Lineare Algebra I für Informatiker, Statistiker und Wirtschaftsmathematiker

Wintersemester 2000/2001

Prof. H.-J. Schneider
Mitschrift: J. Krebs

LMU München

Inhaltsverzeichnis

Kapitel I: Matrizen und Spaltenvektoren

Beispiele 0.1.

1. In einem Hof leben Hühner und Hasen. Insgesamt sind es 10 Tiere mit 26 Beinen. Wieviele Hühner und Hasen leben dort?

   Sei
   $$x_1 = \text{Zahl der Hühner},$$
   $$x_2 = \text{Zahl der Hasen}.$$
   Dann gilt:
   $$\begin{aligned} x_1 + x_2 &= 10 \\ 2x_1 + 4x_2 &= 26 \end{aligned}$$
   Aus der ersten Gleichung folgt
   $$x_1 = 10 - x_2.$$
   Einsetzen in die zweite Gleichung liefert
   $$20 - 2x_2 + 4x_2 = 26,$$
   also
   $$2x_2 = 6 \quad \text{bzw.} \quad x_2 = 3.$$
   Aus der ersten Gleichung ergibt sich dann
   $$x_1 = 7.$$

2. Ein Minister des Kaisers von China (Han-Zeit, 202 v.Chr. - 9 n.Chr.) namens Chang Tsang verfaßte das Buch "Mathematik in 9 Büchern", in dem allgemeine Probleme wie das folgende behandelt werden.

   Kauf und Verkauf von Tieren:
   (a) Verkauft man 2 Rinder und 5 Schafe und kauft vom Erlös 13 Schweine, so bleiben 1000 Geldstücke übrig.
   (b) Verkauft man 3 Rinder und 3 Schweine und kauft vom Erlös 9 Schafe, so reicht das Geld gerade.
   (c) Verkauft man 6 Schafe und 8 Schweine und kauft vom Erlös 5 Rinder, so hat man 600 Geldstücke zuwenig.

   Was kostet ein Rind, ein Schwein, bzw. ein Schaf?

   Sei
   $$x_1 = \text{Preis eines Rindes},$$
   $$x_2 = \text{Preis eines Schafes},$$
   $$x_3 = \text{Preis eines Schweines}.$$

Dann gilt wegen (a)-(c):

$$\begin{array}{rcrcrcr} 2x_1 & + & 5x_2 & - & 13x_3 & = & 1000 \\ 3x_1 & - & 9x_2 & + & 3x_3 & = & 0 \\ -5x_1 & + & 6x_2 & + & 8x_3 & = & -600 \end{array}$$

Auf die Berechnung von x_1, x_2, x_3 wird später eingegangen.

1. Mengen und Abbildungen

Es wird die naive Mengenlehre verwendet.

Spezielle Mengen:

$\mathbb{N} = \{0, 1, 2, 3, ...\}$ Menge der natürlichen Zahlen.
$\mathbb{Z} = \{0, \pm 1, \pm 2, ...\}$ Menge der ganzen Zahlen.
$\mathbb{Q} = \{\frac{a}{b} \mid a, b$ ganze Zahlen, $b \neq 0\}$ Menge der rationalen Zahlen.
$\mathbb{R} =$ Menge der reellen Zahlen.
$\emptyset =$ leere Menge.

Abkürzungen:

$x \in X$	x Element der Menge X.
$x \notin X$	x nicht Element der Menge X.
$Y \subset X$	Y Teilmenge von X, d.h jedes Element von Y liegt in X.
$Y \subsetneqq X$	Y Teilmenge von X und $Y \neq X$.

Wichtiges Beweisprinzip: Gleichheit von Mengen.
Seien X, Y Mengen. $X = Y$ gilt genau dann, wenn $X \subset Y$ und $Y \subset X$, d.h. für jedes Element $x \in X$ gilt $x \in Y$, und für jedes Element $y \in Y$ gilt $y \in X$.

Beschreibung von Mengen:

1. Aufzählen aller Elemente, z.B. $X = \{3, 5\} = \{5, 3\}$.
2. Sei $E(x)$ eine "zulässige Eigenschaft" von Elementen x. Dann ist $\{x \mid E(x) \text{ gilt}\}$ die Menge aller x mit der Eigenschaft $E(x)$, z.B.
$$\{x \mid x \in \mathbb{N},\ 3 \leq x \leq 6,\ x \text{ Primzahl}\} = \{3, 5\}.$$

Bildung neuer Mengen: Seien X, Y Mengen.

1. Vereinigung: $X \cup Y = \{x \mid x \in X \text{ oder } x \in Y\}$.
2. Durchschnitt: $X \cap Y = \{x \mid x \in X \text{ und } x \in Y\}$.
3. Falls $X \subset Y$ gilt, ist $Y \setminus X = \{y \mid y \in Y \text{ und } y \notin X\}$ das Komplement von X in Y.
4. Produktmenge:
$$X \times Y = \{(x, y) \mid x \in X, y \in Y\} = \text{ Menge aller geordneten Paare } (x, y).$$
Hierbei gilt für $x, x' \in X,\ y, y' \in Y:\ (x, y) = (x', y')$ genau dann, wenn $x = x'$ und $y = y'$.

   Ebenso kann man für Mengen X_1, X_2, X_3 geordnete Tripel (x_1, x_2, x_3) mit $x_i \in X_i$ für $i = 1, 2, 3$ bilden.

   Allgemeiner: Sind $X_1, X_2, \ldots, X_n$ Mengen für $n \in \mathbb{N}$, so heißt $(x_1, \ldots, x_n)$ mit $x_i \in X_i$ für alle $1 \leq i \leq n$ geordnetes n-Tupel. Für $x_i, x_i' \in X_i,\ 1 \leq i \leq n$ gilt wieder: $(x_1, \ldots, x_n) = (x_1', \ldots, x_n')$ genau dann, wenn $x_i = x_i'$ für alle $1 \leq i \leq n$.

Abkürzungen:

$\forall$	für alle
$\exists$	es gibt
$\Rightarrow, \Leftarrow, \Leftrightarrow$	logische Implikationen
$:\Leftrightarrow$	nach Definition äquivalent
$:=$	nach Definition gleich

Definition 1.1. Seien X, Y Mengen.
1. Für eine Teilmenge $R \subset X \times Y$ heißt das Tripel (X, Y, R) **Relation** zwischen X und Y.
2. Eine Relation $f = (X, Y, \Gamma)$, also $\Gamma \subset X \times Y$ Teilmenge, heißt **Abbildung** von X nach Y $:\Leftrightarrow$ $\forall x \in X \; \exists$ genau ein $y \in Y: \; (x, y) \in \Gamma$.
3. Ist $f = (X, Y, \Gamma)$ eine Abbildung, so heißt Γ der **Graph** der Abbildung f.

Schreibweisen für Abbildungen: Sei $f = (X, Y, \Gamma)$ eine Abbildung.
1. $f(x) := y$, falls $(x, y) \in \Gamma$. $f(x)$ heißt Bild von x bei f.
2. Für f schreiben wir auch $f : X \to Y,\; x \mapsto f(x)$ oder $X \xrightarrow{\mathrm{f}} Y,\; x \mapsto f(x)$.

Insbesondere gilt für Abbildung $f : X \to Y,\; g : A \to B$:

$$f = g \;\Leftrightarrow\; X = A,\, Y = B,\, \forall x \in X : \; f(x) = g(x).$$

Definition 1.2. Seien $f : X \to Y$ eine Abbildung und $U \subset X,\; V \subset Y$ Teilmengen.
1. $f(U) := \{y \in Y \mid \exists\, x \in U : f(x) = y\} = \{f(x) \mid x \in U\}$ heißt **Bild** von U bei f.
2. $f^{-1}(V) := \{x \in X \mid f(x) \in V\}$ heißt **Urbild** von V bei f.

Eine spezielle Abbildungen ist für jede Menge X

$$\mathrm{id}_X : X \to X,\; x \mapsto x \quad \forall x \in X,$$

d.h. $\mathrm{id}_X(x) := x \quad \forall x \in X$. Diese Abbildung heißt **identische Abbildung** von X.

Seien $f : X \to Y,\; g : Y \to Z$ Abbildungen. Die **Hintereinanderausführung** oder **Komposition** von g und f ist die Abbildung $g \circ f : X \to Z$, definiert durch

$$(g \circ f)(x) := g(f(x)) \quad \forall x \in X.$$

Wir schreiben auch $gf = g \circ f$.

Seien $f : X \to Y,\; g : Y \to Z,\; h : Z \to U$ Abbildungen. Dann ist offenbar:

$$h \circ (g \circ f) = (h \circ g) \circ f \qquad (\textbf{assoziativ}),$$
$$\mathrm{id}_Y \circ f = f \circ \mathrm{id}_X \qquad (\textbf{neutrales Element}).$$

2. Matrizen und Spaltenvektoren

Definition 2.1. Seien $m, n \in \mathbb{N}$.
1. Ein rechteckiges Schema reeller Zahlen $a_{ij} \in \mathbb{R},\; 1 \le i \le m,\, 1 \le j \le n$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

heißt $m \times n$-**Matrix** mit **Koeffizienten** $a_{ij},\ 1 \leq i \leq m, 1 \leq j \leq n$. Hierfür schreibt man auch

$$A = (a_{ij})_{1\leq i\leq m, 1\leq j\leq n} = (a_{ij})_{ij} = (a_{ij}).$$

Für alle $1 \leq i \leq m$ und $1 \leq j \leq n$ heißt

$$\big(a_{i1}, a_{i2}, \cdots, a_{in}\big)\ i\text{-te } \textbf{Zeile} \text{ von A und } \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}\ j\text{-te } \textbf{Spalte} \text{ von A.}$$

2. ${}^m\mathbb{R}^n := \{A \mid A\ m \times n\text{-Matrix mit Koeffizienten in } \mathbb{R}\}$. Eine andere Bezeichnung hierfür ist $\mathbb{R}^{m\times n}$.

   Speziell:

$$\begin{array}{ll} {}^m\mathbb{R} := {}^m\mathbb{R}^1 & (\textbf{Spaltenraum}), \\ \mathbb{R}^n := {}^1\mathbb{R}^n & (\textbf{Zeilenraum}). \end{array}$$

Bemerkung 2.2. Formaler ist eine $m \times n$-Matrix eine Abbildung

$$\{1, 2, \cdots, m\} \times \{1, 2, \cdots, n\} \to \mathbb{R},\ (i, j) \mapsto a_{ij}.$$

Insbesondere gilt für $A = (a_{ij})_{1\leq i\leq m, 1\leq j\leq n}$, $B = (b_{kl})_{1\leq k\leq p, 1\leq l\leq q}$ mit $m, n, p, q \in \mathbb{N}$:

$$A = B \quad \Leftrightarrow \quad m = p,\ n = q,\ \forall\, 1 \leq i \leq m, 1 \leq j \leq n : a_{ij} = b_{ij}.$$

Definition 2.3.

1. Für $m, n \in \mathbb{N}$, $A = (a_{ij})_{1\leq i\leq m, 1\leq j\leq n}$, $B = (b_{ij})_{1\leq i\leq m, 1\leq j\leq n} \in {}^m\mathbb{R}^n$ und $r \in \mathbb{R}$ seien

$$A + B := (c_{ij})_{1\leq i\leq m, 1\leq j\leq n} \in {}^m\mathbb{R}^n,$$

   wobei $\forall\, i, j :\ c_{ij} := a_{ij} + b_{ij}$, und

$$rA := (ra_{ij})_{i\leq m, 1\leq j\leq n} \in {}^m\mathbb{R}^n.$$

2. Für $m, n, p \in \mathbb{N}$, $A = (a_{ij})_{1\leq i\leq m, 1\leq j\leq n} \in {}^m\mathbb{R}^n$ und $B = (b_{kl})_{1\leq k\leq n, 1\leq l\leq p} \in {}^n\mathbb{R}^p$ sei

$$AB := (d_{il})_{1\leq i\leq m, 1\leq l\leq p} \in {}^m\mathbb{R}^p,$$

   wobei $\forall\, i, l$:

$$d_{il} := \sum_{j=1}^{n} a_{ij}\ b_{jl} = a_{i1}\ b_{1l} + a_{i2}\ b_{2l} + \cdots + a_{in}\ b_{nl}.$$

Beispiele 2.4.

1.

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \end{pmatrix}}_{{}^2\mathbb{R}^3} \underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{{}^3\mathbb{R}^3} = \underbrace{\begin{pmatrix} 0 & 0 & 2 \\ 7 & -1 & 5 \end{pmatrix}}_{{}^2\mathbb{R}^3}$$

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

3. Allgemein:
$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix} = \sum_{j=1}^{n} a_{ij}\, b_{jk}$$

4.
$$\underbrace{\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}}_{{}^3\mathbb{R}^1} \underbrace{\begin{pmatrix} 2 & 1 & 3 & 0 \end{pmatrix}}_{{}^1\mathbb{R}^4} = \underbrace{\begin{pmatrix} 2 & 1 & 3 & 0 \\ -4 & -2 & -6 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{{}^3\mathbb{R}^4}$$

Beachte:
$$\begin{aligned} {}^m\mathbb{R}^n \times {}^m\mathbb{R}^n &\xrightarrow{+} {}^m\mathbb{R}^n,\ (A,B) \mapsto A+B \qquad \text{(Addition)} \\ {}^m\mathbb{R}^n \times {}^n\mathbb{R}^p &\xrightarrow{\cdot} {}^m\mathbb{R}^p,\ (A,C) \mapsto AC \qquad \text{(Multiplikation)} \end{aligned}$$

sind nur für die Matrizen der angegebenen Größe definiert.

Beachte: Bei uns ist $\mathbb{R}^n$ der Zeilenraum und ${}^n\mathbb{R}$ der Spaltenraum. Häufig wird der Zeilen- und Spaltenraum gleich bezeichnet, nämlich als $\mathbb{R}^n$ (aus dem Kontext muß klar sein, ob Zeilen oder Spalten gemeint sind).

Satz 2.5 (Rechenregeln für Addition und Skalarmuliplikation)**.** *Seien* $m, n \geq 1$*. Dann gilt* $\forall\ A, B, C \in {}^m\mathbb{R}^n,\ r, s, \in \mathbb{R}$*:*

1. (a) $(A+B)+C = A+(B+C)$ (**assoziativ**)
   (b) $A+B = B+A$ (**kommutativ**)
   (c) $A+0 = A$ (**neutrales Element**)
   (d) $A+(-1)A = 0$ (**inverses Element**)
   Hierbei ist
$$0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$
   die **Nullmatrix**.
2. (a) $(rs)A = r(sA)$ (**assoziativ**)
   (b) $1A = A$ (**neutrales Element**)
   (c) $(r+s)A = rA + sA$ (**distributiv**)
   $r(A+B) = rA + rB$ (**distributiv**)

Beweis. Klar, da komponentenweise richtig. □

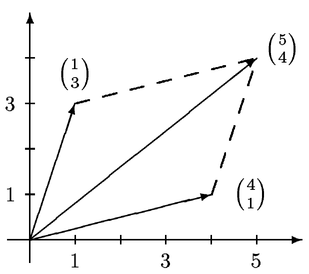
Definition 2.6. Sei $n \geq 1$.

1.
$$E := (\delta_{ij}) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in {}^n\mathbb{R}^n$$
heißt **Einheitsmatrix**. Hierbei ist
$$\delta_{ij} = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{sonst} \end{cases}$$

2. Eine Matrix der Gestalt
$$A = \begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & a_{nn} \end{pmatrix} \in {}^n\mathbb{R}^n,$$
d.h., es gilt $a_{ij} = 0 \; \forall \; i \neq j$, heißt **Diagonalmatrix**.

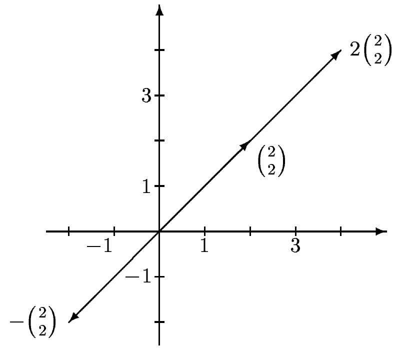
Bemerkungen 2.7.

1. Die Rechengesetze in Satz 2.5 gelten insbesondere für ${}^n\mathbb{R} = {}^n\mathbb{R}^1$, d.h. für den Raum aller Spalten
$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \; a_1, a_2, \cdots, a_n \in \mathbb{R}.$$
Die Elemente in ${}^n\mathbb{R}$ heißen **Spaltenvektoren**.
Für ${}^n\mathbb{R}$ gilt:
$n = 1$: Zahlengerade.
$n = 2$: Euklidische Ebene (Tafelebene).
$n = 3$: Euklidischer 3-dimensionaler Raum.
$n = 4$: Raum der Relativitätstheorie mit 3 Raumkoordianten und einer Zeitkoordinate.



$$\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$
Kräfteparallelogramm

2. Für $A = (a_{ij}) \in {}^m\mathbb{R}^n$ sei $-A := (-1)A = (-a_{ij})$.
Ferner gilt für $k \in \mathbb{N}$: $kA = \underbrace{A + \cdots + A}_{k-\text{mal}}$.



3. Für $a_1, a_2, \cdots, a_n \in \mathbb{R}$ wurde schon die Schreibweise

$$\sum_{i=1}^{n} a_i := \sum_{j=1}^{n} a_j := a_1 + a_2 + \cdots + a_n$$

verwendet. Hierbei ist i (bzw. j) der Laufindex.

Sind nun $a_{ij} \in \mathbb{R}$ für alle $1 \leq i \leq m,\ 1 \leq j \leq n$, so gilt:

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} \right).$$

Dies ist klar, da in

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

bei $\sum_i (\sum_j a_{ij})$ zuerst zeilenweise addiert und bei $\sum_j (\sum_i a_{ij})$ zuerst spaltenweise addiert und dann aufsummiert wird.

Für diese Doppelsummen verwendet man dann auch die Schreibweise

$$\sum_{ij} a_{ij} := \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} \right).$$

Satz 2.8 (Rechenregeln der Multiplikation von Matrizen und Spaltenvektoren)**.** *Seien $m, n, p, q \geq 1$. Dann gilt für alle $A, A' \in {}^m\mathbb{R}^n,\ B, B' \in {}^n\mathbb{R}^p,\ C \in {}^p\mathbb{R}^q,\ r \in \mathbb{R}$:*

1. (a) $(AB)C = A(BC)$ (**assoziativ**)
   (b) $(A + A')B = AB + A'B$ (**distributiv**)
   $A(B + B') = AB + AB'$ (**distributiv**)
   (c) $AE = A\ ,\ EB = B$ (**neutrales Element**)
   Hierbei ist E die Einheitsmatrix in ${}^n\mathbb{R}^n$.
2. $r(AB) = (rA)B = A(rB)$

Beweis. Seien $A = (a_{ij})_{i,j}, A' = (a'_{ij})_{i,j} \in {}^m\mathbb{R}^n,\ B = (b_{jk})_{j,k}, B' = (b'_{jk})_{j,k} \in {}^n\mathbb{R}^p$, $C = (c_{kl})_{k,l} \in {}^p\mathbb{R}^q$.

1. (a) Es ist

$$\begin{aligned} AB &= \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right)_{i,k} \in {}^m\mathbb{R}^p, \\ BC &= \left(\sum_{k=1}^{p} b_{jk} c_{kl}\right)_{j,l} \in {}^n\mathbb{R}^q. \end{aligned}$$

Der Koeffizient von $(AB)C$ an der Stelle (i, l) ist:

$$\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) c_{kl}.$$

Der Koeffizient von $A(BC)$ an der Stelle (i, l) ist:

$$\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk} c_{kl}\right).$$

Wegen

$$\begin{aligned} \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) c_{kl} &= \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij} b_{jk} c_{kl}\right) \\ &= \sum_{j=1}^{n}\left(\sum_{k=1}^{p} a_{ij} b_{jk} c_{kl}\right) \\ &= \sum_{k=1}^{n} a_{ij}\left(\sum_{j=1}^{n} b_{jk} c_{kl}\right) \end{aligned}$$

für alle $1 \leq i \leq n,\ 1 \leq l \leq q$, folgt $(AB)C = A(BC)$.

(b) Es gilt

$$\begin{aligned} (A + A')B &= \left(\sum_j (a_{ij} + a'_{ij}) b_{jk}\right)_{i,k} \\ &= \left(\sum_j (a_{ij} b_{jk} + a'_{ij} b_{jk})\right)_{i,k} \\ &= \left(\sum_j a_{ij} b_{jk}\right)_{i,k} + \left(\sum_j a'_{ij} b_{jk}\right)_{i,k} \\ &= AB + A'B. \end{aligned}$$

Ebenso folgt $A(B + B') = AB + AB'$.

(c) Es ist

$$AE = \left(\sum_j a_{ij} \delta_{ik}\right)_{i,k} = (a_{ij})_{i,j} = A.$$

Ebenso folgt $EB = B$.

2. Es gilt

$$r(AB) = r\left(\sum_j a_{ij}b_{jk}\right)_{i,k} = \left(\sum_j r a_{ij}b_{jk}\right) = (rA)B,$$
$$r(AB) = r\left(\sum_j a_{ij}b_{jk}\right)_{i,k} = \left(\sum_j a_{ij} r b_{jk}\right) = A(rB).$$

□

Wegen der Assoziativität der Matrizenmultiplikation (2.8, 1a) ist folgende Definition sinnvoll.

Definition 2.9. Sei $A \in {}^n\mathbb{R}^n$ eine quadratische Matrix. Dann sei

$$A^0 := E$$

die Einheitsmatrix, und für alle $k \in \mathbb{N}$ mit $k \geq 1$ sei

$$A^k := \underbrace{A \cdot A \cdot \ldots \cdot A}_{k-\text{mal}}.$$

Bemerkung 2.10. Nach 2.5 und 2.8 gelten viele der üblichen Rechenregeln auch für Matrizen. Im allgemeinen ist die Matrizenmultiplikation aber **nicht kommutativ**:

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in {}^2\mathbb{R}^2$$
$$\Rightarrow \quad AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = BA$$

Als Anwendung der Matrizenmultiplikation soll gezeigt werden, dass man mit der Matrizenmultiplikation Aussagen über gerichtete Graphen gewinnen kann.

Definition 2.11. Sei V eine endliche Menge und $E \subset V \times V$.

1. $G := (V, E)$ heißt **gerichteter Graph** $:\Leftrightarrow \quad \forall\, (x, y) \in E : x \neq y$.
2. Sei (V, E) gerichteter Graph. Dann heißen die Elemente aus V **Punkte** oder **Vertices** (vertices) und die Elemente aus E **Kanten** (edges). Für eine Kante $(x, y) \in E$ schreibt man auch $x \to y$.

Bemerkung 2.12. Sei (V, E) ein gerichteter Graph, dessen Punkte numeriert sind, d.h., es ist $V = \{x_1, x_2, \ldots, x_n\}$ mit $n \in \mathbb{N}$ und $x_i \neq x_j$ für alle $i \neq j$. Dann wird durch

$$a_{ij} = \begin{cases} 1, & \text{falls } (x_i, x_j) \in E, \text{ d.h. } x_i \to x_j \\ 0, & \text{sonst} \end{cases}$$

für alle $1 \leq i, j \leq n$ eine Matrix $A := (a_{ij}) \in {}^n\mathbb{R}^n$ definiert. Diese Matrix heißt **Adjazenzmatrix** von (V, E).

Umgekehrt definiert jede Matrix $A = (a_{ij}) \in {}^n\mathbb{R}^n$ mit $a_{ij} \in \{0, 1\}\ \forall\ i, j$ und $a_{ii} = 0\ \forall\ i$ einen gerichteten Graphen mit A als Adjazenzmatrix.

Beispiel 2.13. Der gerichtete Graph (V, E) mit $V = \{x_1, x_2, x_3\}$ und

$$E = \{(x_1, x_3), (x_1, x_2), (x_3, x_1), (x_2, x_3)\}$$

wird durch

$$\begin{tikzcd}
x_1 \arrow[rr] \arrow[dr, shift left] & & x_2 \arrow[dl] \\
& x_3 \arrow[ul, shift left] &
\end{tikzcd}$$

veranschaulicht. Dieser Graph besitzt die Adjazenzmatrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Definition 2.14. Sei (V, E) gerichteter Graph, $k \in \mathbb{N}$ und $v_0, v_1, \ldots, v_k \in V$.

Das $(k+1)$-Tupel $(v_0, v_1, \ldots, v_k)$ heißt **gerichteter Kantenzug** der Länge k von v_0 nach v_k $:\Leftrightarrow$ $(v_i, v_{i+1}) \in E \;\forall\; 0 \leq i \leq k-1$, d.h. $v_0 \to v_1 \to v_2 \to \ldots \to v_{k-1} \to v_k$.

Beispiel 2.15. In

$$\begin{tikzcd}
x_1 \arrow[rr] \arrow[dr, shift left] & & x_2 \arrow[dl] \\
& x_3 \arrow[ul, shift left] &
\end{tikzcd}$$

ist $x_1 \to x_2 \to x_3 \to x_1 \to x_3$ der Kantenzug $(x_1, x_2, x_3, x_1, x_3)$ der Länge 4.

Bemerkungen 2.16 (Beweisprinzip der vollständigen Induktion)**.**

1. Für jede natürliche Zahl $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Dann gilt $A(n)$ für alle $n \in \mathbb{N}$, falls folgendes gilt:
   (a) **Induktionsanfang**:
   $A(0)$ gilt.
   (b) **Induktionsschritt**:
   Für jede natürliche Zahl $n \geq 0$ gilt: Ist $A(n)$ richtig, so gilt auch $A(n+1)$.
2. Variante von 1:
   Sei $n_0 \in \mathbb{Z}$. Ersetze in 1) $\mathbb{N}$ durch $\{m \in \mathbb{Z} \mid m \geq n_0\}$ und 0 durch n_0.

Satz 2.17. *Sei (V, E) gerichteter Graph, $V = \{v_0, v_1, \ldots, v_n\}$ mit $x_i \neq x_j \;\forall\; i \neq j$ eine Numerierung von V und A die zugehörige Adjazenzmatrix.*

Dann gilt für alle $k \in \mathbb{N}$ mit $k \geq 1$ und $1 \leq i, j \leq n$: Die Anzahl der gerichteten Kantenzüge in (V, E) von x_i nach x_j der Länge k ist der Koeffizient der Matrix A^k an der Stelle (i, j).

Beweis. Induktion nach k:

1. Induktionsanfang: $k = 1$.
   Klar nach Definition der Adjazenzmatrix.
2. Induktionsschritt: $k \to k+1$.
   Sei $k \geq 1$, und die Behauptung sei richtig für k.

Um die Behauptung für $k+1$ zu zeigen, werden Kantenzüge der Länge $k+1$ von x_i nach x_j betrachtet. Jeder solche Kantenzug hat die Form

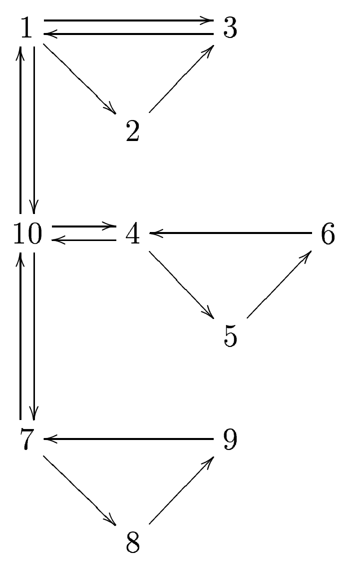
$$x_i \to \underbrace{x_l \to \ldots \to x_j}_{\textit{Kantenzug der Länge } k}$$

für ein x_l, also mit $a_{il} = 1$.

Sei $A^k = (b_{pq})_{p,q} \in {}^n\mathbb{R}^n$. Dann gibt es nach Induktionsvoraussetzung b_{lj} Kantenzüge von x_l nach x_j der Länge k. Also gibt es genau b_{lj} Kantenzüge der Länge $k+1$ von x_i über x_l nach x_j (wie im Bild oben).

Insgesamt gibt es damit $\sum_{l=1}^n a_{il}b_{lj}$ Kantenzüge der Länge $k+1$ von x_i nach x_j. Dies ist aber gerade der Koeffizient von $A^{k+1} = AA^k$ an der Stelle (i,j). □

Beispiel 2.18. 10 Flughäfen mit folgenden direkten Verbindungen:



Es gibt Kantenzüge von i nach j für $(i,j) =$

$$\begin{aligned}
&(1,2), (1,3), (1,10),\\
&(2,3),\\
&(3,1),\\
&(4,5), (4,10),\\
&(5,6),\\
&(6,4)\\
&(7,8), (7,10),\\
&(8,9),\\
&(9,7),\\
&(10,1), (10,4), (10,7).
\end{aligned}$$

Definition 2.19. Seien $1 \le m, n \in \mathbb{N}$.

1. Für eine Matrix $A \in {}^m\mathbb{R}^n$ sei die Abbildung $f_A : {}^n\mathbb{R} \to {}^m\mathbb{R}$ definiert durch $f_A(x) := Ax \;\forall\, x \in {}^n\mathbb{R}$.
2. Eine Abbildung $f : {}^n\mathbb{R} \to {}^m\mathbb{R}$ heißt **linear** $:\Leftrightarrow \forall\, x, y \in {}^n\mathbb{R},\; \alpha \in \mathbb{R}$:

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

d.h., f erhält die algebraischen Operationen mit Spaltenvektoren, nämlich die Addition und die Multiplikation mit Skalaren aus $\mathbb{R}$.

Bemerkungen 2.20.

1. Für $A = (a_{ij}) \in {}^m\mathbb{R}^n$ ist $f_A : {}^n\mathbb{R} \longrightarrow {}^m\mathbb{R}$ wohldefiniert, und es gilt:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R} \;\Rightarrow\; Ax = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

Sei

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in {}^m\mathbb{R}$$

gegeben. Alle $x \in {}^n\mathbb{R}$ mit $f_A(x) = b$, d.h. das Urbild von b bei f_A zu berechnen, bedeutet also, das lineare Gleichungssystem

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

zu lösen.

2. In den Anwendungen kommen häufig "lineare Prozesse" vor, die durch Abbildungen der Form f_A beschrieben werden, z.B.:

   Es sollen n Produkte $P_1, P_2, \dots, P_n$ (z.B. Häusertypen oder Kuchensorten) hergestellt werden. Dazu werden m Rohstoffe oder Zutaten $Z_1, \dots, Z_m$ (z.B. Holz, Glas,... oder Zucker, Eier,...) benötigt.

   Um eine Einheit des Produktes P_j herzustellen, seien a_{ij} Einheiten der Zutaten $Z_i, 1 \le i \le m$ nötig. Um x_j Einheiten des Produktes P_j, $1 \le j \le n$ herzustellen, braucht man also (Linearität) $b_i := \sum_{j=1}^n a_{ij}x_j$ Einheiten von den Zutaten Z_i.

   Somit ist die Abbildung, die jeder Mengenkombination $x \in {}^n\mathbb{R}$ der herzustellenden Produkte $P_1, \dots, P_n$ die Mengenkombination der benötigten Zutaten $Z_1, \dots, Z_m$ zuordnet, durch die lineare Funktion $f_A : {}^n\mathbb{R} \to {}^m\mathbb{R}$, $x \mapsto Ax$ mit $A = (a_{ij})$ gegeben.

Definition 2.21. Sei $1 \le n \in \mathbb{N}$. Für alle $1 \le i \le n$ sei

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in {}^n\mathbb{R},$$

wobei die 1 in der i-ten Zeile steht. $e_1, e_2, \dots, e_n$ heißt **Standardbasis** des ${}^n\mathbb{R}$.

Satz 2.22. *Seien $1 \le m, n, p \in \mathbb{N}$.*

1. *Für alle $A \in {}^m\mathbb{R}^n$ ist $f_A : {}^n\mathbb{R} \to {}^m\mathbb{R}$ eine lineare Abbildung, und für alle $1 \le j \le n$ ist $f_A(e_j)$ die j-te Spalte von A.*
2. *$\forall A \in {}^m\mathbb{R}^n$, $B \in {}^n\mathbb{R}^p$: $f_A \circ f_B = f_{AB}$, also kommutiert*

$$\begin{tikzcd} {}^p\mathbb{R} \arrow[r, "f_B"] \arrow[rr, bend right, "f_{AB}"'] & {}^n\mathbb{R} \arrow[r, "f_A"] & {}^m\mathbb{R} \end{tikzcd}$$

3. *Für jede lineare Abbildung $f : {}^n\mathbb{R} \to {}^m\mathbb{R}$ gibt es genau ein $A \in {}^m\mathbb{R}^n$ mit $f = f_A$.*

Beweis.

1. f_A ist linear, denn für alle $x, y \in {}^n\mathbb{R}$ und $\alpha \in \mathbb{R}$ gilt:
$$f_A(x+y) = A(x+y) = Ax + Ay = f_A(x) + f_A(y),$$
$$f_A(\alpha x) = A(\alpha x) = \alpha(Ax) = \alpha f_A(x).$$
Ferner gilt für $1 \leq j \leq n$:
$$f_A(e_j) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \end{array}$$
$$= \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = j\text{-te Spalte von A}.$$
2. Wegen der Assoziativität der Matrizenmultiplikation nach Satz 2.8 gilt für alle $x \in {}^n\mathbb{R}$
$$(f_A \circ f_B)(x) = f_A(f_B(x)) = f_A(Bx) = A(Bx) = (AB)x = f_{AB}(x).$$
Damit ist $f_A \circ f_B = f_{AB}$ gezeigt.
3. (a) Eindeutigkeit von A:
Sei $A \in {}^n\mathbb{R}^m$ mit $f = f_A$. Dann ist nach 1. für alle $1 \leq j \leq n$ die j-te Spalte von A gleich $f_A(e_j) = f(e_j)$. Damit ist aber A durch f eindeutig festgelegt.
(b) Existenz von A:
Sei $f : {}^n\mathbb{R} \to {}^m\mathbb{R}$ linear. Für alle $1 \leq j \leq n$ sei
$$f(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in {}^n\mathbb{R}.$$
Definiere $A := (a_{ij}) \in {}^m\mathbb{R}^n$, d.h. die j-te Spalte von A ist durch $f(e_j)$ gegeben. Nun ist noch zu zeigen, dass $f(x) = f_A(x)$ für alle $x \in {}^n\mathbb{R}$ gilt. Sei dazu
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}$$
gegeben. Wegen
$$x = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

gilt dann

$$
\begin{aligned}
f(x) &= f(x_1 e_1 + \ldots + x_n e_n) = x_1 f(e_1) + \ldots + x_n f(e_n) \\
&= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix} = (a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax = f_A(x).
\end{aligned}
$$

□

Fazit: Matrizen beschreiben lineare Abbildungen. Dabei entspricht die Matrizenmultiplikation der Hintereinanderausführung linearer Abbildungen.

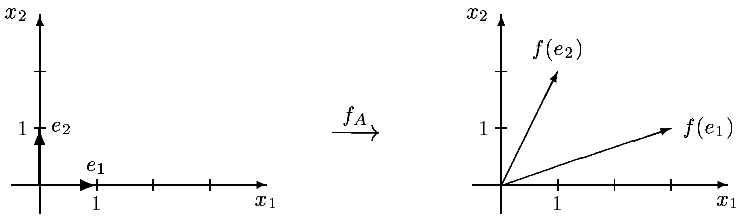
Beispiele 2.23.

1. Sei

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

Dann gilt für $f_A : {}^2\mathbb{R} \to {}^2\mathbb{R}$:

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



2. Für $\varphi \in \mathbb{R}$ sei

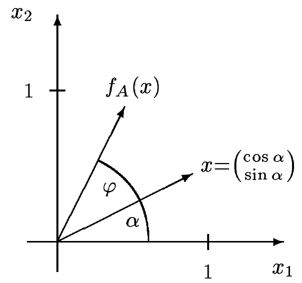
$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Dann ist $f_A : {}^2\mathbb{R} \to {}^2\mathbb{R}$ **Drehung** um den Winkel φ, denn für einen Punkt

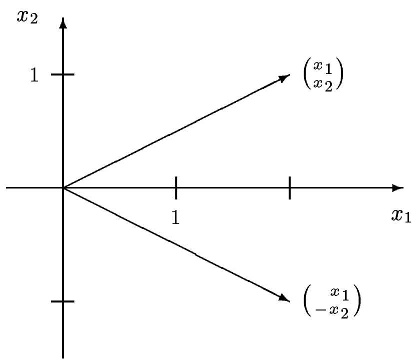
$$x = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \in {}^2\mathbb{R}$$

auf dem Einheitskreis gilt

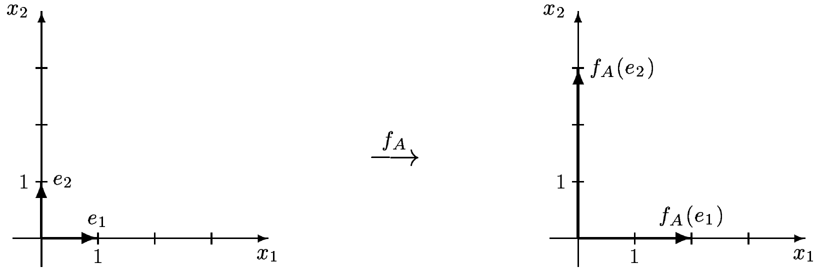
$$
\begin{aligned}
f_A(x) &= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos\varphi\cos\alpha - \sin\varphi\sin\alpha \\ \sin\varphi\cos\alpha + \cos\varphi\sin\alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi+\alpha) \\ \sin(\varphi+\alpha) \end{pmatrix}.
\end{aligned}
$$



3. Für
$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
ist $f_A : {}^2\mathbb{R} \to {}^2\mathbb{R}$ die **Spiegelung** an der x_1-Achse, da:
$$f_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$



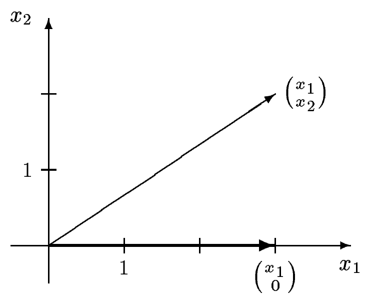
4. Für
$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$
ist $f_A : {}^2\mathbb{R} \to {}^2\mathbb{R}$ **Streckung** in Richtung der x_1- und x_2-Achse, da
$$f_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 3x_2 \end{pmatrix}.$$



5. Für
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

ist $f_A : {}^2\mathbb{R} \to {}^2\mathbb{R}$ **orthogonale Projektion** auf die x_1-Achse, da:

$$f_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}.$$



Beim Rechnen mit Potenzen A^n, $n \geq 1$ von quadratischen Matrizen $A \in {}^m\mathbb{R}^m$ ist der folgende Satz nützlich. Zunächst werden aber noch einige benötigte Begriffe und Hilfsaussagen bereitgestellt.

Definition 2.24.

1. Sei $n \in \mathbb{N}$. Dann heißt

$$n! := \begin{cases} 0, & \text{falls } n = 0 \\ 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n, & \text{sonst} \end{cases}$$

n-Fakultät.

2. Für $n, i \in \mathbb{N}$ mit $0 \leq i \leq n$ heißt

$$\binom{n}{i} := \frac{n!}{i!(n-i)!} \quad \text{(gelesen: } n \text{ über } i\text{)}$$

Binomialkoeffizient.

Lemma 2.25.

1) $\forall\, n \in \mathbb{N} : \binom{n}{0} = 1 = \binom{n}{n}$.

2) $\forall\, n, i \in \mathbb{N},\ n \geq i \geq 1 : \binom{n}{i-1} + \binom{n}{i} = \binom{n+1}{i}$ (**Pascalsches Dreieck**).

Insbesondere sieht man daraus, daß alle Binomialkoeffizienten natürliche Zahlen sind.

Beweis.

1. Klar.
2. Es gilt

$$\begin{aligned} \binom{n}{i-1} + \binom{n}{i} &= \frac{n!}{(i-1)!(n+1-i)!} + \frac{n!}{i!(n-i)!} \\ &= \frac{n!i}{i!(n+1-i)} + \frac{n!(n+1-i)}{i!(n+1-i)!} \\ &= \frac{n!(i+n+1-i)}{i!(n+1-i)!} = \binom{n+1}{i}. \end{aligned}$$

□

Satz 2.26. *Seien* $1 \leq m, n \in \mathbb{N}$ *und* $A, B \in {}^m\mathbb{R}^m$ *(quadratisch) mit* $AB = BA$.

1. **Geometrische Summenformel**:
$$A^{n+1} - B^{n+1} = (A - B)\underbrace{(B^n + AB^{n-1} + A^2B^{n-2} + \ldots + A^n)}_{\sum_{i=0}^n A^iB^{n-i}}$$
2. **Binomische Formel**:
$$(A+B)^n = \underbrace{B^n + \binom{n}{1}AB^{n-1} + \binom{n}{2}A^2B^{n-2} + \ldots + A^n}_{\sum_{i=0}^n \binom{n}{i} A^iB^{n-i}}$$

Beweis.

1. Es gilt
$$\begin{aligned}
& (A-B)(B^n + AB^{n-1} + A^2B^{n-2} + \ldots + A^{n-1}B + A^n) \\
= \; & AB^n + A^2B^{n-1} + \ldots + A^nB + A^{n+1} \\
& -B^{n+1} - AB^n - A^2B^{n-1} - \ldots - A^nB \\
= \; & A^{n+1} - B^{n+1}.
\end{aligned}$$
Hierbei wurde $AB = BA$, also $BAB^{n-1} = AB^n$ usw. verwendet.
2. Induktion nach n:
   (a) Induktionsanfang $n = 1$: Klar.
   (b) Induktionsschritt: $n \to n+1$.
   Sei $n \geq 1$, und sei die Behauptung für n bereits bewiesen. Wir müssen die Behauptung für $n+1$ zeigen:

$$\begin{aligned}
& (A+B)^{n+1} = (A+B)^n(A+B) \\
= \; & \left(\binom{n}{0}B^n + \binom{n}{1}AB^{n-1} + \ldots + \binom{n}{n-1}A^{n-1}B + \binom{n}{n}A^n\right)(A+B) \\
= \; & \binom{n}{0}AB^n + \binom{n}{1}A^2B^{n-1} + \ldots + \binom{n}{n-1}A^nB + \binom{n}{n}A^{n+1} \\
& + \binom{n}{0}B^{n+1} + \binom{n}{1}AB^n + \binom{n}{2}A^2B^{n-1} + \ldots + \binom{n}{n}A^nB \\
= \; & \underbrace{\binom{n}{0}}_{=1}B^{n+1} + \underbrace{\left(\binom{n}{0} + \binom{n}{1}\right)}_{=\binom{n+1}{1}}AB^n + \underbrace{\left(\binom{n}{1} + \binom{n}{2}\right)}_{=\binom{n+1}{2}}A^2B^{n-1} + \ldots + \\
& + \underbrace{\left(\binom{n}{n-1} + \binom{n}{n}\right)}_{=\binom{n+1}{n}}A^nB + \underbrace{\binom{n}{n}}_{=1}A^{n+1} \\
= \; & B^{n+1} + \binom{n+1}{1}AB^n + \binom{n+1}{2}A^2B^{n-1} + \ldots + A^{n+1}.
\end{aligned}$$

□

Potenzen von Matrizen treten in vielen Anwendungen auf: Ein System (Prozeß) werde zur Zeit $k \in \mathbb{N}$ durch den **Zustandsvektor**

$$x(k) = \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{pmatrix} \in {}^n\mathbb{R}$$

beschrieben. Für alle k gelte $x(k+1) = Ax(k)$, wobei $A \in {}^n\mathbb{R}^n$ eine feste Matrix ist. Dann gilt: $x(k) = A^k x(0) \;\forall\; k \in \mathbb{N}$.

Beispiele 2.27.

1. Fibonacci (um 1200):

   Eine Hasenpopulation lebt auf einer Insel. Jedes Paar bekommt jeden Monat ein junges Hasenpaar, wobei jedes Paar erst im Alter von 2 Monaten das erste Mal Junge bekommt.

   Sei F_k die Zahl der Hasenpaare nach k Monaten, wobei $F_0 = 0$ und $F_1 = 1$ sei. Im Monat $k \geq 2$ kommen zu den F_{k-1} Paaren des Vormonats noch F_{k-2} neugeborene Paare hinzu. Somit gilt die Rekursionsformel $F_k = F_{k-1} + F_{k-2}$ für $k \geq 2$.

   Setzt man
   $$x(k) = \begin{pmatrix} F_k \\ F_{k+1} \end{pmatrix}$$
   für $k \geq 0$, so gilt
   $$x(k+1) = \begin{pmatrix} F_{k+1} \\ F_{k+2} \end{pmatrix} = \begin{pmatrix} F_{k+1} \\ F_{k+1} + F_k \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_k \\ F_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} x(k).$$
   Mit
   $$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$
   gilt dann also $x(k) = A^k x(0)$, d.h.
   $$\begin{pmatrix} F_k \\ F_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^k \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
   ist die 2-te Spalte von
   $$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^k.$$

   Ähnlich kann man allgemeine **linear rekursive Folgen** behandeln (siehe Übung).

2. Die Mitarbeiter einer Firma gehen mittags in 2 Restaurants R_1 und R_2 zum Essen. Nach jedem Monat bleiben 68% der Kunden von R_1 bei R_1, und 55% der Kunden von R_2 bleiben R_2 treu.

   Sei $x_i(k)$ der Anteil der Mitarbeiter, die im Monat k im Restaurant R_i essen, d.h.
   $$x_i(k) = \frac{\text{Anzahl der Mitarbeiter die bei } R_i \text{ essen}}{\text{Anzahl aller Mitarbeiter}}$$
   Dann gilt $0 \leq x_i(k) \leq 1$ für $i = 1, 2$ sowie $x_1(k) + x_2(k) = 1$.

Ferner ergeben sich aus den vorgegebenen Daten die Rekursionsgleichungen

$$\begin{aligned} x_1(k+1) &= 0.68x_1(k) + 0.45x_2(k) \\ x_2(k+1) &= 0.32x_1(k) + 0.55x_2(k) \end{aligned}$$

bzw.

$$x(k+1) = \underbrace{\begin{pmatrix} 0.68 & 0.45 \\ 0.32 & 0.55 \end{pmatrix}}_{=:A} x(k) \quad \text{mit} \quad x(k) = \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix},$$

woraus $x(k) = A^k x(0)$ folgt.

Für die Matrix A gilt: Alle Koeffizienten sind ≥ 0, und in jeder Spalte ist die Summe der Einträge $= 1$.

3. Entsprechendes gilt für n Restaurants $R_1, \ldots, R_n$:

$$\begin{aligned} x_i(k) &:= \text{Anteil der Mitarbeiter, die bei } R_i \text{ essen;} \\ a_{ij} &:= \text{Anteil der Kunden von } R_j, \text{ die nach } R_i \text{ wechseln.} \end{aligned}$$

Mit $A := (a_{ij})$ und

$$x(k) := \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix}$$

sind die Rekursionsgleichungen dann durch $x(k+1) = Ax(k)$ gegeben.

Wieder gilt: $a_{ij} \geq 0 \; \forall \; i, j$ und $\sum_{i=1}^n a_{ij} = 1 \; \forall j$.

Andere Interpretation: a_{ij} ist die Wahrscheinlichkeit, dass ein Kunde von R_j nach R_i wechselt.

Definition 2.28. Sei $1 \leq n \in \mathbb{N}$.

1. $A = (a_{ij}) \in {}^n\mathbb{R}^n$ heißt **stochastische Matrix**
$:\Leftrightarrow \; \forall i, j: \; a_{ij} \geq 0$ und $\forall j: \sum_{i=1}^n a_{ij} = 1$.
2. $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}$ heißt **Wahrscheinlichkeitsvektor**
$:\Leftrightarrow \; \forall i: \; x_i \geq 0$ und $\sum_{i=1}^n x_i = 1$.
3. Sei $A \in {}^n\mathbb{R}^n$ stochastische Matrix und $x \in {}^n\mathbb{R}$ Wahrscheinlichkeitsvektor. Dann heißt die Folge $(A^k x)_{k \in \mathbb{N}}$ **Markov-Prozeß**.

Es gilt: (ohne Beweis)

Satz 2.29 (über Markov-Prozesse). *Sei $A \in {}^n\mathbb{R}^n$ eine stochastische Matrix. Es existiere ein $l \in \mathbb{N}$, so dass alle Koeffizienten von $A^l > 0$ sind.*

Dann gibt es ein Wahrscheinlichkeitsvektor

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \textit{mit} \quad \lim_{k \to \infty} A^k = \begin{pmatrix} y_1 & \cdots & y_1 \\ \vdots & & \vdots \\ y_n & \cdots & y_n \end{pmatrix} \quad \textit{(lauter gleiche Spalten).}$$

Hierbei ist für eine Folge $(B(k))_k$ von Matrizen $B(k) = (b_{ij}(k)) \in {}^n\mathbb{R}^n$ gemeint: $\lim_{k\to\infty} B(k) = (\lim_{k\to\infty} b_{ij}(k))_{ij}$. (Zur Definition des Limes siehe Analysis Vorlesung.)

Für $n = 2$ siehe Übungsblatt 3, Aufgabe 4.
Beispiel mit Maple für stochastische 5×5-Matrix.

Bemerkung 2.30. Sei $A \in {}^n\mathbb{R}^n$ stochastische Matrix und

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{Wahrscheinlichkeitsvektor mit} \quad \lim_{k\to\infty} A^k = \begin{pmatrix} y_1 & \cdots & y_1 \\ \vdots & & \vdots \\ y_n & \cdots & y_n \end{pmatrix}.$$

Dann gilt:

1. $\forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ Wahrscheinlichkeitsvektor: $\lim_{k\to\infty} A^k x = y$.
   Beweis. Es gilt
   $$\lim_{k\to\infty} A^k x = (\lim_{k\to\infty} A^k)x = \begin{pmatrix} y_1 & \cdots & y_1 \\ \vdots & & \vdots \\ y_n & \cdots & y_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1(\sum_i x_i) \\ \vdots \\ y_n(\sum_i x_i) \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

2. y ist der eindeutig bestimmte Wahrscheinlichkeitsvektor, der das lineare Gleichungssystems $Ax = x$ löst.
   Beweis.
   (a) Eindeutigkeit: Sei x Wahrscheinlichkeitsvektor mit $Ax = x$. Dann gilt
   $$x = \lim_{k\to\infty} x \overset{\text{Vor.}}{=} \lim_{k\to\infty} A^k x \overset{1.}{=} y.$$
   (b) y erfüllt $Ax = x$, denn aus
   $$A \lim_{k\to\infty} A^k = \lim_{k\to\infty} A^{k+1} = \lim_{k\to\infty} A^k = \begin{pmatrix} y_1 & \cdots & y_1 \\ \vdots & & \vdots \\ y_n & \cdots & y_n \end{pmatrix}$$
   folgt $Ay = y$.

Fazit: Mit Hilfe des obigen Satzes lassen sich die stationären Zustände bei gewissen Markov-Prozessen durch Lösen des linearen Gleichungssystems $Ax = x$ berechnen!

3. Lineare Gleichungssysteme

Seien $1 \le m, n \in \mathbb{N}$, $A = (a_{ij}) \in {}^m\mathbb{R}^n$ und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in {}^m\mathbb{R}$.

Problem: Berechne alle x aus ${}^n\mathbb{R}$ mit $Ax = b$, d.h., es sind alle Lösungen

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}$$

des **linearen Gleichungssystems**

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

zu berechnen.

Lösungsmethode: Betrachte die Matrix

$$(A\ b) := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \vdots & & & \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \in {}^m\mathbb{R}^{n+1}.$$

1. Forme $(A\ b)$ in $(A'\ b')$ um, so dass A' eine einfachere Gestalt hat.
2. Lese die Lösungen x mit $A'x = b'$ direkt ab.

Definition 3.1. Seien $A, A' \in {}^m\mathbb{R}^n$.

A' entsteht aus A durch eine **elementare Zeilenumformung** $:\Leftrightarrow$ A' entsteht aus A durch eine der folgenden Operationen:

① Vertauschen zweier Zeilen.
② Multiplikation einer Zeile mit einem $0 \neq \alpha \in \mathbb{R}$.
③ Addition einer Zeile zu einer anderen.

Lemma 3.2 (Hilfssatz). *Seien $A, A' \in {}^m\mathbb{R}^n$ und $b, b' \in {}^m\mathbb{R}$. $(A'\ b')$ entstehe aus $(A\ b)$ durch endlich viele elementare Zeilenumformungen.*

Dann gilt $\{x \in {}^n\mathbb{R} \mid Ax = b\} = \{x \in {}^n\mathbb{R} \mid A'x = b'\}$.

Beweis. Klar für Operationen ① und ②.

Betrachte Umformung vom Typ ③, z.B. Addition der ersten Zeile zur zweiten: Seien $a_1, \dots, a_m$ die Zeilen von A, also

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Dann gilt $\forall x \in {}^n\mathbb{R}$:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \Leftrightarrow \underbrace{\begin{pmatrix} a_1 \\ a_1 + a_2 \\ a_3 \\ \vdots \\ a_m \end{pmatrix}}_{=A'} x = \underbrace{\begin{pmatrix} b_1 \\ b_1 + b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix}}_{=b'},$$

denn:

$\Rightarrow$: $a_1x = b_1,\ a_2x = b_2 \quad \Rightarrow \quad (a_1 + a_2)x = a_1x + a_2x = b_1 + b_2 \quad \surd$

$\Leftarrow$: $a_1x = b_1,\ \underbrace{(a_1 + a_2)x}_{a_1x + a_2x} = b_1 + b_2 \quad \Rightarrow \quad a_2x = b_2 \quad \surd$

□

Definition 3.3. Sei $A \in {}^m\mathbb{R}^n$.

A heißt **Stufenmatrix** $:\Leftrightarrow$ $A = 0$ oder

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1j_1} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & a_{1n} \\ 0 & \dots & & & \dots & 0 & a_{2j_2} & \dots & \dots & \dots & \dots & a_{2n} \\ \vdots & & & & & & & & & & & \vdots \\ 0 & \dots & & & & & & \dots & 0 & a_{rj_j} & \dots & a_{rn} \\ 0 & \dots & & & & & & & & & \dots & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & \dots & & & & & & & & & \dots & 0 \end{pmatrix}$$

wobei $1 \leq r \leq n, m$ und $1 \leq j_1 < \ldots < j_r \leq n$.

Dann heißen $a_{1j_1}, \ldots, a_{rj_r}$ Angelpunkte (=**pivots**), und die Spalten zu $j_1, \ldots, j_r$ heißen **Stufenspalten**.

Beispiele 3.4.

1.
$$\begin{pmatrix} 0 & -1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$
ist Stufenmatrix mit $j_1 = 2$, $j_2 = 4$, $j_3 = 5$.

2.
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$
ist Stufenmatrix mit $j_1 = 1$, $j_2 = 2$, $j_3 = 3$.

3.
$$\begin{pmatrix} 0 & 1 & 4 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$
ist **keine** Stufenmatrix.

Lemma 3.5. *Sei $A \in {}^m\mathbb{R}^n$.*

Dann existiert eine Stufenmatrix $A' \in {}^m\mathbb{R}^n$, die aus A durch endlich viele elementare Zeilenumformungen entsteht.

Beweis. Falls $A = 0$ ist, so ist die Behauptung mit $A' = A$ erfüllt.

Falls $A \neq 0$ ist, sei $j_1 := \min\{j \mid j\text{-te Spalte von } A \text{ ist } \neq 0\}$ (erste Spalte $\neq 0$) sowie $1 \leq i \leq m$ mit $a_{ij_1} \neq 0$. Anwendung elementarer Zeilenumformungen liefert

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1j_1} & \dots \\ \vdots & & \vdots & \vdots & \\ \vdots & & \vdots & a_{ij_1} & \dots \\ \vdots & & \vdots & \vdots & \\ 0 & \dots & 0 & a_{mj_1} & \dots \end{pmatrix} \quad \underset{\substack{\text{vertausche} \\ \text{1. und } i\text{-te Zeile}}}{\longmapsto}$$

$$A' := \begin{pmatrix} 0 & \dots & 0 & a'_{1j_1} & \dots & a'_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a'_{mj_1} & \dots & a'_{mn} \end{pmatrix} \xmapsto[\substack{\text{addieren } \forall\, l\neq 1, a'_{lj_1}\neq 0}]{-\frac{a'_{lj_1}}{a'_{1j_1}}\cdot(1 \text{ Zeile}) \text{ zur l-ten Zeile}}$$

$$\begin{pmatrix} 0 & \dots & 0 & a'_{1j_1} & \dots & a'_{1n} \\ \vdots & & \vdots & 0 & & \\ \vdots & & \vdots & \vdots & \text{Rest-} & \\ \vdots & & \vdots & \vdots & \text{matrix} & \\ 0 & \dots & 0 & 0 & & \end{pmatrix}$$

Hierbei beachte man, dass $a'_{1j_1} = a_{ij_1} \neq 0$ gilt.

Dann Iteration diese Verfahrens mit der Restmatrix. □

Beispiele 3.6.

1.

$$\begin{pmatrix} 0 & 0 & 2 & 1 & 3 \\ 1 & 0 & 0 & -1 & 2 \\ 1 & 0 & 2 & 0 & 5 \end{pmatrix} \xmapsto[\text{1. und 2. Zeile}]{\text{vertausche}}$$

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & 1 & 3 \\ 1 & 0 & 2 & 0 & 5 \end{pmatrix} \xmapsto[\text{addieren}]{\text{-1}\cdot\text{1. Zeile zur 3. Zeile}}$$

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 2 & 1 & 3 \end{pmatrix} \xmapsto[\text{addieren}]{\text{-1}\cdot\text{2. Zeile zur 3. Zeile}}$$

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

2. Seien

$$A = \begin{pmatrix} 0 & 0 & 2 & 1 & 3 \\ 1 & 0 & 0 & -1 & 2 \\ 1 & 0 & 2 & 0 & 5 \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 2 \\ 8 \end{pmatrix},$$

$$A' = \begin{pmatrix} 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad b' = \begin{pmatrix} 2 \\ 6 \\ 0 \end{pmatrix}.$$

Nach 1. und Lemma 3.2 gilt $\forall\, x \in {}^5\mathbb{R}: \; Ax = b \;\Leftrightarrow\; A'x = b'$.

Mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix}$$

ist $A'x = b'$ aber gleichbedeutend mit

$$\begin{array}{ccccccc} x_1 & & - & x_4 & + & 2x_5 & = & 2 \\ & 2x_3 & + & x_4 & + & 3x_5 & = & 6 \end{array}$$

Die Lösungen dieses Gleichungssystems lassen sich sehr leicht von unten nach oben ausrechnen: Seien $x_2, x_4, x_5 \in \mathbb{R}$ beliebig. Dann berechnet man x_1 und $x_3 \in \mathbb{R}$, so dass

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix}$$

eine Lösung des Gleichungssystems ist, wie folgt:

- Aus der untersten Gleichung ergibt sich $x_3 = \frac{1}{2}(6 - x_4 - 3x_5)$.
- Aus der ersten Gleichung erhält man $x_1 = (2 + x_4 - 2x_5)$.

Ergebnis: Alle Lösungen des Gleichungssystems sind

$$\begin{pmatrix} 2 + x_4 - 2x_5 \\ x_2 \\ \frac{1}{2}(6 - x_4 - 3x_5) \\ x_4 \\ x_5 \end{pmatrix} \in {}^5\mathbb{R}$$

oder mit beliebigen $r = x_2,\ s = x_4,\ t = x_5 \in \mathbb{R}$

$$x = \begin{pmatrix} 2 + s - 2t \\ r \\ \frac{1}{2}(6 - s - 3t) \\ s \\ t \end{pmatrix}.$$

Satz 3.7 (Gauß-Algorithmus)**.** *Sei $A \in {}^m\mathbb{R}^n$ und $b \in {}^m\mathbb{R}$.*

Dann existieren (nach dem Verfahren aus Lemma 3.5) $A' \in {}^m\mathbb{R}^n$ und $b' \in {}^m\mathbb{R}$ mit

- *$(A'\ b')$ entsteht aus $(A\ b)$ durch endlich viele Zeilenumformungen,*
- *A' ist Stufenmatrix,*

und für $L := \{x \in {}^n\mathbb{R} \mid Ax = b\}$ gilt:

1. *$L \neq \emptyset \ \Leftrightarrow\ b'_{r+1} = \ldots = b'_m = 0$.*
   (r wie in der Definition der Stufenmatrix für A'; $r = 0$, falls $A' = 0$.)
2. *Falls $L \neq \emptyset$ ist, so existieren zu beliebig vorgegebenen $x_j \in \mathbb{R}$, $j \notin \{j_1, \ldots, j_r\}$, d.h. j kein Index einer Stufenspalte von A', eindeutig bestimmte $x_{j_1}, \ldots, x_{j_r} \in \mathbb{R}$, so dass*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in L.$$

Hierbei können $x_{j_1}, \ldots, x_{j_r}$ aus $A'x = b'$ von unten nach oben berechnet werden.
Es gibt also $\mathbf{n - r}$ **freie Parameter**.

Beweis. Die Existenz von A' und b' mit den geforderten Eigenschaften ergibt sich aus Lemma 3.5. Somit sind nur noch die beiden Aussagen über die Lösungsmenge L zu zeigen.

1. Nach Lemma 3.2 gilt für alle $x \in {}^n\mathbb{R}:\ Ax = b \ \Leftrightarrow\ A'x = b'$.
   Aus der Form der Stufenmatrix A' folgt:

$$A'x = b' \text{ lösbar} \quad \Leftrightarrow \quad b'_{r+1} = \ldots = b'_m = 0.$$

   Vergleiche hierzu das Beispiel.

2. Folgt wie im Beispiel. □

Folgerung 3.8. *Seien $m, n \in \mathbb{N}$ mit $m < n$ und $A \in {}^m\mathbb{R}^n$.*
Dann existiert $0 \neq x \in {}^n\mathbb{R}$ mit $Ax = 0$, d.h., das Gleichungssystem $Ax = 0$ hat eine nichttriviale Lösung, falls es mehr Unbekannte als Gleichungen gibt.

Beweis. Nach Satz 3.7 ist die Anzahl der freien Parameter der Lösungen gleich $n - r$. Wegen $r \leq m$ und $m < n$ gilt $n - r \geq n - m > 0$, d.h., es gibt mindestens einen freien Parameter. Also existiert eine nichttriviale Lösung. □

Definition 3.9. Sei $A \in {}^n\mathbb{R}^n$. A heißt **invertierbar** $:\Leftrightarrow \exists B \in {}^n\mathbb{R}^n : \; AB = E = BA$. Hierbei bezeichnet E die Einheitsmatrix in ${}^n\mathbb{R}^n$.

Bemerkungen 3.10.

1. Sei $A \in {}^n\mathbb{R}^n$. Falls $B, C \in {}^n\mathbb{R}^n$ existieren mit $AB = E$ und $CA = E$, so gilt $B = C$, und A ist invertierbar.
   Insbesondere ist für invertierbares A die Matrix B in der Definition eindeutig bestimmt, und $A^{-1} = B$ mit $AB = E = BA$ heißt **inverse Matrix** zu A.
   Beweis. Es gilt
   $$B = EB = (CA)B = C(AB) = CE = C.$$
2. Seien $A, B \in {}^n\mathbb{R}^n$ invertierbar. Dann sind auch A^{-1} und AB invertierbar, und es gilt $(A^{-1})^{-1} = A$ und $(AB)^{-1} = B^{-1}A^{-1}$.
   Beweis. Es gilt $A^{-1}A = E = AA^{-1}$, also ist $(A^{-1})^{-1} = A$.
   Ferner gilt
   $$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$$
   und ebenso $(B^{-1}A^{-1})(AB) = E$. Hieraus folgt $(AB)^{-1} = B^{-1}A^{-1}$.
3. Sei $A \in {}^n\mathbb{R}^n$, und es exsitiere ein $k \in \mathbb{N}$ mit $A^{k+1} = 0$. Dann ist $E - A$ invertierbar mit $(E - A)^{-1} = E + A + \ldots + A^k$.
   Beweis. Nach der geometrischer Summenformel (Satz 2.26,1) gilt:
   $$\begin{aligned} E = E^{k+1} - A^{k+1} &= (E - A)(E + A + \ldots + A^k) \\ &= (E + A + \ldots + A^k)(E - A). \end{aligned}$$

Definition 3.11. Sei $A = (a_{ij}) \in {}^m\mathbb{R}^n$. Dann heißt
$$A^T := (a'_{kl}) \in {}^n\mathbb{R}^m \quad \text{mit} \quad a'_{kl} = a_{lk} \; \forall \; 1 \leq k \leq n, \; 1 \leq l \leq m$$
die zu A **transponierte Matrix**.

Beispiele 3.12.

1.
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \\ 3 & 4 & 5 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 0 & 3 \\ 2 & -1 & 4 \\ 3 & 2 & 5 \end{pmatrix}.$$
Somit ergibt sich A^T aus A durch Spiegelung der Einträge an der Hauptdiagonalen.

2.
$$A = \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix}}_{\in {}^2\mathbb{R}^3}, \quad A^T = \underbrace{\begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 2 \end{pmatrix}}_{\in {}^3\mathbb{R}^2}.$$

Bemerkungen 3.13.

1. Seien $A, A' \in {}^m\mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Dann gelten
$$(A^T)^T = A, \quad (A + A')^T = A^T + A'^T, \quad (\alpha A)^T = \alpha A^T.$$
Beweis. Klar, da komponentenweise Struktur.

2. Sei $A \in {}^m\mathbb{R}^n$ und $B \in {}^n\mathbb{R}^p$. Dann gilt $(AB)^T = B^T A^T$.
Beweis. Es ist $B^T \in {}^p\mathbb{R}^n$ und $A^T \in {}^n\mathbb{R}^m$.
Mit $A = (a_{ij})_{i,j}$ und $B = (b_{jk})_{j,k}$ ist der Koeffizient von $B^T A^T$ an der Stelle (k, i), $1 \le k \le p$, $1 \le i \le m$ gleich
$$\sum_{j=1}^{n} b_{jk} a_{ij},$$
da b_{jk} der Koeffizient von B^T an der Stelle (k, j) und a_{ij} der Koeffizient von A^T an der Stelle (j, i) ist.
Andererseits ist
$$\sum_{j=1}^{n} b_{jk} a_{ij} = \sum_{j=1}^{n} a_{ij} b_{jk}$$
das Element an der Stelle (k, i) von $(AB)^T$.

3. Sei $A \in {}^n\mathbb{R}^n$ invertierbar. Dann gilt $(A^T)^{-1} = (A^{-1})^T$.
Beweis. Transponiert man $AA^{-1} = E = A^{-1}A$, so erhält man nach 2. und wegen $E^T = E$
$$(A^{-1})^T A^T = E = A^T (A^{-1})^T,$$
woraus die Behauptung folgt.

Die nächste Folgerung zeigt, dass für $A \in {}^n\mathbb{R}^n$ sogar gilt: A ist invertierbar, falls $B \in {}^n\mathbb{R}^n$ existiert mit $AB = E$ oder $BA = E$.

Folgerung 3.14. *Sei $A \in {}^n\mathbb{R}^n$. Folgende Aussagen sind äquivalent:*

1. $\forall\, x \in {}^n\mathbb{R} : \ Ax = 0 \ \Rightarrow \ x = 0$, *d.h., das Gleichungssystem $Ax = 0$ hat nur die triviale Lösung.*
2. $\exists\, X \in {}^n\mathbb{R}^n : \ AX = E$ *(E = Einheitsmatrix in ${}^n\mathbb{R}^n$).*
3. $\exists\, Y \in {}^n\mathbb{R}^n : \ YA = E$.

Sind diese Aussagen erfüllt, so ist $X = Y = A^{-1}$, *und* A^{-1} *kann wie folgt berechnet werden: Forme* $(A\ E) \in {}^n\mathbb{R}^{2n}$ *durch elementare Zeilenumformungen um in eine Matrix der Form* $(E\ B)$ *mit* $B \in {}^n\mathbb{R}^n$. *Dann gilt:* $B = A^{-1}$.

Beweis.

1. $\Rightarrow$ 2. : Betrachte $(A\ E)$.

Wegen 1. lässt sich A durch elementare Zeilenumformungen in eine Stufenmatrix ohne freie Parameter, d.h. mit $r = n$ umformen:

$$(A\ E) \underset{\text{elem. Zeilumf.}}{\longmapsto} \left(\begin{array}{ccc|c} a'_{11} & & * & \\ & \ddots & & E' \\ 0 & & a'_{nn} & \end{array}\right),$$

wobei $a'_{11}, a'_{22}, \ldots, a'_{nn} \neq 0$. Durch weitere elementare Zeilenumformungen lässt sich erreichen, dass die Einträge $a'_{nn}, a'_{n-1,n-1}, \ldots, a'_{11}$ Eins sowie alle Einträge über diesen Elementen Null werden.

Insgesamt ergibt sich somit

$$(A\ E) \underset{\text{elem. Zeilumf.}}{\longmapsto} (E\ B)$$

für ein $B \in {}^n\mathbb{R}^n$. Dann gilt wegen Lemma 3.2 für $X \in {}^n\mathbb{R}^n$:

$$AX = E \Leftrightarrow EX = B \quad (\text{d.h. } X = B).$$

Also folgt $AB = E$.

2. $\Rightarrow$ 3. : Für alle $x \in {}^n\mathbb{R}$ gilt:

$$A^T x = 0 \underset{\text{Transponieren}}{\Longrightarrow} x^T A = 0 \underset{2.}{\Longrightarrow} 0 = x^T \underbrace{AX}_{E} = x^T, \text{ d.h. } x = 0.$$

Nach 1. $\Rightarrow$ 2. für A^T (statt A) existiert dann $Z \in {}^n\mathbb{R}^n$ mit $A^T Z = E$. Transponieren dieser Gleichung liefert $Z^T A = E$. Mit $Y := Z^T$ gilt somit $YA = E$.

3. $\Rightarrow$ 1. : Sei $x \in {}^n\mathbb{R}$ mit $Ax = 0$. Mit der Matrix Y aus 3. folgt

$$0 = YAx = Ex = x.$$

Außerdem folgt aus 2. und 3. wegen Bemerkung 3.13,1, dass A invertierbar ist mit $X = Y = A^{-1}$. $\square$

Beispiel 3.15. Sei

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

Wegen

$$(A\,|\,E) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right) \underset{II, III \text{ mit } -1 \text{ mult.}}{\overset{III \text{ zu } I \text{ addieren}}{\longmapsto}}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{array}\right) \underset{(-2)\cdot II \text{ zu } I \text{ add.}}{\longmapsto}$$

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

ist

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

Bemerkungen 3.16.

1. In 3.14 haben wir die Matrixgleichung $AX = E,\ A, X \in {}^n\mathbb{R}^n$ durch elementare Zeilenumformungen von $(A\ E)$ berechnet.
2. Allgemeiner kann man für $A \in {}^m\mathbb{R}^n$ und $B \in {}^m\mathbb{R}^p$ eine Matrixgleichung $AX = B$ mit Lösung $X \in {}^n\mathbb{R}^p$ durch simultane Zeilenumformungen von $(A\ B)$ berechnen, denn es gilt:
$$AX = B \ \Leftrightarrow\ \forall\, 1 \le j \le p:\ A \cdot (j\text{-te Spalte von } X) = j\text{-te Spalte von } B.$$

4. Untervektorräume, lineare Abhängigkeit und Basis

Definition 4.1. Sei $k \geq 1$ und $V \subset {}^k\mathbb{R}$ (oder $\mathbb{R}^k$ = Zeilenraum) eine Teilmenge.
V heißt **Untervektorraum** von ${}^k\mathbb{R}$ (bzw. $\mathbb{R}^k$) $:\Leftrightarrow$

1. $0 \in V$.
2. $\forall\, x, y \in V : x + y \in V$.
3. $\forall\, x \in V,\ r \in \mathbb{R}:\ r \cdot x \in V$.

Bemerkungen 4.2.

1. 0 und ${}^k\mathbb{R}$ sind Untervektorräume des ${}^k\mathbb{R}$.

   Beweis. Klar.

2. Seien $m, n \geq 1$ und $f : {}^n\mathbb{R} \to {}^m\mathbb{R}$ linear. Dann gilt:
   (a) Das **Bild** von f
   $$\mathrm{Bi}(f) := \{f(x) \mid x \in {}^n\mathbb{R}\}$$
   ist ein Untervektorraum von ${}^m\mathbb{R}$.
   (b) Der **Kern** von f
   $$\mathrm{Ke}(f) := \{x \in {}^n\mathbb{R} \mid f(x) = 0\}$$
   ist ein Untervektorraum von ${}^n\mathbb{R}$.

   Beweis.
   (a) $\mathrm{Bi}(f)$ ist Untervektorraum von ${}^m\mathbb{R}$, denn:
   (i) $0 \in \mathrm{Bi}(f)$:
   Aus $f(0) = f(0+0) = f(0) + f(0)$ folgt $f(0) = 0$. Somit gilt $0 = f(0) \in \mathrm{Bi}(f)$.
   (ii) $\forall\, u, v \in \mathrm{Bi}(f) : u + v \in \mathrm{Bi}(f)$:
   Seien $u, v \in \mathrm{Bi}(f)$. Dann existieren $x, y \in {}^n\mathbb{R}$ mit $u = f(x)$ und $v = f(y)$. Hieraus folgt
   $$u + v = f(x) + f(y) \underset{f \text{ linear}}{=} f(x+y) \in \mathrm{Bi}(f).$$
   (iii) $\forall\, u \in \mathrm{Bi}(f),\ r \in \mathbb{R} : r \cdot u \in \mathrm{Bi}(f)$:
   Sei $u \in \mathrm{Bi}(f)$. Dann existiert ein $x \in {}^n\mathbb{R}$ mit $u = f(x)$. Es folgt
   $$r \cdot u = r \cdot f(x) \underset{f \text{ linear}}{=} f(r \cdot x) \in \mathrm{Bi}(f).$$
   (b) $\mathrm{Ke}(f)$ ist Untervektorraum von ${}^n\mathbb{R}$, denn:

(i) $0 \in \mathrm{Ke}(f)$, da $f(0) = 0$.

(ii) $\forall\, x, y \in \mathrm{Ke}(f) : x + y \in \mathrm{Ke}(f)$, da

$$f(x+y) = f(x) + f(y) = 0 + 0 = 0.$$

(iii) $\forall x \in \mathrm{Ke}(f),\ r \in \mathbb{R} : r \cdot x \in \mathrm{Ke}(f)$, da

$$f(r \cdot x) = r \cdot f(x) = r \cdot 0 = 0.$$

3. Sei $A \in {}^m\mathbb{R}^n$. Dann ist der Lösungsraum $L = \{x \in {}^n\mathbb{R} \mid Ax = 0\}$ des homogenen linearen Gleichungssystems $Ax = 0$ ein Untervektorraum von ${}^n\mathbb{R}$.

   Beweis. Nach 2. ist $\{x \mid Ax = 0\} = \mathrm{Ke}(f_A)$ ein Untervektorraum.

4. Seien $V_1, \ldots, V_n \subset {}^k\mathbb{R}$ Untervektorräume. Dann sind die **Summe** der V_i's

$$\sum_{i=1}^{n} V_i = V_1 + V_2 + \cdots + V_n := \{v_1 + v_2 + \ldots + v_n \mid \forall\, 1 \le i \le n : v_i \in V_i\}$$

   und der **Durchschnitt** $\bigcap_{i=1}^n V_i$ der V_i's Untervektorräume des ${}^k\mathbb{R}$.

   Beweis. Leicht nachzurechnen.

Definition 4.3. Sei $V \subset {}^k\mathbb{R}$ Untervektorraum, $n \ge 1$ und $v_1, \ldots, v_n \in V$.

1. Für alle $r_1, \ldots, r_n \in \mathbb{R}$ heißt $\sum_{i=1}^n r_i v_i$ **Linearkombination** der v_i's mit Koeffizienten r_i, $1 \le i \le n$.
2. $\mathrm{span}(v_1, \ldots, v_n) := \{\sum_{i=1}^n r_i v_i \mid \forall\ 1 \le i \le n : r_i \in \mathbb{R}\}$ ist die Menge aller Linearkombinationen von $v_1, \ldots, v_n$ mit Koeffizienten in $\mathbb{R}$.
3. $v_1, \ldots, v_n$ heißt **Erzeugendensystem** von V $:\Leftrightarrow$ $V = \mathrm{span}(v_1, \ldots, v_n)$, d.h. jedes $v \in V$ ist Linearkombination von $v_1, \ldots, v_n$.

Bemerkungen 4.4.

1. Seien $v_1, \ldots, v_n \in {}^k\mathbb{R}$. Dann ist $U := \mathrm{span}(v_1, \ldots, v_n)$ der kleinste Untervektorraum von ${}^k\mathbb{R}$, der $v_1, \ldots, v_n$ enthält, d.h., U ist Untervektorraum, und für alle Untervektorräume W mit $v_1, \ldots, v_n \in W$ gilt: $U \subset W$.

   Beweis. Dass U ein Untervektorraum ist, ist klar.

   Sei W Untervektorraum mit $v_1, \ldots, v_n \in W$. Dann gilt für alle $r_1, \ldots, r_n \in \mathbb{R}$:

$$r_1 v_1 + r_2 v_2 + \ldots + r_n v_n \in W.$$

   Also ist $\mathrm{span}(v_1, \ldots, v_n) \subset W$.

2. Sei $A \in {}^m\mathbb{R}^n$ mit Spalten $v_1, \ldots, v_n \in {}^m\mathbb{R}$. Dann gilt $\mathrm{span}(v_1, \ldots, v_n) = \mathrm{Bi}(f_A)$. Dieser Untervektorraum von ${}^m\mathbb{R}$ heißt **Spaltenraum** von A.

   Beweis. Für

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}$$

   gilt $Ax = x_1 v_1 + \ldots + x_n v_1$.

3. Sei $A \in {}^m\mathbb{R}^n$ mit Zeilen $a_1, \ldots, a_m$. Dann heißt $\mathrm{span}(a_1, \ldots, a_m)$ **Zeilenraum** von A.

Definition 4.5. Sei $V \subset {}^k\mathbb{R}$ Untervektorraum und $n \geq 1$, und seien $v_1, \dots, v_n \in V$.

1. $v_1, v_2, \dots, v_n$ heißen **linear unabhängig** $:\Leftrightarrow$
$$\forall\, r_1, \dots, r_n \in \mathbb{R} : \sum_{i=1}^{n} r_i v_i = 0 \quad \Rightarrow \quad r_1 = 0, \dots, r_n = 0.$$
2. $v_1, \dots, v_n$ heißen **linear abhängig** $:\Leftrightarrow$ $v_1, \dots, v_n$ sind nicht linear unabhängig.
3. $v_1, \dots, v_n$ heißt **Basis** von V $:\Leftrightarrow$ $v_1, \dots, v_n$ sind linear unabhängig und bilden ein Erzeugendensystem von V.

Bemerkungen 4.6.

1. Seien $v_1, \dots, v_n \in {}^m\mathbb{R}$, und sei $A \in {}^m\mathbb{R}^n$ Matrix mit den Spalten $v_1, \dots, v_n$. Dann sind äquivalent:
   (a) $v_1, \dots, v_n$ sind linear unabhängig.
   (b) Das homogene lineare Gleichungssystem $Ax = 0$, $x \in {}^n\mathbb{R}$ besitzt nur die triviale Lösung $x = 0$.

   Beweis. Für
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
   gilt $Ax = x_1 v_1 + \ldots + x_n v_n$. Hieraus ergibt sich die Behauptung.

2. Sei $V \subset {}^k\mathbb{R}$ Untervektorraum, und seien $v_1, \dots, v_n \in V$. Dann gilt: $v_1, \dots, v_n$ Basis von V $\Leftrightarrow$
$$\forall\, v \in V\ \exists \text{ eindeutig bestimmte } r_1, \dots, r_n \in \mathbb{R} :\ v = r_1 v_1 + \ldots + r_n v_n.$$

   Beweis.
   $\Rightarrow$: Sei $v_1, \dots, v_n$ Basis von V, und sei $v \in V$.
   (a) Existenz der Darstellung: Klar, da $v_1, \dots, v_n$ Erzeugendensystem von V.
   (b) Eindeutigkeit der Darstellung: Seien $r_1, \dots, r_n$, $s_1, \dots, s_n \in \mathbb{R}$ mit
$$v = \sum_{i=1}^{n} r_i v_i = \sum_{i=1}^{n} s_i v_i.$$
   Dann gilt
$$0 = \sum_{i=1}^{n} r_i v_i - \sum_{i=1}^{n} s_i v_i = \sum_{i=1}^{n} (r_i - s_i) v_i.$$
   Da $v_1, \dots, v_n$ linear unabhängig sind, folgt hieraus $r_i - s_i = 0$, d.h. $r_i = s_i$ für alle $1 \leq i \leq n$.

   $\Leftarrow$: $v_1, \dots, v_n$ ist eine Basis von V, denn:
   (a) Erzeugendensystem: Klar.
   (b) Linear unabhängig: Seien $r_1, \dots, r_n \in \mathbb{R}$ mit $r_1 v_1 + \dots, r_n v_n = 0$. Wegen
$$r_1 v_1 + \ldots + r_n v_n = 0 = 0 v_1 + \ldots, +0 v_n$$
   folgt aus der Eindeutigkeit der Darstellung $r_i = 0$ für alle $1 \leq i \leq n$.

3. Die Standardbasis
$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in {}^n\mathbb{R}$$
ist eine Basis von ${}^n\mathbb{R}$.

Beweis.

(a) Erzeugendensystem: Sei
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}.$$
Dann gilt
$$x = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = x_1 e_1 + \cdots + x_n e_n.$$

(b) Linear unabhängig: Seien $x_1, \ldots, x_n \in \mathbb{R}$ mit
$$0 = x_1 e_1 + \ldots + x_n e_n = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$
Dann gilt $x_i = 0$ für alle $1 \leq i \leq n$.

Satz 4.7 (Invarianz der Basislänge)**.** *Sei $V \subset {}^k\mathbb{R}$ Untervektorraum.*

1. *Sind $v_1, \ldots, v_n \in V$ linear unabhängig, und ist $w_1, \ldots, w_m \in V$ ein Erzeugendensystem von V, so gilt $n \leq m$.*
2. *Je zwei Basen von V haben die gleich Länge, d.h., sind $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ Basen von V, so gilt $n = m$.*

Beweis.

1. Da $w_1, \ldots, w_m$ ein Erzeugendensystem von V ist, gibt es $a_{ij} \in \mathbb{R}$ für $1 \leq i \leq m$, $1 \leq j \leq n$ mit $v_j = \sum_{i=1}^m a_{ij} w_i$ für alle $1 \leq j \leq n$. Sei $A := (a_{ij}) \in {}^m\mathbb{R}^n$.

   Annahme: $n > m$. Dann hat das homogene lineare Gleichungssystem $Ax = 0$ nach 3.8 eine Lösung
$$0 \neq x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}.$$
   Damit gilt
$$\sum_{j=1}^n x_j v_j = \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m \underbrace{\left(\sum_{j=1}^n a_{ij} x_j \right)}_{\substack{=0\ \forall i, \\ \text{da } Ax=0}} w_i = 0.$$

Dies ist ein Widerspruch zu $x \neq 0$ und $v_1, \ldots, v_n$ linear unabhängig. Also ist die Annahme falsch, und es gilt $n \leq m$.

2. Folgt aus 1. $\square$

Satz 4.8 (Basisauswahlsatz). *Sei* $0 \neq V \subset {}^k\mathbb{R}$ *Untervektorraum, und sei* $v_1, \ldots, v_n$ *ein Erzeugendensystem von* V.

Dann erhält man durch Weglassen passender Elemente des Erzeugendensystems eine Basis von V.

Beweis. Falls $v_1, \ldots, v_n$ linear unabhängig sind, so ist man fertig.

Falls $v_1, \ldots, v_n$ linear abhängig sind, existieren $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ und ein $i \in \mathbb{N}$ mit $1 \leq i \leq n$, $\alpha_i \neq 0$ und $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$. Dann gilt

$$v_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(\frac{-\alpha_j}{\alpha_i} \right) v_j.$$

Also ist $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ ein Erzeugendensystem von V. Durch Iteration dieses Verfahrens folgt die Behauptung. $\square$

Satz 4.9 (Basisergänzungssatz). *Sei* $V \subset {}^k\mathbb{R}$ *Untervektorraum und* $n \geq 1$, *und seien* $v_1, \ldots, v_n \in V$ *linear unabhängig.*

Dann lassen sich $v_1, \ldots, v_n$ *zu einer Basis von* V *ergänzen, d.h., es gibt* $l \geq 0$ *und* $v_{n+1}, \ldots, v_{n+l} \in V$, *so dass* $v_1, \ldots, v_n, \ldots, v_{n+l}$ *eine Basis von* V *ist.*

Beweis. Falls $v_1, \ldots, v_n$ ein Erzeugendensystem von V ist, so ist man fertig.

Falls $v_1, \ldots, v_n$ kein Erzeugendensystem von V ist, so existiert ein $v_{n+1} \in V$ mit $v_{n+1} \notin \operatorname{span}(v_1, \ldots, v_n)$. Dann sind $v_1, \ldots, v_{n+1}$ linear unabhängig, und es gilt $n+1 \leq k$, denn:

Seien $\alpha_1, \ldots, \alpha_{n+1} \in \mathbb{R}$ mit $\alpha_1 v_1 + \ldots + \alpha_{n+1} v_{n+1} = 0$. Wegen $v_{n+1} \notin \operatorname{span}(v_1, \ldots, v_n)$ muss $\alpha_{n+1} = 0$ gelten. Somit ist $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$, woraus $\alpha_1 = 0, \ldots, \alpha_n = 0$ folgt, da $v_1, \ldots, v_n$ linear unabhängig sind.

Außerdem folgt aus Satz 4.7 $n + 1 \leq k$, da $e_1, \ldots, e_k$ ein Erzeugendensystem von ${}^k\mathbb{R}$ ist.

Durch Iteration folgt die Behauptung. Dabei kommt es zu einem Abbruch, da nicht mehr als k Vektoren in ${}^k\mathbb{R}$ linear unabhängig sein können. $\square$

Definition 4.10. Sei $V \subset {}^k\mathbb{R}$ Untervektorraum. Dann heißt

$$\dim V := \begin{cases} 0, & \text{falls } V = 0 \\ n, & \text{falls } V \text{ Basis der Länge } n \geq 1 \text{ besitzt} \\ \infty, & \text{sonst} \end{cases}$$

Dimension von V. ($\emptyset$ ist Basis von 0.)

Folgerung 4.11. *Sei* $V \subset {}^k\mathbb{R}$ *Untervektorraum.*

Dann besitzt V *eine Basis der Länge* $n \leq k$, *und* $\dim V = n$ *ist wohldefinierte natürliche Zahl, d.h. unabhängig von der Wahl der Basis.*

Beweis. Für $V = 0$ ist die Behauptung klar.

Falls $V \neq 0$ ist, existiert $0 \neq v \in V$. Dann ist $v = v_1$ linear unabhängig. Nach Satz 4.9 lässt sich v_1 zu einer Basis von V ergänzen, und die Basislänge ist wegen Satz 4.7 eindeutig bestimmt. □

Folgerung 4.12. *Sei $V \subset {}^k\mathbb{R}$ Untervektorraum, und sei $n = \dim V$.*

1. *Für $v_1, \dots, v_n \in V$ sind folgende Aussagen äquivalent:*
   (a) *$v_1, \dots, v_n$ ist ein Erzeugendensystem von V.*
   (b) *$v_1, \dots, v_n$ sind linear unabhängig.*
   (c) *$v_1, \dots, v_n$ ist eine Basis von V.*
2. *Für $U \subset V$ Untervektorraum sind äquivalent:*
   (a) *$U = V$.*
   (b) *$\dim U = \dim V$.*

Beweis. Es wird nur 1. gezeigt, da 2. klar ist.

$(a) \Rightarrow (c)$: Nach Satz 4.8 kann man $l \geq 0$ Vektoren von $v_1, \dots, v_n$ weglassen und erhält dann eine Basis von V. Wegen der Invarianz der Basislänge (Satz 4.7) muss diese Basis aus $\dim V = n$ Elementen bestehen. Also gilt $l = 0$, d.h. $v_1, \dots, v_n$ ist eine Basis von V.

$(b) \Rightarrow (c)$: Nach Satz 4.9 gibt es $l \geq 0$ Vektoren, die zusammen mit $v_1, \dots, v_n$ eine Basis von V bilden. Wegen Satz 4.7 muss wieder $l = 0$ gelten. Somit ist $v_1, \dots, v_n$ eine Basis von V.

$(c) \Rightarrow (a), (b)$: Trivial. □

Problem: Wie findet man eine Basis von $\operatorname{span}(v_1, \dots, v_n)$?

Satz 4.13. *Seien $v_1, \dots, v_m \in \mathbb{R}^n$ (Zeilen), und sei*

$$A := \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \in {}^m\mathbb{R}^n$$

die Matrix mit den Zeilen $v_1, \dots, v_m$. Desweiteren sei A' eine Stufenmatrix, die aus A durch elementare Zeilenumformungen entsteht.

Dann bilden die Zeilen $\neq 0$ von A' eine Basis von $\operatorname{span}(v_1, \dots, v_m)$.

Beweis.

1. Zeige: Entsteht A' aus A durch elementare Zeilenumformungen, so gilt

$$\text{Zeilenraum von } A = \text{Zeilenraum von } A'.$$

Beweis. Dies ist klar für das Vertauschen zweier Zeilen oder die Multiplikation einer Zeile mit $0 \neq \alpha \in \mathbb{R}$.

Betrachte nun die Addition einer Zeile zu einer anderen, z.B.

$$A = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_m \end{pmatrix} \longmapsto A' = \begin{pmatrix} v_1 \\ v_1 + v_2 \\ v_3 \\ \vdots \\ v_m \end{pmatrix}.$$

Dann ist $\mathrm{span}(v_1, \dots, v_m) = \mathrm{span}(v_1, v_1 + v_2, v_3, \dots, v_m)$ zu zeigen. Wegen $v_2 = (v_1 + v_2) - v_1$ gilt $\mathrm{span}(v_1, \dots, v_m) \subseteq \mathrm{span}(v_1, v_1 + v_2, v_3, \dots, v_m)$. Und die umgekehrte Inklusion $\supseteq$ ist klar.

2. Zeige noch: Ist

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1j_1} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & a_{1n} \\ 0 & \dots & & & \dots & 0 & a_{2j_2} & \dots & \dots & \dots & \dots & a_{2n} \\ \vdots & & & & & & & & & & & \vdots \\ 0 & \dots & & & & & & \dots & 0 & a_{rj_j} & \dots & a_{rn} \\ 0 & \dots & & & & & & & & & \dots & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & \dots & & & & & & & & & \dots & 0 \end{pmatrix}$$

eine Stufenmatrix mit $j_1 < \dots < j_r,\ \ a_{1j_1} \neq 0, \dots, a_{rj_r} \neq 0$, so sind die Zeilen $\neq 0$ von A linear unabhängig.

Beweis. Die ersten r Zeilen $v_1, \dots, v_r$ von A sind die Zeilen $\neq 0$.

Seien $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ mit $\alpha_1 v_1 + \dots + \alpha_r v_r = 0$. Da $\alpha_1 v_1 + \dots + \alpha_r v_r$ an der Stelle j_1 den Eintrag $\alpha_1 a_{1j_1}$ besitzt, gilt somit $\alpha_1 a_{1j_1} = 0$. Wegen $a_{1j_1} \neq 0$ folgt hieraus $\alpha_1 = 0$.

Damit lautet nun die ursprüngliche Gleichung $\alpha_2 v_2 + \dots + \alpha_r v_r = 0$. Jetzt kann man ganz analog zu eben nacheinander $\alpha_2 = 0, \dots, \alpha_r = 0$ folgern, indem man diese Gleichung an den Stellen $j_2, \dots, j_r$ betrachtet. $\square$

Beispiel 4.14. Sei

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

Nach Beispiel 3.6,1 ist

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

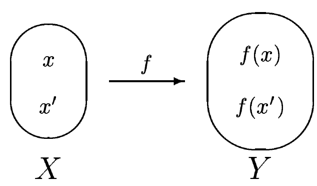
eine Stufenmatrix, die sich aus A durch elementare Zeilenumformungen ergibt. Somit ist $(1\ 0\ 0\ -1\ 2)$, $(0\ 0\ 2\ 1\ 3)$ eine Basis von $\mathrm{span}((0\ 0\ 2\ 1\ 3), (1\ 0\ 0\ -1\ 2), (1\ 0\ 2\ 0\ 5))$.

5. Lineare Abbildungen und Dimensionssatz

Definition 5.1. Seien X, Y Mengen, und sei $f : X \to Y$ eine Abbildung.

1. f heißt **injektiv** $:\Leftrightarrow\ \forall\, x, x' \in X : f(x) = f(x') \ \Rightarrow\ x = x'$.
2. f heißt **surjektiv** $:\Leftrightarrow\ \forall\, y \in Y\ \exists\, x \in X : f(x) = y$, d.h. Y ist gleich dem **Bild** von f: $Y = \mathrm{Bi}(f) := \{f(x) \mid x \in X\}$.
3. f heißt **bijektiv** $:\Leftrightarrow\ f$ ist injektiv und surjektiv.

Bemerkung 5.2. $f : X \to Y$ injektiv bedeutet: $\forall\, x, x' \in X : x \neq x' \ \Rightarrow\ f(x) \neq f(x')$, d.h. f ist "**Einbettung**":



Definition 5.3. Seien $V \subset {}^n\mathbb{R}$, $W \subset {}^m\mathbb{R}$ Untervektorräume, und sei $f : V \to W$ eine Abbildung.

1. f heißt **linear** oder **Homomorphismus** $:\Leftrightarrow$
$$\forall\, x, y \in V,\ \alpha \in \mathbb{R} : f(x+y) = f(x) + f(y),\ f(\alpha x) = \alpha f(x).$$
2. Sei f Homomorphismus.
   (a) f heißt **Monomorphismus** $:\Leftrightarrow$ f injektiv.
   (b) f heißt **Epimorphismus** $:\Leftrightarrow$ f surjektiv.
   (c) f heißt **Isomorphismus** $:\Leftrightarrow$ f bijektiv.
3. Sei f Homomorphismus. Dann heißt $\mathrm{Ke}(f) := \{x \in V \mid f(x) = 0\}$ **Kern** von f.

Bermerkungen 5.4. Sei $f : V \to W$ ein Homomorphismus.

1. Es gilt $f(0) = 0$, da $f(0) = f(0+0) = f(0) + f(0)$.
2. $\mathrm{Bi}(f) \subset W$ und $\mathrm{Ke}(f) \subset V$ sind Untervektorräume.
   Beweis. Wie in Bemerkung 4.2,2.
3. Folgendes Kriterium für f Monomorphismus wird oft angewendet:
   f injektiv, d.h. Monomorphismus $\Leftrightarrow$ $\mathrm{Ke}(f) = 0$.
   Beweis.
   $\Rightarrow$: Sei f injektiv und $v \in \mathrm{Ke}(f)$. Dann gilt $f(v) = 0 = f(0)$. Da f injektiv ist, folgt hieraus $v = 0$.
   $\Leftarrow$: Sei $\mathrm{Ke}(f) = 0$, und seien $v, v' \in V$ mit $f(v) = f(v')$. Dann gilt
   $$0 = f(v) - f(v') = f(v - v').$$
   Also ist $v - v' \in \mathrm{Ke}(f)$. Wegen $\mathrm{Ke}(f) = 0$ folgt hieraus $v - v' = 0$, d.h. $v = v'$.
4. Sei f Isomorphismus und $v_1, \ldots, v_n$ eine Basis von V. Dann ist $f(v_1), \ldots, f(v_n)$ eine Basis von W. Insbesondere ist $\dim V = \dim W$.
   Beweis.
   (a) Erzeugendensystem: Sei $w \in W$. Da f surjektiv ist, existiert dann ein $v \in V$ mit $f(v) = w$. Da $v_1, \ldots, v_n$ ein Erzeugendensystem von V ist, gibt es $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Damit gilt
   $$w = f(v) = f(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n).$$
   (b) Linear unabhängig: Seien $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $\alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) = 0$. Wegen der Linearität von f gilt
   $$0 = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) = f(\alpha_1 v_1 + \ldots + \alpha_n v_n).$$
   Hieraus folgt $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$, da f injektiv ist. Da aber $v_1, \ldots, v_n$ linear unabhängig sind, folgt dann $\alpha_1 = 0, \ldots, \alpha_n = 0$.

Satz 5.5. *Sei $A \in {}^m\mathbb{R}^n$, und sei $A' \in {}^m\mathbb{R}^n$ eine Stufenmatrix, die aus A durch elementare Zeilenumformungen entsteht. Seien $l_1 < \ldots < l_{n-r}$ die Indizes der Spalten von A', die keine Stufenspalten sind, d.h. die Stellen der freien Parameter.*

Dann ist
$$f : \{x \in {}^n\mathbb{R} \mid Ax = 0\} \to {}^{n-r}\mathbb{R},\ x \mapsto \begin{pmatrix} x_{l_1} \\ \vdots \\ x_{l_{n-r}} \end{pmatrix}$$
ein Isomorphismus.

Also ist $\dim\{x \in {}^n\mathbb{R} \mid Ax = 0\} = n-r$, *wobei* r *die Dimension des Zeilenraumes von* A *ist, und man erhält eine Basis von* $\{x \mid Ax = 0\}$ *als Urbilder bei* f *der Standardbasis des* ${}^{n-r}\mathbb{R}$, *d.h. man setze für die freien Parameter der Reihe nach*

$$\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \dots, \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$$

und berechne die zugehörigen Lösungen.

Beweis.

1. Seien $l_1, \dots, l_{n-r}$ die Stellen der freien Parameter. Aus Satz 3.7 folgt, dass
$$f : \{x \in {}^n\mathbb{R} \mid Ax = 0\} \to^{n-r} \mathbb{R},\; f(x) = \begin{pmatrix} x_{l_1} \\ \vdots \\ x_{l_{n-r}} \end{pmatrix} \;\forall\, x \in {}^n\mathbb{R}$$
ein Isomorphismus ist.
2. Wegen Bemerkung 5.4,4 erhält man eine Basis wie angegeben, da f nach 1. ein Isomorphismus ist und da auch die Umkehrabbildung von f ein Isomorphismus ist (siehe Bemerkung 5.7,3).
   Außerdem ist dann $\dim\{x \mid Ax = 0\} = n - r$ mit
$$r = \dim(\text{Zeilenraum von } A') = \dim(\text{Zeilenraum von } A)$$
wegen Satz 4.13. □

Beispiel 5.6. Sei

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

Durch elementare Zeilenumformungen ergibt sich hieraus die Stufenmatrix

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

Also ist

$$\begin{aligned}\{x \in {}^4\mathbb{R} \mid Ax = 0\} &= \{x \in {}^4\mathbb{R} \mid A'x = 0\} \\ &= \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in {}^4\mathbb{R} \mid x_1 - x_4 = 0,\; 2x_3 + x_4 = 0 \right\}.\end{aligned}$$

Die freien Parameter sind an den Stellen 2 und 4. Nun kann wie in Satz 5.5 beschrieben eine Basis von $\{x \in {}^4\mathbb{R} \mid Ax = 0\}$ berechnet werden:

1. Berechne die Lösung mit $x_2 = 1,\; x_4 = 0$: Aus den Gleichungen ergibt sich $x_1 = 0$ und $x_3 = 0$. Somit Lösung:
$$x = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

2. Berechne die Lösung mit $x_2 = 0$, $x_4 = 1$: Es folgt $x_3 = -\frac{1}{2}$ und $x_1 = 1$. Somit Lösung:
$$x = \begin{pmatrix} 1 \\ 0 \\ -\frac{1}{2} \\ 1 \end{pmatrix}.$$

Ergebnis:
$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$
ist eine Basis von $\{x \in {}^4\mathbb{R} \mid Ax = 0\}$.

Bemerkungen 5.7.

1. Seien X, Y, Z Mengen und $f : X \to Y$, $g : Y \to Z$ Abbildungen. Dann gilt:
   (a) gf injektiv $\Rightarrow$ f injektiv.
   (b) gf surjektiv $\Rightarrow$ g surjektiv.
2. Seien X, Y Mengen und $f : X \to Y$ eine Abbildung. Äquivalent sind:
   (a) f bijektiv.
   (b) $\exists\, g : Y \to X$ mit $gf = \mathrm{id}_X$, $fg = \mathrm{id}_Y$.
   In diesem Fall ist g eindeutig bestimmt, und $f^{-1} := g$ heißt **Umkehrabbildung** zu f.
3. Seien $V \subset {}^n\mathbb{R}$, $W \subset {}^m\mathbb{R}$ Untervektorräume und $f : V \to W$ ein Isomorphismus. Dann ist auch f^{-1} ein Isomorphismus.

Beweis. Übungen.

Definition 5.8. Für $V \subset {}^n\mathbb{R}$ Untervektorraum sei
$$V^\perp := \{x \in {}^n\mathbb{R} \mid \forall\, v \in V : v^T x = 0\},$$
gesprochen: V senkrecht.

Beachte: Für
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in {}^n\mathbb{R}$$
gilt
$$y^T x = (y_1, \dots, y_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n y_i x_i.$$

Lemma 5.9. *Sei $V \subset {}^n\mathbb{R}$ Untervektorraum mit Erzeugendemsystem $v_1, \dots, v_m \in V$.*

1. *Dann gilt*
$$V^\perp = \{x \in {}^n\mathbb{R} \mid \underbrace{\begin{pmatrix} v_1^T \\ \vdots \\ v_m^T \end{pmatrix}}_{\in {}^m\mathbb{R}^n} x = 0\}.$$

Insbesondere ist $V^\perp$ ein Untervektorraum von ${}^n\mathbb{R}$.

2. $\dim V^\perp = n - \dim V$.

Beweis.

1. $\subset$: Trivial.
   $\supset$: Sei $x \in {}^n\mathbb{R}$ mit $v_i^T x = 0$ für alle $1 \le i \le m$. Dann ist $v^T x = 0$ für alle $v \in V$ zu zeigen.
   Sei also $v \in V$. Da $v_1, \ldots, v_m$ ein Erzeugendensystem von V ist, existieren dann $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ mit $v = \sum_{i=1}^m \alpha_i v_i$. Damit gilt $v^T x = \sum_{i=1}^m \alpha_i v_i^T x = 0$.
2. Folgt aus 1. und Satz 5.5, da
$$\begin{aligned}\dim V^\perp &\underset{1.}{=} \dim\{x \mid \begin{pmatrix} v_1^T \\ \vdots \\ v_m^T \end{pmatrix} x = 0\} \underset{5.5}{=} n - \dim\ \mathrm{span}(v_1^T, \ldots, v_m^T) \\ &= n - \dim\ \mathrm{span}(v_1, \ldots, v_m) = n - \dim V.\end{aligned}$$

□

Satz 5.10. *Sei $V \subset {}^n\mathbb{R}$ Untervektorraum.*

1. $(V^\perp)^\perp = V$.
2. *Sei $v_1, \ldots, v_d$ ein Erzeugendensystem von V, und sei $a_1, \ldots, a_m$ eine Basis von*
$$\{x \in {}^n\mathbb{R} \mid \begin{pmatrix} v_1^T \\ \vdots \\ v_d^T \end{pmatrix} x = 0\}.$$
   Mit
$$A := \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix} \in {}^m\mathbb{R}^n$$
   *gilt dann: $V = \{x \in {}^n\mathbb{R} \mid Ax = 0\}$, d.h., **jeder** Untervektorraum ist Lösungsraum eines homogenen linearen Gleichungssystems.*

Beweis.

1. Offenbar gilt: $V \subset V^{\perp\perp}$. Wegen
$$\dim(V^\perp)^\perp \underset{5.9}{=} n - dimV^\perp \underset{5.9}{=} n - (n - dimV) = dimV$$
   folgt dann $V = V^{\perp\perp}$ aus 4.12,2.
2. Folgt aus 1. und 5.9,1. □

Bemerkungen 5.11.

1. Seien $A \in {}^m\mathbb{R}^n,\ B \in {}^k\mathbb{R}^n, U := \{x \in {}^n\mathbb{R} \mid Ax = 0\}$ und $V := \{x \in {}^n\mathbb{R} \mid Bx = 0\}$. Dann gilt
$$U \cap V = \{x \in {}^n\mathbb{R} \mid \begin{pmatrix} A \\ B \end{pmatrix} x = 0\} \quad \text{(simultane Lösungen)}.$$
2. Seien $u_1, \ldots, u_m,\ v_1, \ldots, v_k \in {}^n\mathbb{R}$, $U := \mathrm{span}(u_1, \ldots, u_m)$, $V := \mathrm{span}(v_1, \ldots, v_k)$. Dann lässt sich $U \cap V$ mittels 5.10 durch simultanes Lösen des homogen linearen Gleichungssystems wie in 1. berechnen.

Satz 5.12. *Seien $V \subset {}^n\mathbb{R}$, $W \subset {}^m\mathbb{R}$ Untervektorräume, und sei $f : V \to W$ ein Homomorphismus.*

Dann gilt

$$\dim \mathrm{Bi}(f) = \dim V - \dim \mathrm{Ke}(f).$$

(Die Dimension verringert sich um $\dim \mathrm{Ke}(f)$*.)*

Beweis. Sei $v_1, \ldots, v_k$ eine Basis von $\mathrm{Ke}(f)$. Da diese Vektoren linear unabhängig sind, lassen sie sich nach dem Basisergänzungssatz 4.9 zu einer Basis $v_1, \ldots, v_k, v_{k+1}, \ldots, v_l$ von V ergänzen. Sei $U := \mathrm{span}(v_{k+1}, \ldots, v_l)$ (bzw. $U := 0$, falls $k = l$). Betrachtet wird nun die Abbildung $g : U \to \mathrm{Bi}(f)$ mit $g(u) := f(u)$ für alle $u \in U$.

1. Zunächst wird gezeigt, dass g ein Isomorphismus ist.
   (a) g ist Homomorphismus, da g Einschränkung von f ist.
   (b) g Monomorphismus, d.h. $\mathrm{Ke}(g) = 0$:
   Sei $u \in \mathrm{Ke}(g)$, d.h. $u \in U$ mit $g(u) = 0$. Dann gilt $0 = g(u) = f(u)$. Somit ist $u \in \mathrm{Ke}(f) = \mathrm{span}(v_1, \ldots, v_k)$, also insgesamt
   $$u \in \mathrm{span}(v_1, \ldots, v_k) \cap \underbrace{\mathrm{span}(v_{k+1}, \ldots, v_l)}_{=U} = 0.$$
   Hierbei gilt die letzte Gleichung, da $v_1, \ldots, v_k, v_{k+1}, \ldots, v_l$ eine Basis ist. Somit folgt $u = 0$.
   (c) g Epimorphismus:
   Sei $w \in \mathrm{Bi}(f)$. Dann gibt es ein $v \in V$ mit $w = f(v)$. Da $v_1, \ldots, v_l$ eine Basis von V ist, existieren $\alpha_1, \ldots, \alpha_l \in \mathbb{R}$ mit $v = \sum_{i=1}^{l} \alpha_i v_i$. Damit gilt
   $$\begin{aligned} w = f\left(\sum_{i=1}^{l} \alpha_i v_i\right) &\underset{f \text{ linear}}{=} \sum_{i=1}^{l} \alpha_i f(v_i) \\ &\underset{\substack{\text{da } f(v_i)=0 \\ \forall\, 1 \le i \le k}}{=} \sum_{i=k+1}^{l} \alpha_i f(v_i) = f\left(\sum_{i=k+1}^{l} \alpha_i v_i\right) = g(u), \end{aligned}$$
   wobei $u := \sum_{i=k+1}^{l} \alpha_i v_i \in U$.
2. Nach 1. und 5.4,4 folgt
   $$\begin{aligned} \dim \mathrm{Bi}(f) = \dim U &= \dim\ \mathrm{span}(v_{k+1}, \ldots, v_l) \\ &= l - k \qquad (\text{da } v_{k+1}, \ldots, v_l \text{ linear unabhängig}) \\ &= \dim V - \dim \mathrm{Ke}(f). \end{aligned}$$

□

Definition 5.13. Sei $A \in {}^m\mathbb{R}^n$.
1. **Zeilenrang**$(A) := \dim(\text{Zeilenraum von } A)$.
2. **Spaltenrang**$(A) := \dim(\text{Spaltenraum von } A)$.

Folgerung 5.14. *Sei $A \in {}^m\mathbb{R}^n$. Dann gilt*

$$\textit{Zeilenrang}(A) = \textit{Spaltenrang}(A)(= \textit{Zeilenrang}(A^T)).$$

Beweis. Betrachte $f_A : {}^n\mathbb{R} \to {}^m\mathbb{R}$. Nach Satz 5.5 gilt

$$\dim \mathrm{Ke}(f_A) = \dim\{x \in {}^n\mathbb{R} | Ax = 0\} = n - \text{Zeilenrang}(A).$$

Ferner ist $\mathrm{Bi}(f_A) = \mathrm{Spaltenraum}(A)$. Mit dem Dimensionssatz 5.12 folgt damit

$$\begin{aligned}\mathrm{Spaltenrang}(A) = \dim \mathrm{Bi}(f_A) &\underset{5.12}{=} \dim {}^n\mathbb{R} - \dim \mathrm{Ke}(f_A) \\ &= n - (n - \mathrm{Zeilenrang}(A)) = \mathrm{Zeilenrang}(A).\end{aligned}$$

□

Maple-Beispiel zu 5.14

Definition 5.15. Sei $A \in {}^m\mathbb{R}^n$. Dann heißt

$$\mathrm{rg}(A) := \mathrm{Zeilenrang}(A) \underset{5.14}{=} \mathrm{Spaltenrang}(A) = \dim \mathrm{Bi}(f_A)$$

Rang von A.

Folgerung 5.16. *Seien* $V \subset {}^n\mathbb{R}$ *und* $W \subset {}^m\mathbb{R}$ *Untervektorräume mit* $\dim V = \dim W$, *und sei* $f : V \to W$ *ein Homomorphismus.*
Dann sind äquivalent:

1. f *Monomorphismus.*
2. f *Epimorphismus.*
3. f *Isomorphismus.*

Beweis.

1. $\Rightarrow$ 2.: Da f ein Monomorphismus ist, gilt $\mathrm{Ke}(f) = 0$ nach 5.4,3. Der Dimensionssatz 5.12 liefert dann

$$\dim \mathrm{Bi}(f) \underset{5.12}{=} \dim V - \dim \mathrm{Ke}(f) = \dim V = \dim W.$$

Also ist $\mathrm{Bi}(f) \subset W$ ein Untervektorraum mit $\dim \mathrm{Bi}(f) = \dim W$. Nach 4.12,2 gilt dann $Bi(f) = W$. Somit ist f ein Epimorphismus.

2. $\Rightarrow$ 3.: Da f ein Epimorphismus ist, gilt $\mathrm{Bi}(f) = W$. Der Dimensionssatz 5.12 liefert

$$\dim W = \dim \mathrm{Bi}(f) \underset{5.12}{=} \dim V - \dim \mathrm{Ke}(f) = \dim W - \dim \mathrm{Ke}(f),$$

woraus $\dim \mathrm{Ke}(f) = 0$, d.h. $Ke(f) = 0$ folgt. Also ist f nach 5.4,3 auch ein Monomorphismus.

3. $\Rightarrow$ 1., 2.: Trivial. □

Satz 5.17 (Dimensionssatz für Untervektorräume)**.** *Seien* $U, V \subset \mathbb{R}^n$ *Untervektorräume. Dann gilt*

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V).$$

(Ebenso für Spaltenräume.)

Beweis. Auf dem Untervektorraum $U \times V \subset \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ wird die Abbildung

$$f : U \times V \to U + V \quad \text{mit} \quad f(u, v) := u + v \text{ für alle } u \in U \text{ und } v \in V$$

betrachtet.

1. f ist ein Homomorphismus, denn: $\forall\, u, u' \in U,\ v, v' \in V,\ \alpha \in \mathbb{R}$ gilt

$$\begin{aligned}f((u, v) + (u', v')) &= f(u + u', v + v') = u + u' + v + v' \\ &= (u + v) + (u' + v') = f(u, v) + f(u', v'),\\ f(\alpha(u, v)) &= f(\alpha u, \alpha v) = \alpha u + \alpha v = \alpha(u + v) = f(u, v).\end{aligned}$$

2. f ist ein Epimorphismus, denn nach Definition von $U + V$ ist f surjektiv.

3. Sei $g : U \cap V \to \text{Ke}(f)$ mit $g(u) := (u, -u)$ für alle $u \in U \cap V$. Dann ist g ein Isomorphismus, denn:
   (a) Die Abbildung ist wohldefiniert, d.h., für alle $u \in U \cap V$ gilt $g(u) \in \text{Ke}(f)$, da $f(g(u)) = f(u, -u) = u + (-u) = 0$.
   (b) g Homomorphismus: Klar.
   (c) g Monomorphismus: Klar.
   (d) g Epimorphismus: Seien $u \in U$ und $v \in V$ mit $(u, v) \in \text{Ke}(f)$. Dann gilt $0 = f(u, v) = u + v$, woraus $u = -v \in U \cap V$ und $g(u) = (u, -u) = (u, v)$ folgen.
4. Mit dem Dimensionssatz 5.12 folgt:

$$\begin{aligned}\dim(U + V) &\underset{2.}{=} \dim \text{Bi}(f) \underset{5.12}{=} \dim(U \times V) - \dim \text{Ke}(f) \\ &\underset{3.}{=} \dim(U \times V) - \dim(U \cap V).\end{aligned}$$

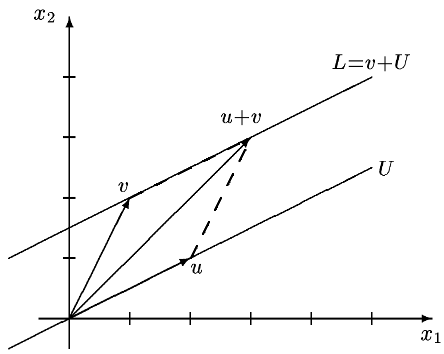
   Somit ist nur noch $\dim(U \times V) = \dim U + \dim V$ zu zeigen.

   Sei $u_1, \ldots, u_k$ eine Basis von U, und sei $v_1, \ldots, v_l$ eine Basis von V. Dann ist $(u_1, 0), \ldots, (u_k, 0), (0, v_1), \ldots, (0, v_l)$ eine Basis von $U \times V$, wie man sich leicht überlegt. Hieraus folgt die Behauptung. □

6. Affine Geometrie

Definition 6.1. Sei $L \subset {}^n\mathbb{R}$ Teilmenge. L heißt **affiner Unterraum** von ${}^n\mathbb{R}$ $:\Leftrightarrow$ $\exists\, U \subset {}^n\mathbb{R}$ Untervektorraum, $v \in {}^n\mathbb{R}$: $L = v + U := \{v + u \mid u \in U\}$ oder $L = \emptyset$.

Beispiel 6.2. Seien $v, u \in {}^2\mathbb{R}$, und sei $U = \text{span}(u) =: \mathbb{R}u \neq 0$.



Man erhält $L = v + U$ durch **Parallelverschiebung** von U (= Gerade durch den Nullpunkt).

Satz 6.3. *Für $A \in {}^m\mathbb{R}^n$ und $b \in {}^n\mathbb{R}$ sei $U := \{x \in {}^n\mathbb{R} \mid Ax = 0\}$ der Lösungsraum des homogenen Gleichungssystems $Ax = 0$ und $L := \{x \in {}^n\mathbb{R} \mid Ax = b\}$ der Lösungsraum des zugehörigen inhomogenen Gleichungssystems. Weiter sei $v \in L$ eine beliebige Lösung des inhomogenen Gleichungssystems. Dann gilt $L = v + U$.*

Beweis.

$\subset$: Sei $x \in L$, d.h. $x \in {}^n\mathbb{R}$ mit $Ax = b$. Dann gilt $A(x - v) = Ax - Av = b - b = 0$. Also ist $x - v \in U$ bzw. $x \in v + U$.

$\supset$: Sei $u \in U$. Dann gilt $A(v + u) = Av + Au = b + 0 = b$. Somit ist $v + u \in L$. □

Folgerung 6.4. *Sei $L \in {}^n\mathbb{R}$ ein affiner Unterraum.*
Dann existieren $A \in {}^m\mathbb{R}^n$ und $b \in {}^n\mathbb{R}$ mit $L = \{x \in {}^n\mathbb{R} \mid Ax = b\}$.

Beweis. Die Behauptung ist klar, falls $L = \emptyset$.

Sei also $L = v + U$ mit $v \in {}^n\mathbb{R}$ und $U \subset {}^n\mathbb{R}$ Untervektorraum. Nach Satz 5.10 gibt es eine Matrix $A \in {}^m\mathbb{R}^n$ mit $U = \{x \in {}^n\mathbb{R} \mid Ax = 0\}$. Definiert man $b := Av$, so gilt $\{x \in {}^n\mathbb{R} \mid Ax = b\} \underset{6.3}{=} v + U = L$. □

Satz 6.5. *Seien $U, U' \subset {}^n\mathbb{R}$ Untervektorräume und $v, v' \in {}^n\mathbb{R}$.*
Dann sind äquivalent:

1. $v + U = v' + U'$.
2. (a) $U = U'$.
   (b) $v - v' \in U$.

Beweis.

1. $\Rightarrow$ 2. : Sei $v + U = v' + U'$. Dann ist 2. zu zeigen.

(a) Es gilt $U = U'$, denn:

$\subset$: Sei $u \in U$. Wegen $v, u + v \in v + U = v' + U'$ existieren $u', u'_1 \in U'$ mit $v = v' + u'$ und $v + u = v' + u'_1$. Hieraus folgt

$$u = v + u - v = v' + u'_1 - (v' + u') = u'_1 - u' \in U'.$$

$\supset$: Ebenso.

(b) Wegen $v \in v + U = v' + U'$ existiert $u' \in U'$ mit $v = v' + u'$. Damit gilt $v - v' = u' \in U' \underset{(a)}{=} U$.

2. $\Rightarrow$ 1. : Sei $U = U'$ und $v + v' \in U$. Dann ist $v + U = v' + U'$ zu zeigen.

$\subset$: Sei $u \in U$. Dann gilt

$$v + u = v' + \underbrace{v - v'}_{\in U} + u \in v' + \underbrace{U'}_{=U \text{ nach } (a)}.$$

$\supset$: Ebenso. □

Also ist es jetzt sinnvoll, zu definieren:

Definition 6.6.

1. Sei $\emptyset \neq L \subset {}^n\mathbb{R}$ ein affiner Unterraum, also $L = v + U$ mit $v \in {}^n\mathbb{R}$ und $U \subset {}^n\mathbb{R}$ Untervektorraum. Dann heißt $\dim L := \dim U$ **Dimension von L**. (Im Fall $L = \emptyset$ setzt man $\dim(\emptyset) := -1$.)
   L heißt **Punkt**, bzw. **Gerade**, bzw. **Ebene**, bzw. **Hyperebene im $^n\mathbb{R}$** $:\Leftrightarrow$ $\dim L = 0$, bzw. 1, bzw. 2, bzw. $n - 1$.
2. Seien $L = v + U$ und $L' = v' + U'$ mit $v, v' \in {}^n\mathbb{R}$ und Untervektorräumen $U, U' \subset {}^n\mathbb{R}$. L heißt **parallel** zu L', oder $L \parallel L'$ $:\Leftrightarrow$ $U \subset U'$ oder $U' \subset U$.

Satz 6.7. *Seien $v_0, v_1, \ldots, v_k \in {}^n\mathbb{R}$. Dann ist*

$$v_0 + \operatorname{span}(v_1 - v_0, \ldots, v_k - v_0) = \left\{ \sum_{i=0}^{k} \alpha_i v_i \mid \forall\, 0 \leq i \leq k : \ \alpha_i \in \mathbb{R},\ \sum_{i=0}^{k} \alpha_i = 1 \right\}$$

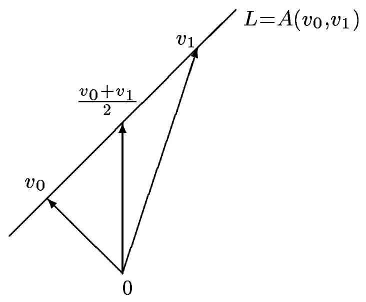
der kleinste affine Unterraum von $^n\mathbb{R}$, der $v_0, v_1, \ldots, v_k$ enthält (= der von $v_0, v_1, \ldots, v_k$ aufgespannte affine Unterraum).

Beweis.

1. Zeige: $v_0 + \operatorname{span}(v_1 - v_0, \dots, v_k - v_0) = \left\{\sum_{i=0}^k \alpha_i v_i \mid \forall\, i\ \alpha_i \in \mathbb{R},\ \sum_{i=0}^k \alpha_i = 1\right\}$.
   Beweis.
   $\subset$: Seien $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. Für
   $$x := v_0 + \sum_{i=1}^k \alpha_i (v_i - v_0) \in v_0 + \operatorname{span}(v_1 - v_0, \dots, v_k - v_0)$$
   gilt dann
   $$x = \underbrace{\left(1 - \sum_{i=1}^k \alpha_i\right)}_{=:\alpha_0} v_0 + \sum_{i=1}^k \alpha_i v_i = \sum_{i=0}^k \alpha_i v_i,$$
   wobei $\sum_{i=0}^k \alpha_i = 1$.
   $\supset$: Seien $\alpha_0, \dots, \alpha_k \in \mathbb{R}$ mit $\sum_{i=0}^k \alpha_i = 1$. Für $x = \sum_{i=0}^k \alpha_i v_i$ gilt dann
   $$x = v_0 + \sum_{i=1}^k \alpha_i (v_i - v_0),$$
   da $\alpha_0 = 1 - \sum_{i=1}^k \alpha_i$.
2. Zeige:
   (a) $v_0 + \operatorname{span}(v_1 - v_0 \dots, v_k - v_0) =: L$ ist affiner Unterraum mit $v_0, \dots, v_k \in L$.
   (b) Sei L' affiner Unterraum mit $v_0, v_1, \dots, v_k \in L'$. Dann gilt $L \subset L'$.
   Beweis.
   (a) L ist nach Definition ein affiner Unterraum, und es gilt $v_0 = v_0 + 0 \in L$ sowie $v_i = v_0 + (v_i - v_0) \in L$ für alle $1 \leq i \leq k$.
   (b) Schreibe $L' = v_0 + U$, wobei U Untervektorraum ist. Dies ist möglich, da $v_0 \in L'$. Wegen $v_1, \dots, v_k \in L'$ existieren dann $u_i \in U$ mit $v_i = v_0 + u_i$ für alle $1 \leq i \leq k$. Also gilt für alle $1 \leq i \leq k$: $v_i - v_0 = u_i \in U$, woraus $\operatorname{span}(v_1 - v_0, \dots, v_k - v_0) \subset U$ folgt, da U ein Untervektorraum ist. Somit gilt $L = v_0 + \operatorname{span}(v_1 - v_0, \dots, v_k - v_0) \subset v_0 + U = L'$. □

Definition 6.8. Seien $v_0, \dots, v_k \in {}^n\mathbb{R}$. Dann heißt
$$A(v_0, v_1, \dots, v_k) := v_0 + \operatorname{span}(v_1 - v_0, \dots, v_k - v_0)$$
affine Hülle von $v_0, v_1, \dots, v_k$.

Beispiel 6.9. Seien $v_0, v_1 \in {}^n\mathbb{R}$ mit $v_1 - v_0 \neq 0$. Dann ist
$$\begin{aligned} A(v_0, v_1) &= v_0 + \mathbb{R}(v_1 - v_0) = v_0 + \operatorname{span}(v_1 - v_0) \\ &= \{v_0 + \alpha(v_1 - v_0) \mid \alpha \in \mathbb{R}\} = \{\gamma v_0 + (1-\gamma) v_1 \mid \gamma \in \mathbb{R}\} \end{aligned}$$
die **Gerade** durch v_0, v_1, und $\{\gamma v_0 + (1-\gamma) v_1 \mid 0 \leq \gamma \leq 1\}$ ist die **Strecke** zwischen v_0 und v_1. Ferner ist $\frac{1}{2} v_0 + \frac{1}{2} v_1 = \frac{v_0 + v_1}{2}$ der **Mittelpunkt** der Strecke zwischen v_0, v_1.

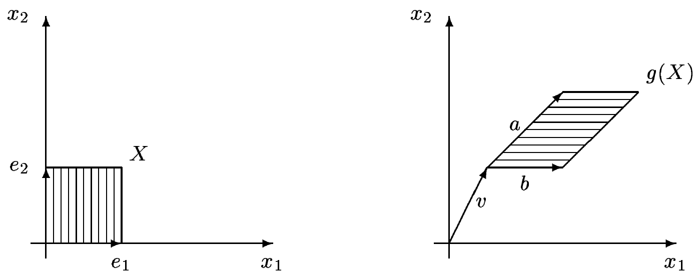


Definition 6.10.

1. Sei $v \in {}^n\mathbb{R}$. Die Abbildung $\tau_v : {}^n\mathbb{R} \to {}^n\mathbb{R}$ mit $\tau_v(x) := v + x$ für alle $x \in {}^n\mathbb{R}$ heißt **Translation** um v.
2. Sei $g : {}^n\mathbb{R} \to {}^n\mathbb{R}$ eine Abbildung.
   g heißt **Affinität** $:\Leftrightarrow \ \exists\, v \in V,\ \exists\, f : {}^n\mathbb{R} \to {}^n\mathbb{R}$ Isomorphismus mit $g = \tau_v \circ f$, d.h. $\forall\, x \in {}^n\mathbb{R} :\ g(x) = v + f(x)$.

Bemerkungen 6.11.

1. Nach Definition sind affine Unterräume von der Form $\tau_v(U) = v + U$, wobei τ_v eine Translation und ein U Untervektorraum ist.
2. In der affinen Geometrie studiert man die Eigenschaften des ${}^n\mathbb{R}$, die invariant sind unter Affinitäten, d.h., man identifiziert für eine "Figur" $X \subset {}^n\mathbb{R}$ die Menge $g(X)$ mit X, falls g eine Affinität ist, z.B.



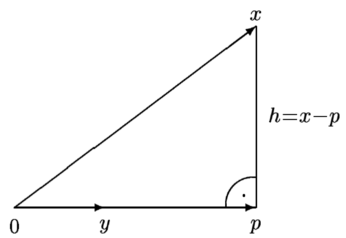
wobei $g = \tau_v f,\ a = f(e_1),\ b = f(e_2)$.

Kapitel II: Das Standard-Skalarprodukt im $\mathbb{R}^n$

1. Standard-Skalarprodukt, Länge, Winkel

Definition 1.1. Seien $x = (x_1, \dots, x_n),\ y = (y_1, \dots, y_n) \in \mathbb{R}^n$.

1. $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$ heißt **(Standard-)Skalarprodukt** von x und y. (Ebenso für Spalten.)
2. $\|x\| := \sqrt{\langle x, x\rangle}$ heißt **Länge** von x.
   $d(x-y) := \|x-y\|$ heißt **Abstand** von x und y.
3. Sei $y \neq 0$. $p \in \mathbb{R}^n$ heißt **orthogonale Projektion von x auf $\mathbb{R}$y** $:\Leftrightarrow$
   (a) $p \in \mathbb{R}y$.
   (b) Für $h := x - p$ gilt: $\langle h, y\rangle = 0$.

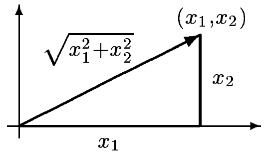


   h heißt **Lot** von x auf $\mathbb{R}y$.
4. x heißt **senkrecht** auf y, geschrieben $x \perp y \ :\Leftrightarrow\ \langle x, y\rangle = 0$.

Bemerkungen 1.2.

1. $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto \langle x, y\rangle$ hat folgende Eigenschaften:
   (a) $\langle , \rangle$ ist **bilinear**, d.h. $\forall\, x, y, z \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$:
$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle, \quad \langle \alpha x, z\rangle = \alpha\langle x, z\rangle,$$
$$\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle, \quad \langle x, \alpha y\rangle = \alpha\langle x, y\rangle.$$
   (b) $\langle , \rangle$ ist **symmetrisch**, d.h. $\forall\, x, y \in \mathbb{R} :\ \langle x, y\rangle = \langle y, x\rangle$.
   (c) $\langle , \rangle$ ist **positiv definit**, d.h. $\forall\, 0 \neq x \in \mathbb{R}^n :\ \langle x, x\rangle\ > 0$.

   Beweis.
   (a), (b) Klar.
   (c) Sei $x = (x_1, \dots, x_n) \neq 0$. $\exists\, 1 \leq j \leq n :\ x_j \neq 0$.
   $\Rightarrow \langle x, x\rangle = \sum_{i=1}^n x_i^2 \geq x_j^2 > 0$.
2. Im folgenden werden vom Standardskalarprodukt meist **nur** die Eigenschaften in 1., d.h. **bilinear, symmetrisch, positiv definit** verwendet!
3. Im $\mathbb{R}^2$ gilt: $\|x\|$ ist die übliche Länge wegen des Satzes von Pythagoras.



Satz 1.3. $x, y \in \mathbb{R}^n,\ y \neq 0.$
$\Rightarrow$ *Es gibt eine eindeutig bestimmte orthogonale Projektion $p \in \mathbb{R}^n$ von x auf $\mathbb{R}y$, und für diese gilt:*

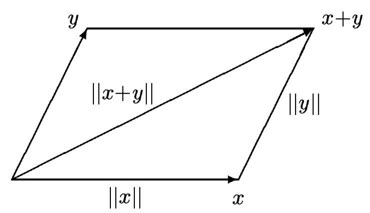
$$p = \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y$$

Beweis.

1. Eindeutigkeit: Sei $p = \alpha y \in \mathbb{R}y$, $\alpha \in \mathbb{R}$ mit $x - p \perp y$, d.h. $\langle x - p, y\rangle = 0$.
   $\Rightarrow \langle x, y\rangle = \langle p, y\rangle = \langle \alpha y, y\rangle = \alpha\langle y, y\rangle$
   $\underset{y\neq 0}{\Rightarrow} \alpha = \frac{\langle x,y\rangle}{\langle y,y\rangle}$
2. Existenz: Definiere $p = \frac{\langle x,y\rangle}{\langle y,y\rangle} \cdot y$. Dies ist möglich, da $y \neq 0$.
   $\Rightarrow \langle p, y\rangle = \left\langle \frac{\langle x,y\rangle}{\langle y,y\rangle} \cdot y, y \right\rangle = \frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y, y\rangle = \langle x, y\rangle$
   $\Rightarrow \langle p - x, y\rangle = \langle p, y\rangle - \langle x, y\rangle = 0$ $\surd$ □

Satz 1.4 (Eigenschaften des Skalarprodukts und der Länge)**.** $x, y \in \mathbb{R}^n,\ \alpha \in \mathbb{R}\ \Rightarrow$

1. $\|x\| = 0 \Leftrightarrow x = 0.$
   $\|x\| \geq 0.$
2. $\|\alpha x\| = |\alpha|\ \|x\|.$
3. **Schwarsche Ungleichung**: $|\langle x, y\rangle| \leq \|x\|\ \|y\|$, *wobei = genau dann gilt, wenn x, y linear abhängig sind.*
4. **Dreiecksungleichung**: $\|x + y\| \leq \|x\| + \|y\|.$



Beweis.

1. Klar wegen Bemerkung 1.2,1, da $\|x\| = \sqrt{\langle x, x\rangle}$.
2. Für $x \in \mathbb{R}^n$ gilt
$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha^2\langle x, x\rangle} = |\alpha|\sqrt{\langle x, x\rangle} = |\alpha|\ \|x\|.$$
3. Sei nach Satz 1.3 p die orthogonale Projektion von x auf $\mathbb{R}y$, wobei $y \neq 0$ sei (Behauptung klar, falls $y = 0$), also: $p = \frac{\langle x,y\rangle}{\langle y,y\rangle}y,\ h := x - p$ Lot. Dann gilt
$$\begin{aligned} 0 \leq \langle h, h\rangle &= \langle x - p, x - p\rangle = \langle x, x\rangle + \langle p, p\rangle - \langle x, p\rangle - \langle p, x\rangle \\ &= \langle x, x\rangle + \langle p, p\rangle - 2\langle x, p\rangle = \langle x, x\rangle + \frac{\langle x, y\rangle^2}{\langle y, y\rangle^2}\langle y, y\rangle - 2\frac{\langle x, y\rangle}{\langle y, y\rangle}\langle x, y\rangle \\ &= \langle x, x\rangle + \frac{\langle x, y\rangle^2}{\langle y, y\rangle} - 2\frac{\langle x, y\rangle^2}{\langle y, y\rangle} = \langle x, x\rangle - \frac{\langle x, y\rangle^2}{\langle y, y\rangle} \end{aligned}$$
   Also: $\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$. Durch Wurzelziehen folgt: $|\langle x, y\rangle| \leq \|x\|\ \|y\|$.
   Außerdem gilt = genau dann, wenn x, y linear abhängig sind, denn:

(a) Sei $|\langle x, y\rangle| = \|x\|\,\|y\|$
$\underset{\text{obiger Beweis}}{\Rightarrow} \|h^2\| = 0$, d.h. $h = 0$
$\Rightarrow x = p = \frac{\langle x,y\rangle}{\langle y,y\rangle} y \;\checkmark$

(b) Seien x, y linear abhängig.
$\Rightarrow\ \exists\, \alpha, \beta \in \mathbb{R} : \alpha x + \beta y = 0$ und $\alpha \neq 0$ oder $\beta \neq 0$.
Sei $\alpha \neq 0$ (ebenso für $\beta \neq 0$).
$\Rightarrow x = \underbrace{-\frac{\beta}{\alpha}}_{=:\gamma} y$
Also: $|\langle x, y\rangle| = |\gamma| \langle y, y\rangle = |\gamma|\,\|y\|^2 = \|\gamma y\|\,\|y\| = \|x\|\,\|y\|$.

4. Es gilt

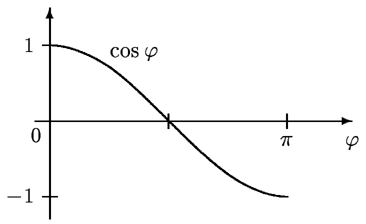
$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle \\ &= \|x\| + 2\langle x,y\rangle + \|y\| \leq \|x\| + 2|\langle x,y\rangle| + \|y\| \\ &\overset{3.}{\leq} \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\end{aligned}$$

Also: $\|x+y\|^2 \leq (\|x\| + \|y\|)^2$, d.h. $\|x+y\| \leq \|x\| + \|y\|$. □

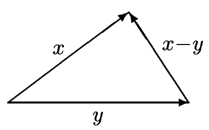
Definition 1.5. Seien $0 \neq x, y \in \mathbb{R}^n$ und $0 \leq \varphi \leq \pi$.
φ heißt **Winkel zwischen $\mathbf{x}, \mathbf{y}$**, geschrieben $\varphi = \angle(x, y)\ :\Leftrightarrow\ \cos\varphi = \frac{\langle x,y\rangle}{\|x\|\,\|y\|}$.

Bemerkungen 1.6.

1. $0 \neq x, y \in \mathbb{R}^n\ \Rightarrow$ Es gibt ein eindeutig bestimmtes φ mit $\varphi = \angle(x, y)$.
Beweis. Nach der Schwarzscher Ungleichung 1.4,3 gilt: $-1 \leq \frac{\langle x,y\rangle}{\|x\|\,\|y\|} \leq 1$.
$\Rightarrow\ \exists$ eindeutig bestimmtes φ mit $0 \leq \varphi \leq \pi$ und $\cos\varphi = \frac{\langle x,y\rangle}{\|x\|\,\|y\|}$, denn:



2. **Cosinus-Satz**:
$0 \neq x, y \in \mathbb{R}^n\ \Rightarrow\ \|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\cos(\varphi)\|x\|\,\|y\|$, wobei $\varphi = \angle(x, y)$.



(Also ist $\cos\varphi$ im $\mathbb{R}^2$ wie üblich.)
Beweis. $\|x-y\|^2 = \langle x-y, x-y\rangle = \langle x,x\rangle + \langle y,y\rangle - 2\underbrace{\langle x,y\rangle}_{=\cos(\varphi)\|x\|\,\|y\|} \;\checkmark$

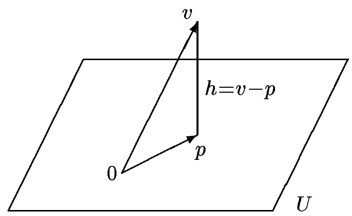
3. $0 \neq x, y \in \mathbb{R}^n$, $\varphi = \angle(x, y)$. Dann: $x \perp y\ \Leftrightarrow\ \varphi = \frac{\pi}{2}$.
Beweis. Es gilt:

$$\begin{array}{ll} & x \perp y \\ \Leftrightarrow & \langle x, y\rangle = 0 \\ \Leftrightarrow & \cos\varphi = \frac{\langle x,y\rangle}{\|x\|\,\|y\|} = 0 \\ \Leftrightarrow & \varphi = \frac{\pi}{2} \end{array}$$

4. $0 \neq x \in \mathbb{R}^n \;\Rightarrow \frac{x}{\|x\|}$ hat Länge 1 (Normierung).
Beweis. Es gilt:
$$\left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|}\|x\| = 1 \;\checkmark$$

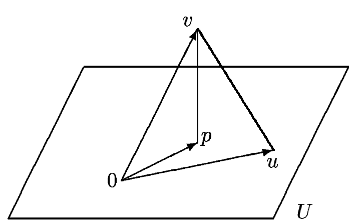
5. **Satz von Pythagoras**: Seien $v_1, \ldots, v_k \in \mathbb{R}^n$ mit $v_i \perp v_j \;\forall\, 1 \leq i, j \leq k,\; i \neq j$.
$\Rightarrow \|v_1 + \ldots + v_k\|^2 = \|v_1\|^2 + \ldots + \|v_k\|^2$.
Beweis. Es gilt:
$$\langle v_1 + \ldots + v_k, v_1 + \ldots + v_k\rangle = \sum_{1 \leq i,j \leq k} \underbrace{\langle v_i, v_j\rangle}_{=0,\text{ falls } i\neq j} = \sum_{i=1}^{k} \langle v_i, v_i\rangle = \sum_{i=1}^{k} \|v_i\|^2$$

Definition 1.7. $U \subset \mathbb{R}^n$ Untervektorraum, $v \in \mathbb{R}^n$.
$p \in \mathbb{R}^n$ heißt **orthogonale Projektion von v auf U** $:\Leftrightarrow\; p \in U,\; h := v - p \perp U$,
d.h. $v - p \perp u \;\forall\, u \in U$.
Dann heißt h das **Lot** von v auf U.



Satz 1.8 (Orthogonale Projektion)**.** *$U \subset \mathbb{R}^n$ Untervektorraum, $v \in \mathbb{R}^n$.*
$\Rightarrow$ Es gibt genau eine orthogonale Projektion von v auf U, und für ein Erzeugendensystem $u_1, \ldots, u_k$ von U und $x_1, \ldots, x_k \in \mathbb{R}$ sind für den Vektor $p = \sum_{i=1}^{k} x_i u_i$ äquivalent:

1. *$v - p \perp U$, d.h. p ist* **orthogonale Projektion** *von v auf U.*
2. $\underbrace{(\langle u_i, u_j\rangle)}_{\in\, {}^k\mathbb{R}^k} \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} \langle u_1, v\rangle \\ \vdots \\ \langle u_k, v\rangle \end{pmatrix}$ (**Normalgleichungen**)
3. $\|p - v\| = \min\{\|v - u\| \mid u \in U\}$, *d.h. p ist die* **beste Approximation von v in U**.



Folgerung 1.9 (Methode der kleinsten Quadrate)**.** $A \in {}^m\mathbb{R}^n,\ b \in {}^m\mathbb{R}$.
Dann sind für

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

folgende Aussagen äquivalent:

1. $\|Ax - b\| = \min\{\|Ay - b\| \mid y \in {}^n\mathbb{R}\}$
2. $A^T A x = A^T b$ (**Normalgleichung**)

Außerdem ist das Gleichungssystem in 2. immer lösbar.

Beweis. Wende Satz 1.8 an mit

- $U =$ Spaltenraum von A, also $U \subset {}^m\mathbb{R}$,
- $v = b$.

Seien $a_1, \dots, a_n$ die Spalten von A, also Erzeugende von U. Wegen $\langle a_i, a_j\rangle = a_i^T a_j$ und $\langle a_i, b\rangle = a_i^T b$ ist

$$(\langle a_i, a_j\rangle) = A^T A \quad \text{bzw.} \quad \begin{pmatrix} \langle a_1, b\rangle \\ \vdots \\ \langle a_n, b\rangle \end{pmatrix} = A^T b.$$

Mit

$$p = \sum_{i=1}^{n} x_i a_i = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \text{orthogonale Projektion von } b \text{ auf } U$$

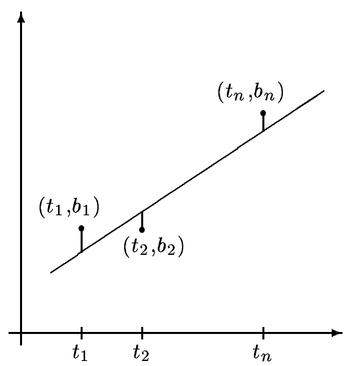
folgt die Behauptung aus Satz 1.8. □

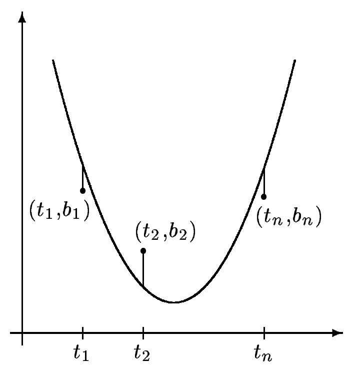
Definition 1.10. Für $\alpha = (\alpha_0, \dots, \alpha_n) \in \mathbb{R}^{n+1}$ sei $P_\alpha : \mathbb{R} \to \mathbb{R}$ definiert durch:

$$P_\alpha(t) = \sum_{i=0}^{n} \alpha_i t^i = (1\ t \dots\ t^n) \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Problem: Bestimme $P_\alpha(t)$ durch Messungen:

- Meßpunkte $t_1, \dots, t_m$;
- beobachtete Ergebnisse $b_1, \dots, b_m$ mit Meßfehlern.





Folgerung 1.11. $m, n \in \mathbb{N},\ t_1, \dots, t_m,\ b_1, \dots, b_m \in \mathbb{R}$.
Dann sind für $\alpha = (\alpha_0, \dots, \alpha_n) \in \mathbb{R}^{n+1}$ mit

$$A := \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^n \end{pmatrix} \qquad \textit{und} \qquad b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

äquivalent:

1. $\sum_{i=1}^m (P_\alpha(t_i) - b_i)^2 = \min\left\{\sum_{i=1}^m (P_\beta(t_i) - b_i)^2 \mid \beta \in \mathbb{R}^{n+1}\right\}$
2. $A^T A \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = A^T b \quad$ (**Normalgleichung**)

Außerdem ist das Gleichungssystem in 2. stets lösbar.

Beweis. Für $\alpha = (\alpha_0, \dots, \alpha_n)$ gilt

$$\begin{pmatrix} P_\alpha(t_1) \\ \vdots \\ P_\alpha(t_m) \end{pmatrix} = A \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}, \text{ da } \quad (1\ t_i \dots\ t_i^n) \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = P_\alpha(t_i).$$

Also folgt 1.11 aus 1.9 mit

$$x = \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

□

Maple-Demonstration mit Polynomfunktion vom Grad 3.

Zusammenhang mit orthogonaler Projektion:
Seien die Bezeichnungen wie in 1.11, und sei

$$P_\alpha(t) = \sum_{i=0}^n \alpha_i t^i = (1\ t \dots t^n) \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}$$

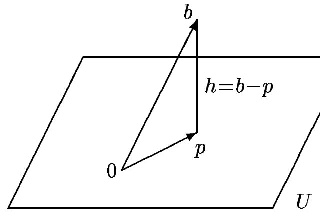
eine Lösung von

$$\sum_{i=1}^m (P_\alpha(t_i) - b_i)^2 = \min\left\{\sum_{i=1}^m (P_\beta(t_i) - b_i)^2 \mid \beta \in \mathbb{R}^{n+1}\right\}.$$

Ferner sei U der Spaltenraum von A. Dann ist

$$p = A \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} P_\alpha(t_1) \\ \vdots \\ P_\alpha(t_m) \end{pmatrix}$$

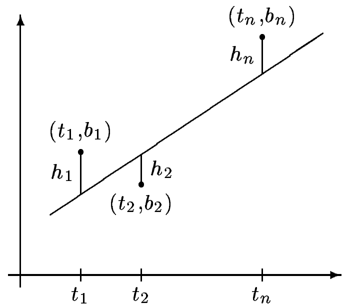
die orthogonale Projektion von b auf U:



Das Lot

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = b - p$$

enthält die Abweichungen der Polynomfunktion $P_\alpha(t)$ von den vorgegebenen Werten b_i an den Stellen t_i:



Bweis von Satz 1.8. Zunächst wird gezeigt:

1. $\Leftrightarrow$ 2.: Es gilt

$$\begin{aligned} v - p \perp U \quad & \underset{\text{Erzeugendensystem}}{\overset{u_1,\dots,u_k}{\Leftrightarrow}} \quad \forall\, 1 \le i \le k : \langle v - p, u_i \rangle = 0 \\ & \Leftrightarrow \quad \forall\, 1 \le i \le k : \langle u_i, v \rangle = \langle u_i, p \rangle \\ & \Leftrightarrow \quad \forall\, 1 \le i \le k : \underbrace{\langle u_i, \underbrace{\sum_{j=1}^{k} x_j u_j}_{=p} \rangle}_{=\sum_{j=1}^{k} \langle u_i, u_j \rangle x_j} = \langle u_i, v \rangle \\ & \Leftrightarrow \quad (\langle u_i, u_j \rangle) \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} \langle u_i, v \rangle \\ \vdots \\ \langle u_k, v \rangle \end{pmatrix} \end{aligned}$$

Mit dieser Äquivalenz kann jetzt die Existenz und Eindeutigkeit der orthogonalen Projektion bewiesen werden. Sei dazu $u_1, \dots, u_k$ eine Basis von U.
$\underset{\text{Lemma 1.14}}{\Rightarrow}$ $(\langle u_i, u_j \rangle)_{1 \le i,j \le k}$ invertierbar.

$\Rightarrow$ Gleichungssystem in 2. hat genau eine Lösung $\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$.

Also folgt wegen 1. $\Leftrightarrow$ 2 : $\exists$ genau eine orthogonale Projektion von v auf U.

Schließlich ist noch die Äquivalenz der Aussagen 1. und 3. zu zeigen.

1. $\Rightarrow$ 3.: Zeige: Für die orthogonale Projektion p gilt: $\forall\, u \in U: \ \|v-p\| \leq \|v-u\|$.

   Beweis. Sei $u \in U$.
   Betrachte: $v-u = \underbrace{v-p}_{\in U^\perp} + \underbrace{p-u}_{\in U}$.
   $\underset{\text{Pythagoras}}{\Rightarrow} \ \|v-u\|^2 = \|v-p\|^2 + \|p-u\|^2 \geq \|v-p\|^2 \ \surd$

3. $\Rightarrow$ 1.: Es gelte $\|v-p\| = \min\{\|v-u\| \mid u \in U\}$.
   Sei q die orthogonale Projektion von v auf U (existiert, wie eben bewiesen).
   Wegen 1. $\Rightarrow$ 3. gilt dann $\|v-q\| = \min\{\|v-u\| \mid u \in U\}$, woraus $\|v-p\| = \|v-q\|$ folgt.
   Betrachte: $v-p = \underbrace{v-q}_{\in U^\perp} + \underbrace{q-p}_{\in U}$.
   $\underset{\text{Pythagoras}}{\Rightarrow} \ \|v-p\|^2 = \|v-q\|^2 + \|q-p\|^2$
   $\underset{\|v-p\|=\|v-q\|}{\Rightarrow} \ \|q-p\|^2 = 0$, d.h. $p = q \ \surd$ $\square$

Definition 1.12. Sei $A \in {}^n\mathbb{R}^n$.

1. A heißt **positiv definit** $:\Leftrightarrow \ \forall\, 0 \neq x \in {}^n\mathbb{R}: \ x^T A x > 0$.
2. A heißt **symmetrisch** $:\Leftrightarrow \ A^T = A$.

Bemerkungen 1.13.

1. Sei $A = (a_{ij}) \in {}^n\mathbb{R}^n$. Für

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in {}^n\mathbb{R}$$

   gilt dann

$$x^T A x = (x_1 \dots x_n) \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j \end{pmatrix} = \sum_{i,j=1}^n x_i a_{ij} x_j =: F(x_1, \dots, x_n).$$

   Also: A ist positiv definit. $\Leftrightarrow$ Für die quadratische Funktion F gilt:

$$\forall\, x_1, \dots, x_n \in \mathbb{R} : F = (x_1, \dots, x_n) \geq 0, \text{ und}$$
$$F(x_1, \dots, x_n) = 0 \Leftrightarrow x_1 = 0, \dots, x_n = 0.$$

2. A positiv definit $\Rightarrow$ A invertierbar.
   Beweis. Wegen Kapitel I,3.14 genügt zu zeigen: $\forall\, x \in {}^n\mathbb{R} : Ax = 0 \Rightarrow x = 0$.
   Dies ist aber erfüllt, denn für $x \in {}^n\mathbb{R}$ gilt:

$$Ax = 0 \ \Rightarrow \ x^T A x = 0 \ \underset{A \text{ positiv definit}}{\Longrightarrow} \ x = 0.$$

Lemma 1.14. *Seien $u_1, \dots, u_k \in \mathbb{R}^n$.*
Dann sind äquivalent:

1. *$u_1, \dots, u_k$ sind linear unabhängig.*
2. *$(\langle u_i, u_j \rangle)_{1 \leq i,j \leq k}$ ist positiv definit.*
3. *$(\langle u_i, u_j \rangle)_{1 \leq i,j \leq k}$ invertierbar.*

Beweis.

1. $\Rightarrow$ 2. : Sei

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \in {}^k\mathbb{R}.$$

Mit $v := \sum_{i=1}^k x_i u_i$ gilt dann

$$x^T(\langle u_i, u_j\rangle)x = \sum_{i,j} x_i \langle u_i, u_j\rangle x_j = \left\langle \sum_{i=1}^n x_i u_i, \sum_{j=1}^n x_j u_j \right\rangle = \|v\|^2 \geq 0$$

sowie

$$x^T(\langle u_i, u_j\rangle)x = 0 \;\Rightarrow\; v = 0 \underset{\substack{u_1,\ldots,u_k \\ \text{linear unabhängig}}}{\Rightarrow} x_1 = 0, \ldots, x_k = 0.$$

2. $\Rightarrow$ 3. : Wegen Bemerkung 1.13,2.
3. $\Rightarrow$ 1. : Seien $x_1, \ldots, x_k \in \mathbb{R}$ mit $\sum_{j=1}^k x_j u_j = 0$. Dann gilt

$$\forall\, 1 \leq i \leq k : \underbrace{\left\langle u_i, \underbrace{\sum_{j=1}^k x_j u_j}_{=0} \right\rangle}_{=\sum_{j=1}^k \langle u_i, u_j\rangle x_j} = 0 \underset{\substack{(\langle u_i,u_j\rangle) \\ \text{invertierbar}}}{\Rightarrow} \forall\, 1 \leq j \leq k :\; x_j = 0$$

□

Bemerkungen 1.15.

1. Sei $A \in {}^m\mathbb{R}^n$ mit linear unabhängigen Spalten. Dann ist die Lösung x der Normalgleichungen $A^TAx = A^Tb$ aus 1.9 **eindeutig bestimmt**, und es gilt: $x = (A^TA)^{-1}A^Tb$.
   Beweis. Seien $a_1, \ldots, a_n$ die Spalten von A. $\underset{1.14}{\Rightarrow}$ $(\langle a_i, a_j\rangle) = A^TA$ invertierbar.
2. Seien $m, n \geq 1$ mit $m \geq n+1$ und $t_1, \ldots, t_m \in \mathbb{R}$ mit $t_i \neq t_j \;\forall\, i \neq j$. Dann sind die Spalten der Matrix

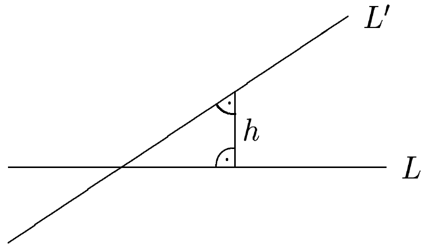
$$A := \begin{pmatrix} 1 & t_1 & t_1^2 & \ldots & t_1^n \\ 1 & t_2 & t_2^2 & \ldots & t_2^n \\ \vdots & & & & \vdots \\ 1 & t_m & t_m^2 & \ldots & t_m^n \end{pmatrix}$$

   linear unabhängig.
   Beweis. Folgt aus Übung.
3. Wegen 1. und 2. ist also die Lösung $\alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{R}^{n+1}$ der Normalgleichung in 1.11, d.h. $\alpha \in \mathbb{R}^{n+1}$ mit

$$\sum_{i=1}^m (P_\alpha(t_i) - b_i)^2 = \min\left\{ \sum_{i=1}^m (P_\beta(t_i) - b_i)^2 \mid \beta \in \mathbb{R}^{n+1} \right\}$$

   eindeutig bestimmt, falls $m \geq n+1$ gilt.

Folgerung 1.16 (Abstand affiner Unterräume). *Seien $L = v + U,\ L' = v' + U' \subset \mathbb{R}^n$ affine Unterräume, $v, v' \in \mathbb{R}^n$, $U, U' \subset \mathbb{R}^n$ Untervektorräume. Sei h das Lot von $v - v'$ auf $U + U'$. Dann ist $\|h\| = \min\{\|x - x'\| \mid x \in L,\ x' \in L'\}$.*



Beweis.

1. Zeige: $\exists\, a \in L,\ a' \in L' :\ h = a - a'$.
   Beweis. Schreibe
$$\begin{aligned} v - v' &= \underbrace{p}_{U+U'} + \underbrace{h}_{(U+U')^\perp} \\ \Rightarrow\quad &\exists\, u \in U,\ u' \in U' : p = u + u' \\ \Rightarrow\quad & h = v - v' - u - u' = \underbrace{(v-u)}_{=:a\in L} - \underbrace{(v'+u')}_{=:a'\in L'} \end{aligned}$$

2. Zeige noch: $\forall\, x \in L,\ x' \in L' :\ \|h\| \le \|x - x'\|$.
   Beweis. Seien $x \in L,\ x' \in L'$. Dann existieren $\tilde{u} \in U$ und $\tilde{u}' \in U'$ mit $x = v + \tilde{u}$ bzw. $x' = v' + \tilde{u}'$. Ist p die orthogonale Projektion von $v - v'$ auf $U + U'$, d.h. $p \in U + U'$ mit $v - v' = p + h$, so folgt
$$\begin{aligned} x - x' &= v + \tilde{u} - (v' + \tilde{u}') = v - v' + \tilde{u} - \tilde{u}' \\ &= p + h + \tilde{u} - \tilde{u}' = \underbrace{p + \tilde{u} - \tilde{u}'}_{\in U+U'} + \underbrace{h}_{\in(U+U')^\perp} \end{aligned}$$
$$\underset{\text{Pythagoras}}{\Rightarrow} \|x - x'\|^2 = \|p + \tilde{u} - \tilde{u}'\|^2 + \|h\|^2 \ge \|h\|^2\ \checkmark$$

□

2. Orthonormalbasis, orthogonale Matrizen und Isometrie

Definition 2.1. Sei $g : {}^n\mathbb{R} \to {}^n\mathbb{R}$ eine Abbildung.
g heißt **Bewegung** oder **Isometrie** $:\Leftrightarrow\ \forall\, x, y \in {}^n\mathbb{R} : \|g(x) - g(y)\| = \|x - y\|$.

Definition 2.2. Sei $A \in {}^n\mathbb{R}^n$.
A heißt **orthogonal** $:\Leftrightarrow$ A ist invertierbar, und $A^{-1} = A^T$.

Bemerkung 2.3. Sei $A \in {}^n\mathbb{R}^n$ mit Zeilen $z_1, \ldots, z_n \in \mathbb{R}^n$ und Spalten $a_1, \ldots, a_n \in {}^n\mathbb{R}$. Dann sind äquivalent:

1. A ist orthogonal.
2. $\forall\, 1 \le i, j \le n : \langle a_i, a_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{sonst} \end{cases}$
3. $A^T A = E$ (Einheitsmatrix).
4. $\forall\, 1 \le i, j \le n :\ \langle z_i, z_j \rangle = \delta_{ij}$.
5. $AA^T = E$.

Beweis.

1\. $\Leftrightarrow$ 3. $\Leftrightarrow$ 5. : Nach Definition und Kapitel I.
2\. $\Leftrightarrow$ 3. : Da $\langle a_i, a_j\rangle = a_i^T a_j$ ist, gilt: $(\langle a_i, a_j\rangle) = (\delta_{ij}) \Leftrightarrow A^T A = E$.
4\. $\Leftrightarrow$ 5. : Wie 2. $\Leftrightarrow$ 3. □

Definition 2.4. Sei $V \subset \mathbb{R}^n$ (oder $\subset {}^n\mathbb{R}$) Untervektorraum, und seien $v_1, \ldots, v_k \in V$.

1. $v_1, \ldots, v_k$ heißt **Orthogonalsystem** (OGS)
   $:\Leftrightarrow\ \forall\, 1 \leq i, j \leq k,\ i \neq j\ : \langle v_i, v_j\rangle = 0$.
2. $v_1, \ldots, v_k$ heißt **Orthonormalsystem** (ONS)
   $:\Leftrightarrow\ \forall\, 1 \leq i, j \leq k,\ i \neq j\ : \langle v_i, v_j\rangle = \delta_{ij}$.
3. $v_1, \ldots, v_k$ heißt **Orthonormalbasis** (ONB) von V
   $:\Leftrightarrow\ v_1, \ldots, v_k$ ist Basis von V und ONS.

Bemerkungen 2.5.

1. Die Standardbasis $e_1, \ldots, e_n$ des $\mathbb{R}^n$ ist eine ONB.
2. Sei $v_1, \ldots, v_k$ OGS mit $v_i \neq 0$ für alle $1 \leq i \leq k$. Dann sind $v_1, \ldots, v_k$ linear unabhängig.
   Beweis. Seien $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ mit $\sum_{i=1}^k \alpha_i v_i = 0$. Dann gilt für alle $1 \leq j \leq k$:
   $$0 = \langle \underbrace{\sum_{i=1}^k \alpha_i v_i}_{=0}, v_j\rangle = \sum_{i=1}^k \alpha_i \langle \underbrace{v_i, v_j}_{\substack{=0,\\ \text{falls } i\neq j}} \rangle = \alpha_j \langle v_j, v_j\rangle.$$
   Wegen $v_j \neq 0$ und damit $\langle v_j, v_j\rangle \neq 0$ folgt $\alpha_j = 0$.
3. Sei $V \subset \mathbb{R}^n$ Untervektorraum mit $\dim V = k$, und sei $v_1, \ldots, v_k \in V$ ein ONS. Dann ist $v_1, \ldots, v_k$ ONB von V.
   Beweis. Nach 2. sind $v_1, \ldots, v_k$ linear unabhängig. Wegen $\dim V = k$ folgt dann die Behauptung aus Kapitel I,4.12.
4. Sei $A \in {}^n\mathbb{R}^n$. Dann sind wegen Bemerkung 2.3 und 3. äquivalent:
   (a) A orthogonal.
   (b) Die Zeilen von A bilden eine ONB von $\mathbb{R}^n$.
   (c) Die Spalten von A bilden eine ONB von ${}^n\mathbb{R}$.

Satz 2.6. *Sei $g : {}^n\mathbb{R} \to {}^n\mathbb{R}$ eine Abbildung. Dann sind äquivalent:*

1. *g ist Bewegung.*
2. *$\exists\, v \in {}^n\mathbb{R},\ A \in {}^n\mathbb{R}^n$ orthogonal: $g = \tau_v f_A$, d.h. $\forall\, x \in {}^n\mathbb{R} :\ g(x) = v + Ax$.*

Beweis.

2\. $\Rightarrow$ 1. : Für alle $x, y \in {}^n\mathbb{R}$ gilt:

$$\begin{aligned} \| \underbrace{g(x)}_{v+Ax} - \underbrace{g(y)}_{v+Ay} \|^2 &= \| \underbrace{Ax - Ay}_{A(x-y)} \|^2 = \langle A(x-y), A(x-y)\rangle \\ &= \underbrace{(A(x-y))^T}_{=(x-y)^T A^T} A(x-y) \underset{A^T A = E}{=} (x-y)^T (x-y) \\ &= \langle x-y, x-y\rangle = \|x-y\|^2. \end{aligned}$$

1. $\Rightarrow$ 2. : Setze $v := g(0)$, $f := \tau_{-v}g$. Dann ist f als Verkettung zweier Bewegungen selbst eine Bewegung, und es gilt $f(0) = 0$, da $f(0) = -v + g(0) = -v + v = 0$. Ferner ist $g = \tau_v f$, da für alle $x \in V$ gilt:

$$(\tau_v f)(x) = v + f(x) = v + (-v) + g(x) = g(x).$$

Damit ist noch zu zeigen: $\exists\, A \in {}^n\mathbb{R}^n$ orthogonal: $f = f_A$.

(a) $\forall\, x \in {}^n\mathbb{R} : \|f(x)| = \|x\|$, denn

$$\|f(x)\| = \|f(x) - \underbrace{f(0)}_{=0}\| \underset{\text{Bewegung}}{=} \|x - 0\| = \|x\|.$$

(b) $\forall\, x, y \in {}^n\mathbb{R} :\ \langle f(x), f(y)\rangle = \langle x, y\rangle$.

Beweis. Es gilt

$$\|x - y\|^2 = \langle x - y, x - y\rangle = \underbrace{\langle x, x\rangle}_{=\|x\|^2} + \underbrace{\langle y, y\rangle}_{=\|y\|^2} - 2\langle x, y\rangle.$$

Ebenso ergibt sich $\|f(x) - f(y)\|^2 = \|f(x)\|^2 + \|f(y)\|^2 - 2\langle f(x), f(y)\rangle$. Außerdem gilt: $\|f(x) - f(y)\| = \|x - y\|$, da f eine Bewegung ist. Mit (a) folgt dann $\langle f(x), f(y)\rangle = \langle x, y\rangle$.

(c) $f(e_1), \ldots, f(e_n)$ ist ONB des ${}^n\mathbb{R}$, wobei $e_1, \ldots, e_n$ die Standardbasis ist.

Beweis. $f(e_1) \ldots, f(e_n)$ ist ONS, da wegen (b) für alle $1 \leq i, j \leq n$ gilt: $\langle f(e_i), f(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}$. Wegen $\dim {}^n\mathbb{R} = n$ folgt die Behauptung dann aus Bemerkung 2.5,3.

(d) $\forall\, x_1, \ldots, x_n \in \mathbb{R}$ und

$$x := \sum_{i=1}^{n} x_i e_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

gilt $f(x) = \sum_{i=1}^n x_i f(e_i)$.

Beweis. Da nach (c) $f(e_1), \ldots, f(e_n)$ eine ONB des ${}^n\mathbb{R}$ ist, existieren $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $f(x) = \sum_{i=1}^n \alpha_i f(e_i)$. Damit gilt für alle $1 \leq j \leq n$:

$$\langle f(x), f(e_j)\rangle = \langle \sum_{i=1}^{n} \alpha_i f(e_i), f(e_j)\rangle = \sum_{i=1}^{n} \alpha_i \underbrace{\langle f(e_i), f(e_j)\rangle}_{\delta_{ij} \text{ wegen } (c)} = \alpha_j.$$

Außerdem ist $\langle f(x), f(e_j)\rangle \underset{(b)}{=} \langle x, e_j\rangle = \langle \sum_{i=1}^n x_i e_i, e_j\rangle \underset{\text{ebenso}}{=} x_j$.

Also folgt $\alpha_j = x_j$ für alle $1 \leq j \leq n$.

(e) $A := (f(e_1), \ldots, f(e_n)) \in {}^n\mathbb{R}^n$ ist orthogonal, und es gilt: $f = f_A$.

Beweis. A ist orthogonal, da $f(e_1), \ldots, f(e_n)$ nach (c) eine ONB des ${}^n\mathbb{R}$ ist. Ferner gilt für alle $x \in {}^n\mathbb{R}$:

$$f(x) \underset{\substack{\text{wegen} \\ (d)}}{=} \sum_{i=1}^{n} x_i f(e_i) = \underbrace{(f(e_1), \ldots, f(e_n))}_{=A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{wobei } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

□

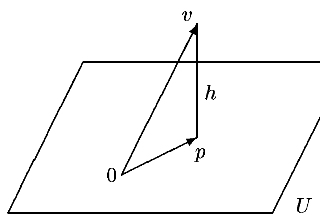
Definition 2.7. Sei $f : {}^n\mathbb{R} \to {}^n\mathbb{R}$ eine lineare Abbildung.
f heißt **orthogonal** $:\Leftrightarrow\ \forall\, x \in {}^n\mathbb{R} :\ \|f(x)\| = \|x\|$.

Folgerung 2.8. *Sei $f : {}^n\mathbb{R} \to {}^n\mathbb{R}$ eine Abbildung. Äquivalent sind:*
1. *f ist orthogonal.*
2. *f ist Bewegung mit $f(0) = 0$.*
3. *$\exists\, A \in {}^n\mathbb{R}^n$ orthogonal: $f = f_A$.*

Beweis.
1. $\Rightarrow$ 2. : Da f linear ist, gilt $f(0) = 0$. Ferner gilt für alle $x, y \in {}^n\mathbb{R}$:
$$\|f(x) - f(y)\| \underset{f \text{ linear}}{=} \|f(x-y)\| \underset{1.}{=} \|x - y\|.$$
2. $\Rightarrow$ 3. : Nach Satz 2.6 existiert ein Vektor $v \in {}^n\mathbb{R}$ sowie eine orthogonale Matrix $A \in {}^n\mathbb{R}^n$ mit $f(x) = v + Ax$ für alle $x \in {}^n\mathbb{R}$. Wegen $0 = f(0) = v + A0 = v$ ist dann $f = f_A$.
3. $\Rightarrow$ 1. : $f = f_A$ ist linear und nach Satz 2.6 eine Bewegung. Somit gilt
$$\|f(x)\| = \|f(x) - f(0)\| = \|x - 0\| = \|x\|$$
für alle $x \in {}^n\mathbb{R}$. □

Problem: Sei $V \subset {}^n\mathbb{R}$ Untervektorraum. Finde in V eine ONB.

Bemerkung 2.9. Sei $U \subset {}^n\mathbb{R}$ ein Untervektorraum mit ONB $u_1, \ldots, u_k$. Sei $v \in {}^n\mathbb{R}$. Dann ist $p := \sum_{i=1}^k \langle v, u_i\rangle u_i$ die orthogonale Projektion von v auf U.



Beweis. Es genügt zu zeigen, dass $h := v - p \perp U$ gilt, d.h. $\forall\, 1 \le j \le k: \ \langle p, u_j\rangle = \langle v, u_j\rangle$, also $h \perp u_j$.

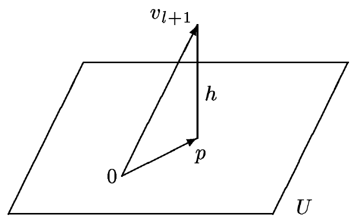
Dies ist aber erfüllt, denn für alle $1 \le j \le k$ gilt
$$\langle p, u_j\rangle = \langle \sum_{i=1}^k \langle v, u_i\rangle u_i, u_j\rangle = \sum_{i=1}^k \langle v, u_i\rangle \underbrace{\langle u_i . u_j\rangle}_{=\delta_{ij}} = \langle v, u_j\rangle.$$
□

Satz 2.10 (Schmidtsches Orthonormalisierungsverfahren)**.** *Sei $V \subset {}^n\mathbb{R}$ ein Untervektorraum mit Basis $v_1, \ldots, v_k$. Dann besitzt V eine ONB $u_1, \ldots, u_k$, so dass für alle $1 \le l \le k$ gilt:* $\operatorname{span}(u_1, \ldots, u_l) = \operatorname{span}(v_1, \ldots, v_l)$.

Beweis. Konstruiere für $1 \le l \le k$ induktiv $u_1, \ldots, u_l$ mit
(a) $u_1, \ldots, u_l$ ONS in V.
(b) $\operatorname{span}(u_1, \ldots, u_l) = \operatorname{span}(v_1, \ldots, v_l)$.

Für $l = 1$ setze $u_1 := \frac{v_1}{\|v_1\|}$. Dann sind (a) und (b) offensichtlich erfüllt.

Schritt von l auf $l+1$: Sei $l < k$, und seien $u_1, \ldots, u_l$ mit (a),(b) bereits konstruiert. Sei h das Lot von v_{l+1} auf $U := \operatorname{span}(v_1, \ldots, v_l) \underset{(b)}{=} \operatorname{span}(u_1, \ldots, u_l)$.



Nach Bemerkung 2.9 ist

$$h = v_{l+1} - p = v_{l+1} - \sum_{i=1}^{l} \langle v_{l+1}, u_i \rangle u_i.$$

Da $v_1, \ldots, v_l, v_{l+1}$ linear unabhängig sind, gilt $v_{l+1} \notin \operatorname{span}(v_1, \ldots, v_l) = U$, woraus $h \neq 0$ folgt. Setze

$$u_{l+1} := \frac{h}{\|h\|}.$$

Nun ist noch zu zeigen, dass (a) und (b) auch für $u_1, \ldots, u_{l+1}$ gelten, d.h.:

(a) $u_1, \ldots, u_{l+1}$ ONS.
(b) $\operatorname{span}(u_1, \ldots, u_{l+1}) = \operatorname{span}(v_1, \ldots, v_{l+1})$.

Beweis.

(a) Klar, da $u_{l+1} \perp \operatorname{span}(u_1, \ldots, u_l)$ und $\|u_{l+1}\| = 1$.
(b) Es gilt $\operatorname{span}(u_1, \ldots, u_{l+1}) = \operatorname{span}(u_1, \ldots, u_l) + \mathbb{R}u_{l+1} = \operatorname{span}(v_1, \ldots, v_l) + \mathbb{R}u_{l+1}$. Außerdem ist $v_{l+1} = \sum_{i=1}^{l} \langle v_{l+1}, u_i \rangle u_i + \|h\| u_{l+1}$. Daraus folgt (b). □

Problem: Was kann man über die "Struktur" orthogonaler Abbildungen im ${}^n\mathbb{R}$ sagen?

Bemerkung 2.11. Seien $A, B \in {}^n\mathbb{R}^n$ orthogonal. Dann sind auch AB und A^{-1} orthogonal.

Beweis. Es gilt

$$(AB)^T(AB) = B^T A^T AB = B^T(A^T A)B \underset{A \text{ orth.}}{=} B^T B \underset{B \text{ orth.}}{=} E$$

$$(A^{-1})^T A^{-1} \underset{A \text{ orth.}}{=} (A^T)^T A^{-1} = AA^{-1} = E$$

□

Beispiele 2.12.

1. Sei $\varphi \in \mathbb{R}$, und sei

$$A := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

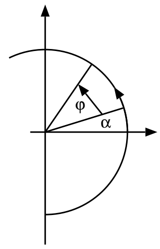
Dann ist A orthogonal, und f_A ist Drehung um den Winkel φ.

Beweis. Mit $c := \cos\varphi$ und $s := \sin\varphi$ gilt

$$AA^T = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} c^2 + s^2 & cs - sc \\ sc - cs & s^2 + c^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Außerdem ist für $\alpha \in \mathbb{R}$:

$$A \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \underset{\substack{\text{Additions-} \\ \text{theoreme}}}{=} \begin{pmatrix} \cos(\varphi+\alpha) \\ \sin(\varphi+\alpha) \end{pmatrix}.$$



2. Für $1 \le i < j \le m$ und $\varphi \in \mathbb{R}$ sei

$$R_{ij}(\varphi) := \begin{array}{@{}c@{\;}c@{}} \begin{pmatrix} 1 & 0 & \cdots & & & & & & & \cdots & 0 \\ 0 & \ddots & & & & & & & & & \vdots \\ \vdots & & 1 & & & & & & & & \vdots \\ 0 & \cdots & 0 & \cos\varphi & 0 & \cdots & 0 & -\sin\varphi & 0 & \cdots & 0 \\ \vdots & & & & 1 & & & & & & \vdots \\ & & & & & \ddots & & & & & \\ \vdots & & & & & & 1 & & & & \vdots \\ 0 & \cdots & 0 & \sin\varphi & 0 & \cdots & 0 & \cos\varphi & 0 & \cdots & 0 \\ \vdots & & & & & & & & 1 & & \vdots \\ \vdots & & & & & & & & & \ddots & 0 \\ 0 & \cdots & & & & & & & \cdots & 0 & 1 \end{pmatrix} & \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array} \end{array}$$

also $R_{ij}(\varphi) = (a_{kl})$ mit

$$a_{ii} := \cos\varphi, \qquad a_{ij} := -\sin\varphi,$$
$$a_{ji} := \sin\varphi, \qquad a_{jj} := \cos\varphi$$

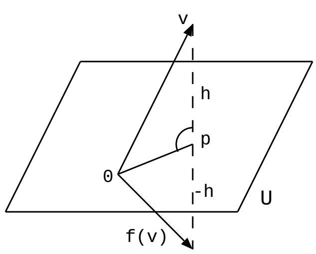
sowie

$$a_{kk} := 1 \quad \forall\, k \neq i, j$$
$$a_{kl} := 0 \quad \forall\, k \neq l,\ (k,l) \neq (i,j), (j,i)$$

Dann ist $R_{ij}(\varphi)$ orthogonal, und es gilt $R_{ij}(\varphi)^{-1} = R_{ij}(-\varphi)$. Die Matrix $R_{ij}(\varphi)$ heißt **ebene Drehmatrix** (oder Jacobi-Drehmatrix).

Beweis. Wie in 1.

3. Sei $U \subset {}^n\mathbb{R}$ ein Untervektorraum. Die Abbildung $f : {}^n\mathbb{R} \to {}^n\mathbb{R}$ sei definiert durch $f(v) := p - h$ für alle $v \in {}^n\mathbb{R}$, wobei p die orthogonale Projektion von v auf U und h das Lot von v auf U ist, also $v = p + h$ mit $p \in U$ und $h \in U^\perp$.



Dann ist f orthogonal, und es gilt $f^2 = \mathrm{id}$.

Beweis. f ist linear, denn: Seien $v, v' \in {}^n\mathbb{R}$, und sei $\alpha \in \mathbb{R}$. Schreibe $v = p + h$ und $v' = p' + h'$, wobei p und p' die orthogonalen Projektionen von v bzw. v' auf U sind. Dann gelten

$$\begin{aligned} v + v' &= \underbrace{p + p'}_{\in U} + \underbrace{h + h'}_{\in U^\perp}, \\ \alpha v &= \underbrace{\alpha p}_{\in U} + \underbrace{\alpha h}_{\in U^\perp}. \end{aligned}$$

Also sind $p + p'$ und αp die orthogonalen Projektionen von $v + v'$ bzw. αv auf U. Damit folgt

$$\begin{aligned} f(v + v') &= p + p' - (h + h') = f(v) + f(v'), \\ f(\alpha v) &= \alpha p - \alpha h = \alpha f(v). \end{aligned}$$

Somit ist die Abbildung f linear.

Ferner gilt für alle $v \in {}^n\mathbb{R}$:

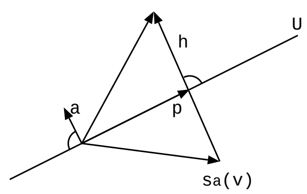
$$\begin{aligned} \|f(v)\|^2 = \|p - h\|^2 &\underset{\text{Pythagoras}}{=} \|p\|^2 + \|-h\|^2 \\ &= \|p\|^2 + \|h\|^2 \underset{\text{Pythagoras}}{=} \|p + h\|^2 = \|v\|^2, \end{aligned}$$

wobei p die orthogonale Projektion und h das Lot von v auf U ist. Somit ist insgesamt gezeigt, dass f orthogonal ist.

Offenbar gilt $f^2 = \mathrm{id}$.

4. Spezialfall von 3.:

Sei $0 \neq a \in {}^n\mathbb{R}$, und sei $U := a^\perp$ Hyperebene, also $U^\perp = \mathbb{R}a$. Für $v = p + h$ mit $p \in U$ und $h \in U^\perp = \mathbb{R}a$ sei $s_a(v) := p - h$. Dann heißt $s_a : {}^n\mathbb{R} \to {}^n\mathbb{R}$ **Hyperebenenspiegelung**.



Für alle $v \in {}^n\mathbb{R}$ gilt

$$s_a(v) = v - 2\frac{\langle v, a\rangle}{\langle a, a\rangle} \cdot a = H_a v,$$

wobei $H_a := E - \frac{2}{\langle a,a\rangle} aa^T \ \in {}^n\mathbb{R}^n$. Beachte hierbei:

$$aa^T = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} (a_1, \ldots, a_n) = (a_i a_j)_{1 \leq i,j \leq n}.$$

Die Matrix H_a heißt **Householder-Matrix**. Sie ist orthogonal und symmetrisch und erfüllt $H_a^2 = E$.

Beweis. Sei $v = p + h$ mit $p \in a^\perp$ und $h \in \mathbb{R}a$. Dann existiert ein $\alpha \in \mathbb{R}$ mit $h = \alpha a$. Wegen $\langle v, a\rangle = \langle p, a\rangle + \langle h, a\rangle = 0 + \langle \alpha a, a\rangle = \alpha\langle a, a\rangle$ gilt $\alpha = \frac{\langle v,a\rangle}{\langle a,a\rangle}$.

Somit ist

$$s_a(v) = p - h = (v - h) - h = v - 2h = v - 2\alpha \cdot a = v - 2\frac{\langle v, a\rangle}{\langle a, a\rangle} \cdot a,$$

und dies stimmt mit $H_a v$ überein, da

$$H_a v = \left(E - \frac{2}{\langle a, a\rangle} a a^T\right) v = v - \frac{2}{\langle a, a\rangle} a \underbrace{a^T v}_{\langle a,v\rangle}.$$

Außerdem ist H_a nach 3. orthogonal mit $H_a^2 = E$. Insbesondere gilt also $H_a^T = H_a^{-1} = H_a$.

5. Beispiel zu 3.: Sei

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

Wegen $\langle a, a\rangle = 5$ ist dann

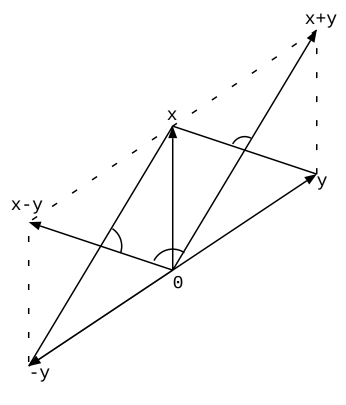
$$H_a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{2}{5}\begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & -\frac{3}{5} \end{pmatrix}.$$

Nächstes Ziel: Multipliziere gegebene Matrix von links mit Householder- oder ebener Drehmatrix, so dass gezielt Einträge zu Null gemacht werden.

Lemma 2.13. *Seien $x, y \in {}^n\mathbb{R}$ mit $\|x\| = \|y\|$.*

1. *Es ist $x - y \perp x + y$.*
2. *Falls $x - y \neq 0$, so gilt $s_{x-y}(x) = y$.*
3. *Falls $x + y \neq 0$, so gilt $s_{x+y}(x) = -y$.*

Beweis.



1. Es gilt

$$\langle x - y, x + y\rangle = \underbrace{\langle x, x\rangle - \langle y, y\rangle}_{=0,\text{ da } \|x\|^2 = \|y\|^2} + \langle x, y\rangle - \langle y, x\rangle = 0.$$

2.,3. Wegen $x = \frac{x-y}{2} + \frac{x+y}{2}$ folgt aus 1.

$$s_{x-y}(x) = \frac{x+y}{2} - \frac{x-y}{2} = y,$$

falls $x - y \neq 0$, und

$$s_{x+y}(x) = \frac{x-y}{2} - \frac{x+y}{2} = -y,$$

falls $x + y \neq 0$. $\square$

Lemma 2.14. *Seien $1 \leq i < j \leq n$ und $x_1, \ldots, x_n \in \mathbb{R}$. Dann existiert ein $\varphi \in \mathbb{R}$, so dass*

$$R_{ij}(\varphi) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix},$$

wobei $x_i' \geq 0$, $x_j' = 0$ und $x_l' = x_l$ für alle $1 \leq l \leq n$ mit $l \neq i$ und $l \neq j$.

Beweis. Für beliebiges φ gilt mit $c := \cos\varphi$ und $s := \sin\varphi$

$$R_{ij}(\varphi) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix},$$

wobei $x_i' = cx_i - sx_j$ und $x_j' = sx_i + cx_j$ sowie $x_l' = x_l$ für alle $l \neq i, j$.

Falls $x_i = 0$ und $x_j = 0$ gilt, so kann $\varphi = 0$ gewählt werden, also $R_{ij}(\varphi) = E$.

Falls $x_i^2 + x_j^2 \neq 0$ ist, setze

$$c := \frac{x_i}{\sqrt{x_i^2 + x_j^2}} \qquad \text{und} \qquad s := \frac{-x_j}{\sqrt{x_i^2 + x_j^2}}.$$

Wegen

$$c^2 + s^2 = \frac{x_i^2 + x_j^2}{\left(\sqrt{x_i^2 + x_j^2}\right)^2} = 1$$

existiert ein Winkel $\varphi \in \mathbb{R}$ mit $c = \cos\varphi$ und $s = \sin\varphi$. Hierfür sind die geforderten Eigenschaften erfüllt, denn:

$$x_i' = cx_i - sx_j = \frac{x_i^2 + x_j^2}{\sqrt{x_i^2 + x_j^2}} > 0,$$

$$x_j' = sx_i + cx_j = \frac{-x_i x_j + x_i x_j}{\sqrt{x_i^2 + x_j^2}} = 0.$$

$\square$

Folgerung 2.15. *Sei $n \geq 2$, und sei $0 \neq x \in {}^n\mathbb{R}$. Dann existiert eine orthogonale Matrix $Q \in {}^n\mathbb{R}^n$ mit*

$$Qx = \|x\| \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hierbei kann Q gewählt werden als:

1. Householder-Matrix oder E

oder 2. Produkt von $n-1$ ebenen Drehmatrizen.

Beweis.

1. Setze
$$y := \|x\| \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$
Dann gilt $\|x\| = \|y\|$, und die Behauptung folgt aus Lemma 2.13.
2. Nach Lemma 2.14 gibt es eine ebene Drehmatrix $R_{n-1,n}$ mit passendem Winkel φ, so dass
$$R_{n-1,n} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-2} \\ x'_{n-1} \\ 0 \end{pmatrix},$$
wobei $x'_{n-1} \geq 0$ gilt. Durch Iteration erhält man dann ebene Drehmatrizen $R_{1,2}, R_{2,3}, \ldots, R_{n-1,n}$ mit
$$\underbrace{R_{1,2} R_{2,3} \cdots R_{n-1,n}}_{=:Q} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x'_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$
wobei $x'_1 \geq 0$ gilt. Da Q orthogonal ist, folgt dann $x'_1 = \|Qx\| = \|x\|$. Also ist
$$Qx = \|x\| \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

□

Satz 2.16 (QR-Zerlegung)**.** *Seien m, n mit $m \geq n$ und $m \geq 2$, und sei $A \in {}^m\mathbb{R}^n$. Dann existieren eine orthogonale Matrix $Q \in {}^m\mathbb{R}^m$ und*
$$R = \begin{pmatrix} \alpha_1 & * & * \\ & \ddots & * \\ & & \alpha_n \\ 0 & & \end{pmatrix} \in {}^m\mathbb{R}^n, \ \alpha_1, \ldots, \alpha_n \in \mathbb{R}$$
mit $A = QR$. Hierbei kann Q gewählt werden als:

1. Produkt von Householder-Matrizen oder E
oder 2. Produkt von ebenen Drehmatrizen.

Beweis. Seien $a_1, \ldots, a_n \in {}^m\mathbb{R}$ die Spalten von A, also $A = (a_1, \ldots, a_n)$. Nach 2.15 gibt es eine orthogonale Matrix Q_1 wie im Satz mit
$$Q_1 A = (Q_1 a_1, Q_1 a_2, \ldots, Q_1 a_n) = \begin{pmatrix} \alpha_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix},$$

wobei $\alpha_1 = \|a_1\|$ und

$$Q_1 a_1 = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Ebenso kann man mit A' verfahren: Nach 2.15 gibt es eine orthogonale Matrix $Q_2' \in {}^{m-1}\mathbb{R}^{m-1}$ mit

$$Q_2' A' = \begin{pmatrix} \alpha_2 & * & \dots & * \\ 0 & & & \\ \vdots & & A'' & \\ 0 & & & \end{pmatrix}.$$

Damit gilt

$$\underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q_2' & \\ 0 & & & \end{pmatrix}}_{=:Q_2} Q_1 A = \begin{pmatrix} \alpha_1 & * & \dots & \dots & * \\ 0 & \alpha_2 & * & \dots & * \\ \vdots & 0 & & & \\ \vdots & \vdots & & A'' & \\ 0 & 0 & & & \end{pmatrix}.$$

Hierbei ist Q_2 wieder orthogonal und Householder- bzw. ebene Drehmatrix, falls dies für Q_2' gilt, denn für Householder-Matrizen gilt z.B.

$$H_{\binom{0}{a}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & H_a & \\ 0 & & & \end{pmatrix}.$$

Durch Iteration folgt: Es gibt Q, R wie im Satz mit

$$QA = R = \begin{pmatrix} \alpha_1 & * & * \\ & \ddots & * \\ & & \alpha_n \\ 0 & & \end{pmatrix}.$$

Hieraus ergibt sich aber $A = Q^T R$, wobei Q^T mit Q orthogonal und von der angegebenen Gestalt ist. □

Bemerkungen 2.17.

1. Eine QR-Zerlegung erhält man auch aus dem Schmidtschen Orthogonalisierungsverfahren, allerdings ohne weitere Informationen über Q. Das Schmidtsche Orthogonalisierungsverfahren ist auch numerisch oft ungünstig.
2. Falls eine QR-Zerlegung $A = QR$ von A bekannt ist, so wird dadurch das Problem häufig auf den Fall einer oberen Dreiecksmatrix reduziert. Insbesondere gilt für die Methode der kleinsten Quadrate mit $b \in {}^m\mathbb{R}$ und $x \in {}^n\mathbb{R}$:

$$\|Ax - b\| = \|QRx - b\| \underset{Q^T \text{ orth.}}{=} \|Rx - Q^T b\|.$$

Lemma 2.18. *Sei $B = (b_{ij}) \in {}^n\mathbb{R}^n$ eine obere Dreiecksmatrix, also*

$$B = \begin{pmatrix} b_{11} & & & * \\ & b_{22} & & \\ & & \ddots & \\ 0 & & & b_{nn} \end{pmatrix}.$$

Falls B orthogonal ist, so ist B eine Diagonalmatrix, d.h.

$$B = \begin{pmatrix} b_{11} & & 0 \\ & \ddots & \\ 0 & & b_{nn} \end{pmatrix},$$

und für alle $1 \leq i \leq n$ gilt $b_{ii} = \pm 1$.

Beweis. Seien $b_1, \dots, b_n$ die Spalten von B. Da B orthogonal ist, gilt dann $\langle b_i, b_j \rangle = \delta_{ij}$ für alle $1 \leq i, j \leq n$.

Insbesondere ist $1 = \langle b_1, b_1 \rangle = b_{11}^2$, woraus $b_{11} = \pm 1$ folgt. Ferner gilt dann $0 = \langle b_1, b_j \rangle = b_{11} b_{1j} = \pm b_{1j}$ für alle $2 \leq j \leq n$. Somit muss $b_{1j} = 0$ sein für alle $2 \leq j \leq n$. Also hat B die Gestalt

$$B = \begin{pmatrix} b_{11} & 0 & \dots & 0 \\ 0 & b_{22} & & * \\ \vdots & 0 & \ddots & \\ 0 & 0 & & b_{nn} \end{pmatrix}.$$

Ebenso kann man nun mit der zweiten Spalte verfahren: Aus $1 = \langle b_2, b_2 \rangle = b_{22}^2$ folgt $b_{22} = \pm 1$, und für alle $3 \leq j \leq n$ gilt dann $0 = \langle b_2, b_j \rangle = b_{22} b_{2j} = \pm b_{2j}$, weshalb $b_{2j} = 0$ sein muss für alle $3 \leq j \leq n$.

Durch Iteration folgt die Behauptung. □

Folgerung 2.19. *Sei $f : {}^n\mathbb{R} \to {}^n\mathbb{R}$ eine orthogonale Abbildung.*

Dann ist f Produkt (= Hintereinanderausführung) von höchstens n Hyperebenenspiegelungen (oder id*).*

Beweis. Nach 2.8 hat f die Form $f = f_A$, wobei $A \in {}^n\mathbb{R}^n$ orthogonal ist. Nach Satz 2.16 gibts es eine orthogonale Matrix Q und eine obere Dreiecksmatrix

$$R = \begin{pmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

mit $A = QR$. Hierbei können Q und R so gewäht werden, dass $Q = Q_1 \cdots Q_{n-1}$ ist, wobei Q_i entweder eine Householder-Matrix oder E ist für alle $1 \leq i < n$, sowie $\alpha_1 > 0, \dots, \alpha_{n-1} > 0$ gelten. (Siehe Beweis von 2.16.)

Da $R = Q^T A$ mit Q^T und A orthogonal ist, folgt dann aus Lemma 2.18:

$$R = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & \pm 1 \end{pmatrix},$$

d.h. $R = E$ oder

$$R = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix} = \text{Householder-Matrix } H_{e_n}.$$

Also ist entweder $A = Q$ oder $A = QH_{e_n}$, und die Behauptung ist gezeigt. □

Bemerkungen 2.20.

1. Ebenso folgt aus Satz 2.16: Jede orthogonale Abbildung ist Produkt von ebenen Drehungen oder Produkt von ebenen Drehungen und einer Hyperebenespiegelung.
2. Für $n = 2$ gilt: Die orthogonalen 2×2-Matrizen haben genau die Form:

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

mit $\varphi \in \mathbb{R}$.

Beweis. Sei $A \in {}^2\mathbb{R}^2$ orthogonal. Nach Lemma 2.14 gibt es dann ein $\varphi \in \mathbb{R}$ mit

$$R_{12}(\varphi)A = \begin{pmatrix} \alpha_1 & \alpha_3 \\ 0 & \alpha_2 \end{pmatrix}, \quad \alpha_1 > 0.$$

Nach Lemma 2.18 folgt dann $\alpha_1 = 1$, $\alpha_3 = 0$ und $\alpha_2 = \pm 1$. Somit ist entweder

$$A = R_{12}(\varphi)^{-1} = R_{12}(-\varphi)$$

oder

$$\begin{aligned} A = R_{12}(-\varphi)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ -\sin\varphi & -\cos\varphi \end{pmatrix} \end{aligned}$$

Dies ist die Behauptung (ersetze φ durch $-\varphi$). □

3. Gruppen, insbesondere Symmetriegruppen

Definition 3.1. Sei G eine Menge und $m : G \times G \to G$ eine Abbildung. Für alle $x, y \in G$ wird folgende Schreibweise verwendet: $xy := x \cdot y := m(x, y)$.
(G, m) oder kurz G heißt **Gruppe** $:\Leftrightarrow$

1. $\forall\, x, y, z \in G : x(yz) = (xy)z$ (**assoziativ**)
2. $\exists\, e \in G :$
   (a) $\forall\, x \in G : xe = x = ex$ (**neutrales Element**)
   (b) $\forall\, x \in G\ \exists\, x^{-1} \in G : xx^{-1} = e = x^{-1}x$ (**inverses Element**)

Hierbei heißt m auch **Produkt** oder **Verknüpfung** von Elementen von G, e **neutrales Element** und x^{-1} **inverses Element** zu x.

Bemerkung 3.2. Sei G Gruppe. Dann gilt:

1. Das neutrale Element in G ist eindeutig bestimmt.
2. Das inverse Element zu x in G ist eindeutig bestimmt.

Beweis.

1. Seien e, e' neutrale Elemente. Dann gilt $e = ee' = e'$. Hierbei wurde bei der ersten Gleichheit verwendet, dass e' ein neutrales Element ist, und bei der zweiten, dass e ein neutrales Element ist.
2. Seien x', x'' inverse Elemente zu x. Dann gilt
$$x' = x'e = x'(xx'') = (x'x)x'' = ex'' = x''.$$

□

Beispiele 3.3.

1. $\mathbb{R}$ mit der Addition als Verknüpfung bildet eine Gruppe mit neutralem Element 0 und inversem Element $-x$ zu $x \in \mathbb{R}$.
2. $\mathbb{R} \setminus \{0\}$ mit der Multiplikation ist eine Gruppe mit neutralem Element 1 und inversem Element $x^{-1} = \frac{1}{x}$ zu $x \in \mathbb{R} \setminus \{0\}$.

   Beweis. Die Gruppenaxiome folgen aus den üblichen Rechenregeln.
3. Seien $m, n \geq 1$. Dann ist ${}^m\mathbb{R}^n$ mit der Addition von Matrizen eine Gruppe mit der Nullmatrix als neutralem Element und inversem Element $-A$ zu $A \in {}^m\mathbb{R}^n$.

   Beweis. Die Gruppenaxiome folgen aus den Rechenregeln für Matrizen.
4. Sei $n \geq 1$. Dann ist
$$\mathrm{GL}_n(\mathbb{R}) := \{A \in {}^n\mathbb{R}^n \mid A \text{ invertierbar}\}$$
(GL = general linear group) mit der Multiplikation von Matrizen eine Gruppe mit der Einheitsmatrix E als neutralem Element und der inversen Matrix als inverses Element.

   Beweis. Bekannt aus Kapitel I.
5. Sei $n \geq 1$. Dann sind
$$\mathrm{Aut}_{\mathbb{R}}({}^n\mathbb{R}) := \{f : {}^n\mathbb{R} \to {}^n\mathbb{R} \mid f \text{ linear und bijektiv}\}$$
(Aut = Automorphismus) und
$$\mathrm{Aff}({}^n\mathbb{R}) := \{g : {}^n\mathbb{R} \to {}^n\mathbb{R} \mid g \text{ Affinität}\}$$
mit der Hintereinanderausführung von Abbildungen Gruppen mit der identischen Abbildung id als neutralem Element und der Umkehrabbildung als inversem Element.

   Beweis. Bekannt aus Kapitel I und den Übungen.

Bemerkungen 3.4.

1. Sei G eine Gruppe, und seien $x, y \in G$. Dann gilt:
   (a) $(xy)^{-1} = y^{-1}x^{-1}$.
   (b) $(x^{-1})^{-1} = x$.

   Beweis.
   (a) Es gilt $(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xx^{-1} = e$. Ebenso erhält man $(y^{-1}x^{-1})(xy) = e$. Also gilt $(xy)^{-1} = y^{-1}x^{-1}$, da das inverse Element eindeutig bestimmt ist.
   (b) Es gilt $xx^{-1} = e = x^{-1}x$. Wegen der Eindeutigkeit des inversen Elementes ist dann $(x^{-1})^{-1} = x$.

2. Sei G eine Gruppe und $x \in G$. Dann sei $x^0 := e$, und für alle $n \in \mathbb{N}$ mit $n \geq 1$ sei
$$x^n := \underbrace{x \cdot x \cdot \ldots \cdot x}_{\text{n-mal}} \qquad \text{und} \qquad x^{-n} := (x^{-1})^n.$$
Dies ist möglich, da die Gruppenverknüpfung assoziativ ist.

Definition 3.5. Sei $n \geq 1$.

1. Sei $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ eine Abbildung.
   σ heißt **Permutation** von $\{1, 2, \ldots, n\}$ $:\Leftrightarrow$ σ ist bijektiv.
2. Sei $S_n := \{\sigma \mid \sigma \text{ Permutationen von } \{1, 2, \ldots, n\}\}$.
   Schreibweise für $\sigma \in S_n$:
$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Bemerkungen 3.6.

1. S_n ist mit der Hintereinanderausführung von Abbildungen eine Gruppe mit neutralem Element id und heißt **symmetrische Gruppe**. Hierbei ist das inverse Element zu σ die Umkehrabbildung.
2. Symmetrische Gruppe S_n für kleine n.
   (a) $n = 1$: Dann ist $S_1 = \{\text{id}\}$ die Gruppe mit einem Element.
   (b) $n = 2$: Dann ist $S_2 = \{\text{id}, \sigma\}$ mit
$$\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$
   d.h. $\sigma(1) = 2$ und $\sigma(2) = 1$. Somit vertauscht σ die Werte 1 und 2. Wegen $\sigma^2 = \text{id}$ ist $\sigma = \sigma^{-1}$.
   (c) $n = 3$: Sei
$$\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \in S_3,$$
   also $\sigma(1) = 1,\ \sigma(2) = 3,\ \sigma(3) = 2$. Dann gilt $\sigma^2 = \text{id}$.
   Sei
$$\rho := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \in S_3,$$
   also $\rho(1) = 3,\ \rho(2) = 1,\ \rho(3) = 2$. Wegen
$$\rho^2(1) = \rho(\rho(1)) = \rho(3) = 2,$$
$$\rho^2(2) = \rho(\rho(2)) = \rho(1) = 3,$$
$$\rho^2(3) = \rho(\rho(3)) = \rho(2) = 1$$
   ist
$$\rho^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$
   Weiter gilt
$$\rho^3(1) = \rho(\rho^2(1)) = \rho(2) = 1,$$
$$\rho^3(2) = \rho(\rho^2(2)) = \rho(3) = 2,$$
$$\rho^3(3) = \rho(\rho^2(3)) = \rho(1) = 3,$$
   woraus $\rho^3 = \text{id}$ und damit $\rho^{-1} = \rho^2$ folgt.

Definition 3.7. $\tau \in S_n$ heißt **Transposition** $:\Leftrightarrow \exists 1 \leq i, j \leq n,\ i \neq j : \ \forall 1 \leq k \leq n :$

$$\tau(k) = \begin{cases} k, & \text{falls } k \neq i, j \\ j, & \text{falls } k = i \\ i, & \text{falls } k = j \end{cases}$$

(τ vertauscht also i, j und lässt alle anderen Elemente fest.)

Satz 3.8. *Sei* $n \geq 1$.

1. S_n *hat* $n!$ *Elemente.*
2. *Für alle* $\sigma \in S_n$ *gibt es* $m \geq 0$ *und Transpositionen* $\tau_1, \ldots, \tau_m$ *mit* $\sigma = \tau_1\, \tau_2 \ldots \tau_m$ *(*$\sigma = \mathrm{id}$*, falls* $m = 0$*).*

Beweis.

1. Induktion nach n:
   (a) Der Induktionsanfang $n = 1$ ist klar.
   (b) Induktionsschritt von $n - 1$ auf n für $n \geq 2$:
   Ein $\sigma \in S_n$ ist gegeben durch
   - $\sigma(n) := i$ für ein beliebiges $i \in \{1, \ldots, n\}$ und
   - verschiedene Werte $\sigma(1), \ldots, \sigma(n-1)$ aus $\{k \in \mathbb{N} \mid 1 \leq k \leq n,\ k \neq i\}$.

   Nach Induktionsvorraussetzung gibt es für die Werte $\sigma(1), \ldots, \sigma(n-1)$ genau $(n-1)!$ viele Möglichkeiten.
   Somit gibt es für σ insgesamt $n(n-1)! = n!$ Möglichkeiten.
2. Induktion nach n:
   (a) Der Induktionsanfang $n = 1$ ist klar.
   (b) Induktionsschritt von $n - 1$ auf n für $n \geq 2$: Betrachte

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

   Da σ bijektiv ist, gibt es ein $i \in \{1, \ldots, n\}$ mit $\sigma(i) = n$.
   1. Fall: $i < n$, also

$$\sigma = \begin{pmatrix} 1 & \ldots & i & \ldots & n \\ \sigma(1) & \ldots & n & \ldots & \sigma(n) \end{pmatrix}.$$

   Mit der Transposition $\tau := (\sigma(i)\ \ \sigma(n)) = (n\ \ \sigma(n))$ gilt dann

$$\tau\sigma = \begin{pmatrix} 1 & \ldots & i & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) & \ldots & n \end{pmatrix},$$

   also $\tau\sigma(n) = n$.
   2. Fall: $i = n$, also $\sigma(n) = n$.
   In beiden Fällen kann also $\tau\sigma$ bzw. σ als Permutation von $\{1, \ldots, n-1\}$ betrachtet werden.
   Nach der Induktionsvoraussetzung gibt es dann Transpositionen $\tau_1, \ldots, \tau_m$ mit $\tau\sigma = \tau_1 \cdots \tau_m$ bzw. $\sigma = \tau_1 \cdots \tau_m$. Wegen $\tau^2 = \mathrm{id}$ folgt im 1. Fall $\sigma = \tau\, \tau_1 \cdots, \tau_m$. □

Definition 3.9. Sei G eine Gruppe und $H \subset G$ eine Teilmenge.
H heißt **Untergruppe** von G $:\Leftrightarrow$

1. $e \in H$.

2. $\forall\, x, y \in H: \; xy \in H$.
3. $\forall\, x \in H: \; x^{-1} \in H$.

Bemerkung 3.10. Sei G eine Gruppe und $H \subset G$ eine Teilmenge.

Dann ist H genau dann eine Untergruppe von G, wenn H mit der Einschränkung der Verknüpfung von G eine Gruppe ist. In diesem Fall stimmt das neutrale Element von H mit dem neutralen Element von G überein, und das inverse Element von $x \in H$ ist gleich dem inversen Element von x in G.

Beweis. Klar. □

Beispiele 3.11.

1. $\mathrm{Is}({}^n\mathbb{R}) := \{g : {}^n\mathbb{R} \to {}^n\mathbb{R} \mid g \text{ Bewegung}\}$ (Is = **Isometriegruppe des** $^{\mathbf{n}}\mathbb{R}$) ist Untergruppe von $\mathrm{Aff}({}^n\mathbb{R})$.

   Beweis.
   (a) id ist offenbar eine Bewegung.
   (b) Seien $g, h \in \mathrm{Is}({}^n\mathbb{R})$. Dann gilt $gh,\; g^{-1} \in \mathrm{Is}({}^n\mathbb{R})$, d.h., gh und g^{-1} sind Bewegungen, denn für alle $x, y \in {}^n\mathbb{R}$ gilt

$$\|g(h(x)) - g(h(y))\| \underset{g \text{ Bewegung}}{=} \|h(x) - h(y)\| \underset{h \text{ Bewegung}}{=} \|x - y\|,$$

$$\|g^{-1}(x) - g^{-1}(y)\| \underset{g \text{ Bewegung}}{=} \|g(g^{-1}(x)) - g(g^{-1}(y))\| = \|x - y\|.$$

2. $\mathrm{O}_n := \{A \in {}^n\mathbb{R}^n \mid A \text{ orthogonal}\}$, genannt **orthogonale Gruppe**, ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$.

   Beweis. Folgt aus Bemerkung 2.11.

Definition 3.12. Sei $M \subset {}^n\mathbb{R}$ eine Teilmenge. Dann heißt

$$G_M := \{g \in \mathrm{Is}({}^n\mathbb{R}) \mid g(M) = M\}$$

Symmetriegruppe von M.

Bemerkungen 3.13.

1. Sei $M \subset {}^n\mathbb{R}$ eine Teilmenge. Dann ist G_M eine Untergruppe von $\mathrm{Is}({}^n\mathbb{R})$.

   Beweis.
   (a) Nach Definition ist $\mathrm{id} \in G_M$.
   (b) Für alle $g, h \in G_M$ gilt $gh,\; g^{-1} \in G_M$, denn:

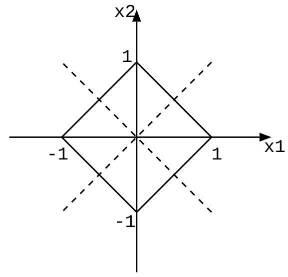
$$g(h(M)) \underset{h \in G_M}{=} g(M) \underset{g \in G_M}{=} M,$$

$$g^{-1}(M) \underset{g \in G_M}{=} g^{-1}(g(M)) = M.$$

2. Für

$$M := \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}$$

   ist G_M die Symmetriegruppe des Quadrates.



Sei
$$A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \text{und} \qquad B := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
Dann ist $r := f_A$ die Drehung um 90° und $s := f_B$ die Spiegelung an der x_1-Achse, und es gilt
$$G_M = \{\overbrace{\mathrm{id}, r, r^2, r^3}^{\text{4 Drehungen}}, \underbrace{s, rs, r^2s, r^3s}_{\text{4 Spiegelungen}}\}.$$

Beweis. Übungen.

Definition 3.14. Seien G und G' Gruppen.

1. Sei $\varphi : G \to G'$ eine Abbildung.
   (a) φ heißt **Homomorphismus** von Gruppen $:\Leftrightarrow$
   $\forall\, x, y \in G: \ \varphi(xy) = \varphi(x)\varphi(y)$.
   (b) φ heißt **Mono-**, bzw **Epi-**, bzw. **Isomorphismus** $:\Leftrightarrow$
   φ ist Homomorphismus, und φ ist injektiv, bzw. surjektiv, bzw. bijektiv.
2. $G \cong G'$ **isomorph** $:\Leftrightarrow$ Es gibt einen Isomorphismus $\varphi : G \to G'$.

Beispiel 3.15. Die Abbildung $\pi : \mathbb{R} \to \mathrm{O}_2$ mit
$$\pi(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$
für alle $\varphi \in \mathbb{R}$ ist ein Homomorphismus von Gruppen, wobei $\mathbb{R}$ Gruppe bezüglich der Addition ist.

Beweis. Seien $\varphi, \psi \in \mathbb{R}$. Mit $c := \cos\varphi$, $s := \sin\varphi$, $c' := \cos\psi$ und $s' := \sin\psi$ gilt
$$\begin{aligned} \pi(\varphi)\pi(\psi) &= \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} c' & -s' \\ s' & c' \end{pmatrix} = \begin{pmatrix} cc' - ss' & -cs' - sc' \\ sc' + cs' & cc' - ss' \end{pmatrix} \\ &= \begin{pmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} = \pi(\varphi+\psi). \end{aligned}$$
Hierbei wurden die Additionstheoreme von sin und cos verwendet. □

Bemerkung 3.16. Sei $\varphi : G \to G'$ ein Homomorphismus von Gruppen. Dann gilt:

1. (a) $\varphi(e) = e$, wobei e das neutrale Element in G bzw. G' bezeichnet.
   (b) $\varphi(x^{-1}) = \varphi(x)^{-1}$ für alle $x \in G$.

Beweis.

(a) Es gilt

$$\varphi(e) = \varphi(e \cdot e) \underset{\varphi \text{ Homo.}}{=} \varphi(e)\varphi(e).$$

Durch Multiplikation mit $\varphi(e)^{-1}$ folgt hieraus

$$e = \varphi(e)^{-1}\varphi(e) = \varphi(e)^{-1}\varphi(e)\varphi(e) = \varphi(e).$$

(b) Es gilt

$$e \underset{(a)}{=} \varphi(e) = \varphi(xx^{-1}) \underset{\varphi \text{ Homo.}}{=} \varphi(x)\varphi(x^{-1}).$$

Ebenso ergibt sich $e = \varphi(x^{-1})\varphi(x)$. Aus der Eindeutigkeit des inversen Elementes folgt dann $\varphi(x^{-1}) = \varphi(x)^{-1}$.

2. $\mathrm{Ke}(\varphi) = \{x \in G \mid \varphi(x) = e\} \subset G$ und $\mathrm{Bi}(\varphi) \subset G'$ sind Untergruppen.

Beweis.

(a) $\mathrm{Ke}(\varphi) \subset G$ ist Untergruppe, denn:
- Es ist $e \in \mathrm{Ke}(\varphi)$, da $\varphi(e) = e$.
- Für alle $x, y \in \mathrm{Ke}(\varphi)$ gilt

$$\varphi(xy) = \varphi(x)\varphi(y) = ee = e$$

und

$$\varphi(x^{-1}) \underset{1.}{=} \varphi(x)^{-1} = e^{-1} = e,$$

d.h., es gilt $xy, x^{-1} \in \mathrm{Ke}(\varphi)$.

(b) $\mathrm{Bi}(\varphi) \subset G'$ ist Untergruppe, denn:
- Es ist $e \in \mathrm{Bi}(\varphi)$, da $\varphi(e) = e$.
- Seien $u, v \in \mathrm{Bi}(\varphi)$. Dann existieren $x, y \in G$ mit $u = \varphi(x)$ und $v = \varphi(y)$. Damit gilt

$$uv = \varphi(x)\varphi(y) = \varphi(xy) \in \mathrm{Bi}(\varphi),$$
$$u^{-1} = \varphi(x)^{-1} \underset{1.}{=} \varphi(x^{-1}) \in \mathrm{Bi}(\varphi).$$

Kapiel III: Determinante, Eigenwerte und abstrakte lineare Algebra

1. Determinanten

Definition 1.1. Sei $n \geq 1$, und sei $d : {}^n\mathbb{R}^n \to \mathbb{R}$ eine Abbildung. Wir betrachten $d(A) = d(a_1, \dots, a_n)$ für $A \in {}^n\mathbb{R}^n$ mit den Spalten $a_1, \dots, a_n$ als Funktion der Spalten.

1. d heißt **multilinear** $:\Leftrightarrow$
   $\forall\, 1 \leq i \leq n,\ a_1, \dots, a_n \in {}^n\mathbb{R}$:
   $${}^n\mathbb{R} \to \mathbb{R},\ x \mapsto d(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n)$$
   ist $\mathbb{R}$-linear.
2. d heißt **alternierend** $:\Leftrightarrow$
   $\forall\ a_1, \dots, a_n \in {}^n\mathbb{R},\ 1 \leq i < n$: $a_i = a_{i+1} \ \Rightarrow\ d(a_1, \dots, a_n) = 0$.
3. d heißt **normiert** $:\Leftrightarrow$ $d(E) = 1$, wobei E die Einheitsmatrix ist.
4. d heißt **Determinante** $:\Leftrightarrow$ d ist multilinear, alternierend und normiert.

Beispiel 1.2. Es gibt genau eine Determinante $d : {}^2\mathbb{R}^2 \to \mathbb{R}$, nämlich:
$$d\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} := \det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \alpha\delta - \beta\gamma$$
für alle $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Beweis. Offenbar ist det eine Determinante.

Für die Eindeutigkeit zeige noch: Ist $d : {}^2\mathbb{R}^2 \to \mathbb{R}$ multilinear und alternierend, so gilt $d(A) = det(A)d(E)$ für alle $A \in {}^2\mathbb{R}^2$.

Sei also $d : {}^2\mathbb{R}^2 \to \mathbb{R}$ multilinear und alternierend.

1. Für alle $a_1, a_2 \in {}^2\mathbb{R}$ gilt $d(a_1, a_2) = -d(a_2, a_1)$, denn
   $$\begin{aligned} 0 &\underset{d \text{ altern.}}{=} d(a_1 + a_2, a_1 + a_2) \underset{d \text{ multilin.}}{=} d(a_1, a_1 + a_2) + d(a_2, a_1 + a_2) \\ &\underset{d \text{ multilin.}}{=} \underbrace{d(a_1, a_1)}_{=0, \text{ da } d \text{ altern.}} + d(a_1, a_2) + d(a_2, a_1) + \underbrace{d(a_2, a_2)}_{=0, \text{ da } d \text{ altern.}} \end{aligned}$$
2. Sei
   $$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in {}^2\mathbb{R}^2$$
   mit Spalten a_1 und a_2. Dann ist $a_1 = \alpha e_1 + \gamma e_2$ und $a_2 = \beta e_1 + \delta e_2$, wobei e_1, e_2 die Standardbasis des ${}^2\mathbb{R}$ ist. Damit gilt
   $$\begin{aligned} d(A) &= d(a_1, a_2) = d(\alpha e_1 + \gamma e_2, \beta e_1 + \delta e_2) \\ &\underset{d \text{ multilin.}}{=} \alpha\beta \underbrace{d(e_1, e_1)}_{=0} + \alpha\delta d(e_1, e_2) + \gamma\beta d(e_2, e_1) + \gamma\delta \underbrace{d(e_2, e_2)}_{=0} \\ &= \alpha\delta d(e_1, e_2) + \gamma\beta d(e_2, e_1) \underset{1.}{=} (\alpha\delta - \beta\gamma) d(e_1, e_2) \\ &= \det(A) d(e_1, e_2) = \det(A) d(E) \end{aligned}$$

Definition 1.3. Sei $A \in {}^n\mathbb{R}^n$. Für $1 \leq i, j \leq n$ sei A_{ij} diejenige Matrix aus ${}^{n-1}\mathbb{R}^{n-1}$, die aus A durch Streichen der i-ten Zeile und der j-ten Spalte entsteht.

Beispiel 1.4. Sei

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ -1 & 2 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 2 & 4 & 5 & 6 \end{pmatrix} \in {}^4\mathbb{R}^4.$$

Dann ist zum Beispiel

$$A_{11} = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \text{und} \quad A_{23} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 2 & 4 & 6 \end{pmatrix}.$$

Satz 1.5. *Definiere $d_n : {}^n\mathbb{R}^n \to \mathbb{R}$ für $n \geq 1$ induktiv wie folgt:*

1. *$d_1(a) := a$ für alle $a \in {}^1\mathbb{R}^1 = \mathbb{R}$.*
2. *Ist d_{n-1} für $n \geq 2$ bereits definiert, so wähle ein $1 \leq i \leq n$ und definiere für alle $A \in {}^n\mathbb{R}^n$:*

$$d_n(A) := \sum_{j=1}^{n} (-1)^{i+j} a_{ij} d_{n-1}(A_{ij}).$$

Dann ist d_n für alle $n \geq 1$ eine Determinante.

Beweis. Induktion nach n.

1. Der Indutionsanfang $n = 1$ ist klar.
2. Induktionsschritt von $n-1$ auf n für $n \geq 2$:
   Sei $1 \leq i \leq n$ fest gewählt. Für $1 \leq j \leq n$ seien Abbildungen $f_j : {}^n\mathbb{R}^n \to \mathbb{R}$ definiert durch

$$d_n(A) = \sum_{j=1} \underbrace{(-1)^{i+j} a_{ij} d_{n-1}(A_{ij})}_{=:f_j(a_1,\ldots,a_n)}.$$

   Hierbei sind $a_1, \ldots, a_n$ die Spalten von A.
   (a) Zeige: d_n ist multilinear.

   Beweis. Es genügt zu zeigen, dass die Abbildung f_j multilinear ist für alle $1 \leq j \leq n$. Sei dazu $1 \leq l \leq n$ beliebig. Betrachte nun $f_j(a_1, \ldots, a_n)$ als Funktion von a_l, wobei alle anderen Spalten fest bleiben.
   1. Fall: $l \neq j$. Dann ist

$$f_j(a_1, \ldots, a_n) = \underbrace{(-1)^{i+j} a_{ij}}_{\substack{\text{unabhängig} \\ \text{von } a_l}} \underbrace{d_{n-1}(A_{ij})}_{\substack{\text{linear in } a_l, \\ \text{da } d_{n-1} \text{ multilin.} \\ \text{nach IV}}}$$

   linear in a_l.
   2.Fall: $l = j$. Dann ist

$$f_j(a_1, \ldots, a_n) = (-1)^{i+j} \underbrace{a_{ij}}_{=a_{il}} \underbrace{d_{n-1}(A_{ij})}_{\substack{\text{unabhängig von } a_l \\ \text{wegen } j=l}}$$

   linear in a_l.

(b) Zeige: d_n ist alternierend

Beweis. Sei $1 \leq l \leq n$ und $a_l = a_{l+1}$. Für $j \neq l,\ l+1$ enthält dann A_{ij} zwei gleiche Spalten. Somit gilt in diesem Fall $d_{n-1}(A_{ij}) = 0$ nach Induktionsvoraussetzung. Dann ist aber

$$d_n(A) = (-1)^{i+l} a_{il} d_{n-1}(A_{il}) + (-1)^{i+l+1} a_{i,l+1} d_{n-1}(A_{i,l+1}) = 0,$$

da $a_{il} = a_{i,l+1}$ und $A_{il} = A_{i,l+1}$ wegen $a_l = a_{l+1}$ gelten.

(c) Zeige: $d_n(E) = 1$.

Beweis. Wegen $E = (\delta_{ij})$ ist

$$d_n(E) = \sum_{j=1}^{n} (-1)^{i+j} \delta_{ij} d_{n-1}(E_{ij}) = (-1)^{2i} d_{n-1}(E_{ii}) \underset{\text{IV}}{=} 1.$$

□

Zum Beweis der Eindeutigkeit der Determinante wird benötigt:

Lemma 1.6. *Sei $d : {}^n\mathbb{R}^n \to \mathbb{R}$ multilinear und alternierend. Dann gilt für alle $a_1, \dots, a_n \in {}^n\mathbb{R}$ und $1 \leq i < j \leq n$:*

1. *Falls $a_i = a_j$ ist, so ist $d(a_1, \dots, a_n) = 0$.*
2. $d(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_n) = -d(a_1, \dots, a_n)$

Beweis.

(a) Zeige: Seien $1 \leq i < j \leq$ fest. Gilt dann 1. für alle $a_1, \dots, a_n$, so ist auch 2. für alle $a_1, \dots, a_n$ erfüllt.

Beweis. Für $a_1, \dots, a_n \in {}^n\mathbb{R}$ fest und $x, y \in {}^n\mathbb{R}$ sei

$$f(x, y) := d(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{j-1}, y, a_{j+1}, \dots, a_n).$$

Dann ist zu zeigen:

$$\forall\, x \in {}^n\mathbb{R} : f(x, x) = 0 \ \Rightarrow\ \forall\, x, y \in {}^n\mathbb{R} : f(y, x) = -f(x, y).$$

Gelte also $f(x, x) = 0$ für alle $x \in {}^n\mathbb{R}$. Dann gilt für alle $x, y \in {}^n\mathbb{R}$:

$$\begin{aligned} 0 = f(x+y, x+y) &\underset{d \text{ multilin.}}{=} f(x, x+y) + f(y, x+y) \\ &\underset{d \text{ multilin.}}{=} \underbrace{f(x, x)}_{=0} + f(x, y) + f(y, x) + \underbrace{f(y, y)}_{=0}, \end{aligned}$$

woraus die Behauptung folgt.

(b) Nun ist noch 1. zu zeigen.

Sind i und j benachbart, also $j = i + 1$, so gilt 1., da d alternierend ist. Dann gilt aber nach (a) auch 2., d.h., beim Vertauchen zweier benachbarter Spalten ändert sich das Vorzeichen von d.

Jetzt kann 1. für beliebige $1 \leq i < j \leq n$ gezeigt werden. Die Matrix

$$(a_1, \dots, a_i, \dots, a_j, \dots, a_n)$$

lässt sich durch Vertauschen aufeinanderfolgender Spalten (vertausche i-te und $i+1$-te Spalte, dann $i+1$-te und $i+2$-te Spalte, usw.) überführen in

$$(a_1, \dots, a_{j-1}, a_i, a_j, a_{j+1}, \dots, a_n).$$

Diese Matrix kann nun wiederum durch Vertauschen aufeinanderfolgender Spalten (vertausche j-te und $j-1$-te Spalte, dann $j-1$-te und $j-2$-te Spalte, usw.) überführt werden in

$$(a_1, \ldots, a_{i-1}, a_j, a_{i+1}, \ldots, a_{j-1}, a_i, a_j, a_{j+1}, \ldots, a_n).$$

Hierbei ist eine Vertauschung mehr notwendig als bei der ersten Umformung.

Insgesamt wird also eine ungerade Anzahl von Vertauschungen durchgeführt. Nach der Vorüberlegung ändert sich bei jeder dieser Vertauschungen das Vorzeichen von d. Hieraus folgt die Behauptung. □

Definition 1.7. Sei $\sigma \in S_n$. Schreibe $\sigma = \tau_1 \cdots \tau_m$ mit $m \geq 0$ und Transpositionen $\tau_1, \ldots, \tau_m$, was nach Satz 3.8 aus Kapitel II möglich ist. Dann heißt $\operatorname{sign}(\sigma) := (-1)^m$ **Vorzeichen** (signum) von σ.

Satz 1.8. $\operatorname{sign} : S_n \to \{\pm 1\} \subset \mathbb{R}$, $\sigma \mapsto \operatorname{sign}(\sigma)$ *ist wohldefinierter Gruppenhomomorphismus.*

Beweis.

1. Zeige: $\operatorname{sign}(\sigma)$ ist wohldefiniert, d.h. unabhängig von der Wahl der Darstellung

$$\sigma = \tau_1 \cdots \tau_m$$

in der Definition 1.7. (Diese Darstellung ist **nicht** eindeutig.)

Beweis. Sei $d : {}^n\mathbb{R}^n \to \mathbb{R}$ eine Determinante. (d existiert nach Satz 1.5.) Ist $\sigma = \tau_1 \cdots \tau_m$ wie in der Definition, so gilt für alle $a_1, \ldots, a_n \in {}^n\mathbb{R}$:

$$\begin{aligned} d(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) &= d(a_{\tau_1(\tau_2 \ldots \tau_m(1))}, \ldots, a_{\tau_1(\tau_2 \ldots \tau_m(n))}) \\ &\underset{1.6}{=} (-1) d(a_{\tau_2 \ldots \tau_m(1)}, \ldots, a_{\tau_2 \ldots \tau_m(n)}) \\ &= \ldots = (-1)^m d(a_1, \ldots, a_n). \end{aligned}$$

Insbesondere ist damit

$$d(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = (-1)^m \underbrace{d(e_1, \ldots, e_n)}_{=d(E)=1} = (-1)^m$$

unabhängig von der Wahl der Darstellung $\sigma = \tau_1 \cdots \tau_m$.

2. Zeige: sign ist Homomorphismus.

Beweis. Seien $\sigma, \rho \in S_n$. Schreibe $\sigma = \tau_1 \cdots \tau_m$ und $\rho = \tau_1' \cdots, \tau_l'$ mit Transpositionen $\tau_1, \ldots, \tau_m, \tau_1', \ldots, \tau_l'$. Dann ist $\sigma\rho = \tau_1 \ldots \tau_m \tau_1' \ldots \tau_l'$, woraus

$$\operatorname{sign}(\sigma\rho) = (-1)^{m+l} = (-1)^m(-1)^l = \operatorname{sign}(\sigma)\operatorname{sign}(\rho)$$

folgt. □

Definition 1.9. Sei $A = (a_{ij}) \in {}^n\mathbb{R}^n$. Dann heißt

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \ldots a_{\sigma(n)n}$$

Determinante von A.

Satz 1.10.

1. *Ist* $d : {}^n\mathbb{R}^n \to \mathbb{R}$ *multilinear und alternierend, so gilt für alle* $A \in {}^n\mathbb{R}^n$*:*
$$d(A) = \det(A) \cdot d(E).$$
2. *Es gibt genau eine Determinante* ${}^n\mathbb{R}^n \to \mathbb{R}$*, nämlich* det.

Beweis.

1. Sei $A = (a_{ij}) \in {}^n\mathbb{R}^n$ mit den Spalten $a_1, \ldots, a_n$. Dann ist $a_j = \sum_{i=1}^n a_{ij} e_i$, wobei $e_1, \ldots, e_n$ die Standardbasis des ${}^n\mathbb{R}$ ist. Damit gilt:
$$\begin{aligned} d(A) &= d(a_1, \ldots, a_n) \\ &= d\left(\sum_{i_1=1}^n a_{i_1 1} e_{i_1}, \sum_{i_2=1}^n a_{i_2 2} e_{i_2}, \ldots, \sum_{i_n=1}^n a_{i_n n} e_{i_n}\right) \\ &\underset{d \text{ multilin.}}{=} \sum_{1 \le i_1, \ldots, i_n \le n} a_{i_1 1} \ldots a_{i_n n} \underbrace{d(e_{i_1}, \ldots, e_{i_n})}_{\underset{1.6}{=} 0, \text{ falls } \exists\, k \neq l:\ i_k = i_l} \\ &= \sum_{\sigma \in S_n} a_{\sigma(1)1} \ldots a_{\sigma(n)n} \underbrace{d(e_{\sigma(1)}, \ldots, e_{\sigma(n)})}_{\underset{\text{Bew. } 1.8}{=} \operatorname{sign}(\sigma) d(e_1, \ldots, e_n)} \\ &= \det(A) d(E). \end{aligned}$$
2. Folgt aus Satz 1.5 und 1. □

Folgerung 1.11. *Sei* $n \geq 1$*, und sei* $A = (a_{ij}) \in {}^n\mathbb{R}^n$*. Dann gilt:*

1. **Leibniz-Formel***:*
$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \det(A^T). \end{aligned}$$
2. **Laplace'scher Entwicklungssatz***:*
   (a) *Entwicklung nach der i-ten Zeile für* $1 \leq i \leq n$*:*
$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$
   (b) *Entwicklung nach der j-ten Spalte für* $1 \leq j \leq n$*:*
$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$
3. $\det(A)$ *ist multilinear als Funktion der Spalten oder Zeilen von* A*.*
4. $\det(A) = 0$*, falls* A *zwei gleiche Spalten oder Zeilen hat.*
5. *Für Spalten (ebenso Zeilen)* $a_1, \ldots, a_n$ *von* A *und* $\sigma \in S_n$ *gilt:*
$$\det(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = \operatorname{sign}(\sigma) \det(A).$$

Beweis.

1. Es gilt

$$\begin{aligned}
\det(A) &\underset{\text{Def.}}{=} \sum_{\sigma\in S_n} \underbrace{\operatorname{sign}(\sigma)}_{\substack{=\operatorname{sign}(\sigma^{-1}),\\ \text{da } \operatorname{sign}(\sigma^{-1})=\operatorname{sign}(\sigma)^{-1}}} \underbrace{a_{\sigma(1)1}\cdots a_{\sigma(n)n}}_{\substack{=a_{1\sigma^{-1}(1)}\cdots a_{n\sigma^{-1}(n)},\\ \text{da } (\sigma(i),i)=(j,\sigma^{-1}(j))\\ \text{mit } j=\sigma(i)}} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma^{-1}) a_{1\sigma^{-1}(1)}\cdots a_{n\sigma^{-1}(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}.
\end{aligned}$$

Hierbei wurde bei der letzten Gleichheit verwendet, dass $S_n \to S_n$, $\sigma \mapsto \sigma^{-1}$ bijektiv ist, mit Umkehrabbildung $\sigma \mapsto \sigma^{-1}$ (wie in jeder Gruppe).

2. (a) Gilt nach Satz 1.5 wegen der Eindeutigkeit der Determinante.
   (b) Sei $B = A^T = (b_{ij})$. Nach 1. gilt dann für alle $1 \leq j \leq n$:

$$\det(A) = \det(B) \underset{(a)}{=} \sum_{k=1}^{n} (-1)^{j+k} \underbrace{b_{jk}}_{=a_{kj}} \underbrace{\det\,\underbrace{(B_{jk})}_{=(A_{kj})^T}}_{=\det(A_{kj}) \text{ wegen } 1.}.$$

3., 4., 5. Klar für Spalten.
Wegen 1. gelten die entsprechenden Aussagen dann aber auch für Zeilen. □

Folgerung 1.12 (Multiplikationssatz)**.** *Seien $A, B \in {}^n\mathbb{R}^n$. Dann gilt*

$$\det(AB) = \det(A)\det(B).$$

Beweis. Sei A fest. Betrachte $d(B) := \det(AB)$ als Funktion der Spalten $b_1, \ldots, b_n$ von B. Wegen $d(B) = \det(Ab_1, \ldots, Ab_n)$ und der Definition der Determinante ist d multilinear und alternierend. Somit gilt nach Satz 1.10

$$\det(AB) = d(B) = \det(B)\, d(E) = \det(B)\det(A).$$

□

Folgerung 1.13 (Kästchenregel)**.** *Sei*

$$A = \begin{pmatrix} A_1 & * & * & * \\ & A_2 & * & * \\ & & \ddots & * \\ 0 & & & A_k \end{pmatrix} \in {}^n\mathbb{R}^n$$

mit quadratischen Matrizen $A_1, \ldots, A_k$. Dann gilt

$$\det(A) = \det(A_1)\det(A_2)\cdots\det(A_k).$$

Beweis. Induktion nach k:

1. $k = 1$: Klar.
2. $k = 2$: Betrachte
$$d(A_1) = \det \begin{pmatrix} A_1 & * \\ & \\ 0 & A_2 \end{pmatrix}$$
als Funktion der Spalten von A_1. Offenbar ist d multilinear und alternierend. Also gilt nach Satz 1.10
$$\det \begin{pmatrix} A_1 & * \\ & \\ 0 & A_2 \end{pmatrix} = d(A_1) = \det(A_1)\, d(E).$$
Dann folgt die Behauptung aus
$$d(E) = \det \begin{pmatrix} 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \\ 0 & \cdots & 0 & A_2 \end{pmatrix} = \det(A_2).$$
Hierbei ergibt sich die letzte Gleichheit durch wiederholtes Entwickeln nach der 1. Spalte.
3. Induktionsschritt von $k-1$ auf k für $k \geq 3$:
$$\det \begin{pmatrix} A_1 & & * & \\ & A_2 & & * \\ 0 & & \ddots & \\ & 0 & & A_k \end{pmatrix} \underset{\substack{\text{Fall}\\ k=2}}{=} \det(A_1) \underbrace{\det \begin{pmatrix} A_2 & & * \\ & \ddots & \\ 0 & & A_k \end{pmatrix}}_{\underset{\text{IV}}{=} \det(A_2)\cdots\det(A_k)}$$

□

Beispiel 1.14. Sei
$$A = \begin{pmatrix} 1 & 0 & 11 & 12 \\ -5 & 2 & 13 & 14 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$
Dann gilt
$$\det(A) = \det \begin{pmatrix} 1 & 0 \\ -5 & 2 \end{pmatrix} \det \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} = 2 \cdot (-3) = -6.$$

Definition 1.15. Sei $A = (a_{ij}) \in {}^n\mathbb{R}^n$. Dann heißt $\tilde{A} = (\tilde{a}_{ij}) \in {}^n\mathbb{R}^n$ mit
$$\tilde{a}_{ij} := (-1)^{i+j} \det(A_{ji})$$
für alle $1 \leq i, j \leq n$ **komplementäre Matrix** (oder **Adjunkte**) zu A. (Hierbei beachte man die Reihenfolge der Indices!)

Beispiel 1.16. Sei
$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in {}^2\mathbb{R}^2.$$

Wegen

$$\begin{aligned} A_{11} &= (\delta), \quad A_{12} = (\gamma), \\ A_{21} &= (\beta), \quad A_{22} = (\alpha) \end{aligned}$$

gilt dann

$$\tilde{A} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$$

wie in Aufgabe 3 von Blatt 3.

Satz 1.17. *Sei $A \in {}^n\mathbb{R}^n$. Dann gilt $\tilde{A}A = A\tilde{A} = \det(A)E$.*

Beweis. Es wird $A\tilde{A} = \det(A)E$, d.h.

$$\sum_{j=1}^{n} a_{ij}\tilde{a}_{jk} = \delta_{ik}\det(A)$$

für alle $1 \leq i, k \leq n$ gezeigt. Ebenso ergibt sich $\tilde{A}A = \det(A)E$.

1. Fall: $i = k$. Dann gilt

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

nach dem Entwicklungssatz.

2. Fall: $i \neq k$. Sei $B = (b_{rs}) \in {}^n\mathbb{R}^n$ diejenige Matrix, die aus A entsteht, indem die k-te Zeile von A durch die i-te Zeile von A ersetzt wird, also mit den Zeilen $a_1, \ldots, a_n$ von A:

$$B := \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_{k-1} \\ a_i \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix}.$$

Dann gilt $b_{kj} = a_{ij}$ und $B_{kj} = A_{kj}$ für alle $1 \leq j \leq n$. Damit ist

$$\begin{aligned} \sum_{j=1}^{n} a_{ij}\tilde{a}_{jk} &= \sum_{j=1}^{n} b_{kj}(-1)^{j+k}\det(A_{kj}) \\ &= \sum_{j=1}^{n}(-1)^{j+k} b_{kj}\det(B_{kj}) \\ &= \det(B) = 0. \end{aligned}$$

Hierbei gilt die vorletzte Gleichheit nach dem Entwicklungssatz und die letzte, da B zwei gleiche Zeilen hat. □

Folgerung 1.18. *Sei $A \in {}^n\mathbb{R}^n$. Äquivalent sind:*

1. *A ist invertierbar.*
2. $\det(A) \neq 0$.

In diesem Fall ist $A^{-1} = \frac{1}{\det(A)}\tilde{A}$.

Beweis.

1. $\Rightarrow$ 2. Aus der Voraussetzung $AA^{-1} = E$ folgt mit dem Multiplikationssatz 1.12:
$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(E) = 1.$$
Also ist $\det(A) \neq 0$.

2. $\Rightarrow$ 1. Setze $A' := \frac{1}{\det(A)}\tilde{A}$. Dies ist wegen $\det(A) \neq 0$ möglich. Nach Satz 1.17 gilt dann
$$AA' = \frac{1}{\det(A)} \underbrace{A\tilde{A}}_{=\det(A)E} = E.$$
Somit ist A invertierbar, und das Inverse ist durch die angegebene Formel gegeben. □

Definition 1.19. Sei $n \geq 1$.

1. $\mathrm{SL}_n(\mathbb{R}) := \{A \in {}^n\mathbb{R}^n \mid A \text{ invertierbar}, \det(A) = 1\}$ heißt **spezielle lineare Gruppe**.
2. $\mathrm{SO}_n := \{A \in {}^n\mathbb{R}^n \mid A \text{ orthogonal}, \det(A) = 1\}$ heißt **spezielle orthogonale Gruppe**.

Folgerung 1.20.

1. $\mathrm{SL}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$ *und* $\mathrm{SO}_n \subset \mathrm{O}_n$ *sind Untergruppen.*
2. *Für alle* $A \in \mathrm{O}_n$ *gilt* $\det(A) = 1$ *oder* $\det(A) = -1$.

Beweis.

1. Folgt entweder aus dem Multiplikationssatz 1.12 oder da nach Definition gilt:
$$\mathrm{SL}_n(\mathbb{R}) = \mathrm{Ke}(\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R} \setminus \{0\}),$$
wobei $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R} \setminus \{0\}$ wegen 1.12 ein Homomorphismus von multiplikativen Gruppen ist.
Ebenso ist $\mathrm{SO}_n = \mathrm{Ke}(\det : \mathrm{O}_n \to \mathbb{R} \setminus \{0\})$.
2. Sei A orthogonal, also $AA^T = E$. Dann gilt
$$\det(A)\det(A^T) \underset{1.12}{=} \det(AA^T) = \det(E) = 1.$$
Außerdem ist nach 1.11: $\det(A^T) = \det(A)$. Somit gilt insgesamt $\det(A)^2 = 1$, woraus $\det(A) = \pm 1$ folgt, da $\det(A) \in \mathbb{R}$. □

Definition 1.21.

1. Sei $A \in \mathrm{O}_n$, also $f_A : {}^n\mathbb{R} \to {}^n\mathbb{R}$ orthogonale Abbildung.
   (a) f_A heißt **eigentlich orthogonal** oder **Drehung** $:\Leftrightarrow \det(A) = 1$.
   (b) f_A heißt **uneigentlich orthogonal** oder **Spiegelung** $:\Leftrightarrow \det(A) = -1$.
2. Sei $A \in {}^n\mathbb{R}^n$, also $f_A : {}^n\mathbb{R} \to {}^n\mathbb{R}$.
   (a) f_A heißt **orientierungserhaltend** $:\Leftrightarrow \det(A) > 0$.
   (b) f_A **orientierungsumkehrend** $:\Leftrightarrow \det(A) < 0$.

Bemerkungen 1.22.

1. Im Fall $n = 2$ sind Drehungen bzw. Spiegelungen die ebenen Drehungen bzw. Spiegelungen an Geraden durch den Nullpunkt wie in Teil II.
2. Nach Folgerung aus der QR-Zerlegung gilt: Jede Matrix $A \in \mathrm{SO}_n$ ist Produkt von $\frac{n(n-1)}{2}$ ebenen Drehmatrizen.

Bemerkungen 1.23 (Berechnung von Determinanten)**.**

1. Wende Entwicklungssatz nach passender Zeile oder Spalte an, falls mehrere Koeffizienten 0 sind.
2. Forme A durch elementare Zeilenumformungen um in eine Stufen-Matrix
$$B = \begin{pmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$
Dann ist $\det(B) = \alpha_1 \cdots \alpha_n$ nach der Kästchenregel 1.13, und $\det(A)$ lässt sich leicht aus $\det(B)$ berechnen, denn:
   (a) Bei Multiplikation einer Zeile mit $\alpha \in \mathbb{R}$ wird det mit α multipliziert.
   (b) Beim Vertauschen von Zeilen ändert det das Vorzeichen.
   (c) Beim Addieren eines Vielfachen einer Zeile zu einer anderen ändert sich det nicht, z.B.:
$$\det \begin{pmatrix} a_1 \\ a_2 + \alpha a_1 \\ a_3 \\ \vdots \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} + \underbrace{\alpha \det \begin{pmatrix} a_1 \\ a_1 \\ a_3 \\ \vdots \end{pmatrix}}_{=0},$$
da det multilinear und alternierend, wobei $a_1, a_2, \ldots, a_n$ die Zeilen von A sind und $\alpha \in \mathbb{R}$.

Index