Kapitel 1 Matrizen

§ 1 Der Vektorbegriff

Erkläre zunächst, was eine **reelle Zahl** ist (mit Dezimalbrüchen).

Eine reelle Zahl kann man sich als Punkt auf einer Geraden vorstellen. Ein **Paar** (x, y) von reellen Zahlen beschreibt einen Punkt in der Ebene, ein **Tripel** (x, y, z) einen Punkt im Raum.

Man kann aber auch z.B. den Einkaufspreis x und den Verkaufspreis y eines Handelsgutes zu einem Paar (x, y), oder die Verkaufszahlen x, y, z dreier Filialen eines Geschäftes zu einem Tripel (x, y, z) zusammenfassen. Befreit man sich also von der Vorstellung, dass die Koordinaten x, y, z unbedingt etwas mit dem Raum zu tun haben müssen, so kann man auch problemlos n Zahlen zu einem n-**Tupel**

$$\mathbf{x} = (x_1, \ldots, x_n)$$

zusammenfassen.

Um z.B. Aussagen über die Gesamtheit aller n-Tupel machen zu können, braucht man ein Verfahren, mathematische Objekte zu neuen Objekten zusammenzufassen.

Begriff der **Menge**, die Beziehung $x \in M$ („x ist *Element* von M“).
Beispiele: $\{1, 2, 3\}$, $\{1, 1, 2, 3, 2, 2, 3\}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Begriff der **Teilmenge**:

$$N \subset M(\text{„}N \text{ Teilmenge von } M\text{“}) :\iff \forall\, x : x \in N \implies x \in M.$$

Die benutzten logischen Symbole werden erklärt.

Die Menge

$$\mathbf{A}^n := \{(x_1, \ldots, x_n) : x_1, \ldots, x_n \in \mathbb{R}\}$$

aller n-Tupel von reellen Zahlen nennen wir den **n-dimensionalen (reellen) affinen Raum**, die Elemente $\mathbf{x} \in \mathbf{A}^n$ nennen wir **Punkte**.

Sind $\mathbf{p}$ und $\mathbf{q}$ zwei Punkte im $\mathbf{A}^n$, so bezeichnen wir das Paar $(\mathbf{p}, \mathbf{q})$ als einen **Pfeil**. Ein solcher Pfeil ist durch seinen Anfangspunkt, seine Richtung und seine Länge festgelegt. In der Physik braucht man Größen, die durch Richtung und Länge bestimmt sind, etwa um Kräfte zu beschreiben. Allerdings müssen diese Größen i.a. frei verschiebbar sein. Dabei soll der Anfangspunkt beliebig gewählt werden können, aber Länge und Richtung sollen gleich bleiben. Unter einem **Vektor** in $\mathbf{A}^n$ verstehen wir deshalb die Klasse (Menge) aller Pfeile, die aus einem festen Pfeil durch beliebige Parallelverschiebung entstehen. Jeder einzelne Pfeil aus der Klasse legt die gesamte Klasse fest und wird deshalb auch als ein **Repräsentant** des Vektors bezeichnet.

Wir bezeichnen hier einen Vektor mit einem Buchstaben und einem Pfeil darüber. Ist der Pfeil $(\mathbf{p}, \mathbf{q})$ ein Repräsentant des Vektors $\vec{v}$, so schreiben wir auch $\vec{v} = \overrightarrow{\mathbf{pq}}$. Ist $\mathbf{0} = (0, \ldots, 0)$ der **Nullpunkt** in $\mathbf{A}^n$, so gibt es genau einen Repräsentanten der Form $(\mathbf{0}, \mathbf{x})$ von $\vec{v}$. Diesen speziellen Pfeil nennt man auch den **Ortsvektor** des Punktes $\mathbf{x}$.

Ist $\mathbf{p} \in \mathbf{A}^n$ ein Punkt und $\vec{v}$ ein Vektor mit dem Repräsentanten $(\mathbf{p}, \mathbf{q})$, so schreibt man:

$$\mathbf{p} + \vec{v} = \mathbf{q} \quad \text{oder } \vec{v} = \mathbf{q} - \mathbf{p}.$$

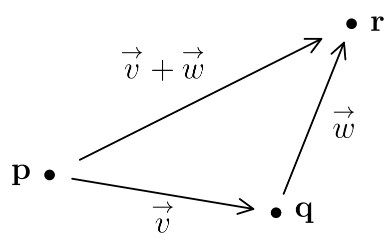
Das bedeutet, dass der Vektor $\vec{v}$ den Punkt $\mathbf{p}$ nach $\mathbf{q}$ verschiebt. Zu je zwei vorgegebenen Punkten $\mathbf{p}$ und $\mathbf{q}$ gibt es genau einen Vektor $\vec{v}$, der $\mathbf{p}$ nach $\mathbf{q}$ verschiebt.

Das Besondere an den Vektoren ist, dass man sie addieren kann: Sind zwei Vektoren $\vec{v}$ und $\vec{w}$ gegeben, so gibt es einen eindeutig bestimmten Vektor $\vec{v} + \vec{w}$, so dass für jeden Punkt $\mathbf{p}$ gilt:

$$(\mathbf{p} + \vec{v}) + \vec{w} = \mathbf{p} + (\vec{v} + \vec{w}).$$

Verwendet man geeignete Repräsentanten, so sieht das folgendermaßen aus:

$$\overrightarrow{\mathbf{pq}} + \overrightarrow{\mathbf{qr}} = \overrightarrow{\mathbf{pr}}.$$



Verschiebt man den Punkt $\mathbf{p}$ zunächst nach $\mathbf{q}$ und dann nach $\mathbf{r}$, so kann man genauso gut gleich $\mathbf{p}$ nach $\mathbf{r}$ verschieben. Die Summe von Vektoren bildet man also, indem man repräsentierende Pfeile aneinander hängt.

Man überzeugt sich leicht davon, dass die folgenden Regeln gelten:

$$\begin{aligned} (\vec{u} + \vec{v}) + \vec{w} &= \vec{u} + (\vec{v} + \vec{w}) \quad \text{(verwende Pfeile)}, \\ \vec{v} + \vec{w} &= \vec{w} + \vec{v} \quad \text{(Kräfteparallelogramm)}. \end{aligned}$$

Eine Ausnahme bildet der *Nullvektor* $\vec{o}$, der zu überhaupt keiner Verschiebung führt. Repräsentanten sind die Paare $(\mathbf{p}, \mathbf{p})$, die schwerlich als Pfeile zu deuten sind, denn ihnen fehlt Richtung und Länge. Trotzdem spricht man von einem Vektor. Es gilt:

$$\vec{v} + \vec{o} = \vec{v}, \text{ für alle Vektoren } \vec{v}.$$

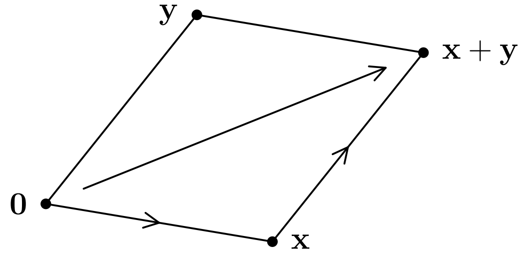
Außerdem gibt es zu jedem Vektor $\vec{v}$ einen Vektor $-\vec{v}$, den *entgegengesetzten oder negativen Vektor*, so dass $\vec{v} + (-\vec{v}) = \vec{o}$ ist. Wenn $(\mathbf{p}, \mathbf{q})$ ein Repräsentant von $\vec{v}$ ist, dann ist $(\mathbf{q}, \mathbf{p})$ ein Repräsentant von $-\vec{v}$.

Da ein Vektor $\vec{x}$ immer in der Form $\vec{x} = \overrightarrow{\mathbf{0x}}$ mit einem geeigneten Endpunkt $\mathbf{x}$ geschrieben werden kann, ist er durch den Ortsvektor $(\mathbf{0}, \mathbf{x})$ und damit durch dessen Endpunkt $\mathbf{x} = (x_1, \ldots, x_n)$ festgelegt. Deshalb schreibt man auch

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

Dies ist die *Spaltenschreibweise* für einen Vektor. Wie man sieht, besteht in Wirklichkeit gar kein großer Unterschied zwischen einem Punkt $\mathbf{x}$, dem Ortsvektor $(\mathbf{0}, \mathbf{x})$ und dem zugehörigen Vektor $\overrightarrow{x} = \overrightarrow{\mathbf{0x}}$.

Sind $\mathbf{x}$ und $\mathbf{y}$ zwei Punkte im $\mathbf{A}^n$, so bezeichnen wir den Endpunkt des Ortsvektors von $\overrightarrow{\mathbf{0x}} + \overrightarrow{\mathbf{0y}}$ mit $\mathbf{x} + \mathbf{y}$. Dann bilden $\mathbf{0}$, $\mathbf{x}$, $\mathbf{y}$ und $\mathbf{x} + \mathbf{y}$ die Ecken eines Parallelogramms.



Ist $\mathbf{x} = (x_1, \ldots, x_n)$ und $\mathbf{y} = (y_1, \ldots, y_n)$, so folgt aus der Konstruktion:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n).$$

Deshalb kann auch die Vektoraddition algebraisch beschrieben werden:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Ist r eine reelle Zahl, so definiert man

$$r \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}.$$

Das bedeutet, dass alle zugehörigen Pfeile um den Faktor r gestreckt werden.

Auch hier gelten diverse Rechenregeln (z.B. $\alpha(\overrightarrow{v} + \overrightarrow{w}) = \alpha\overrightarrow{v} + \alpha\overrightarrow{w}$), die aber nicht alle aufgezählt werden.

Da man mit Paaren, Tripeln, n-Tupeln, Ortsvektoren und freien Vektoren immer auf die gleiche Weise rechnen kann, drängt sich einem der Eindruck auf, dass es bei der Vektorrechnung gar nicht auf die Pfeile oder die Vorstellung davon ankommt. Tatsächlich sind Vektoren einfach Objekte, die ganz bestimmten Rechenregeln genügen, und nur diese Regeln stellen die Verbindung zur anschaulich-physikalischen Vektor-Vorstellung her.

Das führt zu folgender

Definition. Unter einem *reellen Vektorraum* verstehen wir eine Menge V, deren Elemente addiert und mit reellen Zahlen multipliziert werden können, so dass für $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ und $\alpha, \beta \in \mathbb{R}$ folgende Rechenregeln erfüllt sind:

$$
\begin{aligned}
\mathbf{x} + (\mathbf{y} + \mathbf{z}) &= (\mathbf{x} + \mathbf{y}) + \mathbf{z}, \\
\mathbf{x} + \mathbf{y} &= \mathbf{y} + \mathbf{x}, \\
\exists\, \mathbf{0} \in V \,:\, \mathbf{x} + \mathbf{0} &= \mathbf{x}, \\
\forall\, \mathbf{x} \in V\, \exists -\mathbf{x} \in V \,:\, \mathbf{x} + (-\mathbf{x}) &= \mathbf{0}, \\
(\alpha + \beta)\mathbf{x} &= \alpha\mathbf{x} + \beta\mathbf{x}, \\
\alpha(\mathbf{x} + \mathbf{y}) &= \alpha\mathbf{x} + \alpha\mathbf{y}, \\
\alpha(\beta\mathbf{x}) &= (\alpha\beta)\mathbf{x} \\
\text{und } 1 \cdot \mathbf{x} &= \mathbf{x}.
\end{aligned}
$$

Die Elemente von V bezeichnet man als *Vektoren*.
Wichtige Beispiele sind der Vektorraum der Spaltenvektoren und der Vektorraum der n-Tupel.

§ 2 Matrizen und lineare Gleichungen

Matrix, rechteckiges Zahlenschema

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

a_{ij} (oder auch A_{ij}) *Koeffizienten* oder *Einträge*

$$Z_i(A) = (a_{i1}, \ldots, a_{in}) = \text{ i-te Zeile} \quad \text{(Zeilenvektor)}$$

$$S_j(A) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \text{ j-te Spalte} \quad \text{(Spaltenvektor)}$$

$M_{m,n}(\mathbb{R})$ = Menge der $m \times n$-Matrizen

$M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$ = Menge der n-reihigen quadratischen Matrizen

$$\text{speziell} \quad 0 = 0_{mn} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad \text{Nullmatrix,}$$

$$E_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \quad \text{Einheitsmatrix.}$$

Addition von Matrizen:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Multiplikation einer Matrix mit einer reellen Zahl:

$$\alpha \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} := \begin{pmatrix} \alpha \cdot a_{11} & \cdots & \alpha \cdot a_{1n} \\ \vdots & & \vdots \\ \alpha \cdot a_{m1} & \cdots & \alpha \cdot a_{mn} \end{pmatrix}.$$

Dann sind folgende Rechenregeln erfüllt:

$$A + B = B + A, \quad (A + B) + C = A + (B + C) \text{ und } A + 0 = 0 + A = A,$$

sowie

$$1 \cdot A = A, \quad \alpha(\beta A) = (\alpha\beta)A, \quad \alpha(A+B) = \alpha A + \beta B \text{ und } (\alpha+\beta)A = \alpha A + \beta A.$$

Setzt man $-A := (-1)A$, so gilt außerdem: $A + (-A) = 0$.

Damit bilden die Elemente von $M_{m,n}(\mathbb{R})$ einen reellen Vektorraum.

Linearkombination von $A_1, \dots, A_k$ ist $\sum_{i=1}^k \alpha_i A_i = \alpha_1 A_1 + \cdots + \alpha_k A_k$ $(\alpha_i \in \mathbb{R})$

B linear abhängig von $A_1, \dots, A_k$ $:\Longleftrightarrow$

$$B = \alpha_1 A_1 + \cdots + \alpha_k A_k \text{ (mit geeigneten } \alpha_i).$$

Ein System $\mathcal{S} = \{A_1, \dots, A_m\}$ heißt linear abhängig, falls eine der Matrizen A_i von den anderen linear abhängt. Auch $\{0\}$ wird linear abhängig genannt. Ein System heißt *linear unabhängig*, falls es nicht linear abhängig ist.

2.1 Satz. *Das System $\mathcal{S} = \{A_1, \dots, A_k\}$ habe folgende Eigenschaft:*

Zu jedem r gibt es eine Position (i, j), so dass $(A_r)_{ij} \neq 0$ und $(A_l)_{ij} = 0$ für $l \neq r$ ist.

Dann ist $\mathcal{S}$ linear unabhängig.

BEWEIS: Annahme, es sei z.B. A_k Linearkombination von $A_1, \dots, A_{k-1}$. Wähle dann Position (i, j) zu k gemäß der Voraussetzung und stelle fest, dass das einen Widerspruch ergibt. ∎

Beispiel: E_{ij} habe den Eintrag 1 an der Position (i, j) und sonst überall den Eintrag 0. Dann ist das System der E_{ij} mit $i = 1, \dots, m$ und $j = 1, \dots, n$ linear unabhängig (in $M_{m,n}(\mathbb{R})$).

Der *Spann* $\langle A_1, \dots, A_k \rangle$ ist die Menge aller Linearkombinationen von $A_1, \dots, A_k$.

Eine Teilmenge U eines Vektorraumes V heißt *Unterraum* von V, falls gilt:

1. $0 \in U$.
2. $X, Y \in U \implies X + Y \in U$.
3. $\alpha \in \mathbb{R}$, $X \in U \implies \alpha X \in U$.

Es ist $\langle A_1, \dots, A_k \rangle$ ein Unterraum von $M_{m,n}(\mathbb{R})$. Jeder Unterraum ist ein Vektorraum.

Lineares Gleichungssystem:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Eine Lösung des LGS ist ein Spaltenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, dessen Komponenten x_i simultan alle Gleichungen erfüllen. Ist $A = (a_{ij}) \in M_{m,n}(\mathbb{R})$ die Matrix der

Koeffizienten des LGS und $b \in M_{m,1}(\mathbb{R})$ der Spaltenvektor, dessen Einträge b_i die rechten Seiten der Gleichungen bilden, so hat das LGS die Gestalt

$$x_1 \cdot S_1(A) + \cdots + x_n \cdot S_n(A) = b.$$

Führe nun Matrizenprodukt $A \cdot B$ ein, für $A \in M_{m,n}(\mathbb{R})$, $B \in M_{n,r}(\mathbb{R})$:

$$\boxed{(A \cdot B)_{ij} = A_{i1}B_{1j} + \cdots + A_{in}B_{nj}\,.}$$

Dann ist

$$\begin{pmatrix} Z_1(A) \\ \vdots \\ Z_m(A) \end{pmatrix} \cdot (S_1(B), \ldots, S_n(B)) = \begin{pmatrix} Z_1(A) \cdot S_1(B) & \cdots & Z_1(A) \cdot S_n(B) \\ \vdots & & \vdots \\ Z_m(A) \cdot S_1(B) & \cdots & Z_m(A) \cdot S_n(B) \end{pmatrix},$$

wobei $(a_{i1}, \ldots, a_{in}) \cdot \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$ ist.

Für die Matrizenmultiplikation gelten folgende Regeln:

$$\begin{aligned} A(BC) &= (AB)C, \\ A(B+C) &= AB + AC, \\ (A+B)C &= AC + BC, \end{aligned}$$

i.A. ist aber $AB \neq BA$, z.B. ist $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

Insbesondere ist $A \cdot e_j = S_j(A)$ für alle j und

$$A \cdot x = A \cdot (x_1e_1 + \cdots + x_ne_n) = x_1 \cdot S_1(A) + \cdots + x_n \cdot S_n(A).$$

Daraus folgt: Ist $A_1 \cdot x = A_2 \cdot x$ für alle Vektoren x, so ist $A_1 = A_2$.

Das LGS bekommt nun die Gestalt

$$A \cdot x = b.$$

Man nennt $A \cdot x = 0$ ein *homogenes System*, jedes andere ein *inhomogenes System*.

Ziel: Bestimme Lösungsmenge $\text{Lös}(A, b) = \{x \,:\, A \cdot x = b\}$.

2.2 Satz. $\text{Lös}(A, 0)$ *ist ein Unterraum von* $M_{n,1}(\mathbb{R}) = \mathbb{R}^n$.

Beweis: Nachrechnen! ∎

Speziell nennt man 0 die „triviale" Lösung. Eventuell ist das die einzige Lösung.

Die Lösungsmenge eines inhomogenen Systems kann leer sein (führe hier die *leere Menge* $\varnothing$ ein). Ist sie es nicht, so gilt:

2.3 Satz. *Ist x_0 eine spezielle Lösung des LGS $A \cdot x = b$, so ist*

$$\text{Lös}(A, b) = \{x \in \mathbb{R}^n \,:\, x = x_0 + y \textit{ mit } y \in \text{Lös}(A, 0)\}.$$

BEWEIS: Es ist $A \cdot x_0 = b$. Ist auch noch $A \cdot x = b$, so ist $A \cdot (x - x_0) = b - b = 0$, also $x = x_0 + y$ mit $y = x - x_0 \in \text{Lös}(A, 0)$.

Umgekehrt: Ist $x = x_0 + y$ mit einem $y \in \text{Lös}(A, 0)$, so ist $0 = A \cdot y = A \cdot (x - x_0) = A \cdot x - A \cdot x_0 = A \cdot x - b$, also $A \cdot x = b$. ∎

2.4 Folgerung. *$A \cdot x = b$ ist genau dann eindeutig lösbar, wenn $A \cdot x = 0$ nur die triviale Lösung besitzt. (Über die Existenz wird nichts ausgesagt).*

BEWEIS: Ist $\text{Lös}(A, b) = \varnothing$, so ist nichts zu zeigen. Andernfalls folgt die Behauptung aus dem obigen Satz. ∎

§ 3 Das Gaußverfahren

Beginne mit Beispiel:

$$\begin{aligned} x - 2y &= 1, \\ 3x + 2y &= 11. \end{aligned}$$

Eliminiere x aus der 2. Gleichung (Subtrahiere das 3-fache der 1. Gleichung von der 2. Gleichung) und bringe System auf Dreiecksgestalt:

$$\begin{aligned} x - 2y &= 1, \\ 8y &= 8. \end{aligned}$$

Berechne dann $y = 1$ aus der 2. Gleichung und gewinne x durch „Rückwärtseinsetzen“: $x - 2 = 1$, also $x = 3$.

Entwickle nun Gaußverfahren zur Lösung von $A \cdot x = b$. Erlaube spezielle Umformungen, die ein Gleichungssystem in ein „äquivalentes“ System (d.h. ein System mit gleicher Lösungsmenge) überführen:

A) Vertauschung zweier Gleichungen.

B) Addition eines Vielfachen einer Gleichung zu einer anderen Gleichung.

C) Multiplikation einer Gleichung mit einer Konstanten $\lambda \neq 0$.

D) Umbenennung der Variablen.

(A) lässt sich durch eine geschickte Kombination von Umformungen vom Typ (B) erreichen. Der Einfachheit halber lässt man aber (A) als eigenständige Umformung zu.

Auf den Typ (C) verzichten wir erst mal. Auf Typ (D) könnte man auch verzichten, aber dann lässt sich das Verfahren nur mühsam aufschreiben. Wir lassen eine Umbenennung der Variablen zu, allerdings müssen wir darüber Buch führen.

Jede Information über das Gleichungssystem ist in der erweiterten Koeffizientenmatrix enthalten:

$$A_e = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

(A) entspricht der Vertauschung zweier Zeilen von A_e.
(B) entspricht der Addition eines Vielfachen der i-ten Zeile zur k-ten Zeile.
(C) entspricht der Multiplikation einer Zeile mit $\lambda \neq 0$.
(D) entspricht der Vertauschung zweier Spalten von A (aber nicht von A_e). Das hat Einfluss auf den Lösungsraum.

Ziel ist es, für A folgende Dreiecksform zu erreichen:

$$A \to \left(\begin{array}{c|c} D_r & B \\ \hline 0 & 0 \end{array}\right), \text{ mit } D_r = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{rr} \end{pmatrix} \text{ und } a_{ii} \neq 0 \text{ für } i = 1, \ldots, r.$$

Das geht folgendermaßen: $A \in M_{m,n}(\mathbb{R})$ soll r*-speziell* genannt werden, falls gilt:

$$A = \left(\begin{array}{c|c} D_r & B \\ \hline 0 & C \end{array}\right), \text{ mit } D_r = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{rr} \end{pmatrix} \text{ und } a_{ii} \neq 0 \text{ für } i = 1, \ldots, r.$$

Ist $r = 0$, so ist $A = C$. Jede Matrix ist also zumindest 0-speziell. Ist $r = m \leq n$, so ist $A = (D_r|B)$. Ist $r = n \leq m$, so ist $A = \begin{pmatrix} D_r \\ 0 \end{pmatrix}$.

Nun wird folgender **Reduktionsschritt** ausgeführt:

Sei $A_k = \left(\begin{array}{c|c} D_k & B_k \\ \hline 0 & C_k \end{array}\right)$ eine k-spezielle Umformung von A. Ist $C_k = 0$, so hat man die gewünschte Dreiecksform erreicht und ist fertig. Ist $C_k \neq 0$, so gibt es einen Eintrag $a_{ij} \neq 0$ in C_k (ein sogenanntes *Pivot-Element*). Durch eine Reihe von Zeilen- und Spaltenvertauschungen verschiebt man a_{ij} an die Position $(k+1, k+1)$.

Anschließend subtrahiert man geeignete Vielfache der $(k+1)$-ten Zeile von den folgenden Zeilen und erzeugt so Nullen unterhalb der Position $(k+1, k+1)$. Das Ergebnis ist eine $(k+1)$-spezielle Matrix.

Nach endlich vielen Schritten erhält man eine Matrix A_r in der gewünschten Dreiecksgestalt, spätestens, wenn $r = m$ oder $r = n$ erreicht wird. Die Umformungen vom Typ (A) und (B) müssen an der erweiterten Matrix A_e vorgenommen werden. So erhält man ein Gleichungssystem

$$\begin{array}{ccccccccc} a_{11}x_1 & + \cdots + & a_{1r}x_r & + & a_{1,r+1}x_{r+1} & + \cdots + & a_{1,n}x_n & = & b_1 \\ & \ddots & & & & & & & \vdots \\ & & a_{rr}x_r & + & a_{r,r+1}x_{r+1} & + \cdots + & a_{r,n}x_n & = & b_r \end{array}$$

Hinzu kommen noch Gleichungen der Gestalt $0 = b_i$, für $i = r+1, \ldots, m$. Das ursprüngliche Gleichungssystem ist genau dann lösbar, wenn diese zusätzlichen Gleichungen alle erfüllt sind. Man beachte aber, dass die a_{ij} und die b_i nicht mehr die aus dem ursprünglichen System sind. Deshalb kann man Lösbarkeit erst testen, wenn man die Lösung schon fast vorliegen hat.

Endgültig gewinnt man die Lösung durch **Rückwärtseinsetzen:**

$$\begin{aligned} x_r &= \frac{1}{a_{rr}}\Big(b_r - a_{r,r+1}x_{r+1} - \ldots - a_{r,n}x_n\Big) \\ x_{r-1} &= \frac{1}{a_{r-1,r-1}}\Big(b_{r-1} - a_{r-1,r+1}x_{r+1} - \ldots - a_{r-1,n}x_n - a_{r-1,r}x_r\Big) \\ &\ \text{etc.} \end{aligned}$$

a) Eine spezielle Lösung des inhomogenen Systems:
Setzt man $x_{r+1} = \ldots = x_n = 0$ und berechnet dann $x_1, \ldots, x_r$ durch Rückwärtseinsetzen, so erhält man eine spezielle Lösung.

b) Die allgemeine Lösung des zugehörigen homogenen Systems $A \cdot x = 0$:
Fasst man $x_{r+1}, \ldots, x_n$ als freie Parameter auf und berechnet $x_1, \ldots, x_r$ in Abhängigkeit dieser Parameter durch Rückwärtseinsetzen, so erhält man alle Lösungen des homogenen Systems.

Es gibt **Varianten** des Gaußverfahrens:

1. Wenn man keine Spaltenvertauschungen zulassen will, erreicht man nur die **Zeilen-Stufen-Form**:

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

   Das Verfahren ist geringfügig einfacher, weil nur Zeilenoperationen vorkommen, aber dafür ist es unübersichtlicher und mühsamer auszuschreiben.

2. Erlaubt man neben Zeilenoperationen und Spaltenvertauschungen auch noch die Multiplikation einer Zeile mit einem Faktor $\lambda \neq 0$, so kann man das Gaußverfahren u.U. etwas abkürzen. Außerdem kann man erreichen, dass die Elemente a_{ii} alle $= 1$ werden. Das erleichtert das Rückwärts-Einsetzen.

3. Die **reduzierte Zeilen-Stufen-Form** erreicht man, indem man wie in (1) zunächst die Zeilen-Stufen-Form herstellt und dann mittels (2) dafür sorgt, dass die Pivot-Elemente (die vordersten Elemente der Stufen) alle $= 1$ werden und oberhalb eines Pivot-Elementes nur Nullen stehen:

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

4. Die allgemeine Lösung des inhomogenen Systems kann man auch direkt gewinnen, wie beim homogenen System. Allerdings werden wir es bei vielen Anwendungen nur mit homogenen Systemen zu tun haben.

Beispiel.

Betrachte das Gleichungssystem mit der erweiterten Matrix

$$A_e = \left(\begin{array}{rrrrr|r} 1 & 3 & 5 & 2 & 0 & 1 \\ 3 & 9 & 10 & 1 & 2 & 0 \\ 0 & 2 & 7 & 3 & -1 & 2 \\ 2 & 8 & 12 & 2 & 1 & 1 \end{array}\right) .$$

Das Gaußverfahren liefert die Matrix

$$\left(\begin{array}{rrrrr|r} 1 & 3 & 5 & 2 & 0 & 1 \\ 0 & 2 & 2 & -2 & 1 & -1 \\ 0 & 0 & -5 & -5 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) .$$

Mit $x_4 = x_5 = 0$ erhält man: $x_3 = 3/5$, $x_2 = -11/10$, $x_1 = 13/10$.

Zur Lösung des homogenen Systems (rechte Seite = Null): Wähle $x_4 = s$ und $x_5 = t$ als freie Parameter. Dann ist

$$x_3 = -s + \frac{2}{5}t, \quad x_2 = 2s - \frac{9}{10}t \quad \text{und} \quad x_1 = -3s + \frac{7}{10}t.$$

§ 4 Invertierbare Matrizen

Wir betrachten einige spezielle Matrizen:

$E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ ist die (n-reihige) **Einheitsmatrix**. Für $A \in M_n(\mathbb{R})$ ist stets $E_n \cdot A = A \cdot E_n = A$.

E_{ij} sei die Matrix, die an der Position (i, j) eine Eins und sonst lauter Nullen hat. Offensichtlich ist $E_n = E_{11} + \cdots + E_{nn}$. Die Matrix $E_{ij} \cdot A$ weist in der i-ten Zeile die gleichen Elemente wie die in der j-te Zeile von A auf, und sonst lauter Nullen. Die Matrix $A \cdot E_{ij}$ weist in der j-ten Spalte die gleichen Elemente wie die in der i-ten Spalte von A auf, und sonst lauter Nullen.

Nun sei $E_{ij}(\lambda) := E_n + \lambda E_{ji}$. Dann ist $E_{ij}(\lambda) \cdot A = A + \lambda(E_{ji} \cdot A)$ diejenige Matrix, die man aus A erhält, indem man das λ-Fache der i-ten Zeile zur j-ten Zeile addiert. Offensichtlich ist $E_{ij}(-\lambda) \cdot E_{ij}(\lambda) = E_n$ und $E_{ij}(\lambda) \cdot E_{ij}(-\lambda) = E_n$.

Definition: Eine Matrix $A \in M_n(\mathbb{R})$ heißt *invertierbar*, falls es eine Matrix A' mit $AA' = A'A = E_n$ gibt. Die Matrix A' nennt man die *inverse Matrix* zu A und bezeichnet sie mit A^{-1}.

Es gilt:

1. Die inverse Matrix ist (wenn sie existiert) eindeutig bestimmt.

   Sind nämlich A'_1, A'_2 zwei potentielle Inverse, so ist $A'_2 = A'_2 E_n = A'_2(AA'_1) = (A'_2 A)A'_1 = E_n A'_1 = A'_1$.

2. A, B invertierbar $\implies$ AB invertierbar und $(AB)^{-1} = B^{-1}A^{-1}$.

Beispiele.

1. Die „Eliminationsmatrix“ $E_{ij}(\lambda)$ ist invertierbar.

2. Die Einheitsmatrix ist invertierbar.

3. Sei $P_{ij} := E_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}$. Dann ist $A \cdot P_{ij}$ die Matrix, die man aus A erhält, wenn man die i-te und die j-te Spalte vertauscht. Offensichtlich ist $P_{ij} \cdot P_{ij} = E_n$. Man nennt P_{ij} eine „Permutationsmatrix“.

   Das Gaußverfahren bedeutet, dass man eine Matrix A von links mit einer Folge von Eliminationsmatrizen und von rechts mit einer Folge von Permutationsmatrizen multipliziert.

4. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Damit A invertierbar ist, muss es Elemente u, v, x, y geben, so dass $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} u & v \\ x & y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist. Das ergibt folgendes Gleichungssystem:

$$\begin{aligned} au + bx &= 1, \\ av + by &= 0, \\ cu + dx &= 0, \\ cv + dy &= 1. \end{aligned}$$

Subtrahiert man das d-fache der1. Gleichung vom b-Fachen der 3. Gleichung, so erhält man $(ad - bc)u = d$. Analog erhält man auch $(ad - bc)v = -b$.

1. Fall: $ad - bc = 0 \implies b = d = 0$; *dann* $au = 1$, $av = 0$ (also $u \neq 0$, $a \neq 0$ und $v = 0$), sowie $cu = 0$, $cv = 1$ (also $c \neq 0$, $v \neq 0$ und $u = 0$). Das ist ein Widerspruch. Es muss somit gelten:

2. Fall: $ad - bc \neq 0$. Dann ist $u = d/(ad - bc)$ und $v = -b/(ad - bc)$. Einsetzen liefert: $x = -c/(ad - bc)$ und $y = a/(ad - bc)$. Die Zahl $ad - bc$ nennt man die *Determinante* von A. Wir haben gezeigt:

$$A \text{ invertierbar} \quad \Longleftrightarrow \quad \det(A) \neq 0.$$

Wir werden später versuchen, einer beliebigen quadratischen Matrix A eine Determinante zuzuordnen, die darüber entscheidet, ob A invertierbar ist.

4.1 Satz. *Sei A invertierbar. Dann besitzt das Gleichungssystem $A \cdot x = 0$ nur die triviale Lösung $x = 0$, und jedes Gleichungssystem $A \cdot x = b$ ist eindeutig lösbar (durch $x = A^{-1}b$).*

BEWEIS: Ist $A \cdot x = 0$, so ist $x = (A^{-1}A) \cdot x = A^{-1}(Ax) = A^{-1} \cdot 0 = 0$. ∎

4.2 Satz. *Mit $\Delta = \Delta(d_1, \ldots, d_n)$ sei die Diagonalmatrix mit den Elementen $d_1, \ldots, d_n$ bezeichnet. Δ ist genau dann invertierbar, wenn alle $d_i \neq 0$ sind.*

BEWEIS: 1) Ist Δ invertierbar, so gibt es zu jedem j ein x mit $\Delta \cdot x = e_j$. Dann ist $d_j \cdot x_j = 1$, also $d_j \neq 0$.

2) Sind alle $d_j \neq 0$, so ist $\Delta(d_1, \ldots, d_n) \cdot \Delta(d_1^{-1}, \ldots, d_n^{-1}) = E_n$. ∎

4.3 Satz. *Ist $A \in M_n(\mathbb{R})$ invertierbar, so sind die Spalten $S_1(A), \ldots, S_n(A)$ linear unabhängig.*

BEWEIS: Es ist $Ae_j = S_j(A)$. Wäre z.B. Ae_1 Linearkombination der Spalten $Ae_2, \ldots, Ae_n$, so multipliziere mit A^{-1}. Dann ist e_1 Linearkombination von $e_2, \ldots, e_n$. WS! ∎

4.4 Satz. *a) Ist A invertierbar und entsteht A' aus A durch elementare Zeilenumformungen, so ist auch A' invertierbar.*

b) Ist A invertierbar, so kann A durch elementare Zeilenumformungen in die Einheitsmatrix $E = E_n$ überführt werden.

BEWEIS: a) Es gibt eine invertierbare Matrix Q (zusammengesetzt aus Eliminationsmatrizen), so dass $QA = A'$ ist. Dann ist natürlich auch A' invertierbar.

b) Da $S_1(A) \neq 0$ ist, kann man die erste Spalte in e_1 überführen. Das Ergebnis ist wieder invertierbar. Hat man nach endlich vielen Schritten erreicht, dass die ersten k Spalten gerade die ersten k Einheitsvektoren sind, so muss sich in der $(k+1)$-ten Spalte (wegen der linearen Unabhängigkeit) ein Element $a_{i,k+1} \neq 0$ mit $i > k$ finden (denn sonst wäre $S_{k+1}(A)$ Linearkombination von $S_1(A), \ldots, S_k(A)$). Aber dann kann man erreichen, dass $S_{k+1}(A) = e_{k+1}$ ist. ∎

4.5 Satz. *Sei $AB = C$. Entstehen A' aus A und C' aus C durch die gleichen elementaren Zeilenumformungen, so ist auch $A'B = C'$.*

Beweis: Es gibt eine invertierbare Matrix Q mit $QA = A'$ und $QC = C'$. Dann ist $A'B = (QA)B = Q(AB) = QC = C'$. ∎

4.6 Folgerung. *Ist A invertierbar und entstehen E aus A und E' aus E durch die gleichen elementaren Zeilenumformungen, so ist $E' = A^{-1}$.*

Beweis: Es ist $AA^{-1} = E$, also $A^{-1} = EA^{-1} = E'$. ∎

Die letzte Folgerung liefert ein **Verfahren, die inverse Matrix zu berechnen**.

Sei $A \in M_n(\mathbb{R})$ invertierbar. Wir bilden die erweiterte Matrix $(A|E_n)$ und formen diese mit elementaren Zeilenoperationen solange um, bis auf der linken Seite die Einheitsmatrix steht. Man braucht dabei keine Spaltenvertauschungen. Entsteht so aus $(A|E_n)$ die Matrix $(E_n|E')$, so ist $E' = A^{-1}$.

Hier kommt ein Zahlenbeispiel:

$$(A, E) = \begin{pmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 3/4 & 1/2 & 1/4 \\ 0 & 1 & 0 & 1/2 & 1 & 1/2 \\ 0 & 0 & 1 & 1/4 & 1/2 & 3/4 \end{pmatrix} = (E, A^{-1}).$$

Bemerkung. Ist $D \in M_n(\mathbb{R})$ eine obere Dreiecksmatrix, deren Diagonalelemente alle $\neq 0$ sind, so ist D invertierbar. Das folgt daraus, dass D mit elementaren Zeilenumformungen in die Einheitsmatrix überführt werden kann. Es gibt also eine invertierbare Matrix Q, so dass $QD = E$ ist. Dann ist $D = Q^{-1}$ invertierbar.

Jetzt kann man die Ergebnisse auf LGS anwenden. Startet man mit einem LGS $Ax = b$, so liefert das Gaussverfahren ein Gleichungssystem der Gestalt

$$\left(\begin{array}{c|c} D & B \\ \hline 0 & 0 \end{array}\right) \cdot \begin{pmatrix} x' \\ x'' \end{pmatrix} = \begin{pmatrix} b' \\ b'' \end{pmatrix}, \text{ mit invertierbarer Dreiecksmatrix } B.$$

Das führt zu $D \cdot x' + B \cdot x'' = b'$ und $0 = b''$. Die zweite Gleichung entscheidet über die Lösbarkeit des LGS. Ist sie erfüllt, so ist $x' = D^{-1}(b' - B \cdot x'')$.

Uns interessiert besonders das homogene System.

4.7 Satz. *Ist $D \in M_r(\mathbb{R})$ und*

$$(a_1, \ldots, a_{n-r}) = \begin{pmatrix} -D^{-1}B \\ E_{n-r} \end{pmatrix},$$

so lässt sich jede Lösung des homogenen LGS $\begin{pmatrix} D & B \\ 0 & 0 \end{pmatrix} \cdot x = 0$ *eindeutig als Linearkombination von* $a_1, \ldots, a_{n-r}$ *schreiben, und umgekehrt ist jede solche Linearkombination eine Lösung.*

Beweis: 1) Ist x Lösung, so ist $x = \begin{pmatrix} x' \\ x'' \end{pmatrix}$ mit $x' = -D^{-1}B \cdot x''$, also

$$x = \begin{pmatrix} -D^{-1}B \\ E_{n-r} \end{pmatrix} \cdot x'' = x_{r+1}a_1 + \cdots + x_n a_{n-r}.$$

2) Sei $t = \begin{pmatrix} t_{r+1} \\ \vdots \\ t_n \end{pmatrix}$ und

$$x = t_{r+1}a_1 + \cdots + t_n a_{n-r} = (a_1, \ldots, a_{n-r}) \cdot t = \begin{pmatrix} -D^{-1}B \\ E_{n-r} \end{pmatrix} \cdot t\,,$$

also

$$\begin{pmatrix} D & B \\ 0 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} D & B \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -D^{-1}B \\ E_{n-r} \end{pmatrix} \cdot t = \begin{pmatrix} -B + B \\ 0 \end{pmatrix} \cdot t = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Das bedeutet, dass x Lösung ist.

3) Zur Eindeutigkeit: Ist $x = t_{r+1}a_1 + \cdots + t_n a_{n-r} = s_{r+1}a_1 + \cdots + s_n a_{n-r}$, so ist

$$\begin{pmatrix} -D^{-1}Bt \\ t \end{pmatrix} = \begin{pmatrix} -D^{-1}B \\ E_{n-r} \end{pmatrix} \cdot t = \begin{pmatrix} -D^{-1}B \\ E_{n-r} \end{pmatrix} \cdot s = \begin{pmatrix} -D^{-1}Bs \\ s \end{pmatrix},$$

also $s = t$. ∎

Zum Schluss noch einige Bemerkungen über das Transponieren von Matrizen.

Ist $A \in M_{m,n}(\mathbb{R})$, so ist die *transponierte Matrix* die Matrix A^t, die durch $(A^t)_{ij} = A_{ji}$ definiert wird. Dann gilt:

$$\begin{aligned} (A^t)^t &= A, \\ (A + B)^t &= A^t + B^t, \\ (\alpha A)^t &= \alpha \cdot A^t, \\ (AB)^t &= B^t A^t. \end{aligned}$$

Ist $A \in M_n(\mathbb{R})$ invertierbar, so ist auch A^t invertierbar und

$$(A^t)^{-1} = (A^{-1})^t.$$

Ist x ein Spaltenvektor, so ist x^t der entsprechende Zeilenvektor (und umgekehrt).

Kapitel 2 Vektorräume

§ 1 Strukturen

Ergänzungen zur Mengenlehre:

- kartesisches Produkt von Mengen: $A \times B = \{(x, y) : x \in A \text{ und } y \in B\}$.
  Dies sind i.a. Paare von irgendwelchen Objekten.

- Begriff der Abbildung: A, B seien Mengen. Eine **Abbildung** von A nach B ist eine Zuordnung $f : A \to B$, die *jedem* $x \in A$ *genau ein* $y \in B$ zuordnet (man beachte den Unterschied zwischen „ein" und „genau ein"). Die Menge A nennt man den Definitionsbereich der Abbildung, B den Wertebereich.
  Schreibweisen: $y = f(x)$, $f : x \mapsto y$.

  Beispiele:

  1. Student $\mapsto$ Assistent (aber nicht umgekehrt).
  2. $f : M_{m,n}(\mathbb{R}) \to M_{1,n}(\mathbb{R})$, definiert durch $A \mapsto Z_m(A)$.
  3. $f : \mathbb{R} \to \mathbb{R}$, definiert durch $f(x) := x^3$.
  4. $f : \mathbf{A}^n \times \mathbf{A}^n \to \mathbf{A}^n$, definiert durch $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathbf{y}$.

Definition. *Gruppe* = Menge G mit Verknüpfung $G \times G \to G$, $(a, b) \mapsto ab$, mit

1. $a(bc) = (ab)c$ (Assoziativgesetz),
2. $\exists\, e \in G$ mit $ea = ae = a$ für alle $a \in G$ (neutrales Element),
3. $\forall\, a \in G\ \exists\, a' \in G$ mit $aa' = a'a = e$ (inverses Element).

Die Gruppe heißt *kommutativ* oder *abelsch*, falls $ab = ba$ für alle $a, b \in G$.

Bemerkung. Die Verknüpfung wird bei allgemeinen Aussagen über Gruppen multiplikativ geschrieben. In speziellen Fällen benutzt man aber auch + oder andere Symbole.

Beispiele.

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$.
2. $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$ (der Stern bedeutet: die Null wird weggelassen).
3. $\mathrm{GL}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ invertierbar }\}$, mit der Matrizenmultiplikation.

1.1 Satz. *Es reicht, in (2) und (3) nur $ea = a$ und $a'a = e$ zu fordern.*

Beweis: Sei a gegeben, $a'a = e$ und $a''a' = e$. Dann folgt $aa' = (ea)a' = ((a''a')a)a' = \ldots = e$, sowie $ae = a(a'a) = \ldots = a$. ∎

Eine *Untergruppe* ist eine Teilmenge $H \subset G$ mit

$$1 \in H, \quad (a, b \in H \implies ab \in H), \quad (a \in H \implies a^{-1} \in H.$$

Jede Untergruppe ist eine Gruppe. Beispiel: $\mathbb{Z} \subset \mathbb{Q}$.

Definition. *Ring* = Menge R mit zwei Verknüpfungen $R \times R \to R$, $(a, b) \mapsto a + b$ und $(a, b) \mapsto a \cdot b$, mit

1. $(R, +)$ ist abelsche Gruppe,
2. $(a \cdot b) \cdot c = A \cdot (b \cdot c)$,
3. $a \cdot (b + c) = a \cdot b + a \cdot c$ und $(a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivgesetze).

R heißt *kommutativ*, falls zusätzlich $a \cdot b = b \cdot a$ gilt. R heißt *Ring mit Eins*, falls $\exists\, 1 \in R$ mit $1 \cdot a = a \cdot 1 = a$.

Beispiele.

1. $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ sind kommutative Ringe mit 1.
2. $M_n(\mathbb{R})$ ist ein (nicht kommutativer) Ring mit Eins E_n.

Ein Unterring von R ist eine Teilmenge $S \subset R$ mit

$$1 \in S, \quad (a, b \in S \implies a - b,\, ab \in R).$$

Jeder Unterring ist ein Ring.

Definition. Ein *Körper* K ist ein kommutativer Ring mit Eins, in dem gilt:

1. $1 \neq 0$.
2. $\forall\, a \in K^* = K \setminus \{0\}\ \exists\, a' \in K$ mit $aa' = 1$.

Beispiele.

1. $\mathbb{Z}$ und $M_n(\mathbb{R})$ sind **keine** Körper!
2. $\mathbb{Q}$ und $\mathbb{R}$ sind Körper.
3. $\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} \,:\, a, b \in \mathbb{Q}\}$ ist ein kommutativer Ring mit Eins (leicht nachzurechnen). Es ist auch $1 \neq 0$. Multiplikatives Inverses:

   $$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} \in \mathbb{Q}(\sqrt{2}).$$

   Annahme, $\sqrt{3} = a + b\sqrt{2}$. Dann $\sqrt{2} = (3 - a^2 - 2b^2)/(2ab)$ rational, WS.

Zwei umfangreichere Beispiele:

$$\begin{aligned}\mathbb{C} &:= \{Z = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} : x, y \in \mathbb{R}\} \\ &= \{xE + yI : x, y \in \mathbb{R}\}, \text{ mit } I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.\end{aligned}$$

Man rechnet leicht nach, dass $\mathbb{C}$ ein Unterraum und ein Unterring von $M_2(\mathbb{R})$ ist. Außerdem ist $I^2 = -E$.

Für $Z = aE + bI$ setze $\overline{Z} := aE - bI$ und $|Z| := \sqrt{a^2 + b^2}$. Dann ist

$$Z\overline{Z} = (aE + bI)(aE - bI) = a^2E - abI + baI + b^2E = |Z|^2 \cdot E, \text{ also } Z^{-1} = \frac{1}{|Z|^2}\overline{Z}.$$

Damit ist $\mathbb{C}$ ein Körper, der *Körper der komplexen Zahlen*. Man schreibt 1 statt E und i statt I, dann ist $\mathbb{C} = \{a + b\,\mathsf{i} : a, b \in \mathbb{R}\}$. Man kann $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen, wenn man a mit aE „identifiziert".

Für $m \in \mathbb{N}$, $m \geq 2$, sei $\mathbb{Z}_m := \{0, 1, 2, \ldots, m-1\}$. Zwei Zahlen $a, b \in \mathbb{Z}$ heißen *kongruent modulo* m, falls m ein Teiler von $a - b$ ist.

Beispiel: $7 \equiv 19 \mod 3$ (denn $7 - 19 = -12$) und $-12 \equiv 3 \mod 5$ (denn $-12 - 3 = -15$).

Zu jeder ganzen Zahl n gibt es genau ein $r \in \mathbb{Z}_m$, so dass $n = qm + r$ ist (mit $q \in \mathbb{Z}$), also $n \equiv r \mod m$. Schreibe dann $[n]$ für die Zahl r und setze

$$\begin{aligned} x \oplus y &:= [x + y] \\ \text{und} \quad x \otimes y &:= [x \cdot y]. \end{aligned}$$

Weil $x \oplus (m - x) = [x + (m - x)] = [m] = 0$ ist, gibt es zu jedem $x \in \mathbb{Z}_m$ ein Negatives. Man rechnet nun leicht nach, dass $\mathbb{Z}_m$ ein kommutativer Ring mit Eins und $1 \neq 0$ ist.

In $\mathbb{Z}_m$ ist $\underbrace{1 + 1 + \cdots + 1}_{m\text{-mal}} = 0$. Nun sei m eine **Primzahl** und $a \in \mathbb{Z}_m$, $a \neq 0$. Betrachte die Menge $P := \{a \otimes b : b \in \mathbb{Z}_m\} \subset \mathbb{Z}_m$. Wäre $a \otimes b_1 = a \otimes b_2$ für ein $b_1 \geq b_2$, so wäre $b := b_1 - b_2 \in \mathbb{Z}_m$ und $a \otimes b = 0$, also ab durch m teilbar. Weil m Primzahl ist, muss m dann a oder b teilen. Das geht nur, wenn $b = 0$ ist. Also besteht P aus m verschiedenen Elementen und es gibt ein $a' \in \mathbb{Z}_m$, so dass $a \otimes a' = 1$ ist. Also ist $\mathbb{Z}_m$ ein Körper.

Im Falle $m = 2$ erhält man einen Körper mit 2 Elementen.

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Definition. Sei K ein Körper. Ein K-*Vektorraum* ist eine Menge V mit Verknüpfungen $V \times V \to V$, $(x, y) \mapsto x + y$, und $K \times V \to V$, $(\alpha, x) \mapsto \alpha x$, so dass gilt:

1. $(V, +)$ ist eine abelsche Gruppe,
2. für $\alpha, \beta \in K$, $x, y \in V$, gelten die Regeln
   (a) $(\alpha\beta)x = \alpha(\beta x)$,
   (b) $1x = x$,
   (c) $\alpha(x + y) = \alpha x + \alpha y$,
   (d) $(\alpha + \beta)x = \alpha x + \beta x$.

Beispiele.

1. Der Raum $M_{m,n}(K)$ der $m \times n$-Matrizen mit Koeffizienten in K.
2. Der Spaltenraum $K^m := M_{m,1}(K)$ und der Zeilenraum $M_{1,n}(K)$.
3. Der Raum der Abbildungen $f : M \to K$, mit $(f + g)(x) := f(x) + g(x)$ und $(\alpha f)(x) := \alpha \cdot f(x)$.
4. Ist V ein K-Vektorraum, so heißt $U \subset V$ *Unterraum*, falls

   $$0 \in U,\ (x, y \in U \implies x + y \in U), (\alpha \in K, x \in U \implies \alpha x \in U).$$

   U ist dann selbst ein Vektorraum.

   Beispiele von Unterräumen:

   - $\{0\}$,
   - Geraden durch den Nullpunkt,
   - Der Durchschnitt $U_1 \cap U_2$ zweier Unterräume $U_1, U_2 \subset V$.

§ 2 Basis und Dimension

Sei V ein K-Vektorraum. Wie bei reellen Vektorräumen definiert man:

Linearkombination: $\sum_{i=1}^{k} \alpha_i x_i = \alpha_1 x_1 + \cdots + \alpha_k x_k$, $\alpha_i \in K$, $x_i \in V$.

Lineare Abhängigkeit und Unabhängigkeit von Vektoren.

2.1 Satz. *Äquivalente Aussagen:*

1. *$a_1, \ldots, a_k$ linear unabhängig.*
2. *Ist $\sum_{i=1}^{k} \alpha_i a_i = 0$, so ist $\alpha_1 = \ldots = \alpha_k = 0$.*
3. *$\forall\, a$: Ist $a = \sum_{i=1}^{k} \alpha_i a_i$, so sind die α_i eindeutig bestimmt.*

BEWEIS: 1) $\Longrightarrow$ 2): Sei $\sum_{i=1}^{k} \alpha_i a_i = 0$, etwa $\alpha_1 \neq 0$. Dann a_1 Linearkombination von $a_2, \ldots, a_k$, WS!

2) $\Longrightarrow$ 3): Ist $a = \sum_i \alpha_i a_i = \sum_i \beta_i a_i$, so ist $\sum_i (\alpha_i - \beta_i) a_i = 0$, also $\alpha_i = \beta_i$ für alle i.

3) $\Longrightarrow$ 1): Annahme, $a_1 = \sum_{i=2}^{k} \alpha_i a_i$. Dann $0 = 1 \cdot a_1 + \sum_{i=2}^{k} (-\alpha_i) a_i$. Wegen der Eindeutigkeit muss dann $1 = 0$ sein, WS! ∎

Ergänzungen zur Mengenlehre:

- Indexmengen,
- Familien von Mengen, speziell von Unterräumen,
- beliebige Durchschnitte und Vereinigungen.

Es gilt der **Satz:** Der Durchschnitt einer beliebigen Familie von Unterräumen ist wieder ein Unterraum.

Der Durchschnitt D aller Unterräume $U \subset V$ mit $M \subset U$ ist der „kleinste“ Unterraum von V, der M umfasst, denn:

1. D ist ein Unterraum von V, der M umfasst.
2. Sei U irgend ein Unterraum von V, der M umfasst. Dann ist $D \subset U$.

Definition. Sei V ein K-Vektorraum, $M \subset V$ Teilmenge. Dann heißt

$$\operatorname{Lin}(M) := \text{kleinster Unterraum von } V\text{, der } M \text{ umfasst}$$

die *lineare Hülle*, der *Spann* oder das *Erzeugnis* von M. Speziell ist $\operatorname{Lin}(\varnothing) = \{0\}$.

Bemerkung. Ist $M \neq \varnothing$, so ist $\operatorname{Lin}(M)$ die Menge aller Linearkombinationen von Elementen aus M. Definiert man eine leere Linearkombination als 0, so gilt diese Aussage auch für $M = \varnothing$.

Definition. Eine Teilmenge $M \subset V$ heißt *Erzeugendensystem* von V, wenn $V = \mathrm{Lin}(M)$ ist. V heißt *endlich erzeugt*, falls V ein endliches Erzeugendensystem besitzt.

Beispiele.

1. Sei $e_i := (0, \ldots, 0, 1, 0, \ldots, 0)^t \in K^n$ (mit einer 1 an der i-ten Stelle). Dann ist $E = \{e_1, \ldots, e_n\}$ ein endliches Erzeugendensystem von K^n, denn es gilt $x = (x_1, \ldots, x_n)^t = \sum_{i=1}^n x_i e_i$ für jedes $x \in K^n$.

2. $\{1, \mathsf{i}\}$ ist ein endliches Erzeugendensystem für $\mathbb{C}$ als $\mathbb{R}$-Vektorraum.

3. $\{1, \sqrt{2}\}$ ist ein endliches Erzeugendensystem für $\mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$-Vektorraum.

Definition. Ein n-Tupel $B = (a_1, \ldots, a_n)$ von Elementen eines K-Vektorraumes V heißt eine *(endliche) Basis* von V, falls gilt:

1. Die Vektoren $a_1, \ldots, a_n$ sind linear unabhängig.

2. $\{a_1, \ldots, a_n\}$ ist ein Erzeugendensystem von V.

Eine Familie $B = (a_\iota)_{\iota \in I}$ von Elementen von V heißt linear unabhängig, wenn jede endliche Teilfamilie linear unabhängig ist. Bildet die Menge der a_ι außerdem ein Erzeugendensystem von V, so nennt man B eine Basis von V.

Bemerkung. Man kann zeigen, dass jeder Vektorraum eine Basis besitzt. Wir werden diesen Satz aber hier nicht beweisen und uns nur mit endlich erzeugten Vektorräumen beschäftigen.

2.2 Satz. *$B = (a_1, \ldots, a_n)$ ist genau dann eine Basis von V, wenn sich jeder Vektor $x \in V$ auf eindeutige Weise als Linearkombination der a_i schreiben lässt.*

BEWEIS: 1) Ist B eine Basis, so ist B insbesondere ein Erzeugendensystem, und jedes $x \in V$ lässt sich als Linearkombination der a_i schreiben. Wegen der linearen Unabhängigkeit der a_i ist die Darstellung eindeutig.

2) Das Kriterium sei erfüllt. Da die Null auf eindeutige Weise dargestellt werden kann, sind die a_i linear unabhängig. Da jedes $x \in V$ dargestellt werden kann, bilden sie ein Erzeugendensystem. ∎

Beispiel.

$U := \{(x_1, x_2, x_3)^t \,:\, x_1 + x_2 + x_3 = 0\}$ ist Unterraum von $V = K^3$. Ist $x \in U$, so ist $x = x_1 \cdot (1, 0, -1)^t + x_2 \cdot (0, 1, -1)^t$. Also bilden $a_1 := (1, 0, -1)^t$ und $a_2 := (0, 1, -1)^t$ ein Erzeugendensystem von U. Nach dem Kriterium aus Kapitel 1 sind sie auch linear unabhängig, bilden also eine Basis.

2.3 Satz. *Sei $V \neq (0)$ ein endlich erzeugter K-Vektorraum. Dann gilt:*

1. *V besitzt eine Basis.*

2. *Jede Basis von V ist endlich.*

Beweis: 1) Sei $E = \{a_1, \ldots, a_n\}$ ein endliches Erzeugendensystem von V. Sind die a_i linear abhängig, so kann einer der Vektoren durch die anderen linear kombiniert werden und daher E verkleinert werden. Das macht man so lange, bis E minimal ist. Dann müssen die Vektoren linear unabhängig und $B = (a_1, \ldots, a_n)$ eine Basis sein.

2) Sei B' eine weitere (eventuell unendliche) Basis. Zu jedem i gibt es eine endliche Teilfamilie $B_i \subset B'$, so dass a_i Linearkombination der Elemente von B_i ist. Also gibt es eine endliche Teilfamilie $B'' \subset B'$, die ein Erzeugendensystem von V darstellt. Dieses System kann man so lange verkleinern, bis man ein linear unabhängiges Teilsystem $B''' \subset B''$ erhält, das immer noch V erzeugt.

Gäbe es ein $b \in B' \setminus B'''$, so müsste b immerhin noch Linearkombination der Elemente von B''' sein. Das würde aber bedeuten, dass $\{b\} \cup B''' \subset B'$ nicht linear unabhängig wäre, im Widerspruch dazu, dass B' eine Basis ist. Also muss $B' = B'''$ endlich sein. ■

Beispiel.

Die *Einheitsvektoren* $e_1, \ldots, e_n \subset K^n$ bilden eine Basis von K^n. Jede andere Basis von K^n muss auch endlich sein.

2.4 Satz. *Sei V ein beliebiger K-Vektorraum. Folgende Aussagen über Vektoren $a_1, \ldots, a_n \in V$ sind äquivalent:*

1. *$(a_1, \ldots, a_n)$ ist eine Basis von V.*

2. *$a_1, \ldots, a_n$ bilden ein maximales linear unabhängiges System in V.*

3. *$\{a_1, \ldots, a_n\}$ ist ein minimales Erzeugendensystem von V.*

Beweis: (1) $\Longrightarrow$ (2): Sei $(a_1, \ldots, a_n)$ Basis. Ist $a \in V$ beliebig, so gibt es α_i mit $a = \sum_i \alpha_i a_i$. Dann ist

$$1 \cdot a + (-\alpha_1)a_1 + \cdots + (-\alpha_n)a_n = 0.$$

Also können $a, a_1, \ldots, a_n$ nicht linear unabhängig sein.

(2) $\Longrightarrow$ (3): $a_1, \ldots, a_n$ mögen ein maximales linear unabhängiges System bilden. Für jedes $a \notin \{a_1, \ldots, a_n\}$ sind dann $a, a_1, \ldots, a_n$ linear abhängig. Das bedeutet, dass es eine Darstellung $0 = \alpha a + \alpha_1 a_1 + \cdots + \alpha_n a_n$ mit $(\alpha, \alpha_1, \ldots, \alpha_n) \neq (0, 0, \ldots, 0)$ gibt. Wäre $\alpha = 0$, so müsste auch $\alpha_1 = \ldots = \alpha_n = 0$ sein. Also ist $\alpha \neq 0$ und a Linearkombination von $a_1, \ldots, a_n$. Da a beliebig war, ist $\{a_1, \ldots, a_n\}$ ein Erzeugendensystem.

Wäre $\{a_1, \ldots, a_n\}$ kein minimales Erzeugendensystem, so wären die Vektoren linear abhängig. WS!

(3) $\Longrightarrow$ (1): Sei $\{a_1, \ldots, a_n\}$ ein minimales Erzeugendensystem. Wären die Vektoren linear abhängig, so könnte man einen davon weglassen. Aber das ist nicht möglich. Also ist $(a_1, \ldots, a_n)$ eine Basis. ∎

Wollen nun zeigen: Je zwei Basen haben gleich viel Elemente.

2.5 Austausch-Lemma. *Sei V ein K-Vektorraum, $B = (v_1, \ldots, v_n)$ eine Basis von V und $w \in V$. Ist $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$ und $\lambda_k \neq 0$, so ist auch*

$$B' = (v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$$

eine Basis von V.

Beweis: O.B.d.A. $k = 1$, also $\lambda_1 \neq 0$.

1) Dann ist v_1 Linearkombination von $w, v_2, \ldots, v_n$ und daher $\{w, v_2, \ldots, v_n\}$ ein Erzeugendensystem.

2) Sei $\mu w + \sum_{i=2}^{n} \mu_i v_i = 0$. Dann ist

$$0 = \mu \cdot \sum_{i=1}^{n} (\lambda_i v_i) + \sum_{i=2}^{n} \mu_i v_i = (\mu\lambda_1) v_1 + \sum_{i=2}^{n} (\mu\lambda_i + \mu_i) v_i,$$

also $\mu\lambda_1 = \mu\lambda_2 + \mu_2 = \ldots = \mu\lambda_n + \mu_n = 0$. daraus folgt: $\mu = 0$ und dann $\mu_2 = \ldots = \mu_n = 0$. ∎

2.6 Austauschsatz von Steinitz. *Sei $B = (v_1, \ldots, v_n)$ eine Basis von V. Sind die Vektoren $w_1, \ldots, w_r \in V$ linear unabhängig, so gilt:*

1. *$r \leq n$.*
2. *Es gibt Indizes $i_1, \ldots, i_{n-r}$, so dass $B' = (w_1, \ldots, w_r, v_{i_1}, \ldots, v_{i_{n-r}})$ wieder eine Basis von V ist.*

Beweis: Induktion nach r. Ist $r = 0$, so ist nichts zu zeigen.

Der Satz sei nun für $r - 1$ bewiesen.

Sind $w_1, \ldots, w_r \in V$ linear unabhängig, so sind erst recht $w_1, \ldots, w_{r-1}$ linear unabhängig. Nach Induktionsvoraussetzung ist $r - 1 \leq n$, und nach geeigneter Nummerierung ist $(w_1, \ldots, w_{r-1}, v_r, \ldots, v_n)$ eine Basis von V.

Wäre $r - 1 = n$, so wäre $(w_1, \ldots, w_{r-1})$ eine Basis von V, aber kein maximales linear unabhängiges System. Das kann nicht sein. Also ist $r - 1 < n$ und damit $r \leq n$.

Es gibt eine Darstellung $w_r = \sum_{i=1}^{r-1} \lambda_i w_i + \sum_{j=r}^{n} \mu_j v_j$. Wäre $\mu_r = \ldots = \mu_n = 0$, so wären $w_1, \ldots, w_{r-1}, w_r$ linear abhängig, WS! Also gibt es einen Index ϱ, so dass $\mu_\varrho \neq 0$ ist. O.B.d.A. sei $\varrho = r$. Dann kann man nach dem Austausch-Lemma v_r durch w_r ersetzen und erhält wieder eine Basis von V, nämlich $(w_1, \ldots, w_r, v_{r+1}, \ldots, v_n)$.
∎

2.7 Folgerung. *Je zwei endliche Basen von V besitzen gleich viele Elemente.*

BEWEIS: Sind $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_r)$ zwei Basen von V, so folgt aus dem Austauschsatz: $r \le n$ und $n \le r$, also $r = n$. ∎

Definition. Sei V ein K-Vektorraum. Dann heißt

$$\dim_K(V) := \begin{cases} n & \text{falls } V \text{ Basis mit } n \text{ Elementen besitzt,} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

die *Dimension* von V (über K).

2.8 Satz. *Sei V ein endlich erzeugter K-Vektorraum, $U \subset V$ ein Unterraum. Dann ist auch U endlich erzeugt und* $\dim_K(U) \le \dim_K(V)$.

Ist sogar $\dim_K(U) = \dim_K(V)$, *so ist $U = V$.*

BEWEIS: Annahme, U ist nicht endlich erzeugt. Dann gibt es in U kein maximales linear unabhängiges System (denn das wäre ja eine endliche Basis). Ist $n = \dim_K(V)$, so kann man also $n + 1$ linear unabhängige Vektoren in U finden. Aber die sind auch in V linear unabhängig, und das ist ein WS. Also ist U doch endlich erzeugt. Ist B eine Basis von U mit r Elementen, so sind diese Elemente auch in V linear unabhängig und es muss $r \le n$ gelten.

Ist $r = n$, aber $U \ne V$, so gibt es ein $x \in V \setminus U$, $x \ne 0$. Es muss x von den Elementen einer Basis von U linear unabhängig sein (denn U enthält alle Linearkombinationen der Basiselemente). Das ergibt ein System von $n+1$ linear unabhängigen Vektoren in V, aber das kann nicht sein. ∎

2.9 Folgerung. *Jeder Unterraum U von K^n besitzt eine (endliche) Basis mit höchstens n Elementen.*

2.10 Basisergänzungssatz. *Sei V ein endlich-dimensionaler K-Vektorraum. Sind die Vektoren $a_1, \ldots, a_r \in V$ linear unabhängig, so gibt es Vektoren*

$$v_{r+1}, v_{r+2}, \ldots, v_n,$$

so dass $\{a_1, \ldots, a_r, v_{r+1}, \ldots, v_n\}$ eine Basis von V ist.

BEWEIS: Sei $\{v_1, \ldots, v_n\}$ eine Basis von V. Nach dem Austauschsatz ist $r \le n$, und nach geeigneter Nummerierung ist $\{a_1, \ldots, a_r, v_{r+1}, \ldots, v_n\}$ eine Basis von V.
∎

§ 3 Lineare Abbildungen

Ergänzungen zur Theorie der Abbildungen:

Sei $f : A \to B$ eine Abbildung.

- Ist $M \subset A$, so heißt $f(M) := \{f(x) : x \in M\}$ die *Bildmenge* von M.
- Ist $N \subset B$, so heißt $f^{-1}(N) := \{x \in A : f(x) \in N\}$ die *Urbildmenge* von N.
- f heißt *injektiv*, falls gilt: $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$ („verschiedene Bogenschützen treffen verschiedene Ziele").

  Äquivalent dazu: $f(x_1) = f(x_2) \implies x_1 = x_2$.
- f heißt *surjektiv*, falls gilt: $f(A) = B$, d.h. $\forall\, y \in B\ \exists\, x \in A$ mit $f(x) = y$.

Definition. V, W seien K-Vektorräume. Eine Abbildung $f : V \to W$ heißt K-linear, falls gilt:

$$\begin{aligned} f(x+y) &= f(x) + f(y) \text{ für alle } x, y \in V, \\ \text{und } f(\alpha x) &= \alpha \cdot f(x) \text{ für } \alpha \in K,\ x \in V. \end{aligned}$$

Beispiele.

1. V Vektorraum mit Basis $(a_1, \ldots, a_n)$, $f : K^n \to V$ mit $f(\alpha_1, \ldots, \alpha_n) := \alpha_1 a_1 + \cdots + \alpha_n a_n$.
2. Sei V ein K-Vektorraum, $v \in V$, $v \neq 0$, $f : K \to V$ definiert durch $f(\alpha) := \alpha v$.
3. $f : \mathbb{R} \to \mathbb{R}$ mit $f(t) = at$ ist linear, $f(t) = at + b$ mit $b \neq 0$ ist **nicht** linear, $f(t) = t^2$ ist auch nicht linear.
4. Besonders wichtiges Beispiel:

   Sei $A \in M_{m,n}(K)$, $F_A : K^n \to K^m$ definiert durch $F_A(x) := A \cdot x$.
   Dann ist $F_A(e_i) = A \cdot e_i = S_i(A)$.

   Achtung! Nicht jede lineare Abbildung kann so beschrieben werden!

3.1 Satz. *Sei $(a_1, \ldots, a_n) \in V$ eine Basis, $b_1, \ldots, b_n \in W$ beliebig. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit $f(a_i) = b_i$ für $i = 1, \ldots, n$.*

Beweis: Erst Eindeutigkeit, dann Existenz. $f(\sum_i x_i a_i) := \sum_i x_i b_i$. ∎

Definition. Sei $f : V \to W$ linear. Dann definiert man:

$$\operatorname{Ker} f := \{x \in V : f(x) = 0\} \quad \text{und} \quad \operatorname{Im} f := \{f(x) \,.\, x \in V\}.$$

Man zeigt leicht: $\operatorname{Ker} f \subset V$ und $\operatorname{Im} f \subset W$ sind Unterräume.

3.2 Satz. *Sei $f : V \to W$ linear. Dann gilt:*

$$\begin{aligned} f \textit{ injektiv} &\iff \operatorname{Ker} f = \{0\}, \\ f \textit{ surjektiv} &\iff \operatorname{Im} f = W. \end{aligned}$$

BEWEIS einfach.

3.3 Lemma. *Sei $f : V \to W$ linear,* $\operatorname{Ker} f = \{0\}$. *Sind $v_1, \dots, v_n \in V$ linear unabhängig, so sind auch $f(v_1), \dots, f(v_n) \in W$ linear unabhängig.*

BEWEIS trivial.

3.4 Die Dimensionsformel. *V, W seien endlich-dimensionale K-Vektorräume, $f : V \to W$ linear. Dann ist*

$$\dim V = \dim \operatorname{Ker} f + \dim \operatorname{Im} f.$$

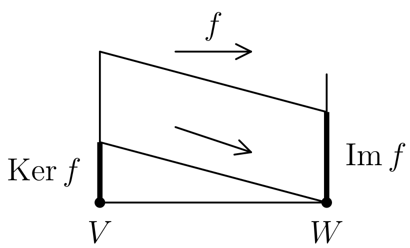
BEWEIS: Für $f = 0$ ist die Behauptung trivial, also sei o.B.d.A. $f \neq 0$, d.h. $\dim \operatorname{Im} f > 0$.

Sei $(w_1, \dots, w_s)$ Basis von $\operatorname{Im} f$ und $(u_1, \dots, u_r)$ Basis von $\operatorname{Ker} f$ (evtl. $r = 0$). Wähle $v_1, \dots, v_s \in V$ mit $f(v_i) = w_i$ für $i = 1, \dots, s$. Zeige dann, dass $(u_1, \dots, u_r, v_1, \dots, v_s)$ Basis von V ist.

a) Erzeugendensystem: $v \in V \implies f(v) = f(x_1 v_1 + \cdots + x_s v_s)$, also

$$v - (x_1 v_1 + \cdots + x_s v_s) \in \operatorname{Ker} f, \text{ d.h. } v \in \operatorname{Lin}(v_1, \dots, v_s, u_1, \dots, u_r).$$

b) Lineare Unabhängigkeit: Sei $\sum_i \alpha_i v_i + \sum_j \beta_j u_j = 0$. Wende f an, dann folgt: $\sum_i \alpha_i w_i = 0$, also alle $\alpha_i = 0$. Dann sind aber auch alle $\beta_j = 0$. ∎



Definition. $f : V \to W$ heißt *Isomorphismus*, falls f linear und bijektiv ist.

3.5 Satz. *Ist $f : V \to W$ ein Isomorphismus, so ist f^{-1} linear und* $\dim V = \dim W$.

BEWEIS: 1. Teil trivial, 2. Teil mit $\operatorname{Ker} f = 0 \implies \dim V = \dim \operatorname{Im} f = \dim W$. ∎

3.6 Satz. *Sei $f : V \to W$ linear,* $\dim V = \dim W$. *Dann ist äquivalent:*

a) f Isomorphismus,
b) f injektiv,
c) f surjektiv.

Beweis: Einfach, arbeite mit der Dimensionsformel. ∎

3.7 Satz. *Ist $f : V \to W$ und $g : W \to U$ linear, so ist auch $g \circ f : V \to U$ linear.*

Beweis trivial.

Definition. Sei V ein K-Vektorraum, $U_1, U_2 \subset V$ zwei Unterräume. Dann nennt man

$$U_1 + U_2 := \{x + y \, : \, x \in U_1 \text{ und } y \in U_2\}.$$

die *Summe* von U_1 und U_2.
$U_1 + U_2$ ist natürlich wieder ein Unterraum von V.

Definition. Sind V, W zwei beliebige K-Vektorräume, so nennt man $V \times W$ das *Produkt* von V und W.

3.8 Satz. *Mit den obigen Bezeichnungen und den Rechenvorschriften*

$$(v, w) + (v', w') := (v + v', w + w') \quad \text{und} \quad \alpha \cdot (v, w) := (\alpha v, \alpha w)$$

ist $V \times W$ wieder ein K-Vektorraum.

Beweis: Trivial. Aber man muss hier **alle** VR-Axiome überprüfen. ∎

3.9 Lemma. *Sei V ein K-Vektorraum, $U_1, U_2 \subset V$ zwei Unterräume. Die Abbildungen $f : U_1 \cap U_2 \to U_1 \times U_2$ und $g : U_1 \times U_2 \to V$ seien definiert durch $f(u) := (u, u)$ und $g(u_1, u_2) := u_1 - u_2$.*

Dann sind f und g linear, und es gilt:

1. $\operatorname{Ker} f = \{0\}$ *und* $\operatorname{Im} f = \operatorname{Ker} g$.
2. $\operatorname{Im} g = U_1 + U_2$.

Beweis: Die Linearität ist trivial.

Offensichtlich ist $(u, u) = f(u) = (0, 0)$ genau dann, wenn $u = 0$ ist, also $\operatorname{Ker} f = \{0\}$. Weiter ist $\operatorname{Im} f = \{(u_1, u_2) \in U_1 \times U_2 \, : \, u_1 = u_2\} = \{(u_1, u_2) \in U_1 \times U_2 \, : \, u_1 - u_2 = 0\}$, also $\operatorname{Im} f = \operatorname{Ker} g$. Schließlich gilt $v = u_1 + u_2$ mit $u_1 \in U_1, u_2 \in U_2$ genau dann, wenn $v = g(u_1, -u_2)$ ist. Also ist $U_1 + U_2 = \operatorname{Im} g$. ∎

3.10 Folgerung. *Sei V ein K-Vektorraum, $U_1, U_2 \subset V$ zwei Unterräume. Dann ist*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

BEWEIS: Klar, weil $\dim(U_1 \times U_2) = \dim U_1 + \dim U_2$ ist. ∎

Definition. Sei V ein K-Vektorraum, $U_1, U_2 \subset V$ zwei Unterräume, $U := U_1 + U_2$. Man nennt U die *direkte Summe* von U_1 und U_2, falls $U_1 \cap U_2 = \{0\}$ ist. Man schreibt dann: $U = U_1 \oplus U_2$.
Offensichtlich ist

$$\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2.$$

Definition. Sei V ein K-Vektorraum und $U \subset V$ ein Unterraum. Ein *Komplementärraum* zu U ist ein Unterraum $U' \subset V$, so dass $U \oplus U' = V$ ist.

3.11 Satz. *Sei V ein endlich-dimensionaler K-Vektorraum und $U \subset V$ ein Unterraum. Dann gibt es in V einen Komplementärraum U' zu U.*

BEWEIS: Sei $(v_1, \ldots, v_k)$ Basis von U. Nach Basisergänzungssatz gibt es $v_{k+1}, \ldots, v_n \in V$, so dass $(v_1, \ldots, v_n)$ Basis von V ist. Man setze $U' := \mathrm{Lin}(\{v_{k+1}, \ldots, v_n\})$. Dann ist offensichtlich $U + U' = V$. Es ist aber auch $U \cap U' = \{0\}$, wegen der linearen Unabhängigkeit der v_i. ∎

Zurück zu den linearen Gleichungssystemen:

Sei $A \in M_{m,n}(K)$ und $b \in K^m$. Dann kann man das Gleichungssystem $A \cdot x = b$ auch in der folgenden Form schreiben: $F_A(x) = b$, wobei $F_A : K^n \to K^m$ die durch $F_A(x) := A \cdot x$ definierte lineare Abbildung ist. Noch allgemeiner:

V, W seien zwei (endlich-dimensionale) K-Vektorräume, $f : V \to W$ linear. Unter einem *linearen Gleichungssystem* versteht man eine Gleichung der Form

$$f(x) = b, \text{ mit } b \in W.$$

Unter dem *zugehörigen homogenen System* versteht man die Gleichung

$$f(x) = 0.$$

Mit $\mathrm{Lös}(f, b)$ bzw. $\mathrm{Lös}(f, 0)$ seien die jeweiligen Lösungsmengen bezeichnet.

3.12 Satz.

1. $\mathrm{Lös}(f, 0) = \mathrm{Ker}(f)$.
2. *Die Gleichung $f(x) = b$ ist genau dann lösbar, wenn $b \in \mathrm{Im}\, f$ ist. Ist $x_0 \in V$ eine Lösung, so ist* $\mathrm{Lös}(f, b) = \{x_0 + x : x \in \mathrm{Ker}\, f\}$.

BEWEIS trivial.

Definition. Die Zahl $\mathrm{rg}(f) := \dim \mathrm{Im}\, f$ nennt man den *Rang* von f.

Ist $A \in M_{m,n}(K)$, so versteht man unter dem *Rang der Matrix A* den Rang der linearen Abbildung $F_A : K^n \to K^m$.

3.13 Satz. $\mathrm{rg}(A)$ *ist die (maximale) Anzahl linear unabhängiger Spalten von* A.

Beweis: Ist $x = \sum_i x_i e_i \in K^n$, so ist $A \cdot x = A \cdot (\sum_i x_i e_i) = \sum_i x_i(A \cdot e_i)$. Also wird $\mathrm{Im}(F_A)$ von den Spalten $A \cdot e_i$ der Matrix A erzeugt. Daraus folgt die Behauptung. ∎

3.14 Satz. *Das Gleichungssystem* $A \cdot x = b$ *ist genau dann lösbar, wenn* $\mathrm{rg}(A) = \mathrm{rg}(A|b)$ *ist.*

Beweis: Es ist $\mathrm{rg}(A) = \mathrm{rg}(A|b)$ genau dann, wenn b linear abhängig von den Spalten $Ae_1, \ldots, Ae_n$ ist, wenn es also Skalare x_i mit $b = \sum_i x_i A e_i = A \cdot (\sum_i x_i e_i)$ gibt. Aber das ist gleichbedeutend damit, dass $Ax = b$ für $x := \sum_i x_i e_i$ ist. ∎

Man bezeichnet den Rang einer Matrix A auch als *Spaltenrang* von A und schreibt $\mathrm{rg}_s(A)$ dafür. Unter dem *Zeilenrang* (in Zeichen $\mathrm{rg}_z(A)$) von A versteht man die (maximale) Anzahl linear unabhängiger Zeilen von A. Offensichtlich ist

$$\mathrm{rg}_z(A) = \mathrm{rg}(A^t).$$

3.15 Satz. *Sei* $A \in M_{m,n}(K)$ *und* $B \in M_{n,r}(K)$, *so ist*

$$F_{AB} = F_A \circ F_B.$$

Beweis: $F_{AB}(x) = (AB)x = A(Bx) = F_A(Bx) = F_A(F_B(x)) = (F_A \circ F_B)(x)$. ∎

3.16 Satz. U, V, W, S *seien* K*-Vektorräume,* $f : V \to W$ *eine lineare Abbildung,* $g : U \to V$ *und* $h : W \to S$ *Isomorphismen. Dann ist* $\mathrm{rg}(h \circ f \circ g) = \mathrm{rg}(f)$.

Beweis: Es ist $g(U) = V$, also $\mathrm{rg}(f \circ g) = \dim(f \circ g)(U) = \dim f(g(U)) = \dim f(V) = \mathrm{rg}(f)$. Es reicht also, $\dim(h \circ f)(V) = \dim f(V)$ zu beweisen.

Sei $T := h(f(V))$. Weil $h : W \to S$ ein Isomorphismus ist, ist $h : f(V) \to T$ injektiv. Außerdem ist diese Abbildung definitionsgemäß surjektiv. Es folgt, dass $\dim T = \dim f(V)$ ist, und damit $\dim(h \circ f)(V) = \dim T = \dim f(V)$. ∎

3.17 Satz. *Spaltenrang und Zeilenrang einer Matrix sind invariant unter den elementaren Umformungen (A) bis (D).*

Beweis: Sei $A \in M_{m,n}(K)$. Geht A durch elementare Umformungen in eine Matrix B über, so gibt es invertierbare Matrizen $Q \in M_m(K)$ und $P \in M_n(K)$, so dass $B = QAP$ ist. Dann ist $\mathrm{rg}_s(B) = \mathrm{rg}(B) = \mathrm{rg}(QAP) = \mathrm{rg}(A) = \mathrm{rg}_s(A)$.

Da auch Q^t und P^t invertierbar sind, folgt:

$$\begin{aligned}\mathrm{rg}_z(B) &= \mathrm{rg}_z(QAP) = \mathrm{rg}((QAP)^t)\\ &= \mathrm{rg}(P^t \cdot A^t \cdot Q^t)\\ &= \mathrm{rg}(A^t) = \mathrm{rg}_z(A).\end{aligned}$$

■

3.18 Satz. *Zeilenrang = Spaltenrang.*

BEWEIS: Jede Matrix $A \in M_{m,n}(K)$ kann mit dem Gauß-Verfahren in die Gestalt $A' = \left(\begin{array}{c|c} D & B \\ \hline 0 & 0 \end{array}\right)$ gebracht werden, wobei $D \in M_r(K)$ eine obere Dreiecksmatrix mit nicht-verschwindenden Diagonalelementen ist, und $B \in M_{r,n-r}(K)$.

Die ersten r Spalten von A' sind linear unabhängig, und da alle Spalten von A' in dem r-dimensionalen Raum $V = \{x \in K^m : x_{r+1} = \ldots = x_m = 0\}$ liegen, ist $\mathrm{rg}_s(A') = r$. Offensichtlich sind auch die ersten r Zeilen von A' linear unabhängig, und da alle anderen Zeilen verschwinden, ist $\mathrm{rg}_z(A') = r$.

Aus dem vorigen Satz folgt: $\mathrm{rg}_z(A) = \mathrm{rg}_z(A') = r = \mathrm{rg}_s(A') = \mathrm{rg}_s(A)$. ■

Das Gaußverfahren kann also auch benutzt werden, um den Rang einer Matrix zu bestimmen.

Ist $A \in M_{m,n}(K)$, $F_A : K^n \to K^m$ die zugehörige lineare Abbildung, so folgt:

$$\dim \mathrm{Lös}(A,0) = \dim \mathrm{Ker}\, F_A = n - \dim \mathrm{Im}\, F_A = n - \mathrm{rg}(A).$$

Eine Basis des Lösungsraumes muss $n - \mathrm{rg}(A)$ Elemente enthalten.

$$\begin{aligned}\text{Es ist}\quad x \in \mathrm{Lös}(A,0) &\iff A'x = 0\\ &\iff x = \begin{pmatrix} x' \\ x'' \end{pmatrix} \text{ mit } Dx' + Bx'' = 0\\ &\iff x = \begin{pmatrix} x' \\ x'' \end{pmatrix} \text{ und } x' = -D^{-1}Bx''.\end{aligned}$$

Setzt man $a_j := -D^{-1}Be_j$, für $j = 1, \ldots, n-r$, so sind die Vektoren

$$s_j := \begin{pmatrix} a_j \\ e_j \end{pmatrix}$$

linear unabhängige Lösungen, bilden also eine Basis von $\mathrm{Lös}(A,0)$.

§ 4 Skalarprodukte

Definition. Sind $\mathbf{a} = (a_1, \ldots, a_n)^t$ und $\mathbf{b} = (b_1, \ldots, b_n)^t$ Vektoren im $\mathbb{R}^n$, so versteht man unter ihrem *Skalarprodukt* die Zahl $\mathbf{a} \bullet \mathbf{b} := a_1 b_1 + \cdots + a_n b_n$.
Man beachte, dass das Skalarprodukt zweier Vektoren immer eine reelle Zahl, also ein Skalar ist!

4.1 Satz.

1. *Es ist* $\mathbf{a} \bullet \mathbf{b} = \mathbf{b} \bullet \mathbf{a}$*.*
2. *Für drei Vektoren* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *gilt:* $\mathbf{a} \bullet (\mathbf{b} + \mathbf{c}) = \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \bullet \mathbf{c}$*.*
3. *Ist* $\alpha \in \mathbb{R}$*, so gilt:* $(\alpha \mathbf{a}) \bullet \mathbf{b} = \mathbf{a} \bullet (\alpha \mathbf{b}) = \alpha \cdot (\mathbf{a} \bullet \mathbf{b})$*.*
4. *Es ist* $\mathbf{a} \bullet \mathbf{a} = \sum_{i=1}^n a_i^2 \geq 0$ *und sogar* > 0*, wenn* $\mathbf{a} \neq \mathbf{0}$ *ist.*

BEWEIS: Die Eigenschaften ergeben sich ganz leicht aus der Definition und den Rechenregeln für reelle Zahlen. ∎

Definition. Unter der *Länge* (oder *Norm*) eines Vektors $a = (a_1, \ldots, a_n)^t$ versteht man die Zahl $\|a\| := \sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{a_1^2 + \cdots + a_n^2}$.

4.2 Satz. *Die Norm von Vektoren hat folgende Eigenschaften:*

1. *Es ist stets* $\|x\| \geq 0$*.*
2. $\|x\| = 0 \iff x = 0$*.*
3. $\|\alpha x\| = |\alpha| \cdot \|x\|$*.*

BEWEIS: 1) Ist klar nach Definition.

2) Ist $\|\mathbf{x}\| = 0$, so ist $x_1^2 + \cdots + x_n^2 = 0$. Das geht nur, wenn alle $x_i = 0$ sind. Die Umkehrung ist trivial.

3) $\|\alpha \mathbf{x}\| = \sqrt{(\alpha x_1)^2 + \cdots + (\alpha x_n)^2} = \sqrt{\alpha^2 \cdot (x_1^2 + \cdots + x_n^2)} = |\alpha| \cdot \|\mathbf{x}\|$. ∎

Wir wollen den Begriff des Skalarproduktes verallgemeinern. Leider geht das nicht bei beliebigen Körpern.

Definition. Sei K ein beliebiger Körper und V ein K-Vektorraum. Eine *Bilinearform* auf V ist eine Abbildung $\Phi : V \times V \to K$ mit folgenden Eigenschaften:

1. $\Phi(x_1 + x_2, y) = \Phi(x_1, y) + \Phi(x_2, y)$ für $x_1, x_2, y \in V$.
2. $\Phi(x, y_1 + y_2) = \Phi(x, y_1) + \Phi(x, y_2)$ für $x, y_1, y_2 \in V$.
3. $\Phi(\alpha x, y) = \Phi(x, \alpha y) = \alpha \cdot \Phi(x, y)$ für $\alpha \in K$ und $x, y \in V$.

Die Bilinearform Φ heißt *symmetrisch*, falls $\Phi(x, y) = \Phi(y, x)$ für alle $x, y \in V$ gilt.

Definition. Sei V ein **reeller** Vektorraum. Eine symmetrische Bilinearform auf V heißt ein *Skalarprodukt* auf V, falls sie außerdem *positiv definit* ist, d.h., falls $\Phi(x, x) > 0$ für alle $x \neq 0$ ist.

Definition. Sei V ein **komplexer** Vektorraum. Eine *hermitesche Form* auf V ist eine Bilinearform Φ auf dem reellen Vektorraum V, für die außerdem gilt:

1. $\Phi(cx, y) = c \cdot \Phi(x, y)$ für $c \in \mathbb{C}$ und $x, y \in V$.
2. $\Phi(y, x) = \overline{\Phi(x, y)}$.

4.3 Satz. *Ist V ein komplexer Vektorraum und Φ eine hermitesche Form auf V, so gilt:*

1. $\Phi(x, cy) = \overline{c} \cdot \Phi(x, y)$ *für $c \in \mathbb{C}$ und $x, y \in V$.*
2. $\Phi(x, x)$ *ist für jedes $x \in V$ eine reelle Zahl.*

Der BEWEIS ist trivial.

Definition. Sei V ein **komplexer** Vektorraum. Eine hermitesche Form auf V heißt ein *Skalarprodukt* auf V, falls sie außerdem *positiv definit* ist, d.h., falls $\Phi(x, x) > 0$ für alle $x \neq 0$ ist.

Wir können künftig reelle und komplexe Vektorräume mit Skalarprodukt weitgehend gemeinsam behandeln. Für reelle Zahlen α ist $\overline{\alpha} = \alpha$. Im reellen Fall sprechen wir von einem *euklidischen Raum*, im komplexen Fall von einem *unitären Raum*. Das Skalarprodukt zweier Vektoren $x, y \in V$ wird mit $<x\,,\,y>$ bezeichnet.

Im Folgenden sei V immer ein euklidischer oder unitärer Raum.

Definition. Die Zahl $\|x\| := \sqrt{<x\,,\,x>}$ nennt man die *Norm* oder *Länge* von x.

4.4 Eigenschaften der Norm. *Sei $K \in \{\mathbb{R}, \mathbb{C}\}$ und V ein K-Vektorraum.*

1. *Es ist stets* $\|x\| \geq 0$.
2. $\|x\| = 0 \iff x = 0$.
3. $\|\alpha x\| = |\alpha| \cdot \|x\|$ *für $\alpha \in K$ und $x \in V$.*

BEWEIS: 1) Ist klar nach Definition.

2) Ist $x = 0$, so ist offensichtlich $\|x\| = 0$. Ist $x \neq 0$, so ist $\|x\| > 0$, nach Definition des Skalarproduktes.

3) $\|\alpha x\| = \sqrt{<\alpha x\,,\,\alpha x>} = \sqrt{\alpha^2 \cdot <x\,,\,x>} = |\alpha| \cdot \|x\|$. ∎

Das kanonische Skalarprodukt des $\mathbb{R}^n$ in niedrigen Dimensionen:

- Ist $n = 1$, so ist das Skalarprodukt einfach das gewöhnliche Produkt.
- Sei $n = 2$, also $\mathbf{a} = (a_1, a_2)^t$, $\mathbf{b} = (b_1, b_2)^t$ und $\mathbf{a} \bullet \mathbf{b} = a_1 b_1 + a_2 b_2$.

  Ist $\mathbf{a} = \mathbf{0}$ oder $\mathbf{b} = \mathbf{0}$, so ist auch $\mathbf{a} \bullet \mathbf{b} = 0$. Das Skalarprodukt zweier Vektoren kann aber auch dann verschwinden, wenn keiner der beiden Vektoren der Nullvektor ist.

  Sei etwa $\mathbf{a} \neq \mathbf{0}$. Dann muss eine Koordinate von $\mathbf{a} \neq 0$ sein, etwa $a_1 \neq 0$. Ist nun $\mathbf{a} \bullet \mathbf{b} = 0$, also $a_1 b_1 + a_2 b_2 = 0$, so folgt: $b_1 = -a_2 b_2 / a_1$. Setzen wir $\lambda := b_2 / a_1$, so ist $(b_1, b_2)^t = \lambda(-a_2, a_1)^t$.

  Die Vektoren $(a_1, a_2)^t, (b_1, b_2)^t$ stehen aufeinander senkrecht. Daher definiert man:

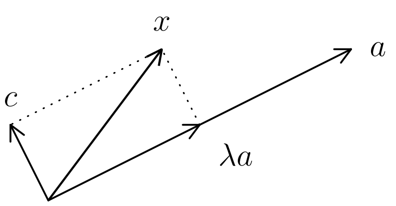
Definition. Sei V ein beliebiger euklidischer oder unitärer Raum. Zwei Vektoren a und b in V heißen *orthogonal*, falls $<a\,,\,b> = 0$ ist.

4.5 Existenz der orthogonalen Projektion. *Es sei $a \neq 0$ ein Vektor in V. Dann gibt es zu jedem $x \in V$ genau ein $\lambda \in K$ und einen Vektor c, so dass $<a\,,\,c> = 0$ und $x = \lambda a + c$ ist.*

Beweis: Aus der Gleichung $0 = <a\,,\,c> = <a\,,\,x - \lambda a> = <a\,,\,x> - \overline{\lambda}<a\,,\,a>$ folgt:

$$\lambda = \frac{<x\,,\,a>}{<a\,,\,a>} \text{ und } c = x - \lambda a.$$

Das zeigt die Eindeutigkeit der Darstellung, aber auch die Existenz. ∎



Den Vektor $\lambda a = \dfrac{<x\,,\,a>}{<a\,,\,a>} \cdot a$ nennt man die *orthogonale Projektion* von x auf a.

4.6 Satz des Pythagoras. *Ist $<a\,,\,b> = 0$, so ist* $\|a + b\|^2 = \|a\|^2 + \|b\|^2$.

Beweis: Es ist

$$\begin{aligned} \|a + b\|^2 &= <a + b\,,\,a + b> \\ &= <a\,,\,a> + <a\,,\,b> + <b\,,\,a> + <b\,,\,b> \\ &= \|a\|^2 + \|b\|^2. \end{aligned}$$

Der Name des Satzes rührt daher, dass a, b und $a+b$ die Seiten eines rechtwinkligen Dreiecks bilden (wenn man die Vektoren durch geeignete Pfeile repräsentiert). ∎

4.7 Cauchy-Schwarz'sche Ungleichung. *Sind $a, b \in V$, so ist*

$$|<a\,,\,b>| \leq \|a\| \cdot \|b\|.$$

BEWEIS: Wir können voraussetzen, dass a und b beide $\neq 0$ sind, denn sonst ist die Aussage trivial. Der Satz von der Existenz der orthogonalen Projektion besagt dann, dass es ein $\lambda \in K$ und einen Vektor c gibt, so dass gilt:

$$<a\,,\,c> = 0 \quad \text{und} \quad b = \lambda a + c, \text{ mit } \lambda = \frac{<b\,,\,a>}{<a\,,\,a>}.$$

Nach dem Satz des Pythagoras ist

$$\|b\|^2 = |\lambda|^2 \cdot \|a\|^2 + \|c\|^2 \geq |\lambda|^2 \cdot \|a\|^2 = \frac{|<b\,,\,a>|^2}{\|a\|^2} = \frac{|<a\,,\,b>|^2}{\|a\|^2}.$$

Das ergibt die Ungleichung $\|a\|^2 \cdot \|b\|^2 \geq |<a\,,\,b>|^2$. Zieht man auf beiden Seiten die Wurzel, so erhält man die Schwarz'sche Ungleichung. ■

4.8 Dreiecks-Ungleichung.

$$\|a + b\| \leq \|a\| + \|b\|.$$

BEWEIS: Ist $z = x + \mathrm{i}\,y \in \mathbb{C}$, so ist $z + \overline{z} = 2x = 2\,\mathrm{Re}(z)$, also

$$\begin{aligned} <a+b\,,\,a+b> &= <a\,,\,a> + 2\,\mathrm{Re}<a\,,\,b> + <b\,,\,b> \\ &\leq \|a\|^2 + 2 \cdot |<a\,,\,b>| + \|b\|^2 \\ &\leq \|a\|^2 + 2 \cdot \|a\| \cdot \|b\| + \|b\|^2 \\ &= (\|a\| + \|b\|)^2. \end{aligned}$$

Wurzelziehen auf beiden Seiten ergibt die Behauptung. ■

Definition. Unter einem *Orthonormalsystem* in V versteht man ein System von Vektoren $\{a_1, \ldots, a_n\}$, so dass gilt:

1. Die Vektoren sind paarweise orthogonal: $<a_i\,,\,a_j> = 0$ für $i \neq j$.

2. Die Vektoren sind normiert: $\|a_i\| = 1$ für alle i.

Führt man das *Kroneckersymbol* ein, $\delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst.} \end{cases}$,
so kann man die beiden Eigenschaften zusammenfassen zu $<a_i\,,\,a_j> = \delta_{ij}$.

Beispiele.

1. Ist n beliebig, so bilden die Einheitsvektoren $e_i = (0, \ldots, 1, \ldots, 0)^t$, $i = 1, \ldots, n$, bezüglich des kanonischen Skalarproduktes ein Orthonormalsystem (im $\mathbb{R}^n$ und im $\mathbb{C}^n$).

2. Sei $\mathbf{x} = (x_1, x_2)^t$ ein beliebiger Vektor $\neq (0,0)^t$ im $\mathbb{R}^2$. Verwenden wir das kanonische Skalarprodukt, so ist

$$\mathbf{a} := \frac{1}{\|\mathbf{x}\|} \cdot \mathbf{x} = \Big(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \frac{x_2}{\sqrt{x_1^2 + x_2^2}}\Big)^t$$

   ein normierter Vektor. Schreibt man $\mathbf{a} = (a_1, a_2)^t$, so steht $\mathbf{b} := (-a_2, a_1)^t$ auf $\mathbf{a}$ senkrecht und ist natürlich ebenfalls normiert. Also bildet $\{\mathbf{a}, \mathbf{b}\}$ ein Orthonormalsystem im $\mathbb{R}^2$.

4.9 Satz. *Bilden die Vektoren $\{a_1, \ldots, a_k\}$ ein Orthonormalsystem, so sind sie linear unabhängig.*

Beweis: Multipliziert man die Gleichung $\alpha_1 a_1 + \cdots + \alpha_k a_k = 0$ (per Skalarprodukt) mit einem a_j, so erhält man $\alpha_j = 0$. ∎

4.10 Satz. *Ist $\dim_K(V) = n$ und $\{a_1, \ldots, a_n\}$ ein Orthonormalsystem in V, so bilden die Vektoren eine Basis von V.*

Beweis: Die Vektoren bilden ein maximales linear unabhängiges System. ∎

4.11 Folgerung. *Ist $\dim_K(V) = n$ und $\{a_1, \ldots, a_n\}$ ein Orthonormalsystem in V, sowie $x \in V$ ein beliebiger Vektor, so ist*

$$x = <x\,,\, a_1> \cdot a_1 + \cdots + <x\,,\, a_n> \cdot a_n\,.$$

Beweis: Es gibt reelle Zahlen α_i mit $x = \alpha_1 a_1 + \cdots + \alpha_n a_n$. Multipliziert man mit a_j, so erhält man $\alpha_j = <x\,,\, a_j>$. ∎

Eine *Orthonormalbasis* (kurz: *ON-Basis*) ist ein maximales Orthonormalsystem.

Beispiele.

1. Die Standardbasis $\{e_1, \ldots, e_n\}$ ist eine ON-Basis des $\mathbb{R}^n$.

2. Die Vektoren $\mathbf{a}_1 := (1,1)^t$ und $\mathbf{a}_2 := (1,-1)^t$ sind bezüglich des kanonischen Skalarproduktes orthogonal, aber nicht normiert.

   Die Vektoren $\mathbf{b}_1 := \dfrac{1}{\sqrt{2}}(1,1)^t$ und $\mathbf{b}_2 := \dfrac{1}{\sqrt{2}}(1,-1)^t$ bilden eine ON-Basis.

4.12 Schmidt'sches Orthogonalisierungsverfahren. *Sei V ein euklidischer oder unitärer K-Vektorraum, $U \subset V$ ein k-dimensionaler Unterraum und $(x_1, \ldots, x_k)$ eine Basis von U.*

Dann besitzt U eine ON-Basis $(a_1, \ldots, a_k)$, so dass gilt:

$$\mathrm{Lin}(a_1, \ldots, a_\varrho) \;=\; \mathrm{Lin}(x_1, \ldots, x_\varrho) \text{ für } \varrho = 1, \ldots, k.$$

Beweis: Wir konstruieren die a_i rekursiv aus den x_i.

Sei $a_1 := x_1/\|x_1\|$. Dann ist $\|a_1\| = 1$ und $\text{Lin}(a_1) = \text{Lin}(x_1)$.

Nun nehmen wir an, wir hätten $a_1, \dots, a_\varrho$ schon mit den gewünschten Eigenschaften konstruiert. Dann sei

$$b_{\varrho+1} := x_{\varrho+1} - \sum_{i=1}^{\varrho} <x_{\varrho+1}\,,\, a_i>a_i.$$

Es gilt:

1. $b_{\varrho+1} \in \text{Lin}(x_1, \dots, x_{\varrho+1})$.
2. Für $j = 1, \dots, \varrho$ ist $<b_{\varrho+1}\,,\, a_j> = <x_{\varrho+1}\,,\, a_j> - <x_{\varrho+1}\,,\, a_j> = 0$.
3. $a_1, \dots, a_\varrho, b_{\varrho+1}$ sind paarweise orthogonal und daher linear unabhängig.

Schließlich setzen wir $a_{\varrho+1} := \dfrac{1}{\|b_{l+1}\|} \cdot b_{\varrho+1}$. Dann bilden $a_1, \dots, a_\varrho, a_{\varrho+1}$ ein ON-System mit $\text{Lin}(a_1, \dots, a_\varrho, a_{l\varrho+1}) = \text{Lin}(x_1, \dots, x_{\varrho+1})$. ∎

Beispiel.

Die Vektoren $\mathbf{x}_1 = (1,1,0)^t$ und $\mathbf{x}_2 = (0,1,0)^t$ bilden eine Basis des Untervektorraumes $U := \{\mathbf{x} \in \mathbb{R}^3 \mid x_3 = 0\}$. Das Schmidt'sche Verfahren liefert:

$$\begin{aligned} \mathbf{a}_1 &= \frac{1}{\sqrt{2}}(1,1,0)^t, \\ \mathbf{b}_2 &= \mathbf{x}_2 - (\mathbf{x}_2 \bullet \mathbf{a}_1)\mathbf{a}_1 \;=\; (-\frac{1}{2}, \frac{1}{2}, 0)^t \\ \text{Dann ist}\quad \|\mathbf{b}_2\| &= \frac{1}{\sqrt{2}}, \quad \text{und wir setzen} \\ \mathbf{a}_2 &= \sqrt{2} \cdot \mathbf{b}_2 \;=\; \frac{1}{\sqrt{2}}(-1,1,0)^t. \end{aligned}$$

Definition. Sei V ein euklidischer oder unitärer K-Vektorraum und $U \subset V$ ein Unterraum. Dann nennt man

$$U^\perp := \{x \in V \;:\; <x\,,\, y> = 0 \text{ für alle } y \in U\}$$

das *orthogonale Komplement* von U.

(Ist $T \subset W$ eine beliebige Teilmenge, so definiert man analog $T^\perp := \{x \in V \;:\; <x\,,\, y> = 0 \text{ für alle } y \in T\}$).

4.13 Satz. *Ist $U \subset V$ ein Untervektorraum, so ist auch $U^\perp \subset V$ ein Untervektorraum.*

BEWEIS: 1) $<0\,,\, y> = 0$ für alle $y \implies 0 \in U^\perp$.
2) $x, x' \in U^\perp \implies <x + x'\,,\, y> = <x\,,\, y> + <x'\,,\, y> = 0 + 0 = 0$ für alle $y \in U$,

also $x + x' \in U^\perp$.
3) $\alpha \in K$, $x \in U^\perp \implies <\alpha x\,,\, y> = \alpha<x\,,\, y> = \alpha \cdot 0 = 0$ für alle $y \in U$, also $\alpha x \in U^\perp$. $\blacksquare$

Es gibt eine Verallgemeinerung des Satzes von der orthogonalen Projektion:

4.14 Satz. *Sei $U \subset V$ endlich-dimensional und $x \in V$ beliebig. Dann gibt es eindeutig bestimmte Vektoren $u \in U$ und $v \in U^\perp$, so dass $x = u + v$ ist.*

Beweis: 1) Eindeutigkeit: Sei $x = u + v = u' + v'$ mit $u, u' \in U$ und $v, v' \in U^\perp$. Dann ist $u - u' = v' - v$, also $<u - u'\,,\, u - u'> = <u - u'\,,\, v' - v> = 0$. Daraus folgt $u - u' = 0$, also $u = u'$ und $v = v'$.

2) Existenz: Sei $(a_1, \ldots, a_k)$ eine ON-Basis von U. Dann setze

$$u := \sum_{i=1}^{k} <x\,,\, a_i>a_i \qquad \text{und} \qquad v := x - u.$$

Dann ist $u \in U$ und $<v\,,\, a_j> = <x\,,\, a_j> - <u\,,\, a_j> = <x\,,\, a_j> - <x\,,\, a_j> = 0$ für alle j, also $v \in U^\perp$. $\blacksquare$

Den eindeutig bestimmten Vektor

$$p_U(x) := u = \sum_{i=1}^{k} <x\,,\, a_i>a_i$$

nennt man die *orthogonale Projektion* von x auf U.

4.15 Satz. *Die orthogonale Projektion $p_U : V \to U$ ist linear und surjektiv, mit $p_U \circ p_U = p_U$ und $\mathrm{Ker}(p_U) = U^\perp$.*

Beweis: 1) Linearität: Sei $x = u+v$ und $x' = u'+v'$, mit $u, u' \in U$ und $v, v' \in U^\perp$. Dann ist $x + x' = (u + u') + (v + v')$ mit $u + u' \in U$ und $v + v' \in U^\perp$. Es folgt:

$$p_U(x + x') = u + u' = p_U(x) + p_U(x').$$

Analog zeigt man, dass $p_U(\alpha x) = \alpha \cdot p_U(x)$ ist.

2) Surjektivität: Ist $u \in U$, so ist auch $u \in V$ und man hat die Zerlegung $u = u+0$. Also ist $p_U(u) = u$.

3) Ist $x = u + v$, so ist $p_U \circ p_U(x) = p_U(u) = p_U(u + 0) = u = p_U(x)$.

4) Kern: Es ist $p_U(u + v) = 0 \iff u = 0 \iff u + v = v \in U^\perp$. $\blacksquare$

4.16 Satz (Eigenschaften des orthogonalen Komplements). *Sei V endlich-dimensional. Dann gilt:*

1. *Ist $U \subset V$ ein Unterraum, so ist*

$$\dim(U^\perp) = \dim(V) - \dim(U).$$

2. Es ist stets $U^{\perp\perp} = U$.

BEWEIS: 1) folgt aus der Dimensionsformel.

2) Offensichtlich ist $U \subset U^{\perp\perp}$. Wegen der Gleichung

$$\dim U^{\perp\perp} = \dim V - \dim U^{\perp} = \dim U$$

folgt die Gleichheit. ∎

Der folgende Satz wurde nicht in der Vorlesung behandelt:

Ist U nicht endlich-dimensional, so braucht die orthogonale Projektion nicht zu existieren. Wenn sie aber existiert, ist sie eindeutig bestimmt, und es gilt:

4.17 Satz. *Sei $U \subset V$ ein beliebiger Unterraum und $x \in V$. Ein Vektor $u \in U$ ist genau dann die orthogonale Projektion von x auf U, wenn $\|x - u\| \leq \|x - u'\|$ für alle $u' \in U$ gilt.*

BEWEIS: 1) Sei $u = p_U(x)$. Dann ist $x - u \in U^{\perp}$. Ist $u' \in U$ und $v := u - u' \in U$, so ist $x - u' = (x - u) + (u - u') = (x - u) + v$, also

$$\|x - u'\|^2 = \|x - u\|^2 + \|v\|^2 \geq \|x - u\|^2.$$

2) Sei $u \in U$ und $\|x - u\| \leq \|x - u'\|$ für alle $u' \in U$. Wir müssen nur zeigen, dass $x - u \in U^{\perp}$ ist.

Annahme, $x - u \notin U^{\perp}$. Dann gibt es ein $v \in U$ mit $\|v\| = 1$ und $c := <x - u\,, v> \neq 0$. Sei $u' := u + cv \in U$, also $x - u' = (x - u) - cv$. Es folgt:

$$\begin{aligned} \|x - u'\|^2 &= <(x - u) - cv\,, (x - u) - cv> \\ &= \|x - u\|^2 + |c|^2 - 2\,\mathrm{Re}<x - u\,, cv> \\ &= \|x - u\|^2 + |c|^2 - 2|c|^2 \\ &= \|x - u\|^2 - |c|^2 \;<\; \|x - u\|^2. \end{aligned}$$

Das ist ein Widerspruch. ∎

Kapitel 3 Determinanten und Koordinaten

§ 1 Determinanten

Unter einer *Permutation* versteht man eine bijektive Abbildung von $\{1, \dots, n\}$ auf sich. Wir schreiben eine Permutation σ in der Form

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

Die Werte $\sigma(1), \dots, \sigma(n)$ bestimmen – in dieser Anordnung – die Abbildung σ. Es gibt genau $n!$ Permutationen der Zahlen $1, \dots, n$. Wir bezeichnen die Menge aller dieser Permutationen mit S_n. Beispiel:

$$S_3 = \{\binom{123}{123}, \binom{123}{132}, \binom{123}{213}, \binom{123}{231}, \binom{123}{312}, \binom{123}{321}\}.$$

Wird n groß, so wird der Umgang mit den Permutationen aus S_n kompliziert. Eine gewisse Hilfe ist dann die „Zykel-Schreibweise“. Beispiel:

$$\text{Sei } \sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} \in S_7.$$

Wir beginnen mit der 1, sie wird auf die 3 abgebildet, die 3 auf die 4, die 4 auf die 5 und die 5 wieder auf die 1. Das ergibt einen abgeschlossenen *Zykel*, den man mit $(1, 3, 4, 5)$ bezeichnet. Damit ist aber die Abbildung σ noch nicht abgehandelt. Die nächste noch nicht berücksichtigte Zahl ist die 2. Mit ihr beginnen wir das Spiel erneut: Die 2 wird auf die 7 abgebildet, die 7 auf die 6 und die 6 wieder auf die 2. Das ergibt den Zykel $(2, 7, 6)$. Da nun alle Zuordnungen von σ berücksichtigt wurden, schreiben wir:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} = (1, 3, 4, 5)(2, 7, 6).$$

Wird eine Zahl i auf sich abgebildet, so ergibt das einen „Einer-Zykel“ (i). Solche Zykel schreibt man normalerweise gar nicht hin. Wenn allerdings σ die Identität ist, so besteht σ nur aus Einer-Zyklen, und man muss wenigstens einen davon hinschreiben:

$$\mathrm{id}_{\{1,\dots,n\}} = (1).$$

Sind $\sigma := \begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix}$ und $\tau := \begin{pmatrix} 1 & \cdots & n \\ \tau(1) & \cdots & \tau(n) \end{pmatrix}$ zwei Permutationen, so kann man sie zu einer neuen Permutation $\sigma \circ \tau$ miteinander verknüpfen, mit

$$\sigma \circ \tau = \begin{pmatrix} 1 & \cdots & n \\ \sigma(\tau(1)) & \cdots & \sigma(\tau(n)) \end{pmatrix}.$$

So ist z.B.

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

Man beachte, dass die Permutationen **von rechts nach links** abzuarbeiten sind! Die rechte (innere) Permutation bildet z.B. die 1 auf die 3 ab, und die linke (äußere) Permutation bildet dann die 3 auf die 2 ab. Im Ergebnis auf der anderen Seite der Gleichung muss deshalb unter der 1 die 2 stehen.

Die Menge S_n aller Permutationen von $\{1,\ldots,n\}$ bildet offensichtlich eine Gruppe. Normalerweise ist die Verknüpfung von Permutationen nicht kommutativ. Wenn wir allerdings die Menge $\{1,\ldots,n\}$ in zwei disjunkte Teilmengen M und N aufteilen können, so dass die Permutation σ nur die Zahlen aus M verändert und τ nur die Zahlen aus N, dann spielt die Reihenfolge keine Rolle, es ist $\sigma \circ \tau = \tau \circ \sigma$. Das hat Konsequenzen für die Zykel-Schreibweise. Die Aufteilung einer Permutation in mehrere (disjunkte) Zykel bedeutet, dass sich die gegebene Permutation aus mehreren einfacheren Permutationen verknüpfen lässt, und die Reihenfolge spielt dabei keine Rolle. Es ist also z.B.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} = (1,3,4,5)(2,7,6) = (2,7,6)(1,3,4,5).$$

Eine Permutation heißt *Transposition* oder *Vertauschung*, wenn nur zwei Zahlen miteinander vertauscht werden und alle anderen fest bleiben, wenn sie also nur aus einem „Zweier-Zykel“ besteht, z.B.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 1 & 5 \end{pmatrix} = (1,4)$$

1.1 Satz. *Jede Permutation lässt sich als endliche Verknüpfung von Transpositionen schreiben.*

BEWEIS: Da wir jede Permutation als Folge von Zyklen schreiben können, reicht es zu zeigen, dass sich jeder beliebige Zykel aus endlich vielen Zweier-Zyklen zusammensetzen lässt.

Betrachten wir also einen beliebigen Zykel $(i_1, i_2, \ldots, i_k)$ in einer Permutation σ.
Behauptung: $(i_1, i_2, \ldots, i_k) = (i_1, i_k) \circ \ldots \circ (i_1, i_3) \circ (i_1, i_2)$.

BEWEIS dazu mit Induktion nach k. Ist $k = 2$, so ist nichts zu zeigen. Die Behauptung sei nun für ein $k \geq 2$ schon bewiesen. Wir untersuchen den Fall $k + 1$.

$(i_1, \ldots, i_k)$ bildet i_1 auf i_2 ab, i_2 auf i_3 usw. Schließlich wird i_k auf i_1 abgebildet. Verknüpfen wir den Zykel mit der Transposition (i_1, i_{k+1}), so wird anschließend i_k auf i_{k+1} und i_{k+1} auf i_1 abgebildet. Also ist

$$\begin{aligned} (i_1, \ldots, i_k, i_{k+1}) &= (i_1, i_{k+1}) \circ (i_1, \ldots, i_k) \\ &= (i_1, i_{k+1}) \circ (i_1, i_k) \circ \ldots \circ (i_1, i_2). \end{aligned}$$

Damit ist alles gezeigt. ∎

Man beachte, dass es hier auf die Reihenfolge ankommt, weil die einzelnen Zweier-Zyklen nicht paarweise disjunkt sind!

Das oben angegebene Verfahren zur Auflösung einer Permutation in Transpositionen liefert ein eindeutiges Ergebnis. Leider sind auch andere Verfahren denkbar, und die Zerlegung einer Permutation in Transpositionen ist i.a. nicht eindeutig bestimmt.

Definition. Ist $\sigma \in S_n$, so nennt man die Zahl

$$\operatorname{sign}(\sigma) := \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}$$

das *Signum* von σ.
Oberhalb und unterhalb des Bruchstriches stehen die gleichen numerischen Werte, nur in anderer Reihenfolge (was keine Rolle spielt) und mit anderen Vorzeichen: Unten stehen nur positive Zahlen, oben kommen Vorzeichen hinzu:

$$s_{i,j}(\sigma) := \begin{cases} +1 & \text{falls } \sigma(i) < \sigma(j) \\ -1 & \text{falls } \sigma(i) > \sigma(j)\,. \end{cases}$$

Also ist $\operatorname{sign}(\sigma) = \prod\limits_{i<j} s_{i,j}(\sigma) = \pm 1$.

Ist z.B. τ eine Transposition mit $\tau(i) = i+1$, $\tau(i+1) = i$ und $\tau(k) = k$ für alle anderen k, so ist $s_{i,i+1}(\tau) = -1$ und $s_{k,l}(\tau) = +1$ in allen anderen Fällen, also insgesamt $\operatorname{sign}(\tau) = -1$.

1.2 Satz (Multiplikativität des Signums). *Für $\sigma, \tau \in S_n$ ist*

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau).$$

BEWEIS: Ist $k \neq l$, so ist $\dfrac{\sigma(k) - \sigma(l)}{k - l} = \dfrac{\sigma(l) - \sigma(k)}{l - k}$. Da bei der Berechnung des Signum-Produktes jedes Paar (k, l) mit $k \neq l$ genau einmal vorkommt, folgt:

$$\begin{aligned} \operatorname{sign}(\sigma \circ \tau) &= \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\ &= \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{i<j} \frac{\tau(j) - \tau(i)}{j - i} \\ &= \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau)\,. \end{aligned}$$

■

Da jede Permutation Produkt von Transpositionen ist, können wir das Signum nun sehr einfach berechnen, und insbesondere folgt:

1.3 Satz. *Auch wenn eine Permutation auf verschiedene Weisen durch eine Folge von Transpositionen erzeugt wird, so ist die Anzahl der dabei benötigten Vertauschungen modulo 2 eindeutig bestimmt.*

BEWEIS: Ist $\sigma = \tau_1 \circ \ldots \circ \tau_k$, so ist $\text{sign}(\sigma) = (-1)^k$. ∎

Eine Permutation $\sigma \in S_n$ heißt *gerade*, wenn sie sich aus einer geraden Anzahl von Vertauschungen zusammensetzt. Andernfalls heißt sie *ungerade*. Ist σ gerade, so ist $\text{sign}(\sigma) = +1$, andernfalls $= -1$. So ist z.B.

$$\begin{aligned}
\text{sign}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix} &= \text{sign}(3,4) \;=\; -1, \\
\text{sign}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 2 & 5 \end{pmatrix} &= \text{sign}\big((1,3) \circ (2,4)\big) \;=\; +1 \\
\text{und} \quad \text{sgn}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix} &= \text{sign}\big((1,3,4) \circ (2,5)\big) \\
&= \text{sign}\big((1,4) \circ (1,3) \circ (2,5)\big) \;=\; -1.
\end{aligned}$$

Für 2-reihige Matrizen haben wir schon Determinanten eingeführt:

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Man kann die Determinante als Funktion $\det : K^2 \times K^2 \to K$ auffassen, die auf den Spalten der Matrix operiert. Dann gilt:

1. $\det : (x, y) \mapsto \det(x, y)$ ist bilinear.
2. $\det(x, x) = 0$ für alle x.
3. $\det(e_1, e_2) = 1$.

BEWEIS: 1) Es ist

$$\begin{aligned}
\det\begin{pmatrix} a + a' & b \\ c + c' & d \end{pmatrix} &= (a + a')d - b(c + c') \;=\; (ad - bc) + (a'd - bc') \\
&= \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det\begin{pmatrix} a' & b \\ c' & d \end{pmatrix}
\end{aligned}$$

und

$$\det\begin{pmatrix} \alpha a & b \\ \alpha c & d \end{pmatrix} = (\alpha a)d - b(\alpha c) \;=\; \alpha(ad - bc) \;=\; \alpha \cdot \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Das liefert die Linearität im 1. Argument. Der Beweis der Linearität im 2. Argument geht genauso.

2) $\det\begin{pmatrix} a & a \\ c & c \end{pmatrix} = ac - ac = 0$, also $\det(x, x) = 0$ für alle x.

3) Es ist $\det(e_1, e_2) = \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$ ∎

Wir wollen nun für beliebiges n eine Funktion

$$\det : M_n(K) = K^n \times \ldots \times K^n \to K$$

suchen, die folgende Eigenschaften besitzt:

D1) det ist in jeder einzelnen Spalte linear, d.h..

$$\begin{aligned} \det(a_1, \ldots, a_i + a_i', \ldots, a_n) &= \det(a_1, \ldots, a_i, \ldots, a_n) \\ &\quad + \det(a_1, \ldots, a_i', \ldots, a_n), \\ \text{und} \quad \det(a_1, \ldots, \alpha a_i, \ldots, a_n) &= \alpha \cdot \det(a_1, \ldots, a_i, \ldots, a_n). \end{aligned}$$

Man sagt dann auch, det ist *n-fach multilinear.*

D2) Sind zwei der Argumente gleich, so ist $\det(a_1, \ldots, a_n) = 0$. Man sagt dann auch, det ist alternierend.

D3) $\det(e_1, \ldots, e_n) = 1$.

Oben wurde schon bewiesen, dass die Determinante im Falle $n = 2$ die Eigenschaften D1, D2 und D3 besitzt.

1.4 Satz. *Sei K ein beliebiger Körper, V ein K-Vektorraum und $f : V \times V \to K$ eine bilineare Funktion.*

1. *Ist $f(x, x) = 0$ für alle $x \in V$, so ist $f(y, x) = -f(x, y)$ für $x, y \in V$.*
2. *Ist $1 + 1 \neq 0$ in K und $f(y, x) = -f(x, y)$ für alle $x, y \in V$, so ist auch $f(x, x) = 0$ für alle $x \in V$.*

Beweis: 1) Sei $f(x, x) = 0$ für alle x. Dann gilt:

$$\begin{aligned} f(x, y) + f(y, x) &= f(x, x) + f(x, y) + f(y, x) + f(y, y) \\ &= f(x, x + y) + f(y, x + y) \\ &= f(x + y, x + y) \;=\; 0\,. \end{aligned}$$

2) Sei $2 := 1 + 1 \neq 0$ in K. Ist $f(y, x) = -f(x, y)$, so ist $2 \cdot f(x, x) = 0$, also auch $f(x, x) = 0$. ∎

Sei $f : K^n \times \ldots \times K^n \to K$ multilinear. Man sagt, f ist *schiefsymmetrisch*, falls für jede Permutation $\sigma \in S_n$ gilt:

$$f(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = \operatorname{sign}(\sigma) \cdot f(a_1, \ldots, a_n).$$

Da jede Permutation in eine Folge von Transpositionen zerlegt werden kann, folgt wie oben: Ist f alternierend, so ist f auch schiefsymmetrisch. Ist $1 + 1 \neq 0$ in K, so folgt aus „schiefsymmetrisch“ auch „alternierend“.

Bemerkung. Da $1 = -1$ in $\mathbb{Z}_2$ ist, folgt, dass $f : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2$ mit $f(x, y) := xy$ zwar schiefsymmetrisch, aber nicht alternierend ist.

1.5 Satz von der Existenz und Eindeutigkeit der Determinante.
Es gibt genau eine Funktion $\det : K^n \times \ldots \times K^n \to K$ *mit den Eigenschaften (D1), (D2) und (D3).*

BEWEIS: Wir beginnen mit dem Beweis der Eindeutigkeit, in der Hoffnung, dass sich daraus auch eine Formel für die Existenz ergibt.

Die Spaltenvektoren $a_1, \ldots, a_n \in K^n$ fassen wir als Spalten einer Matrix

$$A = (a_1, \ldots, a_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_n(K)$$

auf. Dann ist

$$a_j = \sum_{i=1}^{n} a_{ij} e_i, \text{ für } j = 1, \ldots, n.$$

Und nun beginnt die Schwierigkeit: Da wir es mit n Summen zu tun haben, brauchen wir auch n verschiedene Summationsindizes. In Ermanglung geeigneter Buchstaben (und weil die Schreibweise dann auch zu schwerfällig wäre) benutzen wir indizierte Indizes:

$$a_1 = \sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \quad \ldots \quad, a_n = \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}.$$

Wir nehmen nun an, wir hätten schon eine Determinantenfunktion, und setzen diese Vektoren ein. Dann folgt mit der Multilinearität:

$$\begin{aligned} \det(a_1, \ldots, a_n) &= \det\Big(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\Big) \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1 1} \cdot \ldots \cdot a_{i_n n} \det(e_{i_1}, \ldots, e_{i_n}). \end{aligned}$$

Nun ist

$$\det(e_{i_1}, \ldots, e_{i_n}) = \begin{cases} 0 & \text{falls ein Argument doppelt auftritt,} \\ \operatorname{sign}(\sigma) & \text{falls es ein } \sigma \in S_n \text{ gibt, s.d.} \\ & (i_1, \ldots, i_n) = (\sigma(1), \ldots, \sigma(n)) \text{ ist.} \end{cases}$$

Es fallen also alle Summanden weg, bis auf diejenigen, bei denen

$$(i_1, \ldots, i_n) = \big(\sigma(1), \ldots, \sigma(n)\big) \text{ für ein } \sigma \in S_n \text{ ist.}$$

Damit erhalten wir:

$$\boxed{\det(a_1,\ldots,a_n) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1}\cdot\ldots\cdot a_{\sigma(n),n}\,.}$$

Die Determinante muss zwangläufig so aussehen! Das liefert uns die Eindeutigkeit, und auch die Existenz, wenn wir zeigen können, dass der gefundene Ausdruck (den wir vorübergehend mit $D(a_1,\ldots,a_n)$ bezeichnen wollen) die Bedingungen (D1), (D2) und (D3) erfüllt.

Wir beginnen mit der Multilinearität (D1) und beschränken uns dabei aus schreibtechnischen Gründen auf das erste Argument:

$$\begin{aligned}
&D(a_1'+a_1'',a_2,\ldots,a_n) = \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(a'_{\sigma(1),1}+a''_{\sigma(1),1})\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n}\\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(a'_{\sigma(1),1}\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n}+a''_{\sigma(1),1}\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n})\\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a'_{\sigma(1),1}\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n}+\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a''_{\sigma(1),1}\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n}\\
&= D(a_1',a_2,\ldots,a_n)+D(a_1'',a_2,\ldots,a_n).
\end{aligned}$$

Genauso geht es bei den anderen Argumenten. Weiter ist

$$\begin{aligned}
D(\alpha\cdot a_1,a_2,\ldots,a_n) &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(\alpha\cdot a_{\sigma(1),1})\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n}\\
&= \alpha\cdot\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a_{\sigma(1),1}\cdot a_{\sigma(2),2}\cdot\ldots\cdot a_{\sigma(n),n}\\
&= \alpha\cdot D(a_1,a_2,\ldots,a_n).
\end{aligned}$$

Nun nehmen wir an, dass $a_1 = a_2$ ist. $\tau\in S_n$ sei die Permutation, die lediglich 1 mit 2 vertauscht. Dann gilt $\operatorname{sign}(\sigma\circ\tau) = -\operatorname{sign}(\sigma)$ für jedes $\sigma\in S_n$.

Wir definieren

$$\begin{aligned}
&S_n^+ &:=& \{\sigma\in S_n : \operatorname{sign}(\sigma)=1\}\\
\text{und}\quad &S_n^- &:=& \{\sigma\in S_n : \operatorname{sign}(\sigma)=-1\}.
\end{aligned}$$

Offensichtlich ist $S_n^+\cup S_n^- = S_n$ und $S_n^+\cap S_n^- = \varnothing$. Durch $\sigma\mapsto\sigma\circ\tau$ wird eine bijektive Abbildung von S_n^+ auf S_n^- definiert. Daher gilt:

$$\begin{aligned}
D(a_1, a_1, \ldots, a_n) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n^+} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&\quad + \sum_{\sigma \in S_n^-} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n^+} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&\quad + \sum_{\sigma \in S_n^+} \operatorname{sign}(\sigma \circ \tau) a_{\sigma\circ\tau(1),1} \cdot a_{\sigma\circ\tau(2),1} \cdot \ldots \cdot a_{\sigma\circ\tau(n),n} \\
&= \sum_{\sigma \in S_n^+} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&\quad - \sum_{\sigma \in S_n^+} \operatorname{sign}(\sigma) a_{\sigma(2),1} \cdot a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n} \;=\; 0.
\end{aligned}$$

Damit ist die Funktion D auch alternierend.

Wir kommen nun zu (D3): Dabei verwenden wir das Kronecker–Symbol δ_{ij}. Offensichtlich ist δ_{ij} die i-te Komponente des Einheitsvektors e_j, und es gilt:

$$\delta_{\sigma(1),1} \cdot \ldots \cdot \delta_{\sigma(n),n} = \begin{cases} 0 & \text{falls } \sigma \neq \mathrm{id} \\ 1 & \text{falls } \sigma = \mathrm{id} \end{cases}$$

Daraus folgt:

$$\begin{aligned}
D(e_1, \ldots, e_n) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \delta_{\sigma(1),1} \cdot \ldots \cdot \delta_{\sigma(n),n} \\
&= \operatorname{sign}(\mathrm{id}) \;=\; 1.
\end{aligned}$$

Damit ist tatsächlich alles bewiesen! ∎

Bemerkung. Man schreibt auch

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \quad \text{statt} \quad \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

1.6 Satz. *Sei $A \in M_n(K)$. Dann ist* $\det(A^t) = \det(A)$.

BEWEIS:
1) Die Zuordnung $\sigma \mapsto \sigma^{-1}$ liefert eine bijektive Abbildung $S_n \to S_n$.

2) Für alle $\sigma \in S_n$ ist $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$. Man braucht ja nur alle Vertauschungen rückgängig zu machen.

3) σ und τ seien Elemente von S_n. Dann gilt wegen der Kommutativität der Multiplikation:

$$a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} = a_{\tau(1),\sigma\circ\tau(1)} \cdot \ldots \cdot a_{\tau(n),\sigma\circ\tau(n)}.$$

Unter Verwendung von (1) und (2), sowie (3) im Falle $\tau = \sigma^{-1}$ erhält man (mit $b_{ij} := a_{ji}$):

$$\begin{aligned}
\det(A^t) &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) b_{\sigma(1),1} \cdot \ldots \cdot b_{\sigma(n),n} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdot \ldots \cdot a_{\sigma^{-1}(n),n} \\
&= \sum_{\sigma^{-1}\in S_n} \operatorname{sign}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdot \ldots \cdot a_{\sigma^{-1}(n),n} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n} \;=\; \det(A).
\end{aligned}$$

∎

Beispiel.

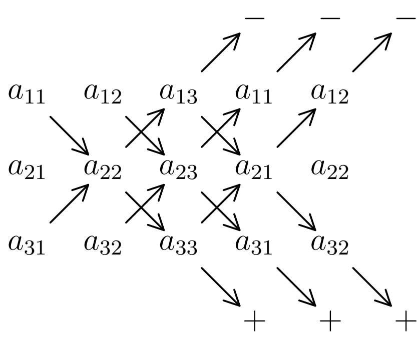
Ist $n = 3$ und $a_j = (a_{1,j}, a_{2,j}, a_{3,j})^t$ für $j = 1, 2, 3$, so ist

$$\begin{aligned}
\det(a_1, a_2, a_3) &= \sum_{\sigma\in S_3} \operatorname{sign}(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot a_{\sigma(3),3} \\
&= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\
&\quad - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}.
\end{aligned}$$

Die so gewonnene Berechnungsformel heißt die *Sarrus'sche Regel*. Es gibt ein Schema, nach dem man sie sich leicht merken kann. Man schreibe die Matrix–Elemente in der folgenden Form auf:

$$\begin{array}{ccc|cc}
a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\
a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\
a_{31} & a_{32} & a_{33} & a_{31} & a_{32}
\end{array}$$

Die Produkte der Elemente in den „Hauptdiagonalen“ werden mit „+“ versehen, die Produkte der Elemente in den „Nebendiagonalen“ werden mit „−“ versehen, und schließlich werden alle Produkte aufsummiert:



Wir fassen zusammen:

1.7 Rechenregeln für Determinanten. *Sei $A \in M_n(K)$.*

1. *Multipliziert man in A eine Zeile oder eine Spalte mit $\lambda \in K$, so muss man auch $\det(A)$ mit λ multiplizieren.*

   Insbesondere ist $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$.

2. *Addiert man das Vielfache einer Zeile (bzw. Spalte) zu einer anderen Zeile (bzw. Spalte), so ändert sich der Wert der Determinante nicht.*

3. *Vertauscht man zwei Zeilen oder Spalten, so wechselt das Vorzeichen der Determinante.*

4. *Ist $\operatorname{rg}(A) < n$, so ist $\det(A) = 0$.*

5. *Es ist $\det(A^t) = \det(A)$.*

6. *Ist $f : M_n(K) \to K$ eine Funktion, die multilinear und alternierend in den Zeilen (bzw. Spalten) von A ist, mit $f(E_n) = 1$, so ist $f(A) = \det(A)$.*

BEWEIS: 1), 2) und 3) folgen daraus, dass det multilinear und alternierend ist.

4) Ist z.B. $a_1 = \alpha_2 a_2 + \cdots + \alpha_n a_n$, so ist

$$\det(a_1, \ldots, a_n) = \sum_{i=2}^{n} \alpha_i \det(a_i, a_2, \ldots, a_n) = 0.$$

5) schon bewiesen.

6) schon bewiesen. ∎

§ 2 Der Entwicklungssatz

Durch elementare Umformungen kann man jede Matrix $A \in M_n(K)$ auf Dreiecksgestalt bringen. Da man dabei die Änderung der Determinante gut kontrollieren kann, ist der folgende Satz sehr nützlich:

2.1 Determinante einer Dreiecksmatrix.

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{vmatrix} = a_{11} \cdot \ldots \cdot a_{nn}.$$

BEWEIS: Die Matrix sei mit A bezeichnet. Wir unterscheiden zwei Fälle:

1. Fall: Ist $a_{ii} = 0$ für ein i, so verschwinden beide Seiten der Gleichung: Bei der rechten Seite ist es klar, und die Determinante auf der linken Seite verschwindet, weil $\text{rg}(A) < n$ ist.

3. Fall: $a_{11}, \ldots, a_{nn} \neq 0$.

Durch elementare Zeilenumformungen, die nicht die Determinante verändern, kann man alle Elemente oberhalb der Diagonalen zum Verschwinden bringen. Wenn aber $a_{ij} = 0$ für $i \neq j$ ist, dann ergibt sich die Behauptung sofort aus der Determinantenformel. ■

Der nächste Satz hat sowohl praktische als auch theoretische Bedeutung:

2.2 Determinanten–Produktsatz. *Es seien $A, B \in M_n(K)$. Dann gilt:*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

BEWEIS:
1) Zunächst sei $\text{rg}(B) < n$. Dann ist $\det(A) \cdot \det(B) = 0$.

Da aber $\text{rg}(A \cdot B) \leq \text{rg}(B)$ ist, ist auch $\det(A \cdot B) = 0$.

2) Ist $\text{rg}(B) = n$, so kann man B durch elementare Zeilenumformungen auf obere Dreiecksgestalt bringen, so dass in der Diagonale nur Elemente $\neq 0$ stehen. Die Determinante ändert sich dabei nur um einen Faktor $\neq 0$. Also ist $\det(B) \neq 0$.

Nun sei $\delta : M_n(K) \to K$ definiert durch

$$\delta(A) := \frac{\det(A \cdot B)}{\det(B)}.$$

Man rechnet leicht nach, dass δ multilinear in den Zeilen von A ist. Und wenn A zwei gleiche Zeilen enthält, dann trifft das auch auf $A \cdot B$ zu, so dass $\delta(A) = 0$ ist. Schließlich ist noch $\delta(E_n) = 1$. Aber dann muss $\delta(A) = \det(A)$ sein, und die Produktformel folgt. ■

2.3 Folgerung. *$A \in M_n(K)$ ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist. In diesem Falle ist*

$$\det(A^{-1}) = \frac{1}{\det(A)}\,.$$

BEWEIS: a) Ist A invertierbar, so ist $1 = \det(E_n) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$. Daraus folgt, dass $\det A \neq 0$ ist, sowie die Formel.

b) Ist $\det A \neq 0$, so ist $\mathrm{rg}(A) = n$. Dann ist F_A surjektiv und (aus Dimensionsgründen) sogar ein Isomorphismus, also A invertierbar. ∎

2.4 Kästchensatz. *Sind $A \in M_r(K)$, $B \in M_{r,n-r}(K)$ und $C \in M_{n-r}(K)$, so ist*

$$\det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det(A) \cdot \det(C).$$

BEWEIS: 1) Mit Hilfe von Zeilenoperationen im Bereich der ersten r Zeilen und Spaltenvertauschungen im Bereich der ersten r Spalten kann man A in eine obere Dreiecksmatrix Δ_1 umformen. Es gibt ein $a \neq 0$ und eine Matrix $B^* \in M_{r,n-r}(K)$, so dass gilt:

$$\det(A) = a \cdot \det(\Delta_1) \quad \text{und} \quad \det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = a \cdot \det\begin{pmatrix} \Delta_1 & B^* \\ 0 & C \end{pmatrix}.$$

2) Mit Hilfe von Zeilenoperationen im Bereich der letzten $n-r$ Zeilen und Spaltenvertauschungen im Bereich der letzten $n-r$ Spalten kann man C in eine obere Dreiecksmatrix Δ_2 umformen. Es gibt dann ein $c \neq 0$ und eine Matrix $B^{**} \in M_{r,n-r}(K)$, so dass gilt:

$$\det(C) = c \cdot \det(\Delta_2) \quad \text{und} \quad \det\begin{pmatrix} \Delta_1 & B^* \\ 0 & C \end{pmatrix} = c \cdot \det\begin{pmatrix} \Delta_1 & B^{**} \\ 0 & \Delta_2 \end{pmatrix}.$$

3) Da die Determinante einer oberen Dreiecksmatrix gleich dem Produkt der Diagonalelemente ist, folgt:

$$\begin{aligned} \det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} &= a \cdot c \cdot \det\begin{pmatrix} \Delta_1 & B^{**} \\ 0 & \Delta_2 \end{pmatrix} \\ &= a \cdot c \cdot \det(\Delta_1) \cdot \det(\Delta_2) \\ &= (a \cdot \det(\Delta_1)) \cdot (c \cdot \det(\Delta_2)) \\ &= \det(A) \cdot \det(C). \end{aligned}$$ ∎

Als nächstes wollen wir den allgemeinen Laplace'schen Entwicklungssatz beweisen, der es erlaubt, die Berechnung einer n-reihigen Determinante auf die von $(n-1)$–reihigen zurückzuführen.

Definition. Sei $A = (a_1, \ldots, a_n) \in M_n(K)$. Dann nennt man

$$A_{ij} := \det(a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n)$$

den *Cofaktor* (oder das *algebraische Komplement* oder die *Adjunkte*) von A zur i–ten Zeile und j–ten Spalte.

Mit $S_{ij}(A)$ wird diejenige Matrix bezeichnet, die man aus A durch Streichen der i–ten Zeile und der j–ten Spalte gewinnt. Man nennt sie auch *Streichungsmatrix.*

2.5 Satz (über Adjunkte und Streichungsmatrix). *Für alle i, j gilt:*

$$A_{ij} = (-1)^{i+j} \det S_{ij}(A).$$

Beweis: Es ist

$$\begin{aligned}
A_{ij} &= \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ & & & 1 & & & \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} \leftarrow i \\
&= (-1)^{(i-1)+(j-1)} \cdot \begin{vmatrix} 1 & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & & & & \vdots \\ 0 & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} \\
&= (-1)^{i+j} \cdot \det \left(\begin{array}{c|c} 1 & * \\ \hline 0 & \\ \vdots & S_{ij}(A) \\ 0 & \end{array} \right) \\
&= (-1)^{i+j} \cdot \det S_{ij}(A).
\end{aligned}$$

Die letzte Gleichung ergibt sich aus dem Kästchensatz! ∎

2.6 Laplace'scher Entwicklungssatz. *Sei $A = (a_{ij}) \in M_n(K)$. Dann gilt für festes j (bzw. i) :*

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det S_{ij}(A) \\
(\textit{bzw.} &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det S_{ij}(A) \quad).
\end{aligned}$$

Man spricht von der Entwicklung nach der j-ten Spalte (bzw. nach der i-ten Zeile).

Beweis: Der zweite Fall folgt aus dem ersten durch Übergang zur transponierten Matrix.

Sind $a_1, \ldots, a_n$ die Spalten von A, so gilt:

$$\begin{aligned}
\det(a_1, \ldots, a_n) &= \det(a_1, \ldots, a_{j-1}, \sum_{i=1}^{n} a_{ij} e_i, a_{j+1}, \ldots, a_n) \\
&= \sum_{i=1}^{n} a_{ij} \det(a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n) \\
&= \sum_{i=1}^{n} a_{ij} A_{ij} \;=\; \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det S_{ij}(A).
\end{aligned}$$

■

Die Vorzeichen sind wie bei einem Schachbrett verteilt:

$$\begin{matrix}
+ & - & + & - & \cdots \\
- & + & - & + & \\
+ & - & + & - & \\
\vdots & & & &
\end{matrix}$$

Beispiel.

Die Entwicklung nach der ersten Zeile ergibt:

$$\begin{aligned}
\det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} &= 0 - 1 \cdot \det \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \\
&= 0 \cdot (-1) - 1 \cdot (-1) + 2 \cdot (3 - 2) \;=\; 3.
\end{aligned}$$

Im allgemeinen wird man eine Zeile oder Spalte suchen, in der möglichst viele Nullen zu finden sind.

Eine Anwendung der Determinantentheorie ergibt sich für Lineare Gleichungen:

2.7 Cramersche Regel. *Sei $A = (a_1, \ldots, a_n) \in M_n(K)$. Dann gilt:*

1. *$A \cdot x = b$ ist genau dann für jedes $b \in K^n$ **eindeutig** lösbar, wenn $\det(A) \neq 0$ ist.*

2. *Ist $\det(A) \neq 0$ und $x = (x_1, \ldots, x_n)^t$ der eindeutig bestimmte Lösungsvektor des LGS $A \cdot x = b$, so ist*

$$x_i = \frac{1}{\det(A)} \cdot \det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n), \text{ für } i = 1, \ldots, n.$$

BEWEIS:

1) Schon bekannt!

2) Ist x der Lösungsvektor des LGS $A \cdot x = b$, so ist

$$b = \sum_{k=1}^{n} x_k a_k.$$

Daraus folgt:

$$\begin{aligned}
\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) &= \\
&= \det(a_1, \ldots, a_{i-1}, \sum_{k=1}^{n} x_k a_k, a_{i+1}, \ldots, a_n) \\
&= \sum_{k=1}^{n} x_k \det(a_1, \ldots, a_{i-1}, a_k, a_{i+1}, \ldots, a_n) \\
&= x_i \cdot \det(a_1, \ldots, a_n).
\end{aligned}$$

■

Beispiel.

Sei $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$ und $b = \begin{pmatrix} -13 \\ -7 \end{pmatrix}$. Dann gilt: $\det(A) = 2 + 15 = 17$, d.h., das LGS ist eindeutig lösbar,

$$\det(b, a_2) = \det \begin{pmatrix} -13 & -3 \\ -7 & 1 \end{pmatrix} = -13 - 21 = -34$$

$$\text{und} \qquad \det(a_1, b) = \det \begin{pmatrix} 2 & -13 \\ 5 & -7 \end{pmatrix} = -14 + 65 = 51.$$

Für den Lösungsvektor $x = (x_1, x_2)^t$ gilt dann:

$$x_1 = \frac{-34}{17} = -2 \text{ und } x_2 = \frac{51}{17} = 3.$$

Als weitere Anwendung ergibt sich eine Berechnungsmöglichkeit für die inverse Matrix:

2.8 Formel für die inverse Matrix. *Sei $A \in \mathrm{GL}_n(K)$. Dann ist $A^{-1} = (y_{ij})$ gegeben durch*

$$y_{ij} = \frac{1}{\det(A)} \cdot A_{ji} = \frac{1}{\det(A)} \cdot (-1)^{i+j} \det S_{ji}(A).$$

Man beachte die Reihenfolge der Indizes!

Beweis: Sei $y_j := (y_{1j}, \ldots, y_{nj})^t$ die j-te Spalte von A^{-1}. Da $A \cdot A^{-1} = E_n$ ist, gilt:

$$A \cdot y_j = e_j \text{ für } j = 1, \ldots, n.$$

Aus der Cramerschen Regel folgt dann:

$$\begin{aligned} y_{ij} &= \frac{1}{\det(A)} \cdot \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) \\ &= \frac{1}{\det(A)} \cdot (-1)^{i+j} \det S_{ji}(A). \end{aligned}$$

∎

Beispiel.

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

Schließlich kann man auch den Rang einer Matrix mit Hilfe von Determinanten bestimmen:

2.9 Rangbestimmung durch Unterdeterminanten. *Sei $A \in M_n(K)$ nicht die Null–Matrix. Dann ist $\mathrm{rg}(A)$ die größte natürliche Zahl r, zu der es eine r-reihige Unterdeterminante $\neq 0$ von A gibt.*

Beweis:

Sei r die Anzahl der Spalten der größten Unterdeterminante $\neq 0$ von A.

1) Sei A' eine r-reihige Untermatrix von A mit $\det(A') \neq 0$. Dann sind die Spalten von A' und damit auch r Spalten von A linear unabhängig. Also ist $\mathrm{rg}(A) \geq r$.

2) Sei $k = \mathrm{rg}(A)$. Dann gibt es k linear unabhängige Spalten in A. Sie bilden eine Matrix $A' \in M_{n,k}(K)$, die ebenfalls den Rang k hat. Aber dann gibt es in A' k linear unabhängige Zeilen. Die bilden eine k-reihige quadratische Untermatrix A'' mit $\det(A'') \neq 0$. Also ist $r \geq \mathrm{rg}(A)$. ∎

Beispiele.

1. Wir betrachten
$$A := \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix}.$$
Es ist $\det(A) = 2+0+8-6-0-4 = 0$, also $\mathrm{rg}(A) < 3$. Links oben findet sich die Unterdeterminante $1 \cdot 1 - 2 \cdot 1 = -1 \neq 0$. Also ist $\mathrm{rg}(A) = 2$.

2. Sei
$$B := \begin{pmatrix} \mathrm{i} & 0 & 2\mathrm{i} \\ 1 & 0 & 2 \\ 0 & \mathrm{i} & \mathrm{i} \\ 1 & 3 & 5 \end{pmatrix} \in M_{4,3}(\mathbb{C}).$$

Man kann nachrechnen, dass alle 3-reihigen Unterdeterminanten Null sind. Man kann aber auch leicht sehen, dass $B = (b_1, b_2, b_3)$ mit $2 \cdot b_1 + b_2 = b_3$ ist. Also ist $\mathrm{rg}(B) < 3$. Rechts unten findet sich die Unterdeterminante $\mathsf{i} \cdot 5 - 3 \cdot \mathsf{i} = 2\mathsf{i} \neq 0$. Also ist $\mathrm{rg}(B) = 2$.

§ 3 Koordinaten und Basiswechsel

A) Basisdarstellung von Vektoren:

Sei V ein beliebiger endlich-dimensionaler K-Vektorraum, $A = (a_1, \ldots, a_n)$ eine Basis von V.

Zu jedem Vektor $x \in V$ gibt es **eindeutig bestimmte** Koeffizienten $x_1, \ldots, x_n \in K$, so dass gilt:

$$x = \sum_{i=1}^{n} x_i a_i.$$

Daher wird durch

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

eine Abbildung $\Phi_A : V \to K^n$ definiert. Wir nennen Φ_A das durch A bestimmte *lineare Koordinatensystem* für V, und die Koeffizienten $x_1, \ldots, x_n$ heißen die *Koordinaten* von x bezüglich der Basis A. Für den Bildvektor $\Phi_A(x)$ schreiben wir auch $[x]_A$.

Speziell ist $[a_\nu]_A = e_\nu$ für $\nu = 1, \ldots, n$.

Beispiele.

1. Ist V selbst der K^n und $E = (e_1, \ldots, e_n)$ die Standardbasis, so ist $\Phi_E = \mathrm{id}_V$, und $[x]_E = x$.

2. Nun sei $V = K^n$, aber $A = (a_1, \ldots, a_n)$ eine beliebige Basis, mit $a_i = (a_{1i}, \ldots, a_{ni})^t$ für $i = 1, \ldots, n$.

   Wir können A auch als Matrix aus $M_n(K)$ auffassen:

   $$A = (a_1, \ldots, a_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

   Sei jetzt $y = (y_1, \ldots, y_n)^t$ ein beliebiger Vektor in V. Dann gilt:

   $$y = \sum_{i=1}^{n} x_i a_i \iff y_k = \sum_{i=1}^{n} x_i a_{ki} \text{ für } k = 1, \ldots, n\,,$$

   also

   $$\Phi_A(y) = x \iff \left\{ \begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ & & \vdots & & & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nn}x_n & = & y_n \end{array} \right\} \iff A \cdot x = y\,.$$

Die Koeffizientenmatrix hat den Rang n, weil ihre Spalten eine Basis des K^n bilden. Die erweiterte Matrix hat dann ebenfalls den Rang n, und das Gleichungssystem ist in jedem Fall eindeutig lösbar. Die Abbildung Φ_A ist durch den Lösungsalgorithmus bestimmt.

Der Vektor $x = [y]_A$ ist die eindeutig bestimmte Lösung des Linearen Gleichungssystems $A \cdot x = y$. Daraus folgt:

$$\boxed{[y]_A = A^{-1} \cdot y\,.}$$

Dies funktioniert nicht mehr, wenn wir einen Vektor y aus einem r-dimensionalen Untervektorraum $V \subset K^n$ bezüglich einer Basis $A = (a_1, \ldots, a_r)$ darstellen wollen, denn in diesem Falle ist $A \in M_{n,r}(K)$ nicht einmal eine quadratische Matrix, kann also nicht invertiert werden. Gültig bleibt aber die Gleichung

$$A \cdot [y]_A = y\,.$$

3. Sei jetzt $V \subset K^n$ ein r-dimensionaler Untervektorraum. Das ist schon fast der allgemeinste Fall, mit dem wir uns zu befassen haben. Die Elemente $a_i = (a_{1i}, \ldots, a_{ni})^t$ einer Basis von V, $i = 1, \ldots, r$, sind immer noch Vektoren im K^n, aber wir haben nur r davon. Also ist

$$A = (a_1, \ldots, a_r) \in M_{n,r}(K).$$

Die Berechnung der Koordinaten funktioniert aber genauso wie oben. Es ist $\mathrm{rg}(A) = r$, und da jeder Vektor $y \in V$ Linearkombination der Spalten von A ist, hat auch die erweiterte Matrix (A, y) den Rang r. Deshalb ist das LGS $A \cdot x = y$ für jedes $y \in V$ lösbar. Weil V und K^r die gleiche Dimension haben, ist das LGS auch eindeutig lösbar, und es gilt:

Der Vektor $x = [y]_A \in K^r$ ist die eindeutig bestimmte Lösung des Linearen Gleichungssystems $A \cdot x = y$.

4. Sei $V = M_2(K)$. Dann bilden die Matrizen

$$E_{11} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_{12} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_{21} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_{22} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

eine Basis A von V, und $\Phi_A : M_{2,2}(K) \to K^4$ ist gegeben durch

$$\Phi_A : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

B) Basiswechsel:

Das lineare Koordinatensystem Φ_A hängt von der Basis $A = (a_1, \ldots, a_n)$ ab. Was passiert, wenn man eine zweite Basis $B = (b_1, \ldots, b_n)$ von V betrachtet. Welcher Zusammenhang besteht – bei festem Vektor $x \in V$ – zwischen $\Phi_A(x)$ und $\Phi_B(x)$?

Φ_A ist ein Isomorphismus, die Umkehrabbildung ist sogar leichter zu verstehen, sie ist durch

$$\Phi_A^{-1} : (x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} x_i a_i$$

gegeben.

Wir betrachten nun das folgende Diagramm:

$$\begin{tikzcd} K^n \arrow[rr, "\Phi_B \circ \Phi_A^{-1}"] & & K^n \\ & V \arrow[ul, "\Phi_A"] \arrow[ur, "\Phi_B"'] & \end{tikzcd}$$

$\Phi_B \circ \Phi_A^{-1} : K^n \to K^n$ wird durch eine Matrix $W_{B,A} \in M_{n,n}(K)$ beschrieben, und es gilt:

$$\Phi_B \circ \Phi_A^{-1}(\Phi_A(x)) = \Phi_B(x).$$

In der Sprache der Matrizen heißt das:

$$\boxed{W_{B,A} \cdot [x]_A = [x]_B.}$$

Man nennt $W_{B,A}$ die *Basiswechsel-Matrix* und die obige Formel die *Basiswechsel-Formel.*

Wie kann man die Basiswechsel-Matrix berechnen? Für die Spalten der Basiswechsel-Matrix gilt:

$$S_j(W_{B,A}) = W_{B,A} \cdot e_j = W_{B,A} \cdot [a_j]_A = [a_j]_B.$$

Also ist

$$W_{B,A} = ([a_1]_B, \ldots, [a_n]_B).$$

Spezialfall:

Sei $V \subset K^n$ ein r-dimensionaler Untervektorraum. In V seien zwei Basen

$$A = (a_1, \ldots, a_r) \quad \text{und} \quad B = (b_1, \ldots, b_r)$$

gegeben. Die zugehörigen Matrizen in $M_{n,r}(K)$ bezeichnen wir ebenfalls mit A und B. Die Basiswechsel-Matrix $W_{B,A}$ liegt in diesem Falle in $M_r(K)$, und es gilt:

$$(B \cdot W_{B,A}) \cdot e_\nu = B \cdot (W_{B,A} \cdot e_\nu) = B \cdot [a_\nu]_B = a_\nu = A \cdot e_\nu,$$

also

$$\boxed{B \cdot W_{B,A} = A.}$$

Die ν-te Spalte von $W_{B,A}$ ist die eindeutig bestimmte Lösung des LGS $B \cdot x = a_\nu$.

Das liefert ein Rechenverfahren für die Bestimmung der Basiswechselmatrix. Ihre Spalten ergeben sich als Lösungen $x_1, \ldots, x_r$ der linearen Gleichungssysteme $B \cdot x_1 = a_1, \ldots, B \cdot x_r = a_r$. Man kann die zugehörigen Gaußverfahren simultan durchführen. Dazu wendet man das Verfahren auf die erweiterte Matrix $(B|a_1|\ldots|a_r)$ an. Hat man B in die gewünschte Dreiecksgestalt gebracht, so erhält man die Lösungen $x_1, \ldots, x_r$ durch Rückwärtseinsetzen, wobei für x_j jeweils die $(r+j)$-te Spalte als rechte Seite benutzt werden muss.

Ist sogar $r = n$, so treibt man die Umformungen weiter, bis man erhält:

$$(B|A) \quad \to \quad (E_n|B^{-1} \cdot A).$$

Auf der rechten Seite steht dann die Basiswechselmatrix, denn es ist ja $B \cdot W_{B,A} = A$. Die Formel $W_{B,A} = B^{-1} \cdot A$ gilt natürlich nicht mehr, wenn wir es mit Basen eines Untervektorraums von K^n zu tun haben.

Allgemein:

Kommen wir zurück zu einem allgemeinen Basiswechsel in einem n-dimensionalen K-Vektorraum V. Es gilt:

$$W_{A,B} \cdot W_{B,A} \cdot [x]_A = W_{A,B} \cdot [x]_B = [x]_A$$

und

$$W_{B,A} \cdot W_{A,B} \cdot [x]_B = W_{B,A} \cdot [x]_A = [x]_B.$$

Weil x völlig beliebig gewählt werden kann, folgt:

$$W_{A,B} \cdot W_{B,A} = E_n \;\text{ und }\; W_{B,A} \cdot W_{A,B} = E_n, \;\text{ also } W_{A,B}^{-1} = W_{B,A}.$$

Beispiele.

1. Die Vektoren $b_1 := (1,1,0)^t$, $b_2 := (1,0,1)^t$ und $b_3 := (0,1,1)^t$ bilden eine Basis B des $\mathbb{R}^3$. Die Matrix

$$B := (b_1, b_2, b_3) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

ist invertierbar, die inverse Matrix ist gegeben durch

$$B^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Wir wollen herausfinden, wie die Darstellung des Vektors $x := (12, 6, 30)^t$ bezüglich der Basis B aussieht. Nach der obigen Formel ist

$$[x]_B = B^{-1} \cdot x = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = \begin{pmatrix} -6 \\ 18 \\ 12 \end{pmatrix}.$$

Probe: Tatsächlich ist

$$-6 \cdot b_1 + 18 \cdot b_2 + 12 \cdot b_3 = \begin{pmatrix} -6 \\ -6 \\ 0 \end{pmatrix} + \begin{pmatrix} 18 \\ 0 \\ 18 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \\ 12 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = x.$$

2. Die Vektoren $a_1 = (1, 3, 2)^t$, $a_2 = (0, -1, 5)^t$ und $a_3 = (4, 2, -3)^t$ bilden ebenfalls eine Basis des $\mathbb{R}^3$. Dann ist

$$\begin{aligned} W_{B,A} &= B^{-1} \cdot A = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 4 \\ 3 & -1 & 2 \\ 2 & 5 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -3 & 9/2 \\ 0 & 3 & -1/2 \\ 2 & 2 & -5/2 \end{pmatrix}. \end{aligned}$$

3. Ein Ausnahmefall ist die Bestimmung der Koordinaten eines Vektors bezüglich einer ON-Basis. Sei V ein euklidischer Vektorraum, $A = (a_1, \ldots, a_n)$ eine ON-Basis von V. Ist $x \in V$ ein beliebiger Vektor, so findet man die Koeffizienten x_i in der Darstellung

$$x = \sum_{i=1}^{n} x_i a_i$$

durch die Formel $\boxed{x_i = <x\,,\, a_i>, \text{ für } i = 1, \ldots, n.}$

Es ist also $[x]_A = (<x\,,\, a_1>, \ldots, <x\,,\, a_n>)^t$.

C) Basisdarstellung von linearen Abbildungen:

Eine lineare Abbildung $f : K^n \to K^m$ wird durch eine eindeutig bestimmte Matrix $M = M(f) \in M_{m,n}(K)$ beschrieben, mit

$$M(f) \cdot x = f(x).$$

Die Spalten von M sind gegeben durch $S_j(M) = M \cdot e_j = f(e_j)$.

Das wollen wir jetzt verallgemeinern.

Gegeben seien

- Zwei endlich-dimensionale K-Vektorräume V und W,
- zwei Basen $A = (a_1, \ldots, a_n)$ von V und $B = (b_1, \ldots, b_m)$ von W,
- eine lineare Abbildung $f : V \to W$.

f ist durch die n Werte $f(a_1), \ldots, f(a_n)$ bereits vollständig festgelegt. Und jeder der Vektoren $f(a_j)$ lässt sich auf eindeutige Weise als Linearkombination der Basisvektoren $b_1, \ldots, b_m$ darstellen:

$$\boxed{f(a_j) = \sum_{i=1}^{m} \alpha_{ij} b_i, \quad \text{mit } \alpha_{ij} \in K, \quad \text{für } j = 1, \ldots, n.}$$

Achtung! Die Position der Indizes ist vorgeschrieben!

Mit $M_{B,A}(f) \in M_{m,n}(K)$ bezeichnen wir die Matrix der α_{ij}.

Mit Hilfe von Koordinaten kann man die Situation folgendermaßen beschreiben:

$$\begin{array}{ccccc} & V & \xrightarrow{f} & W & \\ \Phi_A & \downarrow & & \downarrow & \Phi_B \\ & K^n & \xrightarrow{f_{B,A}} & K^m & \end{array}$$

Die Abbildung $f_{B,A} := \Phi_B \circ f \circ \Phi_A^{-1} : K^n \to K^m$ wird erst durch das vorgelegte Diagramm definiert. Sie ist ebenfalls linear und wird in der üblichen Weise durch eine Matrix $M(f_{B,A})$ beschrieben. Wir zeigen, dass $M(f_{B,A}) = M_{B,A}(f)$ ist. Es gilt nämlich:

$$\begin{aligned} S_j(M(f_{B,A})) &= f_{B,A}(e_j) \\ &= \Phi_B \circ f \circ \Phi_A^{-1}(e_j) \\ &= \Phi_B \circ f(a_j) \\ &= \Phi_B(\alpha_{1j} b_1 + \cdots + \alpha_{mj} b_m) \\ &= (\alpha_{1j}, \ldots, \alpha_{mj})^t \\ &= S_j(M_{B,A}(f)). \end{aligned}$$

Definition. Die Matrix $M_{B,A}(f) = M(f_{B,A})$ zur linearen Abbildung

$$f_{B,A} = \Phi_B \circ f \circ \Phi_A^{-1} : K^n \to K^m$$

nennt man *die Matrix, die $f : V \to W$ bezüglich der Basen A und B beschreibt.*
Da $f_{B,A}(\Phi_A(x)) = \Phi_B(f(x))$ ist, folgt nun:

$$\boxed{M_{B,A}(f) \cdot [x]_A = [f(x)]_B, \quad \text{für } x \in V.}$$

Es ist $[a_j]_A = e_j$, für $j = 1, \ldots, n$, also

$$S_j(M_{B,A}(f)) = M_{B,A}(f) \cdot e_j = M_{B,A}(f) \cdot [a_j]_A = [f(a_j)]_B.$$

Das ergibt folgende Formel zur Berechnung von $M_{B,A}(f)$:

$$\boxed{M_{B,A}(f) = ([f(a_1)]_B, \ldots, [f(a_n)]_B).}$$

Ein Sonderfall liegt vor, wenn f eine lineare Abbildung von K^n nach K^m ist. Dann kann f bezüglich der Standardbasen durch eine Matrix $M := M(f) \in M_{m,n}(K)$ beschrieben werden:

$$f(x) = M \cdot x.$$

Wir wollen f jetzt aber bezüglich anderer Basen $A = (a_1, \ldots, a_n)$ von K^n und $B = (b_1, \ldots, b_m)$ von K^m beschreiben. Dann ist

$$\begin{aligned} M_{B,A}(f) &= ([f(a_1)]_B, \ldots, [f(a_n)]_B) \\ &= ([M \cdot a_1]_B, \ldots, [M \cdot a_n]_B) \\ &= (B^{-1} \cdot M \cdot a_1, \ldots, B^{-1} \cdot M \cdot a_n) \\ &= B^{-1} \cdot M \cdot A. \end{aligned}$$

$$\boxed{f = f_M : K^n \to K^m \text{ wird beschrieben durch } M_{B,A}(f) = B^{-1} \cdot M \cdot A\,.}$$

Sehr häufig wird auch der Fall $V = W$ betrachtet. Eine lineare Abbildung $f : V \to V$ bezeichnet man als einen *Endomorphismus* von V. Man kann dann versuchen, mit einiger einzigen Basis A auszukommen und f durch $M_A(f) := M_{A,A}(f)$ zu beschreiben. Ist $V = K^n$, so liegt es nahe, die Standardbasis zu benutzen. Aber oft zeigt es sich, dass man durch geschickte Wahl der Basis zu einer viel einfacheren Matrix kommen kann. Wie das geht, werden wir im nächsten Abschnitt sehen.

Zum Schluss eine Konsequenz für Determinanten:

Sei V ein beliebiger n-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Wir wählen eine beliebige Basis $A = (a_1, \ldots, a_n)$ von V. Dann wird f durch die Matrix $M = M_A(f) \in M_{n,n}(K)$ beschrieben, mit

$$M_A(f) := M(\Phi_A \circ f \circ \Phi_A^{-1}) = ([f(a_1)]_A, \ldots, [f(a_n)]_A).$$

Ist nun A' eine andere Basis von V, $M' = M_{A'}(f)$ und P die zugehörige Basiswechselmatrix, so ist $M' = P^{-1} \cdot M \cdot P$, also

$$\det(M') = \det(P)^{-1} \cdot \det(M) \cdot \det(P) = \det(M).$$

Daher ist folgendes sinnvoll:

Definition. Sei V ein beliebiger endlich-dimensionaler K-Vektorraum, A eine Basis für V und $f : V \to V$ ein Endomorphismus. Dann setzt man

$$\det(f) := \det(M_A(f)).$$

§ 4 Eigenwerte

Sei K ein beliebiger Körper.

Definition. Ein *Polynom* über K ist eine unendliche Folge $f = (a_0, a_1, a_2, \ldots)$ von Elementen $a_i \in K$, mit $a_i = 0$ für fast alle i. Die Elemente a_i nennt man die *Koeffizienten* von f. Den größten Index n mit $a_n \neq 0$ nennt man den *Grad* von f (in Zeichen $\mathrm{grad}(f)$). Das „Nullpolynom" $0 = (0, 0, 0, \ldots)$ erhält den Grad $-\infty$.
Zwei Polynome sind gleich, wenn alle ihre Koeffizienten übereinstimmen. Man kann zwei Polynome $f = (a_0, a_1, a_2, \ldots)$ und $g = (b_0, b_1, b_2, \ldots)$ addieren und multiplizieren:

$$\begin{aligned} f + g &:= (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots) \\ \text{und} \quad f \cdot g &:= (c_0, c_1, c_2, \ldots), \text{ mit } c_k := \sum_{i+j=k} a_i b_j \,. \end{aligned}$$

In der Gestalt $(c, 0, 0, \ldots)$ kann man jedes Element $c \in K$ als Polynom auffassen.

Das Element $X := (0, 1, 0, 0, \ldots)$ bezeichnet man als die *Unbestimmte*. Es ist

$$\begin{aligned} X^2 &= (0, 0, 1, 0, 0, \ldots), \\ X^3 &= (0, 0, 0, 1, 0, \ldots) \\ \text{und allgemein } X^k &= (0, 0, \ldots, 0, 1, 0, \ldots), \text{ mit einer 1 an der } k\text{-ten Position,} \\ & \quad\; \text{wenn man die vorderste Position als 0. Position bezeichnet.} \end{aligned}$$

Dadurch kann man ein Polynom $f = (a_0, a_1, \ldots, a_n, 0, 0, \ldots)$ mit $a_n \neq 0$ in der Form $f = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n$ schreiben.

Definition. Die Menge aller Polynome über K wird mit $K[X]$ bezeichnet.
Offensichtlich ist $K[X]$ ein kommutativer Ring.

Jedem Polynom $f = \sum_{k=0}^{n} a_k X^k$ ist eine Funktion $F_f : K \to K$ durch $F_f(\alpha) := \sum_{k=0}^{n} a_k \alpha^k$ zugeordnet. Diese Zuordnung ist im Falle $K = \mathbb{R}$ oder $K = \mathbb{C}$ injektiv, nicht aber bei beliebigen Körpern. Über $\mathbb{Z}_2$ stimmen z.B. die Funktionen zu den Polynomen 0 und $X^2 - X$ überein.

4.1 Satz. *Sind f, g zwei Polynome, so ist*

$\mathrm{grad}(f + g) \leq \max(\mathrm{grad}(f), \mathrm{grad}(g))$ *und* $\mathrm{grad}(f \cdot g) = \mathrm{grad}(f) + \mathrm{grad}(g)$.

Der Beweis ist trivial.

4.2 Satz. *Seien f, g zwei Polynome.*

1. *Ist $f \cdot g = 0$, so ist $f = 0$ oder $g = 0$.*
2. *Ist $f \cdot g = 1$, so ist $\mathrm{grad}(f) = \mathrm{grad}(g) = 0$.*

Beweis: 1) Ist $f \neq 0$ und $g \neq 0$, so ist $\mathrm{grad}(f \cdot g) = \mathrm{grad}(f) + \mathrm{grad}(g) \geq 0$, also $f \cdot g \neq 0$.

2) Ist $f \cdot g = 1$, so ist $0 = \mathrm{grad}(1) = \mathrm{grad}(f \cdot g) = \mathrm{grad}(f) + \mathrm{grad}(g)$. Das geht nur, wenn $\mathrm{grad}(f) = \mathrm{grad}(g) = 0$ ist. ∎

Es folgt, dass $K[X]$ kein Körper ist.

4.3 Satz (Division mit Rest). *Seien $f, g \in K[X]$, $g \neq 0$. Dann gibt es eindeutig bestimmte Polynome q, r mit*

1. $f = q \cdot g + r$.
2. $r = 0$ *oder* $\text{grad}(r) < \text{grad}(g)$.

BEWEIS: 1) Existenz: Sei $\text{grad}(f) \geq \text{grad}(g)$ (sonst $q = 0$ und $r = f$). Ist $f = a_n X^n + \cdots + a_0$ und $g = b_m X^m + \cdots + b_0$, so setze man $q_1 := (a_n/b_m)X^{n-m}$. Dann hat $f_1 := f - q_1 g$ einen Grad $< n$. Ist sogar $\text{grad}(f_1) < m$, so ist man fertig. Andernfalls gibt es eine Zerlegung $f_1 = q_2 g + r$ und $f = (q_1 + q_2)g + r$.

2) Eindeutigkeit: Sei $f = q_1 g + r_1 = q_2 g + r_2$. Dann ist $0 = (q_1 - q_2)g + (r_1 - r_2)$, bzw. $(q_1 - q_2)g = r_2 - r_1$. Aus Gradgründen muss $q_1 = q_2$ und $r_1 = r_2$ sein. ∎

Schreibweise: Ist $f \in K[X]$ und $\alpha \in K$, so setzen wir $f(\alpha) := F_f(\alpha)$.

4.4 Satz. *Sei $f \in K[X]$, $\text{grad}(f) \geq 1$ und $f(\alpha) = 0$. Dann gibt es ein $g \in K[X]$, so dass $f = (X - \alpha) \cdot g$ ist.*

BEWEIS: Nach Division mit Rest ist $f = q \cdot (X - \alpha) + r$ mit $r = 0$ oder $\text{grad}(r) = 0$. Setzt man α ein, so erhält man $r = r(\alpha) = 0$. ∎

Definition. Ist $f = (X - \alpha)^k \cdot g$ und $g(\alpha) \neq 0$, so nennt man α eine Nullstelle der *Vielfachheit* k von f.
Nach dem **Fundamentalsatz der Algebra** (hier ohne Beweis) besitzt jedes Polynom über $\mathbb{C}$ von positivem Grad mindestens eine (komplexe) Nullstelle. Daher kann man jedes komplexe Polynom vollständig in Linearfaktoren zerlegen. Über $\mathbb{R}$ gibt es Polynome ohne Nullstellen. Ist f ein Polynom mit reellen Koeffizienten und α eine komplexe Nullstelle von f, so ist auch die konjugierte Zahl $\overline{\alpha}$ eine Nullstelle von f (Beweis trivial).

Definition. Sei V ein K-Vektorraum, $f : V \to V$ ein Endomorphismus. Ein Vektor $x \in V$ heißt *Eigenvektor* von f, falls gilt:

1. x ist nicht der Nullvektor.
2. Es gibt ein $\lambda \in K$, so dass $f(x) = \lambda x$ ist.

Der Skalar λ heißt *Eigenwert* von f zum Eigenvektor x.

Ist $A \in M_n(K)$ eine quadratische Matrix, so setzen wir $f_A(x) := A \cdot x$. Unter einem Eigenvektor (Eigenwert) von A versteht man einen Eigenvektor (Eigenwert) von $f_A : K^n \to K^n$.

Beispiel.

Ist $A := \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ und $x_0 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, so ist

$$A \cdot x_0 = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 \\ 3 \cdot 2 + 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4 \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 4 \cdot x_0,$$

also 4 ein Eigenwert von A zum Eigenvektor x_0.

4.5 Der Raum $E(\lambda)$. *Sei V ein K-Vektorraum, $f : V \to V$ linear und $\lambda \in K$. Dann ist*

$$E(\lambda) := \{x \in V \;:\; f(x) = \lambda x\}$$

ein Untervektorraum von V.

BEWEIS:
1) Offensichtlich liegt der Nullvektor in $E(\lambda)$. Man beachte allerdings, dass 0 kein Eigenvektor ist!
2) Ist $x \in E(\lambda)$, also $f(x) = \lambda x$, so ist auch

$$f(\alpha x) = \alpha \cdot f(x) = \alpha(\lambda x) = \lambda(\alpha x),$$

also $\alpha x \in E(\lambda)$, für $\alpha \in K$.
3) Sind $x, y \in E(\lambda)$, so ist $f(x) = \lambda x$ und $f(y) = \lambda y$. Also ist

$$f(x + y) = f(x) + f(y) = \lambda x + \lambda y = \lambda(x + y),$$

also $x + y \in E(\lambda)$. ∎

Definition. Ist $f : V \to V$ linear, $\lambda \in K$, so heißt $E(\lambda)$ der *Eigenraum* von f zu λ.

Bemerkung. Die Menge $E(\lambda) \setminus \{0\}$ besteht aus Eigenvektoren zum Eigenwert λ. Im allgemeinen wird sie natürlich leer sein.

Wozu sind Eigenvektoren gut?

Definition. Sei V ein endlich-dimensionaler K-Vektorraum. Eine lineare Abbildung $f : V \to V$ heißt *diagonalisierbar*, wenn es eine Basis $(b_1, \ldots, b_n)$ von V gibt, bzgl. der f durch eine Diagonalmatrix beschrieben wird.

Eine Matrix $A \in M_n(K)$ heißt diagonalisierbar, wenn $f_A : K^n \to K^n$ diagonalisierbar ist.

- Sei f ein Endomorphismus von V und $B = (b_1, \ldots, b_n)$ eine Basis, bezüglich der f durch eine Diagonalmatrix $D = \Delta(\lambda_1, \ldots, \lambda_n)$ beschrieben wird. Dann ist $f(b_j) = \lambda_j b_j$ für alle j. Also sind die b_j Eigenvektoren von f (mit zugehörigen Eigenwerten λ_j).

- Nun sei speziell $f = f_A : K^n \to K^n$, mit $A \in M_n(K)$. Bezüglich einer Basis $B = (b_1, \ldots, b_n)$ wird f dann durch die Matrix $M_B(f_A) = B^{-1} \cdot A \cdot B = M_{B,B}(f_A)$ beschrieben, mit der invertierbaren Matrix $B = (b_1, \ldots, b_n)$.

Damit haben wir gezeigt:

4.6 Diagonalisierbarkeit von Matrizen. *Folgende Aussagen über eine Matrix $A \in M_n(K)$ sind äquivalent:*

1. *A ist diagonalisierbar.*
2. *Im K^n gibt es eine Basis von Eigenvektoren von A.*
3. *Es gibt eine Diagonalmatrix D und eine invertierbare Matrix B, so dass gilt:*

$$D = B^{-1} \cdot A \cdot B.$$

Wie findet man Eigenvektoren? Es ist einfacher, mit den Eigenwerten zu beginnen. Sei $f : V \to V$ ein Endomorphismus.

$$\begin{aligned}
\lambda \text{ Eigenwert von } f \quad &\iff \quad \exists\, v \neq 0 \text{ mit } f(v) = \lambda \cdot v \\
&\iff \quad \exists\, v \neq 0 \text{ mit } f(v) - \lambda v = 0 \\
&\iff \quad \exists\, v \neq 0 \text{ mit } (f - \lambda \cdot \mathrm{id}_V)(v) = 0 \\
&\iff \quad \mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V) \neq \{0\} \\
&\iff \quad \mathrm{Im}(f - \lambda \cdot \mathrm{id}_V) \neq V \\
&\iff \quad \mathrm{rg}(f - \lambda \cdot \mathrm{id}_V) < \dim(V) \\
&\iff \quad \det(f - \lambda \cdot \mathrm{id}_V) = 0.
\end{aligned}$$

Wir betrachten diese Aussagenkette für den Spezialfall $f = F_A : K^n \to K^n$.

$$\begin{aligned}
\lambda \text{ Eigenwert von } A \quad &\iff \quad \exists\, v \neq 0 \text{ mit } A \cdot v = \lambda v \\
&\iff \quad \exists\, v \neq 0 \text{ mit } A \cdot v - \lambda \cdot v = 0 \\
&\iff \quad \exists\, v \neq 0 \text{ mit } (A - \lambda \cdot E_n) \cdot v = 0 \\
&\iff \quad \mathrm{Ker}(F_{A - \lambda \cdot E_n}) \neq 0 \\
&\iff \quad \mathrm{rg}(A - \lambda \cdot E_n) < n \\
&\iff \quad \det(A - \lambda \cdot E_n) = 0 \\
&\iff \quad \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.
\end{aligned}$$

Wie sieht diese Determinante aus? Wir setzen $\widetilde{a}_{ij} := \begin{cases} a_{ij} & \text{für } i \neq j \\ a_{ii} - \lambda & \text{für } i = j \end{cases}$.

Außerdem beachten wir folgende Tatsache:

Ist $\sigma \in S_n$ eine Permutation $\neq$ id, so gibt es ein i mit $j := \sigma(i) \neq i$. Weil Permutationen bijektive Abbildungen sind, kann auch nicht $\sigma(j) = j$ gelten, denn dann hätten ja zwei Zahlen das gleiche Bild. Damit folgt:

$$\begin{aligned}
\det(A - \lambda \cdot E_n) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\widetilde{a}_{1,\sigma(1)} \cdot \ldots \cdot \widetilde{a}_{n,\sigma(n)} \\
&= (a_{11} - \lambda) \cdot \ldots \cdot (a_{nn} - \lambda) \; + \text{ Terme, bei denen} \\
&\qquad\quad \lambda \text{ in höchstens } n-2 \text{ Faktoren vorkommt} \\
&= (-1)^n\lambda^n + (-1)^{n-1}(\sum_{i=1}^{n} a_{ii})\lambda^{n-1} + \text{Terme vom Grad} \leq n-2.
\end{aligned}$$

Bei der Definition der Determinante durch die Leibniz-Formel kommen keine Divisionen vor. Deshalb kann die Determinante auch für Matrizen definiert werden, deren Einträge (statt in einem Körper) in einem kommutativen Ring mit 1 liegen, also z.B. in $K[X]$.

Definition. $p_A(X) := \det(A - X \cdot E_n)$ heißt *charakteristisches Polynom* von A. Damit gilt:

Das charakteristische Polynom einer Matrix $A \in M_n(K)$ ist ein Polynom vom Grad n über K, von der Gestalt $p_A(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1} + c_nX^n$, mit

$$\begin{aligned}
c_n &= (-1)^n, \\
c_{n-1} &= (-1)^{n-1} \cdot (a_{11} + \cdots + a_{nn}), \\
&\vdots \\
c_0 &= p_A(0) \;=\; \det(A).
\end{aligned}$$

Zur Abkürzung setzt man $\operatorname{Spur}(A) := \sum_{i=1}^{n} a_{ii}$ (*Spur* von A).

Oben haben wir gesehen:
Die Eigenwerte von A sind die Nullstellen des charakteristischen Polynoms $p_A(X)$.
Das liefert eine praktische Berechnungsmöglichkeit.

Beispiele.

1. Wir betrachten noch einmal die Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Es gilt:

$$\det\begin{pmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4.$$

Also ist

$$p_A(\lambda) = 0 \iff \lambda = \frac{3 \pm \sqrt{9+16}}{2} = \begin{cases} 4 \\ -1 \end{cases}$$

Den einen dieser beiden Eigenwerte hatten wir schon kennengelernt. Nun suchen wir die zugehörigen Eigenvektoren. Dazu muss man bei gegebenem λ die Gleichung $(A - \lambda \cdot E_2) \cdot x = 0$ lösen:

$$\lambda = -1: \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{hat z.B.} \quad b_1 := \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ als Lösung.}$$

$$\lambda = 4: \begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{hat} \quad b_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix} \text{ als Lösung.}$$

b_1 und b_2 sind Eigenvektoren zu den Eigenwerten -1 bzw. 4. Da $\det \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = 5 \neq 0$ ist, bilden sie auch eine Basis.

Damit ist bewiesen, dass A diagonalisierbar ist. Wir machen noch die Probe. Dazu setzen wir $B := (b_1, b_2) = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. Dann ist $B^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$ und

$$\begin{aligned} B^{-1} \cdot A \cdot B &= \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 8 \\ 1 & 12 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} -5 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Es kommt tatsächlich eine Diagonalmatrix heraus, und die Einträge darin sind gerade die Eigenwerte. Das muss natürlich so sein.

2. $R(\varphi)$ sei die „Drehmatrix" $R(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. Dann ist

$$\begin{aligned} \det(R(\varphi) - \lambda \cdot E_2) &= \det \begin{pmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{pmatrix} \\ &= (\cos\varphi - \lambda)^2 + \sin^2\varphi \\ &= \lambda^2 - 2\lambda\cos\varphi + 1. \end{aligned}$$

Dieser Ausdruck wird genau dann Null, wenn

$$\lambda = \frac{1}{2} \cdot (2\cos\varphi \pm \sqrt{4\cos^2\varphi - 4}) = \cos\varphi \pm \sqrt{-\sin^2\varphi} \quad \text{ist.}$$

Das ist nur möglich, wenn $\varphi = 0$ oder $\varphi = \pi$ ist, also $\sin\varphi = 0$. Das bedeutet, dass unter allen Drehmatrizen nur die speziellen Matrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 \quad \text{und} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E_2$$

Eigenwerte besitzen (und zwar $\lambda = 1$ bzw. $\lambda = -1$). Andere Drehungen besitzen keinen Eigenwert und daher auch keinen Eigenvektor.

Spaßeshalber kann man mit $R(\varphi)$ auch einen Endomorphismus des $\mathbb{C}^n$ beschreiben und nach *komplexen* Eigenwerten suchen. Auf Grund der obigen

Berechnung ist sofort klar, dass $\lambda_{1,2} = \cos\varphi \pm \mathsf{i}\ \sin\varphi =: e^{\pm\mathsf{i}\varphi}$ komplexe Eigenwerte sind.

Die Gleichung für einen Eigenvektor $\begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2$ zum Eigenwert $\lambda_1 = e^{\mathsf{i}\varphi}$ lautet:

$$\begin{pmatrix} -\mathsf{i}\ \sin\varphi & -\sin\varphi \\ \sin\varphi & -\mathsf{i}\ \sin\varphi \end{pmatrix} \cdot \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Man sieht, dass $\begin{pmatrix} 1 \\ -\mathsf{i} \end{pmatrix}$ eine Lösung ist. Genauso ist $\begin{pmatrix} 1 \\ \mathsf{i} \end{pmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_2 = e^{-\mathsf{i}\varphi}$. Offensichtlich sind $\begin{pmatrix} 1 \\ \mathsf{i} \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -\mathsf{i} \end{pmatrix}$ linear unabhängig (über $\mathbb{C}$), es gibt also eine Basis von Eigenvektoren. Man sieht, dass eine Matrix über $\mathbb{R}$ nicht diagonalisierbar, aber über $\mathbb{C}$ diagonalisierbar sein kann.

4.7 Lineare Unabhängigkeit von Eigenvektoren. *Sind $v_1, \ldots, v_k \in V$ Eigenvektoren eines Endomorphismus $f : V \to V$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k \in K$, so sind sie linear unabhängig.*

BEWEIS: Wir führen Induktion nach der Anzahl k.

$\underline{k = 1}$: **Ein** einzelner Eigenvektor v_1 ist immer linear unabhängig, weil er $\neq 0$ ist.

$\underline{k - 1 \to k}$ (mit $k \geq 2$): Es seien Eigenvektoren $v_1, \ldots, v_k$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$ gegeben. Wir nehmen an, sie wären linear abhängig, und versuchen, daraus einen Widerspruch zu konstruieren:

Nach Induktionsvoraussetzung sind $v_1, \ldots, v_{k-1}$ linear unabhängig. Wenn sie durch Hinzunahme von v_k linear abhängig werden, muss es Koeffizienten $\alpha_j \in K$ geben, so dass gilt: $v_k = \sum\limits_{j=1}^{k-1} \alpha_j v_j$. Daraus folgt:

$$\begin{aligned} 0 &= (f - \lambda_k \cdot \mathrm{id}_V)(v_k) = (f - \lambda_k \cdot \mathrm{id}_V)(\sum_{j=1}^{k-1} \alpha_j v_j) \\ &= \sum_{j=1}^{k-1} \alpha_j \cdot (f - \lambda_k \cdot \mathrm{id}_V)(v_j) = \sum_{j=1}^{k-1} \alpha_j \cdot (\lambda_j - \lambda_k) \cdot v_j. \end{aligned}$$

Da $v_1, \ldots, v_{k-1}$ linear unabhängig sind, muss $\alpha_j(\lambda_j - \lambda_k) = 0$ für $j = 1, \ldots, k-1$ gelten. Da die $\lambda_j \neq \lambda_k$ sind, also $\lambda_j - \lambda_k \neq 0$ für $j = 1, \ldots, k-1$, muss $\alpha_1 = \ldots = \alpha_{k-1} = 0$ sein. Aber dann ist $v_k = 0$, und das kann bei einem Eigenvektor nicht sein. WS! ∎

4.8 Folgerung 1. *Hat $f : V \to V$ n verschiedene Eigenwerte, so ist f diagonalisierbar.*

4.9 Folgerung 2. *$\lambda_1, \ldots, \lambda_r$ seien paarweise verschiedene Eigenwerte von $A \in M_n(K)$. Ist $w_i \in E(\lambda_i)$ für $i = 1, \ldots, r$ und $w_1 + \cdots + w_r = 0$, so ist $w_1 = \ldots = w_r = 0$.*

BEWEIS: Wir nehmen an, dass es ein k mit $1 \leq k \leq r$ gibt, so dass – nach geeigneter Numerierung – die ersten k Vektoren $\neq 0$ sind, während $w_{k+1} = \ldots = w_r = 0$ ist. Dann ist auch $w_1 + \cdots + w_k = 0$. Ist $k = 1$, so ist nichts zu zeigen. Sei also $k \geq 2$ und

$$0 = A \cdot \big(\sum_{i=1}^{k} w_i\big) = \sum_{i=1}^{k} A \cdot w_i = \sum_{i=1}^{k} \lambda_i \cdot w_i.$$

Da Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, muss $\lambda_1 = \ldots = \lambda_k = 0$ sein. Aber das ist ein Widerspruch, denn die λ_i sollten paarweise verschieden sein. ∎

Das charakteristische Polynom p_A einer Matrix $A \in M_n(K)$ hängt in Wirklichkeit nur von der linearen Abbildung $f = f_A$ ab, ganz egal, bezüglich welcher Basis man f beschreibt:

4.10 Satz. *Ist B invertierbar, so ist*

$$p_{B^{-1}AB}(X) = p_A(X).$$

BEWEIS: Es ist

$$\begin{aligned} p_{B^{-1}AB}(X) &= \det(B^{-1} \cdot A \cdot B - X \cdot E_n) \\ &= \det(B^{-1} \cdot A \cdot B - B^{-1} \cdot (X \cdot E_n) \cdot B) \\ &= \det(B^{-1} \cdot (A - X \cdot E_n) \cdot B) \\ &= (\det B)^{-1} \cdot \det(A - X \cdot E_n) \cdot \det B \\ &= \det(A - X \cdot E_n) \;=\; p_A(X). \end{aligned}$$

∎

Insbesondere kann man vom charakteristischen Polynom eines Endomorphismus sprechen. Eine weitere Folgerung betrifft die Spur. Es ist $\text{Spur}(B^{-1}AB) = \text{Spur}(A)$.

Leider kann man nicht immer erwarten, dass das charakteristische Polynom in lauter verschiedene Linearfaktoren zerfällt.

Beispiel.

Sei $A := \begin{pmatrix} 6 & 4 & 4 \\ -1 & 0 & -1 \\ -2 & 0 & 0 \end{pmatrix}$. Dann gilt:

$$\begin{aligned}
\det(A - \lambda \cdot E_3) &= \det \begin{pmatrix} 6-\lambda & 4 & 4 \\ -1 & -\lambda & -1 \\ -2 & 0 & -\lambda \end{pmatrix} \\
&= (6-\lambda)\lambda^2 + 8 + 0 - 8\lambda - 0 - 4\lambda \\
&= -\lambda^3 + 6\lambda^2 - 12\lambda + 8 \\
&= (2-\lambda)^3.
\end{aligned}$$

Also ist $\lambda = 2$ die einzige Nullstelle von $p_A(x)$. Sie hat die Vielfachheit 3. Mehr Nullstellen kann es aus Gradgründen selbst im Komplexen nicht geben. Wie steht es nun mit den Eigenvektoren? Die zugehörige Gleichung hat die Form

$$\begin{pmatrix} 4 & 4 & 4 \\ -1 & -2 & -1 \\ -2 & 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Aus der Koeffizientenmatrix gewinnt man durch elementare Umformungen zunächst die Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & -1 \\ 1 & 0 & 1 \end{pmatrix}, \text{ dann } \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ und schließlich } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Das resultierende Gleichungssystem

$$\begin{aligned}
x_1 + x_2 + x_3 &= 0, \\
x_2 &= 0
\end{aligned}$$

liefert als Lösungsraum den Eigenraum

$$E(2) = \{\alpha \cdot (-1, 0, 1) \mid \alpha \in \mathbb{R}\}.$$

Der ist nur 1-dimensional. Also ist A **nicht** diagonalisierbar.

Definition. Sei $\lambda \in K$ ein Eigenwert der Matrix $A \in M_n(K)$.

$a(\lambda)$:= Vielfachheit der Nullstelle λ des charakteristischen Polynoms $p_A(x)$ heißt *algebraische Vielfachheit von* λ.

$g(\lambda)$:= $\dim_K E(\lambda)$ heißt *geometrische Vielfachheit von* λ.

Das Verhalten im vorigen Beispiel war kein Zufall:

4.11 Satz. *Ist λ Eigenwert von $A \in M_n(K)$, so ist $1 \leq g(\lambda) \leq a(\lambda) \leq n$.*

BEWEIS: Sei $k := g(\lambda)$.
Dann gibt es eine Basis $(v_1, \ldots, v_k)$ von $E(\lambda)$. Wir ergänzen sie zu einer Basis $(v_1, \ldots, v_n)$ des ganzen K^n. Dann ist $B := (v_1, \ldots, v_n)$ eine invertierbare Matrix, und es gilt:

$$\begin{aligned}
B^{-1} \cdot A \cdot B &= B^{-1} \cdot (A \cdot (v_1, \ldots, v_k), A \cdot (v_{k+1}, \ldots, v_n)) \\
&= B^{-1} \cdot (\lambda v_1, \ldots, \lambda v_k, A \cdot (v_{k+1}, \ldots, v_n)) \\
&= (\lambda \cdot B^{-1} \cdot (v_1, \ldots, v_k), B^{-1} \cdot A \cdot (v_{k+1}, \ldots, v_n)) \\
&= \left(\begin{array}{ccc|c} \lambda & \cdots & 0 & \\ \vdots & \ddots & \vdots & C \\ 0 & \cdots & \lambda & \\ \hline & 0 & & D \end{array}\right),
\end{aligned}$$

mit irgendwelchen Matrizen $C \in M_{k,n-k}(K)$ und $D \in M_{n-k,n-k}(K)$. Daraus folgt:

$$\begin{aligned}
p_A(X) &= p_{B^{-1}AB}(X) = \det \left(\begin{array}{ccc|c} \lambda - X & \cdots & 0 & \\ \vdots & \ddots & \vdots & C \\ 0 & \cdots & \lambda - X & \\ \hline & 0 & & D - X \cdot E_{n-k} \end{array}\right) \\
&= (\lambda - X)^k \cdot \det(D - X \cdot E_{n-k}).
\end{aligned}$$

Also ist λ eine Nullstelle von mindestens k-ter Ordnung, es ist $a(\lambda) \geq g(\lambda)$.
Da $p_A(X)$ höchstens n Nullstellen haben kann, ist $a(\lambda) \leq n$, und da λ nur dann Eigenwert ist, wenn es dazu mindestens einen Eigenvektor gibt, ist $g(\lambda) \geq 1$. ∎

4.12 Satz (Charakterisierung der Diagonalisierbarkeit). *$A \in M_n(K)$ ist genau dann diagonalisierbar, wenn gilt:*

1. *$p_A(X)$ zerfällt vollständig in Linearfaktoren.*
2. *Für jeden Eigenwert λ ist $a(\lambda) = g(\lambda)$.*

BEWEIS: 1) Ist A diagonalisierbar, so gibt es eine Basis $B = (v_1, \ldots, v_n)$ von Eigenvektoren. Wir können die Basisvektoren so anordnen, dass gilt:

Die ersten k_1 von ihnen sind Eigenvektoren zum Eigenwert λ_1, die nächsten k_2 sind Eigenvektoren zum Eigenwert λ_2 usw., wobei die Eigenwerte $\lambda_1, \ldots, \lambda_r$ paarweise verschieden sind. Dann wird F_A bezüglich B durch die Diagonalmatrix

$$\Delta := \begin{pmatrix} \lambda_1 E_{k_1} & & & 0 \\ & \lambda_2 E_{k_2} & & \\ & & \ddots & \\ 0 & & & \lambda_r E_{k_r} \end{pmatrix}.$$

beschrieben, d.h. es ist $\Delta = M_B(F_A) = B^{-1} \cdot A \cdot B$. Daher gilt:

$$p_A(X) = p_\Delta(X) = (\lambda_1 - X)^{k_1} \cdots (\lambda_r - X)^{k_r}.$$

Das charakteristische Polynom zerfällt in Linearfaktoren, und es ist $a(\lambda_i) = k_i$ für $i = 1, \ldots, r$.

Wir wissen schon, dass $k_i \geq g(\lambda_i)$ ist. Aber $B_i := B \cap E(\lambda_i)$ besteht aus k_i linear unabhängigen Vektoren, die alle in $E(\lambda_i)$ liegen. Also ist auch $k_i \leq \dim E(\lambda_i) = g(\lambda_i)$.

2) Sei umgekehrt das Kriterium erfüllt, $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenwerte mit den Vielfachheiten $k_i = a(\lambda_i) = g(\lambda_i)$ und $k_1 + \cdots + k_r = n$.

Dann wählen wir Basen B_i von $E(\lambda_i)$, für $i = 1, \ldots, r$, und setzen

$$B := B_1 \cup \ldots \cup B_r.$$

Offensichtlich besteht B aus n Elementen. Eine Linearkombination w von Elementen von B kann in der Form $w = w_1 + \cdots + w_r$ geschrieben werden, wobei $w_i \in E(\lambda_i)$ jeweils Linearkombination von Elementen von B_i ist. Ist $w = 0$, so verschwinden nach dem Hilfssatz auch alle w_i, und da die Elemente von B_i linear unabhängig sind, kann w nur die triviale Linearkombination sein. Also sind die Elemente von B linear unabhängig und bilden somit eine Basis von Eigenvektoren von A, d.h. A ist diagonalisierbar. ■

Bemerkung. Über $\mathbb{C}$ zerfällt jedes Polynom in Linearfaktoren. Da ist die erste Bedingung überflüssig.

Definition. $W_1, \ldots, W_r \subset V$ seien Unterräume. Man sagt, V ist *direkte Summe* von $W_1, \ldots, W_r$, falls gilt:

1. $V = W_1 + \cdots + W_r$.
2. Ist $w_i \in W_i$ für $i = 1, \ldots, r$ und $w_1 + \cdots + w_r = 0$, so ist $w_1 = \ldots = w_r = 0$.

Man schreibt dann: $V = W_1 \oplus \ldots \oplus W_r$.

4.13 Satz. *$W_1, \ldots, W_r \subset V$ seien Unterräume. Dann sind folgende Aussagen äquivalent:*

1. *$V = W_1 \oplus \ldots \oplus W_r$.*
2. *Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung $v = w_1 + \cdots + w_r$ mit Vektoren $w_i \in W_i$, für $i = 1, \ldots, r$.*
3. *Es ist $V = W_1 + \cdots + W_r$ und $W_i \cap (W_{i+1} + \cdots + W_r) = \{0\}$ für $i = 1, \ldots, r-1$.*

Beweis: (1) $\Longrightarrow$ (2): Weil $V = W_1 + \cdots + W_r$ ist, besitzt jeder Vektor $v \in V$ eine Darstellung $v = w_1 + \cdots + w_r$ mit $w_i \in W_i$. Aus der zweiten Bedingung folgt, dass die Darstellung eindeutig ist.

(2) $\Longrightarrow$ (3): Dass $V = W_1 + \cdots + W_r$ ist, ist klar. Sei nun $v \in W_i \cap (W_{i+1} + \cdots + W_r)$. Wegen der Eindeutigkeit der Darstellung muss $v = 0$ sein.

(3) $\Longrightarrow$ (1): Sei $w_1 + \cdots + w_r = 0$, $w_i \in W_i$ für $i = 1, \ldots, r$. Dann liegt $w_1 = -(w_2 + \cdots + w_r)$ in $W_1 \cap (W_2 + \cdots + W_r)$. Also ist $w_1 = 0$. Analog zeigt man, dass $w_2 = \ldots = w_r = 0$ ist. ∎

4.14 Folgerung. *$A \in M_n(K)$ habe die paarweise verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_r$ mit $a(\lambda_i) = g(\lambda_i)$ für $i = 1, \ldots, r$ und $a(\lambda_1) + \cdots + a(\lambda_r) = n$. Dann ist*

$$K^n = E(\lambda_1) \oplus \ldots \oplus E(\lambda_r).$$

Eine kleine **Anwendung:**

Ist $B^{-1}AB = D := \Delta(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix, so ist

$$B^{-1}A^pB = \underbrace{(B^{-1}AB) \cdots (B^{-1}AB)}_{\text{p-mal}} = D^p.$$

Aber die Potenzen einer Diagonalmatrix kann man sofort hinschreiben, es ist $\Delta(\lambda_1, \ldots, \lambda_n)^p = \Delta(\lambda_1^p, \ldots, \lambda_r^p)$. Daraus folgt:

$$A^p = B \cdot \begin{pmatrix} \lambda_1^p & & 0 \\ & \ddots & \\ 0 & & \lambda_n^p \end{pmatrix} \cdot B^{-1}.$$

Das liefert eine leichte Berechnungsmöglichkeit für Potenzen von (diagonalisierbaren) Matrizen.

§ 5 Isometrien und Hauptachsentransformation

Der Körper wird hier immer $\mathbb{R}$ oder $\mathbb{C}$ sein.

Definition. $A \in M_n(\mathbb{R})$ heißt eine *orthogonale Matrix*, falls $A^t \cdot A = E_n$ ist.

5.1 Notwendiges Kriterium für Orthogonalität. *Ist A orthogonal, so ist $|\det A| = 1$.*

Beweis: Klar, denn wegen $A^t \cdot A = E_n$ ist $\det(A)^2 = 1$. ∎

Bemerkung. Ist $A^t \cdot A = E_n$, so ist offensichtlich $\det A \neq 0$. Daraus folgt, dass A invertierbar ist, und dann ist $A^{-1} = A^t$ und auch $A \cdot A^t = E_n$.

Die Menge aller orthogonalen Matrizen wird mit $O(n)$.

5.2 Satz. *$O(n)$ ist eine Gruppe.*

Beweis: Offensichtlich liegt die Einheitsmatrix E_n in $O(n)$. Sind A, B orthogonal, so ist

$$(A \cdot B) \cdot (A \cdot B)^t = A \cdot B \cdot B^t \cdot A^t = E_n,$$

also auch $A \cdot B$ orthogonal. Außerdem ist

$$A^{-1} \cdot (A^{-1})^t = A^t \cdot A^{tt} = A^t \cdot A = E_n,$$

also auch A^{-1} orthogonal. ∎

Definition. Eine lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ heißt eine *Isometrie*, falls gilt:

$$\|f(\mathbf{x})\| = \|\mathbf{x}\| \text{ für alle } \mathbf{x} \in \mathbb{R}^n.$$

5.3 Satz (Charakterisierung von Isometrien). *Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ linear und $A = M(f)$. Dann sind die folgenden Aussagen äquivalent:*

1. *f ist eine Isometrie.*
2. *f ist ein Isomorphismus, und es ist*
   $$f(x) \bullet f(y) = x \bullet y \text{ für alle } x, y \in \mathbb{R}^n.$$
3. *A ist orthogonal, d.h. es ist $A^t \cdot A = A \cdot A^t = E_n$.*
4. *Die Zeilen von A bilden eine ON-Basis.*
5. *Die Spalten von A bilden eine ON-Basis.*

BEWEIS:

Sind $x, y \in \mathbb{R}^n$, so ist $x \bullet y = x^t \cdot y = x_1 y_1 + \cdots + x_n y_n$.

(1) $\Longrightarrow$ (2): Ist $f(x) = 0$, so ist $\|x\| = \|f(x)\| = 0$, also auch $x = 0$. Damit ist f injektiv, und eine injektive lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ ist automatisch ein Isomorphismus.

Aus der Beziehung

$$\begin{aligned} \|x - y\|^2 &= (x - y) \bullet (x - y) \\ &= x \bullet x + y \bullet y - 2x \bullet y \\ &= \|x\|^2 + \|y\|^2 - 2x \bullet y, \\ \text{folgt:} \qquad x \bullet y &= \frac{1}{2} \cdot (\|x - y\|^2 - \|x\|^2 - \|y\|^2). \end{aligned}$$

Es ist $\|f(x)\|^2 = \|x\|^2$, $\|f(y)\| = \|y\|^2$ und

$$\|f(x) - f(y)\|^2 = \|f(x - y)\|^2 = \|x - y\|^2.$$

Daraus folgt: $f(x) \bullet f(y) = x \bullet y$.

(2) $\Longrightarrow$ (3): Weil $x \bullet y = f(x) \bullet f(y) = x^t \cdot A^t \cdot A \cdot y$ für alle x, y gilt, muss $A^t \cdot A = E_n$ sein.

Nach der Vorbemerkung ist dann aber auch $A \cdot A^t = E_n$.

(3) $\Longrightarrow$ (4): Sind $a_1, \ldots, a_n$ die Zeilen von A, so ist

$$E_n = A \cdot A^t = \Big(a_i \bullet a_j \ : \ i, j = 1, \ldots, n \Big),$$

also $(a_1, \ldots, a_n)$ eine ON-Basis.

(4) $\Longrightarrow$ (5): Bilden die Zeilen von A eine ON-Basis, so ist $A \cdot A^t = E_n$, und nach der Vorbemerkung ist dann auch $A^t \cdot A = E_n$. Daraus folgt wie beim vorigen Schritt, dass die Spalten von A eine ON-Basis bilden.

(5) $\Longrightarrow$ (1): Nach Voraussetzung ist $\{f(e_1), \ldots, f(e_n)\}$ eine ON-Basis. Ist $x = (x_1, \ldots, x_n)^t$, so folgt aus der Linearität von f und der Bilinearität des Slalarproduktes die Beziehung

$$\begin{aligned} \|f(x)\|^2 &= f\Big(\sum_{i=1}^n x_i e_i\Big) \bullet f\Big(\sum_{j=1}^n x_j e_j\Big) \\ &= \sum_{i,j} x_i x_j f(e_i) \bullet f(e_j) \\ &= \|x\|^2. \end{aligned}$$

Damit ist f eine Isometrie. ∎

Im Falle $n = 2$ wollen wir alle orthogonalen Matrizen bestimmen:

Sei also $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ orthogonal. Da dann die Spaltenvektoren ein ON-System bilden, gilt:

$$\|(a,c)^t\|^2 = 1 \quad \text{und} \quad (a,c)^t \bullet (b,d)^t = 0.$$

Also ist $a^2 + c^2 = 1$, und es gibt ein $\varphi \in [0, 2\pi)$, so dass $a = \cos(\varphi)$ und $c = \sin(\varphi)$ ist.

Weil $(a,c) \neq (0,0)$ ist, ist das orthogonale Komplement zu $\mathbb{R}(a,c)^t$ im $\mathbb{R}^2$ 1-dimensional, und weil $(a,c)^t \bullet (-c,a)^t = 0$ ist, gibt es ein $\lambda \in \mathbb{R}$ mit $(b,d)^t = \lambda \cdot (-c,a)^t$.

Wir müssen nun zwei Fälle unterscheiden:

1) Ist $\det(A) = 1$, so ist $1 = ad - bc = \lambda(a^2 + c^2) = \lambda$, also

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

Das ist die *Drehmatrix zum Winkel* φ, die uns schon früher begegnet ist.

$R(\varphi)$ bildet z.B. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ auf $\begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ auf $\begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}$ ab.

2) Sei $\det(A) = -1$.

Mit A und $S := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist auch $A \cdot S$ orthogonal, aber die neue Matrix besitzt die Determinante $+1$. Also gibt es ein φ, so dass $A = R(\varphi) \cdot S^{-1} = R(\varphi) \cdot S$ ist. Dabei ist $f_S(x,y) = (x,-y)$ die Spiegelung an der x–Achse.

Definition. Eine orthogonale Matrix $A \in M_n(\mathbb{R})$ heißt *Drehung*, falls $\det(A) = 1$ ist, und *Drehspiegelung*, falls $\det(A) = -1$ ist.

Die Menge aller Drehungen wird mit $SO(n)$ bezeichnet.

5.4 Satz. *$SO(n)$ eine Untergruppe von $O(n)$.*

Beweis: Offensichtlich liegt die Einheitsmatrix E_n in $SO(n)$. Wegen $\det(A \cdot B) = \det(A) \cdot \det(B)$ und $\det(A^{-1}) = \det(A)^{-1}$ folgt sofort, dass $SO(n)$ eine Untergruppe von $O(n)$ ist. ∎

Ist $A \in O(n)$, so gibt es zwei Möglichkeiten: Entweder ist schon $A \in SO(n)$, oder es ist $A \cdot S \in SO(n)$, mit

$$S := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & & & \\ 0 & 0 & 1 & & \\ \vdots & & & \ddots & 0 \\ 0 & & & 0 & 1 \end{pmatrix}.$$

Das heißt, jede orthogonale Transformation ist eine Drehung oder eine Drehspiegelung.

Im Komplexen sehen die Isometrien etwas anders aus.

Wir nennen $f : \mathbb{C}^n \to \mathbb{C}^n$ eine *Isometrie*, wenn $<f(z)\,,\, f(w)> = <z\,,\, w>$ für alle $z, w \in \mathbb{C}^n$ ist. Da $< z, w >= z^t \cdot \overline{w}$ ist, folgt:

5.5 Satz. *Sei $A \in M_n(\mathbb{C})$, $f = f_A : \mathbb{C}^n \to \mathbb{C}^n$ die zugehörige $\mathbb{C}$-lineare Abbildung. f ist genau dann eine Isometrie (bezüglich des kanonischen hermiteschen Skalarproduktes), wenn $\overline{A}^t \cdot A = E_n$ ist.*

Der BEWEIS funktioniert genauso wie beim euklidischen Skalarprodukt.

Definition. $A \in M_n(\mathbb{C})$ heißt eine *unitäre Matrix*, falls $\overline{A}^t \cdot A = E_n$ ist.

Die Gruppe der unitären Matrizen wird mit $U(n)$ bezeichnet, die Untergruppe aller unitären Matrizen A mit $\det(A) = 1$ wird mit $SU(n)$ bezeichnet.
Es gilt:

$$A \text{ unitär } \implies |\det(A)|^2 = 1.$$

Aber VORSICHT! Eine komplexe Zahl vom Betrag 1 hat die Gestalt $e^{\mathsf{i}\,t} = \cos t + \mathsf{i}\,\sin t$. Daher kann man die Unterscheidung zwischen Drehungen und Drehspiegelungen **nicht** auf unitäre Matrizen übertragen!

Es ist $SU(1) = \{1\}$ und $U(1) = \{z \in \mathbb{C}^* \,:\, |z| = 1\} = \{\cos(\varphi) + \mathsf{i}\,\sin(\varphi)\}$, also $U(1) \cong SO(2)$. Ebenso besteht ein enger Zusammenhang (aber keine Isomorphie) zwischen $SO(3)$ und $SU(2)$, auf den wir hier aber nicht eingehen können.

Wir wollen uns jetzt mit den Eigenwerten spezieller Matrizen befassen:

5.6 Eigenwerte von orthogonalen und unitären Matrizen.

Sei $A \in M_n(K)$ orthogonal (im Falle $K = \mathbb{R}$) oder unitär (im Falle $K = \mathbb{C}$). Dann gilt:

1. *Ist $\lambda \in K$ ein Eigenwert von A, so ist $|\lambda| = 1$.*
2. *Eigenvektoren zu verschiedenen Eigenwerten von A sind zueinander orthogonal.*

BEWEIS: Wir betrachten nur den komplexen Fall. Es sei $f := f_A$.

1) Sei $x \neq 0$, $f(x) = \lambda \cdot x$. Da f eine Isometrie ist, gilt:

$$\lambda\overline{\lambda} < x, x >=< \lambda x, \lambda x >=< f(x), f(x) >=< x, x > .$$

Daraus folgt, dass $|\lambda|^2 = 1$ ist.

2) Ist $f(x) = \lambda x$ und $f(y) = \mu y$, mit $\lambda \neq \mu$, so folgt:

$$\lambda\overline{\mu} < x, y >=< f(x), f(y) >=< x, y > .$$

Also ist $(\lambda\overline{\mu} - 1)\cdot < x, y >= 0$. Wäre $\lambda\overline{\mu} = 1$, so wäre $\overline{\mu} = \lambda^{-1} = \overline{\lambda}$ (weil $|\lambda| = 1$ ist). Aber dann wäre auch $\lambda = \mu$, im Widerspruch zur Voraussetzung. Also muss $< x, y >= 0$ sein. ∎

5.7 Eigenwerte von symmetrischen und hermiteschen Matrizen.

Sei $A \in M_n(\mathbb{R})$ symmetrisch (d.h. $A^t = A$) oder $A \in M_n(\mathbb{C})$ hermitesch (d.h. $\overline{A}^t = A$). Dann gilt:

1. *Das charakteristische Polynom von A zerfällt in Linearfaktoren.*
2. *Alle Eigenwerte von A sind reell.*
3. *Eigenvektoren zu verschiedenen Eigenwerten von A sind zueinander orthogonal.*

Beweis: Eine reelle symmetrische Matrix ist natürlich auch hermitesch, und über $\mathbb{C}$ zerfällt das charakteristische Polynom in Linearfaktoren. Wir brauchen also nur zu zeigen, dass alle Eigenwerte einer hermiteschen Matrix A reell sind, dann folgt (1) und (2) automatisch für $\mathbb{R}$ und $\mathbb{C}$. Es ist $<v\,,\,w> = v^t \cdot \overline{w}$ (in Matrizenschreibweise).

Wieder arbeiten wir mit der zugehörigen linearen Abbildung $f = f_A$. Diesmal ist f „selbstadjungiert", d.h. $< f(z), w >=< z, f(w) >$ für alle $z, w \in \mathbb{C}^n$, denn es ist

$$\begin{aligned} < f(z), w > &= (A \cdot z)^t \cdot \overline{w} \\ &= z^t \cdot A^t \cdot \overline{w} \\ &= z^t \cdot \overline{A} \cdot \overline{w} \\ &= z^t \cdot \overline{A \cdot w} = z^t \cdot \overline{f(w)} \\ &= < z, f(w) > . \end{aligned}$$

Ist $\lambda \in \mathbb{C}$ Eigenwert von f, so gibt es einen Vektor $z \in \mathbb{C}^n$ mit $z \neq 0$, so dass $f(z) = \lambda z$ ist. Dann folgt:

$$\begin{aligned} \lambda < z, z > &= < \lambda z, z > = < f(z), z > \\ &= < z, f(z) > \\ &= < z, \lambda z > = \overline{\lambda} < z, z > . \end{aligned}$$

Da $< z, z >$ reell und positiv ist, folgt: $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.

Zur Orthogonalität der Eigenvektoren:

Seien $\lambda \neq \mu$ zwei Eigenwerte, z bzw. w zugehörige Eigenvektoren. Dann gilt:

$$\begin{aligned} \lambda\cdot < z, w > &= < \lambda z, w > \\ &= < f(z), w > \\ &= < z, f(w) > \\ &= < z, \mu w > \\ &= \mu\cdot < z, w > \quad \text{(weil } \mu \text{ reell).} \end{aligned}$$

Also ist $(\lambda - \mu)\cdot < z, w >= 0$, d.h. $< z, w >= 0$. ∎

5.8 Satz von der Hauptachsentransformation. *Ist $A \in M_n(K)$ eine symmetrische (bzw. hermitesche) Matrix, so gibt es eine orthogonale (bzw. unitäre) Matrix $S \in \mathrm{GL}(n, K)$, so dass $S^{-1} \cdot A \cdot S$ eine reelle Diagonalmatrix ist. Die Einträge in der Diagonalmatrix sind die Eigenwerte der Matrix A.*

BEWEIS: Wir betrachten den komplexen Fall, also eine hermitesche Matrix A.

Es gibt mindestens einen Eigenwert λ und dazu einen Eigenvektor v_1. Diesen kann man normieren, so dass $\|v_1\| = 1$ ist. Schließlich setzen wir $U := \mathbb{C}v_1$.

Offensichtlich ist $f_A(U) \subset U$, und wir wollen zeigen, dass auch $f_A(U^\perp) \subset U^\perp$ ist. Dazu geben wir uns einen beliebigen Vektor $x \in U^\perp$ vor. Dann ist

$$<v_1\,,\, f_A(x)> \;=\; <f_A(v_1)\,,\, x> \;=\; <\lambda \cdot v_1\,,\, x> \;=\lambda \cdot <v_1\,,\, x> \;=0,$$

also auch $f_A(x) \in U^\perp$.

Wir wählen nun eine ON-Basis $\{v_2, \ldots, v_n\}$ von $U^\perp$. Dann ist $B := (v_1, v_2, \ldots, v_n)$ eine ON-Basis des $\mathbb{C}^n$ und damit auch eine unitäre Matrix. Sei

$$M := M_B(f_A) = B^{-1} \cdot A \cdot B = \overline{B}^{\,t} \cdot A \cdot B.$$

Dann ist $\overline{M}^{\,t} = \overline{B}^{\,t} \cdot \overline{A}^{\,t} \cdot B = \overline{B}^{\,t} \cdot A \cdot B = M$, also M wieder hermitesch.

Wir wollen M genauer berechnen. Wegen $f_A(U^\perp) \subset U^\perp$ gibt es Zahlen α_{ji}, so dass $w_i := A \cdot v_i = \sum_{j=1}^{n} \alpha_{ji} v_j$ ist, für $i = 2, \ldots, n$. Daraus folgt:

$$\begin{aligned} M &= \begin{pmatrix} \overline{v}_1^{\,t} \\ \vdots \\ \overline{v}_n^{\,t} \end{pmatrix} \cdot A \cdot (v_1, \ldots, v_n) \\ &= \begin{pmatrix} \overline{v}_1^{\,t} \\ \vdots \\ \overline{v}_n^{\,t} \end{pmatrix} \cdot (\lambda v_1, w_2, \ldots, w_n) \\ &= \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right). \end{aligned}$$

Weil M hermitesch ist, muss auch A_1 hermitesch sein. Wir werden sehen, dass wir damit das Problem um eine Dimension reduziert haben. Zu diesem Zweck nehmen wir an, es gebe schon eine unitäre Matrix S_1, so dass

$$\Delta_1 := \overline{S}_1^{\,t} \cdot A_1 \cdot S_1$$

eine Diagonalmatrix ist. Anschließend setzen wir $S := B \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array}\right)$. Dann ist auch S unitär, und es ist

$$\begin{aligned} S^{-1} \cdot A \cdot S &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \overline{S}_1^{\,t} \end{array}\right) \cdot \overline{B}^{\,t} \cdot A \cdot B \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array}\right) \\ &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \overline{S}_1^{\,t} \end{array}\right) \cdot \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & A_1 \end{array}\right) \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array}\right) \\ &= \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \overline{S}_1^{\,t} \cdot A_1 \cdot S_1 \end{array}\right) \\ &= \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \Delta_1 \end{array}\right). \end{aligned}$$

Das ist die gewünschte Diagonalisierung von A. Analog führt man die Diagonalisierung von A_1 auf eine $(n-2)$-reihige hermitesche Matrix A_2 zurück, usw. ∎

Um den Begriff „Hauptachsentransformation“ zu erläutern, sei noch ein Beispiel angegeben:

Beispiel.

Ist $A \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix, so nennt man $q_A : \mathbb{R}^n \to \mathbb{R}$ mit $q_A(x) := x^t \cdot A \cdot x$ die *zugehörige quadratische Form*. Ist z.B. $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, so ist

$$\begin{aligned} q_A(x_1, x_2) &= (x_1, x_2) \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1, x_2) \cdot \begin{pmatrix} ax_1 + bx_2 \\ bx_1 + cx_2 \end{pmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2. \end{aligned}$$

Für $c \in \mathbb{R}$ bezeichnet man die Menge $Q_A(c) := \{x : q_A(x) = c\}$ als *Quadrik* oder *quadratische Hyperfläche*.

Wir betrachten speziell $A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$. Dann ist

$$p_A(X) = \det \begin{pmatrix} 5 - X & -2 \\ -2 & 8 - X \end{pmatrix} = X^2 - 13X + 36 = (X-4)(X-9).$$

Einen Eigenvektor zum Eigenwert $\lambda_1 = 4$ finden wir durch die Gleichung

$$\begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eine Lösung ist der Vektor $v_1 = (2,1)^t$, mit $\|v_1\| = \sqrt{5}$.

Einen Eigenvektor zum Eigenwert $\lambda_2 = 9$ erhalten wir aus der Gleichung

$$\begin{pmatrix} -4 & -2 \\ -2 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eine Lösung ist $v_2 = (-1,2)^t$, mit $\|v_2\| = \sqrt{5}$.

Dann bilden die Vektoren $w_1 := \frac{1}{\sqrt{5}}(2,1)^t$ und $w_2 := \frac{1}{\sqrt{5}}(1,-2)^t$ eine ON-Basis von Eigenvektoren von A. Die zugehörige Matrix

$$S = \begin{pmatrix} \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \\ \dfrac{1}{\sqrt{5}} & \dfrac{-2}{\sqrt{5}} \end{pmatrix}$$

definiert eine Drehung f_S der Ebene um den Nullpunkt, und es ist

$$S^t \cdot A \cdot S = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}.$$

Nun gilt (mit $y := (f_S)^{-1}(x) = S^{-1} \cdot x = S^t \cdot x$) :

$$\begin{aligned} x \in Q_A(c) &\iff q_A(x) = c \\ &\iff x^t \cdot A \cdot x = c \\ &\iff x^t \cdot (S \cdot \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \cdot S^t) \cdot x = c \\ &\iff y^t \cdot \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \cdot y = c. \end{aligned}$$

Sei etwa $c = 36$. Dann ist

$$\begin{aligned} Q_A(c) &= \{x \,:\, 5x_1^2 - 4x_1x_2 + 8x_2^2 = 36\} \\ &= f_S(\{y \,:\, 4y_1^2 + 9y_2^2 = 36\}) \\ &= f_S(\{y \,:\, (\frac{y_1}{3})^2 + (\frac{y_2}{2})^2 = 1\} \end{aligned}$$

das Bild einer achsenparallelen Ellipse unter der Drehung f_S. So haben wir die „Hauptachsen“ von $Q_A(c)$ bestimmt, sie zeigen in die durch die Spalten von S vorgegebenen Richtungen und haben die Halbachsen 2 und 3.