Skript zur Vorlesung

Lineare Algebra 1

Sommersemester 2015

Prof. Dr. Annette Werner

Inhaltsverzeichnis

1 Matrizenkalkül

Mit Matrizen lassen sich lineare Phänomene berechnen. Ferner liefern sie wichtige Beispiele für die Theorie, die wir später behandeln werden. Daher beginnen wir mit Matrizen und ihren Rechenregeln. Wir verwenden generell folgende Bezeichnungen:

Es sei

$\mathbb{N} = \{1, 2, 3, \ldots\}$ die Menge der natürlichen Zahlen,

$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ die Menge der natürlichen Zahlen mit 0,

$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ die Menge der ganzen Zahlen,

$\mathbb{Q} = \left\{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\right\}$ die Menge der rationalen Zahlen,

$\mathbb{R}$ die Menge der reellen Zahlen.

Definition 1.1 *Seien m und n natürliche Zahlen. Eine $m \times n$-**Matrix** über $\mathbb{R}$ besteht aus mn reellen Zahlen, die in einem rechteckigen Schema angeordnet sind:*

$$\underbrace{\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}}_{n \text{ Spalten}} \quad m \text{ Zeilen}$$

*Die Zahlen a_{ij} für $i = 1, \ldots, m$ und $j = 1, \ldots, n$ heißen **Matrixeinträge**. Der Index i heißt **Zeilenindex**, der Index j heißt **Spaltenindex**. Also ist a_{ij} der Eintrag, der in der i-ten Zeile und der j-ten Spalte steht.*

Wir schreiben eine Matrix auch als $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ oder einfach als $A = (a_{ij})$, wenn klar ist, welche Größe A hat.

Wir werden in § 1 erst einmal Matrizen über $\mathbb{R}$ behandeln. Alle Definitionen und Ergebnisse gelten aber auch, wenn man statt reeller Zahlen Elemente eines beliebigen Körpers (in vielen Fällen sogar eines beliebigen Ringes) betrachtet. Was ein Körper ist, werden wir aber erst in § 4 lernen.

Eine $1 \times n$-Matrix heißt n-dimensionaler **Zeilenvektor**, wir schreiben sie auch als $A = (a_1 \dots a_n)$ oder $A = (a_1, \dots, a_n)$. Analog heißt eine $m \times 1$-Matrix auch m-dimensionaler **Spaltenvektor**, und wir schreiben sie auch als

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

Definition 1.2 *i) Für eine $m \times n$-Matrix $A = (a_{ij})$ ist $-A$ definiert als die $m \times n$-Matrix $-A = (-a_{ij})$.*

ii) Für zwei $m \times n$-Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ ist die Summe $A + B$ definiert als die $m \times n$-Matrix

$$A + B = (a_{ij} + b_{ij}).$$

Bei der Addition von Matrizen werden also jeweils die Einträge an gleicher Position addiert. Die Addition ist nur definiert für Matrizen mit derselben Zeilen- und Spaltenzahl.

Beispiel:

$$\begin{pmatrix} 1 & 0 & 2 \\ 3 & -1 & 7 \end{pmatrix} + \begin{pmatrix} -5 & 3 & 1 \\ 0 & 2 & -1 \end{pmatrix} = \begin{pmatrix} -4 & 3 & 3 \\ 3 & 1 & 6 \end{pmatrix}$$

Wir schreiben für $m \times n$-Matrizen A und B auch $A - B = A + (-B)$. Offenbar ist $A - A = 0$, wobei 0 hier die $m \times n$-Matrix bezeichnet, deren Einträge alle 0 sind. Diese heißt Nullmatrix. Um Mehrdeutigkeiten zu vermeiden, müssten wir eigentlich $0_{m \times n}$ schreiben. Das ist uns aber meist zu umständlich. Aus dem Kontext wird klar werden, welche Nullmatrix wir mit 0 jeweils meinen.

Definition 1.3 *Es sei $A = (a_{ij})$ eine $m \times n$-Matrix über $\mathbb{R}$ und $c \in \mathbb{R}$. Die **skalare Multiplikation** von A mit der Zahl (man sagt auch Skalar) c ist definiert als die $m \times n$-Matrix*

$$cA = (ca_{ij}).$$

Beispiel: Es ist

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

Zusätzlich zu der skalaren Multiplikation kann man auch das Produkt zweier Matrizen A und B definieren, wenn ihre Größen zusammenpassen.

Definition 1.4 *i) Es sei $n \in \mathbb{N}$ und $A = (a_1 \dots a_n)$ ein n-dimensionaler Zeilenvektor (d.h. eine $1 \times n$-Matrix) und $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ ein n-dimensionaler Spaltenvektor (d.h. eine $n \times 1$ Matrix). Dann ist das **Produkt** AB definiert als die 1×1-Matrix mit dem (einzigen) Eintrag*

$$a_1 b_1 + a_2 b_2 + \dots a_n b_n.$$

Diesen Ausdruck können wir unter Verwendung des Summenzeichens $\sum$ auch schreiben als $\sum\limits_{i=1}^{n} a_i b_i$.

*ii) Es seien l, m und n natürliche Zahlen und A eine $l \times m$ sowie B eine $m \times n$ Matrix. Dann ist das **Produkt** AB definiert als die $l \times n$-Matrix $AB = (c_{ij})_{\substack{i=1\dots l \\ j=1\dots n}}$ mit den Einträgen*

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{im} b_{mj}.$$

c_{ij} ist also das Produkt im Sinne von i) der i-ten Zeile $(a_{i1} \dots a_{im})$ von A mit der j-ten Spalte von B.

Beispiel:

i) $\begin{pmatrix} 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} = -3 + 0 + 6 = 3$

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

iii) $\begin{pmatrix} 3 & 1 \\ 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 4 & 1 \\ 0 & -1 & -1 \\ 2 & 3 & 2 \end{pmatrix}$

iv) $\begin{pmatrix} 3 & 1 \\ 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$ ist nicht definiert!

Lemma 1.5 *Die Addition und Multiplikation von Matrizen erfüllt folgende Regeln:*

i) *Für $l \times m$-Matrizen A und A' sowie $m \times n$-Matrizen B und B' gelten die Distributivgesetze*

$$\begin{aligned} A(B+B') &= AB + AB' \quad \textit{und} \\ (A+A')B &= AB + A'B. \end{aligned}$$

ii) *Ist A eine $l \times m$-, B eine $m \times n-$ und C eine $n \times p$-Matrix, so gilt das Assoziativgesetz*

$$(AB)C = A(BC).$$

Bei der Matrizenmultiplikation kann man also beliebig klammern. Daher schreiben wir auch einfach ABC für $(AB)C$.

