Kapitel 2

Matrizen und Vektoren

Eine Matrix ist ein rechteckiges Schema, in dem Zahlen zusammengefasst werden. Matrizen haben vielfältige Anwendungen, etwa bei der Beschreibung von Produktionsprozessen oder dem Lösen linearer Gleichungssysteme. Mit Hilfe der sogenannten Matrix-Vektor-Notation können komplizierte Zusammenhänge oftmals in einfacher und übersichtlicher Form dargestellt werden.

Wir beginnen dieses Kapitel mit zwei Beispielen, an denen wir die Nützlichkeit der Matrix-Vektor-Schreibweise illustrieren.

Beispiel 2.1 (Lineare Produktionsfunktion). Ein Betrieb stellt ein Gut her. Für die Produktion dieses Gutes werden n verschiedene Faktoren benötigt. Der Output y des Gutes bei gegebenen Inputs $x_1, \ldots, x_n$ werde beschrieben durch

$$y = f(x_1, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \,.$$

Die Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist eine sogenannte **lineare Produktionsfunktion**.

Beispiel 2.2 (Offenes Input-Output-Modell von Leontief[1]). Wir betrachten eine Volkswirtschaft, die aus n Sektoren $1, \ldots, n$ (Landwirtschaft, Energiewirtschaft, Baugewerbe, Bankgewerbe, usw.) besteht. In jedem Sektor wird ein einziges Gut als **Output** erzeugt. Die Sektoren sind miteinander verflochten, sodass zur Produktion eines Gutes **Inputs** aus den anderen Sektoren benötigt werden. Es bezeichne x_i den Output des i-ten Sektors und d_i die Endnachfrage nach dem i-ten Gut, die sogenannte **exogene Nachfrage**.

[1] Wassily Leontief (1906-1999); Wirtschaftsnobelpreis 1973.

K. Mosler et al., *Mathematische Methoden für Ökonomen*, 2. Aufl., Springer-Lehrbuch,

Mit a_{ij} werde die Menge des i-ten Gutes bezeichnet, die zur Produktion einer Einheit des j-ten Gutes benötigt wird. Die Menge des Gutes i, das als Input für die Sektoren $1, \ldots, n$ verwendet wird, die sogenannte **endogene Nachfrage**, ist also insgesamt

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \,.$$

Damit die Nachfrage d_i gedeckt werden kann, muss der Output des i-ten Sektors gleich der Summe aus exogener und endogener Nachfrage, d.h.

$$x_i = (a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n) + d_i$$

sein. Es stellt sich nun die Frage, ob eine gegebene exogene Nachfrage befriedigt werden kann und wie groß ggf. die Outputs x_i der einzelnen Sektoren sein müssen. Formal führt dies auf das Problem, ob das folgende Gleichungssystem eine Lösung besitzt, und, wenn ja, für welche Werte $x_1, \ldots, x_n$ es erfüllt ist:

$$\begin{array}{ccccccccccc}
x_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & + & d_1 \\
x_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & + & d_2 \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\
x_n & = & a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & + & d_n
\end{array}$$

Um Gleichungssysteme wie dieses in übersichtlicher Form zu schreiben und zu lösen, verwendet man Matrizen und Vektoren. Im Folgenden definieren wir Matrizen und Vektoren und geben Rechenregeln für ihre Addition und Multiplikation an.

2.1 Matrizen

Definition 2.1: Matrix, transponierte Matrix

Seien $m, n \in \mathbb{N}$. Ein rechteckiges Zahlenschema

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

von Zahlen $a_{ij} \in \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, heißt **Matrix** mit m **Zeilen** und n **Spalten** oder Matrix des **Formats** $m \times n$, kurz $m \times n$ Matrix. Die Menge aller $m \times n$ Matrizen bezeichnet man mit $\mathbb{R}^{m \times n}$.

Man schreibt dafür auch $\boldsymbol{A} = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ bzw. $\boldsymbol{A} = (a_{ij})$ oder $\boldsymbol{A} = (a_{ij})_{mn}$.

Die $n \times m$-Matrix

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

nennt man die zu A **transponierte Matrix**.

Wir bezeichnen Matrizen mit großen lateinischen Buchstaben in Fettdruck, also $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, usw. Beim Übergang zur transponierten Matrix werden Zeilen und Spalten vertauscht: Die Spalten der transponierten Matrix sind gerade die Zeilen der ursprünglichen Matrix und umgekehrt.

Beispiel 2.3.

$$\boldsymbol{A} = \begin{bmatrix} 9 & 10 \\ 6 & 4 \\ 2 & 0 \end{bmatrix}, \qquad \boldsymbol{A}^T = \begin{bmatrix} 9 & 6 & 2 \\ 10 & 4 & 0 \end{bmatrix}.$$

Transponiert man eine bereits transponierte Matrix erneut, so erhält man wieder die ursprüngliche Matrix; es gilt $(\boldsymbol{A}^T)^T = \boldsymbol{A}$. Die Elemente $a_{11}, a_{22}, \dots$ einer Matrix bilden die sogenannte **Hauptdiagonale**. Die Hauptdiagonalen einer Matrix und ihrer transponierten Matrix sind identisch.

Die Matrix, deren Komponenten alle gleich null sind, heißt **Nullmatrix**. Wir bezeichnen die Nullmatrix vom Format $m \times n$ mit $\mathbf{0}_{m\times n}$,

$$\mathbf{0}_{m\times n} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.$$

Eine Matrix $\boldsymbol{A}$ mit gleich vielen Zeilen und Spalten, d.h. mit $m = n$, heißt **quadratische Matrix**. Sie heißt **symmetrisch**, wenn $\boldsymbol{A} = \boldsymbol{A}^T$ ist, also wenn $a_{ij} = a_{ji}$ für alle Indizes i, j gilt. Beispielsweise ist

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 3 & 8 \\ 5 & 8 & 4 \end{bmatrix}.$$

eine symmetrische Matrix. Die quadratische Matrix, deren Hauptdiagonalelemente alle gleich eins und deren restliche Elemente alle gleich null sind, heißt **Einheitsmatrix**. Wir bezeichnen die Einheitsmatrix vom Format $n \times n$ mit $\boldsymbol{I}_n$,

$$\boldsymbol{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Eine quadratische Matrix heißt **Diagonalmatrix**, wenn die Elemente außerhalb der Hauptdiagonalen alle gleich null sind ($a_{ij} = 0$, falls $i \neq j$), während in der Hauptdiagonalen beliebige Zahlen $a_{11}, a_{22}, \ldots, a_{nn}$ stehen. Für eine solche Matrix schreibt man kurz $\mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.

Eine quadratische Matrix heißt **obere Dreiecksmatrix**, wenn die Elemente unterhalb der Hauptdiagonalen alle gleich null sind, d.h. $a_{ij} = 0$, falls $i > j$. Sie heißt **untere Dreiecksmatrix**, wenn die Elemente oberhalb der Hauptdiagonalen alle gleich null sind, d.h. $a_{ij} = 0$, falls $i < j$.

Beispiel 2.4.

$$\boldsymbol{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{U} = \begin{bmatrix} 2 & 6 & 4 \\ 0 & 5 & 3 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 5 & 0 \\ 4 & 6 & 1 \end{bmatrix}.$$

$\boldsymbol{D}$ ist Diagonalmatrix, $\boldsymbol{U}$ ist obere und $\boldsymbol{L}$ ist untere Dreiecksmatrix.

Definition 2.2: Spaltenvektor, Zeilenvektor

Eine Matrix des Formats $n \times 1$ wird als **Spaltenvektor** mit n Komponenten bezeichnet,

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \cdots x_n]^T.$$

x_i nennt man die **i-te Komponente** von $\boldsymbol{x}$, $i = 1, \ldots, n$. Entsprechend wird eine Matrix $\boldsymbol{y}$ des Formats $1 \times n$ als **Zeilenvektor** mit n Komponenten bezeichnet,

$$\boldsymbol{y} = [y_1, y_2, \cdots, y_n].$$

Wenn es auf den Unterschied zwischen Zeilen- und Spaltenform nicht ankommt, sagt man einfach **Vektor**. Wir bezeichnen Vektoren mit kleinen lateinischen Buchstaben in Fettdruck, $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{a}, \boldsymbol{b}$, usw. Das gleiche Symbol ohne Fettdruck mit angehängtem Index j bezeichnet dann die j-te Komponente des Vektors. Beispielsweise ist u_4 die vierte Komponente des Vektors $\boldsymbol{u}$.

Beispiel 2.5. Beispiele für Vektoren sind

$$\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \in \mathbb{R}^3, \quad \boldsymbol{a} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix} \in \mathbb{R}^4, \quad \boldsymbol{a}^T = [2,1,3,0] = (2,1,3,0) \in \mathbb{R}^4 .$$

Der Spaltenvektor mit n Komponenten, $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$, entspricht einem Punkt $(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$. (Das Gleiche gilt natürlich auch für den Zeilenvektor $[x_1, x_2, \cdots, x_n]$.) Im Folgenden identifizieren wir die Punkte des n-dimensionalen Euklidischen Raumes $\mathbb{R}^n$ mit den Spaltenvektoren des Formats $n \times 1$. Der Raum $\mathbb{R}^n$ ist also die Menge aller Spaltenvektoren mit n Komponenten.

Der Ursprung des $\mathbb{R}^n$ entspricht dem **Nullvektor** $[0, 0, \cdots, 0]^T$. Der Vektor, dessen i-te Komponente gleich eins ist und dessen andere Komponenten alle den Wert null haben, heißt ***i*-ter Einheitsvektor**. Er wird mit $\boldsymbol{e}_i$ bezeichnet. Im $\mathbb{R}^n$ gibt es n solche Einheitsvektoren $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, etwa im $\mathbb{R}^3$ die Einheitsvektoren

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

2.2 Addition und Skalarmultiplikation von Matrizen

Mit Matrizen kann man fast wie mit Zahlen rechnen. Die einfachsten Operationen sind die Addition und die Subtraktion zweier Matrizen sowie die Multiplikation einer Matrix mit einer reellen Zahl.

Die Addition zweier Matrizen erfolgt in sehr naheliegender Weise. Man addiert zwei Matrizen, indem man die Komponenten an den entsprechenden Positionen addiert.

Definition 2.3: Addition von Matrizen

Seien $\boldsymbol{A}$ und $\boldsymbol{B}$ $m \times n$ Matrizen, $\boldsymbol{A} = (a_{ij})$, $\boldsymbol{B} = (b_{ij})$. Die Summe von $\boldsymbol{A}$ und $\boldsymbol{B}$ ist so definiert:

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij}) = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

Bemerkung: Zu beachten ist, dass zwei Matrizen nur dann addiert werden können, wenn sie *dasselbe Format*, d.h. die gleiche Anzahl von Zeilen und die gleiche Anzahl von Spalten besitzen!

Beispiel 2.6.

$$\boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ 1 & 5 \\ 3 & 2 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} -2 & 0 \\ 7 & 1 \\ -5 & 2 \end{bmatrix}, \quad \boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 8 & 6 \\ -2 & 4 \end{bmatrix}.$$

Addiert man zu einer Matrix des Formats $m \times n$ die Nullmatrix desselben Formats, so erhält man wieder die ursprüngliche Matrix, d.h. $\boldsymbol{A} + \mathbf{0}_{m\times n} = \mathbf{0}_{m\times n} + \boldsymbol{A} = \boldsymbol{A}$.

Die Subtraktion zweier Matrizen erfolgt wie die Addition komponentenweise.

Definition 2.4: Subtraktion von Matrizen

Seien $\boldsymbol{A}$ und $\boldsymbol{B}$ $m \times n$ Matrizen, $\boldsymbol{A} = (a_{ij})$, $\boldsymbol{B} = (b_{ij})$. Die Differenz von $\boldsymbol{A}$ und $\boldsymbol{B}$ ist so definiert:

$$\boldsymbol{A} - \boldsymbol{B} = (a_{ij} - b_{ij}) = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2n} - b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \cdots & a_{mn} - b_{mn} \end{bmatrix}.$$

Subtrahiert man eine Matrix von sich selbst, so erhält man die Nullmatrix des entsprechenden Formates: $\boldsymbol{A} - \boldsymbol{A} = \mathbf{0}_{m\times n}$.

Eine weitere Rechenoperation ist die Multiplikation einer Matrix mit einer reellen Zahl. Man bezeichnet sie als **Skalarmultiplikation**. Die Skalarmultiplikation einer Matrix erfolgt, indem jede Komponente der Matrix mit der reellen Zahl multipliziert wird.

Definition 2.5: Skalarmultiplikation

Sei $\boldsymbol{A}$ eine $m \times n$ Matrix und $\lambda \in \mathbb{R}$. Das Skalarprodukt aus λ und $\boldsymbol{A}$ ist so definiert:

$$\lambda\boldsymbol{A} = (\lambda a_{ij}) = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{bmatrix}.$$

Beispiel 2.7 (Fortsetzung von Beispiel 2.6).

$$\left(-\frac{1}{2}\right) \cdot \boldsymbol{A} = \begin{bmatrix} -1 & 0 \\ -\frac{1}{2} & -\frac{5}{2} \\ -\frac{3}{2} & -1 \end{bmatrix}$$

Unter der Matrix $-\boldsymbol{A}$ versteht man die Matrix $(-1) \cdot \boldsymbol{A}$.

Bemerkung: Offensichtlich ist $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-1) \cdot \boldsymbol{B}$. Man hätte also die Differenz zweier Matrizen auch mithilfe von skalarer Multiplikation und Addition definieren können.

Satz 2.1: Rechenregeln für Matrizen

Seien $\boldsymbol{A}, \boldsymbol{B}$ und $\boldsymbol{C}$ $m \times n$ Matrizen, $\lambda, \mu \in \mathbb{R}$. Dann gelten die folgenden Aussagen:

(i)	$\boldsymbol{A} + \boldsymbol{B}$	$=$	$\boldsymbol{B} + \boldsymbol{A}\,,$	kommutativ (für „+“)
(ii)	$(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$	$=$	$\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})\,,$	assoziativ (für „+“)
(iii)	$\boldsymbol{A} + \boldsymbol{0}_{m\times n}$	$=$	$\boldsymbol{A}\,,$	neutrales Element (für „+“)
(iv)	$\boldsymbol{A} + (-\boldsymbol{A})$	$=$	$\boldsymbol{0}_{m\times n}\,,$	negatives Element (für „+“)
(v)	$\lambda(\boldsymbol{A} + \boldsymbol{B})$	$=$	$\lambda\boldsymbol{A} + \lambda\boldsymbol{B}\,,$	distributiv
(vi)	$(\lambda + \mu)\boldsymbol{A}$	$=$	$\lambda\boldsymbol{A} + \mu\boldsymbol{A}\,,$	distributiv
(vii)	$\lambda(\mu\boldsymbol{A})$	$=$	$(\lambda\mu)\boldsymbol{A}\,,$	assoziativ (für „·“)
(viii)	$1 \cdot \boldsymbol{A}$	$=$	$\boldsymbol{A}\,.$	neutrales Element (für „·“)

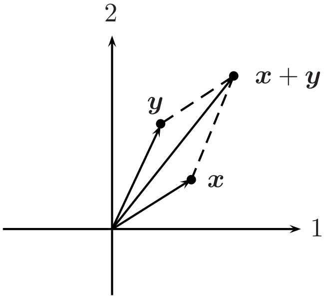
Addition und Skalarmultiplikation von Punkten des $\mathbb{R}^n$ Da Vektoren spezielle Matrizen sind (nämlich solche mit nur einer Spalte bzw. Zeile) sind Addition, Subtraktion und Skalarmultiplikation von Vektoren wie bei Matrizen definiert. Die obigen Rechenregeln gelten daher auch für Vektoren und in gleicher Weise für Punkte des $\mathbb{R}^n$. Insbesondere werden zwei Punkte $\boldsymbol{x} = (x_1, \cdots, x_n)$ und $\boldsymbol{y} = (y_1, \cdots, y_n)$ komponentenweise addiert,

$$\boldsymbol{x} + \boldsymbol{y} = (x_1, \cdots, x_n) + (y_1, \cdots, y_n) \;=\; (x_1 + y_1, \cdots, x_n + y_n)\,.$$

Die Addition von $\boldsymbol{x}$ und $\boldsymbol{y}$ entspricht dem Aneinandersetzen der beiden zugehörigen Pfeile von $\boldsymbol{0}$ bis $\boldsymbol{x}$ bzw. $\boldsymbol{y}$ (Abbildung 2.1).

Komponentenweise geht man auch vor, wenn man den Punkt $\boldsymbol{y}$ vom Punkt $\boldsymbol{x}$ subtrahiert und wenn man den Punkt $\boldsymbol{x}$ mit einer Zahl $\lambda \in \mathbb{R}$ multipliziert,

$$\begin{aligned} \boldsymbol{x} - \boldsymbol{y} = (x_1, \cdots, x_n) - (y_1, \cdots, y_n) &= (x_1 - y_1, \cdots, x_n - y_n), \\ \lambda \boldsymbol{x} = (\lambda x_1, \cdots, \lambda x_n). \end{aligned}$$



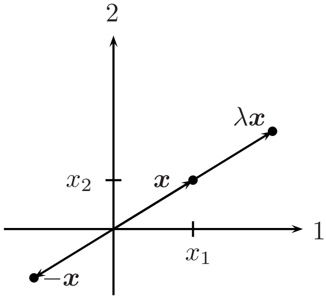


Abbildung 2.1: Summe von Vektoren, $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$, und Vielfaches eines Vektors, $\boldsymbol{z} = \lambda \boldsymbol{x}$.

Beispiel 2.8 (Aggregation von Produktbündeln). Ein Unternehmen produziert vier Produkte, die an drei verschiedenen Standorten gelagert werden. Zu einem bestimmten Zeitpunkt befanden sich $[100, 2000, 300, 0]$ Stück in Lager A, $[40, 0, 0, 700]$ in Lager B und $[400, 200, 0, 0]$ in Lager C. Hierbei sind für jedes Lager die Stückzahlen der vier Produkte zu einem Stückvektor im $\mathbb{R}^4$ zusammengefasst. Insgesamt waren also

$$[100, 2000, 300, 0] + [40, 0, 0, 700] + [400, 200, 0, 0] = [540, 2200, 300, 700]$$

Stück vorhanden. Durch ein Unwetter wurden sodann 80% der Produkte in Lager A zerstört, das sind

$$0.8 \cdot [100, 2000, 300, 0] = [80, 1600, 240, 0]$$

Stück. Insgesamt sind danach noch

$$[540, 2200, 300, 700] - [80, 1600, 240, 0] = [460, 600, 60, 700]$$

Stück der vier Produkte auf Lager.

Beispiel 2.9. In Abbildung 2.2 ist $n = 2$, $\boldsymbol{x} = (3, 1)$, $\boldsymbol{y} = (2, 2)$, $\boldsymbol{z} = (-2, 6)$, also $\boldsymbol{x} + \boldsymbol{y} = (5, 3)$, $\boldsymbol{x} - \boldsymbol{z} = (5, -5)$. Für beliebige Zahlen λ betrachten wir die Vielfachen $\lambda \boldsymbol{x}$. Sie bilden die Menge

$$U = \{(3\lambda, \lambda) : \lambda \in \mathbb{R}\} = \left\{ (x_1, x_2) : x_2 = \frac{1}{3} x_1 \right\}.$$

Dies ist die Gerade durch den Nullpunkt und den Punkt $(3, 1)$ im $\mathbb{R}^2$.

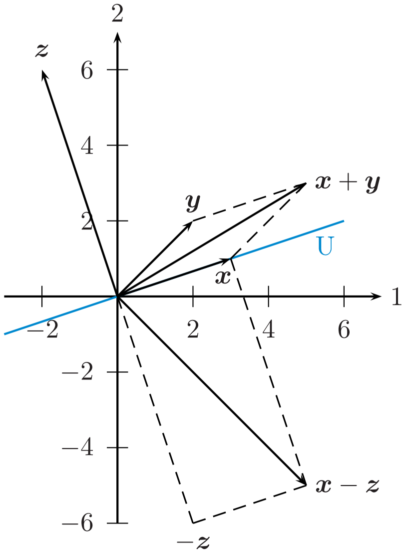


Abbildung 2.2: Zu Beispiel 2.9 .

2.3 Multiplikation von Matrizen

Für Matrizen sind nicht nur die linearen Operationen – Multiplikation mit einer Zahl und Addition von Matrizen gleichen Formats – definiert. Man kann Matrizen auch miteinander multiplizieren. Voraussetzung ist, dass sie im Format zusammenpassen, indem die linke Matrix so viele Spalten hat wie die rechte Matrix Zeilen.

Definition 2.6: Matrizenprodukt

Sei $\boldsymbol{A} \in \mathbb{R}^{m\times k}$ und $\boldsymbol{B} \in \mathbb{R}^{k\times n}$ mit $\boldsymbol{A} = (a_{il})$ und $\boldsymbol{B} = (b_{lj})$. Das **Produkt** von $\boldsymbol{A}$ und $\boldsymbol{B}$ ist die $m \times n$ Matrix $\boldsymbol{AB} = (c_{ij})$ mit den Elementen

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj} = \sum_{l=1}^{k} a_{il}b_{lj}.$$

Das Element c_{ij} ergibt sich also durch Multiplikation der i-ten Zeile der Matrix $\boldsymbol{A}$ mit der j-ten Spalte der Matrix $\boldsymbol{B}$.

Beispiel 2.10. Gegeben seien die Matrizen

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{und} \quad \boldsymbol{B} = \begin{bmatrix} -1 & 2 \\ -3 & 0 \end{bmatrix}.$$

Dann ist $n = 2$, $m = 3$, $k = 2$ und

$$\begin{array}{ccc|ccl}
 & & & -1 & 2 & \boldsymbol{B} \\
 & & & -3 & 0 & \\ \cline{2-5}
 & 1 & 0 & -1 & 2 & \\
\boldsymbol{A} & 0 & 1 & -3 & 0 & \\
 & 2 & 1 & -5 & 4 & \\
 & & & & & \boldsymbol{AB}
\end{array} \quad , \quad \text{also} \quad \boldsymbol{AB} = \begin{bmatrix} -1 & 2 \\ -3 & 0 \\ -5 & 4 \end{bmatrix}.$$

Das Element in der ersten Zeile und der ersten Spalte der Matrix $\boldsymbol{AB}$ berechnet man beispielsweise als

$$1 \cdot (-1) + 0 \cdot (-3) = -1\,.$$

Entsprechend berechnet man die anderen Elemente von $\boldsymbol{AB}$.

Ferner sei $\boldsymbol{x} = [3, 7]^T$. Wir erhalten

$$\boldsymbol{Ax} = \begin{bmatrix} 3 \\ 7 \\ 13 \end{bmatrix} \in \mathbb{R}^3, \; \boldsymbol{Bx} = \begin{bmatrix} 11 \\ -9 \end{bmatrix} \in \mathbb{R}^2, \; \boldsymbol{x}^T\boldsymbol{x} = 3^2 + 7^2 = 58\,.$$

Das Produkt $\boldsymbol{BA}$ ist nicht definiert, da die Anzahl der Spalten von $\boldsymbol{B}$ nicht mit der Anzahl der Zeilen von $\boldsymbol{A}$ übereinstimmt. Wir berechnen $\boldsymbol{AA}^T$ und $\boldsymbol{A}^T\boldsymbol{A}$. Beide Produkte sind definiert, aber im Allgemeinen verschieden, wie das folgende Zahlenbeispiel zeigt. Es gilt

$$\begin{array}{ccc|cccl}
 & & & 1 & 0 & 2 & \boldsymbol{A}^T \\
 & & & 0 & 1 & 1 & \\ \cline{2-6}
 & 1 & 0 & 1 & 0 & 2 & \\
\boldsymbol{A} & 0 & 1 & 0 & 1 & 1 & \\
 & 2 & 1 & 2 & 1 & 5 & \\
 & & & & & & \boldsymbol{AA}^T
\end{array} \quad , \quad \text{aber} \quad
\begin{array}{cccc|ccl}
 & & & & 1 & 0 & \\
 & & & & 0 & 1 & \boldsymbol{A} \\
 & & & & 2 & 1 & \\ \cline{2-6}
\boldsymbol{A}^T & 1 & 0 & 2 & 5 & 2 & \\
 & 0 & 1 & 1 & 2 & 2 & \\
 & & & & & & \boldsymbol{A}^T\boldsymbol{A}
\end{array} \quad .$$

Satz 2.2: Rechenregeln für die Matrizenmultiplikation

Sei $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B}$ und $\boldsymbol{C} \in \mathbb{R}^{n\times k}$, $\boldsymbol{D} \in \mathbb{R}^{k\times r}$, $\lambda \in \mathbb{R}$. Dann gelten die folgenden Aussagen:

$$\begin{array}{rrcll}
\text{(i)} & \boldsymbol{A}(\boldsymbol{B}\boldsymbol{D}) & = & (\boldsymbol{A}\boldsymbol{B})\boldsymbol{D}\,, & \text{assoziativ} \\
\text{(ii)} & \boldsymbol{A}\boldsymbol{I}_{n\times n} & = & \boldsymbol{A}\,, & \multirow{2}{*}{$\left.\begin{array}{l}\\ \\ \end{array}\right\}$ neutrales Element} \\
\text{(iii)} & \boldsymbol{I}_{m\times m}\boldsymbol{A} & = & \boldsymbol{A}\,, & \\
\text{(iv)} & \boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) & = & \boldsymbol{A}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{C}\,, & \multirow{2}{*}{$\left.\begin{array}{l}\\ \\ \end{array}\right\}$ distributiv} \\
\text{(v)} & (\boldsymbol{B}+\boldsymbol{C})\boldsymbol{D} & = & \boldsymbol{B}\boldsymbol{D}+\boldsymbol{C}\boldsymbol{D}\,, & \\
\text{(vi)} & \boldsymbol{A}(\lambda\boldsymbol{B}) & = & \lambda(\boldsymbol{A}\boldsymbol{B})\,, & \\
\text{(vii)} & (\boldsymbol{A}\boldsymbol{B})^T & = & \boldsymbol{B}^T\boldsymbol{A}^T\,. &
\end{array}$$

Bemerkung: Während Addition und Subtraktion von Vektoren im Wesentlichen dieselben Eigenschaften besitzen wie Addition und Subtraktion reeller Zahlen, gilt dies nicht für die Multiplikation von Matrizen. Wir wollen an dieser Stelle zwei Punkte hervorheben, in denen sich die Matrizenmultiplikation grundlegend von der Multiplikation reeller Zahlen unterscheidet.

- Die Matrizenmultiplikation ist **nicht kommutativ**, d.h. es kommt sehr wohl auf die Reihenfolge der Faktoren an. Wenn man das Produkt $\boldsymbol{AB}$ zweier Matrizen bilden kann, heißt dies nicht, dass auch das Produkt $\boldsymbol{BA}$ gebildet werden kann. Aber selbst wenn das Produkt $\boldsymbol{BA}$ definiert ist, müssen die Matrizen $\boldsymbol{AB}$ und $\boldsymbol{BA}$ nicht übereinstimmen. Wir geben dazu ein weiteres Beispiel:

$$\begin{aligned}
\boldsymbol{AB} &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 7 & 11 \end{bmatrix}, \\
\boldsymbol{BA} &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 7 & 10 \end{bmatrix}.
\end{aligned}$$

- Ist das Produkt zweier reeller Zahlen gleich null, so bedeutet dies, dass einer der Faktoren gleich null ist. Das Produkt zweier Matrizen kann jedoch die Nullmatrix ergeben, ohne dass einer der Faktoren eine Nullmatrix ist. So ist z.B.

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{0}_{2\times 2}\,,$$

  obwohl weder $\boldsymbol{A}$ noch $\boldsymbol{B}$ eine Nullmatrix ist.

2.4 Inverse Matrizen

Jede Zahl x, die ungleich null ist, hat einen sogenannten Kehrwert, nämlich die Zahl $1/x = x^{-1}$. Multipliziert man x von rechts oder links mit seinem Kehrwert, so ist das Ergebnis immer gleich eins; es gilt $x \cdot x^{-1} = x^{-1} \cdot x = 1$.

Die Frage ist, ob es etwas Vergleichbares bei Matrizen gibt. Hat jede Matrix, die nicht die Nullmatrix ist, einen „Matrixkehrwert", sodass die Multiplikation der Matrix mit ihrem „Kehrwert" die Einheitsmatrix ergibt? Wenn eine Matrix $\boldsymbol{A}$ nicht quadratisch ist, so lässt sich in der Tat keine Matrix $\boldsymbol{B}$ finden, für die sowohl $\boldsymbol{AB}$ als auch $\boldsymbol{BA}$ gleich der Einheitsmatrix ist. In einer großen Klasse quadratischer Matrizen existiert jedoch ein Analogon zum Kehrwert, die sogenannte **inverse Matrix**.

Definition 2.7: Inverse Matrix

Sei $\boldsymbol{A}$ eine quadratische $n \times n$-Matrix. Gibt es eine $n \times n$-Matrix $\boldsymbol{B}$, sodass $\boldsymbol{AB} = \boldsymbol{I}_n$ ist, dann heißt $\boldsymbol{B}$ **inverse Matrix** zu $\boldsymbol{A}$ oder **Inverse** von $\boldsymbol{A}$. Man schreibt $\boldsymbol{A}^{-1} = \boldsymbol{B}$. Es gilt dann

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_n \,.$$

Existiert zu einer Matrix $\boldsymbol{A}$ die inverse Matrix $\boldsymbol{A}^{-1}$, so heißt $\boldsymbol{A}$ **invertierbar** oder **regulär**. Existiert zu $\boldsymbol{A}$ keine inverse Matrix, so heißt $\boldsymbol{A}$ **nicht invertierbar** oder **singulär**.

Beispiel 2.11.

- Die Matrix $\boldsymbol{A} = \left[\begin{smallmatrix} 4 & 7 \\ 1 & 2 \end{smallmatrix}\right]$ ist invertierbar mit

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix}.$$

  Man prüft dies durch Ausmultiplizieren nach:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

- Die Matrix $\boldsymbol{B} = \left[\begin{smallmatrix} 1 & 2 \\ 2 & 4 \end{smallmatrix}\right]$ ist nicht invertierbar.
- Die Einheitsmatrix $\boldsymbol{I}_n$ ist regulär. Es gilt $\boldsymbol{I}_n^{-1} = \boldsymbol{I}_n$.
- Die Nullmatrix $\boldsymbol{0}_{n\times n}$ ist nicht invertierbar.

Eine 1×1-Matrix $\boldsymbol{A} = [a_{11}]$ ist offensichtlich genau dann invertierbar, wenn $a_{11} \neq 0$ ist. Die Inverse ist dann $\boldsymbol{A}^{-1} = [a_{11}^{-1}]$. Auch für 2×2-Matrizen kann man leicht entscheiden, ob die Matrix invertierbar ist, und ggf. die inverse Matrix in sehr einfacher Weise bestimmen.

Satz 2.3: Inverse, $n = 2$

Für $\boldsymbol{A} = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ existiert $\boldsymbol{A}^{-1}$ genau dann, wenn $ad - bc \neq 0$ ist. In diesem Fall ist

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

Die Größe $ad - bc$ ist die sogenannte Determinante[2] der Matrix $\boldsymbol{A}$. Damit lautet die Merkregel für die Inversion einer 2×2-Matrix:

Merkregel: Inverse, $n = 2$

Hauptdiagonalelemente vertauschen, bei den Nebendiagonalelementen das Vorzeichen umkehren, und alle Elemente jeweils durch die Determinante teilen.

Auch für größere Matrizen kann man – zumindest im Prinzip – Formeln für die Inverse angeben. So ist die Inverse einer 3×3-Matrix

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^{-1} = \frac{1}{\lambda} \begin{bmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix},$$

wobei $\lambda = aei + bfg + cdh - gec - hfa - idb$ die entsprechende Determinante darstellt. Die Inverse existiert, wenn $\lambda \neq 0$ ist. Offensichtlich kann man sich diese Formel aber nur schwer merken, was in der Tat auch nicht nötig ist. Wir werden in Abschnitt 8.4 ein einfaches Verfahren kennenlernen, mit dem man nicht nur entscheiden kann, ob eine Matrix invertierbar ist oder nicht, sondern auch im Fall der Existenz direkt die Inverse berechnet.

Es folgen weitere Eigenschaften der Inversen einer Matrix.

Satz 2.4: Inverse eines Produktes

Seien $\boldsymbol{A}$ und $\boldsymbol{B}$ invertierbare $n \times n$-Matrizen. Dann ist auch $\boldsymbol{AB}$ invertierbar und es gilt

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}.$$

Beweis: Zu zeigen ist, dass $(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{I}_n$ gilt:

$$(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{A}(\boldsymbol{BB}^{-1})\boldsymbol{A}^{-1} = \boldsymbol{AI}_n\boldsymbol{A}^{-1} = \boldsymbol{AA}^{-1} = \boldsymbol{I}_n.$$

□

[2] Zu Determinanten siehe Kapitel 10.

Diesen Satz kann man sich leicht mit der **Socke-Schuh-Regel** merken: Beim Anziehen zieht man erst die Socke an ($\boldsymbol{A}$) und dann den Schuh ($\boldsymbol{B}$). Das Inverse vom Anziehen ($\boldsymbol{AB}$) ist das Ausziehen ($(\boldsymbol{AB})^{-1}$). Dabei muss man zuerst den Schuh ausziehen ($\boldsymbol{B}^{-1}$) und dann die Socke ausziehen ($\boldsymbol{A}^{-1}$). Also ist $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, siehe Abbildung 2.3.

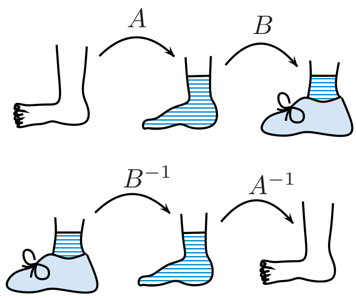


Abbildung 2.3: Illustration $(AB)^{-1} = B^{-1}A^{-1}$.

Satz 2.5: Rechenregeln für die Inversion

Falls $\boldsymbol{A}$ invertierbar ist, gilt:

(i) $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$,

(ii) $(\boldsymbol{A}^{T})^{-1} = (\boldsymbol{A}^{-1})^{T}$,

(iii) Ist $\boldsymbol{A}$ symmetrisch, so ist auch $\boldsymbol{A}^{-1}$ symmetrisch,

(iv) Für jede reelle Zahl $\lambda \neq 0$ ist $(\lambda\boldsymbol{A})^{-1} = \frac{1}{\lambda}\boldsymbol{A}^{-1}$.

2.5 Lineare Abbildungen

Wir wissen also nun, was Matrizen bzw. Vektoren sind und wie man mit ihnen rechnet. Vektoren dienen insbesondere dazu, Punkte im mehrdimensionalen Raum zu beschreiben und mit ihnen zu rechnen. Worin liegt aber nun der Nutzen von Matrizen? Um dies zu sehen, betrachten wir wieder die Beispiele 2.1 und 2.2 vom Beginn dieses Kapitels.

Beispiel 2.12 (Lineare Produktionsfunktion). Fasst man die Koeffizienten $a_1, \ldots, a_n$ der linearen Produktionsfunktion zu einem Zeilenvektor $\boldsymbol{a} = [a_1, \cdots, a_n]$ und die Inputs $x_1, \ldots, x_n$ zu einem Spaltenvektor $\boldsymbol{x} = [x_1, \cdots, x_n]^T$ zusammen, so kann man die lineare Produktionsfunktion in der Form

$$y = f(\boldsymbol{x}) = \boldsymbol{ax}$$

schreiben. Hier wird die Analogie zu linearen Funktionen von $\mathbb{R}$ nach $\mathbb{R}$, die bekanntlich die Form $y = ax$ besitzen, deutlich.

Beispiel 2.13 (Offenes Input-Output-Modell von Leontief). Wie im vorigen Beispiel fassen wir die Produktionskoeffizienten zusammen, jedoch statt zu einem Zeilenvektor zu einer Matrix $\boldsymbol{A} = (a_{ij})$. Die Matrix $\boldsymbol{A}$ heißt **Produktionsmatrix**. Entsprechend fassen wir die exogene Nachfrage $d_1, \ldots, d_n$ und die Outputs $x_1, \ldots, x_n$ zu Spaltenvektoren $\boldsymbol{d} = [d_1, \cdots, d_n]^T$ bzw. $\boldsymbol{x} = [x_1, \cdots, x_n]^T$ zusammen. In dieser Notation wird das Gleichungssystem des Leontief-Modells eine Gleichung zwischen Vektoren des $\mathbb{R}^n$,

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{d}\,.$$

Wegen $\boldsymbol{I}\boldsymbol{x} = \boldsymbol{x}$ kann man die Gleichung auch als

$$\boldsymbol{I}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{d}$$

schreiben und mit Hilfe der Rechenregeln für Matrizen zu

$$(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{d}$$

umformen. Dies ist die bekannte Gleichung des offenen Input-Output-Modells in Matrix-Vektor-Schreibweise. Es handelt sich um eine sogenannte **lineare Gleichung**. Um zu prüfen, ob eine gegebene exogene Nachfrage $\boldsymbol{d}$ gedeckt werden kann und welcher Gesamtoutput $\boldsymbol{x}$ dazu nötig ist, muss lediglich diese Gleichung gelöst werden. Wie man lineare Gleichungen löst, werden wir ausführlich in Kapitel 8 behandeln.

Der tiefere Nutzen von Matrizen liegt darin begründet, dass sie lineare Funktionen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ beschreiben.

Definition 2.8: Lineare Funktion

Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt **linear**, wenn für alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt:

Lin1: $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$,

Lin2: $f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$.

Statt linearer Funktion sagt man auch **lineare Abbildung**.

Durch eine gegebene $m \times n$ Matrix $\boldsymbol{A}$ wird eine Funktion $f_{\boldsymbol{A}} : \mathbb{R}^n \to \mathbb{R}^m$,

$$\boldsymbol{x} \mapsto f_{\boldsymbol{A}}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}\,,$$

definiert. Aus den Rechenregeln für Matrizen ergibt sich unmittelbar, dass $f_{\boldsymbol{A}}$ eine lineare Abbildung ist. Die grundlegende Bedeutung von Matrizen rührt nun daher, dass sogar *jede* lineare Abbildung die Form $f_{\boldsymbol{A}}$ mit einer geeignet gewählten Matrix $\boldsymbol{A}$ hat. Mit Hilfe von Matrizen lassen sich also alle linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ in einfacher Form beschreiben.

Wie findet man aber zu einer gegebenen linearen Funktion f die zugehörige Matrix $\boldsymbol{A}$? Ist $\boldsymbol{A}$ die zu f gehörige Matrix, und sind $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ die Spalten der Matrix $\boldsymbol{A}$, so folgt, dass der Funktionswert des j-ten Einheitsvektors $\boldsymbol{e}_j$ gleich

$$f(\boldsymbol{e}_j) = \boldsymbol{A}\boldsymbol{e}_j = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_m]\, \boldsymbol{e}_j = \boldsymbol{a}_j$$

ist. Mit anderen Worten: Die j-te Spalte der Matrix $\boldsymbol{A}$ ist das Bild des j-ten Einheitsvektors. Wir formulieren dies als Merksatz.

Merkregel: Matrix und zugehörige lineare Abbildung

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung und $\boldsymbol{A}$ die zugehörige $m \times n$ Matrix. Dann sind die Spalten $\boldsymbol{a}_j$ von $\boldsymbol{A}$ die Bilder der Einheitsvektoren, $\boldsymbol{a}_j = f(\boldsymbol{e}_j), j = 1, \ldots, n$.

Hieraus folgt unmittelbar, dass eine lineare Funktion durch die Funktionswerte der Einheitsvektoren eindeutig festgelegt ist.

Beispiel 2.14. Sei $f : \mathbb{R}^3 \to \mathbb{R}^4$ eine lineare Abbildung und gelte

$$f(\boldsymbol{e}_1) = [1,\ 3,\ 4,\ 2]^T, \quad f(\boldsymbol{e}_2) = [2,\ 0,\ 5,\ 1]^T, \quad f(\boldsymbol{e}_3) = [0,\ 3,\ 2,\ 4]^T.$$

Dann ist $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ mit

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 3 \\ 4 & 5 & 2 \\ 2 & 1 & 4 \end{bmatrix}.$$

Lineare Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ haben die Form $f(x) = ax$. Diese haben – sofern $a \neq 0$ ist – immer eine Umkehrfunktion, nämlich $f^{-1}(x) = \frac{1}{a}x$. Unter welchen Bedingungen besitzt nun eine lineare Funktion $f_{\boldsymbol{A}} : \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$, eine Umkehrfunktion und wie sieht diese aus? Wenn $m = n$ und $\boldsymbol{A}$ invertierbar ist, dann ist die lineare Funktion $f_{\boldsymbol{A}^{-1}}$ offenbar die Umkehrfunktion von $f_{\boldsymbol{A}}$, denn es gilt

$$\begin{aligned} f_{\boldsymbol{A}^{-1}}\big(f_{\boldsymbol{A}}(\boldsymbol{x})\big) &= \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{x}, \\ f_{\boldsymbol{A}}\big(f_{\boldsymbol{A}^{-1}}(\boldsymbol{x})\big) &= \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{x} = \boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{x}. \end{aligned}$$

Dass die lineare Abbildung auch nur in diesem Fall eine Umkehrfunktion besitzt, ist die Aussage des folgenden Satzes.

Satz 2.6: Umkehrfunktion einer linearen Funktion

Ist $f_{\boldsymbol{A}} : \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$, eine lineare Funktion, so besitzt $f_{\boldsymbol{A}}$ genau dann eine Umkehrfunktion, wenn $m = n$ und $\boldsymbol{A}$ invertierbar ist. In diesem Fall ist die Umkehrfunktion gegeben durch

$$f_{\boldsymbol{A}^{-1}} : \mathbb{R}^n \to \mathbb{R}^n\,, \quad \boldsymbol{x} \mapsto \boldsymbol{A}^{-1}\boldsymbol{x}\,.$$

Eine weitere wichtige Klasse von Abbildungen sind die sogenannten **affin-linearen** Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$. Diese haben die Form

$$\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$$

mit geeigneter Wahl einer $m \times n$ Matrix $\boldsymbol{A}$ und eines Spaltenvektors $\boldsymbol{b} \in \mathbb{R}^m$. Geometrisch stellt die Addition des Vektors $\boldsymbol{b}$ eine Verschiebung dar.

2.6 Geometrie des $\mathbb{R}^n$

Den $\mathbb{R}^2$ kann man sich anschaulich als eine zweidimensionale Zeichenebene vorstellen, den $\mathbb{R}^3$ als den uns umgebenden dreidimensionalen Raum. Für die Ebene und den dreidimensionalen Raum wissen wir, was unter Begriffen wie „Länge“, „Abstand“, „Winkel“ oder auch „senkrecht“ zu verstehen ist. Der $\mathbb{R}^n$ ist die natürliche Verallgemeinerung des $\mathbb{R}^2$ und des $\mathbb{R}^3$, doch fehlt ihm im Gegensatz zu diesen beiden die direkte Anschaulichkeit. Einige geometrische Eigenschaften des $\mathbb{R}^2$ und des $\mathbb{R}^3$ lassen sich allerdings auf den allgemeinen $\mathbb{R}^n$ übertragen.

In diesem Abschnitt definieren wir diese Begriffe allgemein für den Euklidischen Raum $\mathbb{R}^n$. Zunächst definieren wir das **innere Produkt** zweier Vektoren. Es ist, wie wir zeigen werden, grundlegend für die Verallgemeinerung der eingangs genannten geometrischen Begriffe.

Definition 2.9: Inneres Produkt

Für $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ heißt $\boldsymbol{x}^T\boldsymbol{y}$ **inneres Produkt**. Es ist

$$\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{x} = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i \in \mathbb{R}\,.$$

Das innere Produkt von $\boldsymbol{x}$ und $\boldsymbol{y}$ ist gleich dem Matrizenprodukt des Zeilenvektors $\boldsymbol{x}^T$ mit dem Spaltenvektor $\boldsymbol{y}$.

Beispiel 2.15. Sei

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix} \quad \boldsymbol{c} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Dann ist $\boldsymbol{a}^T\boldsymbol{b} = 1 \cdot (-2) + 2 \cdot 0 + 3 \cdot 2 = 4$ und $\boldsymbol{b}^T\boldsymbol{c} = 0$.

Beispiel 2.16 (Fortsetzung: Beispiel 2.8). Die vier Produkte des Unternehmens A haben pro Stück einen Wert von 500, 700, 200 bzw. 100 €. Der Gesamtwert der durch das Unwetter zerstörten Produkte beträgt

$$\begin{aligned}[80,\ 1600,\ 240,\ 0] \begin{bmatrix} 500 \\ 700 \\ 200 \\ 100 \end{bmatrix} &= 80 \cdot 500 + 1600 \cdot 700 + 240 \cdot 200 + 0 \cdot 100 \\ &= 1\,208\,000 \text{ €}.\end{aligned}$$

Der Lagerbestand hat nach dem Unwetter nur noch einen Wert von

$$[460,\ 600,\ 60,\ 700][500,\ 700,\ 200,\ 100]^T = 732000 \text{ €}.$$

Ist $\boldsymbol{x} = (x_1, x_2)$ ein Punkt im $\mathbb{R}^2$, so kann man den Abstand von $\boldsymbol{x}$ zum Ursprung $\mathbf{0}$ mit dem Satz von Pythagoras[3] berechnen; er beträgt $\sqrt{x_1^2 + x_2^2}$. Analog gilt für einen Punkt $\boldsymbol{x} \in \mathbb{R}^3$, dass der Abstand des Punktes von $\mathbf{0}$ durch $\sqrt{x_1^2 + x_2^2 + x_3^2}$ gegeben ist. Den Abstand eines Punktes oder Vektors $\boldsymbol{x}$ vom Nullpunkt nennt man auch **Norm** von $\boldsymbol{x}$. Auch wenn wir für den $\mathbb{R}^4$ und höhere Dimensionen keine anschauliche Vorstellung mehr besitzen, so lässt sich doch der Begriff der Norm durch eine entsprechende Formel verallgemeinern.

Definition 2.10: Norm

Für $\boldsymbol{x} \in \mathbb{R}^n$ heißt

$$\|\boldsymbol{x}\| := \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$$

(Euklidische) Norm von $\boldsymbol{x}$.

[3] Pythagoras von Samos (ca. 570 v. Chr. - ca. 500 v. Chr.)

Im Fall $n = 1$, also wenn x eine reelle Zahl ist, gilt $\|x\| = \sqrt{x^2} = |x|$. Die Norm ist in diesem Fall gleich dem Betrag. Man kann die Norm also auch als Verallgemeinerung des Betrages ansehen. Betrachtet man im $\mathbb{R}^2$ oder $\mathbb{R}^3$ einen Vektor als Pfeil, der vom Ursprung $\mathbf{0}$ zum Punkt $\boldsymbol{x}$ führt, so entspricht $\|\boldsymbol{x}\|$ der Länge des Pfeiles.

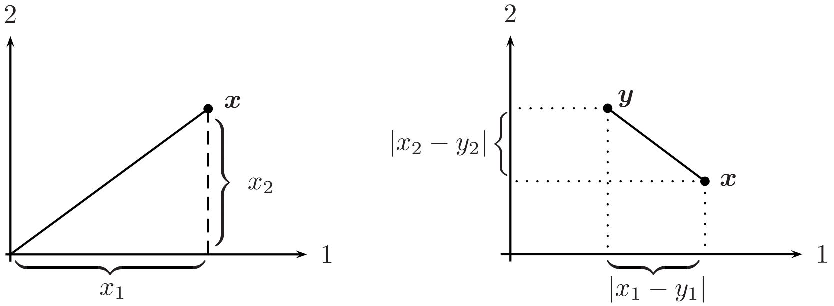


Abbildung 2.4: Norm und Distanz von Vektoren im $\mathbb{R}^2$.

Satz 2.7: Eigenschaften der Norm

Für alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ gilt

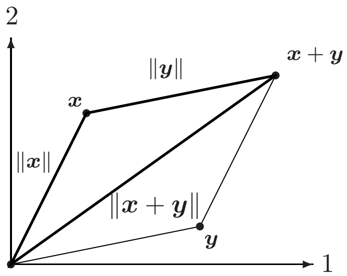
(i) $\|\boldsymbol{x}\| \geq 0$,

(ii) $\|\boldsymbol{x}\| = 0$ genau dann, wenn $\boldsymbol{x} = \mathbf{0}$,

(iii) $\|\lambda\boldsymbol{x}\| = |\lambda|\|\boldsymbol{x}\|$ für jede reelle Zahl λ ,

(iv) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

Eigenschaft (iv) bezeichnet man als **Dreiecksungleichung**. Geometrisch bedeutet sie, dass in einem Dreieck eine Seite höchstens so lang ist, wie die Summe der beiden anderen Seiten, vgl. die folgende Abbildung.



Mit dem Begriff der Norm haben wir festgelegt, was wir im $\mathbb{R}^n$ unter dem

Abstand eines Punktes vom Ursprung verstehen wollen. Wie soll man aber nun den Abstand zwischen zwei Punkten $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ definieren? Die Abbildung 2.4 zeigt, dass im $\mathbb{R}^2$ der Abstand zwischen zwei Punkten $\boldsymbol{x}$ und $\boldsymbol{y}$ durch $\sqrt{(x_1-y_1)^2+(x_2-y_2)^2} = \|\boldsymbol{x}-\boldsymbol{y}\|$ gegeben ist. Entsprechendes gilt für den Abstand zweier Punkte im $\mathbb{R}^3$.

Allgemein definiert man deshalb den Abstand zweier Punkte $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ als die Norm ihrer Differenz,

$$\|\boldsymbol{x}-\boldsymbol{y}\| = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\ldots+(x_n-y_n)^2}\,.$$

Man bezeichnet $\|\boldsymbol{x}-\boldsymbol{y}\|$ als **Euklidische Distanz** oder **Euklidischen Abstand**. Offenbar ist dieser Abstand symmetrisch; es gilt $\|\boldsymbol{x}-\boldsymbol{y}\| = \|\boldsymbol{y}-\boldsymbol{x}\|$. Für $n = 1, 2, 3$ stimmt der Begriff der Euklidischen Distanz mit unserer gewöhnlichen Anschauung überein.

Wir kommen nun zum Begriff des **Winkels** zwischen zwei Vektoren. Grundlegend hierfür ist der folgende Satz.

Satz 2.8: Cauchy-Schwarzsche Ungleichung[4]

Für $x, y \in \mathbb{R}^n$ gilt

$$|\boldsymbol{x}^T\boldsymbol{y}| \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|\,.$$

Aus der Cauchy-Schwarzschen Ungleichung folgt, dass für Vektoren $\boldsymbol{x}, \boldsymbol{y} \neq \boldsymbol{0}$ stets

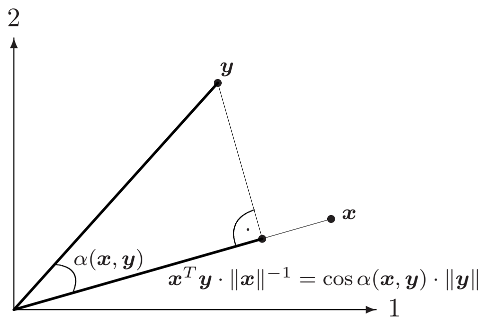
$$-1 \leq \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|} \leq 1$$

gilt. Man verwendet diesen Quotienten, um den **Öffnungswinkel** $\alpha(\boldsymbol{x}, \boldsymbol{y})$ zweier Vektoren $\boldsymbol{x}, \boldsymbol{y} \neq \boldsymbol{0}$ zu definieren:

$$\boxed{\cos\alpha(\boldsymbol{x}, \boldsymbol{y}) = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|}\,.}$$

Diese Definition entspricht im $\mathbb{R}^2$ der klassischen Definition des Kosinus (Ankathete durch Hypotenuse) am rechtwinkligen Dreieck. In der folgenden Abbildung hat die Hypotenuse die Länge $\|\boldsymbol{y}\|$, die Ankathete die Länge $\boldsymbol{x}^T\boldsymbol{y}/\|\boldsymbol{x}\|$.

[4] Augustin Louis Baron Cauchy (1789-1857) und Hermann Amandus Schwarz (1843-1921)



Die Definition der Länge eines Vektors sowie des Öffnungswinkels zweier Vektoren ermöglicht es, das Skalarprodukt zweier Vektoren geometrisch zu interpretieren. Es gilt

$$\boxed{\boldsymbol{x}^T\boldsymbol{y} = \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| \cdot \cos\alpha(\boldsymbol{x}, \boldsymbol{y})\,.}$$

Das Skalarprodukt ist also das Produkt der Längen (Normen) der Vektoren mit dem Kosinus des Öffnungswinkels.

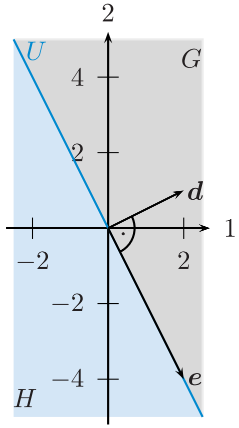
Wenn der von zwei Vektoren $\boldsymbol{x}$ und $\boldsymbol{y}$ gebildete Öffnungswinkel 90° beträgt, sagt man, dass die Vektoren zueinander orthogonal sind, d.h. senkrecht aufeinander stehen. Das innere Produkt $\boldsymbol{x}^T\boldsymbol{y}$ und damit der Kosinus sind in diesem Fall null. Man nutzt diesen Umstand zur formalen Definition des Begriffes „orthogonal".

Definition 2.11: Orthogonale Vektoren

Ist $\boldsymbol{x}^T\boldsymbol{y} = 0$, so heißen $\boldsymbol{x}$ und $\boldsymbol{y}$ **orthogonal** oder auch **senkrecht** zueinander, kurz: $\boldsymbol{x} \perp \boldsymbol{y}$.

Beispiel 2.17. In Beispiel 2.15 ist $\boldsymbol{b} \perp \boldsymbol{c}$.

Sei $\boldsymbol{d} = [2, 1]^T$ und $\boldsymbol{e} = [2, -4]^T$, dann ist $\boldsymbol{d} \perp \boldsymbol{e}$. Die folgende Abbildung zeigt $\boldsymbol{d}$ und $\boldsymbol{e}$ sowie die beiden Mengen $G = \{\boldsymbol{x} \in \mathbb{R}^2 \,:\, \boldsymbol{d}^T\boldsymbol{x} \geq 0\}$ (grau) und $H = \{\boldsymbol{x} \in \mathbb{R}^2 \,:\, \boldsymbol{d}^T\boldsymbol{x} \leq 0\}$ (blau). Die Menge $U = \{\boldsymbol{x} \in \mathbb{R}^2 \,:\, \boldsymbol{x} = \lambda\boldsymbol{e}, \lambda \in \mathbb{R}\}$ ist die Gerade, auf der alle Vektoren, die orthogonal zu $\boldsymbol{d}$ sind, liegen.



Die folgende Tabelle fasst die in diesem Abschnitt entwickelten geometrischen Begriffe sowie ihre Formalisierung im $\mathbb{R}^n$ zusammen.

Geometrischer Begriff	Formalisierung
Länge eines Vektors $\boldsymbol{x}$	$\\|\boldsymbol{x}\\|$,
Abstand zwischen $\boldsymbol{x}$ und $\mathbf{0}$	$\\|\boldsymbol{x}\\|$,
Abstand zwischen $\boldsymbol{x}$ und $\boldsymbol{y}$	$\\|\boldsymbol{x}-\boldsymbol{y}\\|$,
Winkel zwischen $\boldsymbol{x}$ und $\boldsymbol{y}$	$\cos\alpha(\boldsymbol{x},\boldsymbol{y}) = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\\|\boldsymbol{x}\\|\,\\|\boldsymbol{y}\\|}$,
$\boldsymbol{x}$ und $\boldsymbol{y}$ sind orthogonal	$\boldsymbol{x}^T\boldsymbol{y} = 0$.

2.7 Weitere Eigenschaften von Mengen im $\mathbb{R}^n$

Die Menge aller Punkte, deren Norm höchstens 1 beträgt, bildet die **abgeschlossene Einheitskugel** des $\mathbb{R}^n$, das ist die „Kugel“ mit „Radius“ 1 um den Nullpunkt des $\mathbb{R}^n$,

$$K(\mathbf{0},1) = \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \,:\, ||\boldsymbol{x}|| \leq 1\}\,.$$

Offenbar ist dies für $n = 2$ die abgeschlossene Kreisscheibe mit Radius 1,

$$\left\{(x_1, x_2) \in \mathbb{R} \,:\, \sqrt{x_1^2 + x_2^2} \leq 1\right\}.$$

Für $n = 3$ ist es die gewöhnliche Kugel mit Radius 1, und für $n = 1$ ist es das Intervall $[-1, 1]$.

Allgemein betrachtet man die **Kugel** $K(\boldsymbol{a}, \varepsilon)$ mit Mittelpunkt $\boldsymbol{a} \in \mathbb{R}^n$ und Radius $\varepsilon > 0$,

$$K(\boldsymbol{a}, \varepsilon) = \{\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \,:\, ||\boldsymbol{x} - \boldsymbol{a}|| \leq \varepsilon\},$$

das ist die Menge aller Punkte im $\mathbb{R}^n$, die vom Punkt $\boldsymbol{a}$ höchstens den Abstand ε besitzen.

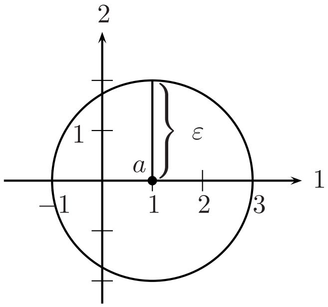


Abbildung 2.5: Kugel $K(\boldsymbol{a}, \varepsilon)$ im $\mathbb{R}^2$ (= Kreisscheibe) mit $\boldsymbol{a} = (0, 1), \varepsilon = 2$.

Eine Menge $A \subset \mathbb{R}^n$ heißt **beschränkt**, wenn sie in eine hinreichend große Kugel um $\mathbf{0}$ passt, d.h. wenn es eine Zahl M gibt, sodass

$$A \subset K(\mathbf{0}, M).$$

Beispielsweise ist jede Kugel $K(\boldsymbol{a}, \varepsilon)$ mit gegebenem Mittelpunkt $\boldsymbol{a} \in \mathbb{R}^n$ und Radius ε beschränkt. Für jedes $\boldsymbol{x} \in K(\boldsymbol{a}, \varepsilon)$ gilt nämlich $||\boldsymbol{x} - \boldsymbol{a}|| \leq \varepsilon$, also

$$||\boldsymbol{x}|| = ||\boldsymbol{x} - \boldsymbol{a} + \boldsymbol{a}|| \leq ||\boldsymbol{x} - \boldsymbol{a}|| + ||\boldsymbol{a}|| \leq \varepsilon + ||\boldsymbol{a}||.$$

Demnach liegt $K(\boldsymbol{a}, \varepsilon)$ in der Kugel um $\mathbf{0}$ mit Radius $M = \varepsilon + ||\boldsymbol{a}||$.

Auch jedes endliche n-dimensionale Intervall $[\boldsymbol{a}, \boldsymbol{b}]$ mit $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ ist eine beschränkte Menge.

Als nächstes betrachten wir die Randpunkte einer Menge $A \subset \mathbb{R}^n$. Ein Punkt $\boldsymbol{z}$ heißt **Randpunkt** von A, wenn jede noch so kleine Kugel um $\boldsymbol{z}$ sowohl Punkte von A als auch Punkte des Komplements von A enthält. Die Menge aller Randpunkte wird als **Rand** $\mathrm{rd}(A)$ bezeichnet.

Beispielsweise hat die Einheitskugel $K(\mathbf{0}, \mathbf{1})$ im $\mathbb{R}^2$ den Rand

$$\{(x_1, x_2) \,:\, x_1^2 + x_2^2 = 1\},$$

das ist die Kreislinie um $\mathbf{0}$ mit Radius 1. Für $a, b \in \mathbb{R}$ besitzt das gewöhnliche abgeschlossene Intervall $[a, b] \subset \mathbb{R}$ die Randpunkte a und b. Ebenso haben

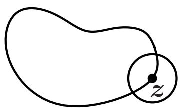


Abbildung 2.6: Randpunkt einer Menge.

das offene Intervall $]a, b[$ und die halboffenen Intervalle $[a, b[$ und $]a, b]$ die Randpunkte a und b. Offenbar können Randpunkte zur Menge dazugehören oder auch nicht.

Eine Menge $A \subset \mathbb{R}^n$ heißt **abgeschlossen**, wenn alle Randpunkte von A auch Elemente von A sind. Die Menge heißt **offen**, wenn keiner der Randpunkte Element von A ist. $A \setminus \mathrm{rd}(A)$ wird als **Inneres** von A bezeichnet, $A \cup \mathrm{rd}(A)$ als **abgeschlossene Hülle** von A. Eine Menge $A \subset \mathbb{R}^n$, die abgeschlossen und beschränkt ist, nennt man **kompakt**.

So ist zum Beispiel das gewöhnliche abgeschlossene Intervall $[a, b] \subset \mathbb{R}$ eine abgeschlossene Menge, das entsprechende offene Intervall $]a, b[$ eine offene Menge. Die halboffenen Intervalle $[a, b[$ und $]a, b]$ sind weder abgeschlossene noch offene Mengen. Die Einheitskugel $K(\mathbf{0}, \mathbf{1})$ im $\mathbb{R}^2$ ist eine abgeschlossene Menge. Allgemeiner ist jede Kugel $K(\boldsymbol{a}, \varepsilon)$ im $\mathbb{R}^n$ abgeschlossen. Der nichtnegative Orthant $\mathbb{R}^n_+$ ist abgeschlossen.

Eine Menge $A \subset \mathbb{R}^n$ nennt man **konvex**, wenn zu je zwei Punkten in A auch deren Verbindungsstrecke ganz in A liegt, d.h. wenn für alle $\boldsymbol{x}, \boldsymbol{y} \in A$ und alle $\lambda \in [0, 1]$ gilt

$$\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in A\,.$$

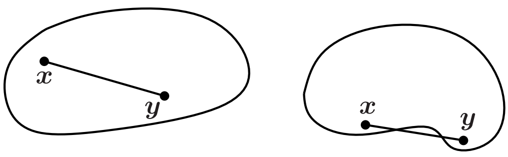


Abbildung 2.7: Eine konvexe und eine nichtkonvexe Menge.

Beispiele für konvexe Mengen sind die Intervalle in $\mathbb{R}$, aber auch alle n-dimensionalen Intervalle $[\boldsymbol{a}, \boldsymbol{b}]$. Weiter ist jede Kugel $K(\boldsymbol{a}, \varepsilon)$ im $\mathbb{R}^n$ konvex. Auch der $\mathbb{R}^n_+$ und der $\mathbb{R}^n$ sind konvex.

2.8 Orthogonale Matrizen und Abbildungen

Eine wichtige Klasse von Matrizen sind die sogenannten orthogonalen Matrizen.

Definition 2.12: Orthogonale Matrix

Eine $n \times n$-Matrix $\boldsymbol{Q}$ heißt **orthogonal**, wenn alle Spalten von $\boldsymbol{Q}$ die Norm eins haben und je zwei verschiedene Spalten orthogonal sind.

Bezeichnet $\boldsymbol{q}_i$ die i-te Spalte von $\boldsymbol{Q}$, so gilt

$$\boldsymbol{Q} \text{ orthogonal} \quad \Longleftrightarrow \quad \boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 1\,, & \text{falls } i = j, \\ 0\,, & \text{falls } i \neq j. \end{cases}$$

Orthogonale Matrizen haben einige sehr schöne Eigenschaften. Betrachten wir zunächst das Produkt $\boldsymbol{Q}^T\boldsymbol{Q}$. Da sich das Element an der Position (i, j) des Produktes als inneres Produkt der i-ten Zeile von $\boldsymbol{Q}^T$ und der j-ten Spalte von $\boldsymbol{Q}$ ergibt, ist dies offensichtlich gleich $\boldsymbol{q}_i^T\boldsymbol{q}_j$. (Die Spalten von $\boldsymbol{Q}$ sind die Zeilen von $\boldsymbol{Q}^T$!) Aus der Definition folgt dann, dass $\boldsymbol{Q}^T\boldsymbol{Q}$ auf der Hauptdiagonalen Einsen und an allen anderen Positionen Nullen enthält. Mit anderen Worten: $\boldsymbol{Q}^T\boldsymbol{Q}$ ist die Einheitsmatrix. Das bedeutet, dass $\boldsymbol{Q}^T$ die Inverse von $\boldsymbol{Q}$ ist. Dann ist aber auch $\boldsymbol{Q}\boldsymbol{Q}^T$ die Einheitsmatrix. Hieraus folgt, dass auch $\boldsymbol{Q}^T$ eine orthogonale Matrix ist. Also sind auch die Spalten von $\boldsymbol{Q}^T$ orthogonal zueinander und haben die Norm Eins. Da die Spalten von $\boldsymbol{Q}^T$ die Zeilen von $\boldsymbol{Q}$ sind, gilt dasselbe auch für die Zeilen von $\boldsymbol{Q}$. Wir fassen diese Überlegungen in einem Satz zusammen:

Satz 2.9: Eigenschaften orthogonaler Matrizen

Sei $\boldsymbol{Q}$ eine $n \times n$-Matrix. Dann sind die folgenden Aussagen äquivalent:

(i) $\boldsymbol{Q}$ ist orthogonal,

(ii) $\boldsymbol{Q}^T$ ist orthogonal,

(iii) $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_n$,

(iv) $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}_n$,

(v) $\boldsymbol{Q}$ ist invertierbar und $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^T$,

(vi) die Spalten von $\boldsymbol{Q}$ sind orthogonal und haben die Norm eins,

(vii) die Zeilen von $\boldsymbol{Q}$ sind orthogonal und haben die Norm eins.

Wir haben in Abschnitt 2.5 gesehen, dass die linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^n$ gerade die Abbildungen $f_{\boldsymbol{A}} : \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ sind, wobei $\boldsymbol{A}$ eine $n \times n$-Matrix ist. Lineare Abbildungen der Form $f_{\boldsymbol{Q}} : \boldsymbol{x} \mapsto \boldsymbol{Q}\boldsymbol{x}$, wobei $\boldsymbol{Q}$ eine **orthogonale** Matrix ist, heißen **orthogonale Abbildungen**. Die Bedeutung der orthogonalen Abbildungen liegt in der Tatsache begründet, dass sie **längen-** und **winkeltreu** sind.

Satz 2.10: Längen- und Winkeltreue orthogonaler Abbildungen

Sei $\boldsymbol{Q}$ eine orthogonale Matrix. Dann ist die Abbildung $f_{\boldsymbol{Q}}$

(i) **längentreu**, d.h. für beliebige Vektoren $\boldsymbol{x}, \boldsymbol{y}$ gilt

$$\|\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{Q}\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\|,$$

(ii) **winkeltreu**, d.h. für beliebige Vektoren $\boldsymbol{x}, \boldsymbol{y} \neq \boldsymbol{0}$ gilt

$$\alpha(\boldsymbol{Q}\boldsymbol{x}, \boldsymbol{Q}\boldsymbol{y}) = \alpha(\boldsymbol{x}, \boldsymbol{y}).$$

Beweis: Eigenschaft (i) ergibt sich aus

$$\begin{aligned}\|\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{Q}\boldsymbol{y}\| = \|\boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{y})\| &= \sqrt{[\boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{y})]^T [\boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{y})]} \\ &= \sqrt{(\boldsymbol{x} - \boldsymbol{y})^T \boldsymbol{Q}^T \boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{y})} = \sqrt{(\boldsymbol{x} - \boldsymbol{y})^T (\boldsymbol{x} - \boldsymbol{y})} = \|\boldsymbol{x} - \boldsymbol{y}\|.\end{aligned}$$

Für (ii) überlegt man sich zunächst, dass

$$(\boldsymbol{Q}\boldsymbol{x})^T (\boldsymbol{Q}\boldsymbol{y}) = \boldsymbol{x}^T \boldsymbol{Q}^T \boldsymbol{Q}\boldsymbol{y} = \boldsymbol{x}^T \boldsymbol{y}.$$

Wegen (i) ist aber auch $\|\boldsymbol{Q}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ und $\|\boldsymbol{Q}\boldsymbol{y}\| = \|\boldsymbol{y}\|$. Eigenschaft (ii) folgt dann unmittelbar aus der Formel für den Öffnungswinkel

$$\cos\alpha(\boldsymbol{Q}\boldsymbol{x}, \boldsymbol{Q}\boldsymbol{y}) = \frac{(\boldsymbol{Q}\boldsymbol{x})^T (\boldsymbol{Q}\boldsymbol{y})}{\|\boldsymbol{Q}\boldsymbol{x}\| \, \|\boldsymbol{Q}\boldsymbol{y}\|} = \frac{\boldsymbol{x}^T \boldsymbol{y}}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|} = \cos\alpha(\boldsymbol{x}, \boldsymbol{y}).$$

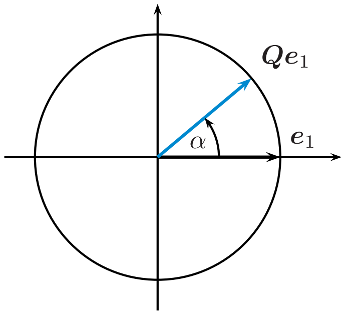
□

Insbesondere folgt aus der Längentreue, dass $\|\boldsymbol{Q}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ ist.

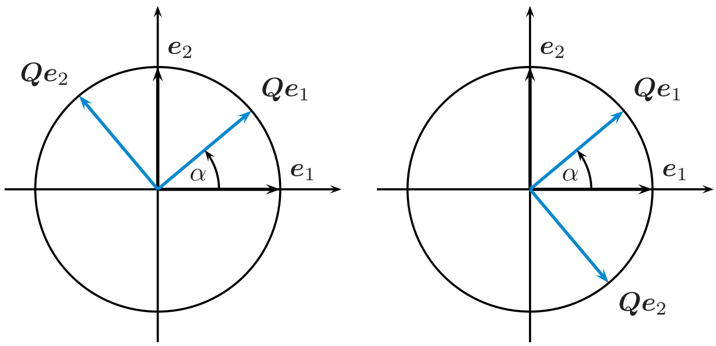
Orthogonale Abbildungen lassen also die Abstände sowie die Winkel zwischen zwei Punkten unverändert. Die orthogonalen Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}^2$ sind gerade die **Drehungen** um den Koordinatenursprung sowie die **Spiegelungen** an den Ursprungsgeraden.

Geometrische Interpretation orthogonaler Matrizen

Wie sehen nun die orthogonalen Abbildungen des $\mathbb{R}^2$ aus? Da zu einer orthogonalen Abbildung eine orthogonale Matrix $\boldsymbol{Q}$ gehört, müssen wir uns lediglich überlegen, wie die orthogonalen 2×2-Matrizen aussehen. Da orthogonale Abbildungen längentreu sind, muss das Bild des ersten Einheitsvektors den Abstand eins vom Ursprung haben, also irgendwo auf dem Einheitskreis liegen. Ist α der Winkel zwischen $\boldsymbol{e}_1$ und dem Bild $\boldsymbol{Q}\boldsymbol{e}_1$, so gilt also $\boldsymbol{Q}\boldsymbol{e}_1 = [\cos\alpha, \sin\alpha]^T$.



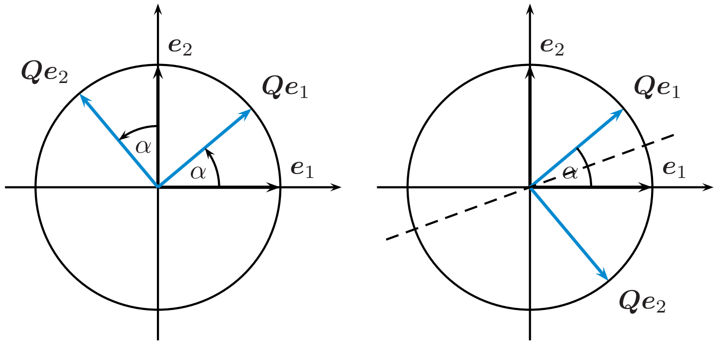
Was ist jetzt aber das Bild des zweiten Einheitsvektors? Da orthogonale Abbildungen winkeltreu sind und die beiden Einheitsvektoren einen rechten Winkel bilden, müssen auch die Bilder der Einheitsvektoren einen rechten Winkel bilden. Da das Bild des zweiten Einheitsvektors wegen der Längentreue ebenfalls auf dem Einheitskreis liegen muss, gibt es nur die folgenden beiden Möglichkeiten:



Im ersten Fall ist $\boldsymbol{Q}\boldsymbol{e}_2 = [-\sin\alpha, \cos\alpha]^T$ und im zweiten Fall entsprechend $\boldsymbol{Q}\boldsymbol{e}_2 = [\sin\alpha, -\cos\alpha]^T$. Da in den Spalten der Matrix die Bilder der Einheitsvektoren stehen, muss eine orthogonale 2×2-Matrix $\boldsymbol{Q}$ also eine der folgenden beiden Gestalten haben:

$$\boldsymbol{Q} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \quad \text{oder} \quad \boldsymbol{Q} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}.$$

Geometrisch beschreibt eine Matrix der ersten Gestalt eine Drehung um den Winkel α im Nullpunkt des Koordinatensystems. Eine Matrix der zweiten Gestalt liefert eine Spiegelung an der Ursprungsgeraden, die mit der x-Achse den Winkel $\alpha/2$ bildet.



Damit haben wir die orthogonalen 2×2-Matrizen vollständig klassifiziert. Wir fassen die Ergebnisse im folgenden Satz zusammen.

Satz 2.11: Klassifikation der orthogonalen 2×2-Matrizen

Ist $\boldsymbol{Q}$ eine orthogonale 2×2-Matrix, so gibt es einen Winkel $\alpha \in]-\pi, \pi]$, sodass $\boldsymbol{Q}$ eine der folgenden Gestalten hat:

$$\boldsymbol{Q} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \quad \text{oder} \quad \boldsymbol{Q} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$$

Im ersten Fall beschreibt $\boldsymbol{Q}$ eine **Drehung** um den Winkel α, im zweiten Fall eine **Spiegelung** an der um $\alpha/2$ gedrehten x-Achse.

In höheren Dimensionen ist die geometrische Interpretation orthogonaler Matrizen deutlich komplizierter. Allerdings gilt allgemein, dass jede orthogonale Matrix einer Hintereinanderausführung von Drehungen und Spiegelungen entspricht.

Wichtige Begriffe zur Wiederholung

Nach der Lektüre dieses Kapitels sollten folgende Begriffe geläufig sein:

- Matrix
- Einheitsmatrix, Nullmatrix
- transponierte Matrix
- symmetrische Matrix

- Skalarmultiplikation einer Matrix
- Addition und Multiplikation zweier Matrizen
- Zeilenvektor, Spaltenvektor
- inneres Produkt von Vektoren, orthogonal
- Norm
- inverse Matrix
- lineare Abbildung

Selbsttest

Anhand folgender Ankreuzaufgaben können Sie Ihre Kenntnisse zu diesem Kapitel überprüfen. Beurteilen Sie dazu, ob die Aussagen jeweils wahr (W) oder falsch (F) sind. Kurzlösungen zu diesen Aufgaben finden Sie in Anhang E.

I. Bei der Multiplikation ...

	W	F
... zweier Diagonalmatrizen entsteht eine Diagonalmatrix.	☐	☐
... einer Diagonalmatrix mit einer unteren Dreiecksmatrix entsteht eine Diagonalmatrix.	☐	☐
... zweier oberer Dreiecksmatrizen entsteht eine obere Dreiecksmatrix.	☐	☐

II. Gegeben sind die Zahl $\lambda \in \mathbb{R}$ und die Matrizen $\boldsymbol{A}, \boldsymbol{B}$ und $\boldsymbol{C}$ von jeweils geeignetem Format. Sind folgende Aussagen wahr oder falsch?

	W	F
$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{CA}$	☐	☐
$\boldsymbol{A}(\boldsymbol{B}+\lambda\boldsymbol{C}) = \boldsymbol{AB}+\lambda\boldsymbol{AC}$	☐	☐
$((\boldsymbol{BA})^{-1})^T = (\boldsymbol{B}^{-1})^T(\boldsymbol{A}^T)^{-1}$	☐	☐
$(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{C}+\boldsymbol{D}) = \boldsymbol{AC}+\boldsymbol{BC}+\boldsymbol{AD}+\boldsymbol{BD}$	☐	☐

III. Gegeben sind die Zahlen $\lambda, \mu \in \mathbb{R}$ und die Vektoren $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^3$. Sind folgende Aussagen wahr oder falsch?

	W	F
$\lvert\lvert\lambda\boldsymbol{x}+\mu\boldsymbol{y}\rvert\rvert \leq \lvert\lambda\rvert \cdot \lvert\lvert\boldsymbol{x}\rvert\rvert + \lvert\mu\rvert \cdot \lvert\lvert\boldsymbol{y}\rvert\rvert$	☐	☐
Aus $\boldsymbol{x}\perp\boldsymbol{y}$ und $\boldsymbol{y}\perp\boldsymbol{z}$ folgt $\boldsymbol{x}\perp\boldsymbol{z}$.	☐	☐

IV. Sind folgende Aussagen wahr oder falsch?

	W	F
Die Vereinigung zweier konvexer Mengen ist wieder konvex.	☐	☐
Die Einheitsmatrix ist die einzige Diagonalmatrix, die orthogonal ist.	☐	☐
Die Matrix $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ ist orthogonal.	☐	☐

Aufgaben

Aufgabe 1

Gegeben seien folgende Vektoren im $\mathbb{R}^4$: $\boldsymbol{x} = [3, 4, 5, 1]^T$, $\boldsymbol{y} = [-1, 4, 3, -2]^T$ und $\boldsymbol{z} = [1, 1, -1, -2]^T$. Berechnen Sie

a) $\boldsymbol{x} + 3\boldsymbol{y} - \boldsymbol{z}$,

b) $\boldsymbol{x}^T \cdot (\lambda \boldsymbol{z}), \quad \lambda \in \mathbb{R}^+ \setminus \{0\}$,

c) die Entfernung zwischen dem Punkt $\mathbf{y}$ und dem Nullpunkt.

Aufgabe 2

Gegeben seien die Vektoren $\boldsymbol{a} = [1, 2, 3, 4, 5]^T$, $\boldsymbol{b} = [0, 1, 0, 1, 0]^T$, $\boldsymbol{c} = [0, -1, -2, 3]^T$ und $\boldsymbol{d} = [0, 2, 4, -1]^T$. Berechnen Sie, falls möglich,

a) $\frac{1}{2}\boldsymbol{a} + \boldsymbol{b}$,

b) $\boldsymbol{a} + \boldsymbol{c}$,

c) $(\boldsymbol{a}^T \cdot \boldsymbol{b})\boldsymbol{c}$,

d) $\boldsymbol{a}(\boldsymbol{b}^T \cdot \boldsymbol{c})$,

e) $\boldsymbol{a}^T \cdot \boldsymbol{b} + \boldsymbol{c}^T \cdot \boldsymbol{d}$.

Aufgabe 3

Gegeben sind die Matrizen und Vektoren

$$\boldsymbol{A} = \begin{bmatrix} 1 & 6 \\ 2 & 5 \\ 3 & 4 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 4 & 1 & 2 \\ 5 & -1 & 3 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 1 & -2 \\ 3 & -2 & 1 \end{bmatrix},$$

$$\boldsymbol{d} = \begin{bmatrix} -2, & 0, & 1 \end{bmatrix}, \quad \boldsymbol{e} = \begin{bmatrix} 5 \\ -1 \\ -3 \end{bmatrix}.$$

Bestimmen Sie falls möglich

a) $\boldsymbol{AB}$, b) $\boldsymbol{BA}$, c) $\boldsymbol{ABC}$, d) $3\boldsymbol{CAB}$, e) $5\boldsymbol{de}$, f) $\boldsymbol{e}^T\boldsymbol{A}$,

g) $\boldsymbol{dC}$, h) $\boldsymbol{eC}$, i) $\boldsymbol{A}+\boldsymbol{B}^T$, j) $\boldsymbol{AB}+\boldsymbol{C}-\boldsymbol{ed}$, k) $\boldsymbol{C}(\boldsymbol{d}^T+\boldsymbol{e})$.

Aufgabe 4

Nehmen Sie an, dass die im Folgenden betrachteten Matrizen alle ein geeignetes Format besitzen. Berechnen Sie

a) $(\boldsymbol{AB})^T(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})^T$,

b) $(\boldsymbol{A}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})\boldsymbol{B})(\boldsymbol{B}+\boldsymbol{A})^{-1}$,

c) $(\boldsymbol{I}^{-1})^T\boldsymbol{A}-(\boldsymbol{I}^T)^{-1}\boldsymbol{A}+0\cdot\boldsymbol{I}$.

Aufgabe 5

Gegeben sind die Matrizen $\boldsymbol{A}=\begin{bmatrix}1 & 2\\ 3 & 5\end{bmatrix}$ und $\boldsymbol{B}=\begin{bmatrix}1 & 2\\ 1 & 4\end{bmatrix}$. Berechnen Sie

a) $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$,

b) $(\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}$,

c) $((\boldsymbol{A}^{-1}\boldsymbol{B})^T)^{-1}$.

Aufgabe 6

Gegeben sind die Vektoren $\boldsymbol{x}^T=[1,2,3]$, $\boldsymbol{y}^T=[0,-3,2]$, $\boldsymbol{z}^T=[-1,1,1]$.

a) Schätzen Sie $||\boldsymbol{x}||+||\boldsymbol{y}||$ und $||\boldsymbol{x}||\cdot||\boldsymbol{z}||$ nach unten ab, d.h. wie groß sind die Ausdrücke mindestens? Berechnen Sie auch die exakten Werte.

b) Wie stehen die Vektoren $\boldsymbol{x}$ und $\boldsymbol{y}$ im Raum zueinander?

Aufgabe 7

Prüfen Sie, ob die folgenden Matrizen orthogonal sind:

$$\text{a)}\ \begin{bmatrix}\frac{\sqrt{3}}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{\sqrt{3}}{2}\end{bmatrix}, \qquad \text{b)}\ \begin{bmatrix}1 & 0 & 0\\ 0 & \frac{3}{5} & -\frac{4}{5}\\ 0 & \frac{4}{5} & \frac{3}{5}\end{bmatrix}.$$