

Kurzfassung

# Mathematische Grundlagen der Geowissenschaften

## Lineare Algebra

Stefan Hergarten  
Institut für Geo- und Umweltnaturwissenschaften  
Albert-Ludwigs-Universität Freiburg



Wintersemester 2019/20  
Stand: 21. Januar 2020



## Inhaltsverzeichnis

1	Vektoren und Vektorräume	4
2	Längen und Abstände	7
3	Polar- und Kugelkoordinaten	9
4	Skalarprodukte	11
5	Das Vektorprodukt in $\mathbb{R}^3$	13
6	Parallel- und Zentralprojektionen	15
7	Projektionen der Einheitskugel in $\mathbb{R}^3$	16
8	Linearkombinationen und lineare Abhängigkeit	19
9	Unterräume	21
10	Dimension	22
11	Basen	23
12	Reziproke Basen	24
13	Lineare Abbildungen	25
14	Lineare Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}^2$ und von $\mathbb{R}^3$ nach $\mathbb{R}^3$	27
15	Matrizen	29
16	Verkettung linearer Abbildungen und Matrixmultiplikation	30



17 Bild, Kern und Rang von linearen Abbildungen	32
18 Lineare Gleichungssysteme	33
19 Die Methode der kleinsten Quadrate	34
20 Linearer Fit	35
21 Matrizen unter Basiswechsel	36
22 Spur und Determinante	37
23 Eigenwerte	40
24 Die einfache Scherung	42
25 Die adjungierte Abbildung	43
26 Orthogonale Abbildungen	44
27 Selbstadjungierte Abbildungen	46
28 Singulärwertzerlegung	48
29 Polarzerlegung	49
30 Hauptstreckungsachsen	49



## 1 Vektoren und Vektorräume

Es gibt verschiedene Vorstellungen von Vektoren, z. B.

**in der Schule:** Pfeile, später Tupel von Zahlen der Form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

**in der Informatik:** eine Ansammlung von  $n$  Objekten (können auch Zahlen sein) sein, welche durchnummeriert sind:

$$\begin{pmatrix} O_1 \\ O_2 \\ \dots \\ O_n \end{pmatrix}$$

**in der Mathematik:** Elemente eines Vektorraums

Ebenfalls gibt es verschiedene Schreibweisen, z. B. mit Vektorpfeil oder Unterstreichung. In der Mathematik werden Vektoren normalerweise überhaupt nicht explizit gekennzeichnet, da immer aus dem Zusammenhang klar ist, ob es um Zahlen oder Vektoren geht. Oftmals verwendet man dann für Vektoren lateinische Buchstaben und für Zahlen griechische Buchstaben.

Die mathematische Definition von Vektoren fasst die Eigenschaften zusammen, welche allen Vorstellungen von Vektoren gemeinsam sind. Diese beruht nicht darauf, wie Vektoren aussehen, sondern darauf, was man mit Vektoren machen kann: Man kann Vektoren addieren und strecken (bzw. stauchen).

Für die Addition von Vektoren gelten dieselben Regeln wie für die Addition von reellen Zahlen:

- A1** Zu jedem Paar von Vektoren  $a$  und  $b$  gibt es einen Vektor  $a + b$ .
- A2**  $a + b = b + a$  für alle Vektoren  $a$  und  $b$ .
- A3**  $(a + b) + c = a + (b + c)$  für alle Vektoren  $a$ ,  $b$  und  $c$ .
- A4** Es gibt einen Nullvektor  $0$ , sodass  $a + 0 = a$  für alle Vektoren  $a$ .
- A5** Zu jedem Vektor  $a$  gibt es einen inversen (umgekehrten) Vektor  $-a$ , sodass  $a + (-a) = 0$ .



In Analogie zur Subtraktion von Zahlen verwendet man die Kurzschreibweise  $a - b$  für  $a + (-b)$ .

Die Streckung (bzw. Stauchung) stellen wir als Multiplikation von reellen Zahlen mit Vektoren dar. Für diese gelten zumindest einige der Regeln, die wir von der Multiplikation reeller Zahlen untereinander kennen:

- S1** Zu jeder reellen Zahl  $\lambda$  und zu jedem Vektor  $a$  gibt es einen Vektor  $\lambda a$ .
- S2**  $1a = a$  für alle Vektoren  $a$ .
- S3**  $\lambda(\mu a) = (\lambda\mu)a$  für alle Zahlen  $\lambda$  und  $\mu$  und alle Vektoren  $a$ .

Darüber hinaus „vertragen“ sich Vektoraddition und Streckung miteinander:

- AS1**  $(\lambda + \mu)a = \lambda a + \mu a$  für alle Zahlen  $\lambda$  und  $\mu$  und alle Vektoren  $a$ .
- AS2**  $\lambda(a + b) = \lambda a + \lambda b$  für alle Zahlen  $\lambda$  und alle Vektoren  $a$  und  $b$ .

Diese Regeln entsprechen ebenfalls denen, die wir von Addition und Multiplikation reeller Zahlen untereinander kennen. Damit ist das Rechnen mit Vektoren nicht komplizierter als das Rechnen mit reellen Zahlen.

Eine Menge aus Objekten, welche sich nach den Regeln A1–A5, S1–S3 und AS1–AS2 addieren und strecken lassen, wird als **Vektorraum** bezeichnet. Die Elemente eines Vektorraumes heißen **Vektoren**.

Die Eigenschaften A1–A5, S1–S3 und AS1–AS2, welche zur Definition eines Vektorraumes verwendet werden, stellen einen minimalen Satz von Eigenschaften dar. Alle weiteren bekannten Rechenregeln für Vektoren, z. B.

$$0a = 0 \quad \text{und} \quad (-1)a = -a$$

lassen sich aus diesen ableiten.

Im Folgenden besprechen wir vier Beispiele von Vektorräumen im Sinne der obigen Definition:

**Die reellen Zahlen** selbst bilden einen Vektorraum. Abgesehen von dem (trivialen) Vektorraum, der nur aus einem Nullvektor besteht, ist dies der einfachste Vektorraum.



**Pfeile in der Ebene** bilden einen Vektorraum, wenn wir die übliche Addition und Streckung verwenden:

- Zwei Pfeile werden addiert, indem der Anfang des zweiten an die Spitze des ersten gesetzt wird. Das Ergebnis ist ein Pfeil vom Anfang des ersten zur Spitze der zweiten Pfeils.
- Ein Pfeil wird um einen positiven Faktor  $\lambda$  gestreckt, indem seine Länge auf das  $\lambda$ -fache erhöht wird, während die Richtung beibehalten wird. Bei einer Streckung um einen negativen Faktor  $\lambda$  wird die Länge um den Faktor  $|\lambda|$  erhöht, und die Richtung umgekehrt.

**Tupel aus  $n$  reellen Zahlen** bilden einen Vektorraum, wenn wir die Addition und Streckung komponentenweise definieren:

- Zwei Tupel werden addiert, indem die einzelnen Komponenten addiert werden.
- Ein Tupel wird um einen Faktor  $\lambda$  gestreckt, indem die einzelnen Komponenten mit  $\lambda$  multipliziert werden.

Dieser Vektorraum heißt  $\mathbb{R}^n$  und wird speziell für  $n = 2$  und  $n = 3$  zur Beschreibung von Punkten in der Ebene bzw. im Raum verwendet.

**Funktionen** können ebenfalls (wenn auch nicht sehr anschauliche) Vektorräume bilden. Nehmen wir z. B. die Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$ , d. h. die Funktionen, deren Definitionsmenge die gesamten reellen Zahlen umfasst, und deren Wertemenge  $\mathbb{R}$  oder eine Teilmenge von  $\mathbb{R}$  ist. Mit den naheliegenden Definitionen

- Addition:  $(f + g)(x) = f(x) + g(x)$
- Streckung:  $(\lambda f)(x) = \lambda f(x)$

bildet die Menge dieser Funktionen einen Vektorraum.

Es gibt also Vektoren, die auf dem ersten Blick nicht viel mit der Vorstellung von Vektoren aus der Schule gemeinsam haben. Die meisten Eigenschaften von Vektoren, die wir demnächst kennenlernen werden, gelten für Vektoren im allgemeinen Sinn (nach obiger Definition eines Vektorraumes) und sind nicht auf die Vorstellung von Vektoren aus der Schulmathematik beschränkt. Daher ist es durchaus von Vorteil, sich mit der allgemeinen Definition von Vektoren zu beschäftigen. Veranschaulichen kann man sich die meisten Ergebnisse allerdings gut am Beispiel von Pfeilen in der Ebene bzw.  $\mathbb{R}^2$  oder  $\mathbb{R}^3$ .

---

**Box 1** Vektoren in MATLAB

In MATLAB bestehen Vektoren aus  $n$  Zahlen, welche von 1 bis  $n$  durchnummeriert sind. Ist  $a$  ein Vektor, wird mit  $a(1)$ ,  $a(2)$ ,  $\dots$ ,  $a(n)$  auf die einzelnen Elemente (Komponenten) zugegriffen. Die mathematischen Rechenoperationen

Vektor+Vektor = Vektor

Zahl\*Vektor = Vektor

sind direkt in MATLAB implementiert und müssen demnach nicht einzeln Element für Element durchgeführt werden. Darüber hinaus implementiert MATLAB auch weitere naheliegende Rechenoperationen direkt, z. B.

Zahl+Vektor = Vektor

Vektor+Zahl = Vektor

Vektor\*Zahl = Vektor

Vektor.\*Vektor = Vektor (elementweise Multiplikation)

sowie die entsprechenden Versionen mit  $-$  statt  $+$  und  $/$  statt  $*$ .

---

Etwas gewöhnungsbedürftig mag es zunächst sein, dass Vektoren nicht unbedingt eine Länge haben, nicht wissen, was es heißt, senkrecht aufeinander zu stehen, und dass man Vektoren nicht unbedingt miteinander multiplizieren kann.

## 2 Längen und Abstände

Der Begriff der Länge eines Vektors gehört nicht zur minimalen mathematischen Definition eines Vektorraums, sondern ist als Zusatz zu betrachten. Dies kommt daher, dass es viele verschiedene Vorstellungen gibt, was unter der Länge eines Vektors zu verstehen ist, und es stark von der jeweiligen Anwendung abhängt, welcher Längenbegriff der sinnvollste ist.

Für Vektoren aus  $\mathbb{R}^2$  und  $\mathbb{R}^3$  haben wir bereits eine Vorstellung, was ihre

Länge ist. Diese lässt sich leicht auf die Länge eines Vektors  $a = \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}$

aus  $\mathbb{R}^n$  verallgemeinern:

$$\|a\| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} \quad (1)$$

Auf allgemeine Vektoren, z. B. auf die im letzten Abschnitt erwähnte Betrachtung von Funktionen als Vektoren, lässt sich dies jedoch nicht direkt übertragen.

Allgemein wird eine **Norm** oder **Länge** (aber nicht Betrag) definiert als eine Abbildung von einem Vektorraum in die reellen Zahlen mit den Eigenschaften

- N1** Für alle Vektoren  $a$  und  $b$  gilt die **Dreiecksungleichung**:  $\|a + b\| \leq \|a\| + \|b\|$ .
- N2** Für alle Vektoren  $a$  und alle reellen Zahlen  $\lambda$  gilt:  $\|\lambda a\| = |\lambda| \|a\|$ .
- N3** Nur der Nullvektor hat die Länge null, d. h.  $\|a\| = 0$  gilt nur, wenn  $a = 0$ .

Aus den Eigenschaften N1 und N2 folgt die wichtige Eigenschaft, dass die Länge eines Vektors niemals negativ ist.

Als **Abstand** zweier Punkte  $a$  und  $b$  bezeichnet man die Länge der Differenz  $a - b$ , d. h.  $\|a - b\|$ .

Nicht jeder Vektorraum besitzt eine Norm. Ein Vektorraum, der eine Norm besitzt, heißt **normierter** Vektorraum.  $\mathbb{R}^n$  ist demnach ein normierter Vektorraum, allerdings gibt es außer der oben erwähnten **euklidischen Norm** noch beliebig viele weitere Normen in  $\mathbb{R}^n$ , z. B.

- Die **Manhattan-Norm**  $\|a\| = |a_1| + |a_2| + \dots + |a_n|$ . Deren Name geht darauf zurück, dass sie in  $\mathbb{R}^2$  die Länge des Weges beschreibt, den man in einer Stadt mit rechtwinklig angeordneten Straßen zurücklegt.
- Die **Maximumsnorm**  $\|a\| = \max\{|a_1|, |a_2|, \dots, |a_n|\}$ ,

Bei der euklidischen Norm wird statt  $\|a\|$  häufig die Schreibweise  $|a|$  verwendet.

Die drei oben diskutierten Normen lassen sich verallgemeinern zur sogenannten **p-Norm**

$$\|a\| = \sqrt[p]{|a_1|^p + \dots + |a_n|^p} \quad (2)$$

für  $p \geq 1$ . Für  $p = 1$  erhalten wir die Manhattan-Norm, für  $p = 2$  die euklidische Norm, und für  $p = \infty$  die Maximumsnorm. Werte  $p < 1$  sind nicht zulässig, weil dann die Dreiecksungleichung nicht erfüllt wäre. Grundsätzlich führen große Werte von  $p$  dazu, dass die (betragsmäßig) größte Komponente von  $a$  einen großen Einfluss auf die Länge hat, während für  $p = 1$  alle Komponenten gleich stark beitragen. Dies ist insbesondere wichtig, wenn theoretische Kurven an Messdaten angepasst werden (Aufgabe 1).

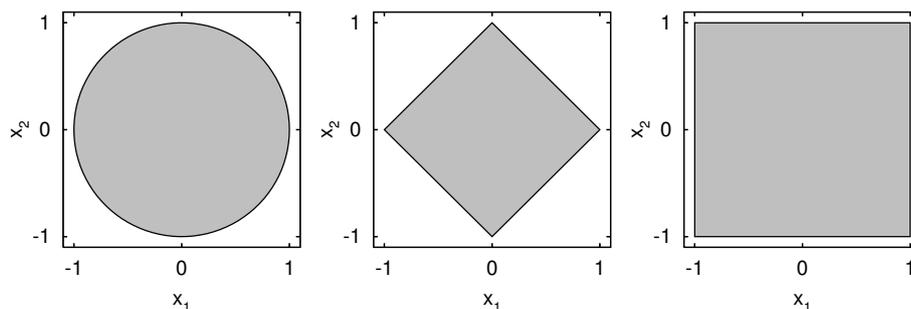


Abbildung 1: Die Einheitskugel in  $\mathbb{R}^2$  (Einheitskreis) mit euklidischer Norm (links), Manhattan-Norm (Mitte) und Maximumnorm (rechts)

---

**Box 2** Die Länge von Vektoren in MATLAB

Die Funktion zur Berechnung der Länge eines Vektors heißt `norm`, nicht `length`. Letztere liefert die Anzahl der Komponenten eines Vektors. Neben dem einfachen Aufruf `norm(a)`, welcher die euklidische Norm liefert, lässt sich eine beliebige  $p$ -Norm durch `norm(a,p)` berechnen (die Maximumnorm mit `norm(a,inf)`).

---

Ein Vektor der Länge 1 heißt **Einheitsvektor** oder **normierter** Vektor. Aus jedem Vektor  $a \neq 0$  eines normierten Vektorraums lässt sich durch geeignete Streckung ein Einheitsvektor machen:  $e = \frac{1}{\|a\|} a$ . Man bezeichnet diesen Vorgang als **Normierung** der Vektors  $a$ .

Die Menge aller Punkte  $x$ , deren Abstand von einem gegebenen Punkt  $a$  kleiner oder gleich einer gegebenen Zahl  $r$  ist, d. h.

$$\|x - a\| \leq r \quad (3)$$

wird als **Kugel** mit Mittelpunkt  $a$  und Radius  $r$  bezeichnet. Die **Einheitskugel** ist die Kugel mit Mittelpunkt 0 und Radius 1. Diese sieht in verschiedenen Normen völlig unterschiedlich aus (Abb. 1).

### 3 Polar- und Kugelkoordinaten

In  $\mathbb{R}^2$  (mit euklidischer Norm) lässt sich jeder Einheitsvektor durch einen Winkel darstellen:  $e = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$  mit  $\varphi \in [0, 2\pi[$ . Hieraus folgt sofort, dass

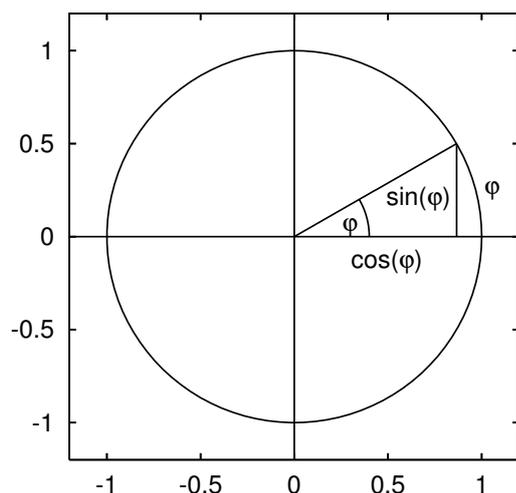


Abbildung 2: Definition von  $\varphi$  im Bogenmaß,  $\cos(\varphi)$  und  $\sin(\varphi)$  am Einheitskreis

---

### Box 3 Sinus und Cosinus in MATLAB

Wie in allen Programmiersprachen erwarten die Sinus- und Cosinusfunktion in MATLAB (`sin` und `cos`) den Winkel als Argument im Bogenmaß. Zusätzlich stellt MATLAB aber auch entsprechende Funktionen `sind` und `cosd` für Winkelangaben in Grad bereit. So ist `sind(phi)` dasselbe wie `sin(phi*pi/180)`.

---

sich jeder Vektor der  $\mathbb{R}^2$  als

$$a = r \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \quad \text{mit } r = |a| \geq 0 \quad \text{und } \varphi \in [0, 2\pi[ \quad (4)$$

darstellen lässt. Diese Darstellung wird als **ebene Polarkoordinaten** bezeichnet.

Winkel werden in der Mathematik meist im Bogenmaß statt in Grad angegeben (Abb. 2). Ein voller Kreis ( $360^\circ$ ) entspricht dabei dem Umfang des Einheitskreises, also  $2\pi$ . Das Bogenmaß ist eine dimensionslose Größe, jedoch wird häufig die Einheit Radian (abgekürzt rad) verwendet.

Die Schreibweisen  $\sin \varphi$  und  $\cos \varphi$  sind Kurzformen von  $\sin(\varphi)$  bzw.  $\cos(\varphi)$ .

In  $\mathbb{R}^3$  werden zwei Winkel benötigt, wobei die Definition der Winkel in

---

**Box 4** Polar- und Kugelkoordinaten in MATLAB

MATLAB stellt die Funktionen `cart2pol` und `pol2cart` für die Transformation zwischen den üblichen kartesischen Koordinaten und ebenen Polarkoordinaten sowie die Funktionen `cart2sph` und `sph2cart` für die Transformation zwischen kartesischen Koordinaten und Kugelkoordinaten bereit. Auch wenn in der Dokumentation die Winkel anders heißen, entsprechen diese der geographischen Definition (Breitengrad und Längengrad).

---

Mathematik und Geowissenschaften meist unterschiedlich ist:

$$e = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} = \begin{pmatrix} \cos \phi \cos \lambda \\ \cos \phi \sin \lambda \\ \sin \phi \end{pmatrix} \quad (5)$$

mit

$$\begin{aligned} \vartheta &\in [0, \pi] \text{ (bzw. } [0, 180^\circ]) &&= \text{Polarwinkel} \\ \varphi &\in [0, 2\pi[ \text{ (bzw. } [0, 360^\circ]) &&= \text{Azimutwinkel} \\ \phi &\in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \text{ (bzw. } [-90^\circ, 90^\circ]) &&= \text{Breitengrad} \\ \lambda &\in ]-\pi, \pi] \text{ (bzw. } [-180^\circ, 180^\circ]) &&= \text{Längengrad} \end{aligned}$$

sodass sich jeder Vektor  $a$  als

$$a = r \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} = r \begin{pmatrix} \cos \phi \cos \lambda \\ \cos \phi \sin \lambda \\ \sin \phi \end{pmatrix} \quad (6)$$

darstellen lässt. Dies bezeichnet man als **Kugelkoordinaten**.

Azimutwinkel und Längengrad sind in ihrer Bedeutung bis auf den Wertebereich identisch, während Polarwinkel und Breitengrad über die Relation  $\vartheta = \frac{\pi}{2} - \phi$  zusammenhängen.

## 4 Skalarprodukte

Das Standard-Skalarprodukt zweier Vektoren  $a = \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}$  und  $b =$

$\begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}$  aus  $\mathbb{R}^n$  wird üblicherweise definiert als

$$a \cdot b = a_1 b_1 + a_2 b_2 + \dots + a_n b_n \quad (7)$$

---

**Box 5** Das Standard-Skalarprodukt in MATLAB

---

In MATLAB lässt sich das Standard-Skalarprodukt zweier Vektoren  $a$  und  $b$  elementar nach Gleichung 7 als  $\text{sum}(a.*b)$  berechnen oder alternativ durch Anwendung der Funktion  $\text{dot}$  als  $\text{dot}(a,b)$ .

---

Die Bezeichnung Skalarprodukt kommt daher, dass es eine Zahl ist.

Wie schon bei der Länge von Vektoren, gibt es auch bei Skalarprodukten verschiedene Definitionen. Ebenfalls analog zur Länge von Vektoren wird das Skalarprodukt über seine Eigenschaften definiert. Ein **Skalarprodukt** ist eine Abbildung, die jedem Paar  $a$  und  $b$  von Vektoren eines Vektorraums eine Zahl  $a \cdot b$  zuordnet und die folgenden Eigenschaften hat:

**SP1** Für alle Vektoren  $a$  und  $b$  ist  $a \cdot b = b \cdot a$ .

**SP2** Für alle Vektoren  $a$ ,  $b$  und  $c$  gilt:

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

**SP3** Für alle Vektoren  $a$  und  $b$  und alle reellen Zahlen  $\lambda$  gilt:

$$(\lambda a) \cdot b = \lambda(a \cdot b)$$

**SP4** Für alle Vektoren  $a \neq 0$  ist  $a \cdot a > 0$ .

Im Gegensatz zur Streckung von Vektoren werden Skalarprodukte mit einem (dicken) Punkt geschrieben.

Nicht jeder Vektorraum besitzt ein Skalarprodukt. Ein Vektorraum, welcher ein Skalarprodukt besitzt, wird als **euklidischer** Vektorraum bezeichnet. Zur Unterscheidung von anderen Skalarprodukten wird in Gleichung 7 definierte Skalarprodukt in  $\mathbb{R}^n$  als **Standard-Skalarprodukt** bezeichnet.

Zwei Vektoren  $a$  und  $b$  werden als **orthogonal** oder **senkrecht** (zu einander) bezeichnet, wenn  $a \cdot b = 0$  ist. Mehrere Vektoren heißen orthogonal, wenn jedes Paar von ihnen orthogonal ist.

Der Nullvektor steht auf allen Vektoren des jeweiligen Vektorraums senkrecht (folgt direkt aus SP3), und er ist auch der einzige Vektor, der diese Eigenschaft hat (folgt direkt aus SP4).

Mittels der Definition

$$\|a\| = \sqrt{a \cdot a} \quad (8)$$

lässt sich aus jedem Skalarprodukt eine Norm bilden. Im Fall des Standard-Skalarprodukts in  $\mathbb{R}^n$  ist dies die euklidische Norm. Manhattan-Norm und Maximumnorm (sowie alle  $p$ -Normen für  $p \neq 2$ ) lassen sich hingegen nicht auf ein Skalarprodukt zurückführen.

Für die aus einem Skalarprodukt gebildete Norm gilt die Cauchy-Schwarz-Ungleichung

$$|a \cdot b| \leq \|a\| \|b\|. \quad (9)$$

Somit liegt der Quotient  $\frac{a \cdot b}{\|a\| \|b\|}$  für alle Vektoren  $a \neq 0$  und  $b \neq 0$  immer im Bereich von  $-1$  bis  $1$ . Hieraus definiert man den **Winkel**  $\alpha$  zwischen den Vektoren  $a$  und  $b$  gemäß

$$\cos \alpha = \frac{a \cdot b}{\|a\| \|b\|} \quad (10)$$

Man wählt für  $\alpha$  immer den Bereich zwischen  $0$  und  $\pi$  (bzw.  $0^\circ$  und  $180^\circ$ ).

Damit definiert das Skalarprodukt einen Längen- und Winkelbegriff.

Ein Satz von Vektoren  $x_1, \dots, x_m$  heißt **orthonormal**, wenn die Vektoren

- paarweise orthogonal, d. h.  $x_i \cdot x_j = 0$  wenn  $i \neq j$ , und
- normiert, d. h.  $\|x_i\| = 1$ ,

sind.

Ist  $b$  ein gegebener Vektor, so lässt sich jeder Vektor  $a$  in einen Anteil

$$a_{\parallel} = (a \cdot e) e \quad (11)$$

in Richtung von  $b$  und einen Anteil

$$a_{\perp} = a - a_{\parallel} = a - (a \cdot e) e \quad (12)$$

senkrecht zu  $b$  zerlegen, wobei  $e$  ein Einheitsvektor in Richtung von  $b$  ist.

## 5 Das Vektorprodukt in $\mathbb{R}^3$

In  $\mathbb{R}^3$  gibt es eine Möglichkeit, Vektoren so miteinander zu multiplizieren, dass das Ergebnis wieder ein Vektor ist (im Gegensatz zum Skalarprodukt,

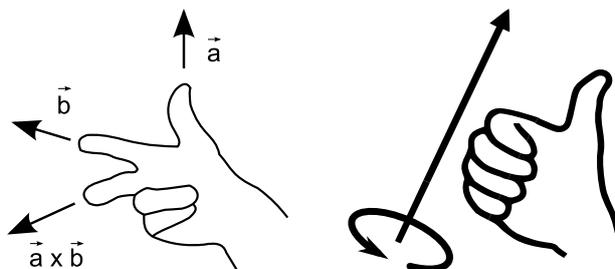


Abbildung 3: Dreifingerregel (links) und Rechte-Faust-Regel (rechts). Quellen: <https://de.wikipedia.org/wiki/Drei-Finger-Regel>, <https://de.wikipedia.org/wiki/Korkenzieherregel>

welches eine reelle Zahl, also ein Skalar ist):

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \quad (13)$$

Dieses Produkt wird als **Vektorprodukt** oder **Kreuzprodukt** bezeichnet. Seine wichtigsten Eigenschaften sind

**VP1** Für alle Vektoren  $a$  und  $b$  ist  $a \times b = -b \times a$ .

**VP2** Für alle Vektoren  $a$ ,  $b$  und  $c$  gilt:

$$(a + b) \times c = a \times c + b \times c$$

**VP3** Für alle Vektoren  $a$  und  $b$  und alle reellen Zahlen  $\lambda$  gilt:

$$(\lambda a) \times b = \lambda(a \times b)$$

**VP4** Der Vektor  $a \times b$  steht senkrecht auf den Vektoren  $a$  und  $b$ .

Die Richtung des Vektors  $a \times b$  lässt sich mit der **Dreifingerregel** veranschaulichen (Abb. 3).

Besondere Relevanz in den Geowissenschaften haben die beiden folgenden Anwendungen des Vektorprodukts:

- Die Länge des Vektors  $a \times b$  ist die Fläche des von den Vektoren  $a$  und  $b$  aufgespannten Parallelogramms. Hieraus folgt, dass die Fläche eines Dreiecks in  $\mathbb{R}^3$  mit den Ecken  $a$ ,  $b$  und  $c$  sich als  $\frac{1}{2} \|(b-a) \times (c-a)\|$  berechnet.

---

**Box 6** Das Vektorprodukt in MATLAB

---

In MATLAB lässt sich das Vektorprodukt zweier Vektoren  $a$  und  $b$  aus  $\mathbb{R}^3$  leicht mit der Funktion `cross` als `cross(a,b)` berechnen.

---

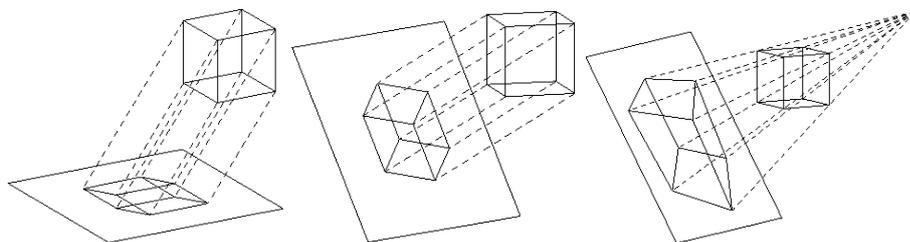


Abbildung 4: Allgemeine Parallelprojektion (links), orthogonale Parallelprojektion (Mitte) und Zentralprojektion (rechts). Quelle: [http://www.mathematik.tu-darmstadt.de/~ehartmann/html\\_cdg/node1.html](http://www.mathematik.tu-darmstadt.de/~ehartmann/html_cdg/node1.html)

- Die Geschwindigkeit  $v$  der Partikel am Ort  $x$  bei einer Rotationsbewegung um den Koordinatenursprung ist

$$v = \omega \times x, \quad (14)$$

wobei  $\omega$  der Vektor der Winkelgeschwindigkeit ist. Die Einheit der Winkelgeschwindigkeit ist Bogenmaß pro Zeit – nicht Grad pro Zeit, sonst stimmt Gleichung 14 nicht. Der Vektor  $\omega$  zeigt in Richtung der Drehachse, und der Drehsinn kann mit der **Rechte-Faust-Regel** (Abb. 3) veranschaulicht werden.

## 6 Parallel- und Zentralprojektionen

Bei einer **Parallelprojektion** werden die Punkte des  $\mathbb{R}^3$  (oder eines anderen Vektorraums) entlang paralleler Linien auf eine gegebene Ebene projiziert (Abb. 4).

Ist  $p$  ein Vektor in Projektionsrichtung, so ist der projizierte Vektor

$$P(x) = x + \lambda p, \quad (15)$$

wobei der Faktor  $\lambda$  so gewählt werden muss, dass  $P(x)$  in der Projektionsebene liegt.

Nehmen wir der Einfachheit halber an, dass die Projektionsebene durch den Koordinatenursprung geht, muss der projizierte Vektor die Bedingung

$P(x) \cdot n = 0$  erfüllen, wobei  $n$  ein Vektor ist, der senkrecht auf der Projektionsebene steht. Hieraus folgt

$$x \cdot n + \lambda p \cdot n = 0, \quad (16)$$

sodass  $\lambda = -\frac{x \cdot n}{p \cdot n}$ , und damit

$$P(x) = x - \frac{x \cdot n}{p \cdot n} p \quad (17)$$

Im Spezialfall der **orthogonalen Projektion** ist die Projektionsrichtung  $p$  senkrecht zur Projektionsebene. Damit reduziert sich Gleichung 17 zu

$$P(x) = x - \frac{x \cdot p}{\|p\|^2} p \quad (18)$$

Eine **Zentralprojektion** unterscheidet sich von einer Parallelprojektion dadurch, dass die Projektionslinien verschiedener Punkte nicht parallel sind, sondern sich in einem **Projektionszentrum**  $p$  treffen. Eine orthogonale Parallelprojektion kann als Spezialfall einer Zentralprojektion mit einem unendlich weit entfernten Projektionszentrum betrachtet werden. Allerdings ist die Zentralprojektion im Gegensatz zur Parallelprojektion keine Projektion im Sinne der linearen Algebra.

Geht die Projektionsebene durch den Koordinatenursprung, so ist der projizierte Vektor

$$P(x) = p + \lambda(x - p), \quad (19)$$

wobei der Faktor  $\lambda$  so gewählt werden muss, dass  $P(x) \cdot n = 0$  ist. Hieraus folgt sofort

$$\lambda = -\frac{p \cdot n}{(x - p) \cdot n} = \frac{1}{1 - \frac{x \cdot n}{p \cdot n}}, \quad (20)$$

und damit

$$P(x) = p + \frac{1}{1 - \frac{x \cdot n}{p \cdot n}} (x - p). \quad (21)$$

Falls  $p$  senkrecht auf der Projektionsebene steht, reduziert sich dies zu

$$P(x) = p + \frac{1}{1 - \frac{x \cdot p}{\|p\|^2}} (x - p). \quad (22)$$

## 7 Projektionen der Einheitskugel in $\mathbb{R}^3$

Projektionen von Kugeloberflächen auf Ebenen haben sowohl eine Bedeutung für die Erstellung von Karten als auch für die Darstellung von

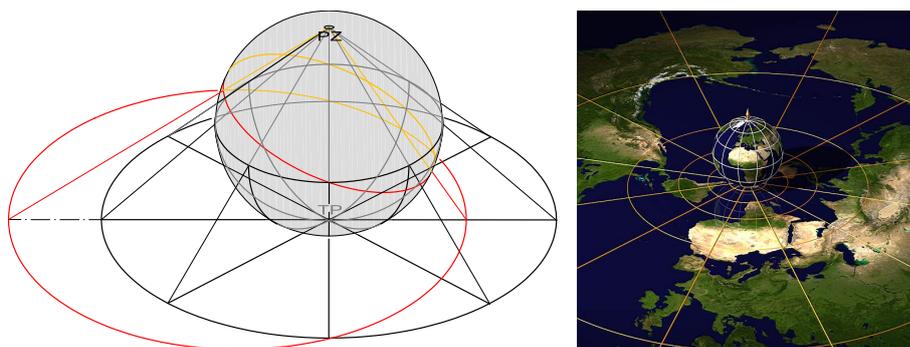


Abbildung 5: Stereografische Projektion. Quellen: [https://de.wikipedia.org/wiki/Stereografische\\_Projektion](https://de.wikipedia.org/wiki/Stereografische_Projektion) (© Analemma), [http://images.math.cnrs.fr/IMG/jpg/TN\\_Stereographic\\_projection.jpg](http://images.math.cnrs.fr/IMG/jpg/TN_Stereographic_projection.jpg)

Richtungs- und Orientierungsdaten in der Geologie. Der Einfachheit halber betrachten wir nur die Einheitskugel. Im allgemeinen Fall werden die Daten zunächst von einer beliebigen Kugel auf die Einheitskugel projiziert.

Die einfachste Projektion von der Einheitskugel auf eine Ebene ist die orthogonale Projektion aus dem letzten Abschnitt. Hierbei ist nur zu beachten, dass fast immer zwei Punkte der Kugel auf denselben Punkt der Ebene projiziert werden, sodass nur jeweils eine Halbkugel betrachtet werden sollte.

Wesentlich weiter verbreitet ist die **stereografische Projektion** (Abb. 5), bei der das Projektionszentrum  $p$  auf der Einheitskugel liegt, und die Projektionsebene gegenüber (also bei  $-p$ ) tangential an der Kugel anliegt.

Etwas einfacher zu berechnen ist die Version, bei der die Projektionsebene nicht tangential anliegt, sondern durch den Mittelpunkt der Einheitskugel geht. Hierbei ist Gleichung 22 mit  $\|p\| = 1$  direkt anwendbar. Von der Anschauung her angenehmer ist es allerdings, wenn wir statt des Projektionszentrums  $p$  (welches bei der Projektion im Unendlichen landet) den gegenüberliegenden Punkt  $s = -p$ , welcher im Zentrum der Projektionsfläche landet, verwenden. Dann ändert sich die Gleichung zu

$$P(x) = -s + \frac{1}{1 + x \cdot s} (x + s). \quad (23)$$

Die stereografische Projektion ist **winkeltreu** und **kreistreu**, d. h. Winkel zwischen zwei Linien auf der Kugel sind in der Projektion dieselben, und Kreise auf der Kugel sind in der Projektion auch Kreise. Als geometrisches

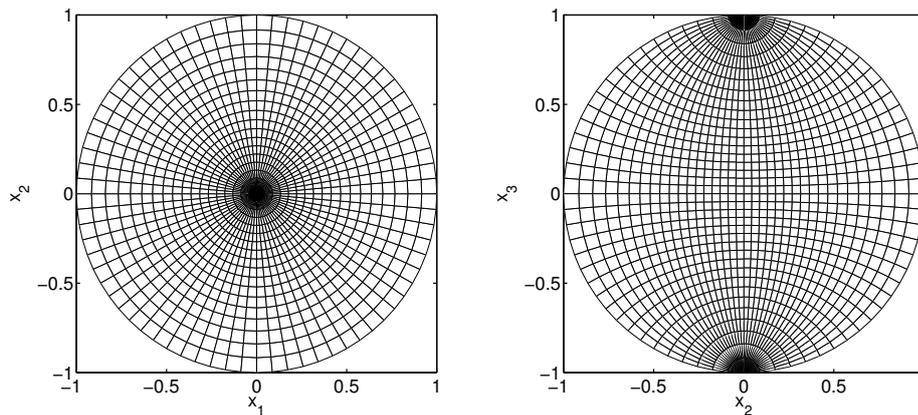


Abbildung 6: Stereografische Projektion der Hälfte des Einheitskreises (Längen- und Breitengrade in  $5^\circ$ -Schritten) mit Zentrum auf der  $x_3$ -Achse (links) und auf der  $x_1$ -Achse (rechts)

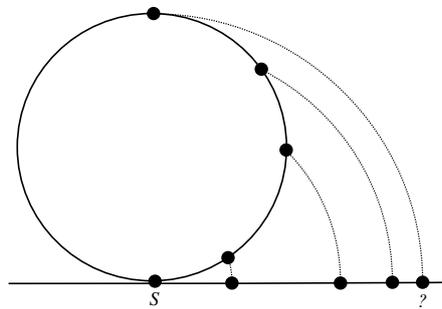


Abbildung 7: Flächentreue Lambert-Azimuthalprojektion. Quelle: [https://en.wikipedia.org/wiki/Lambert\\_azimuthal\\_equal-area\\_projection](https://en.wikipedia.org/wiki/Lambert_azimuthal_equal-area_projection)

Hilfsmittel zur Konstruktion von stereografischen Projektionen wurde bis vor einer halben Ewigkeit das **Wulfsche Netz** (Abb. 6) verwendet, welches nicht nach dem Freiburger Geologen Gerwin Wulf benannt ist.

In der Strukturgeologie und der Seismologie wird allerdings meist die flächentreue Lambert-Azimuthalprojektion (Abb. 7) verwendet. Hierbei wird die Einheitskugel um einen gegebenen Punkt  $s$  herum unter Erhaltung der Oberfläche ausgerollt.

Zunächst wird der Punkt mittels einer orthogonalen Parallelprojektion auf die (im Koordinatenursprung liegende) Projektionsebene projiziert, und

dann wird der resultierende Vektor passend gestreckt:

$$P(x) = \lambda (x - (x \cdot s) s). \quad (24)$$

Da eine Kappe der Einheitskugel mit halbem Öffnungswinkel  $\alpha$  eine Oberfläche von  $2\pi(1 - \cos \alpha)$  hat, und  $\alpha$  der Winkel zwischen  $x$  und  $s$  ist, muss für eine flächentreue Projektion  $\lambda$  so bestimmt werden, dass

$$\pi \|P(x)\|^2 = 2\pi(1 - \cos \alpha) = 2\pi(1 - x \cdot s) \quad (25)$$

ist.

Für die Darstellung von Daten ist es allerdings oftmals sinnvoll, dass die dem Punkt  $s$  zugewandte Halbkugel in den Einheitskreis projiziert wird. Da die Fläche des Einheitskreises halb so groß ist wie die einer Halbkugel ( $\pi$  gegenüber  $2\pi$ ), müssen wir hierfür den Faktor 2 in Gleichung 25 weglassen, also  $\lambda$  so bestimmen werden, dass

$$\pi \|P(x)\|^2 = \pi(1 - x \cdot s) \quad (26)$$

ist. Hieraus folgt

$$\lambda = \frac{\sqrt{1 - x \cdot s}}{\|x - (x \cdot s) s\|} \quad (27)$$

und schließlich

$$P(x) = \frac{\sqrt{1 - x \cdot s}}{\|x - (x \cdot s) s\|} (x - (x \cdot s) s). \quad (28)$$

Diagramme zur Konstruktion der Lambert-Azimutalprojektion werden als **Schmidtsches Netz** (Abb. 8) bezeichnet.

## 8 Linearkombinationen und lineare Abhängigkeit

Die Kombination von Addition und Streckung von Vektoren wird als Linearkombination bezeichnet. Sind  $a_1, \dots, a_m$  Vektoren (d. h. Elemente eines Vektorraumes) und  $\lambda_1, \dots, \lambda_m$  reelle Zahlen, wird der Vektor

$$x = \lambda_1 a_1 + \dots + \lambda_m a_m \quad (29)$$

als **Linearkombination** der Vektoren  $a_1, \dots, a_m$  bezeichnet.

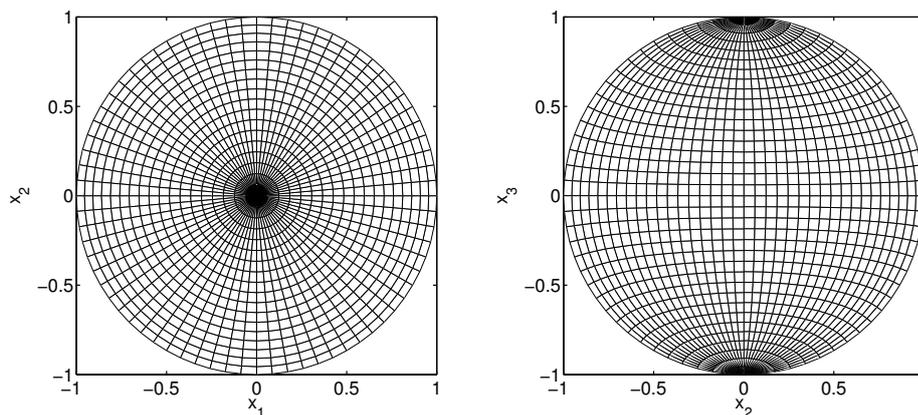


Abbildung 8: Lambert-Azimalprojektion der Hälfte des Einheitskreises (Längen- und Breitengrade in  $5^\circ$ -Schritten) mit Zentrum auf der  $x_3$ -Achse (links) und auf der  $x_1$ -Achse (rechts)

Ein Satz von Vektoren  $a_1, \dots, a_m$  heißt **linear abhängig**, wenn es eine nichttriviale Linearkombination aus ihnen gibt (nicht alle Zahlen  $\lambda_i = 0$ ), welche den Nullvektor ergibt:

$$\lambda_1 a_1 + \dots + \lambda_m a_m = 0 \quad (30)$$

Andernfalls, d. h. wenn sich diese Bedingung nur durch  $\lambda_1 = \dots = \lambda_m = 0$  erfüllen lässt, werden die Vektoren als **linear unabhängig** bezeichnet.

Es gilt: Vektoren sind genau dann linear abhängig, wenn sich mindestens einer von ihnen als Linearkombination der restlichen darstellen lässt. Dies könnte als alternative (vielleicht anschaulichere) Definition von linearer Abhängigkeit benutzt werden und ist evtl. in der Praxis etwas leichter zu überprüfen als die ursprüngliche Definition.

Linear unabhängige Vektoren sind natürlich nicht unbedingt paarweise senkrecht zueinander. Umgekehrt gilt aber:

- Jeder Satz von orthogonalen Vektoren ist linear unabhängig, sofern keiner der Vektoren der Nullvektor ist.
- Jeder Satz von orthonormalen Vektoren ist linear unabhängig.



---

**Box 7** Darstellung eines Vektors als Linearkombination anderer Vektoren in MATLAB

---

Angenommen, dass  $m$  Vektoren  $a_1, \dots, a_m$  aus  $\mathbb{R}^n$  gegeben sind, und ein weiterer Vektor  $b$  als Linearkombination dieser Vektoren dargestellt werden soll. Dann schreibt man die  $m$  Vektoren spaltenweise nebeneinander in eine  $n \times m$ -Matrix  $A$  und berechnet  $A \setminus b$ . Das Ergebnis ist ein Vektor aus  $m$  Komponenten, welcher die Koeffizienten  $\lambda_1, \dots, \lambda_m$  der Linearkombination enthält. Warum das so ist, kommt in Abschnitt 18. Achtung: Diese Vorgehensweise funktioniert nur, wenn sich der Vektor  $b$  tatsächlich als Linearkombination der anderen Vektoren schreiben lässt. Andernfalls erhält man ein falsches Ergebnis ohne Warnung. Wenn hingegen eine Warnung der Form `Warning: Rank deficient, ...` erscheint, so ist das kein Problem, sondern deutet nur darauf hin, dass die Vektoren  $a_1, \dots, a_m$  linear abhängig sind.

---

## 9 Unterräume

Eine Teilmenge eines Vektorraums  $V$ , welche selbst wieder ein Vektorraum ist, wird als **Unterraum** (des Vektorraums  $V$ ) bezeichnet.

Wenn überprüft werden soll, ob eine Teilmenge eines Vektorraums ein Unterraum ist, müssen nicht alle Eigenschaften A1–A5, S1–S3 und AS1–AS2 überprüft werden. Es genügen die Eigenschaften A1 (d. h. durch Addition zweier Vektoren aus der Teilmenge verlässt man diese nicht) und S1 (d. h. durch Streckung eines Vektors aus der Teilmenge verlässt man diese nicht).

In jedem Vektorraum gibt es die trivialen Unterräume:

- der Vektorraum selbst
- der Unterraum, der nur aus dem Nullvektor besteht.

Die Menge aller Linearkombinationen aus Vektoren  $a_1, \dots, a_m$  bildet einen Unterraum. Dieser wird **lineare Hülle** der Vektoren  $a_1, \dots, a_m$  oder als der von  $a_1, \dots, a_m$  **aufgespannte** Unterraum, kurz  $\text{span}(a_1, \dots, a_m)$ , bezeichnet.

## 10 Dimension

Die maximale Anzahl linear unabhängiger Vektoren in einem Vektorraum  $V$  wird als **Dimension** des Vektorraumes bezeichnet, abgekürzt als  $\dim(V)$ .

**Beispiele:**

- $\mathbb{R}^n$  ist  $n$ -dimensional.
- Die Menge aller Vektoren der Form  $\begin{pmatrix} x_1 \\ 0 \end{pmatrix}$  ist ein eindimensionaler Unterraum des  $\mathbb{R}^2$ .
- Die Menge der Vektoren  $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ , für die  $x_1 + x_2 + x_3 = 0$  gilt, ist ein zweidimensionaler Unterraum des  $\mathbb{R}^3$ .
- Sind  $a_1, \dots, a_m$  Vektoren, ist  $\text{span}(a_1, \dots, a_m)$  ein  $m$ -dimensionaler Unterraum, wenn die Vektoren linear unabhängig sind. Sind die Vektoren linear abhängig, ist seine Dimension kleiner als  $n$ .
- Der Vektorraum aller Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  ist unendlichdimensional und damit wesentlich komplizierter als alle Vektorräume endlicher Dimension.

Aus der Definition der Dimension ergibt sich unmittelbar, dass die Dimension eines Unterraums nicht größer als die des ursprünglich betrachteten Vektorraums sein kann. Hieraus lässt sich beispielsweise direkt festlegen, welche Unterräume  $\mathbb{R}^3$  besitzt:

- Einen 0-dimensionalen Unterraum, der nur aus dem Nullvektor  $0$  besteht.
- Unendlich viele Unterräume, die von einem Vektor aufgespannt werden, also eindimensional sind. Dies sind beliebige Geraden durch den Nullpunkt  $0$ .
- Unendlich viele Unterräume, die von zwei linear unabhängigen Vektoren aufgespannt werden, also zweidimensional sind. Dies sind beliebige Ebenen durch den Nullpunkt  $0$ .
- Den gesamten  $\mathbb{R}^3$ .

---

**Box 8** Dimension und lineare Abhängigkeit in MATLAB

---

Die Dimension des von  $m$  Vektoren aus  $\mathbb{R}^n$  aufgespannten Unterraums lässt sich in MATLAB leicht berechnen, indem die Vektoren nebeneinander in eine  $n \times m$ -Matrix  $A$  geschrieben werden. Dann liefert `rank(A)` (der Rang der Matrix) die Dimension des von den Vektoren aufgespannten Unterraums. Ist diese Dimension gleich der Anzahl der Vektoren  $m$ , sind die Vektoren linear unabhängig, andernfalls linear abhängig.

---

Unterräume der Dimension  $n-1$  eines  $n$ -dimensionalen Vektorraums werden als **Hyperebenen** bezeichnet. In jedem euklidischen Vektorraum bildet die Menge aller Vektoren, welche auf einem gegebenen Vektor  $a \neq 0$  senkrecht stehen, eine Hyperebene. Umgekehrt kann man jede Hyperebene durch einen solchen Normalenvektor darstellen.

## 11 Basen

Ein Satz von Vektoren  $x_1, x_2, x_3, \dots$  (es dürfen unendlich viele sein, streng genommen müssen sie sich nicht einmal nummerieren lassen) heißt **Basis** (des Vektorraumes, zu dem sie gehören), wenn sich jeder Vektor  $a$  (des Vektorraumes) in eindeutiger Weise als Linearkombination aus diesen darstellen lässt:

$$a = \lambda_1 x_1 + \dots + \lambda_n x_n \quad (31)$$

(Achtung: Eine Linearkombination besteht immer aus endlich vielen Vektoren, auch wenn die Basis aus unendlich vielen besteht).

Die Zahlen  $\lambda_1, \dots, \lambda_n$  heißen **Koordinaten** des Vektors  $a$  bzgl. der Basis. Man fasst diese oft zu einem Tupel  $\begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix}$  zusammen, welches als **Koordinatenvektor** bezeichnet wird.

Die einfachste Basis des  $\mathbb{R}^n$  besteht aus den  $n$  Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{pmatrix} \quad (32)$$

Sie heißt **Standardbasis** des  $\mathbb{R}^n$ .

Eigenschaften von Basen:

- Jede Basis ist linear unabhängig.
- In einem Vektorraum der (endlichen) Dimension  $n$  ist jeder Satz von  $n$  linear unabhängigen Vektoren eine Basis.

Eine Basis aus orthonormalen Vektoren wird als **Orthonormalbasis** bezeichnet. Orthonormalbasen sind in zweierlei Hinsicht einfacher als andere Basen:

1. Während die Berechnung der Koordinaten  $\lambda_1, \dots, \lambda_n$  eines Vektors  $a$  nach Gleichung 31 im Allgemeinen die Lösung eines linearen Gleichungssystems mit  $n$  Gleichungen (siehe Abschnitt 18) erfordert, gilt für Orthonormalbasen

$$a \cdot x_j = \lambda_1 x_{1j} \cdot x_j + \dots + \lambda_n x_{nj} \cdot x_j = \lambda_j. \quad (33)$$

Die Koordinate  $\lambda_j$  ergibt sich also direkt durch Berechnung des Skalarprodukts des Vektors  $a$  mit dem jeweiligen Basisvektor  $x_j$ .

2. Das Skalarprodukt zweier Vektoren berechnet sich wie das Standard-Skalarprodukt in  $\mathbb{R}^n$ :

$$a \cdot b = \lambda_1 \mu_1 + \dots + \lambda_n \mu_n,$$

wenn  $\lambda_1, \dots, \lambda_n$  die Koordinaten des Vektors  $a$  und  $\mu_1, \dots, \mu_n$  die Koordinaten des Vektors  $b$  sind. Für nicht-orthonormale Basis enthält sie Summe  $n \times n$  Summanden bzw.  $\frac{n \times (n+1)}{2}$  Summanden, wenn man berücksichtigt, dass einige gleich sind.

## 12 Reziproke Basen

Zu jeder Basis aus den Vektoren  $x_1, x_2, x_3, \dots, x_n$  gibt es (genau) eine **reziproke Basis** aus Vektoren  $r_1, r_2, r_3, \dots, r_n$ , welche die Bedingungen

$$x_i \cdot r_j = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j \end{cases} \quad (34)$$

erfüllt (Abb. 9).

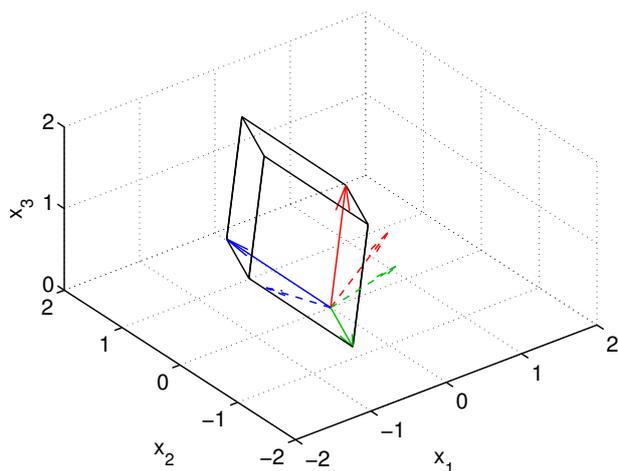


Abbildung 9: Beispiel einer Basis des  $\mathbb{R}^3$  (durchgezogene Pfeile), der von den Basisvektoren gebildeten Elementarzelle (alle Linearkombinationen mit 0 oder 1 als Skalierungsfaktoren) und der reziproken Basis (gestrichelte Pfeile)

Mit Hilfe der reziproken Basis lassen sich Skalarprodukte leicht berechnen. Wird ein Vektor  $a$  bzgl. der ursprünglichen Basis dargestellt und ein Vektor  $b$  bzgl. der reziproken Basis, also

$$\begin{aligned} a &= \lambda_1 x_1 + \dots + \lambda_n x_n \\ b &= \mu_1 r_1 + \dots + \mu_n r_n \end{aligned}$$

so folgt:

$$a \cdot b = \lambda_1 \mu_1 + \dots + \lambda_n \mu_n$$

Bei einer Orthonormalbasis ist die reziproke Basis identisch mit der ursprünglichen Basis. Bei anderen Basen ist für die Bestimmung der reziproken Basis hingegen ein lineares Gleichungssystem zu lösen (siehe Abschnitt 18). In  $\mathbb{R}^3$  lässt sich dies mit Hilfe des Vektorprodukts vereinfachen.

## 13 Lineare Abbildungen

Eine Abbildung  $F$  von einem Vektorraum in einen anderen Vektorraum (es kann auch derselbe Vektorraum wie der erste sein) heißt **linear**, wenn sie

sich in folgender Weise mit der Addition und der Streckung von Vektoren „verträgt“:

**L1** Für alle Vektoren  $a$  und  $b$  ist  $F(a + b) = F(a) + F(b)$ .

**L2** Für alle Vektoren  $a$  und alle reellen Zahlen  $\lambda$  ist  $F(\lambda a) = \lambda F(a)$ .

Hieraus folgen einige weitere Eigenschaften, z. B.  $F(0) = 0$ .

Man kann auch beide Kriterien zusammenfassen: Für alle Vektoren  $a$  und  $b$  und alle reellen Zahlen  $\lambda$  und  $\mu$  ist

$$F(\lambda a + \mu b) = \lambda F(a) + \mu F(b) \quad (35)$$

Dies bedeutet, dass diejenigen Abbildungen linear sind, die sich mit der Linearkombination von Vektoren „vertragen“.

Grundsätzlich können wir natürlich Abbildungen von irgendeinem Vektorraum in irgendeinen Vektorraum betrachten. Unter diesen sind zwei Spezialfälle besonders wichtig:

- Lineare Abbildungen eines Vektorraums in den einfachsten nichttrivialen (d. h. der nicht nur aus einem Nullvektor besteht) Vektorraum, nämlich in die reellen Zahlen selbst. Diese werden als **Linearformen** bezeichnet.
- Lineare Abbildungen, welche von einem Vektorraum in diesen selbst gehen (also nicht von einem Vektorraum in einen anderen). Diese werden als **Endomorphismen** bezeichnet und sind ein Beispiel dafür, dass nicht nur Geologen furchtbare Fachwörter verwenden.

#### Beispiele:

- $F\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$  definiert eine lineare Abbildung in  $\mathbb{R}^2$ .
- $F\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + x_3 \\ x_2 \\ x_3 \end{pmatrix}$  definiert eine lineare Abbildung in  $\mathbb{R}^3$ .
- Ist  $a$  ein gegebener Vektor eines euklidischen Vektorraums, so ist  $F(x) = a \cdot x$  eine Linearform, nicht jedoch die Abbildung  $F(x) = x \cdot x$ .

- Die Abbildung  $F(x) = \|x\|$  von einem normierten Vektorraum in die reellen Zahlen ist nicht linear.
- Die **Nullabbildung**, welche jeden Vektor auf den Nullvektor abbildet, und die **Identität**, die jeden Vektor auf sich selbst abbildet (also unverändert lässt), sind die einfachsten linearen Abbildungen in jedem Vektorraum. Die werden oftmals mit den Symbolen **0** und **1** gekennzeichnet. Da aber ihre Wirkung dieselbe wie die Streckung um einen Faktor 0 oder 1 ist, kann man genausogut die Symbole 0 und 1 verwenden.
- Die Streckung um einen festen Faktor  $\lambda$ , d. h.  $F(x) = \lambda x$  ist ebenfalls eine lineare Abbildung in jedem Vektorraum.
- Die Verschiebung um einen festen Vektor  $a \neq 0$ , d. h.  $F(x) = x + a$  ist keine lineare Abbildung.
- Ist  $\omega$  ein gegebener Vektor in  $\mathbb{R}^3$ , so ist  $F(x) = \omega \times x$  eine lineare Abbildung. Diese bildet den Ort  $x$  auf die Geschwindigkeit  $v$  einer Drehbewegung bei gegebenem Winkelgeschwindigkeitsvektor  $\omega$  ab.
- Die Parallelprojektion (Gl. 17) ist linear, alle anderen in den Abschnitten 6 und 7 behandelten Projektionen sind allerdings nicht linear.
- Im Vektorraum der Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  ist die Ableitung (die also eine Funktion  $f(x)$  auf ihre Ableitung  $f'(x)$  abbildet) eine lineare Abbildung (stimmt aber nicht ganz).

Die Abbildungen (nicht nur die linearen, sondern alle) eines Vektorraums  $V$  in denselben Vektorraum oder in einen anderen Vektorraum  $W$  bilden wieder einen Vektorraum. Die linearen Abbildungen bilden einen Unterraum des Vektorraums aller Abbildungen. Lineare Abbildungen stellen also selbst wieder Vektoren dar.

## 14 Lineare Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}^2$ und von $\mathbb{R}^3$ nach $\mathbb{R}^3$

Lineare Abbildungen von  $\mathbb{R}^2$  nach  $\mathbb{R}^2$  bzw.  $\mathbb{R}^3$  nach  $\mathbb{R}^3$  haben eine direkte geometrische Bedeutung. In der Strukturgeologie beschreiben diese eine homogene Deformation eines Gesteins.

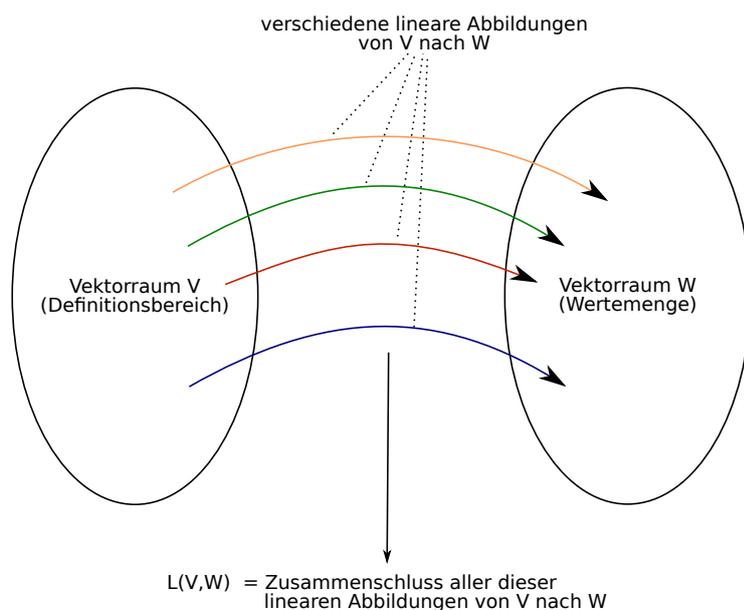


Abbildung 10: Illustration des Vektorraums der linearen Abbildungen von einem Vektorraum in einen anderen Vektorraum. Quelle: [https://de.wikipedia.org/wiki/Lineare\\_Abbildung](https://de.wikipedia.org/wiki/Lineare_Abbildung)

Grundsätzlich würde eine homogene Deformation eines Gesteins zusätzlich eine Translation um einen beliebigen Vektor  $a$  beinhalten, also die Form  $F(x) + a$  haben und damit keine lineare Abbildung sein. Hier vernachlässigen wir aber die Translation, weil sie sich in der Geologie nicht rekonstruieren lässt und für die eigentliche Deformation nicht relevant ist.

Da der Fall  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  weniger Schreibarbeit als der Fall  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$  bedeutet und anschaulicher ist, beschränken wir uns zunächst auf diesen. Alles weitere gilt aber analog in  $\mathbb{R}^n$ .

Im Folgenden leiten wir eine einfache Darstellung beliebiger linearer Abbildungen von  $\mathbb{R}^2$  nach  $\mathbb{R}^2$  her. Jeder Vektor  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  lässt sich als Linearkombination der Vektoren  $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  und  $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  darstellen:

$$x = x_1 e_1 + x_2 e_2$$

Für jede lineare Abbildung  $F$  gilt daher

$$F(x) = F(x_1 e_1 + x_2 e_2) = x_1 F(e_1) + x_2 F(e_2) \quad (36)$$

Nennen wir

$$F(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \quad \text{und} \quad F(e_2) = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \quad (37)$$

können wir schreiben:

$$F(x) = x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \quad (38)$$

Jede lineare Abbildung von  $\mathbb{R}^2$  nach  $\mathbb{R}^2$  kann also durch Angabe von vier Zahlen charakterisiert werden.

## 15 Matrizen

Für die Darstellung einer linearen Abbildung aus dem letzten Abschnitt gibt es eine Kurzschreibweise. Man ordnet die vier Zahlen  $a_{11}$ ,  $a_{12}$ ,  $a_{21}$  und  $a_{22}$  als  $2 \times 2$ -**Matrix** an:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (39)$$

und definiert

$$Ax = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \quad (40)$$

$A$  heißt **Matrix zu  $F$** . Die Anwendung von  $F$  auf einen Vektor  $x$  kann also dargestellt werden als Multiplikation der Matrix  $A$  mit  $x$ :

$$F(x) = Ax \quad (41)$$

Die erste Spalte der Matrix ist also der Vektor, der herauskommt, wenn man  $F$  auf  $e_1$  anwendet, die zweite Spalte ist  $F(e_2)$ .

Zumindest in  $\mathbb{R}^2$  können wir also lineare Abbildungen und Matrizen miteinander identifizieren und müssen diese nicht strikt unterscheiden. Die Begriffe, die wir im Folgenden kennenlernen, sind auf lineare Abbildungen und Matrizen gleichwertig anwendbar.

Zur Nullabbildung gehört natürlich die **Nullmatrix**  $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ . Zur Identität gehört die **Einheitsmatrix**  $E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ .

In beliebigen (endlichdimensionalen) Vektorräumen ist der Begriff der Matrix etwas komplizierter und an eine konkrete Basis gebunden. Angenommen, diese Basis besteht aus den  $x_1, \dots, x_n$ . Dann wird die Abbildung  $F$  auf den ersten Basisvektor ( $x_1$ ) angewandt und das Ergebnis als Linearkombination der Basisvektoren geschrieben:

$$F(x) = \mu_1 x_1 + \dots + \mu_n x_n \quad (42)$$

Die Koordinaten  $\mu_1, \dots, \mu_n$  bilden dann die erste Spalte der Matrix  $A$ . Die Anwendung von  $F$  auf die weiteren Basisvektoren liefert entsprechend die weiteren Spalten von  $A$ .

Die Bedeutung der Matrix ist grundsätzlich dieselbe wie in  $\mathbb{R}^2$ . Hat ein Vektor  $x$  die Koordinaten  $\lambda_1, \dots, \lambda_n$  bzgl. einer gegebenen Basis  $x_1, \dots, x_n$ , so gilt für die Koordinaten  $\mu_1, \dots, \mu_n$  des Ergebnisvektors  $F(x)$ :

$$\begin{pmatrix} \mu_1 \\ \dots \\ \mu_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} = A \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} \quad (43)$$

Zu beachten ist, dass die Matrix  $A$  sowohl von der linearen Abbildung  $F$  als auch von der gewählten Basis abhängt.

Grundsätzlich lässt sich dies auch auf lineare Abbildungen von einem Vektorraum in einen anderen Vektorraum verallgemeinern, wobei dann zwei Basen benötigt werden.

## 16 Verkettung linearer Abbildungen und Matrixmultiplikation

$2 \times 2$ -Matrizen lassen sich (wie Matrizen beliebiger Größe) nach naheliegenden Regeln addieren und mit Zahlen multiplizieren:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$
$$\lambda \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix}$$

Da diese Addition und Multiplikation mit Zahlen alle in Abschnitt 1 aufgestellten Bedingungen an einen Vektorraum erfüllt, bilden die  $2 \times 2$ -Matrizen selbst wieder einen Vektorraum. Seine Dimension ist vier. Dasselbe gilt für

die linearen Abbildungen des  $\mathbb{R}^2$  mit der Addition und Multiplikation mit Zahlen:

$$(F + G)(x) = F(x) + G(x) \quad \text{und} \quad (\lambda F)(x) = \lambda F(x)$$

Mit linearen Abbildungen kann man aber mehr machen als Addition und Multiplikation mit Zahlen: Man kann zwei (oder auch mehrere) lineare Abbildungen hintereinander ausführen, wobei sich wieder eine lineare Abbildung ergibt. Im Fall von zwei linearen Abbildungen  $F$  und  $G$  lässt sich dies leicht einsehen:

$$F(G(\lambda a + \mu b)) = F(\lambda G(a) + \mu G(b)) = \lambda F(G(a)) + \mu F(G(b))$$

Die Hintereinanderausführung – auch als **Verkettung** bezeichnet – kann als eine Multiplikation angesehen werden. Diese wird entweder mit einem Kreis  $F \circ G$  oder ohne Multiplikationssymbol  $FG$  geschrieben. Lineare Abbildungen sind also Vektoren, die sich miteinander so multiplizieren lassen, dass das Ergebnis wieder ein Vektor ist. Für diese Multiplikation gelten einige der Regeln, die für die Multiplikation von Zahlen gelten. Zwei Eigenschaften der Multiplikation reeller Zahlen gelten für die Verkettung linearer Abbildungen allerdings nicht:

- Die Verkettung linearer Abbildungen ist nicht kommutativ, d. h. i. A. ist  $FG \neq GF$  ( $FG$  heißt zuerst  $G$ , danach  $F$ ).
- Während es zu jeder reellen Zahl  $x \neq 0$  eine inverse Zahl  $\frac{1}{x}$  (auch als  $x^{-1}$  bezeichnet) gibt, sodass  $x \frac{1}{x} = 1$ , gibt es nur für einige lineare Abbildungen  $F$  eine **Inverse**  $F^{-1}$  sodass  $GF = FG = 1$  ( $1 = \text{Identität}$ ). Solche Abbildungen werden als **invertierbar** bezeichnet.

Da die Verkettung zweier linearer Abbildungen wieder linear ist, besitzt (zumindest) in  $\mathbb{R}^2$  auch diese eine Matrix. Sei  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$  die Matrix zur Abbildung  $F$ , und  $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$  die Matrix zur Abbildung  $G$ . Dann gilt:

$$\begin{aligned} F(G(x)) &= A(Bx) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11}x_1 + b_{12}x_2 \\ b_{21}x_1 + b_{22}x_2 \end{pmatrix} \\ &= \begin{pmatrix} (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2 \\ (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} x \end{aligned}$$



Die Matrix der Verkettung  $FG$  ist also das folgende Produkt der Matrizen  $A$  und  $B$ :

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Diese Multiplikation wird als **Matrixmultiplikation** bezeichnet. Man kann sie sich beispielsweise auf folgende Arten einprägen:

- $i$ -te Zeile von  $A$  mal  $j$ -te Spalte von  $B$  ergibt das Element in der  $i$ -ten Zeile und der  $j$ -ten Spalte. Kurz: Zeile mal Spalte.
- Multiplikation von  $A$  mit der  $j$ -ten Spalte von  $B$  liefert die  $j$ -te Spalte der Produktmatrix.

Für die Matrixmultiplikation gelten natürlich dieselben Regeln wie für die Verkettung linearer Abbildungen, d. h. also auch dieselben wie für die Multiplikation reeller Zahlen, außer:

- Die Matrixmultiplikation ist nicht kommutativ, d. h. i. A. ist  $AB \neq BA$ .
- Es gibt nicht zu jeder Matrix  $A \neq 0$  eine Inverse.

## 17 Bild, Kern und Rang von linearen Abbildungen

Ist  $F$  eine lineare Abbildung von einem Vektorraum  $V$  in einen anderen (oder denselben) Vektorraum  $W$ , wird die Teilmenge von  $W$ , welche tatsächlich von  $F$  erreicht wird, als das **Bild** von  $F$  bezeichnet, abgekürzt als  $\text{im}(F)$  (leider dasselbe Symbol wie für den sogenannten Imaginärteil von komplexen Zahlen).

Das Bild einer linearen Abbildung,  $\text{im}(F)$ , bildet nicht nur eine Teilmenge von  $W$ , sondern einen Unterraum von  $W$ . Seine Dimension wird als **Rang** von  $F$  bezeichnet, abgekürzt als  $\text{rank}(F)$ .

Der **Kern** einer linearen Abbildung  $F$ ,  $\text{ker}(F)$ , ist die Teilmenge von  $V$ , welche auf den Nullvektor abgebildet wird:  $F(x) = 0$ . Alternativ wird der Kern auch als **Nullraum** bezeichnet.

Für das Bild und den Kern einer linearen Abbildung gilt die Dimensionsformel

$$\dim(V) = \dim(\text{ker}(F)) + \dim(\text{im}(F)) = \dim(\text{ker}(F)) + \text{rank}(F) \quad (44)$$

---

**Box 9** Bild, Rang und Kern von Matrizen in MATLAB

---

Der Rang einer Matrix  $A$  lässt sich mittels `rank(A)` berechnen, was wir bereits vorher für die Bestimmung linearer Unabhängigkeit verwendet haben. Darüber hinaus liefert `orth(A)` eine Orthonormalbasis (bzgl. der Standard-Skalarprodukts) des Bildes von  $A$  und `null(A)` eine Orthonormalbasis des Kerns von  $A$ . Auf diese Art lässt sich u. a. leicht herausfinden, wie man linear abhängige Vektoren zum Nullvektor kombinieren kann.

---

Nehmen wir  $V = \mathbb{R}^n$  und  $W = \mathbb{R}^m$  und die jeweiligen Standardbasen, lassen sich die Begriffe Bild, Rang und Kern sowie die Dimensionsformel sofort auf  $m \times n$ -Matrizen ( $m$  Zeilen,  $n$  Spalten) übertragen. Dabei kann der Rang der Matrix als maximale Anzahl linear unabhängigen Spalten, die man aus der Matrix auswählen kann, interpretieren.

## 18 Lineare Gleichungssysteme

Jedes lineare Gleichungssystem mit  $m$  Gleichungen der Form

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ \dots + \dots + \dots &= \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m \end{aligned} \quad (45)$$

für die  $n$  Unbekannten  $x_1, \dots, x_n$  lässt sich in Matrixform

$$Ax = b \quad (46)$$

schreiben mit der  $m \times n$ -Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad (47)$$

und der rechten Seite

$$b = \begin{pmatrix} b_1 \\ \dots \\ b_m \end{pmatrix}. \quad (48)$$

Ist  $b = 0$ , wird das Gleichungssystem als **homogen** bezeichnet, ansonsten als **inhomogen**.

Ob das lineare Gleichungssystem eine Lösung besitzt, hängt nicht nur von der Matrix  $A$  ab, sondern auch von der rechten Seite  $b$ . Falls  $\text{rank}(A) = m$

---

**Box 10** Lösung linearer Gleichungssysteme in MATLAB

---

Der Standardoperator für das Lösen eines linearen Gleichungssystems der Form  $Ax = b$  ist die sogenannte Linksdivision  $\backslash$ . Die Lösung wird bestimmt durch  $x = A \backslash b$ . Als Eselsbrücke kann man sich das so merken, dass beide Seiten des Gleichungssystems durch die Matrix geteilt werden, die Division aber von links erfolgen muss, weil die Matrix im Ausdruck  $Ax$  links steht. Bei der Lösung eines linearen Gleichungssystems auf diese Weise sind aber zwei Dinge zu beachten:

- (i) Es wird immer nur eine Lösung geliefert, auch wenn das Gleichungssystem mehrere Lösungen besitzt. Die Gesamtheit aller Lösungen lässt sich bei Bedarf über den Kern von  $A$  bestimmen (`null(A)`).
  - (ii) Auch wenn das Gleichungssystem keine Lösung besitzt, liefert  $x = A \backslash b$  ein Ergebnis (siehe Abschnitt 19). Jedenfalls empfiehlt es sich, im Zweifelsfall z. B. mit `norm(A*x-b)` zu überprüfen, ob wir eine (hinreichend genaue) Lösung erhalten haben.
- 

(Anzahl der Zeilen) ist, umfasst das Bild von  $A$  den gesamten  $\mathbb{R}^m$ , d. h. das lineare Gleichungssystem besitzt dann für jede rechte Seite  $b$  eine Lösung.

Die Eindeutigkeit der Lösung (sofern es eine Lösung gibt) hängt dagegen nur von der Matrix  $A$  und nicht von der rechten Seite  $b$  ab. Besteht der Kern  $A$  nur aus dem Nullvektor, besitzt das homogene Gleichungssystem  $Ax = 0$  nur den Nullvektor als Lösung. In diesem besitzt das inhomogene Gleichungssystem  $Ax = b$  höchstens eine Lösung.

Umfasst der Kern von  $A$  nicht nur den Nullvektor, d. h. wenn es Vektoren  $\tilde{x}$  gibt, sodass  $A\tilde{x} = 0$ , so kann man diese Vektoren zu jeder Lösung  $x$  hinzuaddieren und erhält wieder eine Lösung, denn

$$A(x + \tilde{x}) = Ax + A\tilde{x} = b + 0 = b \quad (49)$$

Die Eindeutigkeit der Lösung (sofern es eine Lösung gibt) ist also an die Bedingung gebunden, dass der Kern von  $A$  nur den Nullvektor enthält. Nach der Dimensionsformel (Gl. 44) bedeutet dies  $\text{rank}(A) = n$ .

## 19 Die Methode der kleinsten Quadrate

Nach den Betrachtungen zur Lösbarkeit aus dem letzten Abschnitt kann ein lineares Gleichungssystem  $Ax = b$  nicht für jede rechte Seite lösbar sein, wenn die Anzahl der Gleichungen größer als die Anzahl der Unbekannten ist ( $m > n$ , überbestimmtes Gleichungssystem). In diesem Fall kann es

---

**Box 11** Überbestimmte Gleichungssysteme in MATLAB

---

Besitzt ein lineares Gleichungssystem der Form  $Ax = b$  keine Lösung, so liefert die Linksdivision  $x = A \setminus b$  automatisch die beste Lösung im Sinne der euklidischen Norm. Man muss sich also nicht selbst um die Normalgleichungen kümmern.

---

trotzdem sinnvoll sein, einen Vektor  $x$  so zu bestimmen, dass die Norm der verbleibenden Abweichung  $Ax - b$  minimal wird.

Dies geht am einfachsten, wenn wir die euklidische Norm verwenden. In diesem Fall ergibt sich die Lösung dadurch, dass die verbleibende Abweichung  $Ax - b$  im Sinne der euklidischen Norm senkrecht auf dem Bild von  $A$  steht. Dies lässt sich formal durch die Normalgleichungen

$$A^T Ax = A^T b \quad (50)$$

ausdrücken, wobei  $A^T$  die **Transponierte** der Matrix  $A$  ist (Zeilen und Spalten vertauscht, d. h.,  $(A^T)_{ij} = A_{ji}$ ). Dieses Gleichungssystem besteht aus  $n$  Gleichungen für  $n$  Unbekannte und besitzt immer (mindestens) eine Lösung.

Da die euklidische Norm die (Wurzel aus der) Summe der Quadrate der Abweichungen beschreibt, bezeichnet man diese Methode auch als Methode der kleinsten Quadrate.

## 20 Linearer Fit

Angenommen, wir haben  $m$  Messwerte  $u_1, u_2, \dots, u_m$ , welche zu verschiedenen Zeiten  $t_1, t_2, \dots, t_m$  (oder an verschiedenen Orten oder unter irgendwie verschiedenen Bedingungen) gemessen wurden. Nehmen wir weiter an, dass wir eine Vorstellung davon haben, wie die Daten zustande gekommen sind, also dass diese theoretisch durch eine Funktion  $f(t)$  beschrieben werden. Mögliche Funktionen könnten beispielsweise sein:

**eine Gerade**  $f(t) = at + b$

**ein Polynom**  $n$ -ten Grades  $f(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$

**eine Exponentialfunktion**  $f(t) = a e^{-\lambda t}$

**eine harmonische Schwingung**  $f(t) = a + b \cos(\omega t) + c \sin(\omega t)$

Alle diese Funktionen enthalten anpassbare **Parameter**. Meist ist jedoch die Anzahl dieser Parameter viel geringer als die Anzahl  $m$  der Messwerte, sodass das Gleichungssystem

$$f(t_1) = u_1, \quad f(t_2) = u_2, \quad \dots, \quad f(t_m) = u_m \quad (51)$$

überbestimmt und damit i. A. nicht lösbar ist. Demnach besteht das Ziel darin, die Parameter so zu bestimmen, dass die Abweichung zwischen den theoretischen Werten  $f(t_i)$  und den Messwerten  $u_i$  möglichst gering ist. Hierfür fasst man die theoretischen Werte und die Messwerte jeweils zu Vektoren  $f$  und  $u$  aus  $\mathbb{R}^m$  zusammen und stellt die Bedingung, dass  $\|f - u\|$  minimal werden soll. Hierbei kann eine beliebige Norm verwendet werden. Die Eigenschaften verschiedener Normen in diesem Zusammenhang wurden in Aufgabe 2 untersucht.

Im einfachsten Fall hängen die theoretischen Werte linear von den Parametern ab, d. h. die theoretischen Werte sind eine Linearkombination gegebener Funktionen  $\varphi_1(t), \dots, \varphi_n(t)$ :

$$f(t) = \lambda_1 \varphi_1(t) + \dots + \lambda_n \varphi_n(t) \quad (52)$$

wobei  $\lambda_1, \dots, \lambda_n$  die zu bestimmenden Parameter sind. In diesem Fall spricht man von einem **linearen Fit**. Man beachte, dass dies nicht bedeutet, dass auch die Funktionen  $\varphi_i(t)$  linear sein müssen. Auch die Anpassung eines Polynoms beliebigen Grades ist ein lineares Minimierungsproblem.

Dies lässt sich auch schreiben als

$$A\lambda = u \quad (53)$$

mit

$$A = \begin{pmatrix} \varphi_1(t_1) & \dots & \varphi_n(t_1) \\ \dots & \dots & \dots \\ \varphi_1(t_m) & \dots & \varphi_n(t_m) \end{pmatrix} \quad \text{und} \quad \lambda = \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} \quad (54)$$

Somit wäre im Idealfall das lineare Gleichungssystem  $A\lambda = u$  zu lösen bzw. da dies i. A. nicht geht, ist  $\lambda$  so zu bestimmen, dass  $\|A\lambda - u\|$  minimal wird. Die Vorgehensweise hierfür bei Verwendung der euklidischen Norm wurde in Abschnitt 19 beschrieben. Bei Verwendung anderer Normen ist der lineare Fit allerdings komplizierter (siehe Aufgabe 14).

## 21 Matrizen unter Basiswechsel

Angenommen, eine lineare Abbildung  $F$  des  $\mathbb{R}^n$  hat die Matrix  $A$  bzgl. der Standardbasis, und wir wollen deren Matrix  $\tilde{A}$  bzgl. einer gegebenen

anderen Basis bestimmen. Wie in Abschnitt 15 beschrieben, müssen wir hierfür  $F$  auf die jeweiligen Basisvektoren  $b_1, \dots, b_n$  anwenden und die Koordinaten der erhaltenen Vektoren  $F(b_1), \dots, F(b_n)$  bestimmen. Diese Koordinaten bilden schließlich die Spalten der neuen Matrix  $\tilde{A}$ . Schreiben wir die Vektoren  $b_1, \dots, b_n$  nebeneinander in eine Matrix  $B$ , so lässt sich die Bestimmung der Koordinaten  $\lambda_1, \dots, \lambda_n$  eines beliebigen Vektors  $x$  ein lineares Gleichungssystem der Form

$$B\lambda = x. \quad (55)$$

Da die Basisvektoren linear unabhängig sind, ist die Matrix  $B$  invertierbar, und die Lösung des Gleichungssystems ist

$$\lambda = B^{-1}x. \quad (56)$$

Dies müssen wir nun auf die Vektoren  $F(b_1) = Ab_1, \dots, F(b_n) = Ab_n$  anwenden, woraus sich die  $i$ -te Spalte der neuen Matrix  $\tilde{A}$  als  $B^{-1}Ab_i$  ergibt. Schreiben wir nun alle Spalten nebeneinander, so erhalten wir die **Basiswechselformel**

$$\tilde{A} = B^{-1}AB. \quad (57)$$

## 22 Spur und Determinante

Die **Spur** einer quadratischen  $n \times n$ -Matrix  $A$  ist definiert als die Summe der Diagonalelemente

$$\text{tr}(A) = a_{11} + a_{22} + \dots + a_{nn}, \quad (58)$$

wobei  $\text{tr}$  für trace steht. Die Spur ist eine Linearform des Vektorraumes der  $n \times n$ -Matrizen, d. h. es gilt:

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B) \quad \text{und} \quad \text{tr}(\lambda A) = \lambda \text{tr}(A) \quad (59)$$

Darüber hinaus gilt:

$$\text{tr}(AB) = \text{tr}(BA) \quad (\text{aber} \neq \text{tr}(A)\text{tr}(B)) \quad (60)$$

Mit Hilfe der Basiswechselformel (Gl. 57) folgt daraus

$$\text{tr}(\tilde{A}) = \text{tr}(B^{-1}AB) = \text{tr}(ABB^{-1}) = \text{tr}(A) \quad (61)$$

Dies bedeutet, dass die Spur der Matrix zu einer linearen Abbildung von der gewählten Basis unabhängig ist. Daraus definiert man die Spur einer

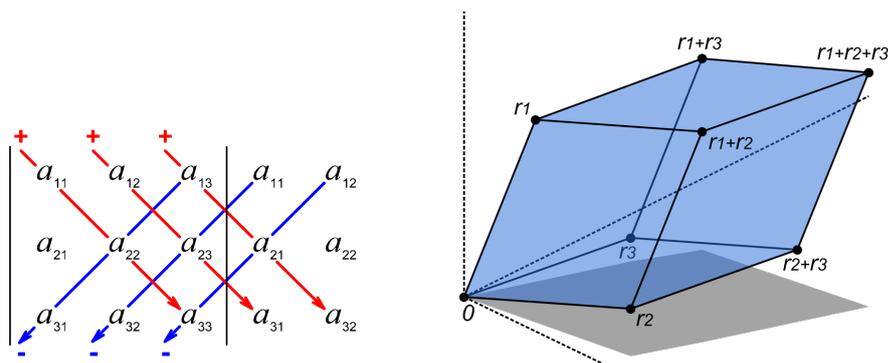


Abbildung 11: Berechnung der Determinante einer  $3 \times 3$ -Matrix und Interpretation als Volumen eines Parallelepipeds. Quellen: [https://de.wikipedia.org/wiki/Regel\\_von\\_Sarrus](https://de.wikipedia.org/wiki/Regel_von_Sarrus) (von Eisenbahn%*s* - Eigenes Werk, CC-BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=46339927>), <https://en.wikipedia.org/wiki/Determinant> (by Claudio Rocchini - Own work, CC BY 3.0, <https://commons.wikimedia.org/w/index.php?curid=2788190>).

linearen Abbildung  $F$  in einem (endlichdimensionalen) Vektorraum als die Spur ihrer Matrix bzgl. einer beliebigen Basis.

Die **Determinante** einer  $2 \times 2$ -Matrix ist definiert als

$$\det A = a_{11}a_{22} - a_{12}a_{21} \quad (62)$$

und die einer  $3 \times 3$ -Matrix als

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \quad (63)$$

(Abb. 11). Dies lässt sich auch auf größere quadratische Matrizen verallgemeinern, wobei dann alle Permutationen der Spalten zu betrachten sind (was bei einer  $4 \times 4$ -Matrix schon  $4! = 24$  sind und schnell sehr viel wird).

Die Eigenschaften einer Matrix hängen fundamental davon ab, ob ihre Determinante ungleich 0 ist. Hierbei sind die folgenden Kriterien äquivalent:

1.  $\det(A) \neq 0$
2.  $\text{rank}(A) = n$
3.  $\ker(A) = \{0\}$

---

**Box 12** Spur und Determinante in MATLAB

MATLAB besitzt entsprechende Funktionen mit den Namen `trace` und `det`.

---

4.  $A$  ist invertierbar.
5. Die Spalten von  $A$  sind linear unabhängig.

Die wichtigsten Eigenschaften der Determinante sind:

- Die Determinante einer Matrix ändert sich nicht, wenn man ein Vielfaches einer Spalte zu einer anderen Spalte addiert.
- Für  $2 \times 2$  Matrizen ist  $|\det A|$  die Fläche des von den beiden Spalten (als Vektoren des  $\mathbb{R}^2$  interpretiert) aufgespannten Parallelogramms, und für  $3 \times 3$ -Matrizen das Volumen des von den drei Spalten aufgespannten Parallelepipeds (Abb. 11).
- $\det(AB) = \det(A) \det(B)$

Aus der letzten Eigenschaft folgt, wie schon bei der Spur, dass die Determinante der Matrix zu einer linearen Abbildung von der gewählten Basis unabhängig ist:

$$\det(\tilde{A}) = \det(B^{-1}AB) = \det(B^{-1}) \det(A) \det(B) = \det(B^{-1}B) \det(A) = \det(A). \quad (64)$$

Daraus definiert man die Determinante einer linearen Abbildung  $F$  in einem (endlichdimensionalen) Vektorraum als die Determinante ihrer Matrix bzgl. einer beliebigen Basis. Damit ist  $|\det(F)|$  die relative Volumenänderung, die  $F$  verursacht. Ist  $|\det(F)| < 1$ , verkleinert  $F$  Volumina, ist  $|\det(F)| > 1$ , vergrößert  $F$  Volumina. Entsprechend sind die linearen Abbildungen mit  $\det(F) = \pm 1$  diejenigen, bei denen Volumina unverändert bleiben. In der Strukturgeologie werden die linearen Abbildungen mit  $\det(F) = 1$  als **Scherungen** bezeichnet.

Im Gegensatz zur Spur ist die Determinante ist allerdings keine lineare Abbildung, i. A. ist

$$\det(A + B) \neq \det(A) + \det(B) \quad \text{und} \quad \det(\lambda A) \neq \lambda \det(A)$$

## 23 Eigenwerte

Besondere Bedeutung für eine lineare Abbildung  $F$  haben die Vektoren  $x \neq 0$ , welche durch die Abbildung nur gestreckt bzw. gestaucht werden (d. h. anschaulich, deren Richtung sich nicht ändert):

$$F(x) = \lambda x \quad (65)$$

Solche Vektoren werden als **Eigenvektoren** von  $F$  (oder der zugehörigen Matrix  $A$ ) bezeichnet. In der Strukturgeologie werden diese als **nichtrotierende Linien** bezeichnet. Der Streckungsfaktor  $\lambda$  wird als **Eigenwert** bezeichnet.

Die Eigenvektoren zu einem gegebenen Eigenwert  $\lambda$  bilden (wenn man den Nullvektor, der streng genommen kein Eigenvektor ist, hinzunimmt) einen Unterraum, der als **Eigenraum** zum Eigenwert  $\lambda$  bezeichnet wird. Diese Eigenräume können ein- oder mehrdimensional sein.

Die Eigenwertbedingung (Gl. 65) lässt sich umschreiben als

$$(F - \lambda \mathbf{1})x = 0, \quad (66)$$

was bedeutet, der Eigenraum zu einem beliebigen Wert  $\lambda$  der Kern der linearen Abbildung  $F(x) - \lambda \mathbf{1}$  ist. Demnach ist  $\lambda$  ein Eigenwert von  $F$ , wenn der Kern von  $F(x) - \lambda \mathbf{1}$  nicht nur aus dem Nullvektor besteht. Dies heißt wiederum, dass

$$\det(F - \lambda \mathbf{1}) = 0 \quad (67)$$

sein muss. Diese Größe wird als **charakteristisches Polynom** von  $F$  bezeichnet:

$$P_F(\lambda) = \det(F - \lambda \mathbf{1}) \quad (68)$$

Die Eigenwerte einer linearen Abbildung  $F$  sind also die Nullstellen ihres charakteristischen Polynoms  $P_F(\lambda)$ .

Ist  $F$  eine lineare Abbildung eines  $n$ -dimensionalen Vektorraums, so ist  $P_F(\lambda)$  ein Polynom  $n$ -ten Grades des Form:

$$P_F(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{tr}(F) \lambda^{n-1} + \dots + \det(F) \quad (69)$$

Die  $n$  Koeffizienten des Polynoms außer  $(-1)^n$  (also  $(-1)^{n-1} \text{tr}(F)$ ,  $\det(F)$ ) und  $n - 2$  kompliziertere Ausdrücke) werden als die **Invarianten** von  $F$  bezeichnet.

Im Idealfall besitzt der betrachtete Vektorraum sogar eine Basis aus Eigenvektoren von  $F$ . Dann ist die Matrix von  $F$  bzgl. dieser Basis eine **Diagonalmatrix**

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix} \quad (70)$$

wobei die Diagonalelemente  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $F$  sind. Hieraus ergibt sich sofort

$$\begin{aligned} \operatorname{tr}(F) &= \lambda_1 + \lambda_2 + \dots + \lambda_n \\ \det(F) &= \lambda_1 \lambda_2 \dots \lambda_n \end{aligned}$$

Nach der Basiswechselformel (Gl. 57) bedeutet Gleichung 70

$$B^{-1}AB = D, \quad (71)$$

wobei  $A$  die Matrix von  $F$  bzgl. der Standardbasis ist, und die Spalten von  $B$  die Basis aus Eigenvektoren von  $F$  enthalten. Solche Matrizen  $A$  werden als **diagonalisierbar** bezeichnet. Demnach lässt sich jede diagonalisierbare Matrix  $A$  darstellen als

$$A = BDB^{-1}. \quad (72)$$

Für lineare Abbildungen des  $\mathbb{R}^2$  (bzw.  $2 \times 2$ -Matrizen) können die folgenden Fälle auftreten:

1.  $P_F(\lambda)$  besitzt zwei voneinander verschiedene Nullstellen  $\lambda_1$  und  $\lambda_2$ , d. h.,  $F$  besitzt zwei verschiedene Eigenwerte. Die entsprechenden Eigenvektoren bilden eine Basis des  $\mathbb{R}^2$  (welche aber nicht unbedingt orthogonal ist), und die Matrix von  $F$  bzgl. dieser Basis hat die Form

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

2.  $P_F(\lambda)$  besitzt keine Nullstellen, d. h.,  $F$  besitzt keine Eigenwerte (z. B. Drehung).
3.  $P_F(\lambda)$  besitzt eine (doppelte) Nullstelle  $\lambda$ . Dann gibt es zwei Fälle:
  - (a) Der Eigenraum zu  $\lambda$  ist zweidimensional, d. h., der gesamte  $\mathbb{R}^2$ . In diesem Fall ist  $F$  einfach eine Streckung um den Faktor  $\lambda$  und hat bzgl. jeder Basis die Matrix

$$D = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

---

**Box 13** Eigenwerte und Eigenvektoren in MATLAB

Eigenwerte und zugehörige Eigenvektoren einer Matrix werden in MATLAB mit der Funktion `eig` berechnet. Diese liefert normierte und nach Möglichkeit orthogonale Eigenvektoren.

---

- (b) Der Eigenraum zu  $\lambda$  nur eindimensional. Dann können wir eine Basis bilden aus einem Eigenvektor von  $F$  und einem weiteren Vektor, und  $F$  hat bzgl. dieser Basis die Matrix

$$D = \begin{pmatrix} \lambda & \mu \\ 0 & \lambda \end{pmatrix}$$

mit irgendeiner Zahl  $\mu \neq 0$ . Die einfache Scherung (Abschnitt 24) fällt in diese Kategorie.

## 24 Die einfache Scherung

Eine **einfache Scherung** in  $\mathbb{R}^2$  ist eine lineare Abbildung  $F$ , welche nur  $\lambda = 1$  als Eigenwert hat, und bei der die Dimension des zugehörigen Eigenraums 1 ist. In  $\mathbb{R}^3$  ist die Definition dieselbe, außer dass der Eigenraum zum Eigenwert  $\lambda = 1$  zweidimensional ist.

Demnach gibt es eine (sogar orthonormale) Basis, sodass die Matrix von  $F$  die Form

$$\tilde{A} = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} \quad \text{bzw.} \quad \tilde{A} = \begin{pmatrix} 1 & 0 & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (73)$$

Die Zahl  $\gamma$  (eigentlich  $|\gamma|$ ) wird als **Shear Strain** bezeichnet.

Aus Gleichung 73 folgt sofort, dass für jede einfache Scherung

$$\begin{aligned} \text{tr}(F) &= n \quad (\text{Dimension des Raumes}) \\ \det(F) &= 1 \quad (\text{wie für jede Scherung}) \end{aligned}$$

gilt. Die Umkehrung gilt allerdings nur in  $\mathbb{R}^2$ . Wenn  $\text{tr}(F) = 2$  ist und  $\det(F) = 1$ , dann handelt es sich um eine einfache Scherung.

Die Matrix einer einfachen Scherung in  $\mathbb{R}^2$  mit gegebenem Shear Strain und gegebener Scherachse bzgl. der Standardbasis lässt sich folgendermaßen

aufschreiben: Ist  $\alpha$  der Winkel der Scherachse gegenüber der  $x_1$ -Achse, so ist  $a = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$  der entsprechende Eigenvektor. Durch Hinzunahme des Vektors  $b = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$  ergibt sich dann eine Orthonormalbasis des  $\mathbb{R}^2$ . Mit der Matrix

$$B = (a, b) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

lässt sich die Matrix der Scherung dann darstellen als

$$A = B\tilde{A}B^{-1} = \begin{pmatrix} 1 - \gamma \cos \alpha \sin \alpha & \gamma \cos^2 \alpha \\ -\gamma \sin^2 \alpha & 1 + \gamma \cos \alpha \sin \alpha \end{pmatrix} \quad (74)$$

Daraus ergibt sich

$$\gamma = a_{12} - a_{21} \quad (75)$$

$$\cos(2\alpha) = \cos^2 \alpha - \sin^2 \alpha = \frac{a_{12} + a_{21}}{a_{12} - a_{21}} \quad (76)$$

$$\sin(2\alpha) = 2 \cos \alpha \sin \alpha = \frac{a_{22} + a_{11}}{a_{12} - a_{21}} \quad (77)$$

## 25 Die adjungierte Abbildung

Ist  $F$  eine lineare Abbildung eines euklidischen Vektorraums, so heißt eine lineare Abbildung  $G$  (desselben Vektorraums) zu  $F$  **adjungiert**, wenn für alle Vektoren  $x$  und  $y$  gilt:

$$F(x) \cdot y = x \cdot G(y) \quad (78)$$

Hierfür wird meist die Bezeichnung  $G = F^*$  gewählt.

In  $\mathbb{R}^n$  mit Standard-Skalarprodukt gibt es zu jeder linearen Abbildung  $F$  genau eine adjungierte Abbildung  $F^*$ , was sich folgendermaßen einsehen lässt: Ist

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

die Matrix von  $F$  bzgl. der Standardbasis, so gilt

$$F(x) \cdot y = (Ax) \cdot y = \sum_{i=1}^n \left( \sum_{j=1}^n A_{ij} x_j \right) y_i = \sum_{j=1}^n x_j \left( \sum_{i=1}^n A_{ij} y_i \right) = x \cdot (A^T y) \quad (79)$$

wobei

$$A^T = \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \dots & \dots & \dots \\ a_{1n} & \dots & a_{nn} \end{pmatrix} \quad (80)$$

die **Transponierte** der Matrix  $A$  ist. Die zu  $F$  adjungierte lineare Abbildung  $F^*$  hat also die Transponierte der Matrix von  $F$  als Matrix. Dies gilt nicht nur bzgl. der Standardbasis in  $\mathbb{R}^n$ , sondern in jedem euklidischen Vektorraum endlicher Dimension, sofern die verwendete Basis eine Orthonormalbasis ist.

Direkt aus der Definition der adjungierten Abbildung einzusehen ist die Beziehung

$$(FG)^* = G^*F^* \quad \text{bzw.} \quad (AB)^T = B^T A^T \quad (81)$$

## 26 Orthogonale Abbildungen

Eine lineare Abbildung  $F$  (oder die entsprechende Matrix) in einem euklidischen Vektorraum heißt **orthogonal**, wenn sie die Längen von Vektoren und die Winkel zwischen Vektoren nicht verändert, also wenn sie das Skalarprodukt von Vektoren nicht ändert:

$$(F(x)) \cdot (F(y)) = x \cdot y \quad (82)$$

Nach Gleichung 78 ist dies dazu äquivalent, dass  $F^*F = \mathbf{1}$  ist bzw. dass

$$F^* = F^{-1} \quad \text{bzw.} \quad A^T = A^{-1} \quad (83)$$

für die entsprechenden Matrizen. Die Bedingung  $A^T A = E_n$  lässt sich auch so interpretieren, dass die Spalten einer orthogonalen Matrix orthonormal sind.

Orthogonale Abbildungen sind Drehungen, Spiegelungen und Kombinationen aus solchen.

Für die Determinante einer orthogonalen Abbildung gilt:

$$(\det(F))^2 = \det(F^*)\det(F) = \det(F^*F) = \det(\mathbf{1}) = 1 \quad (84)$$

sodass

$$\det(F) = \pm 1 \quad (85)$$

sein muss. Umgekehrt ist natürlich nicht jede Abbildung mit  $\det(F) = \pm 1$  orthogonal.

Diejenigen orthogonalen Abbildungen mit  $\det(F) = 1$  werden als **Drehungen** bezeichnet, diejenigen mit  $\det(F) = -1$  enthalten zusätzlich eine Spiegelung.

Drehungen in  $\mathbb{R}^2$  sind sehr einfach darstellbar. Die Matrix  $R$  einer Drehung in  $\mathbb{R}^2$  bzgl. der Standardbasis lautet

$$R = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}, \quad (86)$$

wobei  $\alpha$  der Drehwinkel (gegen den Uhrzeigersinn) ist.

Jede Drehung in  $\mathbb{R}^3$  (grundsätzlich in jedem Raum ungerader Dimension) besitzt eine **Drehachse**, also einen (mindestens) eindimensionalen Eigenraum zum Eigenwert  $\lambda = 1$ . Bei einer Drehung um die  $x_3$ -Achse lautet die Matrix bzgl. der Standardbasis

$$R = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (87)$$

wobei  $\alpha$  wieder der Drehwinkel (gegen den Uhrzeigersinn) ist. Dieselbe Darstellung ergibt sich, wenn bei beliebiger Drehachse eine Orthonormalbasis gewählt wird, bei der der dritte Basisvektor in Richtung der Drehachse zeigt. Bei beliebiger Drehachse und Standardbasis ist die Matrixdarstellung einer Drehung aber wesentlich komplizierter.

Beliebige Drehungen lassen sich mit Hilfe der **Exponentialfunktion** darstellen. Diese ist wie bei reellen Zahlen definiert als

$$\exp(F) = \sum_{n=0}^{\infty} \frac{1}{n!} F^n = \mathbf{1} + F + \frac{1}{2} F^2 + \frac{1}{6} F^3 + \dots \quad (88)$$

Ist  $F$  eine lineare Abbildung mit  $F^* = -F$  (siehe Abschnitt 27), d. h. die Matrix von  $F$  bzgl. einer Orthonormalbasis hat die Eigenschaft  $A^T = -A$ , so ist

$$R = \exp(F) \quad (89)$$

eine Drehung. Die Umkehrung gilt auch, d. h., jede Drehung lässt sich gemäß Gleichung 89 durch eine lineare Abbildung  $F$  mit  $F^* = -F$  darstellen. In  $\mathbb{R}^2$  lautet die entsprechende Matrix von  $F$

$$A = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix}, \quad (90)$$

---

**Box 14** Exponentialfunktionen in MATLAB

Neben der Exponentialfunktion  $\exp$  für reelle Zahlen, welche auch elementweise auf Vektoren und Matrizen angewandt werden kann, besitzt MATLAB eine Funktion  $\expm$  für quadratische Matrizen, welche der in diesem Abschnitt verwendeten Definition entspricht.

---

wobei  $\alpha$  der Drehwinkel (im Bogenmaß, gegen den Uhrzeigersinn) ist. In  $\mathbb{R}^3$  benötigen wir einen Vektor  $\alpha$ , der in Richtung der Drehachse zeigt, und dessen Länge dem Drehwinkel entspricht. Dann ist

$$F(x) = \alpha \times x \quad (91)$$

bzw. die Matrix von  $F$

$$A = \begin{pmatrix} 0 & -\alpha_2 & \alpha_2 \\ \alpha_3 & 0 & -\alpha_1 \\ -\alpha_2 & \alpha_1 & 0 \end{pmatrix} \quad (92)$$

## 27 Selbstadjungierte Abbildungen

Lineare Abbildungen, deren Adjungierte gleich der ursprünglichen Abbildung ist, d. h.  $F^* = F$ , heißen **selbstadjungiert** oder **symmetrisch**.

Dieselbe Definition wird auch für Matrizen angewandt, d. h. quadratische Matrizen mit  $A^T = A$  werden als **symmetrisch** bezeichnet. Die Bezeichnung selbstadjungiert ist im Zusammenhang mit Matrizen allerdings nicht gebräuchlich.

Ist  $F$  eine selbstadjungierte lineare Abbildung, sind Eigenvektoren zu verschiedenen Eigenwerten nicht nur linear unabhängig, sondern sogar orthogonal. Ist  $a$  Eigenvektor zum Eigenwert  $\lambda$  und  $b$  Eigenvektor zum Eigenwert  $\mu$ , gilt:

$$\lambda a \cdot b = F(a) \cdot b = a \cdot F(b) = \mu a \cdot b, \quad (93)$$

was für  $\lambda \neq \mu$  nur sein kann, wenn  $a \cdot b = 0$  ist.

Darüber hinaus sind alle selbstadjungierten linearen Abbildungen diagonalisierbar (der Beweis ist etwas aufwändiger), sodass es eine Orthonormalbasis aus Eigenvektoren von  $F$  gibt.

Hieraus folgt, dass sich jede symmetrische Matrix  $A$  darstellen lässt als

$$A = BDB^T, \quad (94)$$

wobei  $D$  eine Diagonalmatrix und  $B$  eine orthogonale Matrix ist (eigentlich sogar eine Drehmatrix).

Symmetrische Scherungen, also lineare Abbildungen mit  $F^* = F$  und  $\det(F) = 1$ , werden als **reine Scherung** oder als **Plättung** bezeichnet. Diese strecken bzw. stauchen 3 (bzw. 2) zueinander senkrechte Achsen um verschiedene Faktoren, sodass das Produkt der Faktoren 1 ist.

Die wohl wichtigste Anwendung selbstadjungierter Abbildungen (bzw. symmetrischer Matrizen) in der Geologie stellt der **Spannungstensor** dar. Hierbei betrachtet man eine kleine (näherungsweise ebene) Fläche in irgendeinem Medium und charakterisiert ihre Orientierung durch einen Normaleneinheitsvektor  $n$ . Dann ist die Kraft pro Fläche, die auf diese Fläche wirkt, gegeben durch  $\sigma n$ , wobei  $\sigma$  eine selbstadjungierte lineare Abbildung des  $\mathbb{R}^3$  (oder eine symmetrische  $3 \times 3$ -Matrix) ist und als Spannungstensor bezeichnet wird. Aus der Symmetrie von  $\sigma$  folgt, dass es immer drei aufeinander senkrechte Richtungen gibt, bei denen die Kraft in Richtung von  $n$  (oder in Gegenrichtung) zeigt, also senkrecht auf der Fläche steht. Diese drei Richtungen werden als Hauptspannungsachsen bezeichnet, und die zugehörigen Eigenwerte  $\sigma_1$ ,  $\sigma_2$  und  $\sigma_3$  als Hauptspannungen.

Die Komponente von  $\sigma n$ , welche in Richtung von  $n$  zeigt, also  $(\sigma n) \cdot n$ , stellt die Spannung senkrecht zu betrachteter Fläche dar. Diese bezeichnet man als Normalspannung, während die Komponente senkrecht zu  $n$  (also der Rest) als Scherspannung bezeichnet wird.

Allgemein bezeichnet man den Ausdruck  $F(x) \cdot x$  als **quadratische Form** von  $F$ . Ist  $b_1, \dots, b_n$  eine Orthonormalbasis aus Eigenvektoren zu den Eigenwerten  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ , so lässt sich jeder Vektor  $x$  als

$$x = \mu_1 b_1 + \dots + \mu_n b_n, \quad (95)$$

darstellen, und die quadratische Form ist

$$F(x) \cdot x = \mu_1^2 \lambda_1 + \dots + \mu_n^2 \lambda_n. \quad (96)$$

Unter allen Einheitsvektoren, also

$$\|x\|^2 = \mu_1^2 + \dots + \mu_n^2 = 1, \quad (97)$$

Wird die quadratische Form maximal für  $\mu_1 = \pm 1, \mu_2 = \dots = \mu_n = 0$ , also für  $x = \pm b_1$  und hat dabei den Wert  $\lambda_1$ . Sie wird minimal für  $x = b_n$  und hat dann den Wert  $\lambda_n$ . Der minimale und der maximale Wert der quadratischen Form einer selbstadjungierten Abbildung unter allen Einheitsvektoren liefert also den kleinsten und den größten Eigenwert.

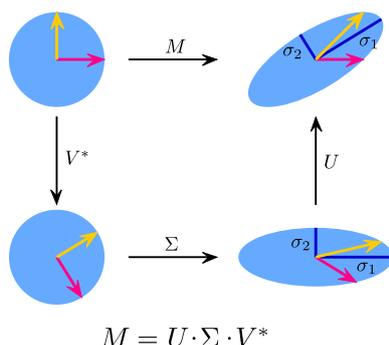


Abbildung 12: Illustration der Singulärwertzerlegung am Beispiel ein einfachen Scherung. Quelle: <https://de.wikipedia.org/wiki/Singulärwertzerlegung> (von Georg-Johann - Eigenes Werk, CC-BY-SA 3.0, <https://commons.wikimedia.org/wiki/User:Georg-Johann>).

### Box 15 Singulärwertzerlegung in MATLAB

MATLAB besitzt eine Funktion `svd` für die Singulärwertzerlegung einer Matrix.

In Bezug auf den Spannungstensor bedeutet dies, dass die minimale bzw. maximale Hauptspannung tatsächlich die minimale bzw. maximale auftretende Normalspannung ist. Die Berechnung der maximalen Scherspannung ist hingegen etwas mühsam. Es lässt sich zeigen, dass diese  $\frac{\lambda_1 - \lambda_n}{2}$  (also  $\frac{\sigma_1 - \sigma_3}{2}$ ) beträgt und in den Richtungen  $\frac{1}{\sqrt{2}} (\pm b_1 \pm b_n)$  auftritt.

## 28 Singulärwertzerlegung

Die Zerlegung in Drehung, Diagonalmatrix und Zurückdrehung nach Gleichung 94,  $A = BDB^T$ , ist nur für symmetrische Matrizen möglich. Für beliebige (auch nicht quadratische) Matrizen  $A$  lässt sich dies verallgemeinern zur **Singulärwertzerlegung**

$$A = UDV^T, \quad (98)$$

wobei  $D$  eine Diagonalmatrix mit nichtnegativen Diagonalelementen ( $\geq 0$ ) ist, während  $U$  und  $V$  orthogonale Matrizen sind (Abb. 12).

## 29 Polarzerlegung

Aus der Singulärwertzerlegung lässt sich ableiten, dass sich jede lineare Abbildung eines endlichdimensionalen Vektorraums in eine orthogonale Abbildung und eine selbstadjungierte Abbildung zerlegen lässt. Für die entsprechenden Matrizen gilt:

$$A = UDV^T = UV^TVDV^T = RS \quad (99)$$

mit der orthogonalen Matrix  $R = UV^T$  und der symmetrischen Matrix (mit nichtnegativen Eigenwerten)  $S = VDV^T$ . Ist  $\det(A) \geq 0$ , so ist  $R$  eine Drehung (ohne Spiegelung). Dies wird als **Polarzerlegung** bezeichnet.

Die Polarzerlegung ist auch in umgekehrter Reihenfolge möglich:

$$A = UDV^T = UDU^TUV^T = \tilde{S}R \quad (100)$$

mit derselben orthogonalen Matrix  $R$ , aber einer anderen symmetrischen Matrix  $\tilde{S} = UDU^T$ .

Aus der Polarzerlegung folgt unmittelbar, dass sich jeder beliebige Typ einer Scherung (insbesondere also die einfache Scherung) als Hintereinanderausführung einer Drehung und einer reinen Scherung (selbstadjungierte Abbildung) darstellen lässt.

## 30 Hauptstreckungsachsen

Das Verhältnis

$$s(x) = \frac{\|F(x)\|}{\|x\|} \quad (101)$$

gibt an, wie sich die Länge des Vektors  $x \neq 0$  durch die Anwendung einer linearen Abbildung  $F$  relativ ändert. Der maximale und der minimale Wert (über alle Vektoren  $x$ ) dieses Verhältnisses,  $s_{\max}$  und  $s_{\min}$ , werden als **Hauptstreckungsfaktoren** von  $F$  bezeichnet. Die zugehörigen Vektoren bilden einen Unterraum und werden als **Hauptstreckungsachsen** bezeichnet.

Die Hauptstreckungsfaktoren und -achsen lassen sich auf zwei Arten berechnen: Liegt bereits eine Polarzerlegung gemäß Gleichung 99 vor, also  $F = RS$ , so gilt

$$s(x) = \frac{\|R(S(x))\|}{\|x\|} = \frac{\|S(x)\|}{\|x\|}, \quad (102)$$

da die orthogonale Abbildung  $R$  die Länge der Vektoren nicht ändert. Mit einem Einheitsvektor  $e = \frac{1}{\|x\|}x$  vereinfacht sich dies zu

$$s(x) = \|S(e)\|. \quad (103)$$

Da  $S$  symmetrisch ist, gibt es eine Orthonormalbasis aus Eigenvektoren  $b_1, \dots, b_n$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$ , sodass

$$e = \mu_1 b_1 + \dots + \mu_n b_n, \quad (104)$$

mit  $\mu_1^2 + \dots + \mu_n^2 = 1$ . Damit ist

$$s(x) = \|\mu_1 \lambda_1 b_1 + \dots + \mu_n \lambda_n b_n\| = \sqrt{\mu_1^2 \lambda_1^2 + \dots + \mu_n^2 \lambda_n^2}. \quad (105)$$

Dieser Ausdruck wird maximal, wenn  $\mu$  des betragsmäßig größten Eigenwerts  $\pm 1$  ist, und der Streckungsfaktor wird

$$s_{\max} = \max\{|\lambda_1|, \dots, |\lambda_n|\}. \quad (106)$$

Die zugehörige Hauptstreckungsachse ist demnach gegeben durch den Eigenraum zu diesem Eigenwert. Entsprechendes gilt für den minimalen Streckungsfaktor:

$$s_{\min} = \min\{|\lambda_1|, \dots, |\lambda_n|\}. \quad (107)$$

Ohne Verwendung der Polarzerlegung lassen sich die Hauptstreckungsfaktoren und -achsen mit Hilfe der Beziehung

$$s(x) = \frac{\|F(x)\|}{\|x\|} = \|F(e)\| = \sqrt{F(e) \cdot F(e)} = \sqrt{F^*(F(e)) \cdot e} \quad (108)$$

Der letzte Ausdruck in der Wurzel ist die quadratische Form der symmetrischen Abbildung  $F^*F$ . Nach Abschnitt 27 sind dann die Hauptstreckungsfaktoren gegeben durch die Wurzeln des größten und kleinsten Eigenwerts von  $F^*F$ , und die zugehörigen Eigenräume sind die Hauptstreckungsachsen.

## Index

- Abstand, 7
- adjungierte Abbildung, 39
- aufgespannt, 20
  
- Basis, 21
- Basiswechselformel, 34
- Bild, 29
  
- charakteristisches Polynom, 36
  
- Determinante, 34
- diagonalisierbar, 37
- Diagonalmatrix, 37
- Dimension, 20
- Drehachse, 41
- Drehung, 40
- Dreiecksungleichung, 7
- Dreifingerregel, 13
  
- Eigenraum, 36
- Eigenvektor, 36
- Eigenwert, 36
- einfache Scherung, 38
- Einheitskugel, 8
- Einheitsmatrix, 27
- Einheitsvektor, 8
- Endomorphismus, 24
- euklidische Norm, 7
- euklidischer Vektorraum, 11
- Exponentialfunktion, 41
  
- Hauptstreckungsachse, 44
- Hauptstreckungsfaktor, 44
- homogenes lineares Gleichungssystem, 31
- Hyperebene, 21
  
- Identität, 25
- inhomogenes lineares Gleichungssystem, 31
- Invariante, 37
  
- Inverse, 29
- invertierbar, 29
  
- Kern, 30
- Koordinaten, 21
- Koordinatenvektor, 21
- kreistreu, 16
- Kreuzprodukt, 13
- Kugel, 8
- Kugelkoordinaten, 10
  
- Länge, 7
- linear abhängig, 19
- linear unabhängig, 19
- lineare Abbildung, 24
- lineare Hülle, 20
- linearer Fit, 33
- Linearform, 24
- Linearkombination, 19
  
- Manhattan-Norm, 8
- Matrix, 27
- Matrixmultiplikation, 29
- Maximumsnorm, 8
  
- nichtrotierende Linie, 36
- Norm, 7
- Normalenvektor, 21
- normierter Vektor, 8
- normierter Vektorraum, 7
- Normierung, 8
- Nullabbildung, 25
- Nullmatrix, 27
- Nullraum, 30
  
- orthogonal, 11
- orthogonale Abbildung, 40
- orthogonale Matrix, 40
- orthogonale Projektion, 14
- orthonormal, 12



p-Norm, 8  
Parallelprojektion, 14  
Parameter, 32  
Plättung, 42  
Polarkoordinaten, 9  
Polarzerlegung, 44  
Projektionszentrum, 15  
  
quadratische Form, 42  
  
Rang, 30  
Rechte-Faust-Regel, 13  
reine Scherung, 42  
reziproke Basis, 22  
  
Scherung, 36  
Schmidtsches Netz, 18  
selbstadjungierte Abbildung, 41  
senkrecht, 11  
Shear Strain, 38  
Singularwertzerlegung, 43  
Skalarprodukt, 11  
Spannungstensor, 42  
Spur, 34  
Standard-Skalarprodukt, 11  
Standardbasis, 22  
stereografische Projektion, 15  
symmetrische Abbildung, 41  
symmetrische Matrix, 42  
  
Transponierte, 32  
transponierte Matrix, 40  
  
Unterraum, 19  
  
Vektor, 5  
Vektorprodukt, 13  
Vektorraum, 5  
Verkettung, 28  
  
Winkel, 12  
winkeltreu, 16  
Wulffsches Netz, 16  
  
Zentralprojektion, 15