Beweis : in den Übungen. □

Lemma 1.6 *Die skalare Multiplikation ist verträglich mit der Addition und der Multiplikation von Matrizen. Für reelle Zahlen (Skalare) α, β und $m \times n$-Matrizen A und B gilt:*

i) $(\alpha\beta)A = \alpha(\beta A)$

ii) $(\alpha + \beta)A = \alpha A + \beta A$

iii) $\alpha(A + B) = \alpha A + \alpha B.$

Für eine $m \times n$-Matrix A und eine $n \times p$-Matrix B gilt ferner

iv) $\alpha(AB) = (\alpha A)B = A(\alpha B).$

Beweis : geduldig prüfen! □

Ist A eine $l \times m$- und B eine $m \times l$-Matrix, so sind beide Produkte AB und BA definiert. AB ist eine $l \times l$- und BA eine $m \times m$-Matrix. Ist $l = m$, so gilt im allgemeinen $AB \neq BA$, die Multiplikation von Matrizen ist also im allgemeinen nicht kommutativ.

Beispiel:

$$\begin{aligned} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}, \\ \text{aber} \quad & \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}. \end{aligned}$$

Die Einträge a_{ii} einer $n \times n$-Matrix $A = (a_{ij})$ heißen **Diagonaleinträge.** A heißt **Diagonalmatrix**, falls für alle $i \neq j$ der Eintrag $a_{ij} = 0$ ist. Eine Diagonalmatrix A hat also die Form

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Wir werden oft mit $*$ einen beliebigen Eintrag einer Matrix bezeichnen und durch eine einzige Null kennzeichnen, dass ein ganzer Bereich der Matrix nur aus Nullen besteht. Mit dieser Notation sieht eine Diagonalmatrix also so aus:

$$A = \begin{pmatrix} * & & 0 \\ & \ddots & \\ 0 & & * \end{pmatrix}.$$

Eine Matrix der Form

$$A = \begin{pmatrix} * & \dots & * \\ & \ddots & \vdots \\ 0 & & * \end{pmatrix}$$

heißt **obere Dreiecksmatrix**. Die $n \times n$-Matrix $A = (a_{ij})$ ist also genau dann eine obere Dreiecksmatrix, wenn für alle $i > j$ der Eintrag $a_{ij} = 0$ ist. Analog definiert man untere Dreiecksmatrizen.

Die $n \times n$-Diagonalmatrix, deren Diagonaleinträge alle 1 sind, heißt n-te Einheitsmatrix:

$$E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Wie man leicht nachrechnet, gilt für jede $m \times n$-Matrix A:

$$E_m A = A \text{ und } AE_n = A.$$

Wir nehmen eine $m \times n$-Matrix A **quadratisch**, falls $m = n$ ist. Eine quadratische Matrix A hat also genauso viele Zeilen wie Spalten.

Definition 1.7 *Sei A eine $n \times n$-Matrix (d.h. A ist quadratisch). A heißt **invertierbar**, falls es eine $n \times n$-Matrix B mit*

$$AB = E_n \textit{ und } BA = E_n$$

gibt.

Die Matrix B ist durch diese Gleichungen eindeutig bestimmt, denn aus

$$AB = E_n \text{ und } B'A = E_n$$

für zwei $n \times n$-Matrizen B und B' folgt

$$B' = B'E_n = B'(AB) \stackrel{1.5}{=} (B'A)B = E_nB = B.$$

Daher schreiben wir auch $B = A^{-1}$ und nennen diese Matrix die **Inverse** zu A.

Wir werden später sehen, dass eine Matrix B, die eine der beiden Gleichungen aus 1.7 erfüllt, automatisch auch die andere Gleichung erfüllt. Da die Matrizenmultiplikation nicht kommutativ ist, ist das nicht offensichtlich.

Beispiel:

i) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, wie man leicht nachrechnet.

ii) Die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist nicht invertierbar, denn für jede 2×2-Matrix B besteht die letzte Spalte von $B \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ nur aus Nullen.

Lemma 1.8 *A und B seien $n \times n$-Matrizen. Sind beide invertierbar, so ist auch ihr Produkt AB invertierbar, und es gilt*

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

Sind allgemeiner die $(n \times n)$-Matrizen $A_1, \ldots, A_m$ invertierbar, so auch das Produkt $A_1 \cdots A_m$, und es gilt $(A_1 \cdots A_m)^{-1} = A_m^{-1} \cdots A_1^{-1}$.

Beweis : Sind A und B invertierbar, so ist

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = (AE_n)A^{-1} \\ &= AA^{-1} = E_n. \end{aligned}$$

Analog rechnet man $(B^{-1}A^{-1})(AB) = E_n$ nach.

Den allgemeinen Fall beweisen wir durch Induktion nach m. Den Anfang $m = 2$ haben wir gerade gezeigt.

Für den Induktionsschluss nehmen wir an, die Behauptung gelte für m Matrizen. Wir müssen sie für $m+1$ Matrizen $A_1, \ldots, A_{m+1}$ zeigen. Nach Induktionsvoraussetzung ist $B := A_1 \cdots A_m$ invertierbar, und es gilt $B^{-1} = A_m^{-1} \cdots A_1^{-1}$. Nach dem schon gezeigten Fall für zwei Matrizen ist $A_1 \cdots A_{m+1} = BA_{m+1}$ invertierbar mit inverser Matrix $A_{m+1}^{-1}B^{-1} = A_{m+1}^{-1}A_m^{-1} \cdots A_1^{-1}$. $\square$

Wir bezeichnen die Menge der invertierbaren $n \times n$-Matrizen über $\mathbb{R}$ als $\mathrm{GL}_n(\mathbb{R})$. Hier steht GL für „general linear group“. Sobald wir wissen, was das ist, werden wir sehen, dass $\mathrm{GL}_n(\mathbb{R})$ eine Gruppe ist.

2 Elementare Zeilenumformungen

Wir werden jetzt ein Verfahren untersuchen, mit dem man eine Matrix vereinfachen kann. Auf diese Weise lösen wir lineare Gleichungssysteme.

Es sei $A = (a_{ij})$ eine $m \times n$-Matrix und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ ein m-dimensionaler Spaltenvektor. Das System von Gleichungen

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

heißt **lineares Gleichungssystem** in den Unbestimmten $x_1, \ldots, x_n$ mit **Koeffizientenmatrix** A. Mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ können wir es kurz als Matrixprodukt

$$Ax = b$$

schreiben.

Ist $b = 0$, so bezeichnet man (1) als **homogenes** lineares Gleichungssystem. Ist $b \neq 0$, so nennt man (1) auch **inhomogenes** lineares Gleichungssystem.

Ein homogenes lineares Gleichungssystem hat stets als triviale Lösung

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

ein inhomogenes lineares Gleichungssystem muss keine Lösung besitzen.

Beispiel

i) Sei $A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ und $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. Das zugehörige lineare Gleichungssystem ist

$$\begin{array}{rrrcl} 2x_1 & & +x_3 & = & b_1 \\ x_1 & -x_2 & +2x_3 & = & b_2. \end{array}$$

Ist $b = 0$, so hat das homogene lineare Gleichungssystem die Lösungsmenge

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_3 = -2x_1, x_2 = -3x_1 \right\},$$

ii) Sei $A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \\ 1 & 2 \end{pmatrix}$ und $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Dann hat das zugehörige inhomogene lineare Gleichungssystem

$$\begin{array}{rrcl} 2x_1 & +x_2 & = & 1 \\ & -x_2 & = & 1 \\ x_1 & +2x_2 & = & 1 \end{array}$$

gar keine Lösung.

Wir definieren jetzt die sogenannten Elementarmatrizen, mit deren Hilfe wir die Koeffizientenmatrix eines linearen Gleichungssystems vereinfachen können.

Mit e_{ij} bezeichnen wir die $n \times n$-Matrix, deren Eintrag an der Stelle (i, j) gleich 1 ist, und deren andere Einträge alle Null sind. Dann erfüllt jede $n \times n$-Matrix $A = (a_{ij})_{i,j=1,\cdots n}$ die Gleichung

$$\begin{aligned} A &= a_{11}e_{11} + a_{12}e_{12} + \cdots + a_{nn}e_{nn} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}e_{ij}. \end{aligned}$$

Definition 2.1 *Sei $n \geq 1$. Eine $n \times n$-Matrix L heißt* ***Elementarmatrix***

i) vom ***Typ I****, falls*

$$L = E_n + ce_{ij}$$

für Indizes $i \neq j$ und einen Skalar c ist, *ii) vom **Typ II**, falls*

$$L = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & \dots & 1 & & & \\ & & & \vdots & & \vdots & & & \\ & & & 1 & \dots & 0 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} = E_n + e_{ij} + e_{ji} - e_{ii} - e_{jj}$$

für Indizes $i \neq j$ ist,

*iii) vom **Typ III**, falls*

$$L = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & c & & \\ & & & & 1 & \\ & & & & & \ddots \\ 0 & & & & & 1 \end{pmatrix} = E_n + (c-1)e_{ii}$$

für einen Index i und einen Skalar $c \neq 0$ ist.

Beispiel: Die 2×2-Elementarmatrizen sind genau folgende

i) $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$, c beliebig

ii) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

iii) $\begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}$, $c \neq 0$.

Lemma 2.2 *Es sei A eine $n \times n$-Matrix und L eine Elementarmatrix. Mit a_i bezeichnen wir die i-te Zeile von A, d.h. es ist $a_i = (a_{i1} \dots a_{in})$.*

i) Ist $L = E_n + ce_{ij}$ vom Typ I, so entsteht die Matrix LA aus A, indem die i-te Zeile a_i durch $a_i + ca_j$ ersetzt wird.

ii) Ist $L = E_n + e_{ij} + e_{ji} - e_{ii} - e_{jj}$ vom Typ II, so entsteht die Matrix LA aus A, indem die Zeilen a_i und a_j vertauscht werden.

iii) Ist $L = E_n + (c-1)e_{ii}$ vom Typ III, so entsteht die Matrix LA aus A, indem die Zeile a_i mit dem Skalar c multipliziert wird.

Beweis : Nach Definition der Matrixmultiplikation ist die i-te Zeile von LA gerade das Produkt der i-ten Zeile von L mit der Matrix A. Die Behauptung lässt sich nun leicht nachrechnen. Sei $L = E_n + ce_{ij}$ eine Elementarmatrix vom Typ I. Da L in allen Zeilen bis auf die i-te mit der Einheitsmatrix übereinstimmt, sind alle Zeilen von LA bis auf die i-te gleich den entsprechenden Zeilen von A.

Die i-te Zeile hat die Gestalt

$$(a_{i1} + ca_{j1} \dots a_{in} + ca_{jn}) = a_i + ca_j,$$

wie behauptet. $\square$

Für Elementarmatrizen vom Typ II und III argumentiert man analog. (Prüfen Sie das nach!) Die in Lemma 2.2 beschriebenen Operationen nennt man auch **elementare Zeilenumformungen** vom Typ I, II bzw.III. Entsteht B also aus A durch Muliplikation mit einer Elementarmatrix vom Typ I, II bzw. III von links, so geht B aus A durch eine elementare Zeilenumformung vom Typ I, II bzw. III hervor.

Lemma 2.3 *Elementarmatrizen sind invertierbar, und ihre Inversen sind wieder Elementarmatrizen.*

Beweis : Wir prüfen nach, dass das Inverse einer Elementarmatrix jeweils die Elementarmatrix zur inversen Zeilenumformung ist. Ist $L = E_n + ce_{ij}$ vom Typ I, so bewirkt L das Ersetzen der i-ten Zeile durch (i-te Zeile) + c(j-te Zeile). Dies kann man rückgängig machen, indem man die i-te Zeile durch (i-te Zeile) $-c$(j-te Zeile) ersetzt. Das entspricht der Linksmultiplikation mit $L' = E_n - ce_{ij}$. L' ist also unser Kandidat für die Inverse. Wir rechnen leicht nach, dass

$$LL' = E_n - ce_{ij} + ce_{ij} = E_n \text{ und } L'L = E_n$$

gilt. Hier geht $e_{ij} \cdot e_{ij} = 0$ $(i \neq j)$ ein. Also ist $L^{-1} = E - ca_{ij}$.

Genauso zeigt man, dass für $L = E_n + e_{ij} + e_{ji} - e_{ii} - e_{jj}$ $(i \neq j)$ gilt $LL = E_n$, d.h. $L = L^{-1}$, und dass für $L = E_n + (c-1)e_{ii}$ $(c \neq 0)$ gilt $L(E_n + (\frac{1}{c} - 1) \cdot e_{ii}) = (E_n + (\frac{1}{c} - 1)e_{ii})L = E$, d.h. $L^{-1} = E_n + (\frac{1}{c} - 1)e_{ii}$. $\square$

Es sei $Ax = b$ ein lineares Gleichungssystem in n Unbestimmten und m Gleichungen, d.h. $A = (a_{ij})$ ist eine $m \times n$-Matrix, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ein unbekannter n-

dimensionaler Spaltenvektor und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ ein gegebener m-dimensionaler Spaltenvektor.

Satz 2.4 *Es sei L eine $m \times m$ Elementarmatrix. Die Lösungen von $Ax = b$ stimmen mit den Lösungen von $(LA)x = Lb$ überein. Wendet man also auf die Koeffizientenmatrix A eines linearen Gleichungssystems und auf den Vektor b dieselben elementaren Zeilenumformungen an, so ändern sich die Lösungen nicht.*

Beweis : Gilt $Ax = b$, so folgt durch Linksmultiplikation mit L, dass auch

$$(LA)x = L(Ax) = Lb$$

gilt. Umgekehrt folgt aus $(LA)x = Lb$ durch Linksmultiplikation mit der Inversen L^{-1}, die nach Lemma 2.3 existiert,

$$Ax = (L^{-1}L)Ax = L^{-1}(LA)x = L^{-1}(Lb) = b.$$

□

Wir können also die Koeffizientennmatrix eines linearen Gleichungssystems durch elementare Zeilenumformungen vereinfachen. Wenn wir diese Umformungen gleichzeitig auf den Vektor b anwenden, ändert sich die Lösungsmenge nicht.

Beispiel: Sei

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 5 & 2 \\ 1 & 2 & 8 & 4 \end{pmatrix} \text{ und } b = \begin{pmatrix} 5 \\ 7 \\ 12 \end{pmatrix},$$

wir suchen also die Lösungen des (inhomogenen) linearen Gleichungssystems

$$\begin{array}{ccccccc} x_1 & & +2x_3 & +x_4 & = & 5 \\ x_1 & +x_2 & +5x_3 & +2x_4 & = & 7 \\ x_1 & +2x_2 & +8x_3 & +4x_4 & = & 12. \end{array}$$

Wir betrachten die Matrix $(Ab) = \begin{pmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{pmatrix}$, die durch Anhängen der Spalte b an die Matrix A entsteht.

Indem wir die 2. Zeile durch (2. Zeile) $-$ (1. Zeile) und die 3. Zeile durch (3. Zeile) $-$ (1. Zeile) ersetzen (das entspricht der Multiplikation mit $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ von links), erhalten wir $\begin{pmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 2 & 6 & 3 & 7 \end{pmatrix}$.

Jetzt ersetzen wir die 3. Zeile durch (3. Zeile) $- 2 \cdot$ (2. Zeile), multiplizieren also mit

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$ von links, und erhalten

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

Jetzt ersetzen wir die 1. Zeile durch (1. Zeile) $-$ (3. Zeile) und die 2. Zeile durch (2. Zeile) $-$ (3. Zeile), multiplizieren also mit

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ von links, und erhalten

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

Die Lösungen von $Ax = b$ sind nach Satz 2.4 genau die Lösungen von

$$\begin{array}{cccccc} x_1 & & +2x_3 & & = & 2 \\ & x_2 & +3x_3 & & = & -1 \\ & & & x_4 & = & 3. \end{array}$$

Dies kann man leicht auflösen, die Lösungsmenge ist

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} : x_3 \text{ beliebig }, x_4 = 3, x_2 = -1 - 3x_3, x_1 = 2 - 2x_3 \right\}.$$

Dieses Verfahren zur Lösung von linearen Gleichungssystemen nennt man **Gauß'sches Eliminationsverfahren**. Wir wollen es jetzt für ein beliebiges Gleichungssystem herleiten.

Definition 2.5 *Eine $m \times n$-Matrix A hat **Zeilenstufenform**, wenn folgende Bedingungen gelten:*

*i) Der erste von Null verschiedene Eintrag in jeder Zeile ist 1. Dieser heißt auch **Leiteintrag**.*

ii) Wenn in Zeile $i+1$ ein Leiteintrag steht, dann auch in Zeile i. In diesem Falle steht der Leiteintrag in Zeile $(i+1)$ rechts vom Leiteintrag in Zeile i.

iii) Alle Einträge oberhalb eines Leiteintrags sind Null.

Eine Matrix in Zeilenstufenform sieht also so aus:

$$\begin{pmatrix}
0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & \dots & 0 & * & \dots & * \\
0 & \dots & & & & \dots & 0 & 1 & * & \dots & * & 0 & \dots & 0 & \vdots & & \vdots \\
0 & \dots & & & & & & & & \dots & 0 & 1 & & 0 & & & \\
\vdots & & & & & & & & & & & & & \vdots & \vdots & & \vdots \\
0 & \dots & & & & & & & & & & & \dots & 1 & * & \dots & * \\
0 & \dots & & & & & & & & & & & \dots & 0 & 0 & \dots & 0 \\
\vdots & & & & & & & & & & & & & & \vdots & & \vdots \\
0 & \dots & & & & & & & & & & & \dots & 0 & 0 & \dots & 0
\end{pmatrix}$$

Sowohl die Nullspalten links als auch die Nullzeilen unten müssen nicht auftreten.

Satz 2.6 *Jede Matrix lässt sich mit elementaren Zeilenumformungen in eine Matrix in Zeilenstufenform überführen. Mit anderen Worten: Ist A eine $m \times n$-Matrix, so gibt es $m \times m$ Elementarmatrizen $L_1, \dots, L_s$, so dass die Matrix $L_s \cdots L_1 A$ Zeilenstufenform hat.*

Beweis : mit Induktion nach der Zeilenzahl m.

Für den Induktionsanfang betrachten wir eine $1 \times n$-Matrix, also einen n-dimensionalen Zeilenvektor $A = (a_1 \dots a_n)$. Sind alle $a_i = 0$, so hat A Zeilenstufenform. Ansonsten ist mindestens ein $a_i \neq 0$. Ist i minimal mit $a_i \neq 0$, so multiplizieren wir A mit a_i^{-1} (das ist eine elementare Zeilenumformung vom Typ III) und erhalten $(0 \dots 0 1 * \cdots *)$, also eine Matrix in Zeilenstufenform.

Für den Induktionsschluss nehmen wir an, dass wir jede Matrix mit höchstens $(m-1)$ Zeilen in Zeilenstufenform überführen können. Ist $A = 0$, dann hat A Zeilenstufenform. Wir können also $A \neq 0$ annehmen. In diesem Fall suchen wir die erste Spalte, die einen Eintrag $\neq 0$ enthält und transportieren diesen durch eine Zeilenvertauschung (Typ II) in die erste Zeile. Wir multiplizieren die erste Zeile mit dem Inversen dieses Eintrags (Typ III) und erhalten eine Matrix der Form

$$\begin{pmatrix} 0 & \dots & 0 & 1 & \\ \vdots & & \vdots & b_2 & * \\ \vdots & & \vdots & \vdots & \\ 0 & \dots & 0 & b_m & \end{pmatrix}$$

Nun ersetzen wir mit Umformungen vom Typ I die 2. Zeile durch (2. Zeile) $-b_2$ (1. Zeile), die 3. Zeile durch (3. Zeile) $-b_3$ (1. Zeile) usw. Das Ergebnis ist eine Matrix der Form

$$A' = \left(\begin{array}{c} \overbrace{\hphantom{0 \quad \dots \quad 0 \quad 1}}^{n-r}\overbrace{\hphantom{c_1 \quad \dots \quad c_r}}^{r} \\ \begin{matrix} 0 & \dots & 0 & 1 & c_1 & \dots & c_r \\ \vdots & & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & D & \\ 0 & \dots & 0 & 0 & & & \end{matrix} \end{array} \right)$$

mit unbekannten Einträgen $c_1, \dots, c_r$ und einer $(m-1) \times r$-Matrix D für ein geeignetes $r < n$. Nach Induktionsvoraussetzung lässt sich D mit elementaren Zeilenumformungen auf Zeilenstufenform bringen. Diese Zeilenumformungen wenden wir auch auf A' an. Sie lassen die ersten $(n-r)$ Spalten und die erste Zeile unverändert. Wir erhalten eine Matrix der Form

$$A'' = \begin{pmatrix} 0 & \dots & 0 & 1 & c_1 & \dots & c_r \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & D' & & \\ 0 & \dots & 0 & 0 & & & \end{pmatrix},$$

wobei D' Zeilenstufenform hat. Der Leiteintrag aus der ersten Zeile von D' stehe unter dem Eintrag c_i. Dann ersetzen wir mit einer Umformung vom Typ I die erste Zeile von A'' durch (1. Zeile) $-c_i$ (2. Zeile) und haben c_i eliminiert. Dieses Verfahren führen wir für alle Leiteinträge von D' durch. Das Ergebnis ist eine Matrix in Zeilenstufenform. □

Man kann zeigen, dass die Zeilenstufenform, in die man eine Matrix durch elementare Zeilenumformungen bringen kann, eindeutig bestimmt ist. Sie hängt also nicht von der gewählten Folge elementarer Umformungen ab. Wie wir in dem obigen Beispiel schon gesehen haben, lässt sich ein lineares Gleichungssystem, dessen Koeffizientenmatrix Zeilenstufenform hat, leicht lösen. Das formulieren wir jetzt ganz allgemein.

Satz 2.7 *Es sei $Ax = b$ ein lineares Gleichungssystem in m Gleichungen und n Variablen, so dass die $m \times (n+1)$-Matrix (Ab), die durch Anhängen der Spalte b an A entsteht, Zeilenstufenform hat. Dann besitzt $Ax = b$ genau dann eine Lösung, wenn in der letzten Spalte b kein Leiteintrag steht. In diesem Fall kann man für jede Unbestimmte x_j, so dass die j-te Spalte keinen Leiteintrag enthält, einen beliebigen Wert vorschreiben.*

Beweis : Es sei $A = (a_{ij})$ und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. Hat (Ab) Zeilenstufenform, so sei $J = \{j_1, \ldots, j_r\} \subset \{1, \ldots, n+1\}$ die Menge aller Spaltenindizes j, so dass die j-te Spalte von (Ab) einen Leiteintrag enthält. Falls in der letzten Spalte ein Leiteintrag steht, d.h. falls $n+1 \in J$ ist, so lautet eine Gleichung des linearen Gleichungssystems $0 = 1$. Das ist natürlich unerfüllbar.

Falls in der letzten Spalte von (Ab) kein Leiteintrag steht, d.h. falls $n+1 \notin J$ ist, so sieht das Gleichungssystem $Ax = b$ folgendermaßen aus:

$$\begin{aligned} x_{j_1} + \sum_{j \notin J} a_{1j} x_j &= b_1 \\ x_{j_2} + \sum_{j \notin J} a_{2j} x_j &= b_2 \\ &\vdots \\ x_{j_r} + \sum_{j \notin J} a_{rj} x_j &= b_r. \end{aligned}$$

Die restlichen Gleichungen sind von der Form $0 = 0$. Die Lösungsmenge von $Ax = b$ ist also folgende:

$$\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_j \text{ beliebig für } j \notin J, x_{j_1} = b_1 - \sum_{j \notin J} a_{1j} x_j, \ldots, x_{j_r} = b_r - \sum_{j \notin J} a_{rj} x_j \right\}.$$

Insbesondere ist sie nicht leer. □

Korollar 2.8 *Jedes lineare Gleichunssystem $Ax = 0$ in m Gleichungen und n Unbekannten mit $m < n$ hat eine nicht-triviale Lösung, d.h. eine Lösung $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, für die nicht alle x_i Null sind.*

Beweis : Mit elementaren Zeilenumformungen lässt sich A in eine Matrix A' in Zeilenstufenform überführen. Nach Satz 2.4 hat $Ax = 0$ dieselben Lösungen wie $A'x = 0$. In jeder Matrix in Zeilenstufenform ist die Anzahl der Leiteinträge höchstens gleich der Zeilenzahl. Also ist die Anzahl der Leiteinträge in A' höchstens gleich m, nach Voraussetzung also $< n$. Es existieren also Spalten, die keinen Leiteintrag enthalten. Daher besitzt $A'x = 0$ nach Satz 2.7 eine nichttriviale Lösung. □

Elementare Zeilenumformungen lassen sich auch anwenden, um Matrizen auf Invertierbarkeit zu prüfen.

Satz 2.9 *Es sei A eine quadratische Matrix. Dann sind die folgenden Bedingungen äquivalent:*

i) A lässt sich durch elementare Zeilenumformungen in die Einheitsmatrix überführen.

ii) A lässt sich als Produkt von Elementarmatrizen schreiben.

iii) A ist invertierbar.

iv) Das homogene lineare Gleichungssystem $Ax = 0$ besitzt nur die triviale Lösung $x = 0$.

Beweis : Wir zeigen die Kette von Implikationen i) ⇒ ii) ⇒ iii) ⇒ iv) ⇒ i).

i) ⇒ ii): Lässt sich A durch elementare Zeilenumformungen auf die Einheitsmatrix E reduzieren, so gibt es nach Lemma 2.2 Elementarmatrizen $L_1, \ldots, L_r$ mit

$$L_r \cdots L_1 A = E$$

Nach Lemma 2.3 sind Elementarmatrizen invertierbar, wir können diese Gleichung also von links mit $L_1^{-1} \cdots L_r^{-1}$ multiplizieren und erhalten $A = L_1^{-1} \cdots L_r^{-1} E = L_1^{-1} \cdots L_r^{-1}$. Da nach Lemma 2.3 die Inversen von Elementarmatrizen wieder Elementarmatrizen sind, folgt ii).

ii) $\Rightarrow$ ii): Ist $A = L_1 \cdots L_r$ das Produkt von Elementarmatrizen, so ist nach Lemma 2.3 A das Produkt invertierbarer Matrizen, also selbst invertierbar nach Lemma 1.8.

iii) $\Rightarrow$ iv): Es sei x eine beliebige Lösung des linearen Gleichungssystems $Ax = 0$. Ist A invertierbar, so folgt aus dieser Gleichung durch Linksmultiplikation mit A^{-1}, dass $x = Ex = (A^{-1}A)x = A^{-1}0 = 0$ gilt. Also besitzt $Ax = 0$ nur die triviale Lösung $x = 0$.

iv) $\Rightarrow$ i): Wir nehmen an, dass $Ax = 0$ nur trivial lösbar ist und verwandeln die Matrix A mit elementaren Zeilenumformungen in eine Matrix A' in Zeilenstufenform. Nach Satz 2.4 besitzt dann auch $A'x = 0$ nur die triviale Lösung. Nun ist A' quadratisch. Es sei n die Anzahl der Zeilen und Spalten. Angenommen, es gibt eine Spalte in A, in der kein Leiteintrag steht. Dann ist die Anzahl der Leiteinträge echt kleiner als n. Also kann in der letzten Zeile kein Leiteintrag stehen, d.h. die letzte Zeile ist eine Nullzeile. Das System $A'x = 0$ besteht höchstens aus $(n-1)$ nicht-trivialen Gleichungen. Daher besitzt es nach Korollar 2.8 eine nicht-triviale Lösung, im Widerspruch zu unserer Voraussetzung.

Unsere Annahme ist somit falsch, d.h. jede Spalte von A' enthält einen Leiteintrag. A besitzt also n Leiteinträge. Definitionsgemäß steht ein Leiteintrag auf der Diagonalen oder rechts davon, daher müssen alle Diagonaleinträge von A' Leiteinträge sein. Da oberhalb der Leiteinträge nur Nullen stehen, ist $A' = E_n$. Das beweist i). $\square$

Korollar 2.10 *Besitzt eine Matrix eine Nullzeile, so ist sie nicht invertierbar.*

Beweis : Besitzt die $n \times n$-Matrix A eine Nullzeile, so besteht das lineare Gleichungssystem $Ax = 0$ aus höchstens $n-1$ nicht-trivialen Gleichungen. Also hat es nach Korollar 2.8 eine nicht-triviale Lösung. Daher ist nach Satz 2.9 A nicht invertierbar. $\square$

Wir können nun ein Verfahren angeben, mit dem man die Inverse einer Matrix berechnen kann.

Korollar 2.11 *Es sei A eine invertierbare $n \times n$-Matrix. Es seien $U_1, \ldots, U_r$ elementare Zeilenumformungen, mit deren Hilfe man A in die Einheitsmatrix umformen kann. Wendet man $U_1, \ldots, U_r$ in derselben Reihenfolge auf die Einheitsmatrix an, so erhält man A^{-1}.*

Beweis : Die elementare Zeilenumformung U_i entspricht der Linksmultiplikation mit einer Elementarmatrix L_i. Durch Anwenden von $U_1, \ldots, U_r$ wird A in die Matrix $L_r \cdots L_1 A$ überführt. Also ist $L_r \cdots L_1 A = E_n$. Da A invertierbar ist, folgt daraus nach Multiplikation mit A^{-1}, dass $L_r \cdots L_1 E_n = L_r \cdots L_1 = E_n A^{-1} = A^{-1}$ gilt. Also entsteht A^{-1} aus E_n durch Anwenden der elementaren Zeilenumformungen $U_1, \ldots, U_r$. □

Beispiel: Wir berechnen die Inverse zu

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix}$$

mit elementaren Zeilenumformungen:

$$\begin{array}{clll}
 & \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \\
\rightsquigarrow & \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} & \text{(2. Zeile } - \text{ 2 (1. Zeile))} \\
\rightsquigarrow & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix} & \text{(1. Zeile } - \text{ 3 (2. Zeile))}
\end{array}$$

Das Inverse von A ist also $\begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$.

Wir haben hier allerdings nicht vorher geprüft, dass A invertierbar ist. Das wäre aber nach Satz 2.9 bei der Umwandlung von A durch elementare Zeilenumformungen in Zeilenstufenform herausgekommen. Für eine nicht-invertierbare Matrix entsteht so nämlich nicht die Einheitsmatrix.

Satz 2.12 *Es seien A und B $n \times n$-Matrizen mit $AB = E_n$. Dann gilt auch $BA = E_n$.*

Beweis : Wir bringen A durch elementare Zeilenumformungen in Zeilenstufenform. Das entspricht der Multiplikation mit $L_k \cdots L_1$. Eine quadratische Matrix in Zeilenstufenform ist entweder die Einheitsmatrix oder enthält eine Nullzeile. Aus $AB = E_n$ folgt $L_1 \cdots L_k AB = L_1 \cdots L_k$, d.h. das Produkt auf der linken Seite ist invertierbar. Also kann $L_1 \cdots L_k A$ keine Nullzeile haben, denn sonst hätte auch $L_1 \cdots L_k AB$ eine, was Korollar 2.10 widerspricht. Also ist $L_1 \cdots L_k A = E_n$, woraus $B = E_n B = L_1 \cdots L_k AB = L_1 \cdots L_k$ folgt. Also ist $BA = L_1 \cdots L_k A = E_n$. □

Korollar 2.13 *Existiert zu einer quadratischen Matrix A eine Matrix B mit $AB = E_n$ oder $BA = E_n$, so ist B das Inverse zu A.*

Beweis : Das folgt sofort aus Satz 2.12. □

Man kann statt elementarer Zeilenumformungen auch elementare Spaltenumformungen betrachten. Das kann man formal aus den Zeilenumformungen herleiten, indem man eine Operation auf Matrizen definiert, die Zeilen und Spalten vertauscht.

Definition 2.14 *Es sei $A = (a_{ij})$ eine $m \times n$-Matrix. Die zugehörige* ***transponierte Matrix*** *A^t ist definiert als die $n \times m$-Matrix*

$$A^t = (b_{ij})_{\substack{i=1...n \\ j=1...m}}$$

mit $b_{ij} = a_{ji}$.

A wird also an der Diagonalen ($\diagdown$) gespiegelt.

Beispiel:

i) $$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$$

ii) $$\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}^t = \begin{pmatrix} 1 & 5 \\ 3 & 7 \end{pmatrix}$$

Lemma 2.15 *Sind A und B $m \times n$-Matrizen und α ein Skalar, so gilt*

i) $(A + B)^t = A^t + B^t$

ii) $(\alpha A)^t = \alpha A^t$

iii) $(A^t)^t = A$

Ist A eine $m \times n$- und B eine $n \times p$-Matrix, so gilt ferner

iv) $(AB)^t = B^t A^t$

Beweis : Übungsaufgabe. □

Lemma 2.16 *Es sei A eine $m \times n$-Matrix und L eine $n \times n$-Elementarmatrix.*

i) Ist $L = E + ce_{ij}$ vom Typ I, so entsteht AL aus A, indem man die j-te Spalte durch (j-te Spalte) $+c$ (i-te Spalte) ersetzt.

ii) Ist $L = E + e_{ij} + e_{ji} - e_{ii} - e_{jj}$ vom Typ II, so entsteht AL aus A, indem die i-te und die j-te Spalte vertauscht werden.

iii) Ist $L = E + (c-1)e_{ii}$ vom Typ III, so entsteht AL aus A, indem die i-te Spalte mit c multipliziert wird.

Beweis : Wir wenden Lemma 2.2 auf die Matrix $(AL)^t = L^t A^t$ an. □

3 Determinanten

Jeder quadratischen Matrix kann man eine Zahl zuordnen, die sogenannte Determinante. Diese wollen wir hier definieren und untersuchen.

Die Determinante einer 1×1-Matrix $A = (a)$ ist einfach ihr einziger Eintrag:

$$\det A = a$$

Die Determinante einer 2×2-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist definiert durch

$$\det A = ad - bc.$$

Um diese Zahl geometrisch zu deuten, betrachten wir den Raum $\mathbb{R}^2$ aller zweidimensionalen Spaltenvektoren über $\mathbb{R}$, d.h. es ist

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

Die 2×2-Matrix A definiert eine (lineare) Abbildung.

$$f : \mathbb{R}^2 \to \mathbb{R}^2,$$

gegeben durch $f \begin{pmatrix} r \\ s \end{pmatrix} = \begin{pmatrix} ar + bs \\ cr + ds \end{pmatrix}$.

Wir betrachten das Bild des Einheitsquadrats

$$\square = \left\{ \begin{pmatrix} r \\ s \end{pmatrix} : 0 \leq r \leq 1, 0 \leq s \leq 1 \right\}$$

unter f. Es ist

$$f(\square) = \left\{ \begin{pmatrix} ar + bs \\ cr + ds \end{pmatrix} : 0 \leq r \leq 1, 0 \leq s \leq 1 \right\},$$

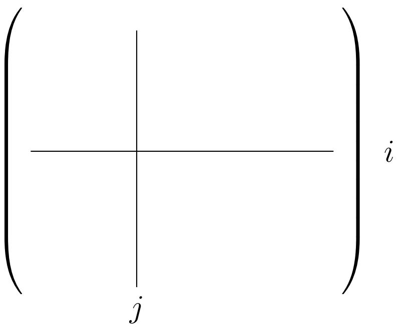
also das von $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ aufgespannte Parallelogramm im $\mathbb{R}^2$.

Der Flächeninhalt dieses Parallelogramms ist gerade $|\det A| = |ad - bc|$.

Insbesondere ist die Determinante $\det A$ genau dann gleich Null, wenn das Parallelprogramm zu einer Strecke entartet ist, d.h. wenn eine der beiden Spalten von A ein Vielfaches der anderen ist.

Auf quadratischen Matrizen beliebiger Größe sind die Formeln für die Determinante komplizierter. Wir werden die Determinante durch **eine** solche Formel definieren, und dann einen Katalog von Eigenschaften herleiten, durch den sie eindeutig bestimmt ist.

Wir bezeichnen den Raum aller $n \times n$-Matrizen über $\mathbb{R}$ mit $\mathbb{R}^{n \times n}$. Für jede Matrix A in $\mathbb{R}^{n \times n}$ und für alle $i, j = 1, \ldots, n$ bezeichnen wir mit A_{ij} diejenige $(n-1) \times (n-1)$-Matrix, die aus A durch Streichen der i-ten Zeile und der j-ten Spalte hervorgeht:

$A_{ij} =$ 

Beispiel: Für $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 2 & 3 & -1 \end{pmatrix}$ ist

$$A_{11} = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} \text{ und } A_{31} = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}$$

Lemma 3.1 *Es gibt genau ein System von Funktionen* $(d_n)_{n\in\mathbb{N}}$, $d_n : \mathbb{R}^{n\times n} \to \mathbb{R}$, *so dass gilt:*

$$i)\ d_1((a)) = a$$

und

$$\begin{aligned} ii)\ d_n(A) &= a_{11}d_{n-1}(A_{11}) - a_{21}d_{n-1}(A_{21}) + a_{31}d_{n-1}(A_{31}) \cdots \pm a_{n1}d_{n-1}(A_{n1}) \\ &= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} d_{n-1}(A_{i1}) \end{aligned}$$

Diese Formel nennt man „Entwicklung nach der 1. Spalte".

Beweis : Die Existenz einer solchen Funktion zeigen wir rekursiv. Für festes n benutzen wir die Formel ii), um $d_n(A)$ mit Hilfe der Funktion d_{n-1} auszudrücken. Dann ersetzen wir alle Terme der Form $d_{n-1}(B)$ mit Hilfe von ii) durch einen Ausdruck, in dem nur noch die Funktion d_{n-2} vorkommt. Nach endlich vielen Schritten haben wir so $d_n(A)$ durch eine Formel ausgedrückt, in der nur noch d_1-Werte auftauchen. Diese können wir mit i) ausrechnen.

Die Eindeutigkeit eines Systems von Funktionen mit i) und ii) zeigt man leicht mit Vollständiger Induktion. (Führen Sie das aus!) □

Definition 3.2 *Für eine* $n \times n$*-Matrix* A *definieren wir die* ***Determinante*** *von* A *als*

$$\det A = d_n(A),$$

wobei d_n *die Funktion aus Lemma 3.1 ist.*

Beispiel:

i) Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}$, so ist

$$\begin{aligned} \det(A) = d_2(A) &\overset{3.1ii)}{=} ad_1((d)) - cd_1((b)) \\ &\overset{3.1i)}{=} ad - bc. \end{aligned}$$

Das stimmt mit der Definition zu Beginn des Paragraphen überein.

ii) Ist $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 2 & 3 & -1 \end{pmatrix}$ wie im obigen Beispiel, so ist

$$\begin{aligned} \det A = d_3(A) &= d_2\left(\begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}\right) + 2d_2\left(\begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}\right) \\ &= -2 - 6 + 4 = -4. \end{aligned}$$

Wir wollen jetzt einige Eigenschaften von Determinanten herleiten.

Proposition 3.3 *Es ist* $\det(E_n) = 1$

Beweis : (mit Induktion nach n):

Induktionsanfang ($n = 1$)**:** Offenbar ist $\det E_1 = 1$.

Induktionsschluss: Wir nehmen an, $\det E_{n-1} = 1$ ist schon gezeigt. Da E_n in der ersten Spalte nur an der ersten Stelle einen Eintrag $\neq 0$ hat, folgt

$$\det E_n = \det(E_n)_{11} = \det(E_{n-1}) = 1.$$

□

Definition 3.4 *Wir nennen eine Funktion*

$$d : \mathbb{R}^{n \times n} \to \mathbb{R}$$

***linear in den Zeilen**, falls für alle* $i = 1, \dots, n$ *gilt:*

i) Ist $A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ *die* $n \times n$*-Matrix mit den Zeilen* $z_1, \dots, z_n$, $B = \begin{pmatrix} z_1 \\ \vdots \\ \tilde{z}_i \\ \vdots \\ z_n \end{pmatrix}$ *die* $n \times n$*-Matrix, deren* i*-te Zeile* $\tilde{z}_i$ *ist und die ansonsten dieselben Zeilen wie* A *besitzt, und* $C = \begin{pmatrix} z_1 \\ \vdots \\ z_i + \tilde{z}_i \\ \vdots \\ z_n \end{pmatrix}$ *die* $n \times n$*-Matrix, deren* i*-te Zeile* $z_i + \tilde{z}_i$ *ist und die ansonsten dieselben Zeilen wie* A *besitzt, so gilt*

$$d(C) = d(A) + d(B).$$

ii) Ist $A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ *die* $n \times n$*-te Matrix mit den Zeilen* $z_1, \ldots, z_n$ *und* $B = \begin{pmatrix} z_1 \\ \vdots \\ cz_i \\ \vdots \\ z_n \end{pmatrix}$ *die Matrix, deren* i*-te Zeile* cz_i *für ein* $c \in \mathbb{R}$ *ist, und die ansonsten dieselben Zeilen wie* A *besitzt, so gilt*

$$d(B) = cd(A).$$

Eine Funktion d ist also linear in den Zeilen, wenn d mit der Addition und skalaren Multiplikation in der i-ten Zeile vertauscht, wobei wir alle anderen Zeilen festlassen.

Satz 3.5 *Die Determinantenfunktion* $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ *ist linear in den Zeilen.*

Beweis : mit Induktion nach n:

Induktionsanfang $(n = 1)$**:** das folgt sofort aus der Definition der Determinante für 1×1-Matrizen.

Induktionsschluss: Angenommen, die Determinante für $(n-1) \times (n-1)$-Matrizen ist linear in den Zeilen. Wir fixieren ein $k \in \{1, \ldots, n\}$. Es seien $A = (a_{ij})_{i,j}$ und $B = (b_{ij})_{i,j}$ zwei $n \times n$-Matrizen, die sich höchstens in der k-ten Zeile unterscheiden, d.h. es ist $a_{ij} = b_{ij}$ für $i \neq k$ und beliebige j.

Dann sei $C = (c_{ij})_{i,j}$ die Matrix, deren k-te Zeile die Summe der k-ten Zeile von A und der k-ten Zeile von B ist, und die ansonsten mit A übereinstimmt. Also ist $c_{ij} = a_{ij} = b_{ij}$ für $i \neq k$ und beliebige j sowie $c_{kj} = a_{kj} + b_{kj}$ für alle j. Dann ist $C_{k1} = A_{k1} = B_{k1}$.

Für $i \neq k$ unterscheiden sich A_{i1} und B_{i1} höchstens in der k-ten Zeile, und C_{i1} ist die Matrix, die außerhalb der k-ten Zeile mit A_{i1} und B_{i1} übereinstimmt und deren k-te Zeile die Summe der k-ten Zeile von A_{i1} und der k-ten Zeile von B_{i1} ist. Nach Induktionsvoraussetzung gilt also $\det(C_{i1}) = \det(A_{i1}) + \det(B_{i1})$.

Wir berechnen die Determinante von C mit der Formel aus Lemma 3.1 als

$$\begin{aligned}
\det(C) &= \sum_{i=1}^{n}(-1)^{i+1}c_{i1}\det(C_{i1}) \\
&= (-1)^{k+1}c_{k1}\det(C_{k1}) + \sum_{i\neq k}(-1)^{i+1}c_{i1}\det(C_{i1}) \\
&= (-1)^{k+1}(a_{k1}+b_{k1})\det C_{k1} + \sum_{i\neq k}(-1)^{i+1}c_{i1}(\det A_{i1}+\det B_{i1}) \\
&= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}\det A_{i1} + \sum_{i=1}^{n}(-1)^{i+1}b_{i1}\det B_{i1} \\
&= \det A + \det B.
\end{aligned}$$

Ferner sei $c \in \mathbb{R}$ und $A = (a_{ij})_{ij}$ eine $n \times n$-Matrix sowie $B = (b_{ij})_{i,j}$ die $n \times n$-Matrix mit $b_{ij} = a_{ij}$ für $i \neq k$ und beliebige j und $b_{kj} = ca_{kