

Mathematik II für Bauwesen

Ivan Izmetiev

TU Darmstadt, SS 2012

Inhaltsverzeichnis

1	Lineare Algebra	1
1	Lineare Gleichungssysteme und Matrizen	1
1.1	Matrizen	1
1.2	Lineare Gleichungssysteme	1
1.3	Elementare Zeilenumformungen	2
1.4	Gauß-Algorithmus	3
2	Die Matrizenmultiplikation	8
2.1	Addition, Subtraktion und Multiplikation mit einer Zahl	8
2.2	Die Matrizenmultiplikation	9
2.3	Invertierbare Matrizen	12
2.4	Die Transponierte einer Matrix	13
2.5	Diagonal- und Dreiecksmatrizen	15
2.6	Berechnung der inversen Matrix mit dem Gauß-Jordan-Verfahren	16
2.7	Begründung des Gauß-Jordan-Verfahrens	18
3	Vektorräume	21
3.1	Linearkombinationen	21
3.2	Der Rang einer Matrix	23
3.3	Lineare Teilräume von \mathbb{R}^n	25
3.4	Vektorraum: Definition	28
3.5	Basis eines Vektorraums	30
3.6	Dimension eines Vektorraums	32
3.7	Basis und Dimension des Zeilenraums und des Kerns einer Matrix	33
3.8	Der Basisauswahlsatz und der Basisergänzungssatz	34
3.9	Zeilenrang = Spaltenrang	36
3.10	Kriterien der Invertierbarkeit einer Matrix	38
4	Lineare Algebra und Statik der Fachwerke	39
4.1	Gleichgewichtsbelastungen	39
4.2	Auflösen der Belastung durch Stabkräfte	40
4.3	Statische Bestimmtheit, Dimension und Rang	41
4.4	Beispiele von Ausnahmefachwerken	44
5	Determinanten	46
5.1	Die Determinante einer 2×2 -Matrix	46
5.2	Die Determinante einer 3×3 -Matrix	47
5.3	Die Determinante einer $n \times n$ -Matrix	48
5.4	Rechenregeln für Determinanten	50
6	Lineare Abbildungen und Eigenwerte	54

6.1	Definition und Beispiele	54
6.2	Lineare Abbildungen und Matrizen	55
6.3	Längentreue Abbildungen und orthogonale Matrizen	57
6.4	Beispiele von linearen Abbildungen und Abbildungsmatrizen	60
6.5	Abbildungsmatrix bezüglich einer beliebigen Basis	62
6.6	Basiswechsel	63
6.7	Eigenwerte und Eigenvektoren einer linearen Transformation	67
6.8	Diagonalisierbare Matrizen	71
7	Symmetrische Matrizen und quadratische Formen	72
7.1	Diagonalisierbarkeit symmetrischer Matrizen	72
7.2	Quadratische Formen	73
7.3	Die Hauptachsentransformation	75
7.4	Quadriken	76
7.5	Positiv definite Matrizen	78
2	Funktionen in mehreren Variablen: Differentiation	79
1	Grundlagen	79
1.1	Reellwertige und vektorwertige Funktionen mehrerer Veränderlicher	79
1.2	Der Graph, Niveaumengen und “partielle” Funktionen	80
1.3	Teilmengen von \mathbb{R}^n	81
1.4	Grenzwerte und Stetigkeit	82
2	Differentiation	83
2.1	Partielle Ableitungen und der Gradient	83
2.2	Differenzierbarkeit und lineare Approximation	85
2.3	Die Richtungsableitung	87
2.4	Die Kettenregel für reellwertige Funktionen	88
3	Anwendungen der Differentiation	90
3.1	Die Bedeutung des Gradienten	90
3.2	Approximation höherer Ordnung: die Taylor-Formel	93
3.3	Lokale Minima und Maxima	96
3.4	Globale Minima und Maxima	97
3.5	Extremwertaufgaben mit Nebenbedingungen	99
4	Vektorwertige Funktionen	103
4.1	Die Differentiation und die Jacobi-Matrix	103
4.2	Die Kettenregel	104
4.3	Skalaren- und Vektorfelder	105
4.4	Divergenz, Rotation, Laplace-Operator	107
3	Funktionen in mehreren Variablen: Integration	110
1	Parameterintegrale	110
2	Integration über ebene Bereiche	112
2.1	Der Flächeninhalt	112
2.2	Definition und Berechnung des Doppelintegrals	113
2.3	Anwendungen des Doppelintegrals	115
2.4	Die Transformationsformel für Gebietsintegrale	117
3	Kurvenintegrale	119
3.1	Definition	119

	3.2	Die Integration eines Vektorfeldes längs einer Kurve . . .	120
	3.3	Das Potential eines Gradientenfeldes	120
	3.4	Der Satz von Green und der ebene Satz von Gauß	123
4		Integration über Flächen im Raum	125
	4.1	Reguläre Flächen	125
	4.2	Berechnung des Flächeninhaltes	126
	4.3	Das Oberflächenintegral	127
	4.4	Der Divergenzsatz von Gauss	127

Kapitel 1

Lineare Algebra

1 Lineare Gleichungssysteme und Matrizen

1.1 Matrizen

Eine Matrix vom Typ $m \times n$ ist ein rechteckiges Zahlenschema mit m Zeilen und n Spalten:

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

Man beachte: die Zahl a_{ij} steht auf der Kreuzung der i -ten Zeile mit der j -ten Spalte. Abkürzende Schreibweise für eine $m \times n$ Matrix mit den Einträgen a_{ij} :

$$A = (a_{ij})_{m \times n}$$

Ein Vektor $v = (v_1, v_2, \dots, v_n)$ kann als eine $1 \times n$ Matrix gesehen werden (Zeilenvektor) oder, in der Spaltenschreibweise für Vektoren

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

als eine $n \times 1$ Matrix.

1.2 Lineare Gleichungssysteme

Betrachten wir ein lineares Gleichungssystem (LGS) aus m Gleichungen für n Unbekannte:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.1}$$

Die komplette Information über das Gleichungssystem ist in ihrer *erweiterten Koeffizientenmatrix* enthalten:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right)$$

Daher können die Matrizen als bequeme Schreibweise bei der Lösung der LGS benutzt werden.

Zuerst aber noch Einiges über die linearen Gleichungssysteme.

- Ein LGS heißt *homogen*, wenn alle Zahlen b_i auf der rechten Seite gleich Null sind, und andernfalls *inhomogen*.

$$3x_1 + 2x_2 = 0$$

$$3x_1 + x_2 = 0$$

homogen

$$3x_1 + 2x_2 = 1$$

$$3x_1 + x_2 = 5$$

inhomogen

- Ein homogenes LGS hat immer mindestens eine Lösung, nämlich die Nulllösung $x_1 = x_2 = \cdots = x_n = 0$.
- Es gibt inhomogene Systeme, die keine Lösungen besitzen, z. B.

$$3x_1 + 2x_2 = 1$$

$$3x_1 + 2x_2 = 2$$

- Ein LGS kann auch unendlich viele Lösungen besitzen, z. B.

$$3x_1 + 2x_2 = 1 \tag{1.2}$$

Eine allgemeine Lösung von (1.2) kann wie folgt beschrieben werden. Die Variable x_2 darf einen beliebigen Wert $\lambda \in \mathbb{R}$ annehmen, und x_1 kann daraus berechnet werden:

$$x_2 = \lambda, \quad x_1 = \frac{1}{3}(1 - 2\lambda) \tag{1.3}$$

Dementsprechend heißt x_2 *unabhängige (oder freie) Variable*, und x_1 heißt *abhängige Variable*.¹

Im Folgenden wollen wir einen Algorithmus beschreiben, der die Lösungen jedes LGS analog zu (1.3) darstellt.

1.3 Elementare Zeilenumformungen

Betrachten wir die folgenden Umformungen eines LGS:

- 1) Vertauschung zweier Gleichungen.
- 2) Multiplikation einer Gleichung mit einer Zahl $\alpha \neq 0$.

¹Welche Variable unabhängig ist, und welche nicht, hängt von unserem Lösungsweg ab. Im obigen Beispiel könnte man x_1 als unabhängige Variable wählen und x_2 durch $x_1 = \lambda$ ausdrücken.

- 3) Addition (bzw. Subtraktion) des Vielfachen einer Gleichung zu (bzw. von) einer anderen.

Bei diesen Umformungen geht ein LGS in ein äquivalentes über, d. h. die Lösungsmenge ändert sich nicht.

Beispiel 1.1. Auf das LGS

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

wird die dritte Elementarumformung angewendet: von der zweiten Zeile wird die erste subtrahiert:

$$\begin{array}{rcl} 3x_1 + 2x_2 = 1 & \longrightarrow & 3x_1 + 2x_2 = 1 \\ 3x_1 + x_2 = 5 & & -x_2 = 4 \end{array}$$

In der Matrizesprache ausgedrückt, entsprechen die obigen Operationen den *elementaren Zeilenumformungen* der erweiterten Koeffizientenmatrix. Wir bezeichnen mit Z_i die i -te Zeile der Matrix und benutzen die folgenden Abkürzungen für diese Operationen:

- | | |
|---|--------------------------------------|
| 1) Vertauschung zweier Zeilen | $Z_i \leftrightarrow Z_j$ |
| 2) Multiplikation einer Zeile mit einer Zahl $\alpha \neq 0$ | $Z_i \rightarrow \alpha Z_i$ |
| 3) Addition (bzw. Subtraktion) des Vielfachen einer Zeile zu (bzw. von) einer anderen | $Z_i \rightarrow Z_i + \alpha Z_j$ |
| | (zu Z_i wird αZ_j addiert) |

1.4 Gauß-Algorithmus

Mit dem Gauß-Algorithmus (auch als Gaußsches Eliminationsverfahren bekannt) wird ein LGS in eine einfache Form gebracht, aus welcher die Lösungen leicht ermittelt werden können.

Gauß-Algorithmus für ein homogenes LGS

Da alle $b_i = 0$ sind, reicht es, die "einfache" Koeffizientenmatrix zu behandeln:

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

Der Gauß-Algorithmus besteht aus zwei Teilen:

- die Vorwärtselimination;
- die Rückwärtssubstitution.

Die Vorwärtselimination Wenn $a_{11} \neq 0$ ist, dann multiplizieren wir die erste Zeile: $Z_1 \rightarrow \frac{1}{a_{11}}Z_1$, sodass die Matrix die folgende Form annimmt:

$$\begin{pmatrix} 1 & * & \cdots & * \\ * & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ * & * & \cdots & * \end{pmatrix} \quad (1.4)$$

(Das *-Zeichen steht für eine beliebige Zahl.)

Wenn $a_{11} = 0$ ist, dann suchen wir eine Zeile, wo die erste Zahl $\neq 0$ ist und tauschen sie mit der ersten Zeile um: $Z_1 \leftrightarrow Z_i$. Nach der Division der (neuen) ersten Zeile durch ihren ersten Eintrag nimmt die Matrix die Form (1.4) an.

Nun *eliminieren* wir alle übrigen Koeffizienten in der ersten Spalte dadurch, dass wir von allen Zeilen ein Vielfaches der ersten Zeile subtrahieren (bzw. dazu addieren): $Z_i \rightarrow Z_i + \alpha_i Z_1$ mit geeigneten α_i . Die Matrix sieht danach aus wie

$$\begin{pmatrix} 1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & * & \cdots & * \end{pmatrix} \quad (1.5)$$

Damit ist der erste Eliminationsschritt vollendet. Die erste Zeile können wir jetzt außer Acht lassen, sie wird nicht mehr angerührt.

Beim zweiten Eliminationsschritt ist Folgendes möglich.

- Die Zahl an der Stelle (2,2) (2-te Zeile, 2-te Spalte) ist $\neq 0$.

Dann machen wir daraus 1: $Z_2 \rightarrow \frac{1}{a_{22}}Z_2$ und eliminieren alle darunter stehenden Zahlen: $Z_i \rightarrow Z_i + \alpha_i Z_2$. Die Matrix wird zu

$$\begin{pmatrix} 1 & * & * & \cdots & * \\ 0 & 1 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & * & \cdots & * \end{pmatrix} \quad (1.6)$$

- Die Zahl an der Stelle (2,2) ist gleich Null, in der Spalte darunter gibt es eine Zahl $\neq 0$.

Wir tauschen die Zeilen: $Z_2 \leftrightarrow Z_i$, multiplizieren die zweite: $Z_2 \rightarrow \frac{1}{a_{22}}Z_2$, eliminieren: $Z_i \rightarrow Z_i + \alpha_i Z_2$, und die Matrix nimmt die Form (1.6) an.

- Die Zahl an der Stelle (2,2) ist gleich Null, sowie alle Zahlen darunter. In diesem Fall verlassen wir die zweite Spalte und bearbeiten die dritte nach demselben Muster. Wenn es in der dritten Spalte eine Zahl $\neq 0$ gibt, dann nimmt die Matrix nach der Elimination die folgende Form an.

$$\begin{pmatrix} 1 & * & * & \cdots & * \\ 0 & 0 & 1 & \cdots & * \\ 0 & 0 & 0 & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix}$$

Nach höchstens $m - 1$ Eliminationsschritten gelangen wir zu einer Matrix in der *Zeilenstufenform*, wie zum Beispiel

$$\begin{pmatrix} 1 & * & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (1.7)$$

Bemerkung 1.2. Ein Eliminationsschritt kann in drei Teile unterteilt werden: "Eliminator" suchen (eine Zahl $\neq 0$ in der aktuellen Spalte) und nach oben bringen; normieren (die Zeile multiplizieren, so dass Eliminator = 1 wird); eliminieren. Bei einer Variante des Algorithmus wird der zweite Schritt ausgelassen. Die Matrix wird dann in die folgende Form gebracht:

$$\begin{pmatrix} \square & * & * & * & * & * & * & * \\ 0 & 0 & \square & * & * & * & * & * \\ 0 & 0 & 0 & \square & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \square & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

wobei \square für eine Zahl $\neq 0$ steht. Diese Form darf ebenfalls Zeilenstufenform genannt werden.

Die Rückwärtssubstitution Das wird auf einem Beispiel erläutert. Es sei durch die Vorwärtselimination die folgende Matrix entstanden:

$$\begin{pmatrix} \boxed{1} & -2 & 3 & 4 & 2 \\ 0 & 0 & \boxed{2} & 1 & -4 \\ 0 & 0 & 0 & \boxed{-1} & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Das entsprechende Gleichungssystem ist

$$\begin{aligned} x_1 - 2x_2 + 3x_3 + 4x_4 + 2x_5 &= 0 \\ 2x_3 + x_4 - 4x_5 &= 0 \\ -x_4 + 3x_5 &= 0 \end{aligned}$$

Die Variablen x_1, x_3, x_4 erklären wir zu den abhängigen. Den unabhängigen Variablen x_2 und x_5 schreiben wir beliebige Werte zu:

$$x_2 = \lambda_1, \quad x_5 = \lambda_2$$

und bringen sie auf die rechten Seiten der Gleichungen:

$$\begin{aligned} x_1 + 3x_3 + 4x_4 &= 2\lambda_1 - 2\lambda_2 \\ 2x_3 + x_4 &= 4\lambda_2 \\ -x_4 &= -3\lambda_2 \end{aligned}$$

Jetzt erfolgt die *Rückwärtssubstitution*. Die letzte Gleichung drückt x_4 durch λ_2 aus, das Einsetzen in die zweite Gleichung drückt x_3 durch λ_2 , und schließlich das Einsetzen in die erste Gleichung drückt x_1 durch λ_1 und λ_2 aus:

$$x_4 = 3\lambda_2, \quad x_3 = 0.5\lambda_2, \quad x_1 = 2\lambda_1 - 15.5\lambda_2$$

Die allgemeine Lösung kann wie folgt aufgeschrieben werden:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2\lambda_1 - 15.5\lambda_2 \\ \lambda_1 \\ 0.5\lambda_2 \\ 3\lambda_2 \\ \lambda_2 \end{pmatrix}$$

Das Prinzip der Rückwärtssubstitution:

- Die Spalten mit \square -Zeichen entsprechen den abhängigen Variablen.
- Die unabhängigen Variablen (also die aus den Spalten ohne \square -Zeichen) werden auf die rechten Seiten der Gleichungen gebracht, und werden gleich $\lambda_1, \lambda_2, \dots$ gesetzt.
- Das entstandene Gleichungssystem wird durch Einsetzen von unten nach oben gelöst.

Beispiel 1.3. Lösen wir das LGS

$$\begin{aligned} x_1 - 4x_2 + 2x_3 &= 0 \\ 2x_1 - 3x_2 - x_3 - 5x_4 &= 0 \\ 3x_1 - 7x_2 + x_3 - 5x_4 &= 0 \\ x_2 - x_3 - x_4 &= 0 \end{aligned}$$

Mit der Elimination bringen wir die Koeffizientenmatrix in die Zeilenstufenform:

$$\begin{aligned} \begin{pmatrix} 1 & -4 & 2 & 0 \\ 2 & -3 & -1 & -5 \\ 3 & -7 & 1 & -5 \\ 0 & 1 & -1 & -1 \end{pmatrix} &\longrightarrow \begin{pmatrix} 1 & -4 & 2 & 0 \\ 0 & 5 & -5 & -5 \\ 0 & 5 & -5 & -5 \\ 0 & 1 & -1 & -1 \end{pmatrix} \longrightarrow \\ &\longrightarrow \begin{pmatrix} 1 & -4 & 2 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -4 & 2 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Die Unbekannten x_3 und x_4 sind frei: $x_3 = \lambda_1$, $x_4 = \lambda_2$. Wir bringen sie auf die rechte Seite und erhalten das Gleichungssystem

$$\begin{aligned} x_1 - 4x_2 &= -2\lambda_1 \\ x_2 &= \lambda_1 + \lambda_2 \end{aligned}$$

Daraus errechnen wir x_1 und schreiben die allgemeine Lösung auf:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2\lambda_1 + 4\lambda_2 \\ \lambda_1 + \lambda_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Gauß-Algorithmus für ein inhomogenes LGS

In diesem Fall besteht der Gauß-Algorithmus aus drei Teilen:

- die Vorwärtselemination;
- die Lösbarkeitsentscheidung;
- die Rückwärtssubstitution.

Die Vorwärtselemination Die erweiterte Koeffizientenmatrix wird in Zeilenstufenform gebracht, bis auf die letzte Spalte:

$$\left(\begin{array}{cccccccc|c} \square & * & * & * & * & * & * & * & c_1 \\ 0 & 0 & \square & * & * & * & * & * & c_2 \\ 0 & 0 & 0 & \square & * & * & * & * & c_3 \\ 0 & 0 & 0 & 0 & 0 & \square & * & * & c_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_6 \end{array} \right)$$

Die Lösbarkeitsentscheidung Ist eine der Zahlen c_5, c_6 im obigen Beispiel von Null verschieden, so entspricht diese Zeile einer Gleichung $0 = c_i$, die keine Lösung besitzt. Gilt $c_5 = c_6 = 0$, so berechnen wir die Lösungen mittels

Rückwärtssubstitution Das funktioniert genau so wie im homogenen Fall: die freien Variablen werden auf die rechte Seite gebracht, zu den c_i ; die Ausdrücke für die abhängigen Variablen werden von unten nach oben durch Einsetzen ermittelt.

Beispiel 1.4. Bei welchem Wert von c schneiden sich die drei Ebenen in \mathbb{R}^3 :

$$x - 2y - z = 3, \quad 2x - 5y + 2z = 2, \quad 4x - 9y = c?$$

Wie sieht bei diesem c die Schnittmenge aus?

Lösung. Der Schnitt dieser Ebenen ist die Lösungsmenge des LGS

$$\begin{aligned} x - 2y - z &= 3 \\ 2x - 5y + 2z &= 2 \\ 4x - 9y &= c \end{aligned}$$

Mit der Elimination bringen wir die erweiterte Matrix des Systems in Zeilenstufenform:

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 3 \\ 2 & -5 & 2 & 2 \\ 4 & -9 & 0 & c \end{array} \right) \longrightarrow \left(\begin{array}{ccc|c} 1 & -2 & -1 & 3 \\ 0 & -1 & 4 & -4 \\ 0 & -1 & 4 & c-12 \end{array} \right) \longrightarrow \left(\begin{array}{ccc|c} 1 & -2 & -1 & 3 \\ 0 & -1 & 4 & -4 \\ 0 & 0 & 0 & c-8 \end{array} \right)$$

Bei $c \neq 8$ besitzt das System keine Lösungen, d. h. die Ebenen schneiden sich nicht. Bei $c = 8$ setzen wir $z = \lambda$ (die freie Variable) und erhalten das System

$$\begin{aligned} x - 2y &= \lambda + 3 \\ -y &= -4\lambda - 4 \end{aligned}$$

Die Substitution ergibt

$$y = 4\lambda + 4, \quad x = 9\lambda + 11$$

Die allgemeine Lösung:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 9\lambda + 11 \\ 4\lambda + 4 \\ \lambda \end{pmatrix} = \begin{pmatrix} 11 \\ 4 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 9 \\ 4 \\ 1 \end{pmatrix}$$

Das heißt, bei $c = 8$ ist die Schnittmenge der Ebenen die Gerade mit Aufpunkt $(11, 4, 0)$ und Richtungsvektor $(9, 4, 1)$. \square

2 Die Matrizenmultiplikation

2.1 Addition, Subtraktion und Multiplikation mit einer Zahl

Seien $A = (a_{ij})_{m \times n}$ und $B = (b_{ij})_{m \times n}$ zwei $m \times n$ Matrizen. Ihre Summe $A + B$ ist definiert als

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

Für jede reelle Zahl $\lambda \in \mathbb{R}$ ist das Produkt λA definiert als

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

Für die $n \times 1$ Matrizen und die $1 \times n$ Matrizen sind es die üblichen Operationen mit den Spaltenvektoren, beziehungsweise Zeilenvektoren:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix} \quad \lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix}$$

$$\begin{aligned} (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) &= (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \\ \lambda(a_1, a_2, \dots, a_n) &= (\lambda a_1, \lambda a_2, \dots, \lambda a_n) \end{aligned}$$

Das Produkt einer Matrix A mit -1 wird mit $-A$ bezeichnet:

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

Die Differenz zweier Matrizen wird definiert als:

$$A - B := A + (-B) = (a_{ij} - b_{ij})_{m \times n}$$

Die *Nullmatrix*

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

spielt dieselbe Rolle, wie 0 bei den reellen Zahlen:

$$\text{es gilt } A + \mathbf{0} = A,$$

wobei A eine beliebige $m \times n$ Matrix ist, und $\mathbf{0}$ die $m \times n$ Nullmatrix.

2.2 Die Matrizenmultiplikation

Zeile mal Spalte

Es seien

$$a = (a_1, a_2, \dots, a_n), \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

ein Zeilenvektor und ein Spaltenvektor der gleichen Länge (Höhe). Ihr Produkt wird definiert als

$$(a_1, a_2, \dots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} := a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

Bemerkung 2.1. Bei $n = 2$ und $n = 3$ erkennt man hier das Skalarprodukt in \mathbb{R}^2 , bzw. in \mathbb{R}^3 .

Beispiel 2.2.

$$(2, 3, 0, 5, 1) \begin{pmatrix} 0 \\ 1 \\ 4 \\ 1 \\ 0 \end{pmatrix} = 2 \cdot 0 + 3 \cdot 1 + 0 \cdot 4 + 5 \cdot 1 + 1 \cdot 0 = 8$$

Jetzt können wir das Produkt zweier Matrizen definieren.

Definition 2.3. *Es seien*

$$A = (a_{ij})_{m \times n} \text{ und } B = (b_{ij})_{n \times p}$$

eine $m \times n$ und eine $n \times p$ Matrix. Das Produkt $C = AB$ ist eine $m \times p$ Matrix definiert durch

$$C = (c_{ij})_{m \times p}, \quad c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{in} b_{nj}$$

In Worten:

Um die Zahl in der i -ten Zeile und j -ten Spalte von AB zu berechnen, multipliziert man die i -te Zeile von A mit der j -ten Spalte von B .

Beispiel 2.4.

$$\begin{pmatrix} 4 & 3 & 0 & 1 \\ 2 & 1 & 4 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 21 & 9 \\ 14 & 8 & 6 \end{pmatrix}$$

Insbesondere, die zweite Zeile von A mal die erste Spalte von B ergibt $2 \cdot 1 + 1 \cdot 0 + 4 \cdot 3 + 0 \cdot 1 = 14$, die Zahl in der zweiten Zeile, ersten Spalte von AB :

$$\begin{pmatrix} 4 & 3 & 0 & 1 \\ \boxed{2} & 1 & 4 & 0 \end{pmatrix} \begin{pmatrix} \boxed{1} & 2 & 0 \\ 0 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 21 & 9 \\ \boxed{14} & 8 & 6 \end{pmatrix}$$

Beachte: AB ist nur erklärt wenn die Zeilen von A dieselbe Länge haben, wie die Spalten von B . Das Produkt

$$\begin{pmatrix} 2 & 1 & 0 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 0 & -1 & 6 \\ 2 & -4 & 5 \end{pmatrix} \text{ ist nicht definiert!}$$

Beispiel 2.5. Für die Matrizen

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

kann man die Produkte AB und BA bilden:

$$AB = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ -1 & -4 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 26 & 15 \end{pmatrix}$$

$$BA = \begin{pmatrix} 0 & 2 \\ -1 & -4 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 6 & 8 & 10 \\ -14 & -17 & -20 \\ 27 & 26 & 25 \end{pmatrix}$$

Beispiel 2.6. Wenn man eine $n \times 1$ Matrix (Spaltenvektor) mit einer $1 \times n$ Matrix (Zeilenvektor) multipliziert, dann erhält man eine $n \times n$ Matrix:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (1 \quad 2 \quad 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$$

Matrizenmultiplikation und lineare Gleichungssysteme

Sei A eine $m \times n$ -Matrix, und sei $x \in \mathbb{R}^n$ ein Spaltenvektor der Höhe n . Das Produkt Ax ist dann ein Spaltenvektor der Höhe m , und zwar:

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

Darin erkennen wir die linken Seiten des linearen Gleichungssystems aus dem Abschnitt 1.2. Und zwar, ein LGS mit Koeffizientenmatrix A kann als

$$Ax = b$$

aufgeschrieben werden, wobei $b \in \mathbb{R}^m$ ein Spaltenvektor ist.

Rechenregeln

Definition 2.7. Die Matrix

$$E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

heißt $n \times n$ -Einheitsmatrix.

Satz 1. Für alle $m \times n$ -Matrizen A, A_1, A_2 , $n \times p$ -Matrizen B, B_1, B_2 und jede $p \times q$ Matrix C gilt:

- $(A_1 + A_2)B = A_1B + A_2B$, $A(B_1 + B_2) = AB_1 + AB_2$
- $(AB)C = A(BC)$
- $E_m A = A = A E_n$
- Im Allgemeinen, $AB \neq BA$. Man sagt dafür, die Matrizenmultiplikation ist nicht kommutativ.

Beweis. a) ist klar und kann einfach nachgerechnet werden. b), mit etwas mehr Mühe, auch.

Zum c): um das (i, j) -Element im Produkt $E_m A$ zu berechnen, multiplizieren wir die i -te Zeile von E_m mit der j -ten Spalte von A . In der i -ten Zeile von E_m steht 1 auf der i -ten Stelle, alle anderen Elemente sind 0. Das Auflegen der i -ten Zeile von E_m auf die j -te Spalte von A "wählt" also das i -te Element aus dieser Spalte, und das ist a_{ij} . Das zeigt $E_m A = A$. Analog wird $A E_n = A$ bewiesen

Für d) genügt es, ein Gegenbeispiel zu geben. Es sei

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (1.8)$$

Dann gilt

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Beachte außerdem, dass für eine $m \times n$ -Matrix A und eine $n \times p$ -Matrix B das Produkt BA nur dann definiert ist, wenn $m = p$ gilt. \square

Bemerkung 2.8. Auf den Matrizen (1.8) kann man weitere Unterschiede zwischen der Matrizenmultiplikation und der Multiplikation reeller Zahlen demonstrieren:

- Es gilt $AB = \mathbf{0}$, während $A \neq \mathbf{0}$ und $B \neq \mathbf{0}$.
- Es gilt $A^2 = \mathbf{0}$, während $A \neq \mathbf{0}$.
- Es gilt $BA = A$, während $B \neq E_2$ und $A \neq \mathbf{0}$.

2.3 Invertierbare Matrizen

In diesem Abschnitt betrachten wir nur quadratische $n \times n$ -Matrizen und bezeichnen $E := E_n$.

Definition 2.9. Eine $n \times n$ -Matrix A heißt invertierbar, wenn es eine $n \times n$ -Matrix B gibt, so dass

$$AB = E = BA$$

In diesem Fall ist die Matrix B eindeutig bestimmt, wird mit A^{-1} bezeichnet, und heißt inverse Matrix von A .

Beweisen wir die Eindeutigkeit der Inversen:

Lemma 1. Gilt $AB = E = BA$ und $AC = E = CA$, so gilt $B = C$.

Beweis. Einerseits gilt

$$BAC = (BA)C = EC = C$$

Andererseits,

$$BAC = B(AC) = BE = B$$

Folglich, $B = C$. □

Lemma 2. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad - bc \neq 0$. Dann ist A invertierbar, und zwar

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Beweis. Setze $B := \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Dann gilt

$$AB = \frac{1}{ad - bc} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = E$$

$$BA = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = E$$

□

Bemerkung 2.10. • Wie wir später feststellen werden, sind 2×2 -Matrizen mit $ad - bc = 0$ nicht invertierbar. Dies ist noch ein Unterschied zu reellen Zahlen, wo nur die Null keine Inverse besitzt.

- Bald werden wir auch sehen, dass aus $BA = E$ auch $AB = E$ folgt (allerdings unter der Voraussetzung, dass A und B quadratische Matrizen sind!). Das heißt, um festzustellen, dass B die Inverse zu A ist, genügt es nur eine der Gleichungen $AB = E$, $BA = E$ zu überprüfen.

Rechenregeln

Satz 2. a) Die Inverse einer invertierbarer Matrix ist invertierbar, und zwar $(A^{-1})^{-1} = A$.

b) Das Produkt zweier invertierbaren Matrizen ist ebenfalls invertierbar, und zwar:

$$(AB)^{-1} = B^{-1}A^{-1} \quad (1.9)$$

Beweis. Zum a): sei A invertierbar, bezeichne A^{-1} mit B . Per Definition $AB = E = BA$. Diese zwei Gleichungen können nach derselben Definition auch als $A = B^{-1}$ interpretiert werden. Also gilt $A = (A^{-1})^{-1}$.

b) wird durch Umformen bewiesen:

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E \\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E\end{aligned}$$

□

Bemerkung 2.11. Der Grund, warum $(AB)^{-1} = A^{-1}B^{-1}$ nicht immer gilt, liegt darin, dass die Matrizenmultiplikation nichtkommutativ ist. Die folgende Erklärung von $AB \neq BA$ und von (1.9) sieht als Witz aus, hat aber einen wahren Kern:²

Sei $A =$ "Socke anziehen", $B =$ "Schuh anziehen". Dann

$AB =$ "erst die Socke anziehen, dann den Schuh"

$BA =$ "erst den Schuh anziehen, dann die Socke" $\neq AB$

Außerdem gilt

$$(AB)^{-1} = AB \text{ rückgängig machen} =$$

$$\text{"erst den Schuh ausziehen, dann die Socke"} = B^{-1}A^{-1}$$

2.4 Die Transponierte einer Matrix

Jeder $m \times n$ -Matrix A ist die $n \times m$ -Matrix A^\top zugeordnet, die durch Spiegelung von A in der Geraden durch die Elemente (a_{ij}) entsteht.

Definition 2.12. Die Transponierte von $A = (a_{ij})_{m \times n}$ ist die Matrix $B = (b_{ij})_{n \times m}$ mit $b_{ij} = a_{ji}$.

Nochmal anders ausgedrückt, aus den Zeilen von A werden Spalten von A^\top .

Beispiel 2.13.

$$\begin{pmatrix} 0 & -1 & 6 \\ 2 & -4 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1 & -4 \\ 6 & 5 \end{pmatrix} \quad (x_1, x_2, \dots, x_n)^\top = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

²Aus dem Buch von Coxeter "Introduction to Geometry". Was die Socken... , nein, was die Matrizen mit Geometrie zu tun haben, wird in einem der folgenden Abschnitte erklärt.

Man benutzt $(x_1, x_2, \dots, x_n)^\top$ häufig als platzsparende Schreibweise für Spaltenvektoren. Zum Beispiel, das Skalarprodukt zweier Spaltenvektoren $x = (x_1, x_2, \dots, x_n)^\top$ und $y = (y_1, y_2, \dots, y_n)^\top$ kann wie folgt geschrieben werden:

$$\langle x, y \rangle = \sum_i x_i y_i = x^\top y$$

Rechenregeln

Satz 3. Für alle $\lambda \in \mathbb{R}$, alle $m \times n$ -Matrizen A, A_1, A_2 und $n \times p$ Matrizen B gilt

$$a) \quad (A_1 + A_2)^\top = A_1^\top + A_2^\top, \quad (\lambda A)^\top = \lambda A^\top$$

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

$$c) \quad (AB)^\top = B^\top A^\top$$

Beweis. Punkte a) und b) sind klar.

Zum c): Man bemerke, dass $(AB)^\top = A^\top B^\top$ kann schon deswegen nicht wahr sein, weil A^\top eine $n \times m$ -, und B^\top eine $p \times n$ -Matrix ist, und sie können in dieser Reihenfolge nicht multipliziert werden, wenn $m \neq p$ ist. Um $(AB)^\top = B^\top A^\top$ zu beweisen, bezeichnen wir mit $Z_i = (a_{i1}, a_{i2}, \dots, a_{in})$ die i -te Zeile von A , und mit S_j die j -te Spalte von B :

$$A = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \end{pmatrix}, \quad B = (S_1, S_2, \dots, S_p)$$

Nach dem Multiplikationsgesetz,

$$AB = \begin{pmatrix} Z_1 S_1 & Z_1 S_2 & \dots & Z_1 S_p \\ Z_2 S_1 & Z_2 S_2 & \dots & Z_2 S_p \\ \dots & \dots & \dots & \dots \\ Z_m S_1 & Z_m S_2 & \dots & Z_m S_p \end{pmatrix}$$

Andererseits,

$$A^\top = (Z_1^\top, Z_2^\top, \dots, Z_m^\top), \quad B^\top = \begin{pmatrix} S_1^\top \\ S_2^\top \\ \vdots \\ S_p^\top \end{pmatrix}$$

(aus Zeilen von A werden Spalten von A^\top , und aus Spalten von B Zeilen von B^\top), und folglich

$$B^\top A^\top = \begin{pmatrix} S_1^\top Z_1^\top & S_1^\top Z_2^\top & \dots & S_1^\top Z_m^\top \\ S_2^\top Z_1^\top & S_2^\top Z_2^\top & \dots & S_2^\top Z_m^\top \\ \dots & \dots & \dots & \dots \\ S_p^\top Z_1^\top & S_p^\top Z_2^\top & \dots & S_p^\top Z_m^\top \end{pmatrix}$$

Wegen $Z_i S_j = S_j^\top Z_i^\top$ folgt $(AB)^\top = B^\top A^\top$. □

Satz 4. Ist A eine invertierbare $n \times n$ -Matrix, so ist A^\top ebenfalls invertierbar, und es gilt

$$(A^\top)^{-1} = (A^{-1})^\top$$

Beweis. Aus

$$\begin{aligned} A^\top(A^{-1})^\top &= (A^{-1}A)^\top = E^\top = E \text{ und} \\ (A^{-1})^\top A^\top &= (AA^{-1})^\top = E^\top = E \end{aligned}$$

folgt, dass $(A^{-1})^\top$ die zu A^\top Inverse ist. \square

Symmetrische und schiefsymmetrische Matrizen

Definition 2.14. Eine $n \times n$ -Matrix A heißt symmetrisch, wenn $A^\top = A$ gilt, und schiefsymmetrisch falls $A^\top = -A$.

Mit anderen Worten,

$$A \text{ symmetrisch} \Leftrightarrow a_{ij} = a_{ji} \text{ für alle } i, j$$

$$A \text{ schiefsymmetrisch} \Leftrightarrow a_{ij} = -a_{ji} \text{ für alle } i, j$$

Insbesondere, bei $i = j$ gilt $a_{ii} = -a_{ii}$, und folglich $a_{ii} = 0$: alle Diagonalelemente einer schiefsymmetrischen Matrix sind gleich Null.

Beispiel 2.15. $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 5 \end{pmatrix}$ symmetrisch $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}$ schiefsymmetrisch

2.5 Diagonal- und Dreiecksmatrizen

Hier betrachten wir nur quadratische Matrizen.

Definition 2.16. Eine $n \times n$ -Matrix wird Diagonalmatrix genannt, wenn alle ihre Elemente außerhalb der Diagonale gleich Null sind:

$$A \text{ ist eine Diagonalmatrix} \Leftrightarrow a_{ij} = 0 \text{ bei } i \neq j$$

Für Diagonalmatrizen wird die folgende Bezeichnung benutzt:

$$\text{Diag}(\alpha_1, \alpha_2, \dots, \alpha_n) := \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_n \end{pmatrix}$$

Diagonalmatrizen sind besonders leicht miteinander zu multiplizieren:

$$\text{Diag}(\alpha_1, \alpha_2, \dots, \alpha_n) \cdot \text{Diag}(\beta_1, \beta_2, \dots, \beta_n) = \text{Diag}(\alpha_1\beta_1, \alpha_2\beta_2, \dots, \alpha_n\beta_n)$$

Sind alle $\alpha_i \neq 0$, dann ist die zugehörige Diagonalmatrix invertierbar, und es gilt

$$\text{Diag}(\alpha_1, \alpha_2, \dots, \alpha_n)^{-1} = \text{Diag}\left(\frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \dots, \frac{1}{\alpha_n}\right)$$

Definition 2.17. Eine $n \times n$ -Matrix wird obere, bzw. untere Dreiecksmatrix genannt, wenn alle ihre Elemente unterhalb, bzw. oberhalb der Diagonale gleich Null sind:

A ist eine obere Dreiecksmatrix $\Leftrightarrow a_{ij} = 0$ bei $i \geq j$

A ist eine untere Dreiecksmatrix $\Leftrightarrow a_{ij} = 0$ bei $i \leq j$

Beispiel 2.18.

$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 5 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 & 0 \\ 4 & 1 & 0 \\ -2 & 3 & 1 \end{pmatrix}$
obere Dreiecksmatrix	untere Dreiecksmatrix

Lemma 3. Das Produkt zweier oberer Dreiecksmatrizen ist eine obere Dreiecksmatrix; das Produkt zweier unterer Dreiecksmatrizen ist eine untere Dreiecksmatrix.

Beweis. Seien A und B obere Dreiecksmatrizen, sei $C = AB$. Es soll gezeigt werden, dass $c_{ij} = 0$ bei $i > j$ gilt. Per Definition,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$$

Es gilt aber $a_{i1} = a_{i2} = \dots = a_{i,i-1} = 0$ und $b_{j+1,j} = \dots = b_{nj} = 0$. Deswegen ist in jedem Produkt $a_{ik}b_{kj}$ mindestens ein Faktor gleich Null. Also gilt $c_{ij} = 0$.

Anschaulich: beim Auflegen der i -ten Zeile von A auf die j -te Spalte von B überlappen sich die Abschnitte aus den Nullen (oder zumindest bedecken die ganze Spalte). □

2.6 Berechnung der inversen Matrix mit dem Gauß-Jordan-Verfahren

Durch elementare Zeilenumformungen kann man feststellen, ob eine gegebene $n \times n$ -Matrix A invertierbar ist, und gegebenenfalls die Inverse finden.

Wir illustrieren das Verfahren auf zwei Beispielen:

$$A_1 = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Das Gauß-Jordan-Verfahren

Als erstes bildet man die $n \times 2n$ -Matrix $(A|E)$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array} \right) \quad \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 1 \end{array} \right)$$

Vorwärtselimination In diesem Schritt wendet man auf die Matrix $(A|E)$ elementare Zeilenumformungen an, bis ihre linke Hälfte Zeilenstufenform annimmt:

$$(A \mid E) \xrightarrow[\text{Zeilenumformungen}]{\text{elementare}} (S \mid P) \quad (1.10)$$

In unseren Beispielen sieht es wie folgt aus:

$$\begin{array}{cc}
 \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array} \right) & \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 1 \end{array} \right) \\
 Z_2 \rightarrow Z_2 + Z_1 & Z_2 \leftrightarrow Z_2 + Z_1; \quad Z_3 \rightarrow Z_3 + Z_1 \\
 \\
 \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array} \right) & \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 & 0 & 1 \end{array} \right) \\
 Z_3 \rightarrow Z_3 + Z_2 & Z_2 \leftrightarrow Z_3 \\
 \\
 \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array} \right) & \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 & 1 \\ 0 & 0 & -2 & 1 & 1 & 0 \end{array} \right) \\
 & Z_2 \rightarrow -\frac{1}{2}Z_2; \quad Z_3 \rightarrow -\frac{1}{2}Z_3 \\
 & \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \end{array} \right)
 \end{array}$$

Invertierbarkeitsentscheidung Enthält die Matrix S in (1.10) eine Null-Zeile, so ist die ursprüngliche Matrix A nicht invertierbar. Das ist der Fall mit der Matrix A_1 in unserem Beispiel.

Enthält die Matrix S keine Null-Zeile, so ist sie eine obere Dreiecksmatrix mit allen Diagonalelementen gleich 1:

$$S = \begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & * \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

In diesem Fall ist die ursprüngliche Matrix A invertierbar, und wir gehen zum nächsten Schritt über.

Rückwärtssubstitution In diesem Schritt wird die linke Hälfte zur Einheitsmatrix umgeformt:

$$(S \mid P) \xrightarrow[\text{Zeilenumformungen}]{\text{elementare}} (E \mid B)$$

Um dies zu erreichen, subtrahieren wir die Vielfachen der letzten Zeile aus den vorherigen Zeilen, so dass in der n -ten Spalte sich Nullen über 1 bilden; danach subtrahieren wir die Vielfachen der vorletzten Zeile aus allen vorherigen,

usw. Auf unserem Beispiel:

$$\left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \end{array} \right)$$

$$Z_1 \rightarrow Z_1 + Z_3$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \end{array} \right)$$

$$Z_1 \rightarrow Z_1 + Z_2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \end{array} \right)$$

Satz 5. Die beim Gauß-Jordan-Verfahren in der rechten Hälfte entstehende Matrix B :

$$(A \mid E) \longrightarrow (S \mid P) \longrightarrow (E \mid B)$$

ist die inverse Matrix von A :

$$B = A^{-1}$$

Diesen Satz beweisen wir im nächsten Abschnitt. Jetzt überzeugen wir uns, dass die Matrix in unserem Beispiel invers zu A_2 ist:

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

2.7 Begründung des Gauß-Jordan-Verfahrens

Um den Satz 5 zu beweisen, stellen wir eine Verbindung zwischen elementaren Zeilenumformungen und Matrizenmultiplikation her.

Elementare Zeilenumformungen und Elementarmatrizen

Definition 2.19. Eine $n \times n$ -Matrix \tilde{E} heißt Elementarmatrix, wenn sie aus der $n \times n$ -Matrix durch eine elementare Zeilenumformung hervorgeht. Wir sagen, \tilde{E} gehört zu dieser Umformung.

Zum Beispiel:

$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ gehört zur Vertauschung der ersten und der zweiten Zeilen;

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ gehört zur Division der zweiten Zeile durch 2;

$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ gehört zur Subtraktion des Zweifachen der dritten Zeile aus der ersten

Lemma 4. Sei A eine $n \times p$ -Matrix. Entsteht \tilde{A} aus A durch eine elementare Zeilenumformung, so gilt

$$\tilde{A} = \tilde{E}A$$

mit der zugehörigen Elementarmatrix \tilde{E} .

Beweis. Wir illustrieren es auf einem Beispiel:

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} a - 2e & b - 2f \\ c & d \\ e & f \end{pmatrix}$$

(Das Auflegen der ersten Zeile der Matrix \tilde{E} auf die Spalten von A subtrahiert aus der ersten Zeile die zweifache dritte; die zweite und die dritte Zeilen von \tilde{E} sind gleich denen der Einheitsmatrix, deswegen bleiben die zweite und die dritte Zeile von A erhalten.) \square

Beweis des Satzes 5

Lemma 5. Für die beim Gauß-Jordan-Verfahren entstehende Matrix B :

$$(A \mid E) \longrightarrow (E \mid B)$$

gilt $BA = E$.

Beweis. Werden auf A elementare Zeilenumformungen angewendet, so transformiert sie sich gemäß

$$A \rightarrow \tilde{E}_1 A \rightarrow \tilde{E}_2 \tilde{E}_1 A \rightarrow \tilde{E}_3 \tilde{E}_2 \tilde{E}_1 A \rightarrow \dots$$

Dementsprechend sieht das Gauß-Jordan-Verfahren wie folgt aus:

$$(A \mid E) \rightarrow (\tilde{E}_1 A \mid \tilde{E}_1) \rightarrow (\tilde{E}_2 \tilde{E}_1 A \mid \tilde{E}_2 \tilde{E}_1) \rightarrow \dots \rightarrow (E \mid B)$$

Es gilt also (s bezeichne die benötigte Anzahl der elementaren Zeilenumformungen):

$$(\tilde{E}_s \cdots \tilde{E}_1)A = E \quad \tilde{E}_s \cdots \tilde{E}_1 = B \quad (1.11)$$

woraus folgt $BA = E$. \square

Lemma 6. Sei $B = \tilde{E}_s \cdots \tilde{E}_1$ Produkt einer Anzahl der Elementarmatrizen. Dann ist B invertierbar.

Beweis. Jede Elementarmatrix ist invertierbar, weil jede elementare Zeilenumformung eine inverse hat. Zum Beispiel, die zu $Z_1 \rightarrow Z_1 - 2Z_3$ inverse Umformung ist $Z_1 + 2Z_3$, und dementsprechend

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_3 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ein Produkt invertierbarer Matrizen ist aber auch invertierbar, und zwar

$$B^{-1} = (\tilde{E}_s \dots \tilde{E}_1)^{-1} = \tilde{E}_1^{-1} \dots \tilde{E}_s^{-1}$$

□

Beweis des Satzes 5. Wir müssen zeigen, dass A invertierbar ist und $A^{-1} = B$ gilt. Nach Lemma 5 gilt $BA = E$ und nach Lemma 6 gibt es eine Matrix C , sodass $CB = E = BC$. Daraus folgt

$$\begin{aligned} CBA &= (CB)A = EA = A \\ CBA &= C(BA) = CE = C \end{aligned}$$

Folglich ist A die Inverse zu B , und das heißt, B ist die Inverse zu A (siehe Satz 2). □

Begründung der Invertierbarkeitsentscheidung

Wir haben gezeigt, dass beim Gauß-Jordan-Verfahren wirklich die inverse Matrix entsteht. Was noch unbegründet bleibt ist warum quadratische Matrizen, die in Zeilenstufenform eine Nullzeile enthalten, nicht invertierbar sind.

Lemma 7. *Sei S eine $n \times n$ -Matrix, die Nullzeile enthält. Dann ist S nicht invertierbar.*

Beweis. Wenn die i -te Zeile von S nur aus Nullen besteht, dann besteht auch die i -te Zeile von ST nur aus Nullen, für jede Matrix T . Deswegen kann ST nicht gleich der Einheitsmatrix sein, d. h. S ist nicht invertierbar. □

Lemma 8. *Wenn nach der Vorwärtselimination $(A | E) \rightarrow (S | P)$ die Matrix S eine Nullzeile enthält, dann ist die Matrix A nicht invertierbar.*

Beweis. Bei der Vorwärtselimination wird die Matrix A von links mit Elementarmatrizen multipliziert:

$$S = \tilde{E}_t \dots \tilde{E}_2 \tilde{E}_1 A = PA$$

Die Matrix P , als Produkt von Elementarmatrizen, ist invertierbar. Wenn die Matrix A invertierbar wäre, dann wäre auch $S = PA$ invertierbar. Nach dem vorherigen Lemma ist das nicht der Fall, also ist A nicht invertierbar. □

Korollar 1. *Jede invertierbare Matrix kann als Produkt von Elementarmatrizen dargestellt werden.*

Beweis. Wenn A invertierbar ist, dann kann ihre Inverse mit dem Gauß-Jordan-Verfahren berechnet werden. Bei diesem Verfahren entsteht die Inverse als Produkt von Elementarmatrizen, siehe (1.11). □

Das Lösen der LGS mit dem Gauß-Jordan-Verfahren

Der letzte Schritt des Gauß-Jordan-Verfahrens heißt Rückwärtssubstitution, genau wie beim Gauß-Verfahren zum Lösen eines LGS. In der Tat kann dieser Schritt auch beim Lösen eines LGS über Matrizen und nicht über lineare Gleichungen ausgeführt werden.

Beim Lösen eines LGS mit dem *Gauß-Jordan-Verfahren* wird die erweiterte $m \times (n + 1)$ -Matrix $(A | b)$ in die folgende Form gebracht:

$$\left(\begin{array}{cccccccc|c} \square & * & 0 & 0 & * & 0 & * & 0 & * \\ 0 & 0 & \square & 0 & * & 0 & * & 0 & * \\ 0 & 0 & 0 & \square & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & \square & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \square & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Im entsprechenden LGS kommt jede abhängige Variable nur einmal vor, und zwar an der \square -Stelle. Um die allgemeine Lösung zu erhalten, reicht es jetzt die unabhängigen Variablen in $\lambda_1, \lambda_2, \dots$ umzubenennen und auf die rechte Seite zu bringen.

Das Lösen der LGS mit Hilfe der inversen Matrix

Besonders einfach sieht das Gauß-Jordan-Verfahren bei einem LGS aus n Gleichungen für n Unbekannten, dessen Matrix invertierbar ist. In diesem Fall entsteht am Ende die Matrix

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & c_1 \\ 0 & 1 & \ddots & \vdots & c_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & c_n \end{array} \right)$$

Das heißt, das System hat eine eindeutige Lösung $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$.

Diese Lösung kann wie folgt beschrieben werden.

Satz 6. Sei $Ax = b$ ein lineares Gleichungssystem aus n Gleichungen für n Unbekannten und einer invertierbaren Matrix A . Dann hat es eine eindeutige Lösung

$$x = A^{-1}b$$

Beweis. Multipliziere beide Seiten der Gleichung $Ax = b$ mit der Matrix A^{-1} von links. \square

3 Vektorräume

3.1 Linearkombinationen

Definition 3.1. Seien $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ beliebige Vektoren, und seien $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ beliebige reelle Zahlen. Dann heißt

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = \sum_{i=1}^k \alpha_i v_i$$

Linearkombination der v_i (mit Koeffizienten α_i).

Beispiel 3.2. $2v_1 + v_2 - 3v_3$ ist eine Linearkombination von Vektoren v_1, v_2, v_3 . Es seien $v_1 = (1, 1, -1)^\top$, $v_2 = (2, 1, 2)^\top$, $v_3 = (1, 1, 0)^\top$. Dann gilt

$$2v_1 + v_2 - 3v_3 = 2 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Man sagt, dass der Vektor $(1, 0, 0)^\top$ als *Linearkombination* von v_1, v_2, v_3 dargestellt werden kann.

Beispiel 3.3. Jeder Vektor in \mathbb{R}^n kann als Linearkombination von Basisvektoren dargestellt werden: bezeichne

$$\begin{aligned} e_1 &= (1, 0, \dots, 0) \\ e_2 &= (0, 1, \dots, 0) \\ &\dots \\ e_n &= (0, 0, \dots, 1) \end{aligned}$$

Sei $v = (x_1, x_2, \dots, x_n)$ beliebig. Dann gilt

$$v = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

Koeffizienten dieser Linearkombination sind also nichts anderes als die Komponenten des Vektors.

Bei den elementaren Zeilenumformungen bilden man Linearkombinationen von Zeilen, wie z. B. $Z_i + \alpha Z_j$. Wir haben gesehen (Lemma 4), dass dies der Multiplikation von links mit einer Elementarmatrix entspricht. Allgemeiner gilt

Lemma 9. Sei A eine $m \times n$ -Matrix, und sei P eine $m \times m$ -Matrix. Betrachten wir das Produkt

$$A' := PA$$

Dann ist jede Zeile von A' eine Linearkombination der Zeilen von A .

Wenn die Matrix P invertierbar ist, dann ist auch jede Zeile von A eine Linearkombination der Zeilen von A' .

Beweis. Bezeichne mit Z_i die i -te Zeile der Matrix A , mit Z'_i die i -te Zeile der Matrix A' :

$$A = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \end{pmatrix} \quad A' = \begin{pmatrix} Z'_1 \\ Z'_2 \\ \vdots \\ Z'_m \end{pmatrix}$$

Dann gilt

$$PA = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \dots & \dots & \dots & \dots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \end{pmatrix} = \begin{pmatrix} p_{11}Z_1 + p_{12}Z_2 + \dots + p_{1m}Z_m \\ p_{21}Z_1 + p_{22}Z_2 + \dots + p_{2m}Z_m \\ \dots \\ p_{m1}Z_1 + p_{m2}Z_2 + \dots + p_{mm}Z_m \end{pmatrix}$$

Also gilt

$$Z'_i = p_{i1}Z_1 + p_{i2}Z_2 + \dots + p_{im}Z_m$$

Wenn P invertierbar ist, dann

$$A' = PA \Rightarrow A = P^{-1}A',$$

woraus folgt, dass auch die Zeilen von A als Linearkombinationen von Zeilen von A' dargestellt werden können. \square

3.2 Der Rang einer Matrix

Definition 3.4. Sei A eine $m \times n$ -Matrix, und sei S eine Zeilenstufenmatrix, die aus A mittels Vorwärtselimination entsteht. Dann heißt die Anzahl der Nichtnullzeilen in S der Rang von A . Bezeichnung: $\text{Rang } A$.

Die Definition gibt auch eine Anleitung zur Berechnung des Rangs von A : die Matrix soll in die Zeilenstufenform gebracht werden, die Anzahl der Nichtnullzeilen ist gleich dem Rang.

Beispiel 3.5.

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

(Zeilenumformung $Z_3 \rightarrow Z_3 - Z_1$ bringt die Matrix in Zeilenstufenform)

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

- Eine Matrix hat Rang 0 genau dann, wenn alle ihre Einträge gleich Null sind. (Die Zeilenstufenmatrix S darf nur aus Nullzeilen bestehen, daraus folgt, dass in A alle Zeilen auch Nullzeilen sind.)

Es gibt ein Problem mit der Definition 3.4. Bei der Vorwärtselimination muss man manchmal eine Wahl treffen, und zwar bei der Zeilenvertauschung, wo eine Zahl ungleich Null in einer Spalte gesucht wird. Und bei unterschiedlichen Wahlen kann man unterschiedliche Zeilenstufenmatrizen am Ende erhalten. Wir müssen zeigen, dass die Anzahl der Nichtnullzeilen in der Zeilenstufenmatrix S unabhängig davon ist, durch welche elementare Zeilenumformungen S aus A entstanden ist.

Lemma 10. Der Rang ist wohldefiniert. Das heißt, wenn S und S' zwei unterschiedliche Zeilenstufenmatrizen sind, die aus A durch elementare Zeilenumformungen entstehen, dann ist die Anzahl der Nichtnullzeilen in S gleich der Anzahl der Nichtnullzeilen in S' .

Beweis. Da S und S' aus A durch elementare Zeilenumformungen entstehen, gilt

$$S = PA, \quad S' = P'A,$$

wobei P und P' Produkte von Elementarmatrizen sind. Das bedeutet insbesondere, dass P und P' invertierbar sind (Lemma 6), sodass

$$S = PA \Rightarrow A = P^{-1}S \Rightarrow S' = P'P^{-1}S$$

Nach dem Lemma 9, die Zeilen von S' sind Linearkombinationen der Zeilen von S . Da $P'P^{-1}$ invertierbar ist: $(P'P^{-1})^{-1} = P(P')^{-1}$, sind auch die Zeilen von S Linearkombinationen der Zeilen von S' .

Jetzt wollen wir die Tatsache ausnutzen, dass S und S' Zeilenstufenform haben. Sagen wir, dass eine Zeile *Leitindex* j hat, wenn ihr erstes Nichtnullelement an der j -ten Stelle vorkommt (für Nullzeilen wird der Leitindex nicht definiert). In Zeilenstufenmatrizen haben wir diese Elemente mit \square -bezeichnet. Es seien

$$j_1 < j_2 < \dots < j_r \quad (1.12)$$

die Leitindizes der Nichtnullzeilen von S .

Behauptung Eine Linearkombination von Zeilen mit Leitindizes (1.12) ist entweder eine Nullzeile oder hat den Leitindex aus derselben Menge (1.12).

Sei $Z = \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_r Z_r$ eine beliebige Linearkombination. Wenn $\alpha_1 \neq 0$ ist, dann hat Z den Leitindex j_1 , denn die Summanden ab $\alpha_2 Z_2$ sind Zeilen, die auf den ersten j_1 Stellen (und eventuell noch weiter) nur Nullen haben. Wenn $\alpha_1 = 0$, aber $\alpha_2 \neq 0$, dann hat Z den Leitindex j_2 , usw. Eine Nullzeile kann nur dann entstehen, wenn $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$ gilt. Hiermit ist die Behauptung bewiesen.

Aus der bewiesenen Behauptung folgt, dass die Leitindizes der Zeilen von S' sind unter den Leitindizes der Zeilen von S enthalten, und umgekehrt. Daraus folgt, dass die "Treppe" von S' dieselben "Stufentiefen" hat, wie die Treppe von S . Insbesondere, sind die Anzahlen der Stufen in beiden Treppen gleich, d. h. die Anzahlen der Nichtnullzeilen. \square

Mit Hilfe des Rangs können wir einige Punkte im Gauß-Verfahren zum Lösen eines LGS neu formulieren.

Lemma 11. Sei $Ax = b$ ein lineares Gleichungssystem aus m Gleichungen für n Unbekannte.

a) Das System ist genau dann lösbar, wenn gilt

$$\text{Rang } A = \text{Rang}(A \mid b) \quad (1.13)$$

b) Wird die Bedingung (1.13) erfüllt, und gilt $\text{Rang } A = r$, so hat die allgemeine Lösung des Systems $n - r$ freie und r abhängige Variablen.

Beweis. Zum a): Bei der Lösbarkeitsentscheidung wird die Matrix $(A \mid b)$ in Zeilenstufenform $(S \mid c)$ gebracht. Hat die Matrix $(S \mid c)$ eine Stufe mehr als die Matrix S , so enthält das umgeformte LGS eine Gleichung $0 = c_i$ mit $c_i \neq 0$ und ist deswegen nicht lösbar. Hat $(S \mid c)$ genauso viele Stufen wie S , dann gilt $\text{Rang } A = \text{Rang}(A \mid b)$, und auch das LGS ist lösbar.

Zum b): $\text{Rang } A$ ist die Anzahl der Nichtnullzeilen in S . Jede Nichtnullzeile entspricht einer abhängigen Variable. Daher ist die Anzahl der abhängigen Variablen gleich r . Die übrigen $n - r$ Variablen sind frei. \square

3.3 Lineare Teilräume von \mathbb{R}^n

Definition 3.6. Eine nichtleere Teilmenge $U \subset \mathbb{R}^n$ heißt linearer Teilraum (oder linearer Unterraum) von \mathbb{R}^n , wenn gilt:

- a) $u, v \in U \Rightarrow u + v \in U$
- b) $u \in U, \lambda \in \mathbb{R} \Rightarrow \lambda u \in U$

Das heißt, mit je zwei Vektoren $u, v \in U$ soll U auch ihre Summe enthalten, und mit jedem Vektor u alle seine Vielfachen.

Beispiel 3.7. Wähle ein $v \in \mathbb{R}^n$ und betrachte die Menge

$$\mathbb{R}v := \{\alpha v \mid \alpha \in \mathbb{R}\}$$

Das ist ein linearer Teilraum, denn

- für alle $\alpha, \beta \in \mathbb{R}$ gilt $\alpha v + \beta v = (\alpha + \beta)v \in \mathbb{R}v$ (Summe zweier Vielfachen von v ist wieder ein Vielfaches von v ;
- für alle $\alpha \in \mathbb{R}, \lambda \in \mathbb{R}$ gilt $\lambda(\alpha v) = (\lambda\alpha)v$ (Vielfaches eines Vielfachen ist wieder ein Vielfaches)

Den Teilraum $\mathbb{R}v$ kann man sich als eine Gerade durch den Koordinatenursprung mit dem Richtungsvektor v vorstellen.

Beispiel 3.8. Insbesondere ist

$$\mathbb{R} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \left\{ \begin{pmatrix} \lambda \\ \lambda \\ \lambda \end{pmatrix}, \lambda \in \mathbb{R} \right\}$$

ein linearer Teilraum von \mathbb{R}^3 .

Beispiel 3.9. Hingegen ist eine Gerade, die nicht durch den Koordinatenursprung geht, kein linearer Teilraum von \mathbb{R}^n . Betrachte z. B. die Gerade $L = \{x + y = 1\}$ in \mathbb{R}^2 . Die Punkte $v = (1, 0)$ und $w = (0, 1)$ liegen auf L , ihre Summe $v + w = (1, 1)$ allerdings nicht. Die Bedingung a) aus der Definition 3.6 wird also nicht erfüllt.

Beachte, dass wir Punkte mit ihren Ortsvektoren identifizieren. Es ist in diesem Sinne, dass wir über die Summe zweier Punkte reden.

Beispiel 3.10. Eine Ebene in \mathbb{R}^3 , die durch den Koordinatenursprung geht, ist ein linearer Teilraum von \mathbb{R}^3 , denn sowohl die Summe zweier Ortsvektoren, als auch das Vielfache eines Ortsvektors bleiben in derselben Ebene liegen.

Speziell ist die xy -Koordinatenebene $\{(\alpha, \beta, 0) \mid \alpha, \beta \in \mathbb{R}\}$ ein linearer Teilraum von \mathbb{R}^3 .

Eine Ebene, die nicht durch den Koordinatenursprung geht, ist kein linearer Teilraum.

Allgemein kann man sich einen linearen Teilraum von \mathbb{R}^n als eine "mehrdimensionale Ebene" durch den Koordinatenursprung im "noch mehrdimensionalen" Raum denken. Während es bei der Dimension des Raums \mathbb{R}^n um die Zahl n geht, soll der Begriff der Dimension eines Teilraums noch definiert werden.

Zuerst aber betrachten wir zwei Möglichkeiten, einen linearen Teilraum von \mathbb{R}^n zu beschreiben. Sie sind analog zu der Parameterdarstellung, bzw. impliziter Darstellung von Geraden und Ebenen.

Lineare Hülle oder Parameterdarstellung eines linearen Teilraums

Definition 3.11. Seien v_1, v_2, \dots, v_k Vektoren aus \mathbb{R}^n . Die Menge aller ihrer Linearkombinationen heißt lineare Hülle der v_i und wird mit $\text{Lin}(v_1, v_2, \dots, v_k)$ bezeichnet:

$$\text{Lin}(v_1, v_2, \dots, v_k) := \{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k \mid \alpha_i \in \mathbb{R}\}$$

Beispiel 3.12. Die lineare Hülle der Vektoren $(1, 1, 1)^\top$ und $(1, 2, 3)^\top$ ist

$$\left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\}$$

d. h. die Ebene durch den Koordinatenursprung und mit Richtungsvektoren $(1, 1, 1)^\top$ und $(1, 2, 3)^\top$.

Beispiel 3.13. Die lineare Hülle eines Vektors v ist die Menge $\mathbb{R}v$, siehe Beispiel 3.7.

Lemma 12. Die lineare Hülle $\text{Lin}(v_1, v_2, \dots, v_k)$ ist ein linearer Teilraum von \mathbb{R}^n .

Beweis. Wir sollen zeigen, dass für die Menge $\text{Lin}(v_1, v_2, \dots, v_k) \subset \mathbb{R}^n$ die beiden Bedingungen aus der Definition 3.6 erfüllt sind.

a): $u, v \in \text{Lin}(v_1, v_2, \dots, v_k)$ bedeutet, dass u und v Linearkombinationen der Vektoren v_1, v_2, \dots, v_k sind:

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k, \quad v = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_k v_k$$

Die Summe zweier Linearkombinationen ist wieder eine Linearkombination:

$$u + v = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \dots + (\alpha_k + \beta_k)v_k$$

Also für alle $u, v \in \text{Lin}(v_1, v_2, \dots, v_k)$ gilt auch $u + v \in \text{Lin}(v_1, v_2, \dots, v_k)$.

b): Wenn $u \in \text{Lin}(v_1, v_2, \dots, v_k)$ und $\lambda \in \mathbb{R}$, dann gilt

$$\lambda u = \lambda(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k) = \lambda \alpha_1 v_1 + \lambda \alpha_2 v_2 + \dots + \lambda \alpha_k v_k$$

(ein Vielfaches einer Linearkombination ist wieder eine Linearkombination). Also aus $u \in \text{Lin}(v_1, v_2, \dots, v_k)$ und $\lambda \in \mathbb{R}$ folgt $\lambda u \in \text{Lin}(v_1, v_2, \dots, v_k)$. \square

Definition 3.14. Der Zeilenraum einer Matrix wird definiert als die lineare Hülle ihrer Zeilen. Der Zeilenraum einer $m \times n$ -Matrix ist hiermit ein linearer Teilraum von \mathbb{R}^n .

Bemerkung 3.15. Die lineare Hülle der Vektoren v_1, v_2, \dots, v_k ist der "kleinste" lineare Teilraum von \mathbb{R}^n , der v_1, v_2, \dots, v_k enthält. In der Tat, wiederholte Anwendung der Eigenschaften a) und b) aus der Definition 3.6 zeigt, dass jeder linearer Unterraum, der v_1, v_2, \dots, v_k enthält, auch jede ihre Linearkombination enthalten soll.

Nullraum oder implizite Darstellung eines linearen Teilraums

Definition 3.16. Sei A eine $m \times n$ -Matrix. Die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ wird der Kern von A oder der rechte Nullraum von A genannt. Bezeichnung:

$$\text{Kern } A := \{x \in \mathbb{R}^n \mid Ax = 0\}$$

Lemma 13. Der Kern einer $m \times n$ -Matrix A ist ein linearer Teilraum von \mathbb{R}^n .

Beweis. Sind $u, v \in \text{Kern } A$, so bedeutet das $Au = Av = 0$. Dann gilt auch

$$A(u + v) = Au + Av = 0 + 0 = 0,$$

und das bedeutet $u + v \in \text{Kern } A$.

Analog, wenn $u \in \text{Kern } A$ und $\lambda \in \mathbb{R}$, dann folgt aus $Au = 0$

$$A(\lambda u) = \lambda(Au) = \lambda \cdot 0 = 0,$$

also $\lambda u \in \text{Kern } A$.

Hiermit sind beide Bedingungen aus Definition 3.6 erfüllt, und $\text{Kern } A \subset \mathbb{R}^n$ ist ein linearer Teilraum. \square

Beispiel 3.17. • Die 1×3 -Matrix $A = (1, 2, 3)$ entspricht der linearen Gleichung $x + 2y + 3z = 0$, deren Lösungsmenge eine Ebene durch den Nullpunkt in \mathbb{R}^3 ist.

- Die 2×3 -Matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ entspricht dem LGS

$$\begin{aligned}x + y + z &= 0 \\x + 2y + 3z &= 0\end{aligned}$$

Das Lösen mit dem Gauß-Verfahren ergibt $z = \lambda$, $y = -2\lambda$, $x = \lambda$, d. h.

$$\text{Kern} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \left\{ \begin{pmatrix} \lambda \\ -2\lambda \\ \lambda \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$$

Dieser lineare Teilraum ist eine Gerade durch den Nullpunkt mit dem Richtungsvektor $(1, -2, 1)$.

Fundamentallösungen eines homogenen LGS

Das Lösen eines homogenen LGS $Ax = 0$ ist nichts anderes als Übergang von einer impliziten Darstellung eines linearen Teilraums zu einer Parameterdarstellung.

In der Tat, wenn die freien Variablen gleich $\lambda_1, \lambda_2, \dots, \lambda_s$ gesetzt werden, dann wird die allgemeine Lösung als eine Linearkombination mit Koeffizienten λ_i dargestellt. Die dabei kombinierte Vektoren heißen *Fundamentallösungen*.

Beispiel 3.18. Die allgemeine Lösung des LGS

$$\begin{aligned}x_1 - 2x_2 + 3x_3 + 4x_4 + 2x_5 &= 0 \\2x_3 + x_4 - 4x_5 &= 0 \\-x_4 + 3x_5 &= 0\end{aligned}$$

haben wir in der ersten Vorlesung wie folgt aufgeschrieben:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2\lambda_1 - 15.5\lambda_2 \\ \lambda_1 \\ 0.5\lambda_2 \\ 3\lambda_2 \\ \lambda_2 \end{pmatrix}$$

Dasgleiche kann auch anders geschrieben werden:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \in \left\{ \lambda_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -15.5 \\ 0 \\ 0.5 \\ 3 \\ 1 \end{pmatrix} \mid \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

Die zwei Lösungen $(2, 1, 0, 0, 0)$ und $(-15.5, 0, 0.5, 3, 1)$ heißen Fundamentallösungen des LGS.

Angenommen, beim Lösen eines homogenen LGS mit dem Gauß-Verfahren erhalten wir s freie Variablen, die gleich $\lambda_1, \lambda_2, \dots, \lambda_s$ gesetzt werden. Die Fundamentallösungen des LGS werden wie folgt ermittelt: für die erste Fundamentallösung setzt man

$$\lambda_1 = 1, \lambda_2 = 0, \dots, \lambda_s = 0$$

ein; für die zweite

$$\lambda_1 = 0, \lambda_2 = 1, \dots, \lambda_s = 0$$

usw. Insgesamt erhält man s Fundamentallösungen, und die Lösungsmenge des Systems kann als ihre lineare Hülle beschrieben werden.

3.4 Vektorraum: Definition

Definition 3.19. Ein Vektorraum ist eine Menge V , auf welcher zwei Operationen definiert sind:

- *Addition:* je zwei Elementen $u, v \in V$ wird ihre Summe $u + v \in V$ zugeordnet;
- *Multiplikation mit einer Zahl:* jedem $u \in V$ und jeder Zahl $\lambda \in \mathbb{R}$ wird ein Element $\lambda v \in V$ zugeordnet.

Für diese Operationen sollen die folgenden Regeln (Vektorraumaxiome) gelten:

$$(A1) \quad \forall u, v \in V \text{ gilt } u + v = v + u$$

$$(A2) \quad \forall u, v, w \in V \text{ gilt } (u + v) + w = u + (v + w)$$

(A3) Es gibt ein Element $\mathbf{0} \in U$, Nullvektor genannt, sodass $u + \mathbf{0} = u$ für alle $u \in U$ gilt.³

(A4) Zu jedem $u \in V$ gibt es genau ein mit $-u$ bezeichnetes Element in V , sodass $u + (-u) = \mathbf{0}$ gilt.

(M1) $\forall u \in V$ gilt $1u = u$

(M2) $\forall \lambda, \mu \in \mathbb{R}$ und $u \in V$ gilt $\lambda(\mu u) = (\lambda\mu)u$

(D1) $\forall \lambda \in \mathbb{R}$ und $u, v \in V$ gilt $\lambda(u + v) = \lambda u + \lambda v$

(D2) $\forall \lambda, \mu \in \mathbb{R}$ und $u \in V$ gilt $(\lambda + \mu)u = \lambda u + \mu u$

Beispiel 3.20. • Das Urbeispiel eines Vektorraums ist natürlich \mathbb{R}^n : die Menge aller n -Tupel reeller Zahlen.

- Die Menge aller $m \times n$ -Matrizen bildet auch einen Vektorraum.
- Ein linearer Teilraum von \mathbb{R}^n ist ein Vektorraum: die Elemente können addiert und mit Zahlen multipliziert werden, die Vektorraumaxiome gelten offensichtlich.

Beispiel 3.21. Die Menge aller Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ bildet einen Vektorraum. Für zwei Funktionen f, g und eine reelle Zahl λ definieren wir neue Funktionen $f + g$ und λf als

$$(f + g)(x) := f(x) + g(x)$$
$$(\lambda f)(x) := \lambda(f(x))$$

(Die Summe $f + g$ kann auch als die "Summe der Formeln" oder die "Summe der Graphen" beschreiben.)

Analog zu Definition 3.6 kann man einen linearen Teilraum eines Vektorraums definieren.

Definition 3.22. Sei V ein Vektorraum. Eine nichtleere Teilmenge $U \subset V$ heißt linearer Teilraum (oder Untervektorraum) von V , wenn gilt:

a) $u, v \in U \Rightarrow u + v \in U$

b) $u \in U, \lambda \in \mathbb{R} \Rightarrow \lambda u \in U$

Das heißt, mit je zwei Elementen $u, v \in U$ soll U auch ihre Summe enthalten, und mit jedem Element u alle seine Vielfachen.

Ein linearer Teilraum ist damit selbst ein Vektorraum (man "vergisst" den umgebenden Raum V und betrachtet U allein).

Beispiel 3.23. • Die Menge aller oberen $n \times n$ -Dreiecksmatrizen ist ein linearer Teilraum des Raums aller $n \times n$ -Matrizen.

- Die Menge aller Polynome ist ein linearer Teilraum des Raums aller Funktionen. In der Tat, die Summe zweier Polynome ist wieder ein Polynom; ein Vielfaches eines Polynoms ist ein Polynom.

³Im Folgenden werde wir den Nullvektor in jedem Vektorraum auch einfach mit 0 bezeichnen.

3.5 Basis eines Vektorraums

Sei V ein Vektorraum. Genau wie für Vektoren in \mathbb{R}^n , kann man *Linearkombinationen* von Elementen aus V definieren: das sind Ausdrücke der Form

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k$$

Mit Linearkombinationen kann man wie gewohnt operieren, dank der Vektorraumaxiomen.

Die Menge aller Linearkombinationen von v_1, v_2, \dots, v_k wird die *lineare Hülle* von v_1, v_2, \dots, v_k genannt.

Beispiel 3.24. • Betrachten wir im Vektorraum aller Funktionen die Funktionen $\sin x$ und $\cos x$. Dann gilt

$$\text{Lin}(\sin x, \cos x) = \{a \sin x + b \cos x \mid a, b \in \mathbb{R}\}$$

(Wir wissen, dass jede solche Funktion auch als $A \sin(x + \omega_0)$ geschrieben werden kann.)

- Die lineare Hülle der Funktionen $1, x, x^2, \dots, x^n$ ist die Menge aller Polynome vom Grad kleiner oder gleich n .

Erzeugendensysteme und lineare Abhängigkeit

Definition 3.25. Man sagt, dass (v_1, v_2, \dots, v_k) ein Erzeugendensystem von V ist, wenn

$$V = \text{Lin}(v_1, v_2, \dots, v_k)$$

Mit anderen Worten, jedes Element von V soll als eine Linearkombination von Elementen v_1, v_2, \dots, v_k darstellbar sein.

Beispiel 3.26. • \mathbb{R}^2 wird erzeugt von $(1, 0)$ und $(0, 1)$, aber auch z. B. von $(1, 1)$ und $(1, -1)$.

- Der Zeilenraum einer Matrix wird erzeugt von ihren Zeilen.
- Die Lösungsmenge eines homogenen LGS wird erzeugt durch seine Fundamentallösungen.
- Die Menge aller Funktionen hat kein endliches Erzeugendensystem.

Definition 3.27. Die Vektoren $v_1, v_2, \dots, v_k \in V$ heißen linear abhängig, wenn es Zahlen $\alpha_1, \alpha_2, \dots, \alpha_k$ gibt, die nicht alle gleich Null sind, sodass gilt

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0$$

(0 auf der rechten Seite bedeutet den Nullvektor)

Die Vektoren v_1, v_2, \dots, v_k heißen linear unabhängig, wenn sie nicht linear abhängig sind, d. h. wenn

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0 \Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$$

Beispiel 3.28. • Die Vektoren in \mathbb{R}^n

$$\begin{aligned}e_1 &= (1, 0, \dots, 0) \\e_2 &= (0, 1, \dots, 0) \\&\dots \\e_n &= (0, 0, \dots, 1)\end{aligned}$$

sind linear unabhängig, denn

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = (\alpha_1, \alpha_2, \dots, \alpha_n),$$

und dies ist ein Nullvektor genau dann wenn alle Koeffizienten α_i gleich Null sind.

- Die Vektoren $v_1 = (1, 1, 0)$, $v_2 = (1, -1, 0)$, $v_3 = (0, 1, 0)$ sind linear abhängig, denn es gilt

$$v_1 - v_2 - 2v_3 = 0$$

- Auch die Vektoren $v_1 = (1, 1, 1)$, $v_2 = (2, 2, 2)$, $v_3 = (3, 3, 3)$ sind linear abhängig. Hier kann man mehrere verschwindende Linearkombinationen angeben, sowie $2v_1 - v_2 = 0$ oder $v_1 + v_2 - v_3 = 0$.
- Die Funktionen $\sin x$, $\cos x$, $\sin(x + \frac{\pi}{3})$ sind linear abhängig.

Definition einer Basis

Definition 3.29. Ein System (v_1, v_2, \dots, v_k) von Vektoren aus V heißt eine Basis von V , wenn gilt:

- a) Die Vektoren v_1, v_2, \dots, v_k erzeugen V , d. h. $V = \text{Lin}(v_1, v_2, \dots, v_k)$.
- b) Die Vektoren v_1, v_2, \dots, v_k sind linear unabhängig.

Beispiel 3.30. • Die Vektoren in \mathbb{R}^n

$$\begin{aligned}e_1 &= (1, 0, \dots, 0) \\e_2 &= (0, 1, \dots, 0) \\&\dots \\e_n &= (0, 0, \dots, 1)\end{aligned}$$

bilden eine Basis von \mathbb{R}^n . Diese Basis heißt die *natürliche Basis* oder die *Standardbasis* von \mathbb{R}^n .

- Der Vektorraum \mathbb{R}^n hat viele andere Basen. Zum Beispiel,

$$v_1 = (1, -1), \quad v_2 = (1, 1)$$

ist eine Basis von \mathbb{R}^2 .

- Die Funktionen $\sin x$ und $\cos x$ sind linear unabhängig (die Funktion $a \sin x + b \cos x$ kann nur bei $a = b = 0$ identisch Null sein). Deswegen bilden sie eine Basis im Vektorraum aller Funktionen der Form $A \sin(x + \omega_0)$.

- Bei manchen Vektorräumen gibt es keine ausgezeichnete Basis: keine Basis ist "besser" als die anderen. Z. B. die wohl "einfachsten" Vektoren in der Ebene $x + y + z = 0$ sind

$$v_1 = (1, -1, 0)$$

$$v_2 = (0, 1, -1)$$

$$v_3 = (1, 0, -1)$$

Jedes Paar (v_1, v_2) , (v_1, v_3) , (v_2, v_3) ist eine Basis von $x + y + z = 0$.

3.6 Dimension eines Vektorraums

Der Vektorraum aller Funktionen hat keine Basis, weil er nicht von einem endlichen System von Funktionen erzeugt werden kann. Man sagt, dieser Raum ist *unendlichdimensional*.

Wenn ein Vektorraum eine Basis hat, dann hat er auch unendlich viele andere Basen. Oft kann keine Basis allen anderen bevorzugt werden. Was aber unabhängig von der Wahl der Basis ist, ist die Anzahl der Basisvektoren.

Definition 3.31. Sei (v_1, v_2, \dots, v_k) eine Basis des Vektorraums V . Dann heißt die Zahl k die Dimension von V , und wird mit $\dim V$ bezeichnet.

Ähnlich wie bei der Definition des Rangs einer Matrix, sollen wir die *Wohldefiniertheit* der Dimension beweisen: die Anzahl der Vektoren ist in jeder Basis dieselbe. Hierfür brauchen wir die folgenden zwei Lemmas.

Lemma 14. Sei (v_1, v_2, \dots, v_k) eine Basis des Vektorraums V . Dann gibt es zu jedem Vektor $u \in V$ genau ein n -Tupel reeller Zahlen $(\alpha_1, \alpha_2, \dots, \alpha_k)$, sodass

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k$$

Die Zahlen $(\alpha_1, \alpha_2, \dots, \alpha_k)$ heißen Koordinaten von u bezüglich der Basis (u_1, u_2, \dots, u_k) .

Beweis. Die Eigenschaft a) aus der Definition einer Basis besagt, dass jeder Vektor $u \in V$ als Linearkombination der Vektoren (v_1, v_2, \dots, v_k) dargestellt werden kann.

Nehmen wir an, es gibt zwei unterschiedliche Wege, den Vektor u als Linearkombination von (v_1, v_2, \dots, v_k) darzustellen:

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k$$

$$u = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_k v_k$$

Die Subtraktion der zweiten Gleichung aus der ersten ergibt

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \dots + (\alpha_k - \beta_k)v_k = 0$$

Das widerspricht aber der linearen Unabhängigkeit der Vektoren v_1, v_2, \dots, v_k (Eigenschaft b) aus der Definition einer Basis). \square

Lemma 15. Sei A eine $m \times n$ -Matrix mit $m > n$. Dann sind ihre Zeilen linear abhängig.

Beweis. Bringen wir die Matrix A in die Zeilenstufenform mit Hilfe der elementaren Zeilenumformungen: $S = PA$. Wegen $m > n$ enthält die Matrix S mindestens eine Nullzeile. Nach dem Lemma 9 ist diese Zeile eine Linearkombination der Zeilen der Matrix A :

$$0 = \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_m Z_m \quad (1.14)$$

Es muss nur gezeigt werden, dass nicht alle α_i gleich Null sind. Diese Zahlen bilden aber eine Zeile der Matrix P (siehe den Beweis des Lemmas 9), und da P invertierbar ist, kann sie keine Zeile aus Nullen enthalten.

Deswegen ist (1.14) eine lineare Abhängigkeit zwischen den Zeilen von A . \square

Satz 7. Sind (v_1, v_2, \dots, v_k) und (w_1, w_2, \dots, w_l) zwei Basen eines Vektorraums V , dann gilt $k = l$.

Beweis. Nehmen wir an, dass $l > k$ gilt. Schreiben wir jeden der Vektoren w_i als Linearkombination der Vektoren (v_1, v_2, \dots, v_k) :

$$\begin{aligned} w_1 &= \alpha_{11}v_1 + \alpha_{12}v_2 + \dots + \alpha_{1k}v_k \\ w_2 &= \alpha_{21}v_1 + \alpha_{22}v_2 + \dots + \alpha_{2k}v_k \\ &\dots \\ w_l &= \alpha_{l1}v_1 + \alpha_{l2}v_2 + \dots + \alpha_{lk}v_k \end{aligned}$$

In der entstehenden $l \times k$ -Matrix $(\alpha_{ij})_{l \times k}$ ist $l > k$. Aus dem Lemma 15 folgt, dass die Zeilen (Z_i) dieser Matrix linear abhängig sind: es gibt $(\lambda_1, \dots, \lambda_l)$ nicht alle gleich Null, so dass $\lambda_1 Z_1 + \dots + \lambda_l Z_l = 0$. Dann gilt aber auch

$$\lambda_1 w_1 + \lambda_2 w_2 + \dots + \lambda_l w_l = 0$$

Das heißt, die Vektoren (w_1, w_2, \dots, w_l) sind linear abhängig. Das widerspricht der Annahmen, dass sie eine Basis bilden.

Im Fall $k > l$ kann man analog zeigen, dass die Vektoren (v_1, v_2, \dots, v_k) linear abhängig sein sollen. Das heißt, beide Annahmen sind falsch und es gilt $k = l$. \square

Bemerkung 3.32. Die Dimension eines Vektorraumes ist gleich Null genau dann, wenn der Vektorraum nur aus dem Nullvektor besteht. Ein nulldimensionaler Vektorraum entsteht z. B. wenn wir die Lösungsmenge eines homogenen LGS betrachten, das nur die Nulllösung besitzt.

3.7 Basis und Dimension des Zeilenraums und des Kerns einer Matrix

Im Folgenden sei A eine $m \times n$ -Matrix.

Lemma 16. Sei S eine Zeilenstufenmatrix, die aus A mittels Vorwärtselimination entsteht. Dann ist der Zeilenraum von S gleich dem Zeilenraum von A .

Beweis. Es gilt $S = PA$ mit einer invertierbaren Matrix P (ein Produkt von Elementarmatrizen). Nach dem Lemma 9 sind die Zeilen (Z'_i) von S Linearkombinationen von Zeilen (Z_i) von A und umgekehrt. Dann ist jede Linearkombination der (Z'_i) auch Linearkombination von (Z_i) und umgekehrt. Folglich

$$\text{Lin}(Z_1, Z_2, \dots, Z_m) = \text{Lin}(Z'_1, Z'_2, \dots, Z'_m)$$

und das Lemma ist bewiesen. \square

Lemma 17. *Sei S eine Zeilenstufenmatrix, die aus A mittels Vorwärtselimination entsteht. Dann bilden die Nichtnullzeilen von S eine Basis des Zeilenraums von A .*

Beweis. Nach dem obigen Lemma ist der Zeilenraum von A gleich dem Zeilenraum von S . Der Zeilenraum von S wird von den Nichtnullzeilen erzeugt, die Nullzeilen können ignoriert werden. Es bleibt zu zeigen, dass die Nichtnullzeilen einer Zeilenstufenmatrix linear unabhängig sind.

Das haben wir allerdings im Beweis des Lemmas 10 festgestellt: eine Linearkombination von Zeilen mit unterschiedlichen Leitindizes ist eine Nullzeile nur wenn alle Koeffizienten der Linearkombination Null sind. \square

Erinnerung: Fundamentallösungen eines homogenen LGS entstehen, wenn eine der freien Variablen gleich 1 gesetzt wird und alle anderen gleich 0.

Lemma 18. *Die Fundamentallösungen bilden eine Basis von Kern A (der Lösungsmenge von $Ax = 0$).*

Beweis. Seien X_1, X_2, \dots, X_s die Fundamentallösungen. Dann ist $\lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_s X_s$ auch eine Lösung, und in dieser Lösung nehmen die freien Variablen die Werte $\lambda_1, \lambda_2, \dots, \lambda_s$. Das zeigt einerseits, dass jede Lösung als Linearkombination fundamentaler Lösungen dargestellt werden kann (denn die Werte freier Variablen bestimmen die Lösung eindeutig), andererseits dass keine Linearkombination der Fundamentallösungen gleich Null ist. \square

Hiermit ist die Dimension des Kerns gleich der Anzahl der freien Variablen. Nach dem Punkt b) des Lemmas 11 ist diese Zahl gleich $n - \text{Rang } A$. Zusammengefasst:

Satz 8. *Sei $Ax = 0$ ein homogenes LGS aus m Gleichungen für n Unbekannte.*

a) *Es gilt*

$$\text{Rang } A + \text{Dim Kern } A = n \quad (1.15)$$

b) *Wird A durch elementare Zeilenumformungen in Zeilenstufenform gebracht, so bilden die Nichtnullzeilen der Zeilenstufenmatrix eine Basis des Zeilenraums von A .*

c) *Die Fundamentallösungen bilden eine Basis des Kerns von A .*

3.8 Der Basisauswahlsatz und der Basisergänzungssatz

Basisauswahl

Sei V ein Vektorraum.

Satz 9. *Aus jedem linear abhängigen System von Vektoren kann eine Basis seiner linearen Hülle gewählt werden.*

Mit anderen Worten, aus jedem linear abhängigen System von Vektoren können einige weggelassen werden, so dass die restlichen linear unabhängig sind und die gleiche lineare Hülle haben wie das ursprüngliche System.

Lemma 19. *Sind Vektoren v_1, v_2, \dots, v_k linear abhängig, so kann einer von ihnen als Linearkombination der anderen dargestellt werden.*

Beweis. Sind v_1, v_2, \dots, v_k linear abhängig, so gibt es Zahlen $\alpha_1, \alpha_2, \dots, \alpha_k$, die nicht alle gleich Null sind, sodass

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0$$

Angenommen, $\alpha_i \neq 0$. Dann folgt aus der obigen Gleichung

$$v_i = -\frac{\alpha_1}{\alpha_i} v_1 - \dots - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \dots - \frac{\alpha_k}{\alpha_i} v_k,$$

also v_i ist eine Linearkombination der anderen Vektoren. □

Bemerkung 3.33. Es gilt allerdings nicht, dass *jeder* der Vektoren eines linear abhängigen Systems als Linearkombination der anderen dargestellt werden kann. Ein Beispiel:

$$v_1 = (1, 0, 0), \quad v_2 = (1, 1, 1), \quad v_3 = (2, 2, 2)$$

Das System (v_1, v_2, v_3) ist linear abhängig: $2v_2 - v_3 = 0$. Allerdings kann v_1 nicht als Linearkombination von v_2 und v_3 dargestellt werden.

Beweis des Satzes 9. Sei (v_1, v_2, \dots, v_k) ein linear abhängiges System von Vektoren. Nach Lemma 19 kann einer von ihnen als Linearkombination der anderen dargestellt werden. Sei v_i so ein Vektor. Man kann leicht sehen, dass sich die lineare Hülle beim Weglassen von v_i nicht ändert:

$$\text{Lin}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k) = \text{Lin}(v_1, \dots, v_k)$$

Sind jetzt die Vektoren $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k$ linear unabhängig, dann haben wir eine Basis der linearen Hülle gewählt. Sind sie abhängig, dann können wir wieder einen Vektor finden, der durch die anderen ausdrückbar ist, ihn weglassen usw. □

Insbesondere kann man aus den Zeilen jeder Matrix eine Basis ihres Zeilenraums wählen.

Beispiel 3.34. Jedes Paar der Zeilen der Matrix

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

bildet eine Basis ihrer Zeilenraums.

Basisergänzung

Hier müssen wir zusätzlich annehmen, dass der Vektorraum V endlichdimensional ist, d. h. eine Basis besitzt.

Satz 10. *Jedes linear unabhängige System von Vektoren kann zu einer Basis ergänzt werden.*

Lemma 20. *Jedes System aus $m > \text{Dim } V$ Vektoren ist linear abhängig.*

Beweis. Das wird mit demselben Argument bewiesen, wie Satz 7. Man zerlegt die m Vektoren bezüglich einer Basis von V und erhält damit eine $m \times n$ -Matrix mit $m > n$. Da die Zeilen einer solchen Matrix immer linear abhängig sind, sind auch die gegebenen Vektoren linear abhängig. \square

Beweis des Satzes 10. Seien Vektoren v_1, v_2, \dots, v_k linear unabhängig. Wenn sie nicht den ganzen Raum V erzeugen, dann gibt es einen Vektor $v_{k+1} \in V$, der nicht in $\text{Lin}(v_1, v_2, \dots, v_k)$ liegt. Es ist leicht zu zeigen, dass dann auch das System $(v_1, v_2, \dots, v_{k+1})$ linear unabhängig ist.

Erzeugt das System $(v_1, v_2, \dots, v_{k+1})$ den Raum V immer noch nicht, so können wir einen weiteren Vektor v_{k+2} hinzunehmen, ohne die lineare Unabhängigkeit zu verlieren, usw. Irgendwann erhalten wir ein linear unabhängiges System, das den ganzen Raum erzeugt (also eine Basis), denn nach Lemma 20 kann es nicht mehr als $\text{Dim } V$ linear unabhängige Vektoren geben. \square

Korollar 2. *Sei U ein linearer Teilraum von V . Dann ist U endlichdimensional und es gilt $\text{Dim } U \leq \text{Dim } V$.*

Beweis. Wähle einen Vektor $u_1 \in U$. Wenn u_1 den U nicht erzeugt, wähle einen weiteren Vektor, usw. (siehe Beweis des Satzes 10) Es kann in U nicht mehr als $\text{Dim } V$ linear unabhängige Vektoren geben, deswegen erhalten wir irgendwann ein den Teilraum U erzeugendes linear unabhängiges System, also eine Basis von U , und diese Basis wird nicht mehr als $\text{Dim } V$ Vektoren enthalten. \square

Korollar 3. *Sei $\text{Dim } V = n$. Dann bildet jedes linear unabhängige System aus n Vektoren eine Basis von V .*

Beweis. Wenn ein linear unabhängiges System keine Basis bildet, dann kann es zu einer Basis ergänzt werden. So entsteht aber eine Basis aus mehr als $\text{Dim } V$ Vektoren, was unmöglich ist. \square

Korollar 4. *Für jede Matrix A ist $\text{Rang } A$ gleich der maximalen Anzahl linear unabhängiger Zeilen von A .*

Beweis. Sei $r = \text{Rang } A$. Nach Satz 9 kann aus den Zeilen eine Basis des Zeilenraums gewählt werden, und das heißt, es gibt r linear unabhängige Zeilen. Andererseits zeigt eine Anwendung des Lemmas 20 auf den Zeilenraum von A , dass es nicht mehr als r linear unabhängige Zeilen geben kann. \square

3.9 Zeilenrang = Spaltenrang

Die Zeilen einer $m \times n$ -Matrix A bilden ein System von Vektoren in \mathbb{R}^n ; die Spalten bilden ein System von Vektoren in \mathbb{R}^m . Wir haben bisher nur mit den Zeilen operiert, allerdings können dieselben Begriffe auch für Spalten betrachtet werden.

Insbesondere kann jede Matrix durch *elementare Spaltenumformungen* in *Spaltenstufenform* gebracht werden.

Definition 3.35. *Sei A' eine Spaltenstufenmatrix, die aus A durch elementare Spaltenumformungen entsteht. Dann heißt die Anzahl der Nichtnullspalten in A' der Spaltenrang von A .*

Mit anderen Worten, der Spaltenrang ist gleich der Dimension des Spaltenraums von A und auch gleich der maximalen Anzahl linear unabhängiger Spalten von A . Der *Spaltenraum* von A besteht aus allen Linearkombinationen der Spalten. Eine Linearkombination von Spalten mit Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_n$ kann als Produkt

$$(S_1 \quad S_2 \quad \dots \quad S_n) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

geschrieben werden (hier bezeichnet S_j die j -te Spalte von A). Deswegen gilt

$$\text{Der Spaltenraum von } A = \{Ax \mid x \in \mathbb{R}^n\}$$

Gerechtigkeithalber benutzen wir jetzt das Wort "Zeilenrang" anstelle des Wortes "Rang". Es wird allerdings nicht mehr nötig sein, nachdem wir den folgenden Satz beweisen.

Satz 11. *Der Zeilenrang jeder Matrix ist gleich ihrem Spaltenrang.*

Es ist klar, dass elementare Zeilenumformungen den Zeilenrang, und elementare Spaltenumformungen den Spaltenrang erhalten. Das folgende Lemma sagt mehr.

Lemma 21. *Elementare Zeilenumformungen ändern den Spaltenrang der Matrix nicht; elementare Spaltenumformungen ändern den Zeilenrang der Matrix nicht.*

Beweis. Eine elementare Zeilenumformung transformiert die Matrix A zu $\tilde{E}A$, wobei \tilde{E} eine Elementarmatrix ist. Für jeden Vektor $x \in \mathbb{R}^n$ gilt

$$Ax = 0 \Leftrightarrow \tilde{E}Ax = 0$$

(das kann als "Lösungsmenge von $Ax = 0$ gleich der Lösungsmenge von $\tilde{E}Ax = 0$ " interpretiert werden). Das heißt,

$$\text{Kern } A = \text{Kern}(\tilde{E}A)$$

Ein Element aus Kern A ist aber nichts anderes als eine lineare Relation zwischen den Spalten von A :

$$Ax = 0 \Leftrightarrow x_1 S_1 + x_2 S_2 + \dots + x_n S_n = 0$$

Die Tatsache, dass sich Kern A bei elementaren Zeilenumformungen nicht ändert, bedeutet also, dass es keine neue lineare Relationen zwischen den Spalten entstehen und keine alte verschwinden. Linear (un)abhängige Systeme von Spalten bleiben linear (un)abhängig. Da der Spaltenrang gleich der maximalen Anzahl linear unabhängiger Spalten ist, folgt daraus, dass der Spaltenrang erhalten bleibt.

Analog wird es gezeigt, dass bei einer elementaren Spaltenumformung der Zeilenrang erhalten bleibt. (Hier muss man die Menge aller $y \in \mathbb{R}^m$ mit $y^\top A = 0$ betrachten.) \square

Lemma 22. Jede $m \times n$ -Matrix A kann mittels elementarer Zeilen- und Spaltenumformungen zu der Form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \quad (1.16)$$

transformiert werden. Hier ist E_r die $r \times r$ -Einheitsmatrix, und 0 steht für Blöcke aus Nullen.

Beweis. Zuerst bringt man die Matrix mittels elementarer Zeilenumformungen in Zeilenstufenform; dann transformiert man die Zeilenstufenmatrix mittels elementarer Spaltenumformungen zum Gestalt (1.16). \square

Beweis des Satzes 11. Transformiert man die Matrix A mittels elementarer Zeilen- und Spaltenumformungen in die Form (1.16), so bleiben ihre Zeilen- und Spaltenrang laut Lemma 21 unverändert. Der Zeilenrang und der Spaltenrang der Matrix (1.16) sind aber beide gleich r . Daraus folgt, dass der Zeilenrang und der Spaltenrang von A beide gleich r sind, also gleich zueinander. \square

Korollar 5. Sei A eine $m \times n$ -Matrix. Dann gilt

$$\text{Rang}(BA) \leq \text{Rang } A \text{ für jede } l \times m\text{-Matrix } B \quad (1.17)$$

$$\text{Rang}(AC) \leq \text{Rang } A \text{ für jede } n \times p\text{-Matrix } C \quad (1.18)$$

Beweis. Jede Zeile der Matrix BA ist eine Linearkombination der Zeilen von A . Deswegen ist der Zeilenraum von BA ein linearer Teilraum des Zeilenraums von A und kann nach Korollar 2 nicht eine größere Dimension haben. \square

3.10 Kriterien der Invertierbarkeit einer Matrix

Satz 12. Sei A eine $n \times n$ -Matrix. Dann sind die folgenden Aussagen äquivalent.

- a) A ist invertierbar.
- b) Es gibt eine Matrix B , sodass $BA = E_n$ gilt.
- c) $\text{Rang } A = n$
- d) Das LGS $Ax = 0$ hat nur die Nulllösung.
- e) Die Zeilen (bzw. die Spalten) der Matrix A bilden eine Basis von \mathbb{R}^n .

Beweis. d) ist gleichbedeutend mit $\text{Dim Kern } A = 0$. Wegen der Formel $\text{Rang } A + \text{Dim Kern } A = n$ ist also c) \Leftrightarrow d).

Es ist auch c) \Leftrightarrow e). In der Tat, $\text{Rang } A = n$ bedeutet wegen Korollar 4 dass die Zeilen von A linear unabhängig sind. Und laut Korollar 3 bilden n linear unabhängige Vektoren in \mathbb{R}^n eine Basis von \mathbb{R}^n , also c) \Rightarrow e). In die andere Richtung: wenn die Zeilen eine Basis bilden, dann sind sie linear unabhängig, und das heißt $\text{Rang } A = n$.

a) \Leftrightarrow c) folgt aus der Definition des Rangs als die Anzahl der Nichtnullzeilen in der Zeilenstufenform: bei der Behandlung des Gauß-Jordan-Verfahrens haben wir festgestellt, dass A genau dann invertierbar ist, wenn sie in der Zeilenstufenform keine Nullzeilen enthält, d. h. genau dann, wenn $\text{Rang } A = n$ gilt.

Nun beweisen wir a) \Rightarrow b) \Rightarrow c). Die erste Implikation ist klar: man nehme einfach $B = A^{-1}$. Aus b) folgt wegen Korollar 5

$$\text{Rang } A \geq \text{Rang } BA = \text{Rang } E_n = n$$

Der Rang einer $n \times n$ -Matrix kann aber nicht größer als n sein, deswegen folgt $\text{Rang } A = n$. \square

Zusammenfassung

Rang A = Anzahl der Nichtnullzeilen in Zeilenstufenform =
 Anzahl der Nichtnullspalten in Spaltenstufenform =
 = Dimension des Zeilenraums = Dimension des Spaltenraums =
 maximale Anzahl linear unabhängiger Zeilen (Spalten)

Eine Basis ist

- ein linear unabhängiges Erzeugendensystem;
- ein System, bezüglich welchem jeder Vektor des Vektorraums eindeutig definierte Koordinaten hat;
- ein maximales linear unabhängiges System von Vektoren

4 Lineare Algebra und Statik der Fachwerke

Wir betrachten ein ebenes oder ein räumliches Fachwerk und studieren seine statischen Reaktionen auf die an die Knoten angreifenden Belastungen. Um nicht die Lagerreaktionen einbeziehen müssen, betrachten wir nur solche Belastungen, für welche die Gleichgewichtsbedingungen erfüllt sind.

Insbesondere, wenn jede Gleichgewichtsbelastung durch Stabkräfte aufgelöst werden kann, dann ist das Fachwerk nicht deformierbar (im Kleinen). Existiert die Auflösung, ist sie aber nicht eindeutig, dann ist das Fachwerk statisch unbestimmt.

4.1 Gleichgewichtsbelastungen

Betrachten wir ein ebenes Fachwerk mit Knoten p_1, p_2, \dots, p_k . Jeder p_k ist ein Punkt in \mathbb{R}^2 :

$$p_i = (x_i, y_i) \in \mathbb{R}^2$$

Sei F_i eine am Knoten p_i angreifende Kraft. Sie wird ebenfalls durch zwei Koordinaten beschrieben:

$$F_i = (f_i, g_i) \in \mathbb{R}^2$$

Definition 4.1. *Wir sagen, dass die Belastung (F_i) eines Fachwerks mit Knoten (p_i) die Gleichgewichtsbedingungen erfüllt, wenn*

$$\sum_{i=1}^k F_i = 0 \quad (\text{die resultierende Kraft gleich Null}) \text{ und} \quad (1.19)$$

$$\sum_{i=1}^k p_i \times F_i = 0 \quad (\text{das resultierende Moment gleich Null}) \text{ gilt.} \quad (1.20)$$

Für zwei Vektoren $v, w \in \mathbb{R}^2$ bezeichnet hier $v \times w := v_x w_y - v_y w_x \in \mathbb{R}$ ihr \mathbb{R}^2 -Kreuzprodukt.

Die Menge aller Belastungen wird mit \mathbb{R}^{2k} identifiziert, indem wir die Kraftkomponenten hintereinander aufschreiben:

$$(f_1, g_1, f_2, g_2, \dots, f_k, g_k) \in \mathbb{R}^{2k} \quad (1.21)$$

Lemma 23. *Die Menge aller Belastungen, die die Gleichgewichtsbedingungen erfüllen, ist ein $(2k - 3)$ -dimensionaler linearer Teilraum von \mathbb{R}^{2k} .*

Beweis. Die Bedingungen (1.19) und (1.20) stellen ein homogenes lineares Gleichungssystem für Unbekannte $f_1, g_1, f_2, g_2, \dots$ dar. Dieses System besteht aus drei Gleichungen und hat als Koeffizientenmatrix die $3 \times 2k$ -Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 1 \\ -y_1 & x_1 & -y_2 & x_2 & \cdots & -y_k & x_k \end{pmatrix}$$

Die Zeilen dieser Matrix sind linear unabhängig (wird das Koordinatensystem so gewählt, dass der Knoten p_1 im Koordinatenursprung liegt, dann hat die Matrix Zeilenstufenform), deswegen gilt $\text{Rang } A = 3$. Aus der Formel $\text{Rang } A + \text{Dim Kern } A = n$ folgt nun

$$\text{Dim Kern } A = n - \text{Rang } A = 2k - 3$$

□

Analog, für räumliche Fachwerke haben wir 6 Gleichgewichtsbedingungen (das resultierende Moment $\sum_{i=1}^k p_i \times F_i$ ist ein Vektor in \mathbb{R}^3), und man kann zeigen, dass die Belastungen unter den Gleichgewichtsbedingungen einen Vektorraum von Dimension $3k - 6$ bilden.

Bemerkung 4.2. Die Gleichgewichtsbedingungen können durch das Einführen von Lagern ersetzt werden: z. B. eines dreiwertigen Lagers in einem Knoten oder eines zweiwertigen und eines einwertigen in zwei unterschiedlichen Knoten. Dann können aus jeder Belastung die Lagerreaktionen eindeutig ermittelt werden. Die Summe Belastung + Lagerreaktionen kann als eine neue Belastung betrachtet werden, die dann die Gleichgewichtsbedingungen erfüllt.

4.2 Auflösen der Belastung durch Stabkräfte

Wir wollen nun die Frage studieren, wann jede Belastung eines Fachwerks in die Stabkräfte aufgelöst werden kann. Die Stabkräfte werden mittels Spannungen beschrieben. Eine Spannung im Stab $p_i p_j$ ist eine Zahl, die positiv ist, wenn auf den Stab eine Zugkraft auswirkt. Wir definieren dabei die Spannung als Kraft pro Längeneinheit.

Definition 4.3. *Wir sagen, dass im Stab $p_i p_j$ eine Spannung ω auftritt, wenn der Stab auf den Knoten p_i die Kraft $\omega(p_i - p_j)$ und auf den Knoten p_j die Kraft $\omega(p_j - p_i)$ ausübt, siehe Abb. 1.1.*

Sei s die Anzahl der Stäbe im Fachwerk. Dann gibt es insgesamt s Spannungen, und wir werden die Spannung im Stab $p_i p_j$ mit ω_{ij} bezeichnen.

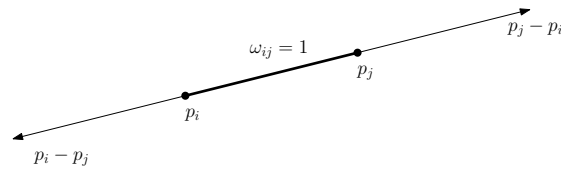


Abbildung 1.1: Stabkräfte bei Spannung = 1

Definition 4.4. Die Spannungen (ω_{ij}) lösen die Belastung (F_i) auf, wenn für jeden Knoten p_i gilt:

$$F_i = \sum_j \omega_{ij}(p_i - p_j) \quad (1.22)$$

(Man beachte, dass in der Summe (1.22) nur diejenigen j auftreten, für welche die Knoten p_i und p_j durch einen Stab verbunden sind.)

Definition 4.5. Ein Fachwerk heißt statisch bestimmt, wenn jede Gleichgewichtsbelastung eindeutig durch Spannungen aufgelöst werden kann.

Wir werden diese Definition auf der Sprache der linearen Algebra umformulieren.

4.3 Statische Bestimmtheit, Dimension und Rang

Welche Belastung wird durch eine einzelne Spannung der Größe 1 im Stab $p_i p_j$ aufgelöst? Graphisch ist diese Belastung auf der Abbildung 1.1 dargestellt, in den Bezeichnungen von (1.21) ist es

$$Z_{ij} := (0, 0, \dots, x_i - x_j, y_i - y_j, \dots, x_j - x_i, y_j - y_i, \dots, 0) \quad (1.23)$$

Die Nichtnulleinträge stehen an den Stellen $2i - 1, 2i$ und $2j - 1, 2j$, sodass $F_i = (x_i - x_j, y_i - y_j)$ und $F_j = (x_j - x_i, y_j - y_i)$ gilt, und die Kräfte an allen anderen Knoten gleich Null sind.

Lemma 24. Ein Fachwerk ist statisch bestimmt genau dann wenn die Belastungen Z_{ij} eine Basis des Vektorraums aller Gleichgewichtsbelastungen bilden.

Beweis. Das ist nur eine Umformulierung der Definition 4.5. Die Gleichung (1.22) ist äquivalent zu

$$F = \sum_{ij} \omega_{ij} Z_{ij} \quad (1.24)$$

Eine Belastung F aufzulösen heißt den Vektor F als Linearkombination von Vektoren Z_{ij} darzustellen. Die Auflösung ist eindeutig genau dann, wenn (Z_{ij}) linear unabhängig sind. Die eindeutige Auflösbarkeit bedeutet also, dass (Z_{ij}) ein linear unabhängiges Erzeugendensystem im Raum aller Gleichgewichtsbelastungen ist, d. h. eine Basis. \square

Nach Lemma 23 hat der Raum von Gleichgewichtsbelastungen Dimension $2k - 3$. Ein n -dimensionaler Vektorraum kann nicht von weniger als n Vektoren erzeugt werden; mehr als n Vektoren in einem solchen Raum sind linear abhängig. Daraus ergibt sich

Korollar 6. 1. Gilt $s < 2k - 3$, so existieren Gleichgewichtsbelastungen, die nicht durch Stabkräfte aufgelöst werden können.

2. Gilt $s > 2k - 3$, so sind die Belastungen Z_{ij} linear abhängig.⁴

3. Eine notwendige Bedingung für die statische Bestimmtheit eines ebenen Fachwerks ist

$$s = 2k - 3$$

Analog, eine notwendige Bedingung der statischen Bestimmtheit für die räumlichen Fachwerke ist

$$s = 3k - 6$$

Bemerkung 4.6. Das Korollar 6 kann auch über lineare Gleichungssysteme erläutert werden.

Bei gegebenen Kräften (F_i) ist die Gleichung (1.24) ein lineares Gleichungssystem aus $2k$ Gleichungen mit s Unbekannten ω_{ij} . Wegen der drei Gleichgewichtsbedingungen auf (F_i) können drei der Gleichungen ausgeschlossen werden; es bleiben $2k - 3$ Gleichungen. Bei $s < 2k - 3$ hat man dann mehr Gleichungen als Unbekannten, und das LGS kann nicht für jede Angabe von (F_i) aufgelöst werden. Analog, bei $s > 2k - 3$ hat das homogene LGS $\sum_{ij} \omega_{ij}(p_i - p_j)$ Nichtnulllösungen.

Die Starrheitsmatrix

Definition 4.7. Die aus den Zeilen Z_{ij} in (1.23) gebildete Matrix nennen wir die Starrheitsmatrix

$$M =_{ij} \begin{pmatrix} & i & & j & & \\ & \vdots & & \vdots & & \\ \cdots & p_i - p_j & \cdots & p_j - p_i & \cdots & \\ & \vdots & & \vdots & & \end{pmatrix}.$$

Die Zeilen der Starrheitsmatrix entsprechen den Stäben des Fachwerks, die Spalten den Knoten; dabei werden jedem Knoten 2 Spalten zugeordnet (für räumliche Fachwerke 3), für die x - und y -Koordinaten der an diesem Knoten angebrachten Stäben. Die Starrheitsmatrix ist also eine $s \times 2k$ -Matrix.

Der Zeilenraum der Starrheitsmatrix ist die lineare Hülle der Vektoren Z_{ij} ; wie im Beweis des Lemmas 24 bemerkt, besteht diese lineare Hülle aus allen auflösbaren Gleichgewichtsbelastungen.

Ein Fachwerk heißt *statisch unterbestimmt* wenn für ihn Gleichgewichtsbelastungen existieren, die nicht durch Stabkräfte aufgelöst werden können.

Im Lemma 24 haben wir eine hinreichende Bedingung $s < 2k - 3$ für statische Unterbestimmtheit gegeben. Eine hinreichende *und* notwendige Bedingung benutzt den Begriff des Rangs.

Satz 13. Ein Fachwerk ist genau dann statisch unterbestimmt, wenn für seine Starrheitsmatrix $\text{Rang } M < 2k - 3$ gilt.

⁴In diesem Fall können im Fachwerk Selbstspannungen auftreten, d. h. Spannungen, die ein Gleichgewichtssystem bilden.

Beweis. Der Vektorraum aller Gleichgewichtsbelastungen hat Dimension $2k - 3$. Die auflösbaren Gleichgewichtsbelastungen bilden seinen linearen Teilraum, der, wie oben bemerkt, auch als der Zeilenraum der Starrheitsmatrix beschrieben werden kann. Ein linearer Teilraum ist gleich dem ganzen Vektorraum genau dann, wenn er dieselbe Dimension hat. Die Dimension des Zeilenraums ist gleich $\text{Rang } M$. Das heißt, jede Gleichgewichtsbelastung ist auflösbar genau dann, wenn $\text{Rang } M = 2k - 3$ gilt. Daraus folgt die Behauptung. \square

Im Korollar 6 wurde die notwendige Bedingung $s = 2k - 3$ für statische Bestimmtheit gefunden. Ein statisch unbestimmtes Fachwerk mit $s = 2k - 3$ heißt ein *Ausnahmefachwerk*.

Lemma 25. *Ein ebenes Fachwerk mit $s = 2k - 3$ ist genau dann ein Ausnahmefachwerk, wenn die Zeilen der Starrheitsmatrix linear abhängig sind, d. h. genau dann wenn $\text{Rang } M < s$ gilt.*

Beweis. $s = 2k - 3$ ist die Dimension des Raums aller Gleichgewichtsbelastungen. Bilden s Vektoren keine Basis des Vektorraums, so sollen sie linear abhängig sein. \square

Statik und Kinematik

Die Anwendung des Satzes über Rang und Kern auf die Starrheitsmatrix M ergibt

$$\text{Rang } M + \text{Dim Kern } M = 2k \quad (1.25)$$

Beschäftigen wir uns näher mit dem Kern M . Ein Element des Kerns ist ein Vektor der Länge $2k$, wir werden es als

$$q = (q_1, q_2, \dots, q_k)$$

schreiben, wobei $q_i \in \mathbb{R}^2$ ist und als eine Verschiebung des Knoten p_i gedeutet wird.

Lemma 26. *Es gilt $q \in \text{Kern } M$ genau dann wenn (q_i) eine infinitesimale Verschiebung des Fachwerks darstellt.*

Beweis. Das Produkt Mq^\top ist ein Vektor der Länge s , dessen Komponenten den Stäben des Fachwerks entsprechen. Die i -te Komponente ist dabei gleich

$$(p_i - p_j) \cdot q_i + (p_j - p_i) \cdot q_j = (p_i - p_j) \cdot (q_i - q_j)$$

Es kann gezeigt werden, dass $(p_i - p_j) \cdot (q_i - q_j) = 0$ Längenerhaltung im Kleinen des Stabes $p_i p_j$ darstellt. Liegt q im Kern M , so werden alle Stablängen erhalten, d. h. q stellt eine infinitesimale Verschiebung des Fachwerks dar. \square

Unter den infinitesimalen Verschiebungen gibt es immer die trivialen: das sind Parallelverschiebungen (alle q_i gleich demselben Vektor) und infinitesimale Drehungen. Die trivialen Verschiebungen der Ebene bilden einen Vektorraum der Dimension 3; die trivialen Verschiebungen des Raums bilden einen Vektorraum der Dimension 6.

Korollar 7. *Ein ebenes Fachwerk ist genau dann unverschieblich, wenn $\text{Rang } M = 2k - 3$ gilt. Ein räumliches Fachwerk ist genau dann unverschieblich, wenn $\text{Rang } M = 3k - 6$ gilt.*

Beweis. Unverschieblichkeit bedeutet, dass jede infinitesimale Verschiebung trivial ist. Das heißt, $\dim \text{Kern } M = 3$ im ebenen Fall und $\dim \text{Kern } M = 6$ im räumlichen Fall. Mit der Formel (1.25) folgt nun die Behauptung. \square

Satz 14. *Ein Fachwerk ist genau dann unverschieblich, wenn jede seine Gleichgewichtsbelastung in Stabkräfte aufgelöst werden kann.*

Beweis. Nach dem Korollar 7 ist Unverschieblichkeit äquivalent zu $\text{Rang } M = 2k - 3$. Andererseits, nach dem Satz 13 ist $\text{Rang } M = 2k - 3$ notwendig und hinreichend dafür, dass jede Gleichgewichtsbelastung auflösbar ist. \square

4.4 Beispiele von Ausnahmefachwerken

Beispiel 4.8. Betrachte ein Fachwerk aus zwei Dreiecken, dessen Knoten paarweise verbunden sind. Ein solches Fachwerk hat $k = 6$ Knoten und $s = 9$ Stäbe, die Bedingung $s = 2k - 3$ ist erfüllt.

Ein Fachwerk von diesem Typ ist ein Ausnahmefachwerk genau dann, wenn die drei verbindenden Stäbe parallel oder zentral sind, siehe Abb. 1.2.

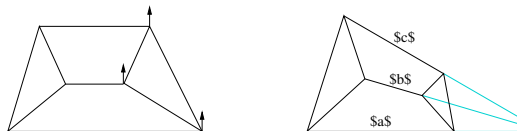


Abbildung 1.2: Ausnahmefachwerke; die Pfeile zeigen die Richtung der infinitesimalen Bewegung an.

Allgemeiner, verbindet man zwei statisch bestimmte Fachwerke durch drei Stäbe, dann entsteht ein statisch bestimmtes Fachwerk genau dann, wenn die verbindenden Stäbe weder parallel noch zentral sind (vgl. das 2. Bildungsgesetz für statisch bestimmte Fachwerke in der Technischen Mechanik).

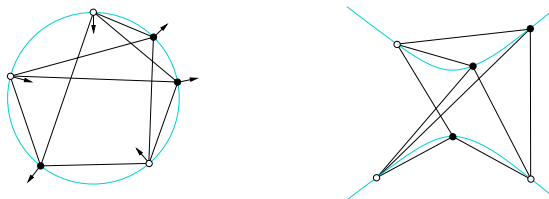


Abbildung 1.3: Ausnahmefachwerke; die Pfeile zeigen die Richtung der infinitesimalen Bewegung an.

Beispiel 4.9. Man kann 6 Knoten mit 9 Stäben auch auf die folgende Weise verbinden: die Knoten werden in zwei Gruppen je 3 aufgeteilt und jeder Knoten aus einer Gruppe wird mit jedem Knoten aus der anderen verbunden.

Ein solches Fachwerk kann nicht ohne Selbüberschneidung realisiert werden. Allerdings kann es aus Papp- oder Metallstreifen konstruiert werden.

Das Fachwerk dieses Typs ist statisch unbestimmt genau dann, wenn alle 6 Knoten auf einem Kegelschnitt liegen.

Betrachten wir nun räumliche Fachwerke.

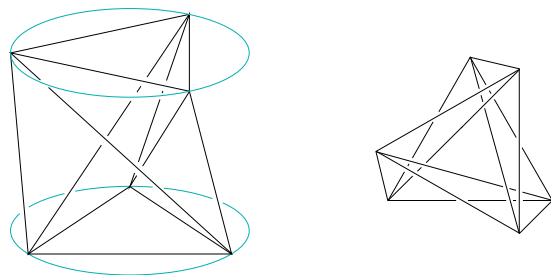


Abbildung 1.4: Das Schönhardt-Oktaeder

Beispiel 4.10. Ein Oktaeder hat $k = 6$ Ecken und $s = 12$ Kanten. Damit ist die Bedingung $s = 3k - 6$ erfüllt.

Das regelmäßige Oktaeder, und auch jedes konvexe Oktaeder ist statisch bestimmt und hiermit unbeweglich. Es gibt aber nichtkonvexe Oktaeder, die statisch unbestimmt sind. Ein Beispiel dafür ist das Schönhardt-Oktaeder, siehe Abb. 1.4 und 1.5.

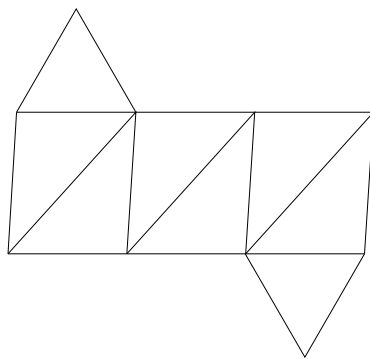


Abbildung 1.5: Entfaltung eines Schönhardt-Oktaeders

Die Seiten des Oktaeders können in schwarz und weiß gefärbt werden, so dass jede schwarze Seite nur an weiße grenzt. Dann kann das Kriterium der Beweglichkeit wie folgt formuliert werden.

Ein Oktaeder ist genau dann infinitesimal beweglich, wenn die Ebenen seiner 4 schwarzen Seiten sich in einem Punkt treffen oder, äquivalent, die Ebenen seiner 4 weißen Seiten sich in einem Punkt treffen.

Beispiel 4.11. Das Jessen-Ikosaeder ist ein Ausnahmefachwerk.

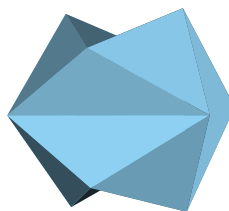


Abbildung 1.6: Das Jessen-Ikosaeder

5 Determinanten

5.1 Die Determinante einer 2×2 -Matrix

Definition 5.1. Die Determinante einer 2×2 -Matrix wird definiert durch

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Die 2×2 -Determinante ist uns schon ein paar Mal begegnet:

- Das Kreuzprodukt zweier Vektoren $v_1 = (x_1, y_1)$ und $v_2 = (x_2, y_2)$ in \mathbb{R}^2 wurde definiert als

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \times \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_1y_2 - x_2y_1 = \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$$

Dabei hat das Kreuzprodukt die folgende geometrische Bedeutung: $v_1 \times v_2$ ist gleich der Fläche des von v_1 und v_2 aufgespannten Parallelograms (mit dem Minus-Vorzeichen, falls (v_1, v_2) ein linkshändiges System bilden).

- In der Formel für die Inverse einer 2×2 -Matrix, Lemma 2, taucht auch die Determinante auf:

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad \text{für } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Insbesondere gilt

$$\det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = 0$$

\Leftrightarrow Vektoren v_1 und v_2 liegen auf einer Gerade
 \Leftrightarrow Vektoren v_1 und v_2 sind linear abhängig
 \Leftrightarrow die Matrix $\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$ ist nicht invertierbar

5.2 Die Determinante einer 3×3 -Matrix

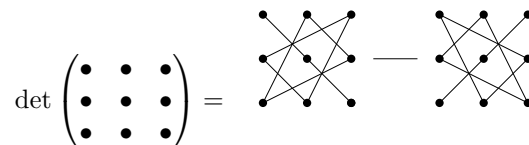
Definition 5.2. Die Determinante einer 3×3 -Matrix wird definiert durch

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} \quad (1.26)$$

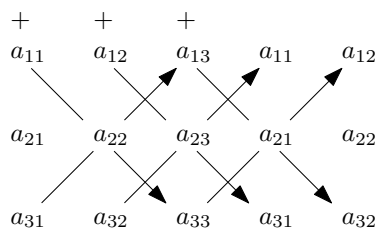
Die Regel von Sarrus

Um sich die Formel (1.26) zu merken, kann man sie graphisch darstellen.

Jedes der dreifachen Produkte in dieser Formel, wie etwa $a_{21}a_{32}a_{13}$, wird aus drei Matrixeinträgen gebildet, von denen keine zwei in der gleichen Zeile oder in der gleichen Spalte stehen. Notieren wir jedes Produkt durch Verbinden der entsprechenden Stellen in der Matrix, so wird es zu einem Dreieck oder einer der Matrixdiagonalen. Dann können wir die Determinante wie folgt beschreiben:



Eine andere graphische Darstellung der Determinante einer 3×3 -Matrix ist die Sarrus-Regel:



Man schreibt die ersten beiden Spalten noch einmal rechts neben A , addiert die längs der \searrow -Diagonalen zu bildenden Produkte und subtrahiert die längs \swarrow berechneten Produkte.

Beispiel 5.3.

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

$$\det A = 0 \cdot 1 \cdot 5 + (-2) \cdot (-2) \cdot 3 + (-2) \cdot 6 \cdot 3 - 3 \cdot 1 \cdot 3 - (-2) \cdot (-2) \cdot 5 - 0 \cdot 6 \cdot (-2) = 0 + 12 - 36 - 9 - 20 - 0 = -53$$

Eigenschaften der 3×3 -Determinanten

Analog zum vorigen Abschnitt, gilt Folgendes:

- Für $v_1 = (x_1, y_1, z_1)^\top$, $v_2 = (x_2, y_2, z_2)^\top$, $v_3 = (x_3, y_3, z_3)^\top$ ist $\det(v_1 \ v_2 \ v_3)$ gleich dem Volumen des von den Vektoren v_1 , v_2 und v_3 aufgespannten Spats (mit dem Minus-Vorzeichen, falls (v_1, v_2, v_3) ein linkshändiges System ist). In der Tat, wir wissen dass dieses Volumen dem Spatprodukt

$$[v_1, v_2, v_3] := v_1 \cdot (v_2 \times v_3)$$

gleich ist, und dieses wird umgeformt zu

$$\begin{aligned} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \cdot \begin{pmatrix} y_2 z_3 - y_3 z_2 \\ z_2 x_3 - z_3 x_2 \\ x_2 y_3 - x_3 y_2 \end{pmatrix} &= x_1(y_2 z_3 - y_3 z_2) + y_1(z_2 x_3 - z_3 x_2) + z_1(x_2 y_3 - x_3 y_2) = \\ &= x_1 y_2 z_3 + y_1 z_2 x_3 + z_1 x_2 y_3 - x_1 y_3 z_2 - y_1 z_3 x_2 - z_1 x_3 y_2 = \\ &= \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} \end{aligned}$$

- Wie man mit der Hilfe der Determinante die inverse Matrix berechnet, werden wir später erfahren, jetzt beweisen wir:

Lemma 27. *Eine 3×3 -Matrix ist genau dann invertierbar, wenn ihre Determinante ungleich Null ist.*

Beweis. In der Tat, das Volumen des von v_1 , v_2 und v_3 aufgespannten Spats ist ungleich Null genau dann, wenn die Vektoren v_1 , v_2 , v_3 eine Basis von \mathbb{R}^3 bilden. Andererseits, nach dem Satz 12 bilden n Vektoren in \mathbb{R}^n genau dann eine Basis wenn die aus ihnen als Spalten gebildete Matrix invertierbar ist. \square

5.3 Die Determinante einer $n \times n$ -Matrix

Für $v_2, v_3 \in \mathbb{R}^3$ können die Komponenten des Kreuzprodukts $v_2 \times v_3$ als 2×2 -Determinanten gedeutet werden. Deswegen kann die Formel $\det(v_1 \ v_2 \ v_3) = v_1 \cdot (v_2 \times v_3)$ auch wie folgt geschrieben werden:

$$\begin{aligned} \det(v_1 \ v_2 \ v_3) &= x_1 \det \begin{pmatrix} y_2 & y_3 \\ z_2 & z_3 \end{pmatrix} + y_1 \det \begin{pmatrix} z_2 & z_3 \\ x_2 & x_3 \end{pmatrix} + z_1 \det \begin{pmatrix} x_2 & x_3 \\ y_2 & y_3 \end{pmatrix} = \\ &= x_1 \det \begin{pmatrix} y_2 & y_3 \\ z_2 & z_3 \end{pmatrix} - y_1 \det \begin{pmatrix} x_2 & x_3 \\ z_2 & z_3 \end{pmatrix} + z_1 \det \begin{pmatrix} x_2 & x_3 \\ y_2 & y_3 \end{pmatrix} \end{aligned}$$

Analog hierzu wird die Determinante einer $n \times n$ -Matrix für alle n definiert.

Definition 5.4. Die Determinante einer $n \times n$ -Matrix A wird definiert durch

$$\det A := a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} - \dots + (-1)^{n+1} a_{n1} \det A_{n1} \quad (1.27)$$

Hier bezeichnet A_{i1} die $(n-1) \times (n-1)$ -Matrix, die aus A durch Streichen der ersten Spalte und der i -ten Zeile entsteht.

Die Formel (1.27) ist *rekursiv*: sie führt die Berechnung einer $n \times n$ -Determinante auf die Berechnung von $(n-1) \times (n-1)$ -Determinanten, die durch dieselbe Formel aus $(n-2) \times (n-2)$ -Determinanten berechnet werden können, usw.

Beispiel 5.5. Berechnen wir die Determinante von $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & -2 & 1 & 0 \\ 0 & 6 & 3 & -2 \\ 2 & 4 & 3 & 1 \end{pmatrix}$.

Es gilt

$$A_{11} = \begin{pmatrix} -2 & 1 & 0 \\ 6 & 3 & -2 \\ 4 & 3 & 1 \end{pmatrix} \quad A_{21} = \begin{pmatrix} 2 & 0 & 1 \\ 6 & 3 & -2 \\ 4 & 3 & 1 \end{pmatrix}$$

$$A_{31} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix} \quad A_{41} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 6 & 3 & -2 \end{pmatrix}$$

Nach der Berechnung der Determinanten dieser Matrizen (wobei es nicht nötig ist, $\det A_{31}$ zu berechnen) erhalten wir

$$\det A = 1 \cdot \det A_{11} - 3 \cdot \det A_{21} + 0 \cdot \det A_{31} - 2 \cdot \det A_{41} = -32 - 3 \cdot 24 - 2 \cdot (-16) = -72$$

Die vollständige Entwicklung der Determinante

Werden in (1.27) alle $\det A_{i1}$ nach der ersten Spalte entwickelt, dann alle entstehenden $(n-2) \times (n-2)$ -Determinanten usw., so erhält man die *vollständige Entwicklung* der Determinante

$$\det A = \sum_{i=(i_1, \dots, i_n)} (-1)^{\varepsilon(i)} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} \quad (1.28)$$

Hier erstreckt sich die Summe über alle Permutationen (i_1, i_2, \dots, i_n) der Zahlen $1, 2, \dots, n$, und die Zahl $\varepsilon(i)$ gleich der Anzahl der Vertauschungen ist, die erforderlich sind, um (i_1, i_2, \dots, i_n) in die natürliche Reihenfolge zu bringen. Die Zahl $(-1)^{\varepsilon(i)}$ ist gleich 1 oder -1 , je nachdem ob $\varepsilon(i)$ gerade oder ungerade ist.

Die vollständige Entwicklung der Determinante bei $n = 3$ ist die Formel (1.26). Die Anzahl der Summanden in (1.28) ist gleich $n!$, d. h. steigt sehr schnell an. Deswegen ist diese Formel bei großen n unpraktikabel.

Die Determinante einer Dreiecksmatrix

Für manche Matrizen kann die Determinante leicht berechnet werden.

Satz 15. Für eine obere Dreiecksmatrix gilt

$$\det \begin{pmatrix} a_1 & * & \dots & * \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & a_n \end{pmatrix} = a_{11}a_{22} \cdots a_{nn}$$

Insbesondere ist die Determinante einer Diagonalmatrix gleich dem Produkt ihrer Diagonalelemente, und $\det E = 1$.

Beweis. Induktionsbeweis: bei $n = 2$ gilt $\det \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = ad$.

Wegen $a_{21} = \dots = a_{n1} = 0$ reduziert sich (1.27) auf

$$\det A = a_{11} \det A_{11}$$

Die Matrix A_{11} ist eine obere $(n - 1) \times (n - 1)$ -Dreiecksmatrix, und ihre Determinante ist nach Induktionsvoraussetzung gleich $a_{22}a_{33} \cdots a_{nn}$. Daraus folgt die Behauptung. \square

5.4 Rechenregeln für Determinanten

Das Verhalten der Determinante bei elementaren Zeilenumformungen

Satz 16. a) Die Determinante ist linear in jeder Zeile:

$$\det \begin{pmatrix} Z_1 \\ \vdots \\ \alpha Z_i \\ \vdots \\ Z_n \end{pmatrix} = \alpha \det \begin{pmatrix} Z_1 \\ \vdots \\ Z_i \\ \vdots \\ Z_n \end{pmatrix} \quad \det \begin{pmatrix} Z_1 \\ \vdots \\ Z'_i + Z''_i \\ \vdots \\ Z_n \end{pmatrix} = \det \begin{pmatrix} Z_1 \\ \vdots \\ Z'_i \\ \vdots \\ Z_n \end{pmatrix} + \det \begin{pmatrix} Z_1 \\ \vdots \\ Z''_i \\ \vdots \\ Z_n \end{pmatrix}$$

b) Die Determinante wechselt das Vorzeichen beim Vertauschen zweier Zeilen:

$$\det \begin{pmatrix} \vdots \\ Z_i \\ \vdots \\ Z_j \\ \vdots \end{pmatrix} = - \det \begin{pmatrix} \vdots \\ Z_j \\ \vdots \\ Z_i \\ \vdots \end{pmatrix}$$

Beweis. Induktionsbeweis basierend auf der Formel (1.27). \square

Korollar 8. Enthält A zwei gleiche Zeilen oder eine Nullzeile, so gilt $\det A = 0$.

Beweis. Nach dem Punkt b) des Satzes 16 wechselt die Determinante beim Vertauschen zweier Zeilen das Vorzeichen. Andererseits, wenn die Zeilen zueinander gleich sind, ändert sich die Matrix dabei nicht. D. h. $\det A = -\det A$, und daraus folgt $\det A = 0$.

Sei $Z_i = 0$. Nach dem Punkt a) gilt $\det \begin{pmatrix} Z_1 \\ \vdots \\ 2Z_i \\ \vdots \\ Z_n \end{pmatrix} = 2 \det \begin{pmatrix} Z_1 \\ \vdots \\ Z_i \\ \vdots \\ Z_n \end{pmatrix}$. Andererseits

$\det \begin{pmatrix} Z_1 \\ \vdots \\ 2Z_i \\ \vdots \\ Z_n \end{pmatrix} = \det \begin{pmatrix} Z_1 \\ \vdots \\ Z_i \\ \vdots \\ Z_n \end{pmatrix}$. Folglich $\det A = 2 \det A$, und das heißt $\det A = 0$. \square

Lemma 28. *Bei Anwendung von elementaren Zeilenumformungen multipliziert sich die Determinante mit der Determinante der entsprechenden Elementarmatrix: entsteht \tilde{A} aus A durch eine elementare Zeilenumformung, so gilt $\det \tilde{A} = \det \tilde{E} \cdot \det A$, wobei \tilde{E} die der Umformung zugehörige Matrix ist.*

Beweis. Die Determinanten von Elementarmatrizen sind leicht zu berechnen, z. B.

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

Für die ersten zwei Typen von elementaren Zeilenumformungen folgt jetzt die Behauptung aus dem Satz 16, Punkt b) und der ersten Formel des Punktes a). Für den dritten Typ haben wir:

$$\det \begin{pmatrix} Z_1 \\ \vdots \\ Z_i + \alpha Z_j \\ \vdots \\ Z_n \end{pmatrix} = \det \begin{pmatrix} Z_1 \\ \vdots \\ Z_i \\ \vdots \\ Z_n \end{pmatrix} + \alpha \det \begin{pmatrix} Z_1 \\ \vdots \\ Z_j \\ \vdots \\ Z_n \end{pmatrix} = \det A + 0,$$

denn der zweite Summand ist eine Matrix mit zwei gleichen Zeilen. Also ändert sich die Determinante bei einer Umformung des dritten Typs nicht. Das entspricht der Multiplikation mit $1 = \det \tilde{E}$ der Elementarmatrix vom dritten Typ. \square

Lemma 28 gibt ein Verfahren zur Berechnung der Determinante. Durch elementare Zeilenumformungen wird die Matrix A in Zeilenstufenform gebracht. Dabei werden die Determinanten der zugehörigen Elementarmatrizen notiert ($-1, \alpha$ oder 1). Erhält man am Ende eine Matrix mit einer Nullzeile, so ist $\det A = 0$. Enthält die Zeilenstufenmatrix keine Nullzeile, so ist sie eine obere Dreiecksmatrix, deren Determinante leicht zu berechnen ist. Die Division durch das Produkt der notierten Zahlen ergibt $\det A$.

Bei großen n läuft dieses Verfahren wesentlich schneller als die vollständige Entwicklung der Determinante.

Rechenregeln

Satz 17. a) Matrix A ist invertierbar genau dann wenn $\det A \neq 0$ gilt.

b) $\det(AB) = \det A \cdot \det B$

c) $\det A^T = \det A$

Beweis. a): Das folgt aus dem oben beschriebenen Verfahren zur Berechnung der Determinante: entsteht beim Umformen auf Zeilenstufenform eine Nullzeile, so bedeutet es, dass A nicht invertierbar ist, und bedeutet auch, dass die Determinante gleich Null ist. Entsteht eine obere Dreiecksmatrix, deren alle Diagonalelemente ungleich Null sind, so war A invertierbar, und die Berechnung der Determinante ergibt auch eine Zahl ungleich Null.

b): Wenn A oder B nicht invertierbar sind, so ist das Produkt AB auch nicht invertierbar. Dann stehen auf beiden Seiten der Gleichung die Nullen. Wenn A und B invertierbar sind, dann können sie mit dem Gauß-Jordan-Verfahren als Produkte von Elementarmatrizen dargestellt werden:

$$A = \tilde{E}_s \cdots \tilde{E}_1, \quad B = \tilde{E}'_t \cdots \tilde{E}'_1$$

Aus Lemma 28 folgt

$$\det A = \det \tilde{E}_s \cdots \det \tilde{E}_1, \quad \det B = \det \tilde{E}'_t \cdots \det \tilde{E}'_1$$

Und aus $AB = \tilde{E}_s \cdots \tilde{E}_1 \tilde{E}'_t \cdots \tilde{E}'_1$ folgt

$$\det(AB) = \det \tilde{E}_s \cdots \det \tilde{E}_1 \cdot \det \tilde{E}'_t \cdots \det \tilde{E}'_1 = \det A \det B$$

c): Ist A nicht invertierbar, so gilt $\det A = 0 = \det A^T$. Ist A invertierbar, dann gilt

$$A = \tilde{E}_s \cdots \tilde{E}_1 \Rightarrow A^T = \tilde{E}_1^T \cdots \tilde{E}_s^T$$

Für jede Elementarmatrix kann $\det \tilde{E} = \det \tilde{E}^T$ leicht nachgewiesen werden. Deswegen gilt

$$\det A = \det \tilde{E}_s \cdots \det \tilde{E}_1 = \det \tilde{E}_1^T \cdots \det \tilde{E}_s^T = \det A^T$$

□

Folgerungen aus dem Multiplikationsgesetz

Korollar 9. a) $\det(AB) = \det(BA)$

b) $\det(A^{-1}) = \frac{1}{\det A}$ (falls A invertierbar ist)

c) $\det(C^{-1}AC) = \det A$ für jede invertierbare Matrix C

Beweis. a): $\det(AB) = \det A \det B = \det B \det A = \det(BA)$

b): $(\det A)(\det A^{-1}) = \det(AA^{-1}) = \det E = 1$

c): $\det(C^{-1}AC) = \det(C^{-1}) \det A \det C = \det A$

□

Die Entwicklung von $\det A$ nach einer beliebigen Zeile oder Spalte

Die Formel (1.27) wird *Entwicklung nach der ersten Spalte* genannt. Mit Hilfe des vorherigen Abschnittes können auch Formeln für Entwicklung nach einer beliebigen Spalte oder Zeile hergeleitet werden.

Erstens gilt $\det A = \det A^\top$. Wird $\det A^\top$ nach der ersten Spalte entwickelt, so entsteht die Formel

$$\begin{aligned}\det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} - \dots + (-1)^{n+1} a_{1n} \det A_{1n} = \\ &= \sum_{j=1}^n (-1)^{j+1} a_{1j} \det A_{1j},\end{aligned}$$

die *Entwicklung nach der ersten Zeile* heißt.

Zweitens, mittels Spaltenvertauschungen kann man die *Entwicklung nach der j -ten Spalte* herleiten:

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$$

und, analog, die *Entwicklung nach der i -ten Zeile*:

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$$

Die Adjunkte und die Inverse

Definition 5.6. Die Matrix $A^* = (a_{ij}^*)$ mit $a_{ij}^* = (-1)^{i+j} \det A_{ji}$ heißt die Adjunkte von A .

Insbesondere, für eine 3×3 -Matrix:

$$A^* = \begin{pmatrix} \det A_{11} & -\det A_{21} & \det A_{31} \\ -\det A_{12} & \det A_{22} & -\det A_{32} \\ \det A_{13} & -\det A_{23} & \det A_{33} \end{pmatrix}$$

Satz 18. Die Inverse ist ein $\frac{1}{\det A}$ -Vielfaches der Adjunkten:

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

Beispiel 5.7. Die Adjunkte einer 2×2 -Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist

$$A^* = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Für die Inverse erhalten wir die uns bereits bekannte Formel

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

6 Lineare Abbildungen und Eigenwerte

6.1 Definition und Beispiele

Definition 6.1. Seien V und W zwei Vektorräume. Eine Abbildung $f: V \rightarrow W$ heißt linear, wenn gilt:

1. $f(\alpha v) = \alpha f(v)$ für alle $\alpha \in \mathbb{R}$, $v \in V$;
2. $f(u + v) = f(u) + f(v)$ für alle $u, v \in V$.

Aus den Eigenschaften 1. und 2. folgt auch

$$f(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 f(v_1) + \dots + \alpha_n f(v_n) \quad (1.29)$$

für alle $\alpha_i \in \mathbb{R}$, $v_i \in V$.

Beispiel 6.2. • Sei V der Raum aller differenzierbaren Funktionen, und sei W der Raum aller Funktionen. Sei $f: V \rightarrow W$ die Abbildung, die jeder Funktion φ ihre Ableitung zuordnet:

$$f(\varphi) = \varphi'$$

Wegen $(\alpha\varphi)' = \alpha\varphi'$ und $(\varphi + \psi)' = \varphi' + \psi'$ ist die Abbildung f linear.

- Sei V der Raum aller stetigen Funktionen auf dem Intervall $[0, 1]$, und sei $W = \mathbb{R}$ der eindimensionale Vektorraum. Die Abbildung

$$f(\varphi) = \int_0^1 \varphi dx$$

ist linear.

Eine lineare Abbildung $f: V \rightarrow V$ (also von einem Vektorraum in sich selbst) heißt auch *lineare Transformation*. Einige Beispiele linearer Transformationen von \mathbb{R}^2 :

Beispiel 6.3. • Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine Drehung um den Nullpunkt mit dem Drehwinkel φ . Dann ist f linear, denn:

1. $f(\alpha v) = \alpha f(v)$ ist erfüllt: zuerst strecken, dann drehen liefert das gleiche Ergebnis wie zuerst drehen, dann strecken.
2. $f(u + v) = f(u) + f(v)$ ist erfüllt: zuerst zwei Vektoren addieren, dann die Summe drehen liefert das gleiche Ergebnis wie zuerst beide Vektoren drehen, dann addieren.

- Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die orthogonale Projektion auf die x -Achse:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} \quad (1.30)$$

Dann ist f linear, was entweder geometrisch wie im vorigen Beispiel oder durch die Formel festgestellt werden kann.

- Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Parallelverschiebung um den Vektor $(1, 0)^\top$:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 1 \\ y \end{pmatrix}$$

Dann ist f *nicht* linear. In der Tat werden beide Eigenschaften nicht erfüllt, z. B. die erste: für $v = (1, 0)^\top$ und $\alpha = 2$ gilt

$$f(2v) = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 2f(v)$$

Beispiel 6.4. Eine Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann linear, wenn $f(x) = ax$ für ein $a \in \mathbb{R}$ ist. In der Tat, wegen der ersten Eigenschaft der linearen Abbildungen gilt $f(x) = xf(1)$. Setzt man $a := f(1)$, so erhält man $f(x) = ax$ für alle x .

Hier gibt es einen Unterschied in der Terminologie zwischen der Analysis und der Linearen Algebra. Die Funktionen der Form $f(x) = ax + b$ werden oft *lineare Funktionen* genannt; bei $b \neq 0$ ist eine solche Funktion aber keine lineare Abbildung im Sinne der Definition 6.1.

6.2 Lineare Abbildungen und Matrizen

Satz 19. Die linearen Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ werden wie folgt charakterisiert:

1. Sei A eine $m \times n$ -Matrix. Dann ist die durch

$$f(x) = Ax$$

gegebene Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear.

2. Auch umgekehrt, für jede lineare Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt es eine $m \times n$ -Matrix A , sodass $f(x) = Ax$ gilt.

Beweis. Für 1: das folgt aus den Eigenschaften der Matrizenmultiplikation

$$f(\alpha x) = A(\alpha x) = \alpha Ax = \alpha f(x) \quad f(x+y) = A(x+y) = Ax + Ay = f(x) + f(y)$$

Für 2 ist der Beweis etwas aufwendiger. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine beliebige lineare Abbildung. Betrachten wir die Bilder $f(e_1), \dots, f(e_n)$ der Vektoren der Standardbasis. Dann gilt Folgendes:

Durch die Vektoren $f(e_1), \dots, f(e_n) \in \mathbb{R}^m$ ist die Abbildung f eindeutig festgelegt und es gilt

$$f(x) = x_1 f(e_1) + \dots + x_n f(e_n) \quad \text{für } x = (x_1, \dots, x_n)^\top \quad (1.31)$$

In der Tat, es gilt $x = x_1 e_1 + \dots + x_n e_n$, und die Anwendung von (1.29) liefert (1.31).

Sei nun

$$f(e_1) = (a_{11}, a_{21}, \dots, a_{m1})^\top, \dots, f(e_n) = (a_{1n}, a_{2n}, \dots, a_{mn})^\top$$

Dann kann (1.31) wie folgt umgeschrieben werden:

$$f(x) = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = Ax$$

mit der $m \times n$ -Matrix $A = (a_{ij})$. □

Die Abbildungsmatrix

Der obige Satz besagt, dass die linearen Abbildungen von \mathbb{R}^n nach \mathbb{R}^m und die $m \times n$ -Matrizen "dasselbe sind". Jede Matrix entspricht einer linearen Abbildung und jede lineare Abbildung f wird durch eine Matrix gegeben, die dann die *Abbildungsmatrix* von f heißt. Der Beweis des Satzes liefert uns auch die Regel zum Bilden der Abbildungsmatrix:

Die Bilder der Standardbasisvektoren sind die Spalten der Abbildungsmatrix.

Beispiel 6.5. Bei der orthogonalen Projektion auf die x -Achse (1.30) wird der erste Basisvektor $(1, 0)^\top$ auf sich selbst abgebildet, und der zweite auf den Nullvektor. Deswegen ist die Abbildungsmatrix gleich $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Das kann auch unmittelbar geprüft werden:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$$

Satz 20. *Die Matrizenmultiplikation entspricht der Verkettung von Abbildungen. Und zwar, sind $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ lineare Abbildungen mit Abbildungsmatrizen A , bzw. B , dann ist $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^p$ ebenfalls linear und hat die Abbildungsmatrix BA .*

Beweis. Die Abbildung f bildet x auf Ax , die Abbildung g bildet Ax auf $B(Ax) = (BA)x$. Also gilt

$$(g \circ f)(x) = (BA)x$$

Das bedeutet, dass $g \circ f$ linear ist und die Abbildungsmatrix BA hat. □

Beachte, dass im obigen Satz die Matrix A eine $m \times n$ Matrix, und B eine $p \times m$ Matrix ist. Deswegen ist das Produkt BA definiert und es ist eine $p \times n$ -Matrix. Solche Matrizen entsprechen linearen Abbildungen von \mathbb{R}^n in \mathbb{R}^p .

Von nun an betrachten wir nur noch lineare Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Sie entsprechen den quadratischen $n \times n$ -Matrizen. Aus den Eigenschaften der Einheitsmatrix und der Inversen, sowie aus dem Satz 20 folgt

Korollar 10. • *Die Einheitsmatrix E ist die Matrix der Identitätsabbildung.*

- *Eine lineare Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist umkehrbar genau dann, wenn ihre Abbildungsmatrix A invertierbar ist. Ist das der Fall, dann ist A^{-1} die Abbildungsmatrix von f^{-1} .*

Die Änderung des Volumens

Betrachten wir den von den Standardbasisvektoren aufgespannten n -Spat Q (bei $n = 2$ Quadrat, bei $n = 3$ Würfel). Eine lineare Transformation $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ bildet diesen orthogonalen Spat auf einen eventuell "schiefen" $f(Q)$. Der Spat $f(Q)$ wird von den Bildern $f(e_1), \dots, f(e_n)$ der Basisvektoren aufgespannt, deswegen ist sein Volumen (bei $n = 2$ sein Flächeninhalt) gleich $|\det(f(e_1) \dots f(e_n))|$. Die Matrix $(f(e_1) \dots f(e_n))$ ist aber die Abbildungsmatrix von f . Deswegen gilt

$$\text{Vol}(Q) = 1, \quad \text{Vol}(f(Q)) = |\det A|,$$

wobei A die Abbildungsmatrix von f bezeichnet.

Satz 21. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Transformation von \mathbb{R}^n mit der Abbildungsmatrix A . Dann ist $|\det A|$ der Volumenverzerrungsfaktor der Abbildung f .

Beispiel 6.6. Die Transformation $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit der Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ bildet e_1 auf sich selbst ab, und e_2 auf $e_1 + e_2$. Das von e_1 und e_2 aufgespannte Quadrat wird auf das von e_1 und $e_1 + e_2$ aufgespannte Parallelogramm abgebildet. Eine solche Abbildung heißt *Scherung*.

Wegen $\det \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = 1$ ist diese Abbildung flächentreu.

6.3 Längentreue Abbildungen und orthogonale Matrizen

Erinnerung über Längen- und Winkelmessung in \mathbb{R}^n

Ähnlich zu \mathbb{R}^2 und \mathbb{R}^3 wird das *Skalarprodukt* zweier Vektoren in \mathbb{R}^n definiert:

$$x \cdot y := x^\top y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

Erstens erlaubt das Skalarprodukt die Längen zu berechnen. So, die Länge (der *Betrag*) des Vektors x :

$$|x| := \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

Und der Abstand zwischen zwei Punkten $x, y \in \mathbb{R}^n$:

$$\text{dist}(x, y) := |x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (1.32)$$

Zweitens kann man auch die Winkel mit Hilfe des Skalarprodukts berechnen. Und zwar wird der Winkel φ zwischen zwei Vektoren $x, y \in \mathbb{R}^2$ definiert/berechnet durch

$$\cos \varphi = \frac{x \cdot y}{|x| |y|}$$

Insbesondere ist der Winkel φ kleiner als 90° genau dann wenn $x \cdot y > 0$ gilt.

Bei $x \cdot y = 0$ gilt $\varphi = 90^\circ$ und die Vektoren x und y heißen *orthogonal* zueinander.

Beispiel 6.7. • Die Vektoren $(1, -1)$ und $(1, 1)$ sind orthogonal. (Das sieht man auch, wenn man diese Vektoren anzeichnet.)

• Die Vektoren $(1, 1, 1, 1)$ und $(1, 1, -1, -1)$ in \mathbb{R}^4 sind orthogonal.

Orthogonale und längentreue Abbildungen

Definition 6.8. Eine lineare Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt orthogonal, wenn sie die Skalarprodukte von Vektoren nicht ändert:

$$f(x) \cdot f(y) = x \cdot y \quad \text{für alle } x, y \in \mathbb{R}^n$$

Direkt aus der Definition folgt, dass eine orthogonale Abbildung die Beträge der Vektoren erhält:

$$|f(x)| = \sqrt{f(x) \cdot f(x)} = \sqrt{x \cdot x} = |x| \quad \text{für alle } x \in \mathbb{R}^n$$

Es gilt auch mehr:

Lemma 29. Eine orthogonale Abbildung ist längen- und winkeltreu. Das heißt, für je zwei Punkte $P, Q \in \mathbb{R}^n$ gilt

$$\text{dist}(f(P), f(Q)) = \text{dist}(P, Q)$$

und für je zwei Vektoren $x, y \in \mathbb{R}^n$ ist der Winkel zwischen $f(x)$ und $f(y)$ gleich dem Winkel zwischen x und y .

Beweis. Wie üblich, identifizieren wir die Punkte mit ihren Ortsvektoren. Deswegen gilt nach der Formel (1.32)

$$\text{dist}(P, Q) = |P - Q| = \sqrt{(P - Q) \cdot (P - Q)}$$

Wegen der Linearität von f gilt $\text{dist}(f(P), f(Q)) = |f(P) - f(Q)| = |f(P - Q)|$, und wegen der Orthogonalität von f gilt

$$|f(P - Q)| = \sqrt{f(P - Q) \cdot f(P - Q)} = \sqrt{(P - Q) \cdot (P - Q)} = |P - Q|$$

Folglich $\text{dist}(f(P), f(Q)) = \text{dist}(P, Q)$ und der erste Teil des Satzes ist bewiesen.

Der zweite Teil ist analog und kann kurz wie folgt zusammengefasst werden: da die Definition des Winkels nur Skalarprodukte benutzt, und die Skalarprodukte bei orthogonalen Transformationen erhalten bleiben, bleiben auch alle Winkel erhalten. \square

Man kann auch zeigen, dass jede längentreue lineare Abbildung orthogonal ist.

Beispiel 6.9. Eine Drehung der Ebene um den Nullpunkt ist längentreu. Deswegen ist das auch eine orthogonale Abbildung.

Orthogonale Abbildungen, orthogonale Matrizen und orthonormale Basen

Hier werden wir die Frage angehen, was die Abbildungsmatrizen der orthogonalen Transformationen unter allen quadratischen Matrizen ausmacht. Zuerst brauchen wir einen neuen Begriff.

Definition 6.10. Eine Basis (v_1, v_2, \dots, v_n) von \mathbb{R}^n heißt orthogonal, wenn die Vektoren v_i paarweise orthogonal sind, d. h.

$$v_i \cdot v_j = 0 \quad \text{für } i \neq j \text{ gilt}$$

Eine Basis heißt orthonormal, wenn sie orthogonal ist und alle Basisvektoren Länge 1 haben:

$$v_i \cdot v_j = 0 \text{ für } i \neq j \quad \text{und} \quad v_i \cdot v_i = 1 \text{ für alle } i \quad (1.33)$$

Beispiel 6.11. • Die Standardbasis von \mathbb{R}^n ist eine orthonormale Basis.

- Die Vektoren $v_1 = (1, -1)$ und $v_2 = (1, 1)$ bilden eine orthogonale Basis von \mathbb{R}^2 , die aber nicht orthonormal ist, da v_1 und v_2 beide Länge $\sqrt{2}$ haben.

Die Vektoren $\frac{1}{\sqrt{2}}v_1$ und $\frac{1}{\sqrt{2}}v_2$ bilden eine orthonormale Basis.

Satz 22. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Transformation mit der Abbildungsmatrix A . Dann sind die folgenden drei Aussagen äquivalent:

- Die Abbildung f ist orthogonal.
- Die Spalten von A bilden eine orthonormale Basis von \mathbb{R}^n .
- Es gilt $A^\top A = E$.

Beweis. a) \Rightarrow b): Wenn die Abbildung f orthogonal ist, dann ist für jeden Standardbasisvektor e_i ihr Bild $f(e_i)$ ein Vektor der Länge 1. Außerdem, da für alle $i \neq j$ die Vektoren e_i und e_j orthogonal zueinander sind, sind auch ihre Bilder $f(e_i)$ und $f(e_j)$ zueinander orthogonal. Kurz gesagt, bildet f die Standardbasis auf eine orthonormale Basis ab. Die Spalten der Abbildungsmatrix A sind aber die Bilder der Standardbasisvektoren unter f . Das beweist die Behauptung.

b) \Rightarrow c): Sei $v_i = f(e_i)$. Die Bedingungen (1.33) können umgeschrieben werden als

$$v_i^\top v_i = 1 \text{ für alle } i \quad \text{und} \quad v_i^\top v_j = 0 \text{ für alle } i \neq j$$

Da v_i^\top die i -te Zeile der Matrix A^\top und v_j die j -te Spalte von A , steht $v_i^\top v_j$ auf der Kreuzung der i -ten Zeile und der j -ten Spalte von $A^\top A$. Aus den obigen Gleichungen folgt nun $A^\top A = E$.

c) \Rightarrow a): Aus $A^\top A = E$ folgt für alle $x, y \in \mathbb{R}^n$

$$(Ax) \cdot (Ay) = (Ax)^\top (Ay) = x^\top A^\top Ay = x^\top Ey = x^\top y = x \cdot y$$

Und das heißt, dass die Abbildung $x \mapsto Ax$ eine orthogonale Abbildung ist. \square

Definition 6.12. Eine quadratische Matrix A heißt orthogonal, wenn gilt

$$A^\top A = E$$

Der obige Satz besagt also, dass die orthogonalen Matrizen genau die Abbildungsmatrizen der orthogonalen Abbildungen sind. Außerdem kann man eine orthogonale Matrix daran erkennen, dass ihre Spalten eine orthonormale Basis von \mathbb{R}^n bilden.

Satz 23. Für jede orthogonale Matrix A gilt

- A ist invertierbar und es gilt $A^{-1} = A^\top$;
- $\det A = \pm 1$.

Beweis. a): Ist A orthogonal, so gilt per Definition $A^\top A = E$. Das heißt, A^\top ist eine Linksinverse zu A . Nach Satz 12 bedeutet es, dass $A^\top = A^{-1}$ gilt.

b): Es gilt

$$1 = \det E = \det(A^\top A) = \det(A^\top) \cdot \det A = \det A \cdot \det A = (\det A)^2$$

Folglich $\det A = \pm 1$. □

6.4 Beispiele von linearen Abbildungen und Abbildungsmatrizen

Beispiel 6.13. • Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Drehung um Nullpunkt mit dem Drehwinkel φ . Um die Abbildungsmatrix zu bestimmen, finden wir die Bilder der Standardbasisvektoren:

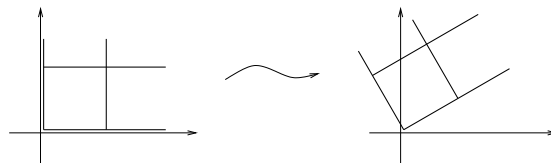
$$f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

Es gilt also

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

und die Abbildung f wird gegeben durch

$$f \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \varphi - y \sin \varphi \\ x \sin \varphi + y \cos \varphi \end{pmatrix}$$



- Sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ die Drehung um die x -Achse mit dem Drehwinkel φ . (Damit wird eine Drehung um φ im Uhrzeigersinn gemeint, beim Blicken in der Richtung der x -Achse.) Dann gilt $f(e_1) = e_1$, und die Vektoren e_2, e_3 werden innerhalb der yz -Ebene in der positiven Richtung um φ gedreht. D. h.

$$f \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \cos \varphi \\ \sin \varphi \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

und die Abbildungsmatrix ist

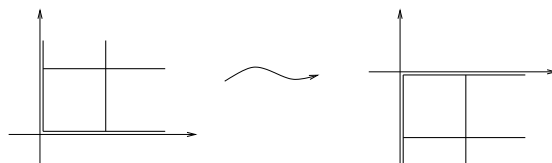
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}$$

- Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Spiegelung an der x -Achse. Diese Abbildung ist linear. Die Bilder der Standardbasisvektoren:

$$f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Die Abbildungsmatrix:

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

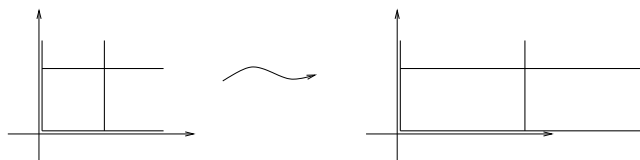


- Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Streckung entlang der x -Achse mit dem Streckungsfaktor 2. Dann

$$f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Die Abbildungsmatrix:

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



- Allgemein, eine Diagonalmatrix

$$A = \text{Diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$$

entspricht den gleichzeitigen Streckungen entlang aller Koordinatenachsen mit dem Faktor α_i für die x_i -Achse (bei $\alpha_i < 0$ wird die Streckung mit der Spiegelung an einer Hyperebene komponiert).

Die folgenden Fragen sind schwieriger zu beantworten:

- Welche Matrix entspricht der Streckung mit dem Faktor 2 entlang der Geraden mit dem Richtungsvektor $(1, -1)$?
- Wie kann die Abbildung mit einer gegebenen Matrix geometrisch beschrieben werden?

6.5 Abbildungsmatrix bezüglich einer beliebigen Basis

Nimmt man anstelle von Standardbasis (e_i) eine andere Basis (e'_i) , so wird jeder linearen Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^n$ die Abbildungsmatrix bezüglich dieser neuen Basis zugeordnet. Die Definition ist analog zur Definition im Falle der Standardbasis.

Definition 6.14. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung und sei e'_1, \dots, e'_n eine Basis von \mathbb{R}^n . Die Abbildungsmatrix von f bezüglich der Basis (e'_i) ist eine $n \times n$ -Matrix A' deren j -te Spalte aus den Koordinaten des Vektors $f(e'_j)$ bezüglich der Basis (e'_i) besteht.

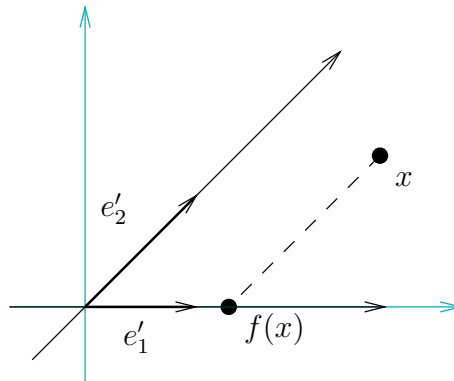
Genau wie mit der Abbildungsmatrix im Standardbasis, folgt aus dieser Definition: sind $(x'_1, \dots, x'_n)^\top$ Koordinaten eines Vektors bezüglich der Basis (e'_i) , so hat das Bild dieses Vektors Koordinaten

$$A' \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} a'_{11}x'_1 + \dots + a'_{1n}x'_n \\ \vdots \\ a'_{n1}x'_1 + \dots + a'_{nn}x'_n \end{pmatrix}$$

bezüglich der Basis (e'_i) .

Beispiel 6.15. Seien $e'_1 = (1, 0)^\top$, $e'_2 = (1, 1)^\top$. Was für eine Abbildung repräsentiert die Matrix $A' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ bezüglich der Basis (e'_1, e'_2) ?

Per Definition stellen die Spalten von A' die Koordinaten der Bilder der Basisvektoren dar. Das heißt, $f(e'_1) = 1 \cdot e'_1 + 0 \cdot e'_2 = e'_1$ und $f(e'_2) = 0$. Diese Abbildung ist die Projektion auf die x -Achse entlang der Geraden mit dem Richtungsvektor $(1, 1)$.



Durch welche Matrix wird diese Abbildung bezüglich der Standardbasis repräsentiert? Dafür brauchen wir die Bilder der Standardbasisvektoren zu finden. Wegen $e_1 = e'_1$ gilt $f(e_1) = e_1$. Und wegen $e_2 = e'_2 - e'_1$ gilt $f(e_2) = f(e'_2) - f(e'_1) = 0 - e_1 = -e_1$. Die Abbildungsmatrix bezüglich der Standardbasis ist also

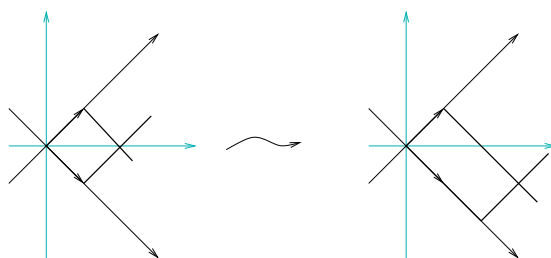
$$A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

Beispiel 6.16. Um die Streckung mit dem Faktor 2 entlang des Vektors $(1, -1)$ (die zu $(1, -1)$ senkrechten Geraden sollen auf $(1, -1)$ senkrecht bleiben, also Streckung ohne Scherung) durch eine Matrix darzustellen, wählen wir eine orthogonale Basis $e'_1 = (1, -1)^\top$, $e'_2 = (1, 1)$. Die Bilder dieser Vektoren unter der Streckung sind

$$f(e'_1) = 2e'_1, \quad f(e'_2) = e'_2$$

Deswegen ist die Abbildungsmatrix *bezüglich der Basis* (e'_1, e'_2) gleich

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



Um die Abbildungsmatrix A bezüglich der Standardbasis zu finden, brauchen wir eine allgemeine Methode, und die wird im nächsten Abschnitt beschrieben.

6.6 Basiswechsel

Änderung der Vektorkoordinaten bei Basiswechsel

Wie üblich, bezeichnen wir mit e_i , $i = 1, \dots, n$ die Standardbasisvektoren von \mathbb{R}^n . Sei e'_1, e'_2, \dots, e'_n eine andere Basis von \mathbb{R}^n . Wir werden (e_i) als die *alte* Basis, und (e'_i) als die *neue* Basis bezeichnen.

Sei $x = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$ ein Vektor; (x_i) sind seine *alten* Koordinaten. Der Vektor x kann aber auch bezüglich der neuen Basis (e'_i) zerlegt werden: es existieren Zahlen $(x'_1, x'_2, \dots, x'_n)$, sodass gilt

$$x = x'_1 e'_1 + x'_2 e'_2 + \dots + x'_n e'_n$$

Wir wollen nun die Regel finden, wie diese *neuen* Koordinaten berechnet werden können.

Definition 6.17. Seien $e'_i = (b_{1i}, b_{2i}, \dots, b_{ni})^\top = b_{1i}e_1 + \dots + b_{ni}e_n$, und allgemein

$$e'_i = \begin{pmatrix} b_{1i} \\ \vdots \\ b_{ni} \end{pmatrix} = b_{1i}e_1 + \dots + b_{ni}e_n \quad (1.34)$$

die Koordinatendarstellungen der neuen Basisvektoren. Die Matrix

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

heißt die Basiswechsellmatrix.

Satz 24. Die Beziehung zwischen den alten und den neuen Koordinaten eines Vektors wird mittels Basiswechselmatrix gegeben:

$$x = Bx' \quad \text{oder, äquivalent} \quad x' = B^{-1}x$$

(Hier sind $x = (x_1, \dots, x_n)^\top$ und $x' = (x'_1, \dots, x'_n)^\top$ die als Spalten geschriebenen Koordinaten eines Vektors in der alten, bzw. neuen Basis.)

Beweis. Die Gleichungen (1.34) können geschrieben werden als

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)B \tag{1.35}$$

Die Tatsache, dass (x_i) , bzw. (x'_i) die Koordinaten eines und desselbes Vektors in zwei Basen sind, führt zur Gleichung

$$(e_1, \dots, e_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (e'_1, \dots, e'_n) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

Das Einsetzen von (1.35) liefert

$$(e_1, \dots, e_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (e_1, \dots, e_n)B \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

Daraus folgt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = B \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

und der Satz ist bewiesen. □

Beispiel 6.18. Für die Basis

$$e'_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad e'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{1.36}$$

haben wir die Basiswechselmatrix

$$B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{1.37}$$

Um die neuen Koordinaten eines Punktes $(x, y) \in \mathbb{R}^2$ berechnen zu können, brauchen wir die Inverse von B :

$$B^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Also gilt

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x-y}{2} \\ \frac{x+y}{2} \end{pmatrix}$$

Zum Beispiel, der Punkt $(2, 1)$ hat in der neuen Basis die Koordinaten $(\frac{1}{2}, \frac{3}{2})$. Um kenntlich zu machen, bezüglich welcher Basis die Koordinaten gemeint sind, schreiben wir einen Strich bei den Koordinaten bezüglich (e'_i) :

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \end{pmatrix}'$$

Wollen wir für den Punkt $\begin{pmatrix} 3 \\ 2 \end{pmatrix}'$ die Koordinaten bzgl. der Standardbasis berechnen, so sollen wir mit der Matrix B multiplizieren:

$$\text{wegen } \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix} \quad \text{gilt } \begin{pmatrix} 3 \\ 2 \end{pmatrix}' = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

Änderung der Abbildungsmatrix bei Basiswechsel

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung mit Abbildungsmatrix A . Das bedeutet, dass die Koordinaten des Punktes $y = f(x)$ aus den Koordinaten des Punktes x durch die Formel

$$y = Ax \tag{1.38}$$

berechnet werden können.

In einer neuen Basis (e'_i) haben die Punkte x und $f(x)$ neue Koordinaten (x'_i) und (y'_i) . Nach dem Satz 24 sind die alten und die neuen Koordinaten durch die Basiswechselmatrix verbunden:

$$x = Bx' \quad y = By'$$

Das Einsetzen in 1.38 liefert

$$By' = ABx'$$

und die Multiplikation von links mit B^{-1} ergibt

$$y' = B^{-1}ABx'$$

Hiermit haben wir den folgenden Satz bewiesen.

Satz 25. *Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung, und sei A ihre Abbildungsmatrix bezüglich der Standardbasis. Wird eine neue Basis durch die Basiswechselmatrix B gegeben, so ist die Abbildungsmatrix A' von f bezüglich der neuen Basis gleich $B^{-1}AB$:*

$$A' = B^{-1}AB$$

Beispiel 6.19. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine lineare Transformation mit der Abbildungsmatrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Finden wir die Abbildungsmatrix bzgl. der Basis $e'_1 = (1, -1)^\top$, $e'_2 = (1, 1)^\top$.

Die Basiswechselmatrix und ihre Inverse sind

$$B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad B^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Also gilt

$$A' = B^{-1}AB = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ -1 & 5 \end{pmatrix}$$

Beispiel 6.20. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Streckung entlang $(1, -1)$ mit Faktor 2. Bezüglich der Basis (1.36) hat diese Abbildung Matrix $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. Wenn A die Abbildungsmatrix von f bezüglich der Standardbasis ist, dann gilt

$$B^{-1}AB = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

mit B aus (1.37). Folglich gilt

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

Beispiel 6.21. Sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ die Drehung um den Vektor $(1, 1, 0)$ mit dem Drehwinkel 45° . Finden wir die Abbildungsmatrix von f bezüglich der Standardbasis.

Der entlang $(1, 1, 0)$ gerichtete Einheitsvektor ist $\frac{1}{\sqrt{2}}(1, 1, 0)^\top$. Ist (e'_1, e'_2, e'_3) eine positiv orientierte orthonormale Basis mit $e'_1 = \frac{1}{\sqrt{2}}(1, 1, 0)^\top$, so hat f bezüglich dieser Basis die Abbildungsmatrix

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

Es gilt $A' = B^{-1}AB$, folglich $A = BA'B^{-1}$ und wir brauchen die Matrix B zu kennen. Die Spalten dieser Matrix sind die Komponenten der Vektoren e'_i .

Eine mit $e'_1 = \frac{1}{\sqrt{2}}(1, 1, 0)^\top$ beginnende orthonormale Basis ist

$$e'_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad e'_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad e'_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Die entsprechende Basiswechselmatrix ist

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

Die Basis (e'_i) ist positiv orientiert, weil $\det B = 1$ gilt. Die Inverse:

$$B^{-1} = B^\top = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

Jetzt ist alles bereit, um die Matrix A zu berechnen:

$$\begin{aligned} A = BA'B^{-1} &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} = \\ &= \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{2} & -1 & 1 \\ \sqrt{2} & 1 & -1 \\ 0 & \sqrt{2} & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \frac{1}{\sqrt{2}} & 1 - \frac{1}{\sqrt{2}} & 1 \\ 1 - \frac{1}{\sqrt{2}} & 1 + \frac{1}{\sqrt{2}} & -1 \\ -1 & 1 & \sqrt{2} \end{pmatrix} \end{aligned}$$

6.7 Eigenwerte und Eigenvektoren einer linearen Transformation

Definition und Beispiele

Hier behandeln wir die Frage, wie man eine Basis findet, in welcher eine gegebene lineare Abbildung eine einfache Gestalt hat.

Ist die Abbildung eine Scherung, so soll die Scherungsachse gefunden werden. Ist es eine gleichzeitige Streckung in zweier Richtungen, so sollen diese Richtungen und die Streckungsfaktoren gefunden werden. Das allgemeine Ziel dabei ist es, einen Vektor zu finden, der durch die Abbildung erhalten oder gestreckt aber nicht gedreht wird. Solche Vektoren heißen *Eigenvektoren* der Abbildung, die zugehörigen Streckungsfaktoren heißen *Eigenwerte*.

Wir betrachten eine lineare Transformation $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, die durch eine $n \times n$ -Matrix A bezüglich der Standardbasis dargestellt wird.

Definition 6.22. Sei A eine $n \times n$ -Matrix. Eine Zahl $\lambda \in \mathbb{R}$ heißt Eigenwert von A , wenn es wenigstens einen Vektor $x \in \mathbb{R}^n, x \neq 0$ gibt mit

$$Ax = \lambda x$$

Jeder Vektor $x \neq 0$, der diese Gleichung erfüllt, heißt Eigenvektor von A zum Eigenwert λ .

Beispiel 6.23. • Für die Matrix $\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ ist $(1, 0)^\top$ ein Eigenvektor zum Eigenwert 2, und $(0, 1)^\top$ ein Eigenvektor zum Eigenwert -1 .

• Für die Matrix $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ ist $(1, 0)^\top$ ein Eigenvektor zum Eigenwert 2.

• Für die Matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ ist $(2, -1)^\top$ ein Eigenvektor zum Eigenwert 0.

Bei einer Drehung behält kein Vektor seine Richtung. Trotzdem hat auch die Drehung Eigenvektoren und Eigenwerte, wenn man komplexe Zahlen als Koordinaten zulässt. Hierfür erweitern wir die obige Definition.

Definition 6.24. Eine Zahl $\lambda \in \mathbb{C}$ heißt Eigenwert von A , wenn es wenigstens einen Vektor $x \in \mathbb{C}^n, x \neq 0$ gibt mit

$$Ax = \lambda x$$

Ein Vektor mit dieser Eigenschaft heißt Eigenvektor von A zum Eigenwert λ .

Hier bedeutet $x \in \mathbb{C}^n$, dass x ein n -Tupel von komplexen Zahlen ist. Matrizen und Vektoren mit komplexen Einträgen werden auf dieselbe Weise multipliziert, wie die mit reellen Einträgen.

Beispiel 6.25. Die Drehmatrix $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ hat $(1, i)^\top$ als Eigenvektor:

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} \cos \varphi - i \sin \varphi \\ \sin \varphi + i \cos \varphi \end{pmatrix} = (\cos \varphi - i \sin \varphi) \begin{pmatrix} 1 \\ i \end{pmatrix}$$

Der zugehörige Eigenwert ist $\cos \varphi - i \sin \varphi$.

Der Vektor $(1, i)^\top$ ist ein Eigenvektor zum Eigenwert $\cos \varphi + i \sin \varphi$:

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} \cos \varphi + i \sin \varphi \\ \sin \varphi - i \cos \varphi \end{pmatrix} = (\cos \varphi + i \sin \varphi) \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

Die Eigenwerte als Wurzeln des charakteristischen Polynoms

Wie findet man die Eigenwerte und die zugehörigen Eigenvektoren einer gegebenen Matrix?

Definition 6.26. Das charakteristische Polynom einer $n \times n$ -Matrix A ist

$$\chi_A(\lambda) := \det(A - \lambda E)$$

Das ist ein Polynom in λ vom Grad n .

Beispiel 6.27. Die Matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ hat das charakteristische Polynom

$$\chi_A(\lambda) = \det \begin{pmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3$$

Satz 26. Eine Zahl $\lambda \in \mathbb{C}$ ist genau dann ein Eigenwert von A , wenn λ eine Nullstelle des charakteristischen Polynoms χ_A ist.

Beweis. Es ist $\lambda \in \mathbb{C}$ genau dann ein Eigenwert, wenn die Gleichung $Ax = \lambda x$ eine Nichtnulllösung besitzt. Wegen $x = Ex$ kann die Gleichung umgeschrieben werden als

$$Ax = \lambda x \Leftrightarrow Ax - \lambda x = 0 \Leftrightarrow (A - \lambda E)x = 0 \quad (1.39)$$

$(A - \lambda E)x = 0$ ist ein homogenes LGS aus n Gleichungen in n Variablen. Es besitzt eine Nichtnulllösung genau dann, wenn $\text{Rang}(A - \lambda E) < n$ gilt. Das wiederum ist äquivalent zum $\det(A - \lambda E) = 0$. Der Satz ist bewiesen. \square

Beispiel 6.28. Finden wir die Eigenwerte der Matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Das charakteristische Polynom

$$\chi_A = \lambda^2 - 2\lambda - 3$$

hat die Nullstellen $\lambda_1 = -1$ und $\lambda_2 = 3$. Das sind also die Eigenwerte der Matrix A .

Die Eigenvektoren als Lösungen eines LGS

Wegen (1.39) ist ein Eigenvektor zum Eigenwert λ eine Lösung des linearen Gleichungssystem $(A - \lambda E)x = 0$. Finden wir die Eigenvektoren in unserem Beispiel.

Beispiel 6.29. Für $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ müssen wir zwei LGS lösen.

$$\text{für } \lambda_1 = -1: \quad \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} x = 0$$

$$\text{für } \lambda_1 = 3: \quad \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} x = 0$$

Eine Lösung der ersten Systems ist $v_1 = (1, -1)^\top$, des zweiten $v_2 = (1, 1)^\top$. Es ist also v_1 ein Eigenvektor zum Eigenwert -1 , und v_2 ein Eigenvektor zum Eigenwert 3 .

Erinnern wir uns an die geometrische Bedeutung der Eigenwerte und Eigenvektoren: $Ax = \lambda x$ bedeutet, dass die lineare Abbildung den Vektor x mit dem Faktor λ streckt (ohne zu drehen). Die Vektoren v_1 und v_2 bilden eine Basis von \mathbb{R}^2 , und die Abbildungsmatrix bezüglich dieser Basis sieht besonders einfach aus: es ist eine Diagonalmatrix mit -1 und 3 auf der Diagonale:

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

Als Probe können wir die Regel der Basiswechsel anwenden: es soll gelten $A' = B^{-1}AB$, wobei B die Basiswechselmatrix ist: Spalten von B sind die neuen Basisvektoren v_1 und v_2 . Also

$$B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad B^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

In der Tat gilt

$$B^{-1}AB = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

Zusammenfassung

Sei A eine $n \times n$ -Matrix. Die Berechnung der Eigenwerte und Eigenvektoren von A erfolgt in 2 Schritten:

1. *Schritt.* Man bestimmt die Nullstellen $\lambda_1, \dots, \lambda_m$ des charakteristischen Polynoms $\chi_A(\lambda) = \det(A - \lambda E)$. Hat eine Nullstelle λ_i die Vielfachheit k_i :

$$\chi_A(\lambda) = \dots (\lambda - \lambda_i)^{k_i} \dots,$$

so heißt k_i die *algebraische Vielfachheit* des Eigenwerts λ_i .

2. *Schritt.* Zu jedem Eigenwert λ_i berechnet man die Lösungsmenge des homogenen LGS $(A - \lambda_i E)x = 0$. Diese Lösungsmenge ist ein linearer Teilraum von \mathbb{C}^n

$$V(\lambda_i) := \{x \in \mathbb{C}^n \mid (A - \lambda_i E)x = 0\} = \text{Kern}(A - \lambda E),$$

der *Eigenraum* zum Eigenwert λ_i heißt. Die Dimension des Eigenraums $V(\lambda_i)$ wird die *geometrische Vielfachheit* des Eigenwerts λ_i genannt. Im Allgemeinen stimmt die geometrische Vielfachheit mit der algebraischen nicht überein, die erstere ist aber immer kleiner oder gleich der letzteren.

Zur Berechnung des charakteristischen Polynoms einer 2×2 -Matrix kann die folgende Regel angewendet werden:

$$\chi_A(\lambda) = \lambda^2 - \text{Spur } A \cdot \lambda + \det A$$

wobei $\text{Spur } A := a_{11} + a_{22}$ die Summe der Diagonalelemente bezeichnet. Die Herleitung der Regel:

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} = \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) \end{aligned}$$

Beispiel 6.30. Bestimmen wir die Eigenwerte und Eigenvektoren der Matrix

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

Aufgrund $\text{Spur } A = 2$, $\det A = 5$ lautet das charakteristische Polynom

$$\chi_A(\lambda) = \lambda^2 - 2\lambda + 5$$

Das Lösen der quadratischen Gleichung ergibt $\lambda_1 = 1 + 2i$, $\lambda_2 = 1 - 2i$. Das sind die Eigenwerte.

Um einen Eigenvektor zum Eigenwert $1 + 2i$ zu finden, lösen wir das LGS $(A - \lambda_1 E)x = 0$:

$$\begin{pmatrix} -1 - 2i & 5 \\ -1 & 1 - 2i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Eine der möglichen Lösungen ist $(x, y) = (1 - 2i, 1)$.

Analog, ein Eigenvektor zum Eigenwert $1 - 2i$ ist $(1 + 2i, 1)$.

Wie wir bereits gesehen haben, hat die Drehmatrix komplexe Eigenwerte; die Komponenten ihrer Eigenvektoren sind ebenfalls komplex. Die Interpretation einer Drehung als gleichzeitiger "Streckung" in komplexen Richtungen mit komplexen Faktoren ist geometrisch nicht sehr ansprechend. Komplexe Eigenwerte spielen aber eine große Rolle (Eigenfrequenzen, Resonanzphänomen).

Beispiel 6.31. Bestimmen wir die Eigenwerte und Eigenvektoren der Matrix

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

Das charakteristische Polynom

$$\chi_A(\lambda) = \lambda^2 - 4\lambda + 4$$

hat nur eine (zweifache) Nullstelle $\lambda_1 = 2$. Es ist also ein Eigenwert mit algebraischer Vielfachheit 2.

Um den entsprechenden Eigenraum zu finden, lösen wir das LGS

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

Der Lösungsraum ist eindimensional und wird beispielsweise von $(1, 1)$ erzeugt. Die geometrische Vielfachheit des Eigenwerts 2 ist also gleich 1.

Bei der Berechnung des charakteristischen Polynoms einer 3×3 -Matrix kann man die folgende Formel benutzen:

$$\chi_A(\lambda) = -\lambda^3 + \text{Spur } A \cdot \lambda^2 - c\lambda + \det A,$$

wobei $c = \det A_{11} + \det A_{22} + \det A_{33}$ ist.

Beispiel 6.32. Bestimmen wir die Eigenwerte und Eigenräume der Matrix

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

Wegen $\text{Spur } A = 5$, $\det A_{11} + \det A_{22} + \det A_{33} = 4 + 2 + 2 = 8$, $\det A = 4$ gilt

$$\chi_A(\lambda) = -\lambda^3 + 5\lambda^2 - 8\lambda + 4$$

Im Allgemeinen ist es keine einfache Aufgabe, die Nullstellen eines Polynoms dritten Grades zu bestimmen. Hier hilft uns die folgende Beobachtung: Da die zweite Spalte der Matrix A gleich $(0, 2, 0)^\top$ ist, ist der zweite Standardbasisvektor ein Eigenvektor von A zum Eigenwert 2. Das bedeutet, $\lambda_1 = 2$ ist eine Nullstelle des charakteristischen Polynoms.

Das charakteristische Polynom ist hiermit durch $\lambda - 2$ teilbar. Die Polynomdivision ergibt

$$-\lambda^3 + 5\lambda^2 - 8\lambda + 4 = -(\lambda - 2)(\lambda^2 - 3\lambda + 2)$$

Nach dem Lösen einer quadratischen Gleichung erhalten wir die Faktorisierung des charakteristischen Polynoms:

$$\chi_A(\lambda) = -(\lambda - 2)^2(\lambda - 1)$$

Also, $\lambda_1 = 2$ ist ein Eigenwert von algebraischer Vielfachheit 2, und $\lambda_2 = 1$ ist ein Eigenwert von algebraischer Vielfachheit 1.

Bestimmen wir den Eigenraum zum Eigenwert $\lambda_1 = 2$. Hierfür lösen wir das LGS mit der Matrix $A - 2E$:

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Die Variablen y und z können als freie Variablen gewählt werden. Das führt zu zwei Fundamentallösungen $(0, 1, 0)$ und $(1, 0, 1)$. Der Eigenraum zum Eigenwert $\lambda_1 = 2$ ist hiermit

$$V(2) = \left\{ a \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} a \\ b \\ a \end{pmatrix} \right\}$$

Der Eigenraum zum Eigenwert $\lambda_2 = 1$ ist eindimensional und wird z. B. durch den Vektor $(1, 1, 0)$ erzeugt.

6.8 Diagonalisierbare Matrizen

Wir haben Eigenwerte und Eigenvektoren einer $n \times n$ -Matrix definiert. Diese Begriffe gehören allerdings eher zu der entsprechenden linearen Transformation $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ als zu der Matrix A .

Lemma 30. Sei $A' = B^{-1}AB$, und sei x ein Eigenvektor der Matrix A . Dann ist $x' := B^{-1}x$ ein Eigenvektor der Matrix A' .

Beweis. Sei $Ax = \lambda x$. Dann gilt

$$A'x' = (B^{-1}AB)(B^{-1}x) = B^{-1}Ax = B^{-1}\lambda x = \lambda(B^{-1}x) = \lambda x'$$

□

Man beachte, dass x' aus dem Lemma einfach die neuen Koordinaten des Vektors x sind, wobei die neue Basis aus den Spalten der Matrix B besteht. Dementsprechend stellen die Matrizen A und A' dieselbe lineare Transformation dar, nur bezüglich unterschiedlicher Basen.

Das erlaubt uns, über die *Eigenvektoren einer linearen Transformation* zu reden: die lineare Abbildung f hat einen Eigenvektor v , der Koordinaten (x_i) in der Standardbasis hat (wo die Abbildung durch die Matrix A dargestellt wird), währenddessen in der neuen Basis die Koordinaten desselben Vektors (x'_i) sind und die Abbildungsmatrix A' ist.

Definition 6.33. Eine $n \times n$ -Matrix A heißt diagonalisierbar, wenn es eine invertierbare $n \times n$ -Matrix B gibt, sodass $B^{-1}AB$ eine Diagonalmatrix ist.

(Die Einträge der Matrix B dürfen auch komplexe Zahlen sein.)

Satz 27. Eine $n \times n$ -Matrix A ist diagonalisierbar genau dann, wenn es in \mathbb{R}^n (oder in \mathbb{C}^n) eine Basis aus Eigenvektoren der Matrix A gibt.

Beweis. Die Matrix $A' := B^{-1}AB$ repräsentiert dieselbe lineare Transformation f wie die Matrix A , allerdings bezüglich einer anderen Basis (e'_i) . Die Matrix A' ist eine Diagonalmatrix genau dann, wenn alle e'_i Eigenvektoren der Transformation f sind. Daraus folgt die Behauptung. \square

Beispiel 6.34. Die Matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ ist diagonalisierbar. Im Beispiel 6.29 haben wir ihre Eigenvektoren gefunden: $v_1 = (1, -1)^\top$ und $v_2 = (1, 1)^\top$. Die Diagonalisierung von A erfolgt mit Hilfe der Matrix $B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ (die erste Spalte ist v_1 , die zweite v_2).

Beispiel 6.35. Eine Drehmatrix wird durch eine Matrix mit komplexen Einträgen diagonalisiert, siehe Beispiel 6.25.

Beispiel 6.36. Die Matrix $\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ ist nicht diagonalisierbar. Sie hat (siehe Beispiel 6.31) nur einen Eigenvektor, bis auf Skalierung. Deswegen gibt es keine Basis aus Eigenvektoren dieser Matrix.

Allgemein gilt:

Satz 28. a) Besitzt eine $n \times n$ -Matrix n unterschiedlichen Eigenwerte, so ist sie diagonalisierbar.

b) Eine Matrix ist diagonalisierbar genau dann, wenn die geometrische Vielfachheit jedes ihres Eigenwerts gleich der algebraischen Vielfachheit ist.

Diagonalisierung kann zur schnellen Berechnung der Potenzen einer Matrix verwendet werden, denn es gilt

$$(BA'B^{-1})^k = B(A')^k B^{-1}$$

und wenn $A' = \text{Diag}(\lambda_1, \dots, \lambda_n)$ gilt, dann $(A')^k = \text{Diag}(\lambda_1^k, \dots, \lambda_n^k)$.

7 Symmetrische Matrizen und quadratische Formen

7.1 Diagonalisierbarkeit symmetrischer Matrizen

Aus dem vorherigen Abschnitt wissen wir, dass nicht jede Matrix diagonalisierbar ist. Der folgende Satz zeigt, dass die symmetrischen Matrizen “besonders

gut" diagonalisierbar sind: erstens braucht man keine komplexe Koordinaten (wie bei Drehmatrizen), und zweitens, die Diagonalisierung kann durch eine orthonormale Basiswechsel erfolgen.

Satz 29. Für jede symmetrische $n \times n$ -Matrix A gilt:

- a) Alle Eigenwerte von A sind reell.
- b) Eigenvektoren zu verschiedenen Eigenwerten von A sind orthogonal.
- c) Algebraische und geometrische Vielfachheiten jedes Eigenwertes sind gleich.

Insbesondere existiert eine orthogonale Matrix B , sodass $B^{-1}AB$ eine Diagonalmatrix ist.

Beweis. Wir beweisen nur Punkt b): Seien x und y zwei Eigenvektoren zu den Eigenwerten λ , bzw. $\mu \neq \lambda$. Also, $Ax = \lambda x$ und $Ay = \mu y$. Dann gilt

$$\lambda x \cdot y = Ax \cdot y = (Ax)^\top y = x^\top A^\top y = x^\top Ay = x \cdot Ay = x \cdot \mu y = \mu x \cdot y$$

Wegen $\lambda \neq \mu$ folgt $x \cdot y = 0$.

Zur Diagonalisierbarkeit bemerken wir nur, dass die (nichtdiagonalisierbare) Scherungsmatrix $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$ nicht symmetrisch ist, sowie auch die (im Reellen nicht diagonalisierbare) Drehmatrix. \square

Beispiel 7.1. Die Eigenwerte der Matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ sind -1 und 3 . Die Eigenvektoren $(1, 1)^\top$ und $(1, -1)^\top$ sind zueinander orthogonal.

7.2 Quadratische Formen

Definition

Bisher haben wir $n \times n$ -Matrizen als "Koordinaten" einer linearen Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^n$ betrachtet. Symmetrische $n \times n$ -Matrizen haben eine andere wichtige Interpretation: sie repräsentieren *quadratische Formen*.

Definition 7.2. Eine Funktion $q: \mathbb{R}^n \rightarrow \mathbb{R}$ der Form

$$q(x) = \sum_{i,j=1}^n a_{ij}x_i x_j \tag{1.40}$$

heißt quadratische Form in n Variablen x_1, \dots, x_n .

Beispiel 7.3. Die Funktionen

$$q_1(x, y) = x^2 + y^2, \quad q_2(x, y) = x^2 - 2xy - 8y^2, \quad q_3(x, y) = xy \tag{1.41}$$

sind quadratische Formen in Variablen x und y .

Lemma 31. Für jede quadratische Form q in n Variablen gibt es genau eine symmetrische $n \times n$ -Matrix A , sodass

$$q(x) = x^\top Ax$$

Beweis. Die Koeffizienten (a_{ij}) in (1.40) bilden eine $n \times n$ -Matrix A , und es gilt

$$q(x) = \sum_{i,j=1}^n a_{ij}x_i x_j = (x_1 \ \dots \ x_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x^\top A x$$

Die Matrix A ist nicht unbedingt symmetrisch, sie kann aber symmetrisiert werden: beim Ersetzen von a_{ij} und a_{ji} durch $\frac{a_{ij}+a_{ji}}{2}$ ändert sich die Funktion $x^\top A x$ nicht. \square

Beispiel 7.4. Die den quadratischen Formen (1.41) entsprechende Matrizen sind

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -1 \\ -1 & -8 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$$

Definition 7.5. Die quadratische Form (1.40) heißt rein quadratisch, falls $a_{ij} = 0$ bei $i \neq j$ gilt.

Anders gesagt, die Matrix einer rein quadratischen Form ist eine Diagonalmatrix. Für $A = \text{Diag}(\lambda_1, \dots, \lambda_n)$ lautet die entsprechende quadratische Form

$$q(x) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$$

Durch ein Variablenwechsel kann eine quadratische Form rein quadratisch gemacht werden. Für die Formen (1.41):

$$\begin{aligned} x^2 + y^2 &\text{ ist eine rein quadratische Form} \\ x^2 - 2xy - 8y^2 &= (x^2 - 2xy + y^2) - 9y^2 = (x - y)^2 - (3y)^2 \quad (1.42) \\ xy &= \frac{1}{4}(x^2 + 2xy + y^2) - \frac{1}{4}(x^2 - 2xy + y^2) = \left(\frac{x+y}{2}\right)^2 - \left(\frac{x-y}{2}\right)^2 \end{aligned}$$

Änderung der Koeffizienten der quadratischen Form bei Basiswechsel

Der Variablenwechsel in (1.42) kann wie folgt beschrieben werden:

$$x^2 - 2xy - 8y^2 = (x')^2 - (y')^2 \quad \text{wobei } x' = x - y, \ y' = 3y \quad (1.43)$$

Wir wollen nun eine allgemeine Regel herleiten, die die Koeffizientenänderung einer quadratischen Form bei einer beliebigen linearen Substitution beschreibt.

Eine lineare Substitution sieht aus wie $x = Bx'$ mit einer $n \times n$ -Matrix B . Man beachte, dass es dieselbe Formel ist, die die alten Koordinaten durch die neuen bei einer Basiswechsel (mit der Basiswechsellmatrix B) ausdrückt.

Satz 30. Sei $q(x) = x^\top A x$ eine quadratische Form in n Variablen. Sei $x = Bx'$ eine lineare Substitution. Dann gilt $q(x) = q'(x')$, wobei q' eine quadratische Form ist, mit der Matrix

$$A' = B^\top A B$$

Beweis.

$$q(x) = x^\top A x = (Bx')^\top A (Bx') = (x')^\top B^\top A B x' = (x')^\top (B^\top A B) x'$$

\square

Beispiel 7.6. Für die Substitution (1.43) gilt

$$\begin{aligned}x &= x' + y = x' + \frac{1}{3}y' \\ y &= \frac{1}{3}y'\end{aligned}$$

Deswegen ist die Basiswechselmatrix $B = \begin{pmatrix} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$. Es gilt in der Tat

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & -8 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

7.3 Die Hauptachsentransformation

Es gibt viele unterschiedlichen Substitutionen, die eine quadratische Form rein quadratisch machen. Z. B. alternativ zu $xy = \left(\frac{x+y}{2}\right)^2 - \left(\frac{x-y}{2}\right)^2$:

$$xy = \left(\frac{x+2y}{2\sqrt{2}}\right)^2 - \left(\frac{x-2y}{2\sqrt{2}}\right)^2$$

Der Satz 29 zeigt die Existenz einer speziellen Substitution, bei welcher die Basiswechselmatrix B orthogonal ist.

Satz 31 (Hauptachsentransformation). *Für jede quadratische Form $q(x) = x^\top Ax$ gibt es eine orthogonale Matrix B , sodass die Substitution $x = Bx'$ die Form rein quadratisch macht.*

Beweis. Nach dem Satz 29 gibt es eine orthogonale Matrix B , sodass $B^{-1}AB$ eine Diagonalmatrix ist. Da aber für eine orthogonale Matrix $B^{-1} = B^\top$ gilt, ist

$$B^{-1}AB = B^\top AB$$

Also, bei der Substitution $x = Bx'$ erhält man eine quadratische Form in x' mit einer Diagonalmatrix, d. h. eine rein quadratische Form. \square

Beispiel 7.7. Die Substitution (1.43) ist nicht orthogonal. Die Substitution $x = \frac{x'+y'}{\sqrt{2}}, y = \frac{x'-y'}{\sqrt{2}}$ ist orthogonal. Sie gibt die Hauptachsentransformation der quadratischen Form $q(x, y) = xy$ an:

$$xy = \frac{x'+y'}{\sqrt{2}} \cdot \frac{x'-y'}{\sqrt{2}} = \frac{1}{2}(x')^2 - \frac{1}{2}(y')^2$$

Definition 7.8. *Ein Hauptachsensystem der quadratischen Form $q(x)$ ist eine Orthonormalbasis (e'_1, \dots, e'_n) , sodass die entsprechende quadratische Form $q'(x')$ rein quadratisch ist.*

Die Berechnung eines Hauptachsensystems der Form $q(x) = x^\top Ax$ besteht im Finden einer orthonormalen Basis aus Eigenvektoren der Matrix A . Sind alle Eigenwerte der Matrix unterschiedlich, so sind die Eigenvektoren orthogonal zueinander, und sollen nur normiert werden. Gibt es ein vielfaches Eigenwert, so soll im entsprechenden Eigenraum eine orthonormale Basis gewählt werden.

Beispiel 7.9. Finden wir das Hauptachsensystem der quadratischen Form

$$q(x, y, z) = -x^2 - y^2 + z^2 + 6xy + 2xz + 2yz$$

Die Matrix der quadratischen Form q :

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

Das charakteristische Polynom

$$\chi_A(\lambda) = -\lambda^3 - \lambda^2 + 12\lambda$$

hat die Nullstellen

$$\lambda_1 = 3, \lambda_2 = -4, \lambda_3 = 0$$

Berechnung der entsprechenden Eigenräume ergibt eine orthonormale Basis

$$e'_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad e'_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad e'_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

Beispiel 7.10. Finden wir die Richtungen der Hauptachsen der quadratischen Form $q(x, y) = x^2 - 2xy - 8y^2$.

Die Matrix dieser quadratischen Form ist $A = \begin{pmatrix} 1 & -1 \\ -1 & -8 \end{pmatrix}$. Das charakteristische Polynom:

$$\chi_A(\lambda) = \lambda^2 + 7\lambda - 9$$

Die Eigenwerte:

$$\lambda_1 = -\frac{7}{2} + \frac{\sqrt{85}}{2}, \quad \lambda_2 = -\frac{7}{2} - \frac{\sqrt{85}}{2}$$

Die Eigenräume:

$$V(\lambda_1) = \text{Kern}(A - \lambda_1 E) = \text{Kern} \begin{pmatrix} \frac{9}{2} - \frac{\sqrt{85}}{2} & -1 \\ -1 & -\frac{9}{2} - \frac{\sqrt{85}}{2} \end{pmatrix} = \mathbb{R} \begin{pmatrix} \frac{9}{2} + \frac{\sqrt{85}}{2} \\ -1 \end{pmatrix}$$

$$V(\lambda_2) = \text{Kern}(A - \lambda_2 E) = \text{Kern} \begin{pmatrix} \frac{9}{2} + \frac{\sqrt{85}}{2} & -1 \\ -1 & -\frac{9}{2} + \frac{\sqrt{85}}{2} \end{pmatrix} = \mathbb{R} \begin{pmatrix} -\frac{9}{2} + \frac{\sqrt{85}}{2} \\ 1 \end{pmatrix}$$

Die Hauptachsen sind parallel zu den Vektoren $v_1 = \begin{pmatrix} \frac{9}{2} + \frac{\sqrt{85}}{2} \\ -1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} -\frac{9}{2} + \frac{\sqrt{85}}{2} \\ 1 \end{pmatrix}$.

7.4 Quadriken

Definition 7.11. Eine Quadrik ist die Menge aller Punkte in \mathbb{R}^n , die eine Gleichung erfüllen der Form

$$p(x) = x^\top Ax + bx + c = 0$$

Hier ist A eine symmetrische $n \times n$ -Matrix, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$.

Beispiel 7.12. $p(x, y) = x^2 + 2xy + 2y^2 + x + 2y + 4 = 0$ beschreibt eine Quadrik in \mathbb{R}^2

Durch die Hauptachsentransformation und eine Parallelverschiebung kann die Gleichung der Quadrik auf Normalform gebracht werden. Das heißt, in einem entsprechenden Koordinatensystem sieht die Quadrik als eine der folgenden Kurven aus ($n = 2$):

- Im Fall $\text{Rang } A = 2$ (alle Eigenwerte $\neq 0$)

$$\begin{array}{l|l} \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 & \text{Ellipse} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0 & \text{leere Menge} \\ \frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0 & \text{Hyperbel} \\ x^2 + a^2y^2 = 0 & \text{Punkt} \\ x^2 - a^2y^2 = 0 & \text{Geradenpaar mit Schnittpunkt } (x - ay = 0 \text{ und } x + ay = 0) \end{array}$$

- Im Fall $\text{Rang } A = 1$ (Ein Eigenwert = 0)

$$\begin{array}{l|l} x^2 - 2py = 0 & \text{Parabel} \\ x^2 - a^2 = 0 & \text{paralleles Geradenpaar} \\ x^2 + a^2 = 0 & \text{leere Menge} \\ x^2 = 0 & \text{Gerade } x = 0 \end{array}$$

Beispiel 7.13. Wie transformiert man die Quadrik $xy - 1 = 0$ auf die Normalform? Ein Hauptachsensystem der quadratischen Form $q(x, y) = xy$ wird z. B. durch die Vektoren $e'_1 = \frac{1}{\sqrt{2}}(1, 1)^\top$ und $e'_2 = \frac{1}{\sqrt{2}}(1, -1)^\top$ gegeben. Die quadratische Form wird in den neuen Koordinaten zu $q'(x', y') = \frac{1}{2}(x')^2 - \frac{1}{2}(y')^2$. Dementsprechend ist die Normalform der Quadrik

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

Das ist eine Hyperbel.

Beispiel 7.14. Zeigen wir, dass die Gleichung $2x^2 + 2xy + 3y^2 = 1$ eine Ellipse beschreibt, und finden wir seine Halbachsen.

Die Eigenwerte der Matrix $\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ sind $\lambda_{1,2} = \frac{5 \pm \sqrt{5}}{2}$. Nach der Hauptachsentransformation hat die Kurve die Gleichung

$$\lambda_1(x')^2 + \lambda_2(y')^2 - 1 = 0$$

Da λ_1 und λ_2 beide positiv sind, handelt es sich um eine Ellipse. Seine Halbachsen sind gleich

$$\frac{1}{\sqrt{\lambda_1}} = \sqrt{\frac{2}{5 + \sqrt{5}}}, \quad \frac{1}{\sqrt{\lambda_2}} = \sqrt{\frac{2}{5 - \sqrt{5}}}$$

Quadriken in \mathbb{R}^3

Wir geben keine vollständige Liste aller Quadriken in \mathbb{R}^3 , sondern erwähnen nur die wichtigsten unter ihnen.

$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$	Ellipsoid mit Halbachsen a , b und c
$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$	einschaliges Hyperboloid
$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0$	zweischaliges Hyperboloid
$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$	Kegel
$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2pz = 0$	elliptisches Paraboloid
$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 2pz = 0$	hyperbolisches Paraboloid

Das einschalige Hyperboloid und das hyperbolische Paraboloid sind *Regelflächen*, d. h. durch jeden Punkt geht eine Gerade, die vollständig in der Fläche enthalten ist. In der Tat, bei diesen zwei Flächen gehen durch jeden Punkt sogar zwei auf der Fläche liegenden Geraden.

7.5 Positiv definite Matrizen

Definition 7.15. Eine quadratische Form $q(x) = x^\top Ax$, bzw. die zugehörige symmetrische Matrix heißt positiv definit (negativ definit), wenn $q(x) > 0$ ($q(x) < 0$) für alle $x \neq 0$ gilt.

Die quadratische Form heißt indefinit, wenn sie sowohl positive als auch negative Werte annimmt. Sie heißt positiv (negativ) semidefinit, wenn stets $q(x) \geq 0$ ($q(x) \leq 0$) gilt.

Beispiel 7.16. • $q(x, y) = xy$ ist indefinit, denn z. B. $q(1, 1) = 1 > 0$ und $q(1, -1) = -1 < 0$.

- $q(x, y) = x^2 + 2y^2$ ist positiv definit.
- $q(x, y) = -x^2 - 2y^2$ ist negativ definit.

Satz 32. (a) Eine Diagonalmatrix $\text{Diag}(\alpha_1, \dots, \alpha_n)$ ist genau dann positiv definit, wenn alle α_i positiv sind.

(b) Eine symmetrische Matrix A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.

Positivitätstest

Satz 33. Eine symmetrische $n \times n$ -Matrix A ist genau dann positiv definit, wenn die Determinanten der n "Hauptmatrizen" H_i positiv sind:

$$H_1 = a_{11}, \quad H_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \dots, \quad H_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}, \dots, \quad H_n = A$$

Beispiel 7.17. Die Matrix $A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 6 & -1 \\ 2 & -1 & 4 \end{pmatrix}$ ist positiv definit, denn $a_{11} = 5 > 0$, $\det \begin{pmatrix} 5 & -2 \\ -2 & 6 \end{pmatrix} = 34 > 0$, $\det A = 83 > 0$.

Korollar 11. Eine symmetrische $n \times n$ -Matrix A ist genau dann negativ definit, wenn die Determinante der Hauptmatrix H_k das Vorzeichen $(-1)^k$ hat:

$$\det H_1 < 0, \quad \det H_2 > 0, \quad \det H_3 < 0, \dots$$

Kapitel 2

Funktionen in mehreren Variablen: Differentiation

1 Grundlagen

1.1 Reellwertige und vektorwertige Funktionen mehrerer Veränderlicher

In Wissenschaft und Technik studiert man oft Größen, die von mehreren Variablen abhängen (von den Ortskoordinaten x, y, z , von der Zeit t , usw.). Betrachtet man z. B. die Strömung einer Flüssigkeit, so wird der Druck im Punkt mit Koordinaten (x, y, z) zum Zeitpunkt t als eine Funktion $p(x, y, z, t)$ beschrieben. Das ist eine *reellwertige Funktion* von 4 Variablen.

Eine *vektorwertige Funktion* nimmt ihre Werte in \mathbb{R}^m . Z. B. die Geschwindigkeit der Strömung im Punkt (x, y, z) zum Zeitpunkt t wird durch einen Vektor $v(x, y, z, t) \in \mathbb{R}^3$ beschrieben.

Im Allgemeinen studieren wir *Abbildungen*

$$f: D \rightarrow \mathbb{R}^m$$

wobei $D \subset \mathbb{R}^n$ eine Teilmenge von \mathbb{R}^n ist. Ist $m = 1$, so schreiben wir

$$f(x) = f(x_1, x_2, \dots, x_n)$$

(reellwertige Funktion, oder Skalarenfeld). Ist $m > 1$, so gilt

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

(vektorwertige Funktion, Vektorfeld).

Beispiel 1.1. Eine lineare Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ wird gegeben durch

$$f(x) = Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

mit einer $m \times n$ -Matrix A .

Wir beschränken uns zunächst auf reellwertige Funktionen $f: D \rightarrow \mathbb{R}$.

1.2 Der Graph, Niveaumengen und “partielle” Funktionen

Der Graph

Sei $D \subset \mathbb{R}^n$ und sei $f: D \rightarrow \mathbb{R}$ eine reellwertige Funktion von n Variablen. Wie im Falle $n = 1$, kann jeder Funktion ihr *Graph* zugeordnet werden:

$$\Gamma_f := \{(x, f(x)) \mid x \in D\} \subset \mathbb{R}^n$$

Mit anderen Worten, Γ_f besteht aus allen Punkten $(x_1, \dots, x_n, x_{n+1})$ mit $x_{n+1} = f(x_1, \dots, x_n)$ (vgl. $y = f(x)$ für $n = 1$).

Bei Funktionen von zwei Variablen werden die Variablen üblicherweise mit x und y bezeichnet; den Graphen ist die Fläche $z = f(x, y)$ in \mathbb{R}^3 .

Beispiel 1.2. Der Graph der Funktion $f(x, y) = x^2 + y^2$ ist die Fläche

$$\{(x, y, z) \mid z = x^2 + y^2\}$$

d. h. das Rotationsparaboloid (entsteht durch die Drehung der Parabel $z = x^2$ um die z -Achse).

Niveaumengen

Der Graph einer Funktion $f(x, y, z)$ ist ein 3-dimensionales Objekt in \mathbb{R}^4 , das nicht einfach zu visualisieren ist. Eine Methode, sich das Verhalten einer solchen Funktion vorzustellen, benutzt die Niveauflächen.

Definition 1.3. Sei $D \subset \mathbb{R}^n$ und sei $f: D \rightarrow \mathbb{R}$ eine Funktion von n Variablen. Die Niveaumenge von f zum Niveau $c \in \mathbb{R}$ wird definiert als

$$N_c := \{x \in D \mid f(x) = c\}$$

Bei $n = 2$ heißen die Niveaumengen *Niveaukurven*, bei $n = 3$ *Niveauflächen*.

Beispiel 1.4. • Die Niveauflächen einer linearen Funktion $f(x, y, z) = ax + by + cz$ sind Ebenen in \mathbb{R}^3 .

- Die Niveaulinie der Funktion $f(x, y) = x^2 - y^2$ zum Niveau c ist

$$\begin{array}{ll} \text{bei } c \neq 0 & \text{eine Hyperbel } x^2 - y^2 = c \\ \text{bei } c = 0 & \text{ein Geradenpaar } x = y \text{ und } x = -y \end{array}$$

Der Graph der Funktion $f(x, y) = x^2 - y^2$ ist ein hyperbolisches Paraboloid $x^2 - y^2 - z = 0$.

Die Niveaumengen sehen nicht immer als glatte Kurven oder Flächen aus. So haben wir im letzten Beispiel ein Geradenpaar als Niveaumenge. Die Niveaumenge der Funktion $f(x, y) = x^2 + y^2$ zum Niveau 0 besteht aus einem Punkt $(0, 0)$. Die Niveaumenge einer konstanten Funktion $f(x, y) = c$ zum Niveau c ist die ganze Ebene.

“Partielle” Funktionen

Definition 1.5. Sei $f: D \rightarrow \mathbb{R}$ eine Funktion von n Variablen, und sei $E \subset D$. Die Abbildung $E \rightarrow \mathbb{R}$, die dieselben Werte annimmt, wie f , heißt Einschränkung von f auf E und wird bezeichnet durch

$$f|_E: E \rightarrow \mathbb{R}$$

Wir bezeichnen die Einschränkung der Funktion f auf eine zu einer Koordinatenachse parallele Gerade als “partielle” Funktion. Eine zur i -ten Achse parallele Gerade besteht aus den Punkten $\{(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n) \mid x_i \in \mathbb{R}\}$ bei konstanten a_1, \dots, a_n . Deswegen kann die “partielle” Funktion als Funktion einer Variable betrachtet werden:

$$x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$$

Beispiel 1.6. Die “partiellen” Funktionen der Funktion $f(x, y) = x^2 - y^2$ sind

$$f(x, a) = x^2 - a^2, \quad f(a, y) = a^2 - y^2$$

Das erlaubt uns, das hyperbolische Paraboloid $z = x^2 - y^2$ als die Familie von der Parabel $z^2 = x^2, y = 0$ “herabhängenden” Parabeln $z = a^2 - x^2, x = a$ zu sehen.

Beispiel 1.7. Die “partiellen” Funktionen der Funktion $f(x, y) = \frac{xy}{x+y}$ sind gebrochen lineare Funktionen:

$$x \mapsto \frac{ax}{x+a}, \quad y \mapsto \frac{by}{y+b}$$

1.3 Teilmengen von \mathbb{R}^n

Definition 1.8. Sei $a \in \mathbb{R}^n$ ein Punkt im n -dimensionalen Euklidischen Raum, und sei $r > 0$. Dann heißt die Menge

$$U_r(a) := \{x \in \mathbb{R}^n \mid \|x - a\| < r\}$$

r -Umgebung von a .

Die r -Umgebung ist also eine Kugel vom Radius r , ohne ihre Randfläche.

Definition 1.9. Sei D eine Teilmenge des \mathbb{R}^n .

- (a) Ein Punkt $a \in D$ heißt innerer Punkt von D , wenn es eine r -Umgebung von a gibt, die ganz in D enthalten ist.
- (b) D heißt offen, wenn jeder Punkt von D ein innerer Punkt ist.

Definition 1.10. (a) Ein Punkt $b \in \mathbb{R}$ heißt Randpunkt von D , wenn jede r -Umgebung von b sowohl mindestens einen Punkt aus D als auch mindestens einen nicht zu D gehörenden Punkt enthält. Die Menge aller Randpunkte von D heißt Rand von D und wird mit ∂D bezeichnet.

- (b) Eine Menge heißt abgeschlossen, wenn sie alle ihre Randpunkte enthält.

Beispiel 1.11. Sei $D = \{(x, y) \mid y > 0\}$ die obere Halbebene in \mathbb{R}^2 , ohne die x -Achse.

Dann ist D offen, denn für jeden Punkt $a = (x_0, y_0) \in D$ gilt $y_0 > 0$, und die $\frac{y_0}{2}$ -Umgebung von a ist in D enthalten.

Der Punkt $b = (1, 0)$ ist ein Randpunkt von D , denn in jeder Umgebung von b gibt es einen Punkt aus D , und auch einen Punkt nicht aus D (z. B. b selbst).

Beispiel 1.12. Die Menge

$$Q = \{(x, y) \mid 2 \leq x < 3, -1 < y \leq 2\}$$

ist weder offen, noch abgeschlossen: $(2, 0)$ liegt in Q , ist aber kein innerer Punkt; und $(3, 0)$ ist ein Randpunkt von Q , liegt in Q aber nicht.

Definition 1.13. Der Abschluss einer Menge $D \subset \mathbb{R}^n$ besteht aus der Menge D zusammen mit allen ihren Randpunkten:

$$\bar{D} := D \cup \partial D$$

Man kann zeigen, dass \bar{D} immer eine abgeschlossene Menge ist.

Beispiel 1.14. Der Abschluss der Menge Q aus dem letzten Beispiel ist

$$\bar{Q} = \{(x, y) \mid 2 \leq x \leq 3, -1 \leq y \leq 2\}$$

Definition 1.15. (a) Eine Menge $D \subset \mathbb{R}^n$ heißt beschränkt, wenn es eine Konstante $C > 0$ gibt, sodass

$$\|x\| < C \text{ für alle } x \in D$$

Mit anderen Worten, D soll in einer Kugel mit Mittelpunkt 0 enthalten sein: $D \subset U_C(0)$

(b) Die abgeschlossenen und beschränkten Mengen nennt man kompakt.

Beispiel 1.16. • Die Menge $U_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| < r\}$ ist beschränkt, denn es gilt

$$U_r(a) \subset U_{r+\|a\|}(0)$$

• Die abgeschlossene Kugel $B_r(a) := \{x \in \mathbb{R}^n \mid \|x - a\| \leq r\}$ ist ebenfalls beschränkt, denn z. B.

$$B_r(a) \subset U_{r+\|a\|+1}(0)$$

• $U_r(a)$ ist nicht kompakt, weil sie keine abgeschlossene Menge ist; die Menge $B_r(a)$ ist kompakt.

1.4 Grenzwerte und Stetigkeit

Definition 1.17. Sei $D \subset \mathbb{R}^n$, $f: D \rightarrow \mathbb{R}$, und $a \in \bar{D}$.

(a) f hat in a den Grenzwert c :

$$\lim_{x \rightarrow a} f(x) = c \quad (\text{oder } f(x) \rightarrow c \text{ für } x \rightarrow a)$$

wenn es zu jeder (beliebig kleinen) Schranke $\varepsilon > 0$ eine r -Umgebung $U_r(a)$ gibt, sodass $|f(x) - c| < \varepsilon$ für alle $x \in D \cap U_r(a)$ gilt.

(b) f heißt in a stetig, wenn $\lim_{x \rightarrow a} f(x) = f(a)$ gilt.

(c) f heißt auf D stetig, wenn f in allen $a \in D$ stetig ist.

Die Definition der Stetigkeit in a kann wie folgt umformuliert werden: kennt man den Punkt a bis auf r genau, so kennt man den Wert $f(a)$ bis auf ε genau, und für jedes noch so kleines ε kann man ein passendes r finden.

Ähnlich wie bei stetigen Funktionen einer Variablen, gilt

Summe, Produkt, Quotient stetiger Funktionen ist stetig.

Verkettung stetiger Funktionen ist stetig.

Beispiel 1.18. Die Funktionen $f_1(x, y) = x$, $f_2(x, y) = y$ sind stetig (für jedes ε kann $r = \varepsilon$ gewählt werden). Daher sind die Funktionen

$$(x, y) \mapsto xy, \quad (x, y) \mapsto x + y$$

stetig; dann ist auch $(x, y) \mapsto \frac{xy}{x+y}$ stetig, und auch $(x, y) \mapsto \sin \frac{xy}{x+y}$.

Satz 34. Seien g_1, \dots, g_r stetige Funktionen auf \mathbb{R}^n . Dann gilt

- Die Menge $\{x \in \mathbb{R}^n \mid g_1(x) < 0, \dots, g_r(x) < 0\}$ ist offen.
- Die Menge $\{x \in \mathbb{R}^n \mid g_1(x) \leq 0, \dots, g_r(x) \leq 0\}$ ist abgeschlossen.

Satz 35. Jede auf einer kompakten Menge $D \subset \mathbb{R}^n$ stetige Funktion nimmt auf D ein Minimum und ein Maximum an; d. h. es gibt $a, b \in D$ mit $f(b) \leq f(x) \leq f(a)$ für alle $x \in D$.

Dieser Satz verallgemeinert den Satz vom Minimum und Maximum für Funktionen einer Variablen. Bei einer Variablen wurde eine Funktion auf einem abgeschlossenen Intervall betrachtet, und ein abgeschlossenes Intervall ist eine kompakte Menge in \mathbb{R} .

2 Differentiation

2.1 Partielle Ableitungen und der Gradient

Partielle Ableitungen

Definition 2.1. Sei $D \subset \mathbb{R}^n$ offen, sei $f: D \rightarrow \mathbb{R}$ eine Funktion von n Variablen, und sei $a = (a_1, a_2, \dots, a_n) \in D$. Existiert die Ableitung der folgenden Funktion einer Variablen x_i :

$$x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$$

an der Stelle $x_i = a_i$, so wird sie die partielle Ableitung von f nach x_i im Punkt a genannt. Bezeichnung:

$$\left. \frac{\partial f}{\partial x_i} \right|_{x=a} \quad \text{oder} \quad \frac{\partial f}{\partial x_i}(a)$$

Um die partielle Ableitung nach x_i zu berechnen, sollen die restlichen Variablen als Konstanten angesehen werden. Dann wird die Differentiation nach den üblichen Regeln (Summe-, Produkt-, Quotient- und Kettenregel) durchgeführt.

Bei konkret angegebenen Funktionen lauten die Variablen normalerweise x, y, z, t, \dots . Zur Abkürzung benutzt man oft die Bezeichnungen

$$f_x := \frac{\partial f}{\partial x} \quad \text{usw.}$$

Man definiert und bezeichnet die höheren partiellen Ableitungen wie folgt:

$$f_{xx} := \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) =: \frac{\partial^2 f}{\partial x^2}, \quad f_{xy} := \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) =: \frac{\partial^2 f}{\partial y \partial x}, \quad \text{usw.}$$

Beispiel 2.2. Sei $f(x, y) = x^2 y^3 + y \ln x$. Dann gilt

$$\begin{aligned} f_x(x, y) &= 2xy^3 + \frac{y}{x}, & f_y(x, y) &= 3x^2 y^2 + \ln x \\ f_{xx}(x, y) &= \frac{\partial}{\partial x} \left(2xy^3 + \frac{y}{x} \right) = 2y^3 - \frac{y}{x^2}, & f_{yy}(x, y) &= \frac{\partial}{\partial y} (3x^2 y^2 + \ln x) = 6x^2 y \\ f_{xy}(x, y) &= \frac{\partial}{\partial y} \left(2xy^3 + \frac{y}{x} \right) = 6xy^2 + \frac{1}{x}, & f_{yx}(x, y) &= \frac{\partial}{\partial x} (3x^2 y^2 + \ln x) = 6xy^2 + \frac{1}{x} \end{aligned}$$

Existieren die partiellen Ableitungen einer Funktion f nach allen Variablen, so heißt f *partiell differenzierbar*. Sind zusätzlich alle partiellen Ableitungen stetig, so heißt f *stetig partiell differenzierbar*.

Der Gradient

Definition 2.3. Sei $D \subset \mathbb{R}^n$ offen, und sei $f: D \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Der Gradient von f ist eine vektorwertige Funktion $\text{grad } f: D \rightarrow \mathbb{R}^n$, wobei die Komponenten von $\text{grad } f(x)$ die partiellen Ableitungen von f sind:

$$\text{grad } f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$$

Eine alternative Bezeichnung für den Gradienten: $\nabla f(x) = \text{grad } f(x)$ (das Symbol ∇ wird als "Nabla" ausgesprochen).

Beispiel 2.4. Der Gradient der Funktion $f(x, y) = x^2 y^3 + y \ln x$ ist

$$\text{grad } f(x, y) = \begin{pmatrix} 2xy^3 + \frac{y}{x} \\ 3x^2 y^2 + \ln x \end{pmatrix}$$

Insbesondere ist der Gradient im Punkt $(1, 1)$ gleich

$$\text{grad } f(1, 1) = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

C^k -Funktionen und Vertauschbarkeit der partiellen Ableitungen

Definition 2.5. Eine Funktion $f: D \rightarrow \mathbb{R}$ heißt k -mal stetig partiell differenzierbar, wenn alle ihre partiellen Ableitungen k -ter Ordnung existieren und stetig sind. Bezeichnung:

$$C^k(D) := \{f: D \rightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal stetig partiell differenzierbar}\}$$
$$C^0(D) := \{f: D \rightarrow \mathbb{R} \mid f \text{ stetig}\}$$

Im Beispiel 2.2 haben wir die Ableitungen $f_{xy} = (f_x)_y$ und $f_{yx} = (f_y)_x$ berechnet und das gleiche Ergebnis erhalten. Das gilt auch allgemein, vorausgesetzt dass f zweimal *stetig* partiell differenzierbar ist.

Satz 36 (Schwarz). Für jede C^2 -Funktion f von mehreren Variablen gilt

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$$

Es gibt aber Beispiele der zweimal partiell differenzierbaren Funktionen, für welche $f_{xy} \neq f_{yx}$ gilt.

Eine mehrfache Anwendung des Satzes von Schwarz auf eine Funktion $f \in C^k(D)$ zeigt, dass es bei einer partiellen Ableitung k -ter Ordnung keine Rolle spielt, in welcher Reihenfolge die partielle Differentiation ausgeführt wird, z. B.

$$\frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x^2}$$

2.2 Differenzierbarkeit und lineare Approximation

Für die Ableitung einer Funktion f einer Variable haben wir die folgende Deutung angegeben:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \text{ ist nahe } x_0 \text{ die beste lineare Approximation an } f \quad (2.1)$$

Geometrisch wird es dadurch ausgedrückt, dass $y = f(x_0) + f'(x_0)(x - x_0)$ die Gleichung der Tangentialgerade zum Graphen von f im Punkt $(x_0, f(x_0))$ ist.

Jetzt wollen wir die mehrdimensionalen Analoga dieser Aussagen formulieren.

Definition 2.6. Für zwei Funktionen $f, g: D \rightarrow \mathbb{R}$ und $k \in \mathbb{N}$ schreibt man

$$f(x) = g(x) + o(|x - x_0|^k) \text{ für } x \rightarrow x_0$$

falls $\lim_{x \rightarrow x_0} \frac{f(x) - g(x)}{|x - x_0|^k} = 0$ gilt.

Wenn $x \in \mathbb{R}$, also wenn es sich um Funktionen einer Variablen handelt, dann kann man $o(x - x_0)$ statt $o(|x - x_0|)$ schreiben.

Beispiel 2.7. • Es gilt

$$\frac{1}{1-x} = 1 + x + o(x) \text{ für } x \rightarrow 0$$
$$\frac{1}{1-x} = 1 + x + x^2 + o(x^2) \text{ für } x \rightarrow 0$$

- Sei $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Für $f(x) := x_1^2 + \dots + x_n^2$ gilt

$$f(x) = o(|x|) \text{ für } x \rightarrow 0$$

denn $|x| = \sqrt{x_1^2 + \dots + x_n^2}$ und folglich $\frac{f(x)}{|x|} = \sqrt{x_1^2 + \dots + x_n^2} \rightarrow 0$ bei $|x| \rightarrow 0$.

Die Aussage (2.1) wird genauer formuliert als

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

Definition 2.8. Sei $D \subset \mathbb{R}^n$ offen. Eine Funktion $f: D \rightarrow \mathbb{R}$ heißt in $x_0 \in D$ total differenzierbar (oder linear approximierbar), wenn es einen Vektor $a \in \mathbb{R}^n$ gibt mit

$$f(x) = f(x_0) + a \cdot (x - x_0) + o(|x - x_0|) \text{ für } x \rightarrow x_0 \quad (2.2)$$

Satz 37. (a) Ist f in $x_0 \in D$ total differenzierbar (2.2), so ist f auch partiell differenzierbar, und es gilt $\text{grad } f(x_0) = a$.

(b) Ist f in x_0 stetig partiell differenzierbar, so ist f in x_0 auch total differenzierbar.

Beweis. (a): Setzen wir in (2.2) $x = x_0 + te_i$ ein, so erhalten wir

$$f(x_0 + te_i) = f(x_0) + a \cdot te_i + o(|te_i|) = f(x_0) + ta_i + o(t)$$

Folglich gilt

$$\frac{\partial f}{\partial x_i}(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + te_i) - f(x_0)}{t} = a_i$$

also die i -te partielle Ableitung existiert und ist gleich der i -ten Komponente des Vektors a .

(b): hier skizzieren wir nur den Beweis. Für zwei Variablen haben wir

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + (f(x, y_0) - f(x_0, y_0)) + (f(x, y) - f(x, y_0)) = \\ &= f(x_0, y_0) + \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} (x - x_0) + \frac{f(x, y) - f(x, y_0)}{y - y_0} (y - y_0) = \\ &= f(x_0, y_0) + \begin{pmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \end{pmatrix} \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + o(\sqrt{(x - x_0)^2 + (y - y_0)^2}) \end{aligned}$$

□

Folglich ist

$$f(x) \approx f(x_0) + \text{grad } f(x_0) \cdot (x - x_0)$$

die beste lineare Approximation von $f(x)$ nahe x_0 . Dementsprechend ist

$$x_{n+1} = f(x_0) + \text{grad } f(x_0) \cdot (x - x_0)$$

die Gleichung der Tangentialhyperebene zum Graphen der Funktion f . Insbesondere, für eine Funktion zweier Variablen:

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \quad (2.3)$$

Beispiel 2.9. Sei $f(x, y) = x^4 + 2x^3y^2 + y$. Die lineare Approximation in der Umgebung des Punktes $(1, 1)$ ist

$$\begin{aligned} f(x, y) &= f(1, 1) + \text{grad } f(1, 1) \cdot \begin{pmatrix} x-1 \\ y-1 \end{pmatrix} + o(\sqrt{(x-1)^2 + (y-1)^2}) = \\ &= 4 + 10(x-1) + 5(y-1) + o(\sqrt{(x-1)^2 + (y-1)^2}) \end{aligned}$$

Die Gleichung der Tangentialebene der Fläche $z = x^4 + 2x^3y^2 + y$ im Punkt $(1, 1, 4)$ ist

$$z = 4 + 10(x-1) + 5(y-1)$$

2.3 Die Richtungsableitung

Per Definition ist die partielle Ableitung gleich

$$\frac{\partial f}{\partial x_i} = \lim_{t \rightarrow 0} \frac{f(x + te_i) - f(x)}{t}$$

Wir können diese Definition verallgemeinern, indem wir statt e_i einen beliebigen Vektor nehmen.

Definition 2.10. Sei $v \in \mathbb{R}^n$. Der Grenzwert

$$\partial_v f(x) := \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}$$

(falls er existiert) heißt die Ableitung von f längs v . Ist v ein Einheitsvektor ($|v| = 1$), dann heißt $\partial_v f(x)$ die Richtungsableitung von f an der Stelle x in Richtung v .

Alternative Bezeichnung:

$$\partial_v f(x) = \frac{\partial f}{\partial v}(x)$$

Die anschauliche Deutung der Richtungsableitung:

- Bewegt sich der Punkt x mit konstanter Geschwindigkeit v startend vom x_0 , so beschreibt $\partial_v f(x_0)$ die momentane Änderung der Funktion f zum Zeitpunkt $t = 0$.
- Die Richtungsableitung $\partial_v f(x_0)$ ist die gewöhnliche Ableitung der Funktion f eingeschränkt auf die Gerade $\{x_0 + tv\}$ (wobei die Richtung auf der Gerade durch den Vektor v angegeben wird).
- Schneidet man den Graphen der Funktion $f(x, y)$ mit der senkrechten Ebene durch die in der xy -Ebene liegende Gerade $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, so beschreibt $\partial_v f(x_0, y_0)$ den Anstieg der auf dem Graphen ausgeschnittenen Kurve.

Satz 38. Ist die Funktion f total differenzierbar, so ist sie in jeder Richtung differenzierbar. Die Ableitung längs v berechnet sich dabei aus dem Skalarprodukt von v mit dem Gradienten von f :

$$\partial_v f(x) = \text{grad } f(x) \cdot v = \sum_{i=1}^n f_{x_i}(x) v_i \quad (2.4)$$

Beweis. Der Beweis ist analog zum Beweis des Satzes 37, Punkt (a): durch Einsetzen von $x = x_0 + tv$ in (2.2) erhalten wir

$$f(x_0 + tv) = f(x_0) + a \cdot tv + o(|tv|) = f(x_0) + ta \cdot v + o(t)$$

Folglich gilt

$$\partial_v f(x) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = a \cdot v$$

Da nach dem Satz 37 $a = \text{grad } f$ gilt, folgt daraus die Behauptung. \square

Beispiel 2.11. Um die Richtungsableitung in einer gegebenen Richtung der Funktion $f(x, y) = x^2 + y^2$ im Punkt $(1, 1)$ berechnen zu können, brauchen wir zuerst den Gradienten:

$$\text{grad } f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \Rightarrow \text{grad } f(1, 1) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Die Richtungsableitung in der Richtung $v = (\cos \alpha, \sin \alpha)$ ist nun

$$\partial_v f(1, 1) = \text{grad } f(1, 1) \cdot v = (2, 2) \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} = 2(\cos \alpha + \sin \alpha)$$

2.4 Die Kettenregel für reellwertige Funktionen

Die Kettenregel $(f \circ g)'(x) = f'(g(x))g'(x)$ für Funktionen einer Variablen kann wie folgt formuliert werden: sind $z = f(y)$ und $y = g(x)$ die Abhängigkeiten z von y und y von x , so gilt $(f \circ g)'(x) = f'(y)g'(x)$, und unter Benutzung der Bezeichnung $f'(x) = \frac{dz}{dx}$:

$$\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}$$

Analog hierzu werden wir jetzt die Kettenregel für eine Funktion mehrerer Variablen formulieren.

Satz 39. Sei $z = f(x, y)$ eine C^1 -Funktion zweier Variablen, und seien

$$x = \varphi(t), \quad y = \psi(t)$$

selbst C^1 -Funktionen einer Variablen. Dann gilt für die Ableitung der zusammengesetzten Funktion $F(t) = f(\varphi(t), \psi(t))$ die Formel

$$F'(t) = \frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}$$

Man schreibt hier $\frac{dz}{dt}$ weil $z = F(t)$ Funktion einer Variablen ist, und $\frac{\partial z}{\partial x}$ weil $z = f(x, y)$ Funktion mehrerer Variablen ist.

Natürlich gilt der Satz auch für eine Funktion $f(x_1, \dots, x_n)$ mehrerer Variablen, die alle nur von einer Variable t abhängen.

Beispiel 2.12. Finden wir $\frac{dz}{dt}$ für

$$z = e^{3x+2y}, \quad x = \cos t, \quad y = t^2$$

Es gilt

$$\begin{aligned}\frac{\partial z}{\partial x} &= 3e^{3x+2y}, & \frac{\partial z}{\partial y} &= 2e^{3x+2y} \\ \frac{dx}{dt} &= -\sin t, & \frac{dy}{dt} &= 2t\end{aligned}$$

Folglich

$$\frac{dz}{dt} = 3e^{3x+2y} \cdot (-\sin t) + 2e^{3x+2y} \cdot 2t = e^{3x+2y}(4t - 3\sin t) = e^{3\cos t + 2t^2}(4t - 3\sin t)$$

Beispiel 2.13. Ein materieller Punkt in \mathbb{R}^2 bewegt sich entlang des Bahnes

$$x = f(t), \quad y = g(t)$$

Mit welcher Geschwindigkeit entfernt er sich vom Koordinatenursprung?

Es soll die Ableitung nach t der Funktion $r(x, y) = \sqrt{x^2 + y^2}$ berechnet werden. Es gilt

$$\frac{\partial r}{\partial x} = \frac{(x^2 + y^2)_x}{2\sqrt{x^2 + y^2}} = \frac{2x}{2\sqrt{x^2 + y^2}} = \frac{x}{\sqrt{x^2 + y^2}}$$

und analog $\frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$. Folglich

$$\frac{dr}{dt} = \frac{x}{\sqrt{x^2 + y^2}} \frac{dx}{dt} + \frac{y}{\sqrt{x^2 + y^2}} \frac{dy}{dt} = \frac{f(t)}{\sqrt{f^2(t) + g^2(t)}} f'(t) + \frac{g(t)}{\sqrt{f^2(t) + g^2(t)}} g'(t)$$

Bei konkret angegebenen Funktionen f und g kann jetzt die Entfernungsgeschwindigkeit berechnet werden. So, z. B. für $x(t) = t, y(t) = t^2$ gilt

$$\sqrt{f^2(t) + g^2(t)} = \sqrt{t^2 + t^4} = |t|\sqrt{1 + t^2}$$

und daher

$$\begin{aligned}\frac{dr}{dt} &= \frac{t}{\sqrt{t^2 + t^4}} + \frac{t^2}{\sqrt{t^2 + t^4}} = \frac{1}{\sqrt{1 + t^2}} \left(\frac{t}{|t|} + 2t^2 \frac{t}{|t|} \right) = \\ &= \begin{cases} \frac{1+2t^2}{\sqrt{1+t^2}} & \text{falls } t > 0 \\ -\frac{1+2t^2}{\sqrt{1+t^2}} & \text{falls } t < 0 \end{cases}\end{aligned}$$

Satz 40. Sei $z = f(x, y)$ eine C^1 -Funktion zweier Variablen. Seien x und y selbst C^1 -Funktionen zweier Variablen

$$x = \varphi(u, v), \quad y = \psi(u, v)$$

Dann berechnen sich die partiellen Ableitungen der zusammengesetzten Funktion $F(u, v) = f(\varphi(u, v), \psi(u, v))$ wie folgt:

$$\begin{aligned}\frac{\partial F}{\partial u} &= \frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u} \\ \frac{\partial F}{\partial v} &= \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v}\end{aligned}$$

Ähnliche Formeln gelten für eine Funktion $f(x_1, \dots, x_n)$, wobei die n Variablen x_i von m Variablen y_1, \dots, y_m abhängen.

Beispiel 2.14. Oft wird es benötigt, eine Funktion $f(x, y)$ in die Polarkoordinaten (r, φ) zu transformieren. D. h. man setzt

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

und definiert

$$F(r, \varphi) = f(x, y) = f(r \cos \varphi, r \sin \varphi)$$

Drücken wir die partiellen Ableitungen $\frac{\partial F}{\partial r}$ und $\frac{\partial F}{\partial \varphi}$ durch $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ aus.

Dafür berechnen wir

$$\begin{aligned} \frac{\partial x}{\partial r} &= \cos \varphi, & \frac{\partial x}{\partial \varphi} &= -r \sin \varphi \\ \frac{\partial y}{\partial r} &= \sin \varphi, & \frac{\partial y}{\partial \varphi} &= r \cos \varphi \end{aligned}$$

Folglich

$$\begin{aligned} \frac{\partial F}{\partial r} &= \frac{\partial f}{\partial x} \cos \varphi + \frac{\partial f}{\partial y} \sin \varphi \\ \frac{\partial F}{\partial \varphi} &= \frac{\partial f}{\partial x} (-r \sin \varphi) + \frac{\partial f}{\partial y} (r \cos \varphi) \end{aligned}$$

3 Anwendungen der Differentiation

3.1 Die Bedeutung des Gradienten

Richtung des stärksten Anstiegs

Die Formel (2.4) und die anschauliche Interpretation der Richtungsableitung bedeuten:

Der Anstieg der Funktion f in der Richtung v ist gleich $\text{grad } f(x) \cdot v$.

Wollen wir die Richtung des stärksten Anstiegs finden, müssen wir unter allen Einheitsvektoren denjenigen finden, dessen Skalarprodukt mit $\text{grad } f$ am größten ist. Es gilt

$$\text{grad } f(x) \cdot v = |\text{grad } f| \cos \alpha \quad (\text{unter Voraussetzung } |v| = 1)$$

wobei α der Winkel zwischen $\text{grad } f$ und v ist. Deswegen ist das Skalarprodukt dann am größten, wenn $\alpha = 0$, d. h. wenn v in die Richtung des Gradienten zeigt.

Der Vektor $\text{grad } f(x)$ zeigt in der Richtung des maximalen Anstiegs der Funktion f im Punkt x .

Analog, $-\text{grad } f$ zeigt in der Richtung des stärksten Abnehmens der Funktion f .

Der Gradient gibt nicht nur die Richtung des stärksten Anstiegs, sondern auch die Geschwindigkeit dieses Anstiegs. Der Einheitsvektor in der Richtung von $\text{grad } f$ ist $v = \frac{\text{grad } f}{|\text{grad } f|}$ und es gilt

$$|\partial_v f| = |v \cdot \text{grad } f| = \left| \frac{|\text{grad } f|^2}{|\text{grad } f|} \right| = |\text{grad } f|$$

Der Betrag des Gradienten ist gleich dem stärksten Anstieg der Funktion.

Beispiel 3.1. Der Gradient der Funktion $f(x, y) = x^2 - y^2$ ist

$$\text{grad } f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$$

Folglich gibt der Vektor $(x, -y)$ die Richtung des stärksten Anstiegs an. Die Stärke des Anstiegs ist gleich $|\text{grad } f(x, y)| = 2\sqrt{x^2 + y^2}$.

Normale, Tangente und Tangentialebene zu Niveaukurve oder Niveaufläche

Wird eine Fläche in \mathbb{R}^3 als Graph einer Funktion $f(x, y)$ dargestellt, so wird die Gleichung der Tangentialebene mittels (2.3) gegeben. Eine Fläche kann aber auch als Niveaufläche einer Funktion dreier Variablen gegeben werden: $F(x, y, z) = c$. So wird z. B. ein Ellipsoid durch $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ beschrieben.

Analog wird eine Kurve in \mathbb{R}^2 oft als die Lösungsmenge einer Gleichung $F(x, y) = c$ beschrieben, d. h. als eine Niveaukurve der Funktion F .

Um die Gleichung der Tangente (bzw. der Tangentialebene) zu solchen implizit gegebenen Kurven oder Flächen schreiben zu können, machen wir die folgende Beobachtung.

Lemma 32. Sei (x_0, y_0) ein Punkt auf der Niveaukurve $N_c = \{F(x, y) = c\}$ der Funktion F . Dann ist der Vektor $\text{grad } F(x_0, y_0)$ orthogonal zu N_c im Punkt (x_0, y_0) .

Analog ist $\text{grad } F(x_0, y_0, z_0)$ orthogonal zur Niveaufläche $N_c = \{F(x, y, z) = c\}$ im Punkt $(x_0, y_0, z_0) \in N_c$.

Beweis. Wir geben eine anschauliche Erklärung. Entlang einer Tangente zur Niveaukurve ändern sich die Funktionswerte nur gering. D. h. $\partial_v F(x_0, y_0) = 0$, sobald v ein Tangentialvektor zu N_c im Punkt (x_0, y_0) ist. Aus der Formel für die Richtungsableitung folgt

$$\text{grad } F(x_0, y_0) \cdot v = 0 \tag{2.5}$$

Also ist $\text{grad } F(x_0, y_0)$ orthogonal zu jedem Tangentialvektor der Niveaukurve, d. h. orthogonal zur Niveaukurve selbst.

Der Fall einer Funktion dreier Variablen ist analog. □

Das gibt uns zuerst eine neue geometrische Deutung des Gradienten:

Der Gradient $\text{grad } F(x)$ ist orthogonal zu der Niveaufläche (bzw. Niveaukurve) von F durch den Punkt x .

Beispiel 3.2. Es hilft hier, einige Niveaukurven und Gradienten der Funktion $f(x, y) = x^2 - y^2$ anzuzeichnen.

Beispiel 3.3. Finden wir einen Normalenvektor zur Ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ im Punkt (x_0, y_0) .

Die Ellipse ist eine Niveaufläche der Funktion $F(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$. Der Gradient von F :

$$\text{grad } F(x_0, y_0) = \left(\frac{2x_0}{a^2}, \frac{2y_0}{b^2} \right)$$

Also ist $(\frac{x_0}{a^2}, \frac{y_0}{b^2})$ ein Normalenvektor zur Ellipse im Punkt (x_0, y_0) .

Wollen wir eine Einheitsnormale finden, so müssen wir den gefundenen Vektor v normieren:

$$n := \frac{v}{|v|} = \frac{1}{\sqrt{\frac{x_0^2}{a^4} + \frac{y_0^2}{b^4}}} \left(\frac{x_0}{a^2}, \frac{y_0}{b^2} \right)$$

Definieren wir den Winkel zwischen zwei Flächen im Punkt P als den Winkel zwischen den Tangentialebenen zu den beiden Flächen im Punkt P . Derselbe Winkel kann auch als der Winkel zwischen den Normalenvektoren zu den beiden Flächen berechnet werden.

Beispiel 3.4. Unter welchem Winkel schneiden sich der Zylinder $x^2 + y^2 = R^2$ und die Sphäre $(x - R)^2 + y^2 + z^2 = R^2$ im Punkt $(\frac{R}{2}, \frac{R\sqrt{3}}{2}, 0)$?

Für $F(x, y, z) = x^2 + y^2$ und $G(x, y, z) = (x - R)^2 + y^2 + z^2$ berechnen wir die Gradienten

$$\text{grad } F(x, y, z) = (2x, 2y, 0), \quad \text{grad } G(x, y, z) = (2(x - R), 2y, 2z)$$

Durch Einsetzen der Koordinaten $(\frac{R}{2}, \frac{R\sqrt{3}}{2}, 0)$ und Skalieren erhalten wir die Einheitsnormalen zu beiden Flächen:

$$n_1 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}, 0 \right), \quad n_2 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0 \right)$$

Der Kosinus des Winkels zwischen n_1 und n_2 ist gleich

$$n_1 \cdot n_2 = \frac{1}{2}$$

Folglich ist der Winkel gleich 60° .

Ein Punkt (x, y) liegt auf der Tangente zu N_c im Punkt (x_0, y_0) genau dann, wenn $v = (x - x_0, y - y_0)$ ein Tangentialvektor zu N_c ist. Wegen (2.5) erhalten wir die *Normalengleichung der Tangente an die Niveaukurve* $F(x, y) = c$ im Kurvenpunkt (x_0, y_0) :

$$F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0) = 0$$

Und analog die *Normalengleichung der Tangentialebene an die Niveaufläche* $F(x, y, z) = c$:

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0 \quad (2.6)$$

Beispiel 3.5. Finden wir die Gleichung der Tangente zur Ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ im Punkt (x_0, y_0) .

Die Antwort:

$$\frac{x_0}{a^2}(x - x_0) + \frac{y_0}{b^2}(y - y_0) = 0$$

Da (x_0, y_0) auf der Ellipse liegt, gilt $\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1$, und die obige Gleichung ist äquivalent zu

$$\frac{x_0}{a^2}x + \frac{y_0}{b^2}y - 1 = 0$$

Normale und Tangentialebene zum Graphen einer Funktion

Im Abschnitt 2.2 haben wir die Gleichung (2.3) der Tangentialebene zum Graphen einer Funktion $f(x, y)$ hergeleitet. Dieselbe Gleichung entsteht, wenn wir den Graphen

$$\Gamma_f = \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\}$$

als die Niveauläche einer Funktion von *drei* Variablen interpretieren. Es gilt nämlich

$$\Gamma_f = \{(x, y, z) \in \mathbb{R}^3 \mid F(x, y, z) = 0\}$$

wobei $F(x, y, z) := f(x, y) - z$ gesetzt wird. Wegen

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x}, \quad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y}, \quad \frac{\partial F}{\partial z} = -1$$

und $z_0 = f(x_0, y_0)$ entsteht beim Einsetzen in (2.6) wieder die Gleichung (2.3).

Die Darstellung des Graphen als Niveauläche erlaubt uns auch, einen Normalenvektor zum Γ_f im Punkt $(x_0, y_0, f(x_0, y_0))$ zu finden. Das ist der Gradient der Funktion $F(x, y, z)$:

$$\text{grad } F(x_0, y_0, z_0) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1 \right)$$

3.2 Approximation höherer Ordnung: die Taylor-Formel

Die Taylor-Formel

Wir wollen die Taylor-Formel für Funktionen einer Variablen auf den Fall mehrerer Variablen verallgemeinern. Sei $f: D \rightarrow \mathbb{R}$ eine C^k -Funktion auf einer offenen Menge $D \subset \mathbb{R}^n$. Sei $x \in D$. Unser Ziel ist es, den Wert von f in einem nahegelegenen Punkt $x + v$ mittels den Werten von f und allen ihren partiellen Ableitungen bis zu Ordnung k im Punkt x zu approximieren. Die lineare Approximation

$$f(x + v) \approx f(x) + \text{grad } f(x) \cdot v = f(x) + \partial_v f(x)$$

ist der erste Schritt.

Mit dem folgenden Trick können wir das Problem wieder auf die Funktion einer Variablen zurückführen. Betrachten wir die Einschränkung von f auf die Gerade durch x und $x + v$ und bezeichnen

$$h(t) := f(x + tv)$$

Die Variable t ist ein Parameter entlang der Gerade, und es gilt

$$h(0) = f(x), \quad h(1) = f(x + v)$$

Die Approximation mit dem Taylor-Polynom vom Grad k lautet $h(t) \approx h(0) + t\dot{h}(0) + \frac{t^2}{2}\ddot{h}(0) + \dots + \frac{t^k}{k!}h^{(k)}(0)$. Insbesondere, bei $t = 1$

$$h(1) \approx h(0) + \dot{h}(0) + \frac{1}{2}\ddot{h}(0) + \dots + \frac{1}{k!}h^{(k)}(0)$$

Wir müssen allerdings voraussetzen, dass die Funktion h auf dem Intervall $[0, 1]$ definiert ist. Das heißt, die die Punkte x und $x + v$ verbindende Strecke soll in der Menge D enthalten sein. Einfachheit halber nehmen wir an, dass es für je zwei Punkte in D gilt.

Definition 3.6. Eine Menge $D \subset \mathbb{R}^n$ heißt konvex, wenn für je zwei Punkte $x, y \in D$ die Strecke zwischen x und y ganz in D enthalten ist.

Die Ableitung $h^{(k)}(0) = \frac{d^k}{dt^k} \Big|_{t=0} f(x + tv)$ wird mit $\partial_v^k f(x)$ bezeichnet:

$$\begin{aligned} \dot{h}(0) &= \frac{d}{dt} \Big|_{t=0} f(x + tv) = \partial_v f(x) \\ \ddot{h}(0) &= \frac{d^2}{dt^2} \Big|_{t=0} f(x + tv) =: \partial_v^2 f(x) \end{aligned}$$

Als Endergebnis erhalten wir

Satz 41. Ist $D \subset \mathbb{R}^n$ ein konvexes Gebiet, $f \in C^k(D)$, $x \in D$, $x + v \in D$, dann gilt

$$f(x + v) = f(x) + \partial_v f(x) + \frac{1}{2!} \partial_v^2 f(x) + \dots + \frac{1}{k!} \partial_v^k f(x) + R_k(x, v) \quad (2.7)$$

mit dem Restglied

$$R_k(x, v) = \frac{1}{(k+1)!} \partial_v^{k+1} f(x + \xi_k v)$$

mit einer Zahl ξ_k zwischen 0 und 1.

Die ersten Terme der Taylor-Formel (2.7) liefern die beste Approximation bis zur k -ten Ordnung der Funktion f nahe x . Das heißt, $R_k(x, v) = o(|v|^k)$.

Höhere Richtungsableitungen und höhere partielle Ableitungen

Die höhere Richtungsableitung $\partial_v^k f$ kann durch die partiellen Ableitungen k -ter Ordnung der Funktion f ausgedrückt werden. Um diese Darstellung zu finden, schreiben wir

$$\partial_v = v_1 \frac{\partial}{\partial x_1} + \dots + v_n \frac{\partial}{\partial x_n}$$

und berechnen ∂_v^k durch das Ausmultiplizieren von

$$\left(v_1 \frac{\partial}{\partial x_1} + \dots + v_n \frac{\partial}{\partial x_n} \right)^k$$

Insbesondere, für eine Funktion zweier Variablen sind die Ableitungen längs $v = (x, y)$ gleich

$$\begin{aligned} \partial_v^2 f &= \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right)^2 f = x^2 \frac{\partial^2 f}{\partial x^2} + 2xy \frac{\partial^2 f}{\partial x \partial y} + y^2 \frac{\partial^2 f}{\partial y^2} \\ \partial_v^3 f &= \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right)^3 f = x^3 \frac{\partial^3 f}{\partial x^3} + 3x^2 y \frac{\partial^3 f}{\partial x^2 \partial y} + 3y^2 x \frac{\partial^3 f}{\partial x \partial y^2} + y^3 \frac{\partial^3 f}{\partial y^3} \end{aligned}$$

Daraus erhalten wir die ersten Terme der Taylor-Entwicklung mit Zentrum $(0, 0)$:

$$\begin{aligned}
 f(x, y) &= f(0, 0) + f_x(0, 0)x + f_y(0, 0)y + \\
 &\quad + \frac{1}{2}f_{xx}(0, 0)x^2 + f_{xy}(0, 0)xy + \frac{1}{2}f_{yy}(0, 0)y^2 + \\
 &\quad + \frac{1}{6}f_{xxx}(0, 0)x^3 + \frac{1}{2}f_{xxy}(0, 0)x^2y + \frac{1}{2}f_{xyy}(0, 0)xy^2 + \frac{1}{6}f_{yyy}(0, 0)y^3 + \dots
 \end{aligned}$$

Beispiel 3.7. Berechnen wir die Taylor-Entwicklung der Funktion

$$f(x, y) = x^3 - 2y^3 + 3xy$$

mit Zentrum $(1, 2)$.

Die Taylor-Entwicklung kann nicht nur durch Berechnen der partiellen Ableitungen gefunden werden, sondern auch durch Operationen mit bekannten Reihenentwicklungen.

Beispiel 3.8.

$$\begin{aligned}
 e^{x+y} &= 1 + (x+y) + \frac{(x+y)^2}{2} + \frac{(x+y)^3}{6} + o((x+y)^3) \\
 &= 1 + x + y + \frac{x^2}{2} + xy + \frac{y^2}{2} + \frac{x^3}{6} + \frac{x^2y}{2} + \frac{xy^2}{2} + \frac{y^3}{6} + o(\sqrt{x^2 + y^2}^3)
 \end{aligned}$$

Die Hesse-Matrix und die quadratische Approximation

Für die zweite Richtungsableitung einer C^2 -Funktion $f(x, y)$ gilt:

$$\partial_v^2 f = f_{xx}v_1^2 + 2f_{xy}v_1v_2 + f_{yy}v_2^2$$

Im Allgemeinen ist die zweite Ableitung längs v eine quadratische Form in v :

$$\partial_v^2 f = q(v) = \sum_{i,j} a_{ij}v_i v_j \quad \text{mit} \quad a_{ij} = f_{x_i x_j}$$

Definition 3.9. Die Matrix

$$H_f(x) = \begin{pmatrix} f_{x_1 x_1}(x) & \cdots & f_{x_1 x_n}(x) \\ \cdots & \cdots & \cdots \\ f_{x_n x_1}(x) & \cdots & f_{x_n x_n}(x) \end{pmatrix}$$

heißt die Hesse-Matrix von f im Punkt x .

Mit diesen Bezeichnungen kann die zweite Richtungsableitung an der Stelle x längs v geschrieben werden wie

$$\partial_v^2 f(x) = v^\top H_f(x)v$$

Approximieren wir den Wert $f(x)$ durch die Taylor-Formel mit Zentrum x_0 , so müssen wir in (2.7) $x = x_0$ und $v = x - x_0$ einsetzen. Das ergibt bei $k = 2$

$$f(x) = f(x_0) + \text{grad } f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)^\top H_f(x_0)(x - x_0) + o(|x - x_0|^2)$$

als die *beste quadratische Approximation* von f nahe x_0 .

Während die lineare Approximation

$$f(x) \approx y = f(x_0) + \text{grad } f(x_0) \cdot (x - x_0)$$

die Tangentialebene zum Graphen von f beschreibt, beschreibt die quadratische Approximation

$$f(x) \approx y = f(x_0) + \text{grad } f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)^\top H_f(x_0)(x - x_0)$$

die *Schmiegequadratik* zum Graphen.

3.3 Lokale Minima und Maxima

Die notwendige Bedingung für lokales Extremum

Analog zu Funktionen einer Variablen,

Definition 3.10. Sei $f: D \rightarrow \mathbb{R}$ eine Funktion von n Variablen. Ein Punkt $a \in D$ heißt lokale Maximalstelle (bzw. Minimalstelle) von f , wenn es eine r -Umgebung $U_r(a)$ von a gibt, so dass für alle $x \in U_r \cap D$ gilt:

$$f(x) \leq f(a) \quad (\text{bzw. } f(x) \geq f(a))$$

Ist $f(x)$ eine Funktion einer Variablen, so haben wir für eine lokale Extremstelle in a

die notwendige Bedingung: $f'(a) = 0$

die hinreichende Bedingung: $f'(a) = 0$ und $f''(a) > 0$ (Maximum)
oder $f''(a) < 0$ (Minimum)

Für Funktionen mehrerer Variablen müssen wir $\text{grad } f$ statt f' betrachten.

Satz 42. Ist $f: D \rightarrow \mathbb{R}$ eine C^1 -Funktion und a ein innerer Punkt von D , so gilt

$$a \text{ ist lokale Extremstelle von } f \Rightarrow \text{grad } f(a) = 0$$

Beweis. Widerspruchsbeweis. Ist $\text{grad } f(a) \neq 0$, so steigt die Funktion in der Richtung des Gradienten und fällt in der Gegenrichtung. Also ist a in diesem Fall keine Extremstelle. \square

Ein Punkt $a \in \mathbb{R}^n$ mit $\text{grad } f(a) = 0$ heißt *stationärer Punkt*. Der obige Satz besagt, dass lokale Extremstellen (im Inneren des Definitionsbereichs) unter den stationären Punkten zu suchen sind. Aber nicht jeder stationäre Punkt ist eine Extremstelle. (Vergleiche die Situation mit dem Fall der Funktion einer Variablen.) Ein stationärer Punkt, der keine Extremstelle ist, heißt *Sattelpunkt*.

Beispiel 3.11. Die Funktion $f(x, y) = x^2 - y^2$ hat den Gradienten $\text{grad } f(x, y) = (2x, -2y)$. Demnach ist $(0, 0)$ ein stationärer Punkt von f . Das ist allerdings keine Extremstelle, denn $f(0, 0) = 0$ und nahe $(0, 0)$ gibt es sowohl Punkte mit negativen Werten von f als auch Punkte mit positiven Werten.

Die hinreichende Bedingung für lokales Extremum

Genau wie bei einer Variablen, liefern die zweiten Ableitungen eine hinreichende Bedingung für lokales Extremum. Die zweiten partiellen Ableitungen bilden die Hesse-Matrix $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$.

Satz 43. Ist $f: D \rightarrow \mathbb{R}$ eine C^2 -Funktion und $a \in D$ ein innerer Punkt, der ein stationärer Punkt von f ist, so gilt:

- a) $H_f(a)$ positiv definit $\Rightarrow a$ ist lokale Minimalstelle,
- b) $H_f(a)$ negativ definit $\Rightarrow a$ ist lokale Maximalstelle,
- c) $H_f(a)$ indefinit $\Rightarrow a$ ist Sattelpunkt.

Beweis. Bei grad $f(a)$ lautet die quadratische Approximation von f in der Nähe von a :

$$f(x) = f(a) + \frac{1}{2}(x-a)^\top H_f(a)(x-a) + o(|x-a|^2)$$

Ist die Matrix $H_f(a)$ positiv (negativ) definit, so gilt per Definition

$$v^\top H_f(a)v > 0 \quad (\text{bzw. } < 0)$$

für alle $v \in \mathbb{R}^n$, und folglich

$$f(a) + \frac{1}{2}(x-a)^\top H_f(a)(x-a) > f(a) \quad (\text{bzw. } < 0)$$

für alle x .

Der Term $o(|x-a|^2)$ ist im Vergleich zum quadratischen Term verschwindend klein, deswegen bleiben diese Ungleichungen in einer kleiner Umgebung von a gültig auch nach dem Addieren von $o(|x-a|^2)$ zur linken Seite. Also ist a eine lokale Minimalstelle (bzw. Maximalstelle).

Ist die Matrix $H_f(a)$ indefinit, so nimmt die quadratische Form $v^\top H_f(a)v$ sowohl negative als auch positive Werte an. Dementsprechend nimmt die Funktion f nahe a sowohl kleinere als auch größere als $f(a)$ Werte an. \square

Der Satz macht keine Aussage im Falle wenn $H_f(a)$ positiv semidefinit oder negativ semidefinit ist.

Beispiel 3.12. Finden wir lokale Extrema der Funktion

$$f(x, y) = y^2 + 2xy + x^3 - 8y - 8x - 8x^2$$

3.4 Globale Minima und Maxima

Die Funktion $f(x, y) = y^2 + 2xy + x^3 - 8y - 8x - 8x^2$ nimmt kein Maximum und kein Minimum an, denn sie ist weder nach oben, noch nach unten beschränkt: z. B. bei $y = 0$ haben wir $f(x, 0) = x^3 - 8x^2 - 8x$, und das geht gegen $+\infty$ bei $x \rightarrow +\infty$ und gegen $-\infty$ bei $x \rightarrow -\infty$.

Anders ist es bei der Funktion $f(x, y) = \frac{1}{x^2+y^2+1}$. Es gilt $f(0, 0) = 1$, und ansonsten

$$\frac{1}{x^2+y^2+1} \leq 1$$

Deswegen ist $(0, 0)$ die globale Maximalstelle von f . Das globale Minimum wird nicht erreicht, denn $\lim_{x^2+y^2 \rightarrow \infty} f(x, y) = 0$, aber nirgendwo gilt $f(x, y) = 0$.

Eine stetige Funktion auf einer kompakten (beschränkten und abgeschlossenen) Menge D nimmt ihren minimalen und maximalen Wert an. Für die globalen Minimalstellen und Maximalstellen gibt es die folgenden Kandidaten:

- die stationären Punkte im Inneren;
- die Punkte auf dem Rand von D .

Ist $D \subset \mathbb{R}^2$ ein Polygon, so kann die Funktion $f(x, y)$ auf jedem geraden Abschnitt des Randes auf eine Funktion einer Variable zurückgeführt werden.

Beispiel 3.13. Bestimmen wir die globalen Minima und Maxima der Funktion

$$f(x, y) = x^2 + y^2 - xy + x + y$$

auf dem Gebiet

$$x \leq 0, \quad y \leq 0, \quad x + y \geq -3$$

Zuerst finden wir die stationären Punkte:

$$\text{grad } f = (2x - y + 1, 2y - x + 1)$$

Das Lösen eines LGS liefert $(x, y) = (-1, -1)$. Der Wert in diesem Punkt ist $f(-1, -1) = -1$. (Keine Untersuchung auf lokales Extremum ist jetzt nötig!)

Nun betrachten wir die Einschränkungen von f auf jede Seite des Dreiecks.

Bei $x = 0$ untersuchen wir $f(0, y) = y^2 + y$ für $y \in [-3, 0]$. Die minimalen und maximalen Werte sind

$$f(0, -\frac{1}{2}) = -\frac{1}{4}, \quad f(0, -3) = 6$$

Analog, bei $y = 0$ erhalten wir die minimalen und maximalen Werte

$$f(-\frac{1}{2}, 0) = -\frac{1}{4}, \quad f(-3, 0) = 6$$

Bei $x + y = -3$ setzen wir $y = -3 - x$ ein und erhalten

$$g(x) = f(x, -3 - x) = 3x^2 + 9x + 6 \quad \text{auf } [-3, 0]$$

Die minimalen und maximalen Werte sind

$$g(-\frac{3}{2}) = -\frac{3}{4}, \quad g(-3) = g(0) = 6$$

Das Vergleichen aller ermittelten Werte liefert:

$$f_{\min} = f(-1, -1) = -1, \quad f_{\max} = f(0, -3) = f(-3, 0) = 6$$

3.5 Extremwertaufgaben mit Nebenbedingungen

Oft soll eine Funktion $f(x_1, \dots, x_n)$ nicht auf einer n -dimensionalen Menge D minimiert, bzw. maximiert werden, sondern auf einer Fläche oder Kurve. Dann spricht man über Extremwertaufgaben unter Nebenbedingungen. Die Nebenbedingungen, die die Menge der zulässigen Werte von x beschreiben, werden oft durch Gleichungen angegeben:

$$g_1(x_1, \dots, x_n) = 0, \dots, g_k(x_1, \dots, x_n) = 0$$

Wir werden zunächst die Nebenbedingungen betrachten, die durch nur eine Gleichung gegeben sind (Extremstellen einer Funktion auf einer Kurve in \mathbb{R}^2 oder auf einer Fläche in \mathbb{R}^3). Ein solches Problem schreiben wir kurz auf als

$$f(x) \rightarrow \text{Extr} \quad \text{mit NB } g(x) = 0$$

Beispiel 3.14. Finde den kleinsten und den größten Wert der Funktion $f(x, y, z) = xy + xz$ auf der Einheitskugel $x^2 + y^2 + z^2 = 1$. D. h.

$$xy + xz \rightarrow \text{Extr} \quad \text{mit NB } x^2 + y^2 + z^2 = 1$$

Wir studieren drei Methoden, Extremwertaufgaben mit Nebenbedingungen zu lösen.

1. Methode (Die explizite Methode). Man löst, falls möglich, $g(x) = 0$ nach einer Variablen auf, z. B. $x_n = h(x_1, \dots, x_{n-1})$. Durch diese Substitution eliminiert man x_n aus f . Das führt zu einem Extremwertproblem ohne Nebenbedingungen

$$f(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1})) \rightarrow \text{Extr}$$

Lösung des Beispielproblems. Das Auflösen der Nebenbedingung nach x :

$$x = \pm \sqrt{1 - y^2 - z^2}$$

Wegen $f(x, y, z) = f(-x, -y, -z)$ genügt es, nur nichtnegative Werte von x zu betrachten. Also setzen wir $x = \sqrt{1 - y^2 - z^2}$ und bestimmen

$$F(y, z) = (y + z)\sqrt{1 - y^2 - z^2} \rightarrow \text{Extr} \quad \text{für } y^2 + z^2 \leq 1$$

Zuerst suchen wir die inneren stationären Punkte

$$\frac{\partial F}{\partial y} = \frac{1 - 2y^2 - z^2 - yz}{\sqrt{1 - y^2 - z^2}}, \quad \frac{\partial F}{\partial z} = \frac{1 - y^2 - 2z^2 - yz}{\sqrt{1 - y^2 - z^2}}$$

Also $\text{grad } F = 0$ genau dann, wenn

$$\begin{aligned} 1 - 2y^2 - z^2 - yz &= 0 \\ 1 - y^2 - 2z^2 - yz &= 0 \end{aligned}$$

Das führt zu $y^2 = z^2$ und schließlich $y = z = \pm \frac{1}{2}$ oder $y = -z = \pm \frac{1}{\sqrt{2}}$. Die letzteren zwei Lösungen sind keine inneren Punkte des Kreises $y^2 + z^2 \leq 1$. Die stationären Punkte im Inneren sind also

$$p_1 = \left(\frac{1}{2}, \frac{1}{2} \right), \quad p_2 = \left(-\frac{1}{2}, -\frac{1}{2} \right)$$

Die Funktionenwerte:

$$F(p_1) = \frac{\sqrt{2}}{2}, \quad F(p_2) = -\frac{\sqrt{2}}{2}$$

Auf dem Rand $y^2 + z^2 = 1$ gilt $F(y, z) = 0$, was kleiner als $F(p_1)$ und größer als $F(p_2)$ ist. Auf dem Rand gibt es also keine Extrema.

Die Antwort: Das Minimum ist

$$f\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right) = f\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right) = -\frac{\sqrt{2}}{2}$$

Das Maximum:

$$f\left(\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right) = f\left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right) = \frac{\sqrt{2}}{2}$$

2. Methode (Parametrisierung der Nebenbedingungen). Im Fall zweier Variablen sucht man eine Parameterdarstellung $x = x(t), y = y(t)$ der Kurve $g(x, y) = 0$. Im Fall dreier Variablen sucht man eine Parameterdarstellung $x = x(u, v), y = y(u, v), z = z(u, v)$ der Fläche $g(x, y, z) = 0$.

Lösung des Beispielproblems. Geeignet in diesem Fall sind die *Kugelkoordinaten* in \mathbb{R}^3 :

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta \end{aligned}$$

($r = \sqrt{x^2 + y^2 + z^2}$ ist der Radius, $\varphi \in [0, 2\pi]$ ist der Asimutwinkel, $\theta \in [0, \pi]$ ist der Polarwinkel)

Die Einheitssphäre entspricht $r = 1$ und wird durch (θ, φ) parametrisiert. Für unseres spezielles Problem ist es bequemer, die Koordinaten zu vertauschen:

$$\begin{aligned} x &= r \cos \theta \\ y &= r \sin \theta \cos \varphi \\ z &= r \sin \theta \sin \varphi \end{aligned}$$

Das Extremwertproblem wird zu

$$F(\theta, \varphi) = \sin \theta \cos \theta (\cos \varphi + \sin \varphi) \rightarrow \text{Extr} \quad \text{für } \varphi \in [0, 2\pi], \theta \in [0, \pi]$$

Das kann wieder durch erst Finden der stationären Punkte im Inneren, dann Studieren der Funktion auf dem Rand gelöst werden. Alternativ kann man $F(\theta, \varphi)$ umformen:

$$F(\theta, \varphi) = \frac{\sqrt{2}}{2} \sin 2\theta \sin\left(\varphi + \frac{\pi}{4}\right)$$

Daran sieht man, dass der maximale Wert $\frac{\sqrt{2}}{2}$ bei $\sin 2\theta = \sin\left(\varphi + \frac{\pi}{4}\right) = 1$ und bei $\sin 2\theta = \sin\left(\varphi + \frac{\pi}{4}\right) = -1$ angenommen wird, und der minimale bei $\sin 2\theta = 1, \sin\left(\varphi + \frac{\pi}{4}\right) = -1$ und bei $\sin 2\theta = -1, \sin\left(\varphi + \frac{\pi}{4}\right) = 1$. Daraus berechnet man auch die (x, y, z) -Koordinaten der Extremstellen.

3. Methode (Die Lagrange-Multiplikatorregel). Diese Methode basiert auf dem folgenden

Satz 44. Sei $f(x) \rightarrow \text{Extr}$ mit $g(x) = 0$ ein Extremwertproblem, wobei f und g C^1 -Funktionen sind. Ist a eine Lösung des Problems, und gilt $\text{grad } g(a) \neq 0$, so gibt es eine Zahl $\lambda_0 \in \mathbb{R}$, sodass gilt

$$\text{grad } f(a) + \lambda_0 \text{grad } g(a) = 0 \quad (2.8)$$

Dieser Satz beschreibt die Kandidaten für die globalen Extremstellen der Funktion f auf der Menge $\{g(x) = 0\}$. Das sind

- Punkte, wo $\text{grad } g = 0$ gilt;
- Punkte, wo (2.8) für ein $\lambda_0 \in \mathbb{R}$ gilt.

Beweis. Wir geben eine anschauliche Begründung im Fall zweier Variablen. Sei N_c die Niveaukurve der Funktion f durch den Punkt a . Ist a eine Extremstelle der Funktion f auf der Kurve $g(x) = 0$, so sollen die Kurven N_c und $g(x) = 0$ im Punkt a parallel sein. Sonst können wir den Punkt a längs der Kurve $g(x) = 0$ verschieben, sodass der Wert von f dabei größer wird. Wenn N_c und $g(x) = 0$ tangential zueinander sind, dann sind ihre Gradienten im Punkt a parallel, und das ist genau die Bedingung (2.8).

Den Fall $\text{grad } g = 0$ muss man berücksichtigen, weil in diesem Fall die Kurve $g(x) = 0$ eventuell keine Tangente im Punkt a besitzt. \square

Lagrange-Multiplikatorregel

1. *Schritt.* Man bildet die Hilfsfunktion $L(x_1, \dots, x_n, \lambda) = f(x_1, \dots, x_n) + \lambda g(x_1, \dots, x_n)$ von $n + 1$ Variablen und berechnet $\text{grad } L$.

2. *Schritt.* Man bestimmt die Lösungen des Gleichungssystems $\text{grad } L(x_1, \dots, x_n, \lambda) = 0$, d. h.

$$\begin{aligned} f_{x_1}(x_1, \dots, x_n) + \lambda g_{x_1}(x_1, \dots, x_n) &= 0 \\ &\dots \\ f_{x_n}(x_1, \dots, x_n) + \lambda g_{x_n}(x_1, \dots, x_n) &= 0 \\ g(x_1, \dots, x_n) &= 0 \end{aligned}$$

3. *Schritt.* Die gefundenen Werte (a_1, \dots, a_n) , sowie die mit $\text{grad } g = 0$, sind Kandidaten für Extremalstellen (der Wert λ_0 spielt keine Rolle mehr). Durch den Vergleich der Werte von f werden der Minimalwert und der Maximalwert ermittelt.

Lösung des Beispielproblems. Es gilt

$$L(x, y, z, \lambda) = xy + xz + \lambda(x^2 + y^2 + z^2 - 1)$$

Berechnung des Gradienten:

$$\begin{aligned} \frac{\partial L}{\partial x} &= y + z + 2\lambda x \\ \frac{\partial L}{\partial y} &= x + 2\lambda y \\ \frac{\partial L}{\partial z} &= x + 2\lambda z \\ \frac{\partial L}{\partial \lambda} &= x^2 + y^2 + z^2 - 1 \end{aligned}$$

Lösen des Gleichungssystems

$$\begin{aligned}y + z + 2\lambda x &= 0 \\x + 2\lambda y &= 0 \\x + 2\lambda z &= 0 \\x^2 + y^2 + z^2 - 1 &= 0\end{aligned}$$

Wir unterscheiden zwei Fälle:

a) $\lambda = 0$. Dann $x = 0$ und $y + z = 0$. Aus der letzten Gleichung findet man $(y, z) = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ oder $(y, z) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. Der Wert von f ist in beiden Punkten gleich 0.

b) $\lambda \neq 0$. Dann folgt aus der zweiten und dritten Gleichung $y = z$. Für x und y ergibt sich ein lineares Gleichungssystem

$$\begin{aligned}\lambda x + y &= 0 \\x + 2\lambda y &= 0\end{aligned}$$

Ist $\det \begin{pmatrix} \lambda & 1 \\ 1 & 2\lambda \end{pmatrix} \neq 0$, so hat das System nur die Lösung $x = y = 0$. Wegen $z = 0$ wird aber die Gleichung $x^2 + y^2 + z^2 - 1 = 0$ nicht erfüllt.

Wir berechnen

$$\det \begin{pmatrix} \lambda & 1 \\ 1 & 2\lambda \end{pmatrix} = 2\lambda^2 - 1$$

Die Determinante verschwindet also bei $\lambda = \frac{\sqrt{2}}{2}$ und bei $\lambda = -\frac{\sqrt{2}}{2}$. Im ersten Fall haben wir

$$\begin{aligned}x = -\sqrt{2}y, \quad y = z, \quad x^2 + y^2 + z^2 = 1 &\Rightarrow \\(x, y, z) = \left(\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right) \text{ oder } (x, y, z) = \left(-\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right)\end{aligned}$$

Der Wert von f in beiden Punkten ist $-\frac{\sqrt{2}}{2}$. Im zweiten Fall haben wir

$$\begin{aligned}x = \sqrt{2}y, \quad y = z, \quad x^2 + y^2 + z^2 = 1 &\Rightarrow \\(x, y, z) = \left(\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right) \text{ oder } (x, y, z) = \left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right)\end{aligned}$$

Der Wert von f in beiden Punkten ist $\frac{\sqrt{2}}{2}$.

Die Antwort: Das Minimum ist

$$f\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right) = f\left(-\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right) = -\frac{\sqrt{2}}{2}$$

Das Maximum:

$$f\left(\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right) = f\left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}, -\frac{1}{2}\right) = \frac{\sqrt{2}}{2}$$

4 Vektorwertige Funktionen

4.1 Die Differentiation und die Jacobi-Matrix

Die Grundlagen

Wie im Abschnitt 1.1 bereits erwähnt, kann eine vektorwertige Funktion $f: D \rightarrow \mathbb{R}^m$ auf $D \subset \mathbb{R}^n$ als ein m -Tupel reellwertiger Funktionen gesehen werden, ihrer *Komponentenfunktionen*:

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

Mittels Komponentenfunktionen können die Grundbegriffe für reellwertige Funktionen auf die vektorwertigen übertragen werden. Und zwar, der Grenzwert, die partielle Differentiation und die "klein-o"-Notation werden komponentenweise erklärt:

$$\lim_{x \rightarrow x_0} f(x) := \begin{pmatrix} \lim_{x \rightarrow x_0} f_1(x) \\ \lim_{x \rightarrow x_0} f_2(x) \\ \dots \\ \lim_{x \rightarrow x_0} f_m(x) \end{pmatrix} \quad \frac{\partial f}{\partial x_i}(x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(x) \\ \frac{\partial f_2}{\partial x_i}(x) \\ \dots \\ \frac{\partial f_m}{\partial x_i}(x) \end{pmatrix}$$

$$f(x) = o(|x - x_0|) \Leftrightarrow f_k(x) = o(|x - x_0|) \text{ für alle } k$$

Definition 4.1. a) Eine Funktion $f: D \rightarrow \mathbb{R}^m$ heißt eine C^r -Funktion genau dann, wenn alle ihre Komponentenfunktionen C^r -Funktionen sind. Insbesondere (der Fall $r = 0$) heißt f stetig genau dann, wenn für jedes k die Funktion f_k stetig ist.

b) f heißt in $x_0 \in D$ total differenzierbar oder linear approximierbar, wenn es eine $m \times n$ -Matrix A gibt, sodass

$$f(x) = f(x_0) + A(x - x_0) + o(|x - x_0|) \quad (2.9)$$

bei $x \rightarrow x_0$ gilt.

Die Jacobi-Matrix

Damit eine vektorwertige Funktion $f: D \rightarrow \mathbb{R}^m$ linear approximierbar wäre, sollen alle ihre Komponenten linear approximierbar sein:

$$f_k(x) = f_k(x_0) + \text{grad } f_k(x_0) \cdot (x - x_0) + o(|x - x_0|), \quad k = 1, \dots, m$$

Diese m Gleichungen ergeben zusammen die Gleichung (2.9). D. h. die k -te Zeile der Matrix A ist $\text{grad } f_k(x_0)$.

Definition 4.2. Die Matrix

$$J_f(x_0) := \begin{pmatrix} \text{grad } f_1(x_0)^\top \\ \text{grad } f_2(x_0)^\top \\ \dots \\ \text{grad } f_m(x_0)^\top \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \dots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}$$

heißt die Jacobi-Matrix (oder auch Funktionalmatrix) von f in x_0 .

Dementsprechend gilt in (2.9) $A = J_f(x_0)$.

Die j -te Spalte der Jacobi-Matrix ist die partielle Ableitung der vektorwertigen Funktion f nach x_j .

Spezialfälle der Jacobi-Matrix:

- $f: D \rightarrow \mathbb{R}$ ist eine reellwertige Funktion,

$$J_f(x_0) = \text{grad } f(x_0)^\top$$

ein Zeilenvektor.

- $D = I \subset \mathbb{R}$ ein Intervall, $f: I \rightarrow \mathbb{R}^m$ eine Kurve,

$$J_f(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x_0) \\ \frac{\partial f_2}{\partial x}(x_0) \\ \dots \\ \frac{\partial f_m}{\partial x}(x_0) \end{pmatrix}$$

der Tangentialvektor zur Kurve im Punkt $f(x_0)$.

Der Fall $m = n$: die Determinante der Jacobi-Matrix

Die Formel (2.9) besagt, dass die Abbildung f in der Nähe von x_0 durch die lineare Abbildung mit Matrix $J_f(x_0)$ (zusammen mit der Verschiebung um den Vektor $f(x_0)$) approximiert werden kann. Deswegen übertragen sich die Eigenschaften der Matrix auf die Abbildung:

- Gilt $\det J_f(x_0) \neq 0$, so ist die Abbildung f in der Nähe von x_0 umkehrbar.
- Die Zahl $|\det J_f(x_0)|$ ist der Volumenverzerrungsfaktor der Abbildung f "im Kleinen" am Punkt x_0 .

4.2 Die Kettenregel

Auf der Sprache der Jacobi-Matrizen hat die Kettenregel der Differentiation eine sehr kompakte Formulierung.

Satz 45. *Bei der Verkettung von vektorwertigen Funktionen multiplizieren sich ihre Jacobi-Matrizen:*

$$J_{g \circ f}(x_0) = J_g(f(x_0))J_f(x_0) \quad (2.10)$$

(Hier sind $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ total differenzierbare Abbildungen.)

Für die Komponenten der Abbildung ausgeschrieben, wird die Formel (2.10) zu

$$\frac{\partial z_i}{\partial x_k} = \sum_{j=1}^m \frac{\partial z_i}{\partial y_j} \frac{\partial y_j}{\partial x_k}$$

wobei wir $y = f(x)$ und $z = g(y)$ setzten. Vergleiche das mit den Formeln im Abschnitt 2.4.

Korollar 12. *Sei $D \subset \mathbb{R}^n$ und sei $f: D \rightarrow \mathbb{R}^n$ eine C^1 -Abbildung. Ist $\det J_f(x_0) \neq 0$, so ist f in einer Umgebung von x_0 invertierbar, und es gilt*

$$J_{f^{-1}}(f(x_0)) = (J_f(x_0))^{-1}$$

Beweis. Die Jacobi-Matrix der Identitätsabbildung ist die Einheitsmatrix E . Deswegen nach dem Kettenregel

$$E = J_{f^{-1}}(f(x_0))J_f(x_0)$$

Daraus folgt die Behauptung. □

Beispiel 4.3. Der Übergang von den Polarkoordinaten zu den kartesischen erfolgt durch die Abbildung

$$\begin{pmatrix} x \\ y \end{pmatrix} = f \begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

Die Jacobi-Matrix von f :

$$J_f = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

Es gilt $\det J_f = r$. Also, bei $r > 0$ gilt $\det J_f \neq 0$, und die Abbildung f ist lokal invertierbar.

Wir wollen nun die Jacobi-Matrix der inversen Abbildung berechnen. Statt die Abbildung zu invertieren ($r = \sqrt{x^2 + y^2}, \dots$) und dann die partiellen Ableitungen zu berechnen, können wir das obige Korollar anwenden:

$$\begin{aligned} J_{f^{-1}} &= \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}^{-1} = \\ &= \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} \end{aligned}$$

4.3 Skalaren- und Vektorfelder

Eine reellwertige Funktion $f: D \rightarrow \mathbb{R}$ wird auch *Skalarenfeld* genannt. Eine vektorwertige Funktion $v: R^n \rightarrow \mathbb{R}^n$ (oder $v: D \rightarrow \mathbb{R}^n$ für $D \subset \mathbb{R}^n$) wird ein *Vektorfeld* genannt. Man beachte, dass hier die Dimension des Definitionsbereichs und des Wertebereichs gleich sein sollen.

Graphisch wird ein Vektorfeld als eine Gesamtheit von "punktgebundener" Vektoren dargestellt: in jedem Punkt x des Bereichs $D \subset \mathbb{R}^n$ wird ein Vektor $v(x) \in \mathbb{R}^n$ angeheftet.

Als Skalaren- oder Vektorfelder werden viele physikalische Begriffe dargestellt: die Temperatur- oder Druckverteilung ist ein Skalarenfeld; das Kraftfeld, das magnetische Feld sind Vektorfelder.

Beispiel 4.4. Sei $c > 0$. Das Gravitationsfeld eines (im Nullpunkt liegenden) Massenpunktes

$$v(x) = -\frac{c}{|x|^3}x \tag{2.11}$$

Der Betrag der Kraft $|v(x)| = \frac{c}{|x|^3}|x| = \frac{c}{|x|^2}$ ist umgekehrt proportional zum Quadrat des Abstandes vom Nullpunkt. Die Kraft ist zum Nullpunkt gerichtet.

Gradientenfeld und Potential

Sei $f: D \rightarrow \mathbb{R}$ eine reellwertige Funktion (Skalarenfeld). Der Gradient von f kann als Vektorfeld auf D betrachtet werden: an jedem Punkt $x \in D$ ist $\text{grad } f$ angeheftet, und der zeigt die Richtung und die Größe des stärksten Anstiegs der Funktion f an.

Definition 4.5. Ein Vektorfeld $v: D \rightarrow \mathbb{R}^n$ heißt Gradientenfeld, wenn es eine Funktion $\Phi: D \rightarrow \mathbb{R}$ gibt, sodass $v = \text{grad } \Phi$ gilt. Die Funktion Φ wird das Potential des Feldes v genannt.

Beispiel 4.6. Das Gravitationsfeld (2.11) hat das Potential

$$\Phi(x) = \frac{c}{|x|}$$

In der Tat, wegen $|x| = \sqrt{\sum_{i=1}^n x_i^2}$ gilt

$$\frac{\partial |x|}{\partial x_i} = \frac{2x_i}{2|x|} = \frac{x_i}{|x|} \quad (2.12)$$

Daraus folgt

$$\frac{\partial \Phi}{\partial x_i} = -\frac{c}{|x|^2} \frac{x_i}{|x|} = -\frac{c}{|x|^3} x_i,$$

mit anderen Worten, $\text{grad } \Phi(x) = v(x)$.

Feldlinien

Definition 4.7. Eine Kurve in $D \subset \mathbb{R}^n$ heißt Feldlinie des Vektorfeldes $v: D \rightarrow \mathbb{R}^n$, wenn für jeden Kurvenpunkt x die Tangente zur Kurve in x den Vektor $v(x)$ enthält.

Durch jeden Punkt $x \in D$ geht eine Feldlinie des Feldes v . Falls $v(x) = 0$, dann entartet die Feldlinie in einen Punkt.

Beispiel 4.8. Die Feldlinien des Feldes (2.11) sind offene vom Nullpunkt ausgehende Halbgeraden, sowie der Nullpunkt selbst.

Beispiel 4.9. Sei $v(x, y) = (-y, x)$. Das ist das Feld einer infinitesimalen Drehung um den Nullpunkt. Die Feldlinien sind Kreise mit Zentrum im Nullpunkt.

Beispiel 4.10. Sei $w \in \mathbb{R}^3$ ein fester vom Null verschiedener Vektor. Betrachte das Vektorfeld

$$v(x) = w \times x$$

Der Vektor $v(x)$ ist orthogonal zu w , also liegt $v(x)$ in einer zur Achse $\mathbb{R}w$ senkrechten Ebene. Die Feldlinien sind in solchen Ebenen enthalten; es sind Kreise mit Zentrum auf der Achse. Das Feld stellt eine infinitesimale Drehung um die Achse $\mathbb{R}w$ dar.

Wir wollen nun zeigen, dass nicht jedes Vektorfeld ein Gradientenfeld ist. Dafür beweisen wir

Lemma 33. Eine abgeschlossene Kurve kann nicht Feldlinie eines Gradientenfeldes sein.

Beweis. Angenommen, Φ ist das Potential eines Feldes v , und das Feld v hat eine abgeschlossene Feldlinie. Bewegen wir uns entlang dieser Feldlinie, so steigt (oder fällt) das Potential streng monoton, denn wir bewegen uns immer in der Richtung (oder Gegenrichtung) des Gradienten von Φ . Ist die Feldlinie abgeschlossen, so kommen wir irgendwann zurück zum Ausgangspunkt x und erhalten somit die Ungleichung $\Phi(x) > \Phi(x)$. Dieser Widerspruch beweist die Behauptung des Lemmas. \square

Es folgt also, dass die Felder aus den letzten zwei Beispielen keine Gradientenfelder sind.

4.4 Divergenz, Rotation, Laplace-Operator

Die Divergenz

Während der Gradient einem Skalarfeld f ein Vektorfeld $\text{grad } f$ zuordnet, ordnet der Divergenzoperator einem Vektorfeld ein Skalarfeld.

Definition 4.11. Sei $v: D \rightarrow \mathbb{R}^n$ ein Vektorfeld auf $D \subset \mathbb{R}^n$. Die Divergenz von v wird definiert als

$$\text{div } v: D \rightarrow \mathbb{R}, \quad \text{div } v(x) = \sum_{i=1}^n \frac{\partial v_i}{\partial x_i}$$

Beispiel 4.12. • Die Divergenz des Vektorfeldes

$$v(x, y) = (-y, x)$$

ist gleich 0.

• Die Divergenz des Vektorfeldes in \mathbb{R}^3

$$v(x) = w \times x = \begin{pmatrix} w_2 x_3 - w_3 x_2 \\ w_3 x_1 - w_1 x_3 \\ w_1 x_2 - w_2 x_1 \end{pmatrix}$$

ist gleich 0, weil $\frac{\partial v_i}{\partial x_i} = 0$ für alle i .

Beispiel 4.13. Berechnen wir die Divergenz des Gravitationsfeldes aus Beispiel 4.4. Es gilt

$$v_i = -\frac{c}{|x|^3} x_i$$

Um $\frac{\partial v_i}{\partial x_i}$ zu berechnen, benutzen wir (2.12).

$$\frac{\partial v_i}{\partial x_i} = -\frac{-3c x_i}{|x|^4 |x|} x_i - \frac{c}{|x|^3} = 3 \frac{c}{|x|^5} x_i^2 - \frac{c}{|x|^3}$$

Folglich

$$\text{div } v = 3 \frac{c}{|x|^5} \sum_{i=1}^3 x_i^2 - 3 \frac{c}{|x|^3} = 3 \frac{c}{|x|^5} |x|^2 - 3 \frac{c}{|x|^3} = 0$$

Man beachte, dass das Gravitationsfeld im Nullpunkt nicht definiert ist, deswegen ist $\text{div } v(0)$ nicht definiert.

Die physikalische Deutung der Divergenz ist die "Quelldichte" des Vektorfeldes. Ist $\operatorname{div} v(x) > 0$, so quillt das Fluid aus dem Punkt x , ist $\operatorname{div} v(x) < 0$, so gibt es im Punkt x eine "Senke". Vektorfelder mit $\operatorname{div} v = 0$ heißen *quellfrei* und können die Bewegung eines inkompressiblen Fluids darstellen.

Beispiel 4.14. Für das Vektorfeld $v(x, y) = (x, y)$ gilt $\operatorname{div} v = 2$. Hier quillt es aus jedem Punkt, mit gleicher Stärke.

Die Rotation

Die Rotation eines Vektorfeldes in \mathbb{R}^2 ist ein Skalarfeld; die Rotation eines Vektorfeldes in \mathbb{R}^3 ist wieder ein Vektorfeld.

Definition 4.15. Die Rotation eines ebenen Vektorfeldes $v: D \rightarrow \mathbb{R}^2$, $v(x_1, x_2) = (v_1, v_2)$ wird definiert als

$$\operatorname{rot} v: D \rightarrow \mathbb{R}, \quad \operatorname{rot} v(x) = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}$$

Die Rotation eines räumlichen Vektorfeldes $v: D \rightarrow \mathbb{R}^3$, $v(x_1, x_2, x_3) = (v_1, v_2, v_3)$ wird definiert als

$$\operatorname{rot} v: D \rightarrow \mathbb{R}^3, \quad \operatorname{rot} v(x) = \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}$$

Die Faustregeln für die Formeln für $\operatorname{div} v$ und $\operatorname{rot} v$ sind

$$\operatorname{div} v = \nabla \cdot v, \quad \operatorname{rot} v = \nabla \times v$$

Hier steht ∇ für den symbolischen Spaltenvektor

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix}$$

Beispiel 4.16. • Die Rotation des Vektorfeldes $v(x, y) = (-y, x)$ ist gleich

$$\operatorname{rot} v = \frac{\partial}{\partial x} x - \frac{\partial}{\partial y} (-y) = 2$$

• Die Rotation des räumlichen Vektorfeldes $v(x) = w \times x$ ist gleich

$$\operatorname{rot} v = \operatorname{rot} \begin{pmatrix} w_2 x_3 - w_3 x_2 \\ w_3 x_1 - w_1 x_3 \\ w_1 x_2 - w_2 x_1 \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} w_2 x_3 - w_3 x_2 \\ w_3 x_1 - w_1 x_3 \\ w_1 x_2 - w_2 x_1 \end{pmatrix} = 2 \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = 2w$$

Die physikalische Deutung der Rotation ist die *Wirbeldichte*. Das kleine Stück des Fluids um den Punkt x wird mit dem Vektorfeld um den Vektor $v(x)$ verschoben; außerdem wird es in sich selbst gedreht. Die Zahl $\operatorname{rot} v$ im ebenen Fall und der Vektor $\operatorname{rot} v$ im räumlichen Fall zeigen die Drehgeschwindigkeit, bzw. auch die Drehachse an.

Man sieht es gut auf den obigen zwei Beispielen. Das erste Vektorfeld ist das Feld einer starren Drehung in positiver Richtung. Daher ist die Rotation in jedem Punkt positiv und konstant. Das zweite Vektorfeld ist das Feld der Drehung um die Achse $\mathbb{R}w$ mit Winkelgeschwindigkeit $|w|$. Dementsprechend zeigt $\operatorname{rot} v = 2w$ dass jedes kleine Stück des Fluids sich um dieselbe Achse mit dergleichen Geschwindigkeit dreht.

Ein Vektorfeld mit $\operatorname{rot} v = 0$ heißt *wirbelfrei*.

Lemma 34. Für jeden C^2 -Skalarenfeld $f: D \rightarrow \mathbb{R}^2$ oder $D \rightarrow \mathbb{R}^3$ gilt

$$\operatorname{rot} \operatorname{grad} f = 0$$

Beweis. Für $v = \operatorname{grad} f$ gilt $v_i = \frac{\partial f}{\partial x_i}$. Daher $\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}$ und folglich $\operatorname{rot} v = 0$. \square

Beispiel 4.17. Das Gravitationsfeld aus Beispiel 4.4 ist wirbelfrei, weil es ein Gradientenfeld ist.

Korollar 13. Ein nichtwirbelfreies Vektorfeld kann kein Gradientenfeld sein.

Der Laplace-Operator

Definition 4.18. Sei $f: D \rightarrow \mathbb{R}$ eine C^2 -Funktion. Der Laplace-Operator Δ ordnet f eine neue Funktion $\Delta f: D \rightarrow \mathbb{R}$ zu, mit den Werten

$$\Delta f(x) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$$

Lemma 35. Es gilt $\Delta f = \operatorname{div} \operatorname{grad} f$.

Beweis. Direkt aus den Definitionen. \square

Eine Funktion f mit $\Delta f(x) = 0$ für alle x heißt *harmonisch*. Die (mit der Zeit sich nicht ändernde) Temperaturverteilung wird durch eine harmonische Funktion beschrieben.

Beispiel 4.19. • Die Funktion $f(x, y) = x^2 - y^2$ ist harmonisch.

- Die Funktion $\Phi(x) = \frac{1}{|x|}$ ist harmonisch. In der Tat, es ist das Potential des Gravitationsfeldes eines Massenpunktes, und wir haben gezeigt, dass dieses Feld divergenzfrei ist. Also gilt $\Delta \Phi = \operatorname{div} \operatorname{grad} \Phi = 0$.

Die Funktion $\Phi(x) = \frac{1}{|x|}$ beschreibt die von einem “unendlich heißen Punkt” erzeugte Temperaturverteilung.

Die harmonischen Funktionen haben keine strikten lokalen Extrema. Man sieht es z. B. daran, dass Δf die Spur der Hesse-Matrix ist, und bei einer positiv definiten Matrix kann die Spur nicht verschwinden.

Kapitel 3

Funktionen in mehreren Variablen: Integration

1 Parameterintegrale

Das bestimmte Integral von a bis b ordnet jeder Funktion $f: [a, b] \rightarrow \mathbb{R}$ eine Zahl zu:

$$\int_a^b f(x) dx$$

Wir werden jetzt zwei Situationen betrachten, wo das Integral von weiteren Parametern abhängt:

1. Die Funktion f hängt nicht von einer, sondern von mehreren Variablen ab. Z. B. beim Integrieren einer Funktion $f(x, y)$ bezüglich y entsteht eine Funktion von x :

$$F(x) := \int_c^d f(x, y) dy$$

Beispiel 1.1.

$$F(x) = \int_0^1 \cos(xy) dy$$

Dieses Integral lässt sich berechnen, sodass wir eine explizite Darstellung der Funktion F erhalten. Wird x als eine Konstante wahrgenommen, so gilt

$$\int \cos(xy) dy = \int \frac{1}{x} \cos(xy) d(xy) = \frac{1}{x} \sin(xy)$$

Folglich

$$\int_0^1 \cos(xy) dy = \left|_{y=0}^{y=1} \frac{\sin(xy)}{x} \right| = \frac{\sin x}{x}$$

2. Die Integrationsgrenzen sind Variabel:

$$F(x) := \int_c^x f(y) dy \quad \text{oder, allgemeiner} \quad F(x) := \int_c^{h(x)} f(y) dy$$

In beiden Fällen wollen wir die Ableitung der Funktion F berechnen.

Fall 1: Ausintegrieren einer von mehreren Variablen

Satz 46. Sei $f: D \rightarrow \mathbb{R}$ eine stetige Funktion auf dem Rechteck $D = [a, b] \times [c, d]$. Dann gilt für die Funktion

$$F(x) := \int_c^d f(x, y) dy, \quad F: [a, b] \rightarrow \mathbb{R}$$

a) F ist in $[a, b]$ stetig;

b) existiert die partielle Ableitung $\frac{\partial f}{\partial x}$ in D und ist dort stetig, dann ist F differenzierbar mit der Ableitung

$$F'(x) = \frac{d}{dx} \int_c^d f(x, y) dy = \int_c^d \frac{\partial f}{\partial x}(x, y) dy$$

Beispiel 1.2. Sei $F(x) = \int_0^1 e^{xy^2} dy$. Dann gilt

$$F'(x) = \int_0^1 \frac{\partial}{\partial x} e^{xy^2} dy = \int_0^1 e^{xy^2} y^2 dy$$

Fall 2: Variable Integrationsgrenzen

Im Fall

$$F(x) = \int_a^x f(y) dy$$

gilt $F'(x) = f(x)$. Das ist ein Lemma zum Hauptsatz der Differential- und Integralrechnung. Analog gilt

$$\frac{d}{dx} \int_x^b f(y) dy = -f(x)$$

Wenn die Integrationsgrenzen Funktionen von x sind, dann gilt

Satz 47. Sei $f: [c, d] \rightarrow \mathbb{R}$ eine stetige Funktion, und seien g, h differenzierbare Funktionen mit Werten in $[c, d]$. Dann ist die Funktion

$$F(x) = \int_{g(x)}^{h(x)} f(y) dy$$

differenzierbar, und es gilt

$$F'(x) = \frac{d}{dx} \int_{g(x)}^{h(x)} f(y) dy = f(h(x))h'(x) - f(g(x))g'(x)$$

Beweis. Führen wir neue Variablen ein:

$$u = g(x), \quad v = h(x)$$

sodass $F(x) = \int_u^v f(y) dy$ gilt. Nach der Kettenregel gilt

$$\frac{dF}{dx} = \frac{\partial F}{\partial u} \frac{du}{dx} + \frac{\partial F}{\partial v} \frac{dv}{dx} = -f(u) \frac{du}{dx} + f(v) \frac{dv}{dx} = f(h(x))h'(x) - f(g(x))g'(x)$$

□

Kombination beider Fälle

Satz 48.

$$\frac{d}{dx} \int_{g(x)}^{h(x)} f(x, y) dy = \int_{g(x)}^{h(x)} f_x(x, y) dy + f(x, h(x))h'(x) - f(x, g(x))g'(x)$$

Beispiel 1.3. Für $I(\alpha) = \int_0^{\alpha^2} \cos(\alpha t^2) dt$ gilt

$$\frac{dI}{d\alpha} = - \int_0^{\alpha^2} t^2 \sin(\alpha t^2) dt + 2\alpha \cos(\alpha^5)$$

2 Integration über ebene Bereiche

2.1 Der Flächeninhalt

Wie misst man den Flächeninhalt einer ebenen Figur? Die folgende Definition basiert auf einem praktischen Verfahren zur approximativen Berechnung der Fläche.

Wähle eine natürliche Zahl k und betrachte in der Ebene das Quadratgitter der Kantenlänge $\frac{1}{k}$. Sei $s_k(M)$ der Flächeninhalt aller Quadrate, die ganz in M liegen, und sei $S_k(M)$ der Flächeninhalt aller Quadrate, die mindestens einen Punkt von M enthalten. Wenn der Flächeninhalt von M überhaupt Sinn macht, dann soll es gelten

$$s_k(M) \leq F(M) \leq S_k(M)$$

Außerdem, bei genügend großen k sollen beide Zahlen $s_k(M)$ und $S_k(M)$ beliebig nah an $F(M)$ liegen. Das führt zur folgenden Definition.

Definition 2.1. Eine Menge $M \subset \mathbb{R}^2$ heißt (Riemann)-meßbar, wenn die beiden Grenzwerte

$$\lim_{k \rightarrow \infty} s_k(M) \quad \text{und} \quad \lim_{k \rightarrow \infty} S_k(M)$$

existieren und zueinander gleich sind. Ihr gemeinsame Wert wird dann der Flächeninhalt von M genannt und mit $F(M)$ bezeichnet.

Es gibt Mengen (mit einem sehr "zerklüfteten" Rand), die keinen Flächeninhalt besitzen. Im Folgenden werden wir nur "gute" Mengen betrachten, die allerdings eine genügend große Klasse bilden.

Definition 2.2. Eine Teilmenge $B \subset \mathbb{R}^2$ heißt regulär, wenn

- B beschränkt und abgeschlossen ist (also kompakt);
- Der Rand ∂B aus endlich vielen regulären Kurvenstücken besteht.

Zur Erinnerung: eine Kurve $\gamma: [a, b] \rightarrow \mathbb{R}^2$ heißt regulär, wenn γ differenzierbar ist und der Geschwindigkeitsvektor $\dot{\gamma}$ nirgendwo verschwindet. Eine reguläre Kurve besitzt in jedem Punkt eine Tangente.

Da der Rand ∂B eines regulären Bereichs aus mehreren regulären Kurvenstücken bestehen darf, sind auf dem Rand Knickpunkte und Spitzen erlaubt. Ein Rechteck und das von der Astroide beschränkte Bereich sind also reguläre Bereiche.

Satz 49. Jeder reguläre Bereich ist meßbar.

2.2 Definition und Berechnung des Doppelintegrals

Definition

Sei B ein regulärer Bereich, und sei $f: B \rightarrow \mathbb{R}$ eine stetige Funktion. Um das Integral von f über B zu definieren,

- teilen wir B in kleinere Bereiche B_i auf (z. B. mittels Quadratgitter mit Kantenlänge $\frac{1}{k}$ wie bei der Definition des Flächeninhaltes),
- wählen in jedem Teilbereich B_i einen Punkt (x_i, y_i) ,
- und bilden die Riemann-Summe

$$Z_k = \sum_i f(x_i, y_i) F(B_i) \quad (3.1)$$

Der Index k in (3.1) bezieht sich auf die Kantenlänge der Gitter: je größer k , desto feiner die Unterteilung von B . Man kann zeigen, dass aufgrund der Stetigkeit von f jede auf diese Weise gebildete Folge (unabhängig von der Wahl der Punkte (x_i, y_i)) gegen dieselbe Zahl konvergiert. Diese Zahl ist das Integral von f über B .

Es gilt

$$F(M) = \int_M dF$$

d. h. der Flächeninhalt von M ist gleich dem Integral der konstanten Funktion $f(x, y) = 1$.

Berechnung des Doppelintegrals

Der vorige Abschnitt gibt eine intuitive Vorstellung vom Doppelintegral. Ist z. B. $\rho(x, y)$ die Dichte einer ebenen Platte B , so wird mit $\int_B \rho dF$ die Masse dieser Platte berechnet: in der Riemann-Summe (3.1) gibt jeder Summand die approximative Masse des Stückes B_i .

Die Riemann-Summe wird auch in der Praxis benutzt, um das Integral numerisch zu berechnen/approximieren. Für eine genaue Berechnung des Doppelintegrals wird es auf die zweifache Anwendung der üblichen Integration (über Intervalle) zurückgeführt. Das wird zunächst für spezielle Bereiche gemacht.

Definition 2.3. Ein regulärer Bereich $B_1 \subset \mathbb{R}^2$ heißt Normalbereich vom Typ I, wenn es $a, b \in \mathbb{R}$ und stetige Funktionen $g, h: [a, b] \rightarrow \mathbb{R}$ gibt mit $g(x) \leq h(x)$ und

$$B_1 = \{(x, y) \mid a \leq x \leq b, g(x) \leq y \leq h(x)\}$$

Analog, ein Normalbereich vom Typ II hat die Form

$$B_2 = \{(x, y) \mid l(y) \leq x \leq r(y), c \leq y \leq d\}$$

mit $c, d \in \mathbb{R}$ und stetigen Funktionen $l, r: [c, d] \rightarrow \mathbb{R}$.

Ein Normalbereich vom Typ I schneidet jede senkrechte Gerade in einem Intervall (oder leerer Menge); ein Normalbereich vom Typ II schneidet jede waagerechte Gerade in einem Intervall (oder leerer Menge).

Ein Normalbereich B_1 vom Typ I wird hiermit als Vereinigung von senkrechten Intervallen dargestellt. Um $\int_{B_1} f dF$ zu berechnen, integrieren wir f zuerst entlang jedes senkrechten Intervalls, was eine Funktion auf dem Intervall $[a, b]$ ergibt:

$$I(x) := \int_{g(x)}^{h(x)} f(x, y) dy$$

und dann integrieren wir I von a bis b :

$$\int_{B_1} f dF = \int_a^b I(x) dx$$

Satz 50. Für jede stetige Funktion $f: B_1 \rightarrow \mathbb{R}$ auf einem Normalbereich B_1 vom Typ I gilt:

$$\int_{B_1} f(x, y) dx dy = \int_a^b \left(\int_{g(x)}^{h(x)} f(x, y) dy \right) dx$$

Für jede stetige Funktion $f: B_1 \rightarrow \mathbb{R}$ auf einem Normalbereich B_2 vom Typ II gilt:

$$\int_{B_2} f(x, y) dx dy = \int_c^d \left(\int_{l(y)}^{r(y)} f(x, y) dx \right) dy$$

Ist B gleichzeitig ein Normalbereich vom Typ I und ein Normalbereich vom Typ II, so gelten beide Formeln. Insbesondere, für ein Rechteck

$$R = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$$

haben wir

$$\int_R f dF = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

Die letzten zwei Gleichungen werden auch *Satz von Fubini* genannt.

Ist ein regulärer Bereich B kein Normalbereich, so kann B in Normalbereiche B_1, B_2, \dots, B_n aufgeteilt, und das Integral mittels

$$\int_B f dF = \int_{B_1} f dF + \dots + \int_{B_n} f dF$$

berechnet werden.

Bemerkung 2.4. Im Spezialfall $f = 1$ wird die Formel für den Integral von f über einen Normalbereich zu uns aus dem ersten Semester bekannter Formel für den Flächeninhalt zwischen den Graphen der Funktionen g und h :

$$F = \int_B dF = \int_a^b \left(\int_{g(x)}^{h(x)} dy \right) dx = \int_a^b (h(x) - g(x)) dx$$

Beispiel 2.5. Berechnen wir den Flächeninhalt des Bereichs B zwischen den Kurven $y = x$, $xy = 1$ und $y = 2$.

Ein Blick auf die Graphen zeigt, dass B sowohl ein Normalbereich vom Typ I als auch vom Typ II ist. Benutzen wir das Verfahren für Typ II. B kann dargestellt werden als

$$B = \{(x, y) \mid 1 \leq y \leq 2, \frac{1}{y} \leq x \leq y\}$$

Also gilt

$$F(B) = \int_B dF = \int_1^2 \left(\int_{\frac{1}{y}}^y dx \right) dy = \int_1^2 \left(y - \frac{1}{y} \right) dy = \left[\frac{y^2}{2} - \ln y \right]_1^2 = \frac{3}{2} - \ln 2$$

Wird B als Normalbereich vom Typ II angesehen, so sieht die Rechnung so aus:

$$F(B) = \int_{\frac{1}{2}}^1 \left(\int_{\frac{1}{x}}^2 dy \right) dx + \int_1^2 \left(\int_x^2 dy \right) dx = \dots$$

2.3 Anwendungen des Doppelintegrals

Volumen unter dem Graphen

Sei $f: B \rightarrow \mathbb{R}$ eine positive Funktion auf einem Normalbereich. Dann ist

$$V = \int_B f dF$$

gleich dem Volumen des Körpers

$$K = \{(x, y, z) \mid (x, y) \in B, 0 \leq z \leq f(x, y)\}$$

Beispiel 2.6. Berechnen wir das Volumen des auf dem Bereich B aus dem Beispiel 2.5 stehenden Zylinderabschnittes mit Deckfläche $z = \frac{y^2}{x^2}$.

$$V = \int_B \frac{y^2}{x^2} dF = \int_1^2 \left(\int_{\frac{1}{y}}^y \frac{y^2}{x^2} dx \right) dy = \int_1^2 \left[-\frac{y^2}{x} \right]_{x=\frac{1}{y}}^{x=y} dy = \int_1^2 (-y + y^3) dy = \frac{9}{4}$$

Der Massenmittelpunkt und der geometrische Schwerpunkt

Die Masse einer Platte B mit Dichteverteilung ρ wird berechnet als

$$M = \int_B \rho dF$$

Der Massenmittelpunkt entsteht durch Berechnung des Durchschnitts der x - bzw. der y -Koordinaten aller Punkte. (Siehe hierzu auch die Motivation durch den Hebelgesetz im ersten Semester.) Die Koordinaten werden mit Funktion ρ gewichtet:

$$x_s = \frac{1}{M} \int_B x \rho(x, y) dF, \quad y_s = \frac{1}{M} \int_B y \rho(x, y) dF$$

sind die Koordinaten des Massenmittelpunktes von B mit Dichteverteilung ρ .

Ist die Platte homogen ($\rho = \text{const}$), so wird der Massenmittelpunkt als der *geometrische Schwerpunkt* bezeichnet. Für ihn erhält man die Koordinaten

$$\bar{x} = \frac{1}{F} \int_B x \, dF, \quad \bar{y} = \frac{1}{F} \int_B y \, dF$$

Beispiel 2.7. Sei B ein Normalbereich vom Typ I:

$$B = \{(x, y) \mid a \leq x \leq b, g(x) \leq y \leq h(x)\}$$

Dann erhält man

$$\begin{aligned} \bar{x} &= \frac{1}{F} \int_a^b \left(\int_{g(x)}^{h(x)} x \, dy \right) dx = \frac{1}{F} \int_a^b x(h(x) - g(x)) \, dx \\ \bar{y} &= \frac{1}{F} \int_a^b \left(\int_{g(x)}^{h(x)} y \, dy \right) dx = \frac{1}{F} \int_a^b \frac{h(x)^2 - g(x)^2}{2} \, dx \end{aligned}$$

Das Trägheitsmoment

Das Trägheitsmoment beschreibt den Widerstand eines starren Körpers gegenüber einer Änderung seiner Rotationsbewegung. Das Trägheitsmoment bezieht sich auf eine Achse (oder auf einen Punkt, wenn wir die Drehung einer Figur in der Ebene betrachten). Je größer das Trägheitsmoment bezüglich einer Achse, desto schwieriger ist es, den Körper zum Drehen um diese Achse zu bringen.

Das Trägheitsmoment einer homogenen Platte $B \subset \mathbb{R}^2$ bezüglich des Ursprungs ist gleich

$$J(0, 0) := \int_B (x^2 + y^2) \, dF$$

Das Trägheitsmoment derselben Platte bezüglich eines Punktes (a, b) ist

$$J(a, b) := \int_B ((x - a)^2 + (y - b)^2) \, dF$$

Satz 51. *Der kleinste Wert des Trägheitsmomentes $J(a, b)$ wird erreicht, wenn (a, b) der geometrische Schwerpunkt der Platte B ist.*

Beweis. Bei $(a, b) \rightarrow \infty$ geht $J(a, b)$ gegen $+\infty$. Deswegen nimmt die Funktion $J: \mathbb{R}^2 \rightarrow \mathbb{R}$ ihren minimalen Wert an, und zwar an einem ihrer stationären Punkte.

Um die stationären Punkte zu bestimmen, differenzieren wir $J(a, b)$:

$$\begin{aligned} \frac{\partial J}{\partial a} &= \frac{\partial}{\partial a} \int_B ((x - a)^2 + (y - b)^2) \, dF = \int_B \frac{\partial}{\partial a} ((x - a)^2 + (y - b)^2) \, dF = \\ &= \int_B 2(a - x) \, dF = 2(aF - \int_B x \, dF) \end{aligned}$$

Analog,

$$\frac{\partial J}{\partial b} = 2(bF - \int_B y \, dF)$$

Es gibt also einen einzigen stationären Punkt, und seine Koordinaten sind

$$a = \frac{1}{F} \int_B x \, dF, \quad b = \frac{1}{F} \int_B y \, dF$$

Das sind aber genau die Koordinaten des geometrischen Schwerpunktes. \square

Die physikalische Bedeutung dieses Satzes (im räumlichen Fall): wird einem starren Körper ein Drehimpuls erteilt (durch Anwendung eines Kräftepaars, z. B.), so beginnt der Körper um eine Achse zu drehen, die durch seinen Massenmittelpunkt geht. (Nach dem Prinzip des kleinsten Widerstandes wählt der Körper von allen parallelen Achsen diejenige, um welche er am leichtesten drehen kann.)

2.4 Die Transformationsformel für Gebietsintegrale

Die Formel

$$\int_B f \, dF = \int_a^b \left(\int_{g(x)}^{h(x)} f(x, y) \, dy \right) dx$$

ermöglicht die Berechnung eines Gebietsintegrals durch wiederholte Anwendung von "einfachen" Integralen (hier ist B ein Normalbereich vom Typ I). Manchmal bietet es sich aber an, statt der kartesischen Koordinaten (x, y) die Polarkoordinaten, oder noch irgendeine andere Parametrisierung des Integrationsbereichs zu wählen.

Unter Parameterdarstellung verstehen wir eine Darstellung von kartesischen Koordinaten x und y als Funktionen zweier anderer Variablen:

$$x = x(u, v), \quad y = y(u, v)$$

Hier durchläuft der Punkt (u, v) einen ebenen Bereich A , sodass wir eine Abbildung (Parametrisierung) $A \rightarrow B$ erhalten.

Beispiel 2.8. Sei B ein Einheitskreis mit Zentrum im Nullpunkt, und sei $x = r \cos \varphi$, $y = r \sin \varphi$. Dann variiert r zwischen 0 und 1, und φ zwischen 0 und 2π . Das heißt, der Kreis B wird durch den Rechteck

$$A = \{(r, \varphi) \mid 0 \leq r \leq 1, 0 \leq \varphi \leq 2\pi\}$$

parametrisiert.

Nicht jede Parameterdarstellung ist für die Berechnung des Integrals geeignet.

Definition 2.9. Eine Parameterdarstellung $x = x(u, v)$, $y = y(u, v)$, $(u, v) \in A$ eines ebenen regulären Flächenstücks B heißt Koordinatentransformation, wenn im Inneren von A gilt:

- $x(u, v)$ und $y(u, v)$ stetig partiell differenzierbar sind;
- aus $x(u, v) = x(u', v')$ und $y(u, v) = y(u', v')$ folgt $(u, v) = (u', v')$
- $x_u y_v - x_v y_u \neq 0$

Der Ausdruck $x_u y_v - x_v y_u$ ist nichts anderes als die Determinante der Jacobi-Matrix der Abbildung $A \rightarrow B$. Wir bezeichnen sie in diesem Fall als

$$\frac{\partial(x, y)}{\partial(u, v)} := \det \begin{pmatrix} x_u(u, v) & x_v(u, v) \\ y_u(u, v) & y_v(u, v) \end{pmatrix}$$

Satz 52. *Entsteht der reguläre Bereich $B \subset \mathbb{R}^2$ unter der Koordinatentransformation $x = x(u, v), y = y(u, v)$ aus A , dann gilt für jede stetige Funktion $f: B \rightarrow \mathbb{R}$*

$$\int_B f(x, y) \, dx dy = \int_A f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| \, dudv$$

Für Polarkoordinaten gilt

$$\frac{\partial(x, y)}{\partial(r, \varphi)} := \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r$$

Folglich

$$\int_S f(x, y) \, dx dy = \int_D f(r \cos \varphi, r \sin \varphi) r \, dr d\varphi$$

Beispiel 2.10. Berechnen wir das Volumen der Einheitskugel.

Dieses Volumen V ist das Doppelte des Volumens der Halbkugel

$$\{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1, z \geq 0\}$$

Es gilt also

$$\frac{V}{2} = \int_B \sqrt{1 - x^2 - y^2} \, dF$$

wobei $B = \{(x, y) \mid x^2 + y^2 \leq 1\}$.

Versucht man dieses Integral in kartesischen Koordinaten zu berechnen, so kommt man auf

$$\frac{V}{2} = \int_{-1}^1 \left(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \sqrt{1-x^2-y^2} \, dx \right) dy$$

Die Stammfunktion von $f(x) = \sqrt{a^2 - x^2}$ kann zwar berechnet werden (durch Substitution $x = a \sin t$), einfacher aber ist es, dieses Integral in Polarkoordinaten zu berechnen.

Integrationsbereich in Polarkoordinaten:

$$B = \{(r, \varphi) \mid 0 \leq r \leq 1, 0 \leq \varphi \leq 2\pi\}$$

Die Funktion in Polarkoordinaten:

$$f(x, y) = \sqrt{1 - x^2 - y^2} = \sqrt{1 - r^2}$$

Das Volumen als Integral in Polarkoordinaten:

$$\frac{V}{2} = \int_0^{2\pi} \left(\int_0^1 \sqrt{1 - r^2} r \, dr \right) d\varphi$$

Die innere Stammfunktion:

$$\int \sqrt{1 - r^2} r \, dr = \frac{1}{2} \int \sqrt{1 - r^2} d(r^2) = -\frac{1}{2} \int (1 - r^2)^{\frac{1}{2}} d(1 - r^2) = -\frac{1}{2} \frac{(1 - r^2)^{\frac{3}{2}}}{3/2} = -\frac{1}{3} (1 - r^2)^{\frac{3}{2}}$$

Integralberechnung:

$$\int_0^{2\pi} \left(\int_0^1 \sqrt{1 - r^2} r \, dr \right) d\varphi = \int_0^{2\pi} \left[-\frac{1}{3} (1 - r^2)^{\frac{3}{2}} \right]_0^1 d\varphi = \int_0^{2\pi} \frac{1}{3} d\varphi = \frac{2\pi}{3}$$

Folglich ist das Volumen der Einheitskugel gleich $\frac{4\pi}{3}$.

3 Kurvenintegrale

3.1 Definition

Unter einer (parametrisierten) Kurve in \mathbb{R}^n verstehen wir eine Abbildung

$$w: [a, b] \rightarrow \mathbb{R}^n$$

Eine parametrisierte Kurve ist also die Bahn eines Punktes; dabei wird es vorge-schrieben, wo sich der Punkt zur Zeit t befindet. Die Parametrisierung w heißt glatt, wenn die Abbildung w stetig differenzierbar ist (eine C^1 -Abbildung). In diesem Fall ist der Geschwindigkeitsvektor

$$\dot{w}(t) = \frac{dw}{dt}$$

in jedem Kurvenpunkt definiert.

Bemerkung 3.1. Eine glatt parametrisierte Kurve muss nicht unbedingt glatt aussehen. Z. B. die Abbildung

$$w(t) = (t^3, t^2)$$

ist glatt, die Kurve hat aber eine Spitze im Koordinatenursprung.

Wird es vorausgesetzt, dass der Geschwindigkeitsvektor nirgendwo verschwin-det, dann heißt die Kurve regulär und hat in jedem Punkt eine Tangente (in Richtung des Geschwindigkeitsvektors).

Definition 3.2. Sei $w: [a, b] \rightarrow D \subset \mathbb{R}^n$ eine glatt parametrisierte Kurve und sei $f: D \rightarrow \mathbb{R}$ eine stetige, auf der Kurve definierte, Funktion. Dann heißt

$$\int_w f ds := \int_a^b f(w(t)) |\dot{w}(t)| dt$$

das Kurvenintegral von f längs w .

Das Symbol $ds = |\dot{w}(t)| dt$ heißt Bogenelement und gibt die Länge des wäh-rend des kurzen Zeitabschnittes dt durchgelaufenen Weges (die Geschwindigkeit des Punktes zur Zeit t ist ja gleich $|\dot{w}(t)|$).

Läuft der Punkt $u(t)$ dieselbe Kurve wie der Punkt $w(t)$ durch, so gilt

$$\int_u f ds = \int_w f ds$$

(das Kurvenintegral ist unabhängig von der Parametrisierung).

Um die Länge einer Kurve zu berechnen, muss man die konstante Funktion $f = 1$ längs der Kurve integrieren:

$$L(w) = \int_w ds = \int_a^b |\dot{w}(t)| dt = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} dt$$

Beispiel 3.3. Bestimmen wir die Länge der Schraubenlinie

$$x(t) = 2 \cos t, \quad y(t) = 2 \sin t, \quad z = \frac{1}{2}t$$

Beispiel 3.4. Berechnen wir das Integral der Funktion $f(x, y, z) = x^2 + y^2 + z^2$ längs der Schraubenlinie aus dem letzten Beispiel.

Definition 3.5. Eine stückweise glatte Kurve besteht aus aneinandergereihten glatten Kurven $w_i: [a_i, b_i] \rightarrow \mathbb{R}^n$. Es muss also $w_i(b_i) = w_{i+1}(a_{i+1})$ gelten.

Das Kurvenintegral von f längs w wird definiert als die Summe der Integrale über w_i :

$$\int_w f ds = \int_{w_1} f ds + \dots + \int_{w_n} f ds$$

3.2 Die Integration eines Vektorfeldes längs einer Kurve

Definition 3.6. Sei $w: [a, b] \rightarrow D \subset \mathbb{R}^n$ eine glatt parametrisierte Kurve, und sei $v: D \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Man nennt

$$\int_w v \cdot dX := \int_a^b v(w(t)) \cdot \dot{w}(t) dt$$

das Kurvenintegral von v längs w .

Hat das Vektorfeld v die Komponenten (v_1, v_2, \dots, v_n) , so benutzt man auch die Schreibweise

$$\int_w v \cdot dX = \int_w v_1 dx_1 + v_2 dx_2 + \dots + v_n dx_n$$

In Dimension 3:

$$\int_w v \cdot dX = \int_w v_1 dx + v_2 dy + v_3 dz$$

Für die Berechnung dieses Kurvenintegrals setzt man die Parameterdarstellungen der Koordinaten $x = x(t), y = y(t), z = z(t)$ in die Formel ein. Dabei gilt, wie üblich

$$dx = \dot{x}(t) dt$$

Also gilt

$$\int_w v \cdot dX = \int_a^b (v_1(w(t))\dot{x}(t) + v_2(w(t))\dot{y}(t) + v_3(w(t))\dot{z}(t)) dt$$

Beispiel 3.7. Berechnen wir das Kurvenintegral

$$A = \int_w x^2 y dx + (x - z) dy + (xyz) dz$$

längs des Graphen $y = x^3, x \in [0, 2]$, in der Ebene $z = 2$.

(Hier steht das Integral des Vektorfeldes $(x^2 y, x - z, xyz)$.)

Eine Parametrisierung der Kurve:

$$x(t) = t, y(t) = t^3, z(t) = 2$$

3.3 Das Potential eines Gradientenfeldes

Sei $v: D \rightarrow \mathbb{R}^n$ ein Vektorfeld auf $D \subset \mathbb{R}^n$. Wir wollen herausfinden, ob v ein Gradientenfeld ist, d. h. ob es eine Funktion $f: D \rightarrow \mathbb{R}$ mit $v(x) = \text{grad } f(x)$ gibt, und wenn ja, dann die Funktion f (ein Potential von v) berechnen.

Das Kurvenintegral eines Gradientenfeldes ist gleich der Potentialdifferenz

Satz 53. Ist $v: D \rightarrow \mathbb{R}^n$ ein stetiges Gradientenfeld mit einem Potential f , dann gilt für jede stückweise glatte Kurve $w: [a, b] \rightarrow D$

$$\int_w v \cdot dX = f(w(b)) - f(w(a))$$

Insbesondere hängt das Integral von v längs w nur vom Anfangspunkt und Endpunkt des Weges ab.

Beweis. Falls w glatt ist und $v = \text{grad } f$ gilt, dann

$$\begin{aligned} \int_w v \cdot dX &= \int_w \text{grad } f \cdot dX = \int_a^b \text{grad } f(w(t)) \cdot \dot{w}(t) dt = \\ &= \int_a^b \left(\frac{d}{dt} f(w(t)) \right) dt = f(w(b)) - f(w(a)) \end{aligned}$$

□

Korollar 14. Das Kurvenintegral eines Gradientenfeldes längs einer geschlossenen Kurve ist gleich Null.

Beispiel 3.8. Das Gravitationsfeld

$$v(x) = -\frac{c}{|x|^3}x, \quad x \in \mathbb{R}^3 \setminus \{0\}$$

hat das Potential $f(x) = \frac{c}{|x|}$. Dementsprechend gilt

$$\int_w v \cdot dX = \frac{c}{|w(b)|} - \frac{c}{|w(a)|}$$

Satz 54. Ein Vektorfeld $v: D \rightarrow \mathbb{R}^n$ ist genau dann ein Gradientenfeld, wenn das Kurvenintegral von v wegunabhängig ist. D. h. genau dann, wenn für alle Kurven w in D der Wert $\int_w v \cdot dX$ nur vom Anfangs- und Endpunkt von w abhängt.

Beweis. Wenn v ein Gradientenfeld ist, dann ist das Integral von v längs w gleich der Potentialdifferenz in End- und Anfangspunkten, ist also wegunabhängig.

Ist das Kurvenintegral von v wegunabhängig, so legen wir einen Punkt $x_0 \in D$ fest und setzen

$$f(x) := \int_{x_0}^x v \cdot dX$$

wobei mit $\int_{x_0}^x$ das Kurvenintegral längs eines beliebigen Weges von x_0 zu x gemeint wird. Ähnlich zum Hauptsatz der Differential- und Integralrechnung wird es gezeigt, dass $\text{grad } f = v$ gilt (man berechne z. B. die partiellen Ableitungen von f). □

Integrabilitätsbedingung für Vektorfelder

Wie bestimmt man, ob ein Vektorfeld v ein Gradientenfeld ist? Im vorigen Abschnitt haben wir eine notwendige und hinreichende Bedingung dazu gegeben: das Kurvenintegral von v soll wegunabhängig sein. Finden wir zwei Wege mit denselben Anfangs- und Endpunkten, längs welchen die Integrale von v ungleich sind, so ist v kein Gradientenfeld. Andererseits, wenn v ein Gradientenfeld ist, dann müssen wir, um es zu zeigen, alle möglichen Wege zwischen allen Paaren von Punkten betrachten, was unmöglich ist.

Es gibt eine einfachere notwendige und hinreichende Bedingung, die aber nur für Gebiete "ohne Löcher" gilt.

Definition 3.9. Eine Teilmenge $D \subset \mathbb{R}^n$ heißt einfach zusammenhängend, wenn jede geschlossene Kurve in D auf einen Punkt zusammengezogen werden kann, ohne D zu verlassen.

Eine Kreisscheibe, ein Rechteck, ein Hufeisen sind einfach zusammenhängend. Ein Ring, ein Brezel sind nicht einfach zusammenhängend.

Satz 55. Sei $D \subset \mathbb{R}^n$ einfach zusammenhängend. Dann ist ein C^1 -Vektorfeld $v: D \rightarrow \mathbb{R}^n$ genau dann ein Gradientenfeld, wenn

$$\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i} \quad (3.2)$$

für alle $i, j = 1, \dots, n$ gilt.

Die Notwendigkeit von (3.2) ist einfach zu zeigen: wenn $v = \text{grad } f$ gilt, dann

$$\frac{\partial v_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial v_j}{\partial x_i}$$

Bei $n = 2$ und $n = 3$ ist uns die Gleichung (3.2) schon mal begegnet, in der Form $\text{rot grad } f = 0$. Die Differenzen $\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i}$ sind nämlich die Komponenten der Rotation.

Berechnung des Potentials eines Vektorfeldes auf einem einfach zusammenhängenden Gebiet D .

1. Prüfen, ob die Integrabilitätsbedingung (3.2) erfüllt ist.
2. Einen festen Punkt x_0 wählen, und für jeden $x \in D$ den Wert $f(x)$ berechnen aus dem Kurvenintegral

$$f(x) = \int_w v \cdot dX$$

für eine geeignete Kurve w , die x_0 mit x verbindet.

Zwei natürliche Wege von x_0 nach x :

- die Strecke $w(t) = x_0 + t(x - x_0)$, $t \in [0, 1]$;
- ein Streckenzug entlang der Koordinatenachsen, z. B. von (x_0, y_0, z_0) über (x, y_0, z_0) und (x, y, z_0) nach (x, y, z) .

Im letzteren Fall vereinfacht sich das Kurvenintegral auf der Strecke w_1 von (x_0, y_0, z_0) zu (x, y_0, z_0) zu

$$\int_{w_1} v_1 dx + v_2 dy + v_3 dz = \int_{w_1} v_1 dx = \int_{x_0}^x v_1(t, y_0, z_0) dt$$

Das ganze Integral schreibt sich dann als

$$\int_w v \cdot dX = \int_{x_0}^x v_1(t, y_0, z_0) dt + \int_{y_0}^y v_2(x, t, z_0) dt + \int_{z_0}^z v_3(x, y, t) dt$$

Beispiel 3.10. Bestimmen Sie, ob das Vektorfeld in \mathbb{R}^3

$$v(x, y, z) = (2xy + z^3, x^2 + 3z, 3z^2x + 3y)$$

ein Gradientenfeld ist, und falls ja, finden Sie sein Potential.

Der Raum \mathbb{R}^3 ist einfach zusammenhängend. Es gilt außerdem $\operatorname{rot} v = 0$, also ist v ein Gradientenfeld.

Wählen wir $(0, 0, 0)$ als Basispunkt und berechnen wir das Potential $f(x, y, z)$ als das Integral $\int_w v \cdot dX$ mit dem Weg w von $(0, 0, 0)$ über $(x, 0, 0)$ und $(x, y, 0)$ nach (x, y, z) .

$$f(x, y, z) = \int_0^x 0 dt + \int_0^y x^2 dt + \int_0^z (3t^2x + 3y) dt = x^2y + xz^3 + 3yz$$

Bemerkung 3.11. Ist D nicht einfach zusammenhängend, so ist die Integrierbarkeitsbedingung (3.2) nicht hinreichend für die Existenz eines Potentials. Ein Beispiel: für das Vektorfeld in der punktierten Ebene

$$v(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right), \quad (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$$

gilt $\frac{\partial v_1}{\partial y} = \frac{\partial v_2}{\partial x}$. Allerdings ist v kein Gradientenfeld, denn sein Integral entlang des geschlossenen Weges

$$w(t) = (\cos t, \sin t), \quad t \in [0, 2\pi]$$

ist ungleich Null.

3.4 Der Satz von Green und der ebene Satz von Gauß

Zirkulation und Fluß

Sei $w: [a, b] \rightarrow \mathbb{R}^n$ ein geschlossener Weg ($w(a) = w(b)$), und sei v ein Vektorfeld in \mathbb{R}^n . Dann wird das Integral von v längs w auch die *Zirkulation* des Feldes v längs w genannt. Diese Benennung wird durch das folgende Lemma erklärt.

Lemma 36. *Das Integral eines Vektorfeldes v längs einer Kurve w ist gleich dem Integral der Tangentialkomponente von v längs w :*

$$\int_w v \cdot dX = \int_w T_w(v) ds$$

wobei $T_w(v): [a, b] \rightarrow \mathbb{R}$ jedem t die Länge der Projektion von $v(t)$ auf $\dot{w}(t)$ zuordnet (mit Minusvorzeichen, falls diese Projektion gegen $\dot{w}(t)$ gerichtet ist).

Beweis. Die Projektionlänge von v auf \dot{w} berechnet sich aus

$$T_w(v) = \frac{v \cdot \dot{w}}{|\dot{w}|}$$

Die Umformung des Kurvenintegrals von v ergibt

$$\int_w v \cdot dX = \int_a^b v(w(t)) \cdot \dot{w}(t) dt = \int_a^b \frac{v \cdot \dot{w}}{|\dot{w}|} |\dot{w}| dt = \int_a^b \frac{v \cdot \dot{w}}{|\dot{w}|} ds$$

□

Sei nun w ein geschlossener Weg in der Ebene. Das Integral der Normalkomponente von v längs w wird der *Fluss* von v durch w genannt. Die Normalkomponente wird dabei in der Richtung der äußeren Normale berechnet. Bezeichnet n die äußere Einheitsnormale zur Kurve w , so ist die Normalkomponente gleich

$$N_w(v) = v \cdot n$$

und der Fluss wird definiert als

$$\int_w v \cdot n ds$$

Beispiel 3.12. Sei $v(x, y) = (y, 0)$ und sei w der positiv durchlaufene Kreis vom Radius 1 mit Zentrum $(0, 1)$. Berechnen wir die Zirkulation von v längs w und seinen Fluss durch w .

Die Parametrisierung des Kreises:

$$w(t) = (\cos t, 1 + \sin t), \quad t \in [0, 2\pi]$$

Der Geschwindigkeitsvektor:

$$\dot{w}(t) = (-\sin t, \cos t)$$

Die Zirkulation:

$$\int_w v \cdot dX = \int_0^{2\pi} \begin{pmatrix} 1 + \sin t \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = \dots$$

Die äußere Einheitsnormale:

$$n(t) = (\cos t, \sin t)$$

Der Fluss:

$$\int_w v \cdot n = \int_0^{2\pi} \begin{pmatrix} 1 + \sin t \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} dt = \dots$$

Der Satz von Green

Sei $D \subset \mathbb{R}^2$ ein beschränkter regulärer Bereich, dessen Rand ∂D aus endlich vielen geschlossenen, stückweise glatten Kurven w_1, w_2, \dots, w_n besteht. Die Parametrisierung jeder Kurve w_i sei so gewählt, dass B links zur Durchlaufrichtung liegt. Der äußere Rand wird also gegen den Uhrzeigersinn durchgelaufen, die Ränder der "Löcher" im Uhrzeigersinn.

Unter dem Integral eines Vektorfeldes v längs ∂D verstehen wir die Summe

$$\int_{\partial D} v \cdot dX := \int_{w_1} v \cdot dX + \dots + \int_{w_n} v \cdot dX$$

Satz 56. Für jedes C^1 -Vektorfeld $v: D \rightarrow \mathbb{R}^2$ gilt

$$\int_{\partial D} v \cdot dX = \int_D \operatorname{rot} v \, dF = \int_D \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dF \quad (3.3)$$

Der Satz von Green rechtfertigt die Bezeichnung von $\operatorname{rot} v$ als Wirbeldichte von v : um die Zirkulation von v längs des Randes von D zu berechnen, integriert man $\operatorname{rot} v$ über D .

Der Satz von Gauss

Satz 57. Sei $D \subset \mathbb{R}^2$ ein beschränkter regulärer Bereich, und sei n die äußere Einheitsnormale zum ∂D . Dann gilt für jedes C^1 -Vektorfeld $v: D \rightarrow \mathbb{R}^2$

$$\int_{\partial D} v \cdot n \, ds = \int_D \operatorname{div} v \, dF$$

Das rechtfertigt die Bezeichnung von $\operatorname{div} v$ als Quelldichte: um den Fluss vom Gebiet D nach außen zu berechnen, muss man $\operatorname{div} v$ über D integrieren.

Korollar 15. Für jede C^2 -Funktion $f: D \rightarrow \mathbb{R}$ gilt

$$\int_D \Delta u \, dF = \int_{\partial D} \partial_n u \, ds$$

Beweis. Es gilt $\Delta u = \operatorname{div} \operatorname{grad} u$, also

$$\int_D \Delta u \, dF = \int_D \operatorname{div} \operatorname{grad} u \, dF = \int_{\partial D} \operatorname{grad} u \cdot n \, ds = \int_{\partial D} \partial_n u \, ds$$

□

4 Integration über Flächen im Raum

4.1 Reguläre Flächen

Flächen im Raum können ähnlich zu Kurven in der Ebene dargestellt werden, durch eine Parametrisierung. Hängt ein Kurvenpunkt von einem Parameter ab, so hängt ein Flächenpunkt von zwei Parametern.

Definition 4.1. Sei $D \subset \mathbb{R}^2$ ein regulärer Bereich. Unter einer Parameterdarstellung eines regulären Flächenstückes verstehen wir eine C^1 -Abbildung $p: D \rightarrow \mathbb{R}^3$, $p(u, v) = (x(u, v), y(u, v), z(u, v))$ mit den Eigenschaften

1. für $(u, v) \neq (u', v')$ gilt stets $p(u, v) \neq p(u', v')$;
2. für jeden $(u, v) \in D$ sind die Vektoren $p_u(u, v) = \frac{\partial p}{\partial u}(u, v)$ und $p_v(u, v) = \frac{\partial p}{\partial v}(u, v)$ linear unabhängig.

Sind die Vektoren p_u und p_v linear unabhängig, so spannen sie die Tangentialebene zur Fläche im Punkt $p(u, v)$ auf.

Beispiel 4.2. Der Graph $z = h(x, y)$ kann mittels x - und y -Koordinaten parametrisiert werden:

$$p(x, y) = (x, y, h(x, y))$$

Dann gilt

$$p_x = (1, 0, h_x), \quad p_y = (0, 1, h_y)$$

Da diese zwei Vektoren immer linear unabhängig sind (warum?), ist der Graph einer C^1 -Funktion eine reguläre Fläche.

4.2 Berechnung des Flächeninhaltes

Wie bestimmt man im Allgemeinen, ob zwei Vektoren in \mathbb{R}^3 linear unabhängig sind? Dafür berechnet man das Kreuzprodukt:

$$p_u, p_v \text{ linear unabhängig} \Leftrightarrow p_u \times p_v \neq 0$$

Das Kreuzprodukt $p_u \times p_v$ hat eine geometrische Bedeutung: das ist der Flächeninhalt des von p_u und p_v aufgespannten Parallelograms. Dieser Flächeninhalt entspricht dem lokalen Flächenverzerrungsfaktor der Abbildung p (vom ebenen Bereich auf das Flächenstück). Das erklärt die Formel im folgenden

Satz 58. Der Flächeninhalt O eines regulären Flächenstücks $p: D \rightarrow \mathbb{R}^3$ ist

$$O = \int_D |p_u \times p_v| \, dudv$$

Beispiel 4.3. Der Oberflächeninhalt des Graphen $z = h(x, y)$

$$O = \int_D \sqrt{1 + h_x^2 + h_y^2} \, dxdy$$

Die Fundamentalgrößen E, F, G

Man bezeichnet

$$E := |p_u|^2, \quad F := p_u \cdot p_v, \quad G := |p_v|^2$$

Dann kann man zeigen:

$$|p_u \times p_v|^2 = EG - F^2$$

Deswegen kann die Formel zur Berechnung des Flächeninhaltes auch wie folgt geschrieben werden:

$$O = \int_D \sqrt{EG - F^2} \, dxdy$$

Beispiel 4.4. Eine Kurve $(x(t), 0, z(t))$ in der xz -Koordinatenebene erzeugt bei Drehung um die z -Achse die Drehfläche

$$p(t, \varphi) = (x(t) \cos \varphi, x(t) \sin \varphi, z(t))$$

Es gilt

$$x_t = (\dot{x} \cos \varphi, \dot{x} \sin \varphi, \dot{z}), \quad x_\varphi = (-x \sin \varphi, x \cos \varphi, 0)$$

Daraus folgt

$$E = \dot{x}^2 + \dot{z}^2, \quad F = 0, \quad G = x^2$$

Folglich berechnet sich der Flächeninhalt einer Drehfläche als

$$O = \int_D \sqrt{x^2(\dot{x}^2 + \dot{z}^2)} dt d\varphi = 2\pi \int_{t_0}^{t_1} x \sqrt{\dot{x}^2 + \dot{z}^2} dt$$

Beispiel 4.5. Die Oberfläche der Kugel.

4.3 Das Oberflächenintegral

Integral einer Funktion

Definition 4.6. Sei $p: D \rightarrow \mathbb{R}^3$ eine Parameterdarstellung eines regulären Flächenstücks S , und sei f eine auf S definierte Funktion. Dann wird das Oberflächenintegral von f über S definiert/berechnet als

$$\int_S f dO = \int_D f(x(u, v), y(u, v), z(u, v)) |p_u(u, v) \times p_v(u, v)| du dv$$

Integral eines Vektorfeldes

Definition 4.7. Sei $S \subset \mathbb{R}^3$ ein reguläres Flächenstück, und sei n das Feld der Einheitsnormalen zu S . Dann definiert man das Oberflächenintegral eines Vektorfeldes v über S als

$$\int_S v \cdot dO = \int_S (v \cdot n) dO$$

Wird das Flächenstück mit $p(s, t)$ parametrisiert, so kann die Einheitsnormale wie folgt berechnet werden:

$$n = \pm \frac{1}{|p_s \times p_t|} p_s \times p_t$$

Bei der Wahl des + Vorzeichen erhalten wir dann

$$\int_S v \cdot dO = \int_D [v, p_s, p_t] ds dt$$

Die physikalische Deutung: dieses Integral gibt den Fluß des Vektorfeldes durch die Fläche S (in der Richtung von n) an.

4.4 Der Divergenzsatz von Gauss

Satz 59. Sei $D \subset \mathbb{R}^3$ ein beschränkter Bereich dessen Oberfläche ∂D aus endlich vielen geschlossenen, stückweise regulären Flächen besteht. Bezeichnet n die nach außen weisende Einheitsnormale, dann gilt für jeden Vektorfeld v auf D die Formel

$$\int_D (\operatorname{div} v) dV = \int_{\partial D} v \cdot dO$$

Mit anderen Worten: der Fluß durch die Fläche ∂D ist gleich dem Integral der Divergenz über D .