ f_{k-1} \circ \dots \circ f_2 \circ f_1)(\vec{x}) \\ &= A_k \cdot A_{k-1} \cdot \dots \cdot A_1 \cdot \vec{x} \\ &= A \cdot \vec{x} \\ &\in \mathbb{R}^{n_k \times n_0} \end{aligned}$$

der Einzelmatrizen, wobei $A := A_k \cdot \dots \cdot A_1$.

Dabei müssen die Dimensionen der Abbildungen so zueinander passen, dass für alle $2 \leq i \leq k$ der Definitionsbereich der Abbildung f_i die gleiche Dimension hat wie der Wertebereich der Abbildung f_{i-1} .

Bemerkung 4.23: Eine weitere Anwendung der Matrizenmultiplikation ist die gleichzeitige Abbildung mehrerer Vektoren. Diese kann man in Form von Spaltenvektoren zu einer Matrix zusammenfassen und erhält als Ergebnis wieder eine Matrix, dessen Spaltenvektoren die abgebildeten Ausgangsvektoren sind

4.2.2.2 Matrizen als lineare Abbildungen

Im letzten Abschnitt wurde gezeigt, dass eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sich als das Produkt

$$f(x) = A \cdot x$$

darstellen lässt, wobei A eine $m \times n$ -Matrix ist. Der folgende Satz besagt, dass auch die Umkehrung gilt.

Satz 4.24: Jede $m \times n$ -Matrix A (wobei A nicht von x abhängt) liefert durch die Zuordnung

$$f(x) = A \cdot x$$

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

BEWEIS $f(x)$ ordnet jedem x einen eindeutigen Wert zu und ist damit eine Abbildung. Ist $f(x)$ auch linear? Dazu sind die Additivität und die Homogenität zu prüfen.

Additivität: Es sei $\vec{u} = A(\vec{x} + \vec{y})$ und $\vec{v} = A\vec{x} + A\vec{y}$. Dann gilt:

$$\begin{aligned} u_i &= \sum_{j=1}^n a_{ij}(x_j + y_j) \\ &= \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n a_{ij}y_j \\ &= v_i \\ \Rightarrow \vec{u} &= \vec{v} \end{aligned}$$

Homogenität: Es sei $\vec{u} = A(\lambda\vec{x})$ und $\vec{v} = \lambda A\vec{x}$. Dann gilt:

$$\begin{aligned} u_i &= \sum_{j=1}^n a_{ij}(\lambda x_j) \\ &= \lambda \sum_{j=1}^n a_{ij} x_j \\ &= v_i \\ \Rightarrow \vec{u} &= \vec{v} \end{aligned}$$

Damit ist f additiv und homogen und damit linear. ■

Bemerkung 4.25: Die Überprüfung, ob eine bestimmte Abbildung linear ist, kann somit auf verschiedenen Wegen erfolgen.

- Ganz allgemein kann man die beiden Axiome (Additivität und Homogenität) überprüfen.
- Bei Abbildungen vom \mathbb{R}^n in den \mathbb{R}^m kann man überprüfen, ob sich die Abbildung als $f(x) = A \cdot x$ ausdrücken lässt. In der Matrix dürfen keine Ausdrücke vorkommen, die x enthalten. Das bedeutet, dass $f(x)$ ein Polynom sein muss, in dem nur der *lineare* Term vorhanden ist, also weder Konstanten, noch quadratische oder höhere Terme.

Es bleibt allerdings die Frage, wie man überprüft, ob sich eine Abbildung als $f(x) = A \cdot x$ ausdrücken lässt. Kann man z.B. die Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$; $f(\vec{x}) = \vec{z} \times \vec{x}$ ausdrücken als

$$\vec{z} \times \vec{x} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \vec{x}$$

und wie sehen die Koeffizienten der Matrix aus? Diese Frage beantwortet der nächste Abschnitt.

4.2.3 Berechnung linearer Abbildungen

Satz 4.26: Es seien V und W Vektorräume über K und v_1, \dots, v_n eine Basis von V . Dann gibt es zu jedem n -Tupel w_1, \dots, w_n von Vektoren in W genau eine lineare Abbildung $f : V \rightarrow W$ mit $f(v_i) = w_i, i = 1, \dots, n$. Dies bedeutet, dass die Abbildung der Basisvektoren die gesamte Abbildung bestimmt.

Erläuterung: Gegeben ist

- eine Basis $\{v_1, \dots, v_n\}$ aus V

- beliebige Vektoren $\{w_1, \dots, w_n\}$ aus W

Daraus soll eine lineare Abbildung konstruiert werden, so dass:

$$\begin{aligned} f(v_1) &= w_1 \\ f(v_2) &= w_2 \\ &\vdots \\ f(v_n) &= w_n \end{aligned}$$

gilt. Der Satz besagt nun, dass es genau eine lineare Abbildung mit dieser Eigenschaft gibt.

BEWEIS

Beweis der Existenz: Sei $v \in V$. Bezüglich der gewählten Basis $\{v_1, \dots, v_n\}$ lässt sich jedes v eindeutig darstellen als Linearkombination $v = \sum_{i=1}^n \lambda_i v_i$.

Wähle nun $f(v) := \sum_{i=1}^n \lambda_i w_i$, wobei die λ_i jene Koordinaten von v bzgl. der Basis $\{v_1, \dots, v_n\}$ sind. Dann ist die Bedingung $f(v_i) = w_i$ erfüllt. Es fehlen noch die Bedingungen der Additivität und der Homogenität.

Homogenität:

$$\begin{aligned} f(\lambda v) &= f\left(\lambda \left(\sum_{i=1}^n \lambda_i v_i\right)\right) \\ &= f\left(\sum_{i=1}^n (\lambda \cdot \lambda_i) v_i\right) \\ &\stackrel{Def.}{=} \sum_{i=1}^n (\lambda \cdot \lambda_i) w_i \\ &= \lambda \left(\sum_{i=1}^n \lambda_i w_i\right) \\ &\stackrel{Def.}{=} \lambda f(v) \end{aligned}$$

Additivität:

Seien $a = \alpha_1 v_1 + \dots + \alpha_n v_n$ und $b = \beta_1 v_1 + \dots + \beta_n v_n$, dann gilt:

$$\begin{aligned} f(a+b) &= f\left(\left(\sum_{i=1}^n \alpha_i v_i\right) + \left(\sum_{i=1}^n \beta_i v_i\right)\right) \\ &= f\left(\sum_{i=1}^n (\alpha_i + \beta_i) v_i\right) \\ &\stackrel{Def.}{=} \sum_{i=1}^n (\alpha_i + \beta_i) w_i \\ &= \sum_{i=1}^n \alpha_i w_i + \sum_{i=1}^n \beta_i w_i \\ &\stackrel{Def.}{=} f(a) + f(b) \end{aligned}$$

Damit hat f die Eigenschaften einer linearen Abbildung und erfüllt die Bedingung $f(v_i) = w_i$.

Mit dem Beweis der Existenz ist aber noch nicht klargelegt, ob die Abbildung auch eindeutig ist. Es könnte ja auch noch eine zweite Abbildung g existieren, die die Bedingungen $g(v_i) = w_i$ ebenfalls erfüllt. Der folgende Beweis soll die Eindeutigkeit zeigen, d.h. es gibt keine zweite Abbildung, die die Bedingungen $g(v_i) = w_i$ erfüllt.

Beweis der Eindeutigkeit: Sei g eine weitere lineare Abbildung, die die Bedingung $g(v_i) = w_i$ ebenfalls erfüllt. Also gilt $f(v_i) = w_i = g(v_i)$, für $i = 1, \dots, n$. Da $\{v_1, \dots, v_n\}$ eine Basis von V ist, lässt sich jedes Element $v \in V$ eindeutig als Linearkombination $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, mit $\lambda_i \in K$, schreiben.

Wegen der Linearität gilt $\forall v \in V$:

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^n \lambda_i v_i\right) \\ &= \sum_{i=1}^n \lambda_i f(v_i) \\ &= \sum_{i=1}^n \lambda_i g(v_i) \\ &= g\left(\sum_{i=1}^n \lambda_i v_i\right) \\ &= g(v) \end{aligned}$$

Also muss $f = g$ gelten und es gibt nur eine eindeutige lineare Abbildung, die die Bedingungen erfüllt. ■

Man kann also die Abbildungsmatrix einer linearen Abbildung berechnen, wenn man die Abbildung der Vektoren einer Basis der Definitionsmenge kennt.

Beispiel 4.27: Es seien $V = W = \mathbb{R}^2$. Die lineare Abbildung $f : V \rightarrow W$ wird eindeutig beschrieben durch die Abbildung der Basisvektoren $(1, 1)$ und $(-1, 0)$:

$$\begin{aligned} f(1, 1) &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ f(-1, 0) &= \begin{pmatrix} -4 \\ -2 \end{pmatrix} \end{aligned}$$

Die Matrixelemente werden wie folgt berechnet:

$$\begin{aligned} \begin{pmatrix} 2 \\ 3 \end{pmatrix} &= f(1, 1) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} + a_{12} \\ a_{21} + a_{22} \end{pmatrix} \\ \begin{pmatrix} -4 \\ -2 \end{pmatrix} &= f(-1, 0) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -a_{11} \\ -a_{21} \end{pmatrix} \\ &\Rightarrow a_{11} = 4, \quad a_{12} = -2, \quad a_{21} = 2, \quad a_{22} = 1 \end{aligned}$$

Also hat f folgende Gestalt:

$$f(v) = \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \end{pmatrix}.$$

Besonders gut lassen sich die Matrixelemente berechnen, wenn man die Bilder der Einheitsvektoren von V kennt:

Bemerkung 4.28: Es seien $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, B_n und B_m beliebige Basen des Raums \mathbb{R}^n bzw. \mathbb{R}^m sowie $\{\vec{e}_1, \dots, \vec{e}_n\}$ die kanonischen Einheitsvektoren bzgl. B_n . Dann sind $f(\vec{e}_1), \dots, f(\vec{e}_n)$ bzgl. B_m die Spaltenvektoren der zu f gehörigen Abbildungsmatrix A bzgl. der Basen B_n und B_m .

Das erkennt man sofort, wenn man das Matrixprodukt $A \cdot \vec{e}_i$ ausrechnet. Der Übersichtlichkeit halber wird das hier am Beispiel einer Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ und \vec{e}_2 gezeigt:

$$A \cdot \vec{e}_2 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$$

Dies entspricht gerade der zweiten Spalte der Matrix A .

Im Folgenden wählen wir für die Basen B_n und B_m - falls nicht anders beschrieben - immer die kanonischen Basen.

Beispiel 4.29: Eine Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ist gegeben durch $f(\vec{x}) = \vec{a} \times \vec{x}$, wobei $\vec{a} \in \mathbb{R}^3$ ein konstanter Vektor und $\vec{a} \times \vec{x}$ das bekannte Vektorprodukt zweier Vektoren ist. Es gilt:

$$\begin{aligned} f(\vec{e}_1) &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ a_3 \\ -a_2 \end{pmatrix} \\ f(\vec{e}_2) &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -a_3 \\ 0 \\ a_1 \end{pmatrix} \\ f(\vec{e}_3) &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_2 \\ -a_1 \\ 0 \end{pmatrix} \end{aligned}$$

Damit ergibt sich

$$A = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

Test:

$$A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{pmatrix} = f(\vec{x})$$

Damit ist $f(\vec{x})$ eine lineare Abbildung.

Der Test am Ende, ob $f(\vec{x}) = A \cdot \vec{x}$ wirklich erfüllt ist, ist wichtig, weil man sich mit der oben beschriebenen Methode auch von nicht linearen Abbildungen eine Abbildungsmatrix erzeugen kann - nur entspricht diese dann nicht der Abbildung.

Beispiel 4.30: Gegeben ist die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1^2 - x_2^2 \end{pmatrix}$$

Wir versuchen, eine Abbildungsmatrix zu konstruieren:

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

Also ist

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

Die Überprüfung ergibt:

$$g(\vec{x}) = A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix} \neq f(\vec{x})$$

Damit ist $f(\vec{x})$ keine lineare Abbildung. Die eindeutige lineare Abbildung, die sich aus der Abbildung von \vec{e}_1 und \vec{e}_2 ergibt, ist $g(\vec{x})$. Da $f(\vec{x}) \neq g(\vec{x})$, kann $f(\vec{x})$ keine lineare Abbildung sein.

Bemerkung 4.31: Das Bild einer linearen Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist die lineare Hülle der Spaltenvektoren seiner Abbildungsmatrix.

BEWEIS Für einen beliebigen Vektor $\vec{v} \in \mathbb{R}^n$ gilt:

$$f(\vec{v}) = f(\lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n) = \lambda_1 f(\vec{e}_1) + \dots + \lambda_n f(\vec{e}_n).$$

$f(\vec{e}_1) \dots f(\vec{e}_n)$ sind aber gerade die Spaltenvektoren $\vec{a}_1 \dots \vec{a}_n$ der Abbildungsmatrix A . Damit gilt:

$$\begin{aligned} f(\vec{v}) &= \lambda_1 \vec{a}_1 + \dots + \lambda_n \vec{a}_n \\ \text{Bild}(f) &= L(\vec{a}_1, \dots, \vec{a}_n). \end{aligned}$$

■

4.2.4 Die Dimensionsformel für lineare Abbildungen

Definition 3.50 besagte, dass n die Dimension eines Vektorraums V mit einer Basis $\{v_1, \dots, v_n\}$ ist. Bezeichnung: $\dim(V) = n$.

Definition 4.32: Es sei $f : V \rightarrow W$ eine lineare Abbildung. Ist das Bild von f endlichdimensional, dann heißt $\text{rg}(f) = \dim(\text{Bild}(f))$ der **Rang von f** .

Satz 4.33: (Dimensionsformel für lineare Abbildungen): Es sei $f : V \rightarrow W$ eine lineare Abbildung von einem n -dimensionalen linearen Vektorraum V in W . Dann ist

$$\dim(\text{Ker}(f)) + \text{rg}(f) = n.$$

Erläuterung: Die Summe der Dimension des Wertebereichs (Bild) und der Dimension des Kerns ergibt die Dimension des Definitionsbereichs.

BEWEIS V ist n -dimensional, d.h. $\dim(V) = n$, und $\text{Ker}(f)$ ist ein Unterraum von V . Also hat $\text{Ker}(f)$ eine Basis $\{v_1, \dots, v_r\}$, $r \leq n$. Sie wird zu einer Basis $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ von V ergänzt. Setzt man $w_i = f(v_{r+i})$ für $i = 1, \dots, n - r$, dann gilt $\forall v \in V$:

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \dots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n) \\ &= \lambda_1 \underbrace{f(v_1)}_{=0} + \dots + \lambda_r \underbrace{f(v_r)}_{=0} + \lambda_{r+1} f(v_{r+1}) + \dots + \lambda_n f(v_n) \\ &= \lambda_{r+1} f(v_{r+1}) + \dots + \lambda_n f(v_n) \\ &= \lambda_{r+1} w_{r+1} + \dots + \lambda_n w_n \end{aligned}$$

Also gilt $\text{Bild}(f) = L[w_{r+1}, \dots, w_n]$. Außerdem sind $\{w_{r+1}, \dots, w_n\}$ linear unabhängig. Dazu ist zu zeigen:

$$\lambda_{r+1} w_{r+1} + \dots + \lambda_n w_n = 0 \Rightarrow \lambda_{r+1} = \dots = \lambda_n = 0.$$

Aus

$$\lambda_{r+1} w_{r+1} + \dots + \lambda_n w_n = f(\lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n) = 0$$

folgt:

$$\lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n \in \text{Ker}(f).$$

Also muss gelten:

$$\lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n = \lambda_1 v_1 + \dots + \lambda_r v_r$$

für gewisse $\lambda_1, \dots, \lambda_r$.

Da aber $\{v_1, \dots, v_n\}$ linear unabhängig ist, muss auch gelten:

$$\lambda_1 = \dots = \lambda_n = 0.$$

Also ist $\{w_{r+1}, \dots, w_n\}$ linear unabhängig und damit eine Basis von $\text{Bild}(f)$.
Wegen

$$\dim(\text{Ker}(f)) = r,$$

gilt

$$\dim(\text{Bild}(f)) = \text{rg}(f) = n - r.$$

Insgesamt erhalten wir dann:

$$\dim(\text{Ker}(f)) + \dim(\text{Bild}(f)) = \dim(\text{Ker}(f)) + \text{rg}(f) = r + n - r = n.$$

Damit ist die Dimensionsformel bewiesen. ■

Beispielaufgaben folgen im nächsten Abschnitt. Zunächst noch eine weitere Bemerkung:

Bemerkung 4.34: Eine lineare Abbildung $f : V \rightarrow W$ zwischen Räumen der gleichen endlichen Dimension n ist genau dann surjektiv, wenn sie injektiv ist.

BEWEIS Wir wissen:

- Injektiv bedeutet, dass $\text{Ker}(f) = \{0\}$ und damit auch $\dim(\text{Ker}(f)) = 0$ gilt.
- Surjektiv bedeutet $\text{Bild}(f) = W$ und $\text{rg}(f) = n$.

Aufgrund der Dimensionsformel $\dim(\text{Ker}(f)) + \text{rg}(f) = n$ sind beide Aussagen äquivalent. ■

4.2.5 Beispielaufgaben zur Berechnung von Kern und Bild

4.2.5.1 Weg über Berechnung der Abbildungsmatrix

Beispielaufgabe zur Berechnung von Kern und Bild:

Es sei

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y \\ x - y + z \\ 4x - y + 2z \end{pmatrix}$$

- a) Ist f linear? Wenn ja, gebe man die entsprechende Abbildungsmatrix an.
 b) Man bestimme Kern und Bild von f sowie deren Dimensionen.

Teil a) kann man in einem Durchgang erledigen, denn wenn sich eine Abbildungsmatrix angeben lässt, dann muss f auch linear sein. Die Spalten der Matrix ergeben sich aus:

$$f \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}, \quad f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

Damit lautet die Matrix:

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & -1 & 2 \end{pmatrix}$$

Test:

$$f(\vec{x}) = A \cdot \vec{x} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y \\ x - y + z \\ 4x - y + 2z \end{pmatrix}$$

Da sich f als $f(\vec{x}) = A \cdot \vec{x}$ ausdrücken lässt, ist f linear.

Teil b): Man beachte die folgende Bemerkung:

Bemerkung 4.35: Am einfachsten hält man bei der Bestimmung folgende Reihenfolge ein:

1. Bestimmung des Kerns
2. Bestimmung der Dimension des Kerns
3. Bestimmung des Rangs (\Rightarrow Dimensionsformel)
4. Bestimmung des Bildes

Bestimmung des Kerns:

Es ist die Gleichung $f(\vec{x}) = A \cdot \vec{x} = 0$ zu lösen, also das Gleichungssystem:

$$\begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 4 & -1 & 2 & 0 \end{array}$$

Die Lösung mit Gauß ergibt:

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

und schließlich $z = \lambda'$; $y = \frac{2}{3} \cdot \lambda'$; $x = -\frac{1}{3} \cdot \lambda'$. Das heißt mit $\lambda = \frac{1}{3} \cdot \lambda'$:

$$\text{Ker}(f) = \left\{ \vec{x}; \vec{x} = \lambda \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}; \lambda \in \mathbb{R} \right\}$$

und $\dim(\text{Ker}(f)) = 1$.

Bestimmung des Bildes:

Die Dimension des Definitionsbereichs n ist gleich der Anzahl der Spalten von A , also $n = 3$. Nach der Dimensionsformel ist

$$\text{rg}(f) = n - \dim(\text{Ker}(f)) = 2.$$

Wir wissen aus Bemerkung 4.31, dass das Bild die lineare Hülle der Spaltenvektoren von A ist. $\text{rg}(f) = 2$ bedeutet also, dass zwei linear unabhängige

Spalten von A die Basis des Bildes bilden und die dritte Spalte linear abhängig davon ist.

Man wählt zwei Spaltenvektoren aus und testet, ob sie linear unabhängig sind.

$$\vec{s}_1 = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} \quad \text{und} \quad \vec{s}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

sind linear unabhängig ($\vec{s}_1 \neq \lambda \vec{s}_2$). Damit ist das Bild gleich

$$\text{Bild}(f) = \left\{ \vec{x}; \vec{x} = \lambda \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}; \lambda, \mu \in \mathbb{R} \right\}.$$

Aufgabe: Es sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ mit

$$f(\vec{x}) = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \cdot \vec{x}.$$

Man bestimme Kern und Bild von f sowie deren Dimensionen.

4.2.5.2 Andere Rechenwege

Beispielaufgabe zur Berechnung von Kern und Bild:

Es sei P^n der Vektorraum der Polynome vom Grad $\leq n$. Zeigen Sie, dass die Abbildung

$$f : P^n \rightarrow P^{n+1}; f(p(x)) = p(x) \cdot x$$

eine lineare Abbildung ist. Bestimmen Sie das Bild und den Kern dieser Abbildung.

Ist die Abbildung linear? Da es sich nicht um eine Abbildung $\mathbb{R}^m \rightarrow \mathbb{R}^n$ handelt, kann man nicht so einfach eine Abbildungsmatrix angeben. Daher überprüft man Additivität und Homogenität.

Additivität:

$$\begin{aligned} f(p_1(x) + p_2(x)) &= (p_1(x) + p_2(x)) \cdot x \\ &= p_1(x) \cdot x + p_2(x) \cdot x \\ &= f(p_1(x)) + f(p_2(x)) \end{aligned}$$

Homogenität:

$$\begin{aligned} f(\lambda p_1(x)) &= (\lambda p_1(x)) \cdot x \\ &= \lambda(p_1(x) \cdot x) \\ &= \lambda f(p_1(x)) \end{aligned}$$

Damit ist f eine lineare Abbildung.

Bestimmung des Kerns: Es ist folgende Gleichung zu lösen:

$$p(x) \cdot x = 0$$

Diese Gleichung ist nur erfüllt für $p(x) = 0$. Die Abbildung ist also nach Bemerkung 4.19 injektiv.

Bestimmung des Bildes: Da der Kern die Dimension 0 hat, sind die Dimension von Definitionsbereich und Wertebereich gleich. Da die Dimension der Zielmenge um 1 größer ist als die Dimension des Definitionsbereichs, kann die Abbildung nicht surjektiv sein. Es gilt:

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \\ f(p(x)) &= a_n x^{n+1} + a_{n-1} x^n + \cdots + a_1 x^2 + a_0 x \end{aligned}$$

Das Bild ist also der Unterraum aller Polynome vom Grad $n+1$, für die das konstante Glied gleich Null ist.

4.2.6 Isomorphismus und Umkehrabbildung

Bemerkung 4.36: Es seien V und W Vektorräume über K und $\{v_1, \dots, v_n\}$ eine Basis von V . Eine lineare Abbildung $f : V \rightarrow W$ ist genau dann ein Isomorphismus, wenn $(f(v_1), \dots, f(v_n))$ eine Basis von W ist, d.h. wenn eine Basis von V in eine Basis von W abgebildet wird.

BEWEIS

\Rightarrow :

Sei f ein Isomorphismus. Als erstes wird gezeigt, dass die Vektoren $f(v_1), \dots, f(v_n)$ linear unabhängig sind. Als zweites wird gezeigt, dass die lineare Hülle von $f(v_1), \dots, f(v_n)$ der Vektorraum W ist. Daraus folgt, dass $f(v_1), \dots, f(v_n)$ eine Basis von W bilden.

1. Lineare Unabhängigkeit: Es sei

$$\lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = 0. \quad (4.1)$$

Wegen der Linearität der Abbildung f gilt:

$$f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0.$$

Da die Abbildung f ein Isomorphismus ist, ist f auch injektiv. Also gilt:

$$f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0 \Leftrightarrow \lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

Die Vektoren v_1, \dots, v_n bilden im Vektorraum V eine Basis und sind daher linear unabhängig. Folglich sind die Faktoren $\lambda_i = 0, \forall i = 1, \dots, n$. Mit (4.1) folgt, dass die Vektoren $f(v_1), \dots, f(v_n)$ ebenfalls alle linear unabhängig sind.

2. Lineare Hülle: Da f surjektiv ist, gibt es zu jedem $w \in W$ ein $v \in V$ mit:

$$w = f(v).$$

Da die Vektoren v_1, \dots, v_n eine Basis im Vektorraum V sind, gibt es Faktoren $\lambda_i \in \mathbb{R}, \forall i = 1, \dots, n$, mit:

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Wegen der Linearität der Abbildung f folgt:

$$w = f(v) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n).$$

Also ist jedes $w \in W$ als Linearkombination von $f(v_1), \dots, f(v_n)$ darstellbar. Damit ist die Lineare Hülle von $\{f(v_1), \dots, f(v_n)\}$ gleich dem Vektorraum W .

3. Damit sind die Vektoren $f(v_1), \dots, f(v_n)$ eine Basis von W .

\Leftarrow :

Nun wird umgekehrt vorausgesetzt, dass die Menge von Vektoren $\{f(v_1), \dots, f(v_n)\}$ eine Basis von W ist. Zu zeigen ist, dass f ein Isomorphismus, d.h. injektiv und surjektiv ist.

Injektivität: Sei $f(v) = 0$. Da $\{v_1, \dots, v_n\}$ eine Basis von V ist, läßt sich v als Linearkombination darstellen:

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Wegen der Linearität von f gilt:

$$0 = f(v) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n).$$

Da die Vektoren $f(v_1), \dots, f(v_n)$ linear unabhängig sind, folgt:

$$\lambda_1 = \cdots = \lambda_n = 0,$$

d.h. $v = 0$, also ist f injektiv.

Surjektivität:

Es sei $w \in W$. Da $\{f(v_1), \dots, f(v_n)\}$ eine Basis von W ist, existieren $\lambda_i \in \mathbb{R}, i = 1, \dots, n$, mit:

$$w = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n).$$

Ist nun $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, dann gilt wegen der Linearität von f :

$$f(v) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = w,$$

d.h. zu jedem $w \in W$ existiert ein $v \in V$ mit $w = f(v)$. Also ist f surjektiv und damit insgesamt bijektiv, folglich ist f ein Isomorphismus. ■

Bemerkung 4.37: Folgende Aussagen sind äquivalent:

1. Eine lineare Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ist ein Isomorphismus.
2. Die zugehörige Abbildungsmatrix A ist quadratisch, d.h. $m = n$, und die Spaltenvektoren von A bilden eine Basis des \mathbb{R}^n .

3. $\det(A) \neq 0$.

BEWEIS (1) \Leftrightarrow (2): Die Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_m$ bilden eine Basis des \mathbb{R}^m . Ihre Abbildungen sind die Spalten $\vec{a}_1, \dots, \vec{a}_m$ der Abbildungsmatrix A . Nach der vorigen Bemerkung ist f genau dann ein Isomorphismus, falls $\vec{a}_1, \dots, \vec{a}_m$ eine Basis des \mathbb{R}^n bilden. Dies ist aber nur möglich, wenn $m = n$ ist.

(2) \Leftrightarrow (3):

Es gilt:

Die Spaltenvektoren von A bilden eine Basis des $\mathbb{R}^n \Leftrightarrow$ Sie sind linear unabhängig $\Leftrightarrow \det(A) \neq 0$ ■

Definition 4.38: Zwei Vektorräume V und W heißen **isomorph**, falls ein Isomorphismus $f : V \rightarrow W$ existiert.

Bemerkung 4.39: Je zwei n -dimensionale Vektorräume über K sind isomorph.

BEWEIS Es seien $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_n\}$ Basen von V bzw. W . Nach Satz 4.26 existiert (genau) eine Abbildung $f : V \rightarrow W$ mit $f(v_1) = w_1, \dots, f(v_n) = w_n$. Da $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_n\}$ Basen sind, ist f nach Bemerkung 4.36 ein Isomorphismus. ■

Definition 4.40: Die **Umkehrabbildung** $f^{-1} : W \rightarrow V$ zu einer Abbildung $f : V \rightarrow W$ wird durch die Eigenschaft $f^{-1}(f(v)) = v$ bzw. $f(f^{-1}(w)) = w$ definiert, wobei $v \in V$ und $w \in W$.

Bemerkung 4.41: Sei $f : V \rightarrow W$ eine lineare Abbildung. Genau dann, wenn $f : V \rightarrow W$ ein Isomorphismus ist, existiert die Umkehrabbildung $f^{-1} : W \rightarrow V$ und ist ebenfalls ein Isomorphismus.

BEWEIS

\Rightarrow :

Ist f als ein Isomorphismus gegeben, dann existiert zu jedem y (surjektiv) ein eindeutiges x (injektiv), so dass $f(x) = y$ gilt. Wegen der eindeutigen Zuordnung gibt es eine Abbildung $g : W \rightarrow V$, mit $g(y) = x$ und g ist ebenfalls eine bijektive Abbildung. Es bleibt nur noch zu zeigen, dass g linear ist.

- Additivität:

Es seien $f(x_1) = y_1$ und $f(x_2) = y_2$.

$\Rightarrow g(y_1) = x_1$ und $g(y_2) = x_2$

$$\begin{aligned} \Rightarrow x_1 + x_2 &= g(f(x_1)) + g(f(x_2)) \\ &= g(y_1) + g(y_2) \end{aligned}$$

Es gilt ebenfalls:

$$\begin{aligned} x_1 + x_2 &= g(f(x_1 + x_2)) \\ &= g(f(x_1) + f(x_2)) \\ &= g(y_1 + y_2) \end{aligned}$$

$$\Rightarrow g(y_1 + y_2) = g(y_1) + g(y_2)$$

- Homogenität:
Es sei $f(x) = y$.
 $\Rightarrow g(y) = x$

$$\begin{aligned}\Rightarrow \lambda x &= \lambda g(f(x)) \\ &= \lambda g(y)\end{aligned}$$

Es gilt ebenfalls:

$$\begin{aligned}\lambda x &= g(f(\lambda x)) \\ &= g(\lambda f(x)) \\ &= g(\lambda y)\end{aligned}$$

$$\Rightarrow g(\lambda y) = \lambda g(y)$$

g ist eine lineare Abbildung und folglich auch ein Isomorphismus. Mit $f^{-1} := g$ ist der erste Punkt der Bemerkung gezeigt.

\Leftarrow :

Ist eine Umkehrabbildung $f^{-1} : W \rightarrow V$ gegeben, dann gilt $f^{-1}(f(x)) = x$. Es ist zu zeigen, dass f ein Isomorphismus ist. Für die Injektivität ist zu zeigen: $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.

$$\begin{aligned}f(x_1) &= f(x_2) \\ \Rightarrow f^{-1}(f(x_1)) &= f^{-1}(f(x_2)) \\ \Rightarrow x_1 &= x_2\end{aligned}$$

$\Rightarrow f$ ist injektiv.

Surjektivität: Wir nehmen an, f sei nicht surjektiv. Daraus folgt, dass ein Element y aus W existiert, dem kein x aus V mit $f(x) = y$ zugeordnet werden kann. Demnach ist $f^{-1}(y)$ nicht definiert und es existiert keine Umkehrabbildung. Dies ist aber ein Widerspruch zur Ausgangsbedingung. Folglich muss f auch surjektiv sein und f ist damit ein Isomorphismus. Nach dem ersten Punkt der Bemerkung ist f^{-1} ebenfalls ein Isomorphismus. ■

Beispiel 4.42: Eine lineare Abbildung $f : V = \mathbb{R}^2 \rightarrow W = \mathbb{R}^2$ ist durch $(x, y) \mapsto (x - y, x + y)$ gegeben. Ist f ein Isomorphismus? Diese Frage ist jetzt leicht zu beantworten. Es ist zu überprüfen, ob eine Umkehrabbildung existiert. Zum Beweis der Aussage ist die eindeutige Lösbarkeit des Gleichungssystem

$$\begin{aligned}a &= x - y \\ b &= x + y\end{aligned}$$

bzw.

$$\begin{array}{c|c} 1 & -1 & a \\ 1 & 1 & b \end{array}$$

zu prüfen. Sie ergibt sich durch die Bedingung $\det(A) \neq 0$:

$$\det A = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2 \neq 0.$$

Aus $\det(A) \neq 0$ folgt, dass das Gleichungssystem eindeutig lösbar ist. Damit existiert eine Umkehrabbildung. Nach der vorherigen Bemerkung ist f ein Isomorphismus und somit injektiv und surjektiv. Der Kern von f ist $\{0\}$. Das Bild ist der \mathbb{R}^2 .

Die Umkehrabbildung selbst erhält man durch die Lösung des Gleichungssystems nach dem Gauß-Algorithmus:

$$\begin{aligned}x &= \frac{a+b}{2} \\y &= \frac{b-a}{2}.\end{aligned}$$

$$\Rightarrow f^{-1}(a, b) = \frac{1}{2}(a+b, b-a)$$

4.3 Matrizen

Wir haben in Bemerkung 3.25 gesehen, dass auch die Matrizen einen Vektorraum bilden.

4.3.1 Allgemeines

Zusätzlich zu den Definitionen und Eigenschaften bezüglich Matrizen aus Kapitel 1 brauchen wir nun noch die folgenden:

Definition 4.43: Eine Matrix mit genau einer Spalte wird **Spaltenvektor** genannt.

Eine Matrix mit genau einer Zeile wird **Zeilenvektor** genannt.

Es gibt auch 1×1 -Matrizen, z.B. (5). Hier lässt man meist die Klammern weg und behandelt sie wie Skalare.

Bemerkung 4.44: Die „alten“ Vektoren werden grundsätzlich als Spaltenvektoren betrachtet.

$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Will man einen Vektor aus Platzgründen in eine Zeile schreiben, muss korrekterweise durch ein hochstehendes T angezeigt werden, dass der transponierte Spaltenvektor gemeint ist:

$$\vec{a} = (1, 2, 3)^T$$

4.3.2 Die Inverse einer Matrix

Definition 4.45: Sei A eine quadratische Matrix. Gibt es eine Matrix A^{-1} mit

$$AA^{-1} = A^{-1}A = E,$$

so heißt A invertierbar. A^{-1} wird als **Inverse** von A bezeichnet.

Bemerkung 4.46: Eine Matrix A ist genau invertierbar, wenn die zugehörige lineare Abbildung ein Isomorphismus ist, bzw. wenn $\det(A) \neq 0$.

Wir hatten in Bemerkung 4.41 gesehen, dass eine lineare Abbildung f genau dann invertierbar ist, wenn sie ein Isomorphismus ist. Falls gilt: $f(x) = Ax$, dann ergibt sich die inverse Matrix A^{-1} aus der Umkehrabbildung $f^{-1}(x) = A^{-1}x$ und

$$\begin{aligned} f^{-1}(f(x)) &= A^{-1}Ax = Ex = x \\ f(f^{-1}(x)) &= AA^{-1}x = Ex = x \end{aligned}$$

In Bemerkung 4.37 wurde auch gezeigt, dass eine Abbildung genau dann ein Isomorphismus ist, falls die Determinante der Abbildungsmatrix ungleich 0 ist.

Bemerkung 4.47: Weiterhin gilt:

1. Jede invertierbare Matrix ist quadratisch.
2. Sind $A, B \in M(n \times n, K)$, dann gilt

$$AB = E \Leftrightarrow BA = E \Leftrightarrow B = A^{-1}$$

3. Sind $A, B \in M(n \times n, K)$ invertierbar, dann auch AB und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.
4. Ist $A \in M(n \times n, K)$ invertierbar, dann auch A^{-1} , und weiterhin gilt $(A^{-1})^{-1} = A$.
5. $(A^T)^{-1} = (A^{-1})^T$
6. $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$

BEWEIS 1. wurde bereits im vorigen Kapitel bewiesen.

2. Dass aus $B = A^{-1}$ die anderen Eigenschaften folgen, folgt aus der Definition von A^{-1} .

Es sei jetzt $AB = E$. Dann ist A surjektiv, denn jedes $y \in K^n$ lässt sich als $A(z)$ ausdrücken, wenn man $z = By$ wählt:

$$A(By) = (AB)y = Ey = y \quad \forall y \in K^n.$$

Eine surjektive Abbildung zwischen Räumen gleicher Dimension ist auch injektiv, also bijektiv. Also existiert A^{-1} . Ist $A^{-1} = B$? Es ist

$$BA = (A^{-1}A)BA = A^{-1}(AB)A = A^{-1}EA = A^{-1}A = E.$$

Damit gilt: $AB = E \Leftrightarrow BA = E$. Nach der Definition von A^{-1} folgt daraus $A^{-1} = B$.

3. Folgt aus:

$$\begin{aligned} B^{-1}A^{-1} &= B^{-1}A^{-1}(AB)(AB)^{-1} \\ &= B^{-1}A^{-1}A(B(AB)^{-1}) \\ &= B^{-1}E(B(AB)^{-1}) \\ &= (B^{-1}B)(AB)^{-1} \\ &= E(AB)^{-1} \\ &= (AB)^{-1} \end{aligned}$$

4. Es gilt $EE = E$, also ist E invers zu sich selbst. Deswegen gilt:

$$E = E^{-1} = (AA^{-1})^{-1} = (A^{-1})^{-1}A^{-1}$$

Damit ist A^{-1} invertierbar und wegen:

$$\begin{aligned} (A^{-1})^{-1} &= (A^{-1})^{-1}(A^{-1}A) \\ &= ((A^{-1})^{-1}A^{-1})A \\ &= EA \\ &= A \end{aligned}$$

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

- 5.

$$\begin{aligned} (A^{-1})^T &= (A^{-1})^T(A^T(A^T)^{-1}) \\ &= ((A^{-1})^T A^T)(A^T)^{-1} \\ &= ((AA^{-1})^T(A^T)^{-1}) \\ &= (A^T)^{-1} \end{aligned}$$

- 6.

$$\begin{aligned} (\lambda A)^{-1} &= (\lambda A)^{-1}(AA^{-1}) \\ &= (\lambda A)^{-1}(\lambda A \frac{1}{\lambda} A^{-1}) \\ &= ((\lambda A)^{-1} \lambda A) \frac{1}{\lambda} A^{-1} \\ &= \frac{1}{\lambda} A^{-1} \end{aligned}$$

■

4.3.3 Wie invertiert man eine Matrix?

4.3.3.1 Erläuterung anhand eines Beispiels

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

Die Matrix A soll invertiert werden. Zunächst wird neben die Matrix eine Einheitsmatrix E vom gleichen Typ geschrieben:

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

Anschließend wird die Matrix A schrittweise durch Zeilenoperationen in die Einheitsmatrix überführt. Die gleichen Operationen werden auch rechts auf der Einheitsmatrix E durchgeführt. Die erlaubten Zeilenoperationen entsprechen den Operationen des Gauß-Algorithmus’.

Z1. Addition eines Vielfachen einer Zeile zu einer anderen.

Z2. Vertauschen zweier Zeilen.

Z3. Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0, \lambda \in K$.

Zunächst wird A in Dreiecksform gebracht:

$$\begin{array}{cccc|cccc} 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 2 & 1 & 0 & 1 & 0 & 0 & | - \textcircled{1} \\ 0 & -1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 & 0 & 1 & | - \textcircled{1} \\ \hline 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 1 & 0 & | + \textcircled{2} \\ 0 & 0 & -1 & 1 & -1 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & 0 & 0 & 1 & | + \textcircled{3} \\ \hline 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -2 & 1 & 1 & 1 \end{array}$$

Die Diagonalelemente werden durch Multiplikation auf den Wert 1 gebracht.

$$\begin{array}{r|cccc}
 1 & 0 & 1 & 1 & \\
 0 & 1 & 1 & 0 & \\
 0 & 0 & 1 & 1 & \\
 0 & 0 & 0 & 2 & \\
 \hline
 1 & 0 & 1 & 1 & \\
 0 & 1 & 1 & 0 & \\
 0 & 0 & 1 & 1 & \\
 0 & 0 & 0 & 1 & \\
 \hline
 & & & & | \cdot 1/2
 \end{array}$$

Durch die entsprechenden Zeilenoperationen werden oberhalb der Diagonalen Nullen erzeugt:

$$\begin{array}{r|cccc|c}
 1 & 0 & 1 & 1 & & \\
 0 & 1 & 1 & 0 & & \\
 0 & 0 & 1 & 1 & & \\
 0 & 0 & 0 & 1 & & \\
 \hline
 1 & 0 & 1 & 0 & & \\
 0 & 1 & 1 & 0 & & \\
 0 & 0 & 1 & 0 & & \\
 0 & 0 & 0 & 1 & & \\
 \hline
 1 & 0 & 0 & 0 & & \\
 0 & 1 & 0 & 0 & & \\
 0 & 0 & 1 & 0 & & \\
 0 & 0 & 0 & 1 & & \\
 \hline
 & & & & & | - \textcircled{4} \\
 & & & & & | - \textcircled{4} \\
 & & & & & | - \textcircled{3} \\
 & & & & & | - \textcircled{3}
 \end{array}$$

Die Matrix auf der rechten Seite ist die Inverse A^{-1} .

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

4.3.3.2 Allgemeine Erläuterung: Elementarmatrizen

Definition 4.48: Eine $n \times n$ -Matrix heißt **Elementarmatrix**, wenn sie durch eine einzige elementare Zeilenoperation aus der Einheitsmatrix E_n hervorgeht.

Alternative Formulierung: Eine Elementarmatrix ist eine quadratische Matrix, welche sich entweder durch die Änderung eines einzigen Eintrags in einen Wert $\neq 0$ oder durch Vertauschen zweier Zeilen von einer $n \times n$ -Einheitsmatrix unterscheidet.

Beispiel 4.49: Elementarmatrizen zu folgenden Operationen:

- Operation „Zeile I \cdot 2“: $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.
- Operation: „1. und 2. Zeile vertauschen“: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
- Operation „Zeile II + 2 \cdot Zeile I“: $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$

Die Multiplikation einer Matrix A von links mit einer der Elementarmatrizen entspricht der dazugehörigen elementaren Zeilenoperation auf A .

Beispiel 4.50: Beispiele für die bekannten Zeilenoperationen:

Z1: „Zeile $i + \alpha \cdot$ Zeile j “. Beispiel:

$$M \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ \alpha a + d & \alpha b + e & \alpha c + f \\ g & h & i \end{pmatrix}$$

Z2: Die Zeilen i und j werden vertauscht:

$$V \cdot A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}$$

Z3: Multiplikation der Zeile i mit dem Wert α . Beispiel:

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ \alpha d & \alpha e & \alpha f \\ g & h & i \end{pmatrix}$$

Eine Zeilenoperation entspricht einer Multiplikation von links mit der entsprechenden Elementarmatrix. Mehrere aufeinanderfolgende Zeilenoperationen lassen sich durch Multiplikation von A mit mehreren Elementarmatrizen ausdrücken:

$$B_n \cdot \dots \cdot B_2 \cdot B_1 \cdot A = \overbrace{(B_n \cdot \dots \cdot B_2 \cdot B_1)}^B \cdot A = B \cdot A$$

Daraus folgt: Erzeugt man durch eine Folge von Zeilenoperationen aus einer Matrix A die Matrix $B \cdot A = E$, dann erzeugt die gleiche Folge von Zeilenoperationen aus einer Einheitsmatrix E die Matrix $B \cdot E = B = A^{-1}$.

Beispiel 4.51: Invertierung der Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$

$$\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 & | & -3 \cdot \textcircled{1} \\ \hline 1 & 2 & 1 & 0 \\ 0 & -1 & -3 & 1 & | & \cdot(-1) \\ \hline 1 & 2 & 1 & 0 & | & -2 \cdot \textcircled{2} \\ 0 & 1 & 3 & -1 \\ \hline 1 & 0 & -5 & -2 \\ 0 & 1 & 3 & -1 \end{array}$$

Auf der linken Seite steht:

$$E = M_2 \cdot F_1 \cdot M_1 \cdot A \quad (\text{Bezeichnung wie in Beispiel 4.50})$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

Auf der rechten Seite steht:

$$A^{-1} = M_2 \cdot F_1 \cdot M_1 \cdot E$$

$$\begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

4.3.4 Matrizen als Vektorräume

Bemerkung 4.52: Die Menge der linearen Abbildungen $\text{Hom}(V, W)$ ist wieder ein Vektorraum.

Erläuterung: Die Vektoren dieses Vektorraums sind lineare Abbildungen, d.h. die Matrizen $A \in M(m \times n, K)$ aus Bemerkung 3.25. Die Operationen \oplus und \odot entsprechen der Addition zweier Matrizen bzw. der Multiplikation einer Matrix mit einem Skalar.

Bemerkung 4.53: Hat man zwei lineare Abbildungen $g : V \rightarrow W$ und $f : W \rightarrow Y$, so ist die zusammengesetzte Abbildung $f \circ g = f(g(x))$ wieder linear.

BEWEIS Es sei $f(y) = Ay$ und $g(x) = Bx$, dann lässt sich $f(g(x))$ darstellen als ABx , also der linearen Abbildung, die durch das Matrixprodukt AB gegeben ist. ■

4.3.5 Der Rang einer Matrix

Im Definition 4.32 haben wir den Rang einer linearen Abbildung definiert. Im Folgenden werden zusätzlich die Begriffe Zeilenrang und Spaltenrang eingeführt.

Definition 4.54: Ist $A \in M(m \times n, K)$, so nennt man $\text{rg}(A) = \dim(\text{Bild}(A))$ den **Rang** der Matrix A . Die maximale Anzahl linear unabhängiger Spaltenvektoren von A nennt man den **Spaltenrang** von A , die maximale Anzahl linear unabhängiger Zeilenvektoren von A den **Zeilenrang** von A .

Es wird sich herausstellen, dass Rang, Zeilenrang und Spaltenrang immer gleich sind. Zunächst wird die Gleichheit von Rang und Spaltenrang gezeigt. Anschließend werden wir zeigen, dass Zeilenrang und Spaltenrang einer Matrix immer gleich sind.

Satz 4.55: Der Rang einer Matrix A ist gleich ihrem Spaltenrang.

BEWEIS Sei $\vec{x} = x_1 \vec{e}_1 + \dots + x_n \vec{e}_n$. Wir wissen, dass für eine lineare Abbildung $f : X \rightarrow Y$ gilt:

$$f(\vec{x}) = x_1 f(\vec{e}_1) + \dots + x_n f(\vec{e}_n)$$

Die Vektoren $\vec{a}_1 = f(\vec{e}_1), \dots, \vec{a}_n = f(\vec{e}_n)$ sind dabei die Spaltenvektoren der Abbildungsmatrix A . Für das Bild von f gilt:

$$f(X) = L(\vec{a}_1, \dots, \vec{a}_n)$$

Die Dimension des Bildes ist somit die Anzahl der linear unabhängigen Spaltenvektoren der Abbildungsmatrix. ■

Um den Zusammenhang von Zeilenrang und Spaltenrang zu untersuchen müssen wir eine Vorüberlegung anstellen. Dazu betrachten wir die bekannten elementaren Zeilenoperationen **Z1**, **Z2**, **Z3** sowie die entsprechenden Spaltenoperationen:

- S1.** Addition eines Vielfachen einer Spalte zu einer anderen.
- S2.** Vertauschen zweier Spalten.
- S3.** Multiplikation einer Spalte mit einem Skalar $\lambda \neq 0, \lambda \in K$.

Wir werden die Auswirkungen dieser Operationen auf die entsprechenden Ränge einer Matrix untersuchen. Zunächst ist zu bemerken, dass jede Operation die entsprechende inverse Operation besitzt. Weiterhin gilt:

Satz 4.56: Zeilen- bzw. Spaltenoperationen ändern nicht den Zeilen- bzw. Spaltenrang einer Matrix.

BEWEIS

Zeilenoperationen/Zeilenrang:

Seien $\vec{z}_1, \dots, \vec{z}_m$ die m Zeilenvektoren einer Matrix A . Die Lineare Hülle dieser Vektoren ist gegeben durch $L(\vec{z}_1, \dots, \vec{z}_m)$. Wir wissen, dass der Zeilenrang gleich der Dimension der linearen Hülle ist.

- Die Zeilenoperation **Z2** ändert trivialerweise L nicht.
- Wegen $L(\vec{z}_1, \dots, \vec{z}_m) = L(\vec{z}_1, \dots, \lambda \vec{z}_i, \dots, \vec{z}_m)$ ändert die Zeilenoperation **Z3** auch die lineare Hülle nicht.
- Da auch $L(\vec{z}_1, \dots, \vec{z}_m) = L(\vec{z}_1, \dots, \vec{z}_j + \lambda \vec{z}_i, \dots, \vec{z}_i, \dots, \vec{z}_m)$ gilt, ändert die Zeilenoperation **Z1** ebenfalls die lineare Hülle nicht.

Folglich ändern die Zeilenoperationen den Zeilenrang nicht.

Zeilenoperationen/Spaltenrang:

Es ist zu zeigen, dass die Zeilenoperationen die Anzahl der linearen unabhängigen Spaltenvektoren nicht ändert. Es seien die Spaltenvektoren $\vec{s}_1, \dots, \vec{s}_i$ eine beliebige Teilmenge aller Spaltenvektoren der Matrix A .

Lineare Unabhängigkeit heißt:

$$\lambda_1 \vec{s}_1 + \dots + \lambda_i \vec{s}_i = 0 \Leftrightarrow \lambda_1 = \dots = \lambda_i = 0$$

Komponentenweise in Matrixschreibweise heißt das:

$$\begin{array}{cccc|c} s_{11} & s_{12} & \cdots & s_{1i} & 0 \\ \vdots & \vdots & & \vdots & \\ s_{n1} & s_{n2} & \cdots & s_{ni} & 0 \end{array} \Leftrightarrow \text{es gibt nur die triviale Lösung}$$

Die Lösung dieses Gleichungssystems mit dem Gaußschen Algorithmus entspricht genau der Anwendung der Zeilenoperationen **Z1-Z3**.

Also ändern die Zeilenoperationen **Z1 - Z3** die Lösung $\lambda_1, \dots, \lambda_i$ nicht.

Die Spaltenvektoren $\vec{s}_1, \dots, \vec{s}_i$ sind nach den Zeilenoperationen **Z1 - Z3** genau dann linear unabhängig, wenn sie es schon vorher waren. Da die Spaltenvektoren $\vec{s}_1, \dots, \vec{s}_i$ eine beliebige Teilmenge aller Spaltenvektoren der Matrix A sind, ändern die Zeilenoperationen **Z1 - Z3** den Spaltenrang nicht.

Analog verfährt man anschließend auch mit Spaltenoperationen/Spaltenrang und Spaltenoperationen/Zeilenrang. ■

Satz 4.57: Es sei A eine Matrix vom Zeilenrang r . Durch eine Folge von Zeilen- und Spaltenoperationen der Form **Z1 - Z3** bzw. **S1 - S3** kann man die Matrix A zu einer Matrix umformen, die in den ersten $r \times r$ Elementen die Form einer Einheitsmatrix hat und deren restliche Elemente 0 sind. Sie hat also die Darstellung

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Die Einsen stehen auf der Diagonalen der ersten r Spalten bzw. Zeilen.

Daraus folgt sofort:

Satz 4.58: Spaltenrang und Zeilenrang einer Matrix A sind gleich.

Beispiel 4.59:

$$\begin{array}{l}
 \\
 \\
 \\
 -2 \cdot \textcircled{1} \\
 -\textcircled{1} \\
 -2 \cdot \textcircled{1} \\
 \\
 -\textcircled{2} \\
 :3 \\
 \cdot(-1) \\
 +\textcircled{3} \\
 +\textcircled{3}
 \end{array}
 \begin{pmatrix}
 1 & 0 & 3 & 2 & -2 \\
 2 & 3 & 1 & 0 & 4 \\
 1 & 0 & 2 & 0 & 1 \\
 0 & 0 & 1 & 2 & -3 \\
 2 & 3 & 2 & 2 & 1 \\
 \\
 1 & 0 & 3 & 2 & -2 \\
 0 & 3 & -5 & -4 & 8 \\
 0 & 0 & -1 & -2 & 3 \\
 0 & 0 & 1 & 2 & -3 \\
 0 & 3 & -4 & -2 & 5 \\
 \\
 1 & 0 & 3 & 2 & -2 \\
 0 & 3 & -5 & -4 & 8 \\
 0 & 0 & -1 & -2 & 3 \\
 0 & 0 & 1 & 2 & -3 \\
 0 & 0 & 1 & 2 & -3 \\
 \\
 1 & 0 & 3 & 2 & -2 \\
 0 & 1 & -\frac{5}{3} & -\frac{4}{3} & \frac{8}{3} \\
 0 & 0 & 1 & 2 & -3 \\
 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 \\
 \\
 1 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0
 \end{pmatrix}$$

4.4 Determinanten

4.4.1 Einführung

Wir beginnen direkt mit der Definition:

Definition 4.60: Die **Determinante** $\det : M(n \times n, K) \rightarrow K$ ist eine Abbildung vom Raum der $(n \times n)$ -Matrizen über dem Körper K nach K mit den folgenden vier Eigenschaften:

1. Vertauscht man in A zwei benachbarte Spalten, dann ändert $\det(A)$ ihr Vorzeichen, d.h.

$$\det(\dots, \vec{a}_i, \vec{a}_{i+1}, \dots) = -\det(\dots, \vec{a}_{i+1}, \vec{a}_i, \dots)$$

2. Multipliziert man eine Spalte der Matrix A mit einer Konstanten $s \in K$, so multipliziert man die Determinante mit s , d.h.

$$\det(\vec{a}_1, \dots, s\vec{a}_i, \dots, \vec{a}_n) = s \cdot \det(\vec{a}_1, \dots, \vec{a}_i, \dots, \vec{a}_n)$$

Diese Eigenschaft nennt man **Homogenität**.

3. Stimmen zwei Matrizen überall bis auf eine Spalte i überein, so addiert man ihre Determinanten, indem man die beiden Spalten i addiert und dann die Determinante der Summe bildet, d.h.

$$\begin{aligned} & \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \quad \vec{a}_i, \quad \vec{a}_{i+1}, \dots, \vec{a}_n) \\ & + \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \quad \vec{a}_i^*, \quad \vec{a}_{i+1}, \dots, \vec{a}_n) \\ & = \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_i + \vec{a}_i^*, \vec{a}_{i+1}, \dots, \vec{a}_n) \end{aligned}$$

Dies ist die **Additivität** der Determinanten.

4. Es ist $\det(E) = 1$, wobei E die Einheitsmatrix ist.

Warum gerade diese (zunächst einmal ziemlich willkürlich erscheinende) Eigenschaften? Historisch war zunächst die Determinante über ihre Rechenvorschrift definiert. Es war bekannt, dass eine bestimmte Rechenvorschrift $\det(A)$ einer quadratischen Matrix (oder einem entsprechenden Gleichungssystem) einen Wert zuordnet, mit dem bestimmte Eigenschaften der Matrix ermittelt werden können. Einige davon haben wir bereits kennengelernt:

- Ein lineares Gleichungssystem ist genau dann eindeutig lösbar, wenn die Determinante ungleich 0 ist.

- Zu einem Gleichungssystem existiert eine eindeutige Umkehrabbildung, wenn die Determinante ungleich 0 ist.
- Durch Berechnen einer Determinante kann festgestellt werden, ob eine Menge reeller Funktionen linear unabhängig ist.

Mit der Determinante kann man also einfach die oben genannten Eigenschaften bestimmen (determinieren = bestimmen, festlegen!).

Die Determinante ist ein unentbehrliches Werkzeug, auf das man bei der Untersuchung und Ermittlung von Eigenschaften eines linearen Operators nicht verzichten kann (Schaum: Lineare Algebra, S. 171).

Mit dieser Definition sind allerdings einige Fragen schwer zu beantworten. Beispielsweise:

- Hat die Determinante noch andere nützliche Eigenschaften?
- Gibt es noch andere Rechenverfahren, die den gleichen Zweck erfüllen?

Daher wurde eine möglichst kurze Definition entwickelt, aus der automatisch die bekannten Eigenschaften (sowie eventuell neue) folgen. Daraus entwickelte man die möglichen Rechenverfahren zur Berechnung der Determinanten. Bei der Entwicklung erkannte man, dass es nur eine Rechenverfahren mit diesen Eigenschaften gibt.

Bemerkung 4.61: Für konkrete Matrizen A wird anstatt $\det(A)$ auch die kürzere Schreibweise $|A|$ verwendet. Beispiel:

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

4.4.2 Eigenschaften der Determinante

Aus den vier Eigenschaften, die laut Definition 4.60 eine Determinante definieren, lassen sich nun weitere nützliche Eigenschaften herleiten. Um hier nicht relevante Ausnahmen nicht diskutieren zu müssen, beschränken wir uns in diesem Unterkapitel auf die Körper $K \in \{\mathbb{R}, \mathbb{C}\}$. Wir beginnen mit einem Ergebnis der Berechnungsmethode (siehe Kapitel 4.4.5), das wir hier ohne Beweis angeben.

Bemerkung 4.62: Es gilt

$$\det(A) = \det(A^T) \quad \forall A \in M(n \times n, K),$$

wobei A^T die zu A transponierte Matrix ist.

Hieraus folgt, dass die Eigenschaften 1-4 einer Determinante auch für Zeilen an Stelle von Spalten gelten. Außerdem können wir uns damit für die Beweise der folgenden Aussagen auf Spalten beschränken ohne die Aussagekraft für Zeilen zu verlieren.

Bemerkung 4.63: Vertauscht man zwei beliebige Spalten/Zeilen, so ändert die Determinante ihr Vorzeichen.

BEWEIS Ist der Abstand zweier Spalten n , so ist die Anzahl der Vertauschungen zweier Nachbarn $2n - 1$, also eine ungerade Zahl. ■

Bemerkung 4.64: Besitzt eine Matrix A zwei gleiche Spalten/Zeilen, so gilt $\det(A) = 0$.

BEWEIS Durch die Spaltenvertauschung erhält man $\det(A) = -\det(A)$. ■

Bemerkung 4.65: Addiert man ein Vielfaches einer Spalte/Zeile zu einer anderen Spalte/Zeile, so ändert sich die Determinante nicht.

BEWEIS Spalte $i + \alpha \cdot$ Spalte j ; $i \neq j$:

$$\begin{aligned} & \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_i + \alpha \vec{a}_j, \vec{a}_{i+1}, \dots, \vec{a}_n) \\ &= \det(\vec{a}_1, \dots, \vec{a}_n) + \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \alpha \vec{a}_j, \vec{a}_{i+1}, \dots, \vec{a}_n) \\ &= \det(\vec{a}_1, \dots, \vec{a}_n) + \alpha \cdot \overbrace{\det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_j, \vec{a}_{i+1}, \dots, \vec{a}_n)}^{=0} \\ &= \det(\vec{a}_1, \dots, \vec{a}_n) \end{aligned}$$

■

Bemerkung 4.66: Addiert man ein Vielfaches einer Spalte/Zeile zu dem Vielfachen einer anderen Spalte/Zeile, so ändert sich die Determinante in der Art, dass der Faktor der Spalte/Zeile erhalten bleibt, in die die Summe eingetragen wird.

BEWEIS Diese Operation entspricht der Hintereinanderausführung einer Multiplikation einer Spalte mit einem Skalar und der anschließenden Addition eines Vielfachen einer anderen Spalte:

$\beta \cdot$ Spalte $i + \alpha \cdot$ Spalte j ; $i \neq j$:

$$\begin{aligned} & \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \beta \cdot \vec{a}_i + \alpha \vec{a}_j, \vec{a}_{i+1}, \dots, \vec{a}_n) \\ &= \beta \cdot \det(\vec{a}_1, \dots, \vec{a}_n) + \alpha \cdot \overbrace{\det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_j, \vec{a}_{i+1}, \dots, \vec{a}_n)}^{=0} \\ &= \beta \cdot \det(\vec{a}_1, \dots, \vec{a}_n) \end{aligned}$$

■

Für die folgende Bemerkung geben wir wieder keinen Beweis an, sie lässt sich aber wieder anhand der Rechenregeln aus Kapitel 4.4.5 nachweisen.

Bemerkung 4.67: Es gilt

$$\det(AB) = \det(A) \cdot \det(B) \quad \forall A, B \in M(n \times n, K).$$

Aus der Herleitung des Rangs einer quadratischen Matrix sieht man, dass die Gleichung $AX = E$ für jede Matrix $A \in M(n \times n, K)$ genau dann eine Lösung $X = A^{-1}$ hat, wenn gilt $\text{rg}(A) = n$. Dies ist genau dann der Fall, wenn gilt $\det(A) \neq 0$. Das entspricht genau der Aussage der wichtigen Bemerkung 4.46, in der wir festgestellt hatten, dass eine Matrix A genau dann invertierbar ist, wenn $\det(A) \neq 0$.

Die Anwendung des Satzes über die Determinante eines Produkts liefert

Bemerkung 4.68: Ist A invertierbar, d.h. $\det(A) \neq 0$, dann gilt

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

BEWEIS

$$\det(A) \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(E) = 1$$

■

4.4.3 Geometrische Interpretation

Interessanterweise hat die Determinante auch eine sinnvolle geometrische Bedeutung. Die Determinante von n Vektoren des \mathbb{R}^n ist gleich dem Betrag des n -dimensionalen „Volumens“, das von den Spaltenvektoren bzw. von den Zeilenvektoren der Matrix eingeschlossen wird. In Kapitel 2.3.1.1 hatten wir diesen Zusammenhang bereits im \mathbb{R}^3 untersucht. Allgemein im \mathbb{R}^n gilt dieser Zusammenhang zwar auch, wir können ihn aber nicht darstellen. Auf den \mathbb{R}^2 lässt er sich dagegen leicht übertragen:

Die Determinante einer 2×2 -Matrix ergibt

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc = \left\| \begin{pmatrix} 0 \\ 0 \\ ad - bc \end{pmatrix} \right\| = \left\| \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \times \begin{pmatrix} c \\ d \\ 0 \end{pmatrix} \right\|.$$

Nach Definition 2.31 gibt die euklidische Norm des Kreuzprodukts zweier Vektoren im \mathbb{R}^n aber genau den Flächeninhalt des aufgespannten Parallelogramms an. Somit gibt die Determinante einer 2×2 -Matrix also das 2-dimensionale „Volumen“, sprich selbigen Flächeninhalt, an.

4.4.4 Schematische Darstellung der Determinante

Die folgende schematische Darstellung der Determinanten kommt praktisch ohne mathematische Formeln aus.

Nehmen wir als Beispiel eine 3×3 -Matrix. Wir bilden Produkte aus je einem Element jeder Zeile, wobei alle Elemente aus unterschiedlichen Spalten stammen müssen. Z.B:

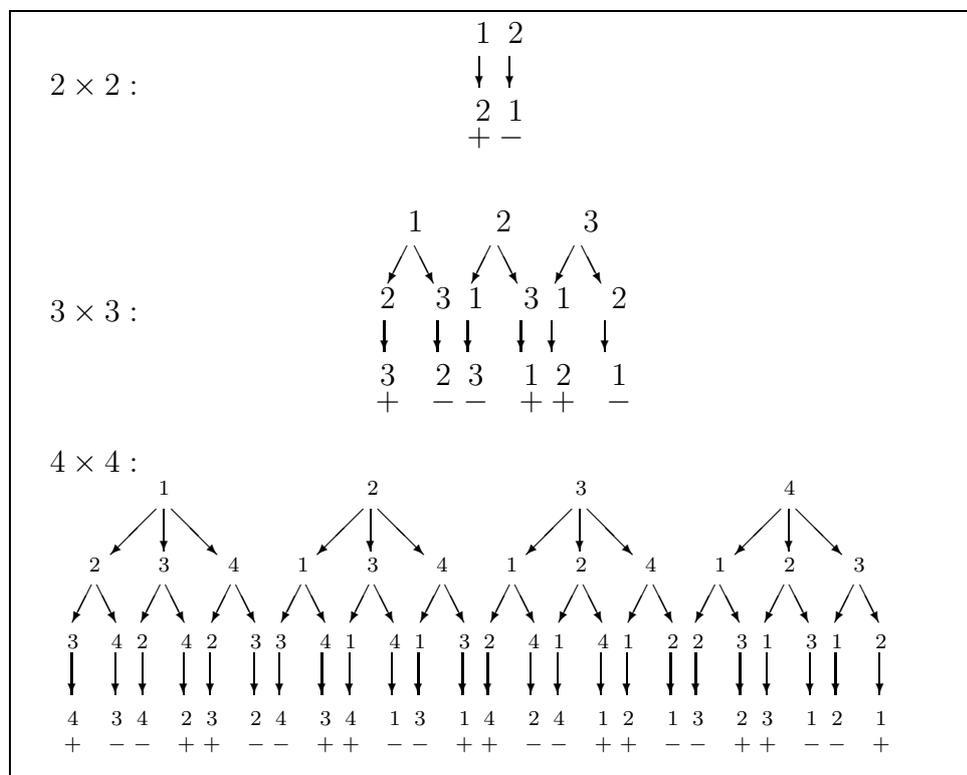
$$\left| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right| \quad \left| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right|$$

Wie viele von diesen Produkten gibt es? Für eine $n \times n$ -Matrix sind es $n!$, also 2 für eine 2×2 -Matrix, 6 für eine 3×3 -Matrix und 24 für eine 4×4 -Matrix. Diese Produkte werden mit einem Vorzeichen versehen (Regel später) und addiert. Das Ergebnis ist die Determinante. Die mathematische Formel dazu (später) ist die Leibnizsche Formel.

Die Produkte kann man in Baumform darstellen. Da der Zeilenindex sich automatisch aus dem Level des Elements im Baum ergibt, kann man sich auf den Spaltenindex beschränken:

$$2 \times 2 : \quad \begin{array}{cc} a_{11} & a_{12} \\ \downarrow & \downarrow \\ a_{22} & a_{21} \end{array} \quad \text{bzw.} \quad \begin{array}{cc} 1 & 2 \\ \downarrow & \downarrow \\ 2 & 1 \end{array}$$

Nachfolgend die entsprechende Bäume für 2×2 , 3×3 und 4×4 -Matrizen. Das Vorzeichen richtet sich nach der Anzahl der Vertauschungen, die nötig sind, um die Elemente des Pfades zwischen der Wurzel und dem entsprechenden Blatt in aufsteigende Reihenfolge zu bringen (Gerade Anzahl = positiv, ungerade Anzahl = negativ).



4.4.5 Rechenverfahren

4.4.5.1 Leibnizsche Formel

Es sollen jetzt die Regeln aus Definition 4.60 und Kapitel 4.4.2 angewandt werden, um zu zeigen, dass damit eine Formel gefunden werden kann, um den Wert der Determinante zu berechnen, und dass damit die Determinante eindeutig bestimmt wird. Wir wollen dies zunächst für zweireihige Determinanten durchführen und daran die Berechnungsmethode erkennen. Es sei also

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Für den 1. Spaltenvektor kann man nun schreiben:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_1 \vec{e}_1 + a_2 \vec{e}_2$$

und damit

$$\det(\vec{a}, \vec{b}) = \det(a_1 \vec{e}_1 + a_2 \vec{e}_2, \vec{b}) = a_1 \cdot \det(\vec{e}_1, \vec{b}) + a_2 \cdot \det(\vec{e}_2, \vec{b}),$$

wegen der Eigenschaften 2 und 3 der Definition. Wendet man die gleiche Methode auf den Vektor \vec{b} an, dann erhält man weiter

$$\begin{aligned} \det(\vec{a}, \vec{b}) &= a_1(b_1 \cdot \det(\vec{e}_1, \vec{e}_1) + b_2 \cdot \det(\vec{e}_1, \vec{e}_2)) + a_2(b_1 \cdot \det(\vec{e}_2, \vec{e}_1) + b_2 \cdot \det(\vec{e}_2, \vec{e}_2)) \\ &= a_1 b_1 \cdot \overbrace{\det(\vec{e}_1, \vec{e}_1)}^{=0} + a_1 b_2 \cdot \overbrace{\det(\vec{e}_1, \vec{e}_2)}^{=1} + a_2 b_1 \cdot \overbrace{\det(\vec{e}_2, \vec{e}_1)}^{=-1} + a_2 b_2 \cdot \overbrace{\det(\vec{e}_2, \vec{e}_2)}^{=0}. \end{aligned}$$

Zwei Determinanten sind gleich 0, weil die Matrizen gleiche Spalten haben. Wegen $\det(E) = 1$ und dem Vorzeichenwechsel beim Tausch von Zeilen gilt $\det(\vec{e}_1, \vec{e}_2) = 1$ und $\det(\vec{e}_2, \vec{e}_1) = -1$ und damit die [bekannte Formel](#)

$$D = \det(\vec{a}, \vec{b}) = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1.$$

Im Falle $n=3$ liefert die Anwendung dieser Eigenschaften die bekannte [Sarrussche Regel](#).

Die Anwendung der selben Regeln auf die 4-reihige Determinante von

$$A := \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix}$$

liefert für $\det(A)$ das Ergebnis:

$$\begin{aligned} &a_1 b_2 c_3 d_4 - a_1 b_2 d_3 c_4 + a_1 b_3 c_4 d_2 - a_1 b_3 c_2 d_4 + a_1 b_4 c_2 d_3 - a_1 b_4 c_3 d_2 \\ &- a_2 b_1 c_3 d_4 + a_2 b_1 d_3 c_4 - a_2 b_3 c_4 d_1 + a_2 b_3 c_1 d_4 - a_2 b_4 c_1 d_3 + a_2 b_4 c_3 d_1 \\ &+ a_3 b_1 c_2 d_4 - a_3 b_1 c_4 d_2 + a_3 b_2 c_4 d_1 - a_3 b_2 c_1 d_4 + a_3 b_4 c_1 d_2 - a_3 b_4 c_2 d_1 \\ &- a_4 b_1 c_2 d_3 + a_4 b_1 c_3 d_2 - a_4 b_2 c_3 d_1 + a_4 b_2 c_1 d_3 - a_4 b_3 c_1 d_2 + a_4 b_3 c_2 d_1. \end{aligned}$$

Dies sind insgesamt 24 Produkte. Man sieht also, dass die Sarrussche Regel für $n = 4$ nicht mehr gilt. Außerdem erkennt man, dass in jedem Produkt aus jeder Zeile und jeder Spalte genau ein Element vorkommt, und dass sämtliche Indexkombinationen vorkommen. Die Anzahl dieser verschiedenen Kombinationen ist genau $n!$. Bei jedem Produkt steht, wie wir im Falle $n = 2$ gesehen haben, zunächst eine Determinante, die nur aus Einheitsvektoren besteht. Ihr Wert ist daher:

0, wenn gleiche Vektoren vorkommen,

1, wenn sie durch eine gerade Anzahl von Vertauschungen in die Determinante der Einheitsmatrix überführt werden kann bzw.

-1, wenn die Anzahl der nötigen Vertauschungen ungerade ist.

Für n -reihige Determinanten sieht die Herleitung formal so aus:

- \vec{a}_1 bis \vec{a}_n als Summe schreiben: $\vec{a}_i = \sum_{k=1}^n a_{ki} \vec{e}_k$.
- Für alle Summen die Determinanten aufspalten. Man erhält eine Summe aus n^n Summanden. Jeder Summand hat n Vorfaktoren (je einen Vorfaktor aus jeder Spalte) und eine Determinante aus Einheitsvektoren.
- Alle Summanden, die doppelte Einheitsvektoren in der Determinante haben, fallen weg. Es bleiben nur noch Summanden, die Permutationen von $(\vec{e}_1, \dots, \vec{e}_n)$ in der Determinante haben ($n!$ Summanden).
- Die Determinanten sind 1 oder -1, je nachdem, ob sie sich durch eine gerade oder eine ungerade Anzahl von Vertauschungen in $\det(\vec{e}_1, \dots, \vec{e}_n)$ überführen lassen.

Satz 4.69: (Leibnizsche Formel): Die Determinante einer n -reihigen Matrix $A = (a_{ij})$ kann dann wie folgt berechnet werden:

$$\det(A) = \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \delta^{i_1, \dots, i_n} a_{i_1, 1} \cdots a_{i_n, n},$$

wobei gilt

$$\delta^{i_1, \dots, i_n} = \begin{cases} 0, & \text{falls gleiche Indizes unter den } i_1, \dots, i_n \text{ vorkommen} \\ 1, & \text{falls } i_1 \dots i_n \text{ durch eine gerade Zahl von Vertauschungen} \\ & \text{in die Reihenfolge } 1, 2, \dots, n \text{ gebracht werden kann} \\ -1, & \text{falls die Zahl der notwendigen Vertauschungen ungerade ist.} \end{cases}$$

Diese Formel liefert eine Berechnungsmethode der Determinanten, die nur aus den Eigenschaften 1-4 aus Definition 4.60 hergeleitet werden kann und zeigt damit, dass eine n -reihige Determinante durch diese vier Eigenschaften eindeutig bestimmt ist. Damit kann man auch den folgenden Entwicklungssatz herleiten.

4.4.5.2 Entwicklungssatz nach Laplace

Definition 4.70: Ist $A \in M(n \times n, K)$, so bezeichnet A_{ij} die $(n-1) \times (n-1)$ -Matrix, die durch Streichen der i -ten Zeile und der j -ten Spalte von A entsteht. Damit ergibt sich folgendes Ergebnis:

Satz 4.71: (Entwicklungssatz nach Laplace): Es sei $A \in M(n \times n, K)$ und j ein beliebiger fester Index mit $1 \leq j \leq n$, dann gilt

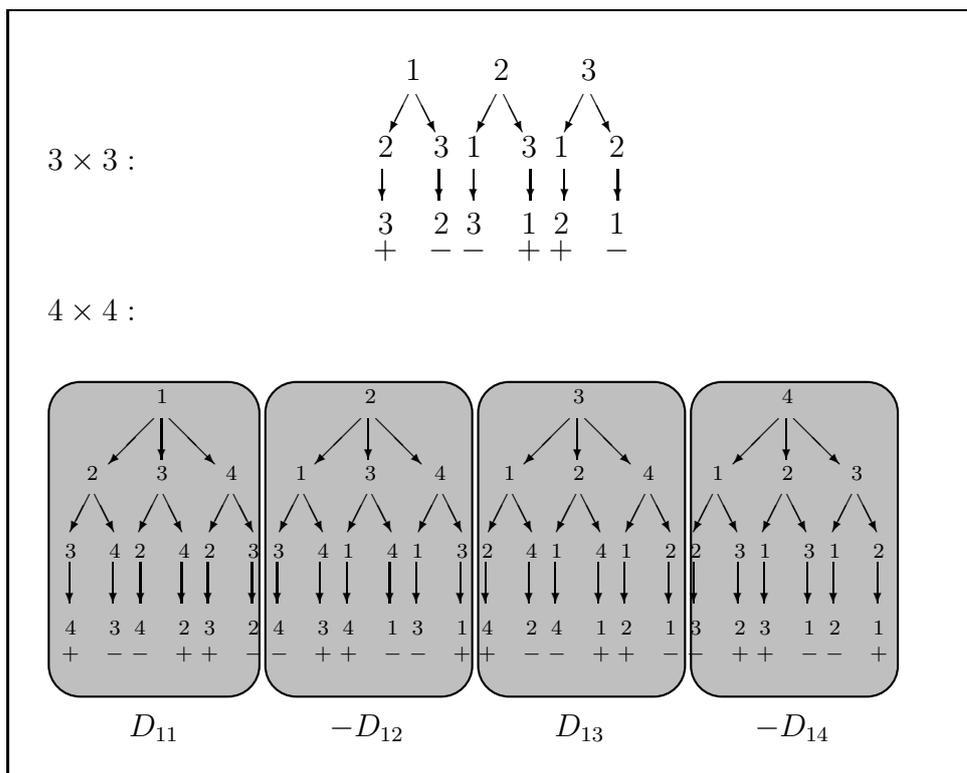
$$\det(A) = \sum_{i=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Hier sind die a_{ij} die Elemente der j -ten Spalte. Das nennt man **Entwicklung der Determinanten nach der j -ten Spalte**.

Statt nach einer Spalte kann ebensogut auch **nach einer Zeile entwickelt** werden:

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Der Entwicklungssatz nach Laplace lässt sich auch in der [schematischen Darstellung](#) der Determinante wiederfinden. Das Schema zur Berechnung der $n \times n$ -Matrix umfasst n unterschiedliche $(n-1) \times (n-1)$ -Untermatrizen:



Diese Berechnungsformel für eine Determinante kann auch als rekursive Definition der Determinanten aufgefasst werden. Dazu muss man dann noch

für $n = 1$ die Determinante definieren durch

$$\det(A) = a_{11}.$$

Beispiel 4.72: Entwicklung nach der 1. Zeile:

$$\begin{aligned} \begin{vmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} &= 2 \cdot \begin{vmatrix} 2 & -3 \\ 3 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 4 & 2 \\ 5 & 3 \end{vmatrix} \\ &= 2(2 + 9) + (12 - 10) = 24 \end{aligned}$$

Aus dem Entwicklungssatz nach Laplace folgt direkt ein einfacher Sonderfall:

Bemerkung 4.73: Ist $A \in M(n \times n, K)$ eine obere oder untere Dreiecksmatrix, dann ist $\det(A)$ das Produkt der Elemente auf der Hauptdiagonalen.

$$\det \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{nn} \end{pmatrix} = b_{11} \cdot b_{22} \cdot \cdots \cdot b_{nn}.$$

4.4.5.3 Gauß-Algorithmus

Dieses Ergebnis motiviert die Berechnung einer Determinante mit Hilfe des Gauß-Algorithmus. Dabei werden die Regeln zur Umformung der Matrix entsprechend der Regeln zur Lösung eines Gleichungssystems nach Gauß angewendet bis eine Dreiecksgestalt erreicht ist. Allerdings muss man aufgrund der bisher gezeigten Eigenschaften der Determinante beachten:

- Die Multiplikation einer Zeile/Spalte mit einem Faktor A bewirkt, dass die Determinante mit dem Faktor A multipliziert wird. Dies muss ausgeglichen werden. Beispiel:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 2a & 2b \\ c & d \end{vmatrix}$$

- Die Vertauschung zweier Zeilen bzw. Spalten verändert das Vorzeichen der Determinante:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = - \begin{vmatrix} c & d \\ a & b \end{vmatrix}$$

- Einzig die Operation „Zeile $a + \alpha \cdot$ Zeile b “ bzw. „Spalte $a + \alpha \cdot$ Spalte b “ verändert den Wert der Determinanten nicht.

Damit bringt man die Matrix auf Dreiecksform. Die Determinante ist das Produkt der Hauptdiagonalelemente.

Beispiel 4.74:

$$\begin{aligned}
 \begin{vmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} &= \frac{1}{2} \begin{vmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 10 & 6 & 2 \end{vmatrix} \\
 &= \frac{1}{2} \begin{vmatrix} 2 & 0 & 1 \\ 0 & 2 & -5 \\ 0 & 6 & -3 \end{vmatrix} \\
 &= \frac{1}{2} \begin{vmatrix} 2 & 0 & 1 \\ 0 & 2 & -5 \\ 0 & 0 & 12 \end{vmatrix} \\
 &= \frac{1}{2} \cdot 2 \cdot 2 \cdot 12 = 24
 \end{aligned}$$

Beispiel 4.75: Geschickte Kombination von Zeilen- und Spaltenoperationen. Erklärung der Kürzel: Z2=Zeile 2; S2=Spalte 2.

$$\begin{aligned}
 &\begin{vmatrix} t-1 & 1 & 1 \\ 1 & t-1 & 1 \\ 1 & 1 & t-1 \end{vmatrix} && |Z2 - Z1 \rightarrow Z2; \quad Z3 - Z1 \rightarrow Z3 \\
 &= \begin{vmatrix} t-1 & 1 & 1 \\ 2-t & t-2 & 0 \\ 2-t & 0 & t-2 \end{vmatrix} && |S1 + S2 + S3 \rightarrow S1 \\
 &= \begin{vmatrix} t+1 & 1 & 1 \\ 0 & t-2 & 0 \\ 0 & 0 & t-2 \end{vmatrix} \\
 &= (t-2)^2 \cdot (t+1)
 \end{aligned}$$

4.4.5.4 Sarrussche Regel

Für 2×2 - und für 3×3 -Matrizen kann man auch die [Sarrussche Regel](#) anwenden:

Beispiel 4.76: $\begin{vmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} = 2 \cdot 2 \cdot 1 + 1 \cdot 4 \cdot 3 - 2 \cdot (-3) \cdot 3 - 1 \cdot 2 \cdot 5 = 2 + 12 + 18 - 10 =$

4.4.5.5 Kombination der Regeln

Es können zur Berechnung der Determinanten auch mehrere Regeln kombiniert werden. Beliebt ist z.B. für 4×4 -Matrizen die Kombination: 1 Schritt Laplace, danach Sarrus. Gut funktioniert auch: Gauß, bis nur noch wenige Werte in einer Zeile/Spalte stehen, danach Laplace.

Beispiel 4.77: Mit der Operation „ $S1 - 2 \cdot S3 \rightarrow S1$ “ wird erreicht, dass nur eine einzige Zahl $\neq 0$ in der ersten Zeile steht. Danach wird mit Laplace fortgesetzt.

$$\begin{vmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ 10 & 2 & -3 \\ 3 & 3 & 1 \end{vmatrix} \\ = 30 - 6 = 24$$

4.4.6 Rechenaufwand zur Berechnung einer Determinanten

Dies nennt man die **Komplexität** des Problems. Es sei

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix}.$$

Es werden nun durch Zeilenoperationen, die den Wert der Determinanten nicht ändern, unterhalb der Diagonalen Nullen erzeugt. Wir nehmen zunächst an, dass $a_{11} \neq 0$ ist. Als Beispiel soll a_{21} Null werden. Dann gilt für die neue 2. Zeile:

$$a_{2i}^{(1)} = a_{2i} - a_{1i} \frac{a_{21}}{a_{11}}, \quad i = 1, \dots, n.$$

$a_{21}^{(1)}$ wird nicht berechnet, da es immer Null wird. Entsprechend kann man überall in der 1. Spalte unterhalb der Diagonalen Nullen erzeugen und dieses Verfahren für die weiteren Spalten so fortführen, bis man eine obere Dreiecksmatrix erhält. Berechnet man die Summe sämtlicher Multiplikationen und Divisionen, die dazu nötig sind, dann erhält man

$$\frac{n^3 - n}{3}.$$

Für $n = 10$ ist der Wert 330. Beim Entwicklungssatz erhält man einen Wert $\geq n!$, für $n = 10$ also $10! = 3628800 \approx 3.6 \cdot 10^6$. Die Leibnizsche Formel braucht $(n-1)n!$ Multiplikationen, also noch mehr.

4.4.7 Berechnung der Inversen mit der Methode der Unterdeterminanten

Definition 4.78: Unter der zu A komplementären Matrix \tilde{A} versteht man die Matrix mit den Elementen

$$\tilde{a}_{ij} = (-1)^{i+j} \cdot \det(A_{ji}).$$

Bemerkung 4.79: Für die inverse Matrix gilt:

$$A^{-1} = \frac{1}{\det(A)} \cdot \tilde{A}.$$

Beispiel 4.80: Die Matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 4 & 1 & 1 \\ 3 & 2 & -7 \end{pmatrix}$$

wird einmal mit dem Gaußschen Algorithmus und einmal mit der Methode der Unterdeterminanten invertiert.

Gaußscher Algorithmus:

$$\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 & & & \\ 4 & 1 & 1 & 0 & 1 & 0 & & | & -4 \cdot \textcircled{1} \\ 3 & 2 & -7 & 0 & 0 & 1 & & | & -3 \cdot \textcircled{1} \\ \hline 1 & 0 & 2 & 1 & 0 & 0 & & & \\ 0 & 1 & -7 & -4 & 1 & 0 & & & \\ 0 & 2 & -13 & -3 & 0 & 1 & & | & -2 \cdot \textcircled{2} \\ \hline 1 & 0 & 2 & 1 & 0 & 0 & & | & -2 \cdot \textcircled{3} \\ 0 & 1 & -7 & -4 & 1 & 0 & & | & +7 \cdot \textcircled{3} \\ 0 & 0 & 1 & 5 & -2 & 1 & & & \\ \hline 1 & 0 & 0 & -9 & 4 & -2 & & & \\ 0 & 1 & 0 & 31 & -13 & 7 & & & \\ 0 & 0 & 1 & 5 & -2 & 1 & & & \end{array}$$

Methode mit Unterdeterminanten:

1. Schritt: Aufstellen einer Matrix B , an deren Position b_{ij} die Unterdeterminante A_{ij} steht.

$$B = \begin{pmatrix} \begin{vmatrix} 1 & 1 \\ 2 & -7 \end{vmatrix} & \begin{vmatrix} 4 & 1 \\ 3 & -7 \end{vmatrix} & \begin{vmatrix} 4 & 1 \\ 3 & 2 \end{vmatrix} \\ \begin{vmatrix} 0 & 2 \\ 2 & -7 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 3 & -7 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} \\ \begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 4 & 1 \end{vmatrix} \end{pmatrix} \\ = \begin{pmatrix} -9 & -31 & 5 \\ -4 & -13 & 2 \\ -2 & -7 & 1 \end{pmatrix}$$

2. Schritt: Verändern des Vorzeichens nach dem Schachbrettmuster.

$$C = \begin{pmatrix} -9 & 31 & 5 \\ 4 & -13 & -2 \\ -2 & 7 & 1 \end{pmatrix}$$

3. Schritt: Transponieren der Matrix. Das Ergebnis ist die komplementäre Matrix \tilde{A} .

$$\tilde{A} = \begin{pmatrix} -9 & 4 & -2 \\ 31 & -13 & 7 \\ 5 & -2 & 1 \end{pmatrix}$$

4. Schritt: Teilen durch $\det(A)$. Das Ergebnis ist die Inverse A^{-1} . Es gilt:

$$\det(A) = -7 + 16 - 2 - 6 = 1$$

Damit ist:

$$A^{-1} = \begin{pmatrix} -9 & 4 & -2 \\ 31 & -13 & 7 \\ 5 & -2 & 1 \end{pmatrix}$$

Für 2×2 -Matrizen ist die Berechnung der Unterdeterminanten aber einfacher und schneller als der Gaußsche Algorithmus.

Beispiel 4.81:

$$\begin{aligned}
 A &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
 B &= \begin{pmatrix} d & c \\ b & a \end{pmatrix} \\
 C &= \begin{pmatrix} d-c & \\ -b & a \end{pmatrix} = \tilde{A}^T \\
 \tilde{A} &= \begin{pmatrix} d-b & \\ -c & a \end{pmatrix} \\
 A^{-1} &= \frac{1}{ad-bc} \begin{pmatrix} d-b & \\ -c & a \end{pmatrix}
 \end{aligned}$$

Beispiel 4.82:

$$\begin{aligned}
 A &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\
 A^{-1} &= -\frac{1}{2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}
 \end{aligned}$$

4.4.8 Unterdeterminanten und Rang einer Matrix

Bemerkung 4.83: Es sei $A \in M(m \times n, K)$. Es sei r die Anzahl der Reihen der größten $(r \times r)$ -Untermatrix von A , deren Determinante ungleich Null ist, dann ist r der Rang der Matrix.

Beispiel 4.84: Es sei

$$A = \begin{pmatrix} 7 & 25 \\ 3 & 14 \\ -2 & 06 \\ 4 & -31 \end{pmatrix}$$

Welchen Rang hat A ? Durch Streichen der Zeile i erhält man vier mögliche 3×3 Untermatrizen A_i :

$$A_1 = \begin{pmatrix} 3 & 14 \\ -2 & 06 \\ 4 & -31 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 7 & 25 \\ -2 & 06 \\ 4 & -31 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 7 & 2 & 5 \\ 3 & 1 & 4 \\ 4 & -3 & 1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 7 & 2 & 5 \\ 3 & 1 & 4 \\ -2 & 0 & 6 \end{pmatrix}$$

Nach der vorigen Bemerkung hat A genau dann den Rang 3, wenn die Determinante von mindestens einer dieser Untermatrizen $\neq 0$ ist. Es ergibt sich:

$$\det(A_1) = 24 + 24 + 54 + 2 = 104 \neq 0.$$

Damit hat A den Rang 3. Da A_1 sich aus den ersten drei Zeilen der Matrix zusammensetzt, sind diese drei Zeilen linear unabhängig.

4.5 Transformationsmatrizen

4.5.1 Basiswechsel

Bisher haben wir die Koordinaten eines Vektors im \mathbb{R}^n immer bezüglich der kanonischen Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_n$ angegeben. Zum Beispiel bedeutete $\vec{x} = (3, 1, 4)^T$, dass $\vec{x} = 3\vec{e}_1 + \vec{e}_2 + 4\vec{e}_3$. Man kann den gleichen Vektor aber auch bezüglich einer beliebigen anderen Basis des \mathbb{R}^n wiedergeben. Nehmen wir z.B. die Basis $A = \{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ mit

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{a}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Dann hat der Vektor \vec{x} bezüglich der Basis A die Koordinaten $(2, 1, 3)^T$. Wir verwenden als Schreibweise

$$K_A(\vec{x}) = (2, 1, 3)^T,$$

wenn explizit unterschieden werden soll, auf welche Basis sich ein Vektor bezieht. Fehlt die Angabe der Basis, sind immer die kanonischen Einheitsvektoren gemeint.

Darüberhinaus möchten wir neben dem \mathbb{R}^n auch andere Vektorräume untersuchen. Deshalb suchen wir eine Möglichkeit, die Koordinaten bzgl. unterschiedlicher Basen in beliebigen Vektorräumen ineinander umzurechnen.

Gegeben sei ein Vektorraum der Dimension n mit den Basen

$$A = \{\vec{a}_1, \dots, \vec{a}_n\} \quad \text{und} \quad A' = \{\vec{a}'_1, \dots, \vec{a}'_n\}.$$

Für einen Vektor \vec{x} existieren also die Darstellungen

$$\vec{x} = \underbrace{(K_A(\vec{x}))_1 \cdot \vec{a}_1 + \dots + (K_A(\vec{x}))_n \cdot \vec{a}_n}_{A \cdot K_A(\vec{x})} = \underbrace{(K_{A'}(\vec{x}))_1 \cdot \vec{a}'_1 + \dots + (K_{A'}(\vec{x}))_n \cdot \vec{a}'_n}_{A' \cdot K_{A'}(\vec{x})}.$$

Daraus lassen sich direkt die **Transformationen**

$$K_A(\vec{x}) = \underbrace{A^{-1}A'}_{T_A^{A'}} \cdot K_{A'}(\vec{x})$$

bzw.

$$K_{A'}(\vec{x}) = \underbrace{A'^{-1}A}_{T_A^A} \cdot K_A(\vec{x})$$

ableiten, wobei $T_A^{A'}$ die **Transformationsmatrix** von A' nach A genannt wird. In Worten ausgedrückt bestehen die Spalten von $T_A^{A'}$ aus den Koordinaten der Einheitsvektoren von A' bezüglich der Basis A . Analoges gilt auch für T_A^A .

Bemerkung 4.85: Die Transformation von A' nach A ist die Umkehrung der Transformation von A nach A' :

$$T_A^{A'} = A^{-1}A' = (A'^{-1}A)^{-1} = (T_A^A)^{-1}$$

Da A und A' Basen bilden, sind sie und somit auch $T_A^{A'}$ und T_A^A invertierbar.

Beispiel 4.86: Bezüglich der Basen

$$A = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\} \quad \text{und} \quad A' = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

hat der Vektor $\vec{x} = (3, 1, 4)^T$ die Koordinaten

$$K_A(\vec{x}) = (2, 1, 3)^T \quad \text{bzw.} \quad K_{A'}(\vec{x}) = (3, -4, 0)^T.$$

Ein beliebiger Vektor \vec{x} lässt sich mit Hilfe der Transformationsmatrix

$$\begin{aligned} T_A^{A'} = A^{-1}A' &= \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 0 & 1 \\ -1 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & -1 \\ -2 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ -1 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -2 & -2 & 1 \\ 3 & 2 & 0 \\ -3 & -3 & 1 \end{pmatrix} \end{aligned}$$

von A' nach A überführen. Weiterhin gilt damit für die Umkehrung:

$$T_A^A = (T_A^{A'})^{-1} = \begin{pmatrix} -2 & -2 & 1 \\ 3 & 2 & 0 \\ -3 & -3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 1 & 2 \\ 3 & -1 & -3 \\ 3 & 0 & -2 \end{pmatrix}.$$

Die Korrektheit lässt sich mit obigem Beispielvektor überprüfen:

$$\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = K_A(\vec{x}) = T_A^{A'} \cdot K_{A'}(\vec{x}) = \begin{pmatrix} -2 & -2 & 1 \\ 3 & 2 & 0 \\ -3 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

Bemerkung 4.87: Ist A die kanonische Einheitsbasis, so ergibt sich als Transformationsmatrix

$$T_A^{A'} = A^{-1} \cdot A' = E^{-1} \cdot A' = A'.$$

Dieses Ergebnis erhält man für den \mathbb{R}^n auch aus Bemerkung 4.28, wenn die Basistransformation als Abbildung interpretiert wird.

4.5.2 Basistransformation bei Abbildungsmatrizen

Das Aussehen der Abbildungsmatrizen hängt nach Bemerkung 4.28 von der Wahl der Basen im Definitionsbereich und im Bild ab.

Definition 4.88: Sei A eine Basis des Vektorraums V sowie B eine Basis des Vektorraums W . Eine Matrix M_B^A repräsentiert die lineare Abbildung $\Phi : V \rightarrow W$ bezüglich der Basen A und B , wenn gilt:

$$K_B(\Phi(x)) = M_B^A(\Phi) \cdot K_A(x)$$

Die Matrix $M_B^A(\Phi)$ heißt **Abbildungsmatrix** von Φ bezüglich A und B .

Wie sieht nun die Abbildungsmatrix der gleichen linearen Abbildung bezüglich der Basen A' und B' aus? Es gilt:

Satz 4.89: Es sei $\Phi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung und $M_B^A(\Phi)$ ihre Abbildungsmatrix bezüglich A und B , dann ist

$$M_{B'}^{A'}(\Phi) = T_{B'}^B \cdot M_B^A(\Phi) \cdot T_A^{A'}$$

$$\begin{array}{ccc} \begin{array}{|c|} \hline K_A(\vec{x}) \\ \hline \end{array} & \xrightarrow{M_B^A} & \begin{array}{|c|} \hline K_B(\Phi(\vec{x})) \\ \hline \end{array} \\ \begin{array}{|c|} \hline T_A^{A'} \\ \hline \end{array} \uparrow & & \downarrow \begin{array}{|c|} \hline T_{B'}^B \\ \hline \end{array} \\ \begin{array}{|c|} \hline K_{A'}(\vec{x}) \\ \hline \end{array} & \xrightarrow{M_{B'}^{A'}} & \begin{array}{|c|} \hline K_{B'}(\Phi(\vec{x})) \\ \hline \end{array} \\ \in \mathbb{R}^m & & \in \mathbb{R}^n \end{array}$$

Die Spalten einer Transformationsmatrix $T_{B'}^B$ sind die in Kapitel 4.5.1 erläuterten Einheitsvektoren der Basis B bezüglich der Basis B' . Analoges gilt für $T_A^{A'}$.

Beispiel 4.90: Die Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ mit $f((x, y, z)^T) = (x, y)^T$ projiziert den \mathbb{R}^3 auf die (x, y) -Ebene. Die Abbildungsmatrix bezüglich der kanonischen Basen A und B ist

$$M_B^A(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Um die Abbildungsmatrix bezüglich der Basen

$$A' = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\} \quad \text{und} \quad B' = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}$$

zu bestimmen, nutzen wir also

$$\begin{aligned} M_{B'}^{A'}(f) &= T_{B'}^B \cdot M_B^A(f) \cdot T_A^{A'} = (T_B^{B'})^{-1} \cdot M_B^A(f) \cdot T_A^{A'} \\ &= \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -4 & 2 & 3 \\ 3 & -1 & -2 \end{pmatrix} \end{aligned}$$

4.5.3 Koordinatentransformationen in der Ebene

Zwei Basen, die durch eine lineare Koordinatentransformation $T_A^{A'} \cdot \vec{x}$ ineinander überführt werden können, sind zueinander gedreht, gespiegelt und/oder verzerrt, jedoch sind die Nullpunkte nicht gegeneinander verschoben. Dies gilt allgemein im \mathbb{R}^n . Der Einfachheit halber beschränken wir uns auf den \mathbb{R}^2 .

Beispiel 4.91: Die folgenden Transformationen spielen aufgrund der geometrischen Vorstellung eine besondere Rolle:

- $T_A^{A'} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \Rightarrow$ **Drehung** des Koordinatensystems um den Winkel ϕ
- $T_A^{A'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \Rightarrow$ **Spiegelung** des Koordinatensystems an der x -Achse

- $T_A^{A'} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow$ **Verzerrung** des Koordinatensystems in y -Richtung

Ist zusätzlich die Determinante von $T_A^{A'}$ gleich 1, dann kommt nur noch eine Drehung in Frage. Ist die Determinante gleich -1 , wurde zusätzlich gespiegelt (ohne Beweis).

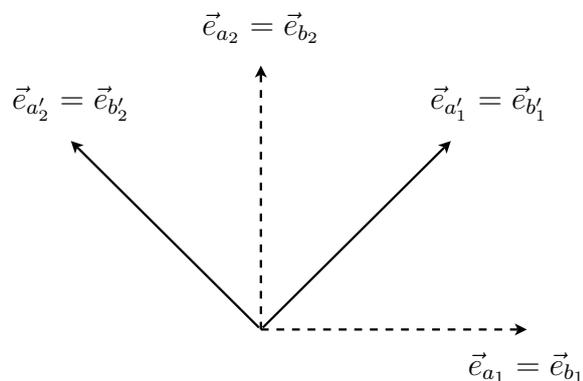
Beispiel 4.92: Eine Abbildung $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ sei für die kanonischen Basen $A = B$ gegeben durch

$$M_B^A(\Phi) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

beschreibt also eine Spiegelung an der y -Achse, wenn sowohl im Definitionsbereich als auch im Bild die kanonischen Basen gewählt werden. Die Einheitsvektoren der Basen $A' = B'$ haben bezüglich der Basen A bzw. B die Werte:

$$\vec{e}_{a'_1} = \vec{e}_{b'_1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{e}_{a'_2} = \vec{e}_{b'_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Das neue Koordinatensystem A' bzw. B' für den Definitionsbereich und das Bild ist damit gegenüber A bzw. B jeweils um 45° gedreht.



Die Transformationsmatrizen haben die Form:

$$T_A^{A'} = T_B^{B'} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Wie sieht die Abbildung Φ bezüglich der Basen A' und B' aus?

$$\begin{aligned} M_{B'}^{A'}(\Phi) &= (T_B^{B'})^{-1} \cdot M_B^A(\Phi) \cdot T_A^{A'} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Die Abbildung ist, wie zu erwarten war, die Spiegelung an der Mitteldiagonalen.

Bemerkung 4.93: Die Invertierung der Transformationsmatrix $(T_B^{B'})^{-1}$ erfordert einen Aufwand, den wir - wie in Bemerkung 6.6 dargestellt - in den sehr häufig auftretenden Spezialfällen der Orthogonalmatrizen aus Kapitel 6.1 deutlich reduzieren können.

Kapitel 5

Lineare Gleichungssysteme

5.1 Lösbarkeit eines linearen Gleichungssystems

Zunächst einmal verallgemeinern wir die Definitionen aus Kapitel 1.2.1:

Definition 5.1: Ist $A = (a_{ij}) \in M(m \times n, K)$ und $b = (b_1, \dots, b_m) \in K^m$, dann heißt

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

ein **lineares Gleichungssystem** bzgl. (x_1, \dots, x_n) mit Koeffizienten in K . x_1, \dots, x_n sind die Unbekannten des Systems. Ist b der Nullvektor, dann nennt man das System **homogen**, sonst **inhomogen**. Das Gleichungssystem kann in der Form

$$Ax = b$$

geschrieben werden.

Definition 5.2: Unter der **Lösungsmenge** des zu (A, b) gehörigen Gleichungssystems verstehen wir

$$L(A, b) = \{x \in K^n \mid Ax = b\}.$$

Die aus dem vorherigen Kapitel bekannten Eigenschaften über lineare Abbildungen und Matrizen führen zu folgenden Ergebnissen.

Satz 5.3: $Ax = b$ ist genau dann lösbar, wenn gilt:

$$\operatorname{rg} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \operatorname{rg} \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

Eine andere Schreibweise ist

$$\operatorname{rg}(A) = \operatorname{rg}(A, b).$$

BEWEIS Es seien $\vec{a}_1, \dots, \vec{a}_n$ die Spaltenvektoren der Matrix A . Es gilt:

$$Ax = (\vec{a}_1, \dots, \vec{a}_n)\vec{x} = x_1 \cdot \vec{a}_1 + \cdots + x_n \cdot \vec{a}_n = \vec{b}$$

Daraus folgt, dass ein Lösungsvektor \vec{x} genau dann existiert, wenn gilt

$$\vec{b} \in L\{\vec{a}_1, \dots, \vec{a}_n\} \Leftrightarrow \operatorname{rg}(A) = \operatorname{rg}(A, b),$$

wobei L die lineare Hülle der Vektoren ist. ■

Satz 5.4: Ist $x_s \in K^n$ eine Lösung, d.h. gilt $Ax_s = b$, dann gilt

$$L(A, b) = x_s + \operatorname{Ker}(A) = \{x_s + x \mid x \in \operatorname{Ker}(A)\}.$$

BEWEIS

$$\begin{aligned} x \in \operatorname{Ker}(A) &\Leftrightarrow Ax = 0 \\ \Leftrightarrow A(x_s + x) &= Ax_s + Ax = Ax_s = b \end{aligned}$$

■

Bemerkung 5.5: Ist x_s eine Lösung und (v_1, \dots, v_r) eine Basis von $\operatorname{Ker}(A)$, dann ist

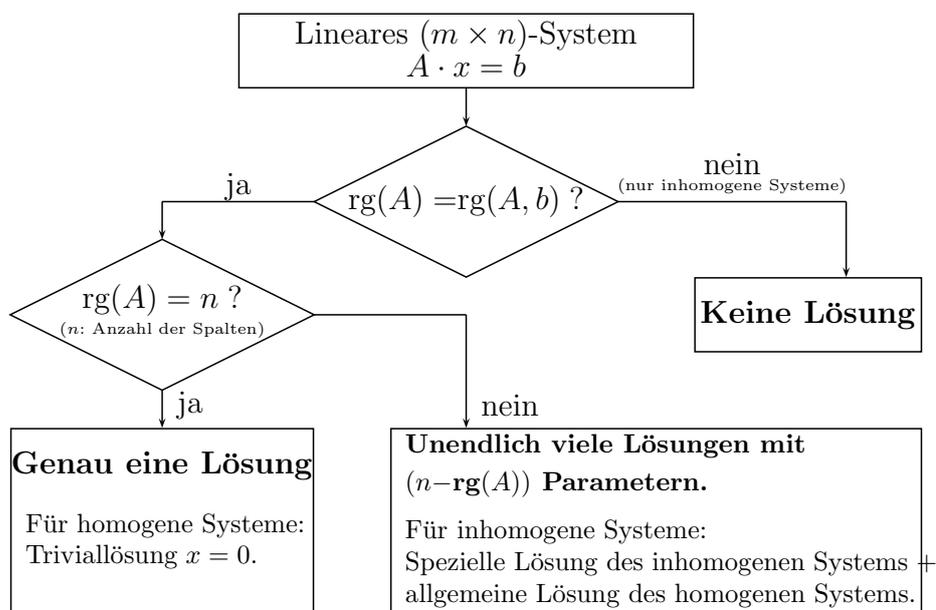
$$L(A, b) = \{x_s + \lambda_1 v_1 + \dots + \lambda_r v_r \mid \lambda_i \in K\}.$$

Dabei gilt wegen der Dimensionsformel für lineare Abbildungen: $r = \dim(\operatorname{Ker}(A)) = n - \operatorname{rg}(A)$.

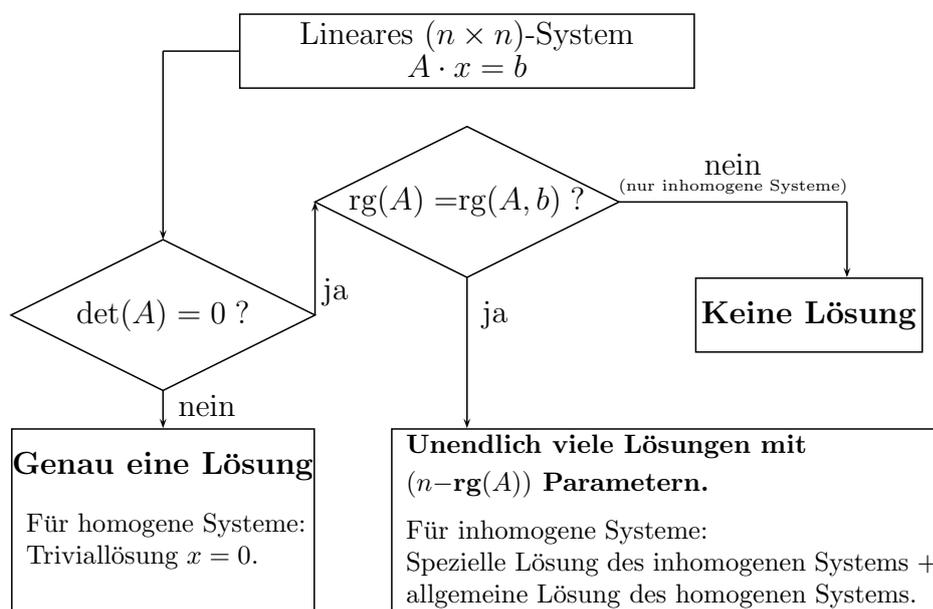
Erläuterung: Ist x_s eine **spezielle Lösung des inhomogenen Gleichungssystems** $Ax = b$ und ist x_h die **allgemeine Lösung des homogenen Systems** $Ax = 0$, dann ist $x_a = x_h + x_s$ die **allgemeine Lösung des inhomogenen Gleichungssystems** $Ax = b$.

Bemerkung 5.6: Ein lösbares Gleichungssystem ist genau dann eindeutig lösbar, wenn $\text{Ker}(A) = 0$, d.h. $\text{rg}(A) = n$ ist.

Eine eindeutige Lösung existiert also genau dann, wenn $\text{rg}(A) = \text{rg}(A, b) = n$ ist.



Bemerkung 5.7: Ist A quadratisch, d.h. $m = n$, so ist das Gleichungssystem $Ax = b$ genau dann eindeutig lösbar, wenn $\det(A) \neq 0$ ist.

**Beispiel 5.8:**

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}; \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Es ist $n = 2$. Da $\det(A) = -2$, ist auch $\operatorname{rg}(A) = 2$ und $\operatorname{rg}(A, b) = 2$. Daher gibt es eine eindeutige Lösung.

Beispiel 5.9:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}; \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Es ist $n = 3$. Da die Zeilen nicht Vielfache voneinander sind, ist $\operatorname{rg}(A) = 2$ und auch $\operatorname{rg}(A, b) = 2$. Daher gibt es unendlich viele Lösungen.

Beispiel 5.10:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}; \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Es ist $n = 3$. Da die beiden Zeilen von A Vielfache voneinander sind, ist $\operatorname{rg}(A) = 1$. Die beiden Zeilen von (A, b) sind keine Vielfachen voneinander, also ist $\operatorname{rg}(A, b) = 2$. Daher existiert keine Lösung.

Beispiel 5.11:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \\ 4 & 1 \end{pmatrix}; \quad b = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}.$$

Es ist $n = 2$. Die beiden Spalten sind keine Vielfachen voneinander, also ist $\text{rg}(A) = 2$. Welchen Wert hat der Rang von (A, b) ? Da $\text{rg}(A, b) \geq \text{rg}(A)$ und $\text{rg}(A, b) \leq (n + 1)$ ist $\text{rg}(A, b)$ gleich 2 oder 3. $\text{rg}(A, b) = 3$ gilt genau dann, wenn $\det(A, b) \neq 0$. Eine Überprüfung ergibt:

$$\det \begin{pmatrix} 1 & 2 & -1 \\ 3 & 2 & 1 \\ 4 & 1 & 3 \end{pmatrix} = 6 + 8 - 3 - 1 - 18 + 8 = 0$$

Also ist $\text{rg}(A, b) = 2$. Daher gibt es eine eindeutige Lösung.

Bemerkung 5.12: Es dürfen zwei unterschiedliche Situationen nicht verwechselt werden:

- Ein Gleichungssystem $Ax = b$ mit quadratischem A ist genau dann eindeutig lösbar, wenn $\det(A) \neq 0$.
- Ist ein Gleichungssystem $Ax = b$ gegeben, wobei
 - A eine Zeile mehr als Spalten hat (Form: $(n + 1) \times n$).
 - die Spalten von A linear unabhängig sind.

ist das System genau dann eindeutig lösbar, wenn $\det(A, b) = 0$ ist.

Beispiel 5.13:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \\ 4 & 1 \end{pmatrix}; \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Es ist $n = 2$ und $\text{rg}(A) = 2$. Weiterhin gilt: $2 \leq \text{rg}(A, b) \leq 3$. Mit

$$\det(A, b) = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 2 & 0 \\ 4 & 1 & 0 \end{vmatrix} = 1 \begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} = -5 \neq 0$$

ergibt sich $\text{rg}(A, b) = 3$. Daher existiert keine Lösung.

5.1.1 Lösbarkeit von quadratischen Systemen mit $\det=0$

Im Falle singularer Matrizen, d.h. $\det(A) = 0$, ist das inhomogene Gleichungssystem $Ax = b$ nicht für alle b lösbar. Die Existenz einer Lösung kann wie folgt charakterisiert werden.

Satz 5.14: Ist $A \in M(n \times n, K)$ und $\det(A) = 0$, dann ist das inhomogene Gleichungssystem $Ax = b$ genau dann lösbar, wenn für alle Lösungen z des transponierten homogenen Systems $A^T z = 0$ gilt $\langle b, z \rangle = 0$.

BEWEIS Wir nehmen an, $Ax = b$ sei lösbar, x sei die Lösung und z sei Lösung von $A^T z = 0$. Dann folgt

$$\langle b, z \rangle = \langle Ax, z \rangle = (Ax)^T z = (x^T A^T) z = x^T (A^T z) = 0.$$

Also gilt $\langle b, z \rangle = 0$.

Nun sei $\langle b, z \rangle = 0$ für alle Lösungen z des transponierten homogenen Systems $A^T z = 0$. Zu zeigen ist, dass $Ax = b$ dann lösbar ist.

Seien \vec{a}_i die Zeilenvektoren der Matrix A , d.h. die Spaltenvektoren der transponierten Matrix A^T und $z = (z_1, \dots, z_n)^T$. Dann gilt

$$\begin{aligned} & A^T z = 0 \\ \Leftrightarrow & (\vec{a}_1, \dots, \vec{a}_n) \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = 0 \\ \Leftrightarrow & \vec{a}_1 z_1 + \dots + \vec{a}_n z_n = 0 \end{aligned}$$

Für diese z gilt nach Voraussetzung außerdem $\langle b, z \rangle = 0$, also

$$b_1 z_1 + \dots + b_n z_n = 0.$$

Dies kann man mit der vorigen Gleichung zusammenfassen zu:

$$\begin{pmatrix} A^T \\ b^T \end{pmatrix} z = 0$$

Diese Gleichung hat die gleichen Lösungen z wie $A^T z = 0$, somit das gleiche Bild und auch den gleichen Rang. Damit ist

$$\operatorname{rg} \begin{pmatrix} A^T \\ b^T \end{pmatrix} = \operatorname{rg}(A)$$

und da für alle Matrizen $\operatorname{rg}(Z) = \operatorname{rg}(Z^T)$ ist, gilt auch:

$$\operatorname{rg}(A, b) = \operatorname{rg}(A)$$

Dies ist genau die Lösbarkeitsbedingung für die Gleichung $Ax = b$. ■

Beispiel 5.15:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 1 \end{pmatrix} = (\vec{a}_1, \vec{a}_2, \vec{a}_3)$$

Da $\vec{a}_2 = 2 \cdot \vec{a}_1$, ist $\operatorname{rg}(A) = 2$ und $\det(A) = 0$. Wir betrachten zwei unterschiedliche rechte Seiten:

1. $\vec{b}_1 = \begin{pmatrix} 3 \\ 6 \\ 8 \end{pmatrix} = \vec{a}_1 + \vec{a}_2 - \vec{a}_3 \Rightarrow \text{rg}(A, b) = 2 \Rightarrow$ unendliche viele Lösungen.
2. $\vec{b}_2 = \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} \Rightarrow \text{rg}(A, b) = 3 \Rightarrow$ keine Lösung

Gemäß des vorigen Satzes müsste für die Lösungen z des homogenen Systems $A^T z = 0$ gelten: $\langle \vec{b}_1, \vec{z} \rangle = 0$ und $\langle \vec{b}_2, \vec{z} \rangle \neq 0$. Es gilt:

$$A^T z = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = 0$$

$\Rightarrow z_3 = 0 \Rightarrow 2z_1 + 4z_2 = 0 \Rightarrow z_1 = -2z_2$. Mit $z_2 = x$ erhält man als Lösungen:

$$\vec{z} = \begin{pmatrix} -2x \\ x \\ 0 \end{pmatrix}.$$

1. $\langle \vec{b}_1, \vec{z} \rangle = \left\langle \begin{pmatrix} 3 \\ 6 \\ 8 \end{pmatrix}, \begin{pmatrix} -2x \\ x \\ 0 \end{pmatrix} \right\rangle = -6x + 6x = 0$
2. $\langle \vec{b}_2, \vec{z} \rangle = \left\langle \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix}, \begin{pmatrix} -2x \\ x \\ 0 \end{pmatrix} \right\rangle = -6x + 5x = -x$

Beispiel 5.16: $\vec{a} \times \vec{x} = \vec{b}$ bzw. in Matrixschreibweise: $Ax = b$ mit

$$A = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}$$

Sind Lösungen vorhanden? Die Lösungen zu:

$$A^T z = -Az = -\vec{a} \times \vec{z} = 0$$

sind alle $\vec{z} \parallel \vec{a}$. Das Gleichungssystem $Ax = b$ hat nur dann Lösungen, wenn für alle diese \vec{z} gilt: $\langle \vec{b}, \vec{z} \rangle = 0$. Dies ist nur dann der Fall, wenn $\vec{b} \perp \vec{a}$.
 $\Rightarrow \vec{a} \times \vec{x} = \vec{b}$ ist nur dann lösbar, wenn $\vec{b} \perp \vec{a}$.

5.1.2 Äquivalente Bedingungen zur Lösbarkeit eines Gleichungssystems

Es sollen jetzt sämtliche äquivalente Bedingungen für die eindeutige Lösbarkeit eines linearen Gleichungssystems angegeben werden.

Satz 5.17: Für eine Matrix $A \in M(n \times n, \mathbb{R})$ und die dadurch gegebene lineare Abbildung L_A sind folgende Bedingungen äquivalent:

1. A ist invertierbar.
2. $Ax = 0$ hat nur die triviale Lösung $x = 0$.
3. Durch Zeilen- und Spaltenumformungen kann A auf die Einheitsmatrix transformiert werden.
4. A ist darstellbar als Produkt von Elementarmatrizen.
5. Die Gleichung $Ax = b$ besitzt für jedes $b \in \mathbb{R}^n$ mindestens eine Lösung.
6. $Ax = b$ hat genau eine Lösung für jedes $b \in \mathbb{R}^n$.
7. $\det(A) \neq 0$.
8. Das Bild von L_A ist \mathbb{R}^n .
9. L_A ist bijektiv.
10. Die Spaltenvektoren von A sind linear unabhängig.
11. Die Zeilenvektoren von A sind linear unabhängig.
12. Die Spaltenvektoren von A sind eine Basis des \mathbb{R}^n .
13. Die Zeilenvektoren von A sind eine Basis des \mathbb{R}^n .
14. $\text{rg}(A) = n$.
15. $\text{Ker}(L_A) = \{0\}$.
16. Das orthogonale Komplement von $\text{Ker}(L_A)$ ist \mathbb{R}^n .
17. Das orthogonale Komplement des von den Zeilen von A aufgespannten Raums ist $\{0\}$.
18. $A^T A$ ist invertierbar.

5.2 Die Cramersche Regel

Es wird jetzt eine Formel für die Lösung eines linearen Gleichungssystems mit Determinanten angegeben, die sogenannte Cramersche Regel:

Satz 5.18: (Cramersche Regel): Es seien $A \in M(n \times n, K)$ und $\vec{x}, \vec{b} \in K^n$ sowie $A\vec{x} = \vec{b}$ ein lineares Gleichungssystem und es gelte $\det(A) \neq 0$. Man definiert Matrizen A_i dadurch, dass die Matrix A_i mit der Matrix A bis auf die i -te Spalte übereinstimmt, in der der Vektor \vec{b} steht. Es gilt z.B.

$$A_i = (\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \dots, \vec{a}_n).$$

Dann gilt

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

BEWEIS Man definiert Matrizen X_i dadurch, dass sich die Matrix X_i von der Einheitsmatrix E nur in der i -ten Spalte durch den Vektor \vec{x} unterscheidet. Es gilt z.B.

$$X_3 := \begin{pmatrix} 1 & 0 & x_1 & 0 & \dots & 0 \\ 0 & 1 & x_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & x_n & 0 & \dots & 1 \end{pmatrix}.$$

Dann gilt:

$$\begin{aligned} A \cdot X_i &= A \cdot (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{i-1}, \vec{x}, \vec{e}_{i+1}, \dots, \vec{e}_n) \\ &= (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_{i-1}, A\vec{x}, \vec{a}_{i+1}, \dots, \vec{a}_n) \\ &= (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \dots, \vec{a}_n) \\ &= A_i \end{aligned}$$

Aufgrund der Darstellung von X_i erkennt man: $\det(X_i) = x_i$. Daraus folgt:

$$\begin{aligned} \det(A_i) &= \det(A \cdot X_i) = \det(A) \cdot \det(X_i) \\ x_i &= \frac{\det(A_i)}{\det(A)} \end{aligned}$$

■

Beispiel 5.19: Lösung des linearen Gleichungssystems

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 2 \\ x_1 - x_2 + 3x_3 &= -7 \\ 5x_1 + 2x_2 + 4x_3 &= 1 \end{aligned}$$

$$D = \det(A) = \begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 3 \\ 5 & 2 & 4 \end{vmatrix} = -2 \neq 0$$

Damit ist A invertierbar. Weiterhin gilt:

$$D_1 = \begin{vmatrix} 2 & 1 & 1 \\ -7 & -1 & 3 \\ 1 & 2 & 4 \end{vmatrix} = -2, \quad D_2 = \begin{vmatrix} 2 & 2 & 1 \\ 1 & -7 & 3 \\ 5 & 1 & 4 \end{vmatrix} = -4, \quad D_3 = \begin{vmatrix} 2 & 1 & 2 \\ 1 & -1 & -7 \\ 5 & 2 & 1 \end{vmatrix} = 4$$

Lösung: $x_1 = 1$, $x_2 = 2$, $x_3 = -2$

Probe mit Gauß:

$$\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ 1 & -1 & 3 & -7 \\ 5 & 2 & 4 & 1 \\ \hline 1 & -1 & 3 & -7 \\ 2 & 1 & 1 & 2 \\ 5 & 2 & 4 & 1 \\ \hline 1 & -1 & 3 & -7 \\ 0 & 3 & 5 & 16 \\ 0 & 7 & -11 & 36 \\ \hline 1 & -1 & 3 & -7 \\ 0 & 1 & -\frac{5}{3} & -\frac{16}{3} \\ 0 & 0 & \frac{2}{3} & -\frac{4}{3} \end{array}$$

$$\frac{2}{3}x_3 = -\frac{4}{3} \Rightarrow x_3 = -2$$

$$x_2 + \frac{10}{3} = \frac{16}{3} \Rightarrow x_2 = 2$$

$$x_1 - 2 - 6 = -7 \Rightarrow x_1 = 1$$

Beispiel 5.20: Berechnen Sie das Gleichungssystem

$$\begin{array}{cc|c} 3 & -5 & -4 \\ 7 & 8 & 3 \end{array}$$

nach Gauß und nach der Cramerschen Regel.

Gauß:

$$\begin{array}{cc|c} 3 & -5 & -4 \\ 7 & 8 & 3 \\ \hline 3 & -5 & -4 \\ 0 & 59 & 37 \end{array} \quad | \cdot 3 - 7 \cdot \textcircled{1}$$

$$\begin{aligned}
 y &= \frac{37}{59} \\
 x &= \frac{1}{3}(5y - 4) = \frac{1}{3}\left(5 \cdot \frac{37}{59} - 4\right) = \frac{1}{3}\left(\frac{185 - 236}{59}\right) \\
 &= -\frac{1}{3} \cdot \frac{51}{59} = -\frac{17}{59}
 \end{aligned}$$

Cramersche Regel:

$$\begin{aligned}
 A &= \begin{vmatrix} 3 & -5 \\ 7 & 8 \end{vmatrix} = 24 + 35 = 59 \\
 A_1 &= \begin{vmatrix} -4 & -5 \\ 3 & 8 \end{vmatrix} = -32 + 15 = -17 \\
 A_2 &= \begin{vmatrix} 3 & -4 \\ 7 & 3 \end{vmatrix} = 9 + 28 = 37 \\
 x &= \frac{A_1}{A} = -\frac{17}{59}, \\
 y &= \frac{A_2}{A} = \frac{37}{59}
 \end{aligned}$$

5.3 Überbestimmte lineare Gleichungssysteme

5.3.1 Orthogonalprojektion auf Unterräume

In Satz 3.97 haben wir die Projektion eines Vektors auf einen Unterraum mit einer Orthonormalbasis kennengelernt. Nun möchten wir die Projektion auf allgemeine Unterräume des \mathbb{R}^m untersuchen.

Bemerkung 5.21: $p_A(\vec{b})$ sei die Projektion eines Vektors $\vec{b} \in \mathbb{R}^m$ auf den von den Vektoren $A = (\vec{a}_1, \dots, \vec{a}_n) \in M(m \times n, \mathbb{R})$ aufgespannten Unterraum $U = L(\vec{a}_1, \dots, \vec{a}_n)$, also das Bild von A . Damit existiert ein $\vec{x} \in \mathbb{R}^n$ mit

$$p_A(\vec{b}) = \sum_{k=1}^n x_k \cdot \vec{a}_k = A \cdot \vec{x}.$$

Dann gilt $\vec{b} - p_A(\vec{b}) \perp U$ bzw. $\vec{b} - A \cdot \vec{x} \in U^\perp$ und damit

$$\begin{aligned}
 \vec{b} - A \cdot \vec{x} &\perp \vec{a}_k \quad \forall k \\
 \Leftrightarrow \langle \vec{a}_k, \vec{b} - A \cdot \vec{x} \rangle &= 0 \quad \forall k \\
 \Leftrightarrow A^T \cdot (\vec{b} - A \cdot \vec{x}) &= 0 \\
 \Leftrightarrow A^T A \cdot \vec{x} &= A^T \cdot \vec{b}.
 \end{aligned}$$

Die Gleichungen $A^T A \vec{x} = A^T \vec{b}$ heißen **Normalgleichungen**.

Bemerkung 5.22: Die Normalgleichungen sind für jede reelle Matrix $A \in M(m \times n, \mathbb{R})$ lösbar. Im Falle $\text{rg}(A) = n$ existiert mit

$$\vec{x} = (A^T A)^{-1} \cdot A^T \cdot \vec{b}$$

sogar eine eindeutige Lösung. Die Matrix $(A^T A)^{-1} \cdot A^T$ heißt dann auch **verallgemeinerte Inverse** von A .

5.3.2 Methode der kleinsten Quadrate

Satz 5.23: (Methode der kleinsten Quadrate): Gegeben ist das Gleichungssystem

$$A\vec{x} = \vec{b}, \quad A \in M(m \times n, \mathbb{R}), \quad \vec{b} \in \mathbb{R}^m, \quad m \geq n.$$

Da mehr Gleichungen als Unbekannte gegeben sind, ist das Gleichungssystem i.A. nicht lösbar. Sind die Spaltenvektoren von A linear unabhängig, d.h. $\text{rg}(A) = n$, dann ist

$$\vec{x} = (A^T A)^{-1} \cdot A^T \cdot \vec{b}$$

die Näherungslösung nach der Methode der kleinsten Quadrate (LSE = Least Squares Estimator), d.h. die Norm des Fehlervektors $\vec{\varepsilon}$ wird für \vec{x} minimal:

$$\|\vec{\varepsilon}\| = \|\vec{b} - A\vec{x}\| = \min_{\vec{z} \in \mathbb{R}^n} \|\vec{b} - A\vec{z}\|$$

BEWEIS Nach Bemerkung 5.22 ist $\vec{x} = (A^T A)^{-1} \cdot A^T \cdot \vec{b}$ dann die eindeutige Lösung der Normalgleichungen und $p_A(\vec{b}) = A\vec{x} = A \cdot (A^T A)^{-1} \cdot A^T \cdot \vec{b}$ ist die orthogonale Projektion von \vec{b} auf das Bild von A . Mit dem Satz des Pythagoras ergibt sich für beliebige $\vec{z} \in \mathbb{R}^n$:

$$\begin{aligned} \|\vec{b} - A\vec{z}\|^2 &= \|(\vec{b} - p_A(\vec{b})) + \underbrace{(p_A(\vec{b}) - A\vec{z})}_{\perp (\vec{b} - p_A(\vec{b}))}\|^2 \\ &= \|\vec{b} - p_A(\vec{b})\|^2 + \|p_A(\vec{b}) - A\vec{z}\|^2 \\ &\geq \|\vec{b} - p_A(\vec{b})\|^2 \\ &= \|\vec{b} - A\vec{x}\|^2 \end{aligned}$$

Somit liefert $p_A(\vec{b})$ den kleinsten euklidischen Abstand aller Elemente des Bildes von A zu \vec{b} . ■

Beispiel 5.24: Das Gleichungssystem

$$x = 1$$

$$x = 2$$

ist überbestimmt und nicht lösbar. Es gilt:

$$\begin{aligned}x - 1 &= \varepsilon_1 \\x - 2 &= \varepsilon_2\end{aligned}$$

Die Aufgabe ist es, eine Zahl x (bzw. einen Vektor \vec{x} des \mathbb{R}^1) zu finden, für den $\|\vec{\varepsilon}\| = \left\| \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \right\|$ minimal ist. In diesem einfachen Fall kann man x leicht direkt bestimmen:

$$\|\vec{\varepsilon}\|^2 = (x - 1)^2 + (x - 2)^2 = x^2 - 2x + 1 + x^2 - 4x + 4 = 2x^2 - 6x + 5$$

$\|\vec{\varepsilon}\|^2$ ist minimal für $4x - 6 = 0$, d.h. für $x = \frac{3}{2}$.

Nun wird der Wert noch einmal ausgerechnet, diesmal unter Verwendung der verallgemeinerten Inversen. In Matrixform hat die Gleichung das Aussehen

$$A\vec{x} = \vec{b},$$

wobei

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{x} = (x), \quad \vec{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Berechnung der besten Näherung \vec{x} :

$$\begin{aligned}A^T A \vec{x} &= A^T \vec{b} \\(1, 1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} x &= (1, 1) \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\2x &= 3 \\x &= \frac{3}{2}\end{aligned}$$

Beispiel 5.25: Das Gleichungssystem

$$\begin{aligned}u &= 1 \\v &= 2 \\u + v &= 4\end{aligned}$$

ist überbestimmt und nicht lösbar. In Matrixform hat die Gleichung das Aussehen

$$A\vec{x} = \vec{b}$$

wobei

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

$\|\vec{\varepsilon}\|^2$ ist minimal für $\vec{x} = (A^T A)^{-1} A^T \vec{b}$.

$$A^T A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$(A^T A)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$A^T \vec{b} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

$$\vec{x} = (A^T A)^{-1} A^T \vec{b} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 4 \\ 7 \end{pmatrix}$$

Die am besten „passende“ Lösung ist $u = \frac{4}{3}$ und $v = \frac{7}{3}$. Der Fehler ε ist:

$$\vec{\varepsilon} = A\vec{x} - \vec{b} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 4 \\ 7 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

$$= \frac{1}{3} \left[\begin{pmatrix} 4 \\ 7 \\ 11 \end{pmatrix} - \begin{pmatrix} 3 \\ 6 \\ 12 \end{pmatrix} \right]$$

$$= \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

$$\|\vec{\varepsilon}\|^2 = \frac{1}{3}$$

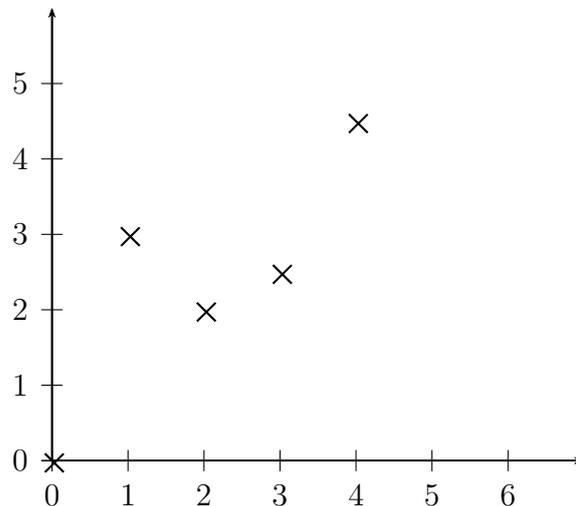
5.3.3 Ausgleichsrechnung mit der Methode der kleinsten Quadrate

Im Rahmen einer **Ausgleichsrechnung** oder **Regression** sind mehrere Datenpunkte gegeben, die physikalische Messwerte, wirtschaftliche Größen oder Ähnliches repräsentieren können. Gesucht sind dabei die Parameter eines als bekannt angenommenen parametrisierbaren funktionalen Zusammenhanges. Dazu verwendet man die vermutete (z.B. physikalische) Beziehung, die der Messung zugrunde liegt. Die Aufgabe ist es, die Parameter

der Funktion zu finden, die möglichst genau durch die Punktwolke passt (Fitting). Unter „passen“ versteht man, dass die Summe der quadratischen Abweichungen der Kurve von den beobachteten Punkten minimiert wird. Beschränkt man sich auf die Funktionenmenge der Polynome, so muss man zunächst festlegen, welchen Grad n das Polynom haben soll. Wählt man $n = 1$, so erhält man eine Gerade und spricht von einer **linearen Regression**.

Beispiel 5.26: Diese Punkte können beispielsweise so aussehen:

$$(0; 0), (1; 3), (2; 2), (3; 2.5), (4; 4.5)$$



In unserem Beispiel verwenden wir zunächst ein Polynom vom Grad $n = 1$, d.h., wir versuchen eine Gerade $y = \alpha \cdot x + \beta$ durch die Punkte zu legen. Das führt zu folgendem Gleichungssystem:

$$\alpha \cdot 0 + \beta = 0$$

$$\alpha \cdot 1 + \beta = 3$$

$$\alpha \cdot 2 + \beta = 2$$

$$\alpha \cdot 3 + \beta = 2.5$$

$$\alpha \cdot 4 + \beta = 4.5$$

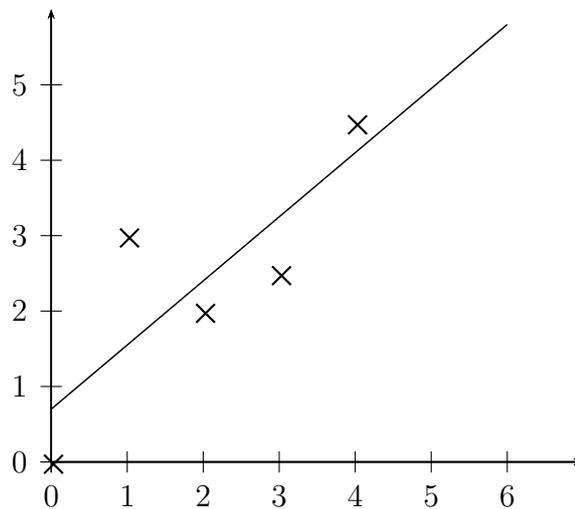
oder in Matrixform:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \\ 2.5 \\ 4.5 \end{pmatrix}$$

α und β ergeben sich aus $(\alpha, \beta)^T = (A^T A)^{-1} A^T \vec{b}$.

$$\begin{aligned} A^T A &= \begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix} \\ (A^T A)^{-1} &= \frac{1}{50} \begin{pmatrix} 5 & -10 \\ -10 & 30 \end{pmatrix} \\ A^T \vec{b} &= \begin{pmatrix} 32,5 \\ 12 \end{pmatrix} \\ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= \frac{1}{50} \begin{pmatrix} 42,5 \\ 35 \end{pmatrix} = \begin{pmatrix} 0,85 \\ 0,7 \end{pmatrix} \end{aligned}$$

Also ist die gesuchte Kurve $y = 0,85 \cdot x + 0,7$.



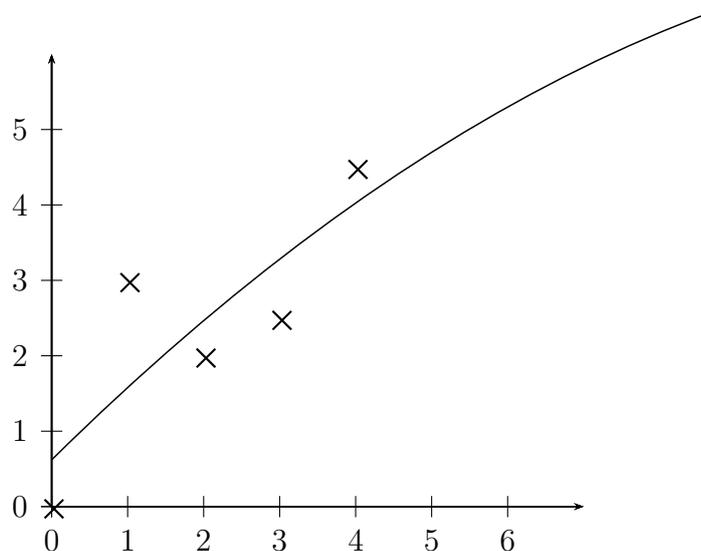
Ein quadratischer Fit ergibt das Gleichungssystem:

$$\begin{aligned}
 \alpha \cdot 0^2 + \beta \cdot 0 + \gamma &= 0 \\
 \alpha \cdot 1^2 + \beta \cdot 1 + \gamma &= 3 \\
 \alpha \cdot 2^2 + \beta \cdot 2 + \gamma &= 2 \\
 \alpha \cdot 3^2 + \beta \cdot 3 + \gamma &= 2.5 \\
 \alpha \cdot 4^2 + \beta \cdot 4 + \gamma &= 4.5
 \end{aligned}$$

In Matrixform:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \\ 2.5 \\ 4.5 \end{pmatrix}$$

was (nach erheblich längerer Rechnung) das Ergebnis $y = -0.0357 \cdot x^2 + 0.9929 \cdot x + 0.6286$ liefert.



Bemerkung 5.27: Gegeben seien die Datenpunkte (x_k, y_k) , $k = 1, \dots, m$. Das Ausgleichspolynom soll den Grad $n = 1$ haben, also eine Gerade sein. Dann ist

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Die Gleichung $A^T A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = A^T \vec{b}$

ergibt ausmultipliziert

$$\begin{pmatrix} \sum_{k=1}^m x_k^2 & \sum_{k=1}^m x_k \\ \sum_{k=1}^m x_k & m \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m x_k y_k \\ \sum_{k=1}^m y_k \end{pmatrix}$$

was genau das Ergebnis aus der Analysis-Vorlesung ist, in der wir über die Extremwerte der Summe der Abweichungsquadrate argumentiert hatten.

5.4 Unterbestimmte lineare Gleichungssysteme

Nun betrachten wir ein ähnliches Ergebnis für unterbestimmte Gleichungssysteme.

Satz 5.28: Gegeben ist ein lineares Gleichungssystem

$$A \vec{x} = \vec{b}, \quad A \in M(m \times n, \mathbb{R}), \quad \vec{x} \in \mathbb{R}^n, \quad \vec{b} \in \mathbb{R}^m, \quad m < n.$$

Dieses Gleichungssystem hat immer Lösungen, wenn $\text{rg}(A) = m$ ist, da dann auch gilt $\text{rg}(A) = \text{rg}(A, \vec{b}) = m$. In diesem Fall hat der Lösungsvektor

$$\vec{x}_s = A^T (A \cdot A^T)^{-1} \vec{b}$$

des Gleichungssystems $A \vec{x} = \vec{b}$ die Eigenschaft, dass er:

- die Lösung mit der kleinsten Norm von allen Lösungen \vec{x} ist, d.h.

$$\|\vec{x}_s\| = \min_{A \vec{x} = \vec{b}} \{\|\vec{x}\|\}.$$

- orthogonal zu sämtlichen Lösungen \vec{z} des zugehörigen homogenen Gleichungssystem $A \vec{z} = 0$ ist, d.h.

$$\langle \vec{x}_s, \vec{z} \rangle = 0$$

In diesem Fall heißt die Matrix

$$A^T (A \cdot A^T)^{-1}$$

die **verallgemeinerte Inverse** der Matrix A .

BEWEIS \vec{x}_s ist definiert, wenn $(A \cdot A^T)$ invertierbar ist. Da $\text{rg}(A) = m$, sind die Zeilen von A linear unabhängig. Also ist $A \cdot A^T$ nach Kapitel 5.1.2 auch invertierbar (siehe auch Bemerkung 6.40).

\vec{x}_s ist eine Lösung von $A \vec{x} = \vec{b}$, denn

$$A \vec{x}_s = A (A^T (A \cdot A^T)^{-1} \vec{b}) = (A \cdot A^T) (A \cdot A^T)^{-1} \vec{b} = \vec{b}.$$

Allgemein gilt: allgemeine Lösung des inhomogenen Gleichungssystems = spezielle Lösung des inhomogenen Gleichungssystems + allgemeine Lösung des homogenen Gleichungssystems. Die Lösungen des homogenen Gleichungssystems

$$A\vec{z} = 0$$

sind allgemein

$$\vec{z} = \lambda_1 \vec{z}_1 + \dots + \lambda_r \vec{z}_r$$

\vec{x}_s liegt senkrecht auf den Lösungen \vec{z} , denn

$$\langle \vec{z}, \vec{x} \rangle = \vec{z}^T A^T (A \cdot A^T)^{-1} \vec{b} = \overbrace{\langle A\vec{z}, (A \cdot A^T)^{-1} \vec{b} \rangle}^{=0} = 0$$

Man kann $\vec{z}_1, \dots, \vec{z}_r$ so wählen, dass sie paarweise orthogonal sind. Mit der speziellen Lösung \vec{x}_s lässt sich die allgemeine Lösung schreiben als:

$$\vec{x} = \vec{x}_s + \mu_1 \vec{z}_1 + \dots + \mu_r \vec{z}_r, \mu_i \in \mathbb{R}.$$

Wegen der Orthogonalität gilt für jede Lösung \vec{x} nach Pythagoras:

$$\|\vec{x}\|^2 = \|\vec{x}_s\|^2 + \|\mu_1 \vec{z}_1 + \dots + \mu_r \vec{z}_r\|^2 \geq \|\vec{x}_s\|^2.$$

Damit ist \vec{x}_s der eindeutig bestimmte Lösungsvektor von $A\vec{x} = \vec{b}$ mit der kleinsten Norm.

■

Beispiel 5.29: Das lineare Gleichungssystem

$$x + 2y - 2z = 5$$

ist unterbestimmt. Die Lösungen bilden eine Ebene im \mathbb{R}^3 . Der gesuchte minimale Lösungsvektor \vec{x} ist der Vektor, der vom Nullpunkt ausgeht und senkrecht auf die Ebene trifft. Aus der Hesseschen Normalform der Ebene

$$\frac{1}{3}(x + 2y - 2z) = \frac{5}{3}$$

erhält man direkt $\vec{x} = \frac{5}{9} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$. Dieses Ergebnis erhält man auch durch die

Berechnung von $\vec{x} = A^T(AA^T)^{-1}\vec{b}$ mit

$$A = (1, 2, -2), \quad \vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \vec{b} = (5)$$

Es gilt:

$$\begin{aligned}\vec{x} &= A^T(AA^T)^{-1}\vec{b} \\ &= \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} ((1, 2, -2) \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix})^{-1} \cdot 5 \\ &= \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} (9)^{-1} \cdot 5 \\ &= \frac{5}{9} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}\end{aligned}$$

Kapitel 6

Spezielle Matrizen

6.1 Orthogonale Matrizen

Definition 6.1: Eine Matrix $A \in M(n \times n, \mathbb{R})$ heißt **orthogonal** bzw. **orthonormal**, falls gilt $A^T A = E$.

Die Menge der orthogonalen Matrizen bezeichnet man mit $O(n)$. Es gilt $O(n) \subset M(n \times n, \mathbb{R})$.

6.1.1 Eigenschaften orthogonaler Matrizen

Bemerkung 6.2: Folgende Aussagen sind äquivalent für eine Matrix $A \in M(n \times n, \mathbb{R})$:

- (a) $A \in O(n)$
- (b) A ist invertierbar und es gilt $A^{-1} = A^T$.
- (c) $A \cdot A^T = E$.
- (d) Die Spaltenvektoren von A bilden ein Orthonormalsystem bzgl. des Standardskalarprodukts im \mathbb{R}^n .
- (e) Die Zeilenvektoren von A bilden ein Orthonormalsystem bzgl. des Standardskalarprodukts im \mathbb{R}^n .

BEWEIS (a) \Leftrightarrow (b): Die Aussage (a) bedeutet nach Definition der inversen Matrix, dass $A^T = A^{-1}$ ist, also die Äquivalenz von (a) und (b).

(a) \Leftrightarrow (c): Da bekanntlich ein linksinverses Element in einer Gruppe gleich dem rechtsinversen ist, folgt aus $A^T \cdot A = E$ auch $A \cdot A^T = E$, also die Äquivalenz von (a) und (c).

(a) \Leftrightarrow (d): Die Aussage (a) bedeutet, dass gilt $A^T \cdot A = E$. Schreibt man dieses Matrix-

produkt explizit hin, dann sieht man, dass (a) äquivalent zu (d) ist:

$$\begin{aligned}
 A &= (\vec{a}_1, \dots, \vec{a}_n), \quad \vec{a}_1, \dots, \vec{a}_n : \text{Spaltenvektoren von } A. \\
 \Rightarrow A^T &= \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} \\
 \Rightarrow A^T \cdot A &= \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} \cdot (\vec{a}_1, \dots, \vec{a}_n) \\
 &= \begin{pmatrix} \langle \vec{a}_1, \vec{a}_1 \rangle & \cdots & \langle \vec{a}_1, \vec{a}_n \rangle \\ \vdots & & \vdots \\ \langle \vec{a}_n, \vec{a}_1 \rangle & \cdots & \langle \vec{a}_n, \vec{a}_n \rangle \end{pmatrix} \\
 &= E \\
 \Leftrightarrow \langle \vec{a}_i, \vec{a}_i \rangle &= 1 \\
 \langle \vec{a}_i, \vec{a}_j \rangle &= 0, \quad i \neq j
 \end{aligned}$$

(a) \Leftrightarrow (e): Beweis ähnlich wie (d) ■

Bemerkung 6.3: Für alle $A \in O(n)$ gilt $|\det(A)| = 1$.

BEWEIS Mit den Eigenschaften der Determinanten gilt

$$|\det(A \cdot A^T)| = |\det(A) \cdot \det(A^T)| = |\det(A)| \cdot |\det(A^T)| = |\det(A)| \cdot |\det(A)|$$

und

$$|\det(A \cdot A^T)| = |\det(E)| = 1.$$

Also gilt insgesamt $|\det(A)|^2 = 1$ und damit gilt die Bemerkung. ■

Die Umkehrung der letzten Bemerkung gilt nicht, denn für die Matrix

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

gilt $|\det(E)| = 1$. A ist offensichtlich keine Orthogonalmatrix.

In Beispiel 4.91 hatten wir Dreh-, Spiegelungs- und Verzerrungsmatrizen im \mathbb{R}^2 kennen gelernt. Diese untersuchen wir nun auf Orthogonalität:

Beispiel 6.4: (Drehmatrix im \mathbb{R}^2):

$$A = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

ist eine Drehmatrix im \mathbb{R}^2 . $\vec{y} = A \cdot \vec{x}$ dreht den Punkt \vec{x} um den Winkel ϕ um den Ursprung.

Das wird deutlich, wenn man die Gleichung $\vec{y} = A \cdot \vec{x}$ in Polarkoordinaten schreibt, also $\vec{x} = r_x \begin{pmatrix} \cos \lambda \\ \sin \lambda \end{pmatrix}$ und $\vec{y} = r_y \begin{pmatrix} \cos \mu \\ \sin \mu \end{pmatrix}$. Dann ist:

$$\begin{aligned} r_y \begin{pmatrix} \cos \mu \\ \sin \mu \end{pmatrix} &= \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \cdot r_x \begin{pmatrix} \cos \lambda \\ \sin \lambda \end{pmatrix} \\ &= r_x \begin{pmatrix} \cos \lambda \cos \phi - \sin \lambda \sin \phi \\ \cos \lambda \sin \phi + \sin \lambda \cos \phi \end{pmatrix} \\ &= r_x \begin{pmatrix} \cos(\phi + \lambda) \\ \sin(\phi + \lambda) \end{pmatrix} \end{aligned}$$

Die Drehmatrix $A = (\vec{v}_1, \vec{v}_2)$ ist orthogonal, denn ihre Spalten bilden ein Orthonormalsystem.

$$\begin{aligned} \|\vec{v}_1\| &= \cos^2 \phi + \sin^2 \phi = 1 \\ \|\vec{v}_2\| &= \sin^2 \phi + \cos^2 \phi = 1 \\ \langle \vec{v}_1, \vec{v}_2 \rangle &= -\cos \phi \sin \phi + \sin \phi \cos \phi = 0 \end{aligned}$$

Bemerkung 6.5: Spiegelungsmatrizen sind ebenfalls Orthogonalmatrizen. Matrizen zu Verzerrungen hingegen nicht.

Nun untersuchen wir die Umkehrung:

Bemerkung 6.6: Ist die Transformationsmatrix $T_G^{G'}$ von Basis G' in Basis G eine Orthogonalmatrix, dann kann die Transformation aus einer Drehung und/oder einer Spiegelung, nicht aber aus einer Verzerrung, bestehen. Das ergibt sich aus der Eigenschaft $\langle T_G^{G'} \cdot x_i, T_G^{G'} \cdot x_j \rangle = \langle x_i, x_j \rangle$ der Orthogonalmatrizen. Da Orthogonalmatrizen außerdem die Eigenschaft $(T_G^{G'})^{-1} = (T_G^{G'})^T$ besitzen, gilt nach Satz 4.89 für Abbildungsmatrizen einer Abbildung Φ bzgl. der Basen G' und H' bzw. G und H

$$M_{H'}^{G'}(\Phi) = (T_H^{H'})^T \cdot M_H^G(\Phi) \cdot T_G^{G'}.$$

Beispiel 6.7: In Beispiel 4.92 können wir somit die Invertierung der Matrix $T_B^{B'}$ durch deren Transponierung ersetzen.

Bemerkung 6.8: Für $A, B \in O(n)$ gilt $AB \in O(n)$.

BEWEIS

$$(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T$$

■

Bemerkung 6.9: $O(n)$ ist eine Gruppe bzgl. der Matrixmultiplikation.

BEWEIS Die Abgeschlossenheit ergibt sich aus der vorigen Bemerkung. Die Assoziativität ergibt sich allgemein aus der Matrixmultiplikation. Das neutrale Element E ist $\in O(n)$, ebenso wie das inverse Element $A^{-1} = A^T$.

■

Bemerkung 6.10:

- (a) $A \in O(n) \Leftrightarrow \langle Ax, Ay \rangle = \langle x, y \rangle \quad \forall x, y \in \mathbb{R}^n$.
 (b) $A \in O(n) \Leftrightarrow \|Ax\| = \|x\| \quad \forall x \in \mathbb{R}^n$

BEWEIS

$$\langle Ax, Ay \rangle = (Ax)^T Ay = x^T A^T Ay = x^T y = \langle x, y \rangle$$

■

Bemerkung 6.11: A ist eine Orthogonalmatrix genau dann, wenn sie jedes Orthonormalsystem in ein anderes transformiert.

BEWEIS Die Bedingungen für ein Orthonormalsystem waren:

- a) Sämtliche Vektoren x_k sind paarweise orthogonal. $\langle x_i, x_j \rangle = 0, \quad i \neq j$.
 b) $\|x_i\| = 1$

Es gilt: $\langle Ax_i, Ax_j \rangle = \langle x_i, x_j \rangle = 0$ und $\|Ax_i\| = \|x_i\| = 1$

■

6.2 Die QR-Zerlegung

Die QR-Zerlegung wird zur Eigenwertberechnung sehr großer Matrizen verwendet (Eigenwerte werden im nächsten Kapitel behandelt). Das komplette Verfahren wird in Numerik 1 behandelt. In diesem Kapitel wird:

- erläutert, was unter der QR-Zerlegung zu verstehen ist sowie
- kurz beschrieben, wie man eine Matrix unter Zuhilfenahme des Gram-Schmidtschen Orthogonalisierungsverfahrens QR-zerlegt. In der Praxis verwendet man statt dessen jedoch die besseren (numerisch stabileren) Verfahren von Givens oder Householder, die ebenfalls in Numerik 1b erläutert werden.

Bemerkung 6.12: Ist A eine $(n \times m)$ - Matrix mit linear unabhängigen Spaltenvektoren (d.h. $m \leq n$), dann gibt es eine $(n \times m)$ - Matrix Q mit orthonormalen Spalten (d.h. $Q^T \cdot Q = E_m$) und eine $(m \times m)$ - Matrix R , die eine invertierbare obere Dreiecksmatrix ist, so dass gilt

$$A = QR,$$

die sogenannte **QR-Zerlegung** der Matrix A . E_m ist die Einheitsmatrix in $M(m \times m, \mathbb{R})$.

Die Matrix Q ist im Falle $m < n$ *keine* Orthogonalmatrix, da sie nicht quadratisch ist. Es gilt nur $Q^T \cdot Q = E_m$, aber $Q \cdot Q^T \neq E_n$!

Erläuterung: Es ist die Gleichung $A = Q \cdot R$ zu lösen. A ist gegeben. Q und R sind gesucht. Nimmt man den Vektorraum $V = \mathbb{R}^n$, dann kann man A, Q und R wie folgt darstellen:

- $A = (\vec{a}_1, \dots, \vec{a}_m)$, Die Vektoren \vec{a}_k sind die Spalten von A .
- $Q = (\vec{q}_1, \dots, \vec{q}_m)$, Die Vektoren \vec{q}_k sind die Spalten von Q und damit paarweise orthonormal.

$$\bullet R = \begin{pmatrix} \lambda_{11} & \dots & \dots & \lambda_{1m} \\ 0 & \lambda_{22} & \dots & \lambda_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{mm} \end{pmatrix}$$

Diese Gleichungen kann man mit dem Orthogonalisierungsverfahren nach Gram-Schmidt lösen. Das wird am folgenden Beispiel gezeigt.

Beispiel 6.13: Die Matrix

$$A = (\vec{a}_1, \vec{a}_2) = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 2 & 2 \end{pmatrix}$$

soll aus dem Orthogonalisierungsverfahren nach Gram-Schmidt QR-zerlegt werden. $A = QR$ ergibt ausgeschrieben:

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 2 & 2 \end{pmatrix} = (\vec{q}_1, \vec{q}_2) \cdot \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ 0 & \lambda_{22} \end{pmatrix}$$

In Gleichungsform ist das:

$$\begin{aligned} \vec{a}_1 &= \lambda_{11} \vec{q}_1 \\ \vec{a}_2 &= \lambda_{12} \vec{q}_1 + \lambda_{22} \vec{q}_2 \end{aligned}$$

Die orthogonalen Vektoren \vec{q}_1 und \vec{q}_2 können mit Gram-Schmidt erzeugt werden. Die λ -Werte erhält man aus den Zwischenergebnissen. Aus Gram-Schmidt ist bekannt:

$$\begin{aligned}\vec{q}_1 &= \frac{\vec{a}_1}{\|\vec{a}_1\|}; & \vec{r}_2 &= \vec{a}_2 - \langle \vec{a}_2, \vec{q}_1 \rangle \vec{q}_1; & \vec{q}_2 &= \frac{\vec{r}_2}{\|\vec{r}_2\|} \\ \Rightarrow \vec{a}_1 &= \vec{q}_1 \|\vec{a}_1\| \\ \vec{a}_2 &= \vec{r}_2 + \langle \vec{a}_2, \vec{q}_1 \rangle \vec{q}_1 = \|\vec{r}_2\| \vec{q}_2 + \langle \vec{a}_2, \vec{q}_1 \rangle \vec{q}_1 \\ \Rightarrow \lambda_{11} &= \|\vec{a}_1\| \\ \lambda_{12} &= \langle \vec{a}_2, \vec{q}_1 \rangle \\ \lambda_{22} &= \|\vec{r}_2\|\end{aligned}$$

Es gilt: $\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ und $\vec{a}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ Daraus errechnet sich

$$\lambda_{11} = \|\vec{a}_1\| = 3$$

$$\vec{q}_1 = \frac{\vec{a}_1}{\|\vec{a}_1\|} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

$$\lambda_{12} = \langle \vec{a}_2, \vec{q}_1 \rangle = \frac{1}{3} \cdot 3 = 1$$

$$\begin{aligned}\vec{r}_2 &= \vec{a}_2 - \langle \vec{a}_2, \vec{q}_1 \rangle \vec{q}_1 \\ &= \vec{a}_2 - 1 \cdot \vec{q}_1 = \frac{1}{3} \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}\end{aligned}$$

$$\lambda_{22} = \|\vec{r}_2\| = \frac{\sqrt{45}}{3} = \sqrt{5}$$

$$\vec{q}_2 = \frac{\vec{r}_2}{\|\vec{r}_2\|} = \frac{1}{\sqrt{45}} \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}$$

Damit ist

$$Q = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{45}} \\ \frac{2}{3} & \frac{-5}{\sqrt{45}} \\ \frac{2}{3} & \frac{4}{\sqrt{45}} \end{pmatrix}$$

$$R = \begin{pmatrix} 3 & 1 \\ 0 & \sqrt{5} \end{pmatrix}$$

Also ist die Matrix A QR-zerlegt.

6.3 Eigenwerte von Matrizen

Wir betrachten jetzt ausschließlich Abbildungen von einem Vektorraum in denselben Vektorraum: $f : V \rightarrow V$. Diese Abbildungen heißen **Endomorphismen**. Die zugehörigen Abbildungsmatrizen sind quadratisch. In diesem Kapitel geht es darum, zu einem gegebenen Endomorphismus eine Basis (d.h. ein Koordinatensystem) zu finden, in dem diese Abbildung durch eine Diagonalmatrix dargestellt werden kann:

$$y = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \cdot x$$

In diesem Fall gilt auch

$$A \cdot e_i = \lambda_i \cdot e_i$$

Ein Vektor auf einer Koordinatenachse bleibt also nach der Transformation auf der Koordinatenachse (möglicherweise an einer anderen Stelle dort). Um die Basisvektoren des neuen Systems zu bestimmen, suchen wir (im alten System) nach den folgenden Vektoren:

Definition 6.14: Sei eine Matrix $A \in M(n \times n, K)$ gegeben, dann heißen alle Vektoren x , für die gilt

$$A \cdot x = \lambda \cdot x,$$

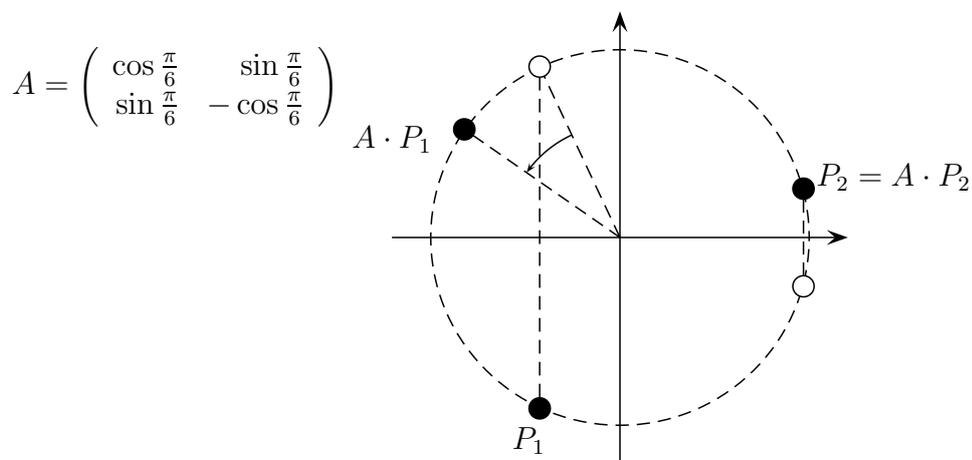
Eigenvektor der Matrix A . Die zugehörigen $\lambda \in \mathbb{K}$ heißen **Eigenwert** oder **charakteristische Zahl** von A zum Eigenvektor x .

Beispiel 6.15:

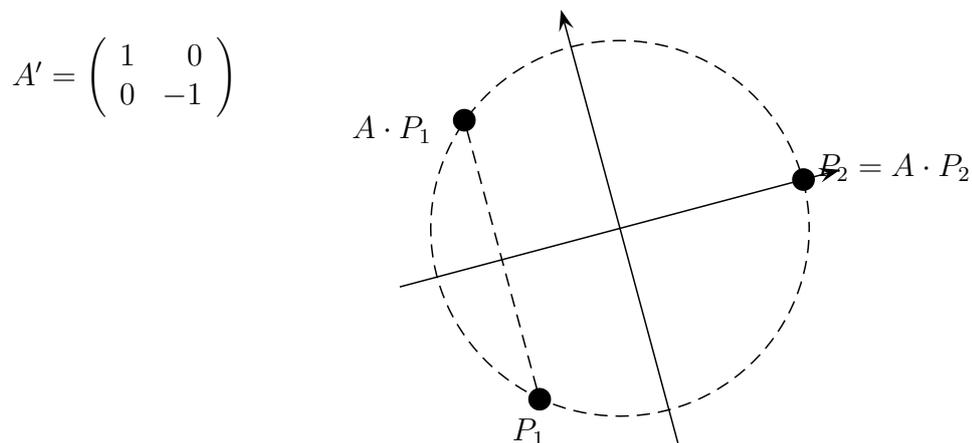
$$A = \begin{pmatrix} \cos \frac{\pi}{6} & \sin \frac{\pi}{6} \\ \sin \frac{\pi}{6} & -\cos \frac{\pi}{6} \end{pmatrix} = \begin{pmatrix} \cos \frac{\pi}{6} & -\sin \frac{\pi}{6} \\ \sin \frac{\pi}{6} & \cos \frac{\pi}{6} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

spiegelt einen Punkt erst an der x -Achse und dreht ihn anschließend um $30^\circ (\frac{\pi}{6})$. Die Abbildungsvorschrift vereinfacht sich wesentlich, wenn man sie in einem neuen Koordinatensystem betrachtet, das gegenüber dem alten um 15° gedreht ist:

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



Konstruktion von $A \cdot P$. Der Punkt P wird erst an der x -Achse gespiegelt und dann um 30° gegen den Uhrzeigersinn gedreht.



Die Abbildungsvorschrift vereinfacht sich wesentlich, wenn man das Koordinatensystem um 15° dreht. Der Punkt P wird nur an der x -Achse gespiegelt. Die neuen Koordinatenachsen sind die Eigenvektoren. Die Eigenwerte sind die Werte auf der Hauptdiagonalen von A .

Gibt es zu einer Matrix $A \in M(n \times n, \mathbb{K})$ eine Zahl λ und einen Vektor $x \neq 0$ mit

$$Ax = \lambda x?$$

Diese Gleichung kann in der Form

$$Ax - \lambda x = 0 \quad \text{bzw.} \quad (\lambda E - A)x = 0 \quad \text{bzw.} \quad (A - \lambda E)x = 0$$

geschrieben werden. Dies ist ein homogenes Gleichungssystem mit der Matrix

$$\lambda E - A.$$

Definition 6.16: Das Polynom

$$p(\lambda) = \det(\lambda E - A)$$

heißt **charakteristische Polynom** einer Matrix $A \in M(n \times n, \mathbb{K})$.

Somit besitzt das lineare homogene Gleichungssystem $(\lambda E - A)x = 0$ genau dann eine nichttriviale Lösung \vec{x} , wenn gilt

$$p(\lambda) = 0.$$

Das charakteristische Polynom $p(\lambda)$ ist ein Polynom n -ten Grades und hat deshalb n Nullstellen λ_k in \mathbb{C} . Es müssen also nicht in jedem Fall reelle Lösungen existieren. Zu jedem Eigenwert λ_k gehört mindestens ein Eigenvektor \vec{x}_k .

6.3.1 Berechnung der Eigenwerte und Eigenvektoren

Ausgangspunkt: Eine Gleichung $Ax = \lambda x$. A ist bekannt, λ und x sind unbekannt. A ist eine $n \times n$ -Matrix.

1. Lösung des charakteristischen Polynoms

$$|\lambda E - A| = 0 \text{ oder } |A - \lambda E| = 0$$

Die Lösungen des Polynoms $\lambda_1, \lambda_2, \dots, \lambda_n$ sind die Eigenwerte. Es kann mehrfache Eigenwerte geben. Eigenwerte können auch komplex sein.

2. Einsetzen der Eigenwerte λ_i in das homogene Gleichungssystem

$$(\lambda_i E - A)x = 0$$

Für jeden Eigenwert findet man unendlich viele Lösungen (sonst hat man sich verrechnet). Es gibt mindestens einen freien Parameter, es sind aber auch mehrere möglich.

3. Findet man einen freien Parameter, z.B. $\vec{x} = \lambda(1, 1, 0)^T$, so bezeichnet man $(1, 1, 0)^T$ als (speziellen) Eigenvektor zum entsprechenden Eigenwert. Der Zusatz „speziell“ wird oft weggelassen. Auch alle Vielfachen dieses Vektors sind Eigenvektoren. Gibt es zwei freie Parameter, z.B. $\vec{x} = \lambda(1, 0, 0)^T + \mu(1, 1, 1)^T$, so erhält man zu dem Eigenwert zwei linear unabhängige (spezielle) Eigenvektoren. Dies bedeutet, dass alle Linearkombinationen davon ebenfalls Eigenvektoren sind. Dies gilt äquivalent auch für drei und mehr Eigenvektoren.

Beispiel 6.17:

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$$

$$\begin{aligned} p(\lambda) = \det(\lambda E - A) &= \begin{vmatrix} \lambda - 1 & 3 & -3 \\ -3 & \lambda + 5 & -3 \\ -6 & 6 & \lambda - 4 \end{vmatrix} \\ &= \begin{vmatrix} \lambda - 1 & 0 & -3 \\ -3 & \lambda + 2 & -3 \\ -6 & \lambda + 2 & \lambda - 4 \end{vmatrix} \\ &= \begin{vmatrix} \lambda - 1 & 0 & -3 \\ -3 & \lambda + 2 & -3 \\ -3 & 0 & \lambda - 1 \end{vmatrix} \\ &= (\lambda + 2)[(\lambda - 1)^2 - 9] \end{aligned}$$

Aus dem Faktor $(\lambda + 2)$ ergibt sich sofort $\lambda_1 = -2$. Die beiden anderen Nullstellen ergeben sich aus $(\lambda - 1)^2 - 9 = \lambda^2 - 2\lambda - 8$. Dies führt zu $\lambda_2 = -2$ und $\lambda_3 = 4$. Ein Eigenwert ist also doppelt.

Berechnung der Eigenvektoren: Einsetzen der Eigenwerte und Lösung des homogenen Systems.

- $\lambda = -2$:

$$\begin{array}{r|l} -3 & 3 & -3 & 0 \\ -3 & 3 & -3 & 0 \\ -6 & 6 & -6 & 0 \\ \hline -1 & 1 & -1 & 0 \end{array}$$

Es gilt also $z = y - x$. Die Lösung hat zwei Freiheitsgrade und ist z.B.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Damit erhält man zu $\lambda = -2$ die Eigenvektoren $(1, 0, -1)^T$ und $(0, 1, 1)^T$. Es ist auch möglich, dass man andere Eigenvektoren erhält. Sie müssen allerdings linear unabhängig und Linearkombinationen der beiden oben genannten Vektoren sein.

- $\lambda = 4$: Analoges Einsetzen ergibt

$$\begin{array}{ccc|c} 3 & 3 & -3 & 0 \\ -3 & 9 & -3 & 0 \\ -6 & 6 & 0 & 0 \\ \hline 3 & 3 & -3 & 0 \\ 0 & 12 & -6 & 0 \\ \hline 1 & 1 & -1 & 0 \\ 0 & 2 & -1 & 0 \end{array}$$

eine Lösung mit einem Freiheitsgrad und einen (speziellen) Eigenvektor von $(1, 1, 2)^T$ (oder ein Vielfaches davon).

Beispiel 6.18: Es sei

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix}$$

Das charakteristische Polynom lautet:

$$\begin{aligned} p(\lambda) &= \det(\lambda E - A) = \det \begin{pmatrix} \lambda - 3 & 0 & 0 \\ 0 & \lambda - 2 & 5 \\ 0 & -1 & \lambda + 2 \end{pmatrix} \\ &= (\lambda - 3)[(\lambda - 2)(\lambda + 2) + 5] = (\lambda - 3)(\lambda^2 + 1) \end{aligned}$$

Die Eigenwerte lauten also $\lambda_1 = 3$, $\lambda_2 = i$, $\lambda_3 = -i$. Eigenwerte können auch komplex sein. Es kann im Reellen kein Koordinatensystem gefunden werden, in dem A durch eine Diagonalmatrix dargestellt wird. Die (komplexen) Eigenvektoren werden wie im vorigen Beispiel durch Lösung des Gleichungssystems $(A - \lambda E)x = 0$ nach dem Gauß-Schema berechnet. Dies wird am Beispiel $\lambda_2 = i$ gezeigt:

$$\begin{array}{ccc|c} i-3 & 0 & 0 & 0 \\ 0 & i-2 & 5 & 0 \\ 0 & -1 & i+2 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & -1 & i+2 & 0 \\ 0 & i-2 & 5 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & -1 & i+2 & 0 \\ 0 & 0 & 0^* & 0 \end{array}$$

$$* \quad 5 + (i+2)(i-2) = 5 + (i^2 - 2^2) = 5 - 1 - 4 = 0$$

Wähle $z = \alpha \Rightarrow y = z(i+2); \quad x = 0; \quad \vec{v}_2 = \alpha \begin{pmatrix} 0 \\ i+2 \\ 1 \end{pmatrix}$

Bemerkung 6.19: Es gilt für das charakteristische Polynom

$$\begin{aligned}
 p(\lambda) &= \det(\lambda E - A) \\
 &= \det(\lambda \vec{e}_1 - \vec{a}_1, \lambda \vec{e}_2 - \vec{a}_2, \dots, \lambda \vec{e}_n - \vec{a}_n) \\
 &= \det(\lambda \vec{e}_1, \lambda \vec{e}_2 - \vec{a}_2, \dots, \lambda \vec{e}_n - \vec{a}_n) + \det(-\vec{a}_1, \lambda \vec{e}_2 - \vec{a}_2, \dots, \lambda \vec{e}_n - \vec{a}_n) \\
 &= \dots = \det(\lambda \vec{e}_1, \dots, \lambda \vec{e}_n) + \dots + \det(-\vec{a}_1, \dots, -\vec{a}_n) \\
 &= \lambda^n + \dots + (-1)^n \cdot \det(A)
 \end{aligned}$$

Andererseits gilt bei entsprechender Wahl des Vorzeichens auch:

$$\begin{aligned}
 p(\lambda) &= (\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_n) \\
 &= \lambda^n + \dots + (-1)^n \prod_{i=1}^n \lambda_i
 \end{aligned}$$

Zusammen erhält man:

$$\det(A) = \prod_{i=1}^n \lambda_i$$

Beispiel 6.20: Im vorigen Beispiel hatte die Matrix

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$$

die Eigenwerte -2 (doppelt) und 4. Daraus ergibt sich

$$\det(A) = (-2)^2 \cdot 4 = 16$$

Nachprüfen ergibt:

$$\det \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} = -20 - 54 - 54 + 18 + 36 + 90 = 16$$

6.3.2 Eigenwerte spezieller Matrizen

Bemerkung 6.21: Die Eigenwerte einer Dreiecksmatrix sind die Werte auf der Hauptdiagonalen.

BEWEIS Ist A eine Dreiecksmatrix, so ist $A - \lambda E$ ebenfalls eine Dreiecksmatrix und das charakteristische Polynom ist

$$p(\lambda) = (a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda) = 0$$

Diese Gleichung ist erfüllt für $\lambda_1 = a_{11}, \dots, \lambda_n = a_{nn}$. ■

Bemerkung 6.22: Ist λ ein Eigenwert von A und x ein zugehöriger Eigenvektor, so gilt:

$$A^k x = \lambda^k x$$

BEWEIS

$$\begin{aligned} A^k x &= A^{k-1} \cdot Ax = A^{k-1} \cdot \lambda x = \lambda A^{k-1} x \\ &= \dots = \lambda^k x \end{aligned}$$

■

Bemerkung 6.23: Der Absolutbetrag der Eigenwerte von Orthogonalmatrizen $A \in O(n)$ ist 1, d.h.

$$Ax = \lambda x \quad \Rightarrow \quad |\lambda| = 1.$$

Damit bleibt bei der Abbildung mit A nicht nur die Richtung, sondern auch die Länge erhalten.

BEWEIS Mit Bemerkung 6.10 erhält man:

$$\|x\| = \|Ax\| = \|\lambda x\| = |\lambda| \cdot \|x\| \quad \Leftrightarrow \quad |\lambda| = 1.$$

■

6.3.3 Diagonalisierbarkeit linearer Abbildungen

Gegeben sei eine beliebige $n \times n$ -Matrix. Hat diese Matrix n linear unabhängige Eigenvektoren $\vec{x}_1, \dots, \vec{x}_n$, so gilt mit den dazugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$:

$$\begin{aligned} A\vec{x}_1 &= \lambda_1\vec{x}_1 \\ &\vdots \\ A\vec{x}_n &= \lambda_n\vec{x}_n \end{aligned}$$

In Matrixschreibweise:

$$A \overbrace{(\vec{x}_1, \dots, \vec{x}_n)}^B = \overbrace{(\vec{x}_1, \dots, \vec{x}_n)}^B \cdot \overbrace{\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}^D,$$

wobei Eigenwerte auch doppelt sein dürfen. B ist invertierbar und D diagonal. Daraus ergibt sich die Diagonalisierung einer linearen Abbildung:

$$AB = BD$$

$$D = B^{-1}AB \quad \text{bzw.} \quad A = BDB^{-1}$$

In den Spalten von B stehen die Eigenvektoren. Die Diagonalelemente von D sind die entsprechenden Eigenwerte.

Manche quadratischen Matrizen können nicht diagonalisiert werden. Andere besitzen komplexe Eigenwerte und -vektoren, so dass die Diagonalisierung keinen praktischen Nutzen hat. Im Einzelnen gelten folgende Punkte:

- Zu jeder $(n \times n)$ -Matrix gehört das charakteristische Polynom

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0.$$

Damit gibt es m Eigenwerte $\lambda_0, \dots, \lambda_m$, die mehrfach sein können. Weiterhin können die Eigenwerte komplex sein. Hat ein Eigenwert λ_i die Vielfachheit v_i , dann ist $\sum_{i=1}^m v_i = n$. v_i heißt die **algebraische Vielfachheit** von λ_i .

- Jeder Eigenwert besitzt eine Anzahl g_i von linear unabhängigen Eigenvektoren. g_i nennt man die **geometrische Vielfachheit** von λ_i . Es gilt:

$$1 \leq g_i \leq v_i.$$

Jeder Eigenwert hat also mindestens einen Eigenvektor. Ein doppelter Eigenwert kann zwei (linear unabhängige) Eigenvektoren, aber auch nur einen Eigenvektor besitzen. Die Menge E_λ aller Eigenvektoren zu λ und des Nullvektors nennt man **Eigenraum** zu λ und bildet offensichtlich einen Unterraum.

- Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.
- Eine Abbildung ist genau dann diagonalisierbar, wenn die geometrische und die algebraische Vielfachheit aller Eigenwerte gleich sind. Anders ausgedrückt: Wenn es zu jedem n -fachen Eigenwert auch n Eigenvektoren gibt. Insbesondere ist eine Abbildung dann diagonalisierbar, wenn jeder Eigenwert nur einfach ist.
- Wie im nächsten Kapitel gezeigt wird, lassen sich reellsymmetrische Matrizen immer diagonalisieren, denn für sie gilt:
 - Eigenwerte und Eigenvektoren sind reell.
 - Jeder n -fache Eigenwert hat n orthonormale Eigenvektoren.

6.4 Symmetrische Matrizen

Symmetrische Matrizen hatten wir in Definition 1.14 definiert.

Satz 6.24: Für reelle symmetrische Matrizen sind die Eigenwerte immer reell und es existiert immer ein reeller Eigenvektor.

BEWEIS Es gelte also $Az = \lambda z$, wobei wir zunächst annehmen, dass der Vektor $z = x + iy$ ($x, y \in \mathbb{R}^n$) und $\lambda = \gamma + i\omega$ ($\gamma, \omega \in \mathbb{R}$) komplex sind:

$$A(x + iy) = (\gamma + i\omega)(x + iy).$$

Da die Elemente von A reell sind, kann man dies nach Real- und Imaginärteil aufteilen und erhält

$$\begin{aligned} Ax &= \gamma x - \omega y, \\ Ay &= \gamma y + \omega x. \end{aligned}$$

Nun ist A symmetrisch. Daher gilt

$$\langle Ax, y \rangle = (Ax)^T y = x^T A^T y = \langle x, A^T y \rangle = \langle x, Ay \rangle.$$

Daraus folgt

$$\langle \gamma x - \omega y, y \rangle = \langle x, \gamma y + \omega x \rangle,$$

also

$$\gamma \langle x, y \rangle - \omega \langle y, y \rangle = \gamma \langle x, y \rangle + \omega \langle x, x \rangle$$

oder

$$\omega(\|x\|^2 + \|y\|^2) = 0$$

und da der Eigenvektor $x + iy = z \neq 0$ ist, folgt $\omega = 0$, also $\lambda \in \mathbb{R}$.

Da λ reell ist, gibt es auch einen reellen Eigenvektor x . ■

Für die Eigenvektoren von A gilt nun:

Bemerkung 6.25: Die zu verschiedenen Eigenwerten einer reellsymmetrischen Matrix gehörigen Eigenvektoren sind orthogonal zueinander.

BEWEIS Für zwei Eigenwerte $\lambda_1 \neq \lambda_2$ einer Matrix A seien x_1 und x_2 zwei (bekanntlich von Null verschiedene) zugehörige Eigenvektoren aus \mathbb{R}^n . Dann gilt

$$\begin{aligned} \langle x_2, Ax_1 \rangle &= x_2^T Ax_1 = (A^T x_2)^T x_1 \\ &= (Ax_2)^T x_1 = \langle Ax_2, x_1 \rangle \\ &= \langle x_1, Ax_2 \rangle \end{aligned}$$

Wegen

$$\begin{aligned} \langle x_2, Ax_1 \rangle &= \lambda_1 \langle x_2, x_1 \rangle \text{ und } \langle x_1, Ax_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle \\ \text{folgt } \lambda_1 \langle x_2, x_1 \rangle &= \lambda_2 \langle x_1, x_2 \rangle, \text{ also } (\lambda_1 - \lambda_2) \langle x_2, x_1 \rangle = 0. \end{aligned}$$

Da nach Voraussetzung $\lambda_1 \neq \lambda_2$ ist, folgt $\langle x_1, x_2 \rangle = 0$ d.h. $\vec{x}_1 \perp \vec{x}_2$. ■

Falls alle Eigenwerte $\lambda_1, \dots, \lambda_n$ unterschiedlich sind, und die Eigenvektoren $\vec{x}_1, \dots, \vec{x}_n$ so gewählt werden, dass $\|\vec{x}_1\| = \dots = \|\vec{x}_n\| = 1$, dann ist $B = (\vec{x}_1, \dots, \vec{x}_n)$ nicht nur invertierbar, sondern sogar orthogonal und es gilt:

$$B^T A B = D$$

Wie wir sehen werden, gilt dies auch, wenn mehrfache Eigenwerte auftreten.

Satz 6.26: Zu jeder reellsymmetrischen Matrix A gibt es eine Orthogonalmatrix U und eine Diagonalmatrix D mit

$$U^T A U = D$$

mit

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix}$$

BEWEIS Um dies zu beweisen, geht man schrittweise vor.

Hilfssatz 1: Ist A eine reellsymmetrische $(n \times n)$ -Matrix und ist x ein Eigenvektor von A zum (reellen) Eigenwert λ von A , dann gilt für das orthogonale Komplement $x^\perp = \{y \in \mathbb{R}^n; y \perp x\}$, dass aus

$$y \in x^\perp \Rightarrow Ay \in x^\perp.$$

Beweis: Es gilt $\langle Ay, x \rangle = \langle y, A^T x \rangle = \langle y, Ax \rangle = \lambda \langle y, x \rangle = 0$ wegen der Symmetrie von A und der Voraussetzung $y \perp x$. Also ist auch $Ay \in x^\perp$.

Weiter kann man, wenn man einen Eigenwert kennt, das Problem um eine Dimension, also von n auf $n-1$ reduzieren.

Hilfssatz 2: Ist x_1 normierter Eigenvektor einer reellsymmetrischen Matrix A zum Eigenwert λ_1 , dann existiert eine Orthogonalmatrix V und eine symmetrische $(n-1) \times (n-1)$ -Matrix B mit

$$V^T \cdot A \cdot V = \begin{pmatrix} \lambda_1 & 0 \\ 0 & B \end{pmatrix}.$$

Beweis: Man definiert die Matrix V durch $V = (x_1 | V_{n-1})$, wobei die $n \times (n-1)$ -Matrix V_{n-1} gerade $(n-1)$ orthonormierte Spaltenvektoren aus dem orthogonalen Komplement des Vektors x_1 enthält. Dann folgt

$$A \cdot V = A \cdot (x_1 | V_{n-1}) = (\lambda_1 x_1 | A \cdot V_{n-1})$$

und

$$\begin{aligned} V^T \cdot A \cdot V &= \left(\begin{array}{c|c} x_1^T & \\ \hline V_{n-1}^T & \end{array} \right) (\lambda_1 x_1 | A \cdot V_{n-1}) \\ &= \left(\begin{array}{c|c} x_1^T \lambda_1 x_1 & x_1^T A V_{n-1} \\ \hline V_{n-1}^T \lambda_1 x_1 & V_{n-1}^T A \cdot V_{n-1} \end{array} \right) \\ &= \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & V_{n-1}^T A \cdot V_{n-1} \end{array} \right). \end{aligned}$$

Die Nullen in der 1. Zeile entstehen wegen $x^T \cdot A V_{n-1} = (0, \dots, 0)$ wegen Hilfssatz 1 und die Nullen in der 1. Spalte wegen $V_{n-1}^T \cdot \lambda_1 \cdot x = (0, \dots, 0)^T$ ebenfalls wegen Hilfssatz 1, da die Zeilenvektoren von V_{n-1}^T das orthogonale Komplement des Vektors x bilden. $B = V_{n-1}^T A V_{n-1}$ ist symmetrisch, da

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

Somit ist B eine symmetrische $(n-1) \times (n-1)$ -Matrix.

Wenn man jetzt annimmt, dass für die $(n-1) \times (n-1)$ -Matrix B ein Orthonormalsystem von Eigenvektoren (im \mathbb{R}^{n-1}) existiert, dann kann man zeigen, dass die $(n \times n)$ -Matrix A ein Orthonormalsystem von Eigenvektoren im \mathbb{R}^n besitzt (**Schluss von $n-1$ auf n**).

Hilfssatz 3: Ist B eine reellsymmetrische $(n-1) \times (n-1)$ -Matrix, für die eine Diagonalmatrix D_{n-1} und eine Orthogonalmatrix $W \in O(n-1)$ existiert, dann gilt dies auch für eine reellsymmetrische $(n \times n)$ -Matrix A .

Beweis: Man definiert eine Matrix U durch

$$U = V \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W \end{array} \right),$$

wobei $V \in O(n)$ wie im Beweis von Hilfssatz 3 definiert ist. Damit gilt auch $U \in O(n)$ und es folgt

$$A \cdot U = A \cdot V \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W \end{array} \right),$$

also

$$U^T \cdot A \cdot U = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W^T \end{array} \right) V^T \cdot A \cdot V \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W \end{array} \right)$$

und nach Hilfssatz 3

$$\begin{aligned} U^T \cdot A \cdot U &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W^T \end{array} \right) \cdot \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & B \end{array} \right) \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W \end{array} \right) \\ &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W^T \end{array} \right) \cdot \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & B \cdot W \end{array} \right) \end{aligned}$$

$$= \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & W^T \cdot B \cdot W \end{array} \right) = \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & D_{n-1} \end{array} \right),$$

wobei die letzte Matrix eine $(n \times n)$ - Diagonalmatrix ist, die auf der Diagonalen die Eigenwerte λ_i der reellsymmetrischen Matrix A enthält. ■

Beispiel 6.27: Gegeben ist die symmetrische Matrix

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

Der vorige Satz sagt: Es existieren drei orthonormale Vektoren $\vec{u}_1, \vec{u}_2, \vec{u}_3$ und drei Eigenwerte $\lambda_1, \lambda_2, \lambda_3$, so dass mit $U = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$ gilt:

$$U^T A U = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Suchen der Eigenwerte:

$$\begin{aligned} \det \begin{pmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{pmatrix} &= (1-\lambda)^3 + 1 + 1 - 3(1-\lambda) \\ &= (3-\lambda)\lambda^2 = 0 \\ \Rightarrow \lambda_1 &= 3, \quad \lambda_{2/3} = 0 \end{aligned}$$

Suchen der orthonormalen Vektoren:

$\lambda_1 = 3$:

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

Die Lösung des Gleichungssystems ergibt $x_1 = x_2 = x_3$. Der normierte Eigenvektor ist

$$\vec{u}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$\lambda_{2/3} = 0$:

$$\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}$$

Die Lösung des Gleichungssystems hat zwei Freiheitsgrade. Wähle zwei linear unabhängige Lösungen und orthonormiere sie (falls sich die Lösungen orthonormieren lassen, waren sie linear unabhängig. Man muss das nicht extra überprüfen). Wähle z.B.:

$$\vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

Aus der Orthonormalisierung nach Gram-Schmidt ergibt sich:

$$\vec{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \vec{u}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

Damit ist

$$U = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix}$$

und

$$D = U^T A U = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Umgekehrt gilt:

Bemerkung 6.28: Gibt es zu einer Matrix A eine Orthogonalmatrix U und eine Diagonalmatrix D mit

$$U^T A U = D,$$

so ist A symmetrisch.

BEWEIS Es gilt: $U^T A U = D$, sowie $D = D^T$ und $U^T = U^{-1}$. Dann ist:

$$\begin{aligned} A &= (U^T)^{-1} D U^{-1} = U D U^T \\ A^T &= (U D U^T)^T = (U^T)^T D^T U^T = U D U^T = A \end{aligned}$$

■

6.5 Definitheit von Matrizen

6.5.1 Definitheit

Es gibt spezielle reellsymmetrische Matrizen mit zusätzlichen Eigenschaften, die bei der Lösung von Gleichungssystemen und Extremwert- bzw. Optimierungsaufgaben wichtig sind, wie wir sie aus der Analysis oder der Numerik kennen.

Zunächst einige Definitionen:

Definition 6.29: Das Produkt $F = \langle x, Ax \rangle$ wird **quadratische Form** genannt.

Beispiel 6.30: Es sei

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Dann ist die quadratische Form:

$$\begin{aligned} \langle x, Ax \rangle &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix} \right\rangle \\ &= x_1^2 + 2x_1x_2 + 3x_2x_1 + 4x_2^2 \\ &= x_1^2 + 5x_1x_2 + 4x_2^2 \end{aligned}$$

Definition 6.31: Die **Hauptunterdeterminanten** $D_k = \det(A_k)$ einer Matrix $A \in M(n \times n, \mathbb{K})$ sind die Determinanten der von links oben beginnenden $k \times k$ -Untermatrizen $A_k = (a_{ij}), 1 \leq i \leq k, 1 \leq j \leq k, k = 1, \dots, n$.

Beispiel 6.32: Die Matrix

$$M = \begin{pmatrix} 6 & 1 & -3 & 2 \\ 1 & 8 & 0 & -2 \\ -3 & 0 & 6 & 3 \\ 2 & -2 & 3 & 7 \end{pmatrix}$$

hat vier Hauptunterdeterminanten $D_1 - D_4$. Es sind die Determinanten der vier möglichen quadratischen Teilmatrizen (1×1 bis 4×4), die das linke

obere Element von M einschließen:

$$\begin{aligned}
 D_1 &= \det(6) = 6 \\
 D_2 &= \det \begin{pmatrix} 6 & 1 \\ 1 & 8 \end{pmatrix} = 48 - 1 = 47 \\
 D_3 &= \det \begin{pmatrix} 6 & 1 & -3 \\ 1 & 8 & 0 \\ -3 & 0 & 6 \end{pmatrix} = 288 - 6 - 72 = 210 \\
 D_4 &= \det \begin{pmatrix} 6 & 1 & -3 & 2 \\ 1 & 8 & 0 & -2 \\ -3 & 0 & 6 & 3 \\ 2 & -2 & 3 & 7 \end{pmatrix} = 375
 \end{aligned}$$

Jetzt zu den speziellen reellsymmetrischen Matrizen:

Definition 6.33: Eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ heißt **positiv definit**, falls die **quadratische Form** $F = \langle x, Ax \rangle > 0$ ist für alle Vektoren $x \in \mathbb{R}^n$ mit $x \neq 0$.

Häufig verwendet man auch die Abkürzung „spd“ für symmetrisch positiv definit.

Satz 6.34: Folgende Aussagen sind äquivalent für symmetrische Matrizen A :

1. A ist positiv definit;
2. A hat nur positive Eigenwerte;
3. sämtliche Hauptunterdeterminanten von A sind positiv;

Der vollständige Beweis des Satzes ist z.B. in *J.H. Wilkinson; The Algebraic Eigenvalue Problem, Oxford 1965, p.28, p.229* zu finden.

Beispiel 6.35: Die Matrix

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

ist positiv definit. Die Hauptunterdeterminanten von A sind $D_1 = 2$ und $D_2 = 4 - 1 = 3$ und somit beide positiv. Die Eigenwerte von A ergeben sich aus:

$$\det(A - \lambda E) = \det \begin{pmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{pmatrix} = \lambda^2 - 4\lambda + 3 = 0$$

und haben die Werte $\lambda_1 = 1$ und $\lambda_2 = 3$, sind also beide positiv. Für die quadratische Form gilt:

$$\begin{aligned}\langle z, Az \rangle &= \left\langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} 2x + y \\ x + 2y \end{pmatrix} \right\rangle \\ &= 2x^2 + 2xy + 2y^2 \\ &= (x + y)^2 + x^2 + y^2 > 0\end{aligned}$$

Bemerkung 6.36: Eine positiv definite Matrix A ist invertierbar.

BEWEIS Aufgrund Eigenschaft 3 ist $\det(A) \neq 0$. Damit ist A invertierbar.

■

Definition 6.37: Eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ heißt

1. **negativ definit**, falls gilt $\langle x, Ax \rangle < 0 \quad \forall x \in \mathbb{R}_{\neq 0}^n$;
2. **positiv semidefinit**, falls gilt $\langle x, Ax \rangle \geq 0 \quad \forall x \in \mathbb{R}_{\neq 0}^n$;
3. **negativ semidefinit**, falls gilt $\langle x, Ax \rangle \leq 0 \quad \forall x \in \mathbb{R}_{\neq 0}^n$;
4. **indefinit**, falls gilt sie weder positiv noch negativ (semi-)definit ist, d.h.

$$\exists x, y \in \mathbb{R}_{\neq 0}^n \quad \text{mit} \quad \langle x, Ax \rangle > 0 \quad \wedge \quad \langle y, Ay \rangle < 0.$$

Bemerkung 6.38: Ist A negativ definit, dann gilt $-\langle x, Ax \rangle = \langle x, -Ax \rangle > 0$ d.h. $-A$ ist positiv definit.

Bemerkung 6.39: Für die Hauptunterdeterminanten einer negativ definiten Matrix gilt:

$$\begin{aligned}\det(A_k) &= \det(-(-A_k)) = (-1)^k \cdot \det(-A_k) \\ (-1)^k \det(A_k) &= \det(-A_k) > 0\end{aligned}$$

Das heißt, die Hauptunterdeterminanten A_k sind für $k = 1, \dots, n$ abwechselnd positiv und negativ. In der Analysis haben wir das mit Hilfe des **Hurwitz-Kriteriums** überprüft.

	$\langle x, Ax \rangle$	D_1	D_2	D_3	D_4	Qualität Kriterium
positiv definit	> 0	> 0	> 0	> 0	> 0	hinreichend
negativ definit	< 0	< 0	> 0	< 0	> 0	hinreichend
positiv semidefinit	≥ 0	≥ 0	≥ 0	≥ 0	≥ 0	notwendig
negativ semidefinit	≤ 0	≤ 0	≥ 0	≤ 0	≥ 0	notwendig
indefinit		sonst				hinreichend

Die letzte Spalte sagt dabei aus, dass man aus dem Vorzeichen der Unterdeterminanten auf positiv oder negativ Definitheit schließen kann, da es sich um ein hinreichendes Kriterium handelt.

Selbiges gilt jedoch nicht für die beiden grau gedruckten Zeilen der Semidefinitheit, da es sich hier nur um ein notwendiges Kriterium handelt.

Das Kriterium für die Indefinitheit ist zwar hinreichend, aber es gibt auch indefinite Matrizen, bei denen Unterdeterminanten gleich 0 sind.

6.5.2 Die Matrix $A^T \cdot A$

Ein spezielles Beispiel einer positiv definiten Matrix wird im folgenden Hilfsatz angegeben.

Bemerkung 6.40: Sind die n Spalten einer Matrix $A \in M(m \times n, \mathbb{R})$ linear unabhängig, dann ist die Matrix $A^T A$ positiv definit.

BEWEIS Die Matrix A hat n Spaltenvektoren $\vec{a}_i \in \mathbb{R}^m$. Also muss $n \leq m$ sein, damit diese Vektoren linear unabhängig sein können. Nun ist

$$\langle x, A^T A x \rangle = x^T A^T A x = (Ax)^T A x = \langle Ax, Ax \rangle = \|Ax\|^2.$$

Weiter gilt

$$Ax = x_1 \vec{a}_1 + \dots + x_n \vec{a}_n$$

und da die Spaltenvektoren \vec{a}_i linear unabhängig sind, ist dieser Ausdruck nur gleich Null, wenn alle x_i gleich Null sind d.h. x der Nullvektor ist. Also gilt für alle $x \neq 0$, dass

$$\langle x, A^T A x \rangle = \|Ax\|^2 > 0$$

ist d.h. dass die Matrix $A^T A$ positiv definit ist. Also ist sie auch regulär und damit invertierbar. ■

Bemerkung 6.41: Vertauscht man Zeilen und Spalten, dann gilt, dass die Matrix AA^T positiv definit ist, falls die Zeilen von A linear unabhängig sind.

Bemerkung 6.42: Ist die Matrix A außerdem quadratisch, dann folgt bekanntlich aus der linearen Unabhängigkeit der Spalten die Bedingung $\det(A) \neq 0$. Wegen

$$\det(A^T \cdot A) = \det(A^T) \cdot \det(A) = \det(A)^2 \neq 0$$

gilt also, dass die Matrix A genau dann invertierbar ist, wenn $A^T \cdot A$ invertierbar ist. Dies sind die äquivalenten Bedingungen 1 und 18 aus Satz 5.17 in Kapitel 5 über lineare Gleichungssysteme.

6.6 Hauptachsentransformationen

6.6.1 Einführung

Ist eine Matrix $A \in M(n \times n, \mathbb{R})$ symmetrisch, dann existiert nach Satz 6.26 eine Orthogonalmatrix Q mit der Eigenschaft

$$A = QDQ^T,$$

wobei D eine Diagonalmatrix ist, deren Diagonalelemente D_{ii} die (reellen) Eigenwerte von A sind. Die Spalten von Q sind die Eigenvektoren von A , wobei die zum gleichen Eigenwert gehörigen Eigenvektoren so zu wählen sind, dass sie zueinander orthogonal stehen. Umgestellt ergibt sich die Gleichung

$$D = Q^T A Q.$$

Das bedeutet, dass man ein Koordinatensystem A' finden kann, in dem die Abbildung A diagonal ist. Die Umrechnungsmatrix nach Kapitel 4.5.1 ist $T_A^{A'} = Q$. Da Q orthogonal ist, sind nur Drehungen und Spiegelungen möglich. Die neuen Koordinatenachsen sind die Spalten von Q , also die Eigenvektoren von A .

Beispiel 6.43: Aus Beispiel 4.92 erhielten wir bei gleichen Basen im Definitionsbereich und Bild $A' = B'$ die Abbildungsmatrix

$$M_{A'}^{A'}(\Phi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Da $M_{A'}^{A'}(\Phi)$ symmetrisch ist, kann $M_{A'}^{A'}(\Phi)$ mit einer Hauptachsentransformation wieder diagonalisiert werden:

$$\det(M_{A'}^{A'}(\Phi) - \lambda E) = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$$

Bei der Bestimmung der Matrizen Q und D gibt es folgende Freiheitsgrade:

- Die Eigenwerte können in ihrer Reihenfolge variiert werden (je zwei Möglichkeiten).
- Die Eigenvektoren können mit -1 multipliziert werden (je vier Möglichkeiten).

Bei zwei Eigenwerten bzw. Eigenvektoren gibt es insgesamt acht Möglichkeiten. Schließt man mit der Bedingung $\det(Q) = 1$ die Spiegelungen aus, verbleiben vier Möglichkeiten, die jeweils um 90° gegeneinander gedreht sind. Die Rücktransformation in die ursprüngliche Basis A erfolgt mit

$$\lambda_1 = -1, \lambda_2 = 1, \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

6.6.2 Hauptachsentransformation von Kegelschnitten

Eine Anwendung des letzten Abschnitts ist, gedrehte Kegelschnitte mit Hilfe einer Hauptachsentransformation zu bestimmen.

6.6.2.1 Hauptachsentransformation von quadratischen Formen

Eine quadratische Form ist definiert durch

$$F(x) = \langle x, Zx \rangle; \quad x \in \mathbb{R}^n, \quad Z \in M(n \times n, \mathbb{R}).$$

Außerdem ist Z als symmetrisch vorausgesetzt. Dann gilt:

$$Z = QDQ^T.$$

Führt man eine neue Basis U ein, so dass $T_A^U = Q$, dann gilt:

$$F(x) = \langle x, QDQ^T x \rangle = x^T QDQ^T x = (Q^T x)^T D(Q^T x) = (T_U^A x)^T D(T_U^A x).$$

Rechts und links steht $T_U^A x = K_U(x)$, also die Darstellung von x bezüglich der Basis U . Schreiben wir diese Vektoren als $u = K_U(x)$, lässt sich die Gleichung schreiben als:

$$F(u) = u^T D u = \langle u, D u \rangle = \lambda_1 u_1^2 + \cdots + \lambda_n u_n^2,$$

d.h. bezüglich der neuen Koordinaten $u^T = (u_1, \dots, u_n)$ besitzt F nur noch Quadrate und die λ_i sind die Eigenwerte der Matrix A .

6.6.2.2 Hauptachsentransformation von Kegelschnitten

Mit den bisherigen Ergebnissen können gedrehte Kegelschnitte bestimmt werden. Wir betrachten quadratische Gleichungen in (x, y) der Form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0, \quad x, y \in \mathbb{R}.$$

Aus Kapitel 2.4.6 wissen wir, dass ein Kegelschnitt gedreht ist, wenn $b \neq 0$ ist. Diese Art von Kegelschnitten konnten wir bisher noch nicht bestimmen. Jetzt können wir allerdings ein Koordinatensystem bestimmen, in dem der Kegelschnitt nicht gedreht, d.h. $b = 0$ ist. Die zugehörige quadratische Form, die zu betrachten ist, ist

$$ax^2 + bxy + cy^2 = \langle \vec{x}, A\vec{x} \rangle,$$

mit

$$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}; \quad A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix};$$

Da A symmetrisch ist, kann diese Gleichung umgeformt werden in:

$$\langle \vec{x}, A\vec{x} \rangle = \langle \vec{x}, QDQ^T, \vec{x} \rangle = \langle \vec{u}, D\vec{u} \rangle,$$

wobei $\vec{u} = (u, v)^T$ die Koordinaten bezüglich eines gedrehten Systems U sind. Dort verschwindet das gemischte Glied und es gilt $\langle \vec{u}, D\vec{u} \rangle = a'u^2 + c'v^2$. Wir haben also als Zwischenergebnis:

$$a'u^2 + c'v^2 + dx + ey + f = 0.$$

a' und c' sind dabei die Diagonalelemente von D . Für den zweiten Teil definiert man $\vec{k} = (d, e)^T$. Damit ist

$$dx + ey = \langle \vec{k}, \vec{x} \rangle = \langle \vec{k}, T_A^U \vec{u} \rangle = \langle \vec{k}, Q\vec{u} \rangle = \vec{k}^T Q\vec{u} = (Q^T \vec{k})^T \vec{u} = \langle Q^T \vec{k}, \vec{u} \rangle$$

Man erhält mit $\vec{k}' = \begin{pmatrix} d' \\ e' \end{pmatrix} = Q^T \vec{k}$:

$$a'u^2 + c'v^2 + d'u + e'v + f = 0.$$

Beispiel 6.44: $F = 5x^2 - 4xy + 8y^2 - 36 = 0$ oder

$$\vec{x}^T A \vec{x} - 36 = 0; \quad A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}.$$

Hier ist also $d = e = 0$. Das charakteristische Polynom von A ist

$$p(\lambda) = \det(A - \lambda E) = (5 - \lambda)(8 - \lambda) - 4 = (\lambda - 9)(\lambda - 4):$$

Die Eigenwerte von A sind also $\lambda_1 = 4$ und $\lambda_2 = 9$ und die zugehörigen Eigenvektoren sind

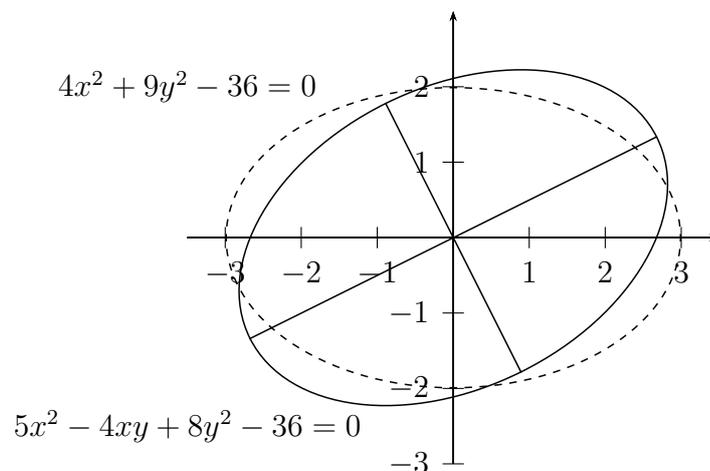
$$\vec{x}^{(1)} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \vec{x}^{(2)} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \Rightarrow Q = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \Rightarrow$$

$$F = 4u^2 + 9v^2 - 36 = 0 \quad \Leftrightarrow \quad \left(\frac{u}{3}\right)^2 + \left(\frac{v}{2}\right)^2 = 1.$$

Die Kurve ist also eine Ellipse mit den Halbachsen 2 und 3. Die genaue Lage der Ellipse ist nicht Bestandteil der Vorlesung, soll aber hier kurz erwähnt werden. Durch die Hauptachsentransformation kann das Koordinatensystem gedreht und an einer Geraden durch den Nullpunkt gespiegelt sein. Da $\det(Q) = 1$, handelt es sich um eine reine Rotation (ohne Beweis). Bei $\det(Q) = -1$ wäre der Kegelschnitt zusätzlich noch gespiegelt worden. Der Drehwinkel ist

$$\cos(\theta) = \langle \vec{x}^{(1)}, \vec{e}_x \rangle = \frac{2}{\sqrt{5}}$$

$$\theta \approx 26,565^\circ$$



Beispiel 6.45: Das 2. Beispiel ist der Graph der Gleichung

$$F = 17x^2 - 12xy + 8y^2 - 68x + 24y - 12 = 0.$$

Es soll nur die Mittelpunktsgleichung, nicht die genaue Lage des Kegelschnitts berechnet werden. Die quadratische Form ist mit $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$\langle x, Ax \rangle = \langle \vec{x}, \begin{pmatrix} 17 & -6 \\ -6 & 8 \end{pmatrix} \vec{x} \rangle = 17x^2 - 12xy + 8y^2$$

Das charakteristische Polynom zu A ist:

$$\begin{aligned} p(\lambda) &= \det(A - \lambda E) = (17 - \lambda)(8 - \lambda) - 36 \\ &= \lambda^2 - 25\lambda + 136 - 36 \\ &= \lambda^2 - 25\lambda + 100 \end{aligned}$$

Daraus ergeben sich die Eigenwerte

$$\begin{aligned} \lambda_{1/2} &= \frac{25}{2} \pm \sqrt{\left(\frac{25}{2}\right)^2 - 100} \\ &= \frac{25}{2} \pm \sqrt{\frac{625 - 400}{4}} \\ &= \frac{25}{2} \pm \frac{15}{2} \end{aligned}$$

Damit ist $\lambda_1 = 20$ und $\lambda_2 = 5$. Der normierte Eigenvektor zu λ_1 ist die Lösung des Gleichungssystems

$$\begin{array}{r|l} -3 & -6 & 0 \\ -6 & -12 & 0 \\ \hline 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}$$

und somit $\vec{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Der Eigenvektor zu λ_2 ergibt sich mit

$$\begin{array}{r|l} 12 & -6 & 0 \\ -6 & 3 & 0 \\ \hline 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}$$

zu $\vec{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Damit erhält man

$$Q = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}, \quad Q^T = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 20 & 0 \\ 0 & 5 \end{pmatrix}$$

Jetzt müssen noch der x - und der y -Term behandelt werden. Es ist

$$\begin{pmatrix} d' \\ e' \end{pmatrix} = Q^T \begin{pmatrix} -68 \\ 24 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -68 \\ 24 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -160 \\ -20 \end{pmatrix}$$

Damit ergibt sich als Zwischengleichung:

$$20v_1^2 + 5v_2^2 - \frac{160}{\sqrt{5}}v_1 - \frac{20}{\sqrt{5}}v_2 - 12 = 0.$$

Da $a \cdot b = 20 \cdot 5 > 0$ und $a \neq b$, handelt es sich um eine Ellipse. Diese muss jetzt noch durch quadratische Ergänzung in die Standardlage gebracht werden. Es ist:

$$\begin{aligned} 20v_1^2 - \frac{160}{\sqrt{5}}v_1 &= 20\left(v_1^2 - \frac{8}{\sqrt{5}}v_1\right) \\ &= 20\left[\left(v_1 - \frac{4}{\sqrt{5}}\right)^2 - \frac{16}{5}\right] \\ &= 20\left(v_1 - \frac{4}{\sqrt{5}}\right)^2 - 64 \\ 5v_2^2 - \frac{20}{\sqrt{5}}v_2 &= 5\left(v_2^2 - \frac{4}{\sqrt{5}}v_2\right) \\ &= 5\left[\left(v_2 - \frac{2}{\sqrt{5}}\right)^2 - \frac{4}{5}\right] \\ &= 5\left(v_2 - \frac{2}{\sqrt{5}}\right)^2 - 4 \end{aligned}$$

Mit $u_1 = v_1 - \frac{4}{\sqrt{5}}$ und $u_2 = v_2 - \frac{2}{\sqrt{5}}$ ergibt sich schließlich die Gleichung

$$\begin{aligned} 20u_1^2 - 64 + 5u_2^2 - 4 - 12 &= 0 \\ 20u_1^2 + 5u_2^2 &= 80 \\ \frac{1}{4}u_1^2 + \frac{1}{16}u_2^2 &= 1 \end{aligned}$$

Es handelt sich also um eine Ellipse mit Hauptachsen der Länge 2 und 4.

6.7 Hadamardsche Ungleichung

Aus dem Schmidtschen Orthogonalisierungsverfahren und damit aus der QR-Zerlegung lassen sich einige einfache Ergebnisse herleiten.

Satz 6.46: (Hadamardsche Ungleichung): Es sei A eine reelle $n \times n$ -Matrix $\neq 0$. Ferner sei

$$H = \frac{|\det(A)|}{\prod_{k=1}^n \|a^{(k)}\|}.$$

Dann gilt: $0 \leq H \leq 1$. Die Grenzfälle bedeuten:

$H = 1$: Die Spalten von A sind orthogonal

$H = 0$: Die Spalten von A sind linear abhängig

Je kleiner H ist, um so weniger weichen die Spaltenvektoren der Matrix A von linear abhängigen Spaltenvektoren ab.

BEWEIS Die Spaltenvektoren $a^{(k)}$ von A sind genau dann linear abhängig, wenn $\det(A) = 0$ und somit $H = 0$. Sind die Spaltenvektoren von A linear unabhängig, dann gilt nach dem Schmidtschen Orthonormierungsverfahren

$$a^{(k)} = \sum_{i=1}^k r_{ik} q^{(i)} \quad (1)$$

wobei $q^{(i)}$ die orthonormierten Vektoren sind. Weiter folgt daraus

$$\begin{aligned} \|a^{(k)}\|^2 &= \langle a^{(k)}, a^{(k)} \rangle \\ &= \left(\sum_{i=1}^k r_{ik} q^{(i)} \right) \cdot \left(\sum_{j=1}^k r_{jk} q^{(j)} \right) \\ &= \sum_{i=1}^k r_{ik}^2 \\ &= \|r^{(k)}\|^2 \quad (2) \end{aligned}$$

Außerdem bedeutet (1), dass gilt

$$A = QR,$$

wobei Q eine orthogonale Matrix ist und R eine obere Dreiecksmatrix mit $\det(R) \neq 0$ d.h. mit nicht verschwindenden Diagonalelementen und es gilt

$$|\det(A)| = |\det(QR)| = |\det(Q)| |\det(R)| = |\det(R)| = |r_{11}| \cdots |r_{nn}| \\ \leq \|r^{(1)}\| \cdots \|r^{(n)}\| = \|a^{(1)}\| \cdots \|a^{(n)}\|$$

d.h.

$$|\det(A)| \leq \prod_{k=1}^n \|a^{(k)}\|.$$

Wegen

$$|r_{kk}| \leq \sqrt{r_{1k}^2 + \cdots + r_{kk}^2}$$

gilt Gleichheit genau dann, wenn R eine Diagonalmatrix ist, also alle Elemente außerhalb der Hauptdiagonalen gleich Null sind. Da

$$a^{(i)} = q^{(i)} \cdot r_{ii},$$

gilt in diesem Fall $(a^{(i)} \parallel q^{(i)})$. Da die $q^{(i)}$ paarweise orthogonal sind, sind auch die $a^{(i)}$ paarweise orthogonal. ■

Beispiel 6.47: Sind die Spalten der Matrix

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

orthogonal? Dies kann mit der Hadamardschen Ungleichung überprüft werden. Es gilt:

$$\det(A) = -4 - 4 - 4 + 1 - 8 - 8 = -27 \\ \prod_{k=1}^3 \|a^{(k)}\| = \sqrt{9} \cdot \sqrt{9} \cdot \sqrt{9} = 27$$

Damit sind die Spalten von A orthogonal.

Beispiel 6.48: Für eine 2×2 -Matrix $A = (\vec{a}, \vec{b})$ gilt:

$$\det(A) = \det(\vec{a}, \vec{b}) = \|\vec{a}\| \|\vec{b}\| \sin \theta$$

und somit

$$H = \frac{|\det(\vec{a}, \vec{b})|}{\|\vec{a}\| \|\vec{b}\|} = |\sin \theta|$$

d.h. H ist gleich dem Sinus des zwischen den beiden Spaltenvektoren eingeschlossenen Winkels. Im \mathbb{R}^3 gilt die Beziehung $H = |\sin \theta \cos \phi|$ (siehe Lineare Algebra 1).

Bekanntlich hat eine orthogonale Matrix Q die Eigenschaft $|\det(Q)| = 1$. Aus der Bedingung $|\det(Q)| = 1$ kann aber nicht geschlossen werden, dass Q orthogonal ist. Es gilt aber

Satz 6.49: Es sei A eine reelle $n \times n$ - Matrix mit $|\det(A)| = 1$. Außerdem seien sämtliche Spalten von A auf 1 normiert, dann ist A eine orthogonale Matrix.

BEWEIS Da die Spalten von A normiert sind, ist

$$\prod_{k=1}^n \|a_{(k)}\| = 1$$

Falls auch $|\det(A)| = 1$, ist $H = 1$ und die Spalten sind auch orthogonal. Daraus folgt, dass A eine Orthogonalmatrix ist. ■

Index

- Abbildung, 125
- Abbildungsmatrix, 174
- Abelsche Gruppe, 80
- Abgeschlossenheit, 80
- Abstand, 60
- Additivität, 157
- Äquivalenzklasse, 80
- Äquivalenzrelation, 79
- Äquivalenzumformungen, 18
- algebraische Vielfachheit, 212
- allgemeine Lösung, 180
- Assoziativität, 80
- Ausgleichsrechnung, 192
- Austauschlemma, 98

- Basis, 97
- Basisergänzungssatz, 97
- Basistransformation, 174
- Basiswechsel, 172
- Betrag eines Vektors, 37
- Betragssummennorm, 37
- Bijektivität, 126
- Bild, 125
- Bildmenge, 125
- Bilinearform, 111

- charakteristische Zahl, 205
- charakteristisches Polynom, 207
- Cramersche Regel, 187

- Definitionsbereich, 125
- Definitionsmenge, 125
- Determinante, 64, 157
 - Entwicklung, 165
 - Matrizenprodukt, 160
- Diagonalisierung, 212, 214
- Diagonalmatrix, 11, 215, 216, 222

- Differenzvektoren, 33
- dim, 97, 139
- Dimension, 97
- Dimensionsformel, 139
- direkte Summe, 89
- Distributivgesetz, 84
- Drehung, 175
- Dreiecksungleichung, 36, 45, 112

- Ebene im \mathbb{R}^3 , 55
- Eigenraum, 212
- Eigenvektor, 205
 - reell symm. Matrizen, 213
- Eigenwert, 205
 - algebraische Vielfachheit, 212
 - geometrische Vielfachheit, 212
 - Orthogonalmatrix, 211
 - Potenzen einer Matrix, 211
 - reell symm. Matrizen, 213
- Eigenwerte
 - von Dreiecksmatrizen, 210
- eindeutige Lösbarkeit, 186
- Eindeutigkeit der Dimension, 99
- Eineindeutigkeit, 126
- Einernorm, 37
- Einheitsmatrix, 12, 150
- Einheitsvektor, 37
- Elementarmatrix, 151
- Elemente, 9
- Ellipse, 224, 227
- Endomorphismen, 205
- Entwicklungssatz, 168
- Entwicklungssatz nach Laplace, 165
- Erzeugendensystem, 91
 - minimal, 97
- euklidische Norm, 37

- euklidische Skalarprodukt, 34
- Euklidischer Vektorraum, 110
- Faktorzerlegung, 102
- Funktion, 125
- Gauß-Algorithmus, 17, 150, 166
- Gauß-Jordan-Algorithmus, 17
- geometrische Vielfachheit, 212
- Gerade im \mathbb{R}^2 , 54
- Gerade im \mathbb{R}^3 , 53
- Gleichungssystem
 - äquivalent, 18
 - Äquivalenzumformungen, 18
 - homogen, 16, 179
 - inhomogen, 16, 179
 - linear, 16, 179
 - Lösbarkeit, 186
 - Lösung mit MATLAB, 20
 - reduzierte Stufenform, 18
 - Stufenform, 18
 - überbestimmtes, 21, 189
 - unterbestimmtes, 21, 196
- Grad, 101
- Gram-Schmidt, 118, 203
- Graph, 69
- Gruppe, 80
- Hadamardsche Ungleichung, 227
- Hauptachsentransformation, 222
- Hauptunterdeterminante, 218, 219
- Hessesche Normalform, 51
- Hom, 153
- homogenes Gleichungssystem, 179
- Homogenität, 157
- Homomorphismus, 128
- Hurwitz-Kriteriums, 220
- Hyperbel, 72
- Hyperebene, 52
- Hyperebene im \mathbb{R}^n , 56
- indefinit, 220
- inhomogenes Gleichungssystem, 179
- Injektivität, 126
- Interpolationspolynom, 103
- Inverse, 148
 - verallgemeinert, 196
 - verallgemeinerte, 190
- inverse Matrix, 150, 169
- inverses Element, 80
- invertierbare Matrix, 186
- isomorph, 145
- Isomorphismus, 130, 143
- Körper, 83
- kanonische Einheitsvektoren, 37
- Kartesisches Produkt, 85
- Kegelschnitt, 69, 223
 - gleichung, 73
 - Mittelpunktsform, 71
 - Standardlage, 71
 - Verschiebungsform, 71
- Kegelschnitte, 223
- Kern, 129
- Koeffizienten, 101
- Koeffizientenmatrix, 17
- Körper, 83
- kommutative Gruppe, 80
- komplementäre Matrix, 169
- Komplexer Vektorraum, 110
- komplexer Vektorraum, 86
- Komplexität, 168
- Komponente, 40
- Kongruenz, 82
- Koordinate, 29, 97
- Koordinatensystem
 - Drehung, 175
 - Spiegelung, 175
 - Verzerrung, 176
- Koordinatentransformationen, 175
- Kreuzprodukt, 46
- Leibnizsche Formel, 164, 168
- Leitkoeffizient, 101
- Lineare Abbildung, 128
- lineare Abbildung, 128
- Lineare Gleichungssysteme, 16, 179
- Lineare Hülle, 90
- lineare Regression, 193

- Lineare Unabhängigkeit, 91
- Lineare Unabhängigkeit
 - Polynome, 104
- linearer Raum, 85
- lineares Gleichungssystem, 67, 179
- Lineares Gleichungssystem, 187, 196
- Linearform, 111
- Linearkombination, 90
- lösbar, 180
- Lösungsmenge, 179
- Lotfußpunkt, 60, 118
- LSE, 190

- $M(m \times n, K)$, 87
- MATLAB, 15
- Matrix, 9, 147
 - Abbildungs-, 174
 - Addition, 10
 - äquivalent, 18
 - Äquivalenzumformungen, 18
 - erweitert, 17
 - Hauptunterdeterminante, 218
 - invertierbar, 186
 - komplementär, 169
 - MATLAB, 14, 20
 - Matrixprodukt, 12
 - Multiplikation mit Skalar, 10
 - orthogonal, 199
 - orthonormal, 199
 - symmetrisch, 12
 - transponierte, 10
- Maximumnorm, 37
- Methode der kleinsten Quadrate, 190
- Metrik, 113
- minimales Erzeugendensystem, 97

- negativ definit, 220
- negativ semidefinit, 220
- Neutralelement, 80
- nichtkommutative Gruppe, 81
- Norm, 36, 112
 - Betragssummen-, 37
 - Einer-, 37
 - euklidische, 37
 - Maximum-, 37
 - Zweier-, 37
- Normalenvektor, 50
- Normalform, 50
- Normalgleichungen, 189
- normierter Vektor, 38
- Nullmatrix, 9
- Nullraum, 129
- Nullstelle, 101

- $O(n)$, 199–202, 215
- Octave, 15
- orthogonale Projektion, 40
- orthogonales Komplement, 115
- Orthogonalität, 56, 114
- Orthogonalität von Vektoren, 39
- Orthogonalmatrix, 199, 202, 229
- Orthogonalsystem, 114
- Orthonormalbasis, 114
- Orthonormalisierungsverfahren, 118
- Orthonormalmatrix, 199
- Orthonormalsystem, 114
- Ortsvektor, 31

- Parabel, 72
- Parallelepiped, 65
- Parallelität, 56
- Parallelität von Vektoren, 33
- Parametergleichungen, 49
- parameterlosen Gleichungen, 50
- Pivot-Element, 22
- Pivot-Spalte, 22
- Polynom, 100, 101
- positiv definit, 219–221
- positiv semidefinit, 220
- Projektion, 40
- Punkt-Richtungsgleichung, 49
- Punktprodukt, 34

- QR-Zerlegung, 203
- quadratische Form, 218, 223

- Rang, 139, 154, 160, 171
- Rechte-Hand-Regel, 46
- Rechtssystem, 46

- Reeller Vektorraum, 110
- reeller Vektorraum, 86
- Reflexivität, 79
- Regression, 192
 - linear, 193
- Regularität, 130
- Repräsentant, 83
- Restklasse, 83
- rg, 139, 181

- Sarrussche Regel, 65, 163, 167
- Satz des Pythagoras, 39
- Schnittmengen, 57
- Schwarzsche Ungleichung, 44, 112
- Scilab, 15
- semidefinit, 220
- senkrechte Projektion, 40
- Singularität, 130
- Skalarprodukt, 34, 107
- Spaltenrang, 154
- Spaltenvektor, 147
- span, 90
- Spat, 65
- Spatprodukt, 65
- spezielle Lösung, 180
- Spiegelpunkt, 60
- Spiegelung, 175
- spitzer Winkel, 42
- Standardmetrik, 113
- Standardnorm, 36, 112
- Standardskalarprodukt, 34
- Stufenform, 18
- stumpfer Winkel, 42
- Surjektivität, 126
- Symmetrie, 79
- symmetrische Matrix, 219, 220

- Transformation, 173
- Transformationsmatrix, 173
- Transitivität, 79
- Transponierte Matrix, 10
- transponierte Matrix, 158
- Tupel, 29

- überbestimmtes Gl.-system, 21, 189

- Umkehrabbildung, 145
- Unitärer Vektorraum, 110
- unterbestimmtes Gl.-system, 21, 196
- Unterdeterminanten, 169
- Unterraum, 88
- Untervektorraum, 88

- Vektor, 27, 86
 - Addition, 29, 31
 - Einheits-, 37
 - gebunden, 31
 - lokalisiert, 31
 - Multiplikation mit Skalar, 29, 31
 - normiert, 39
 - normierter, 38
 - orthogonal, 39
 - Orts-, 31
 - parallel, 33
 - senkrecht, 39
 - Winkel, 41
- Vektorprodukt, 46, 185
- Vektorraum, 27, 85
 - komplexer, 86
 - reeller, 86
- Vektorraum \mathbb{R}^n , 29
- verallgemeinerte Inverse, 190, 196
- Verschiebungsvektor, 30
- Verzerrung, 176

- Wertebereich, 125
- Wertemenge, 125
- Wertevorrat, 125
- windschief, 59
- Winkel zwischen Vektoren, 41

- Zeilenrang, 154
- Zeilenvektor, 147
- Zielmenge, 125
- Zweiernorm, 37
- Zweipunktgleichung, 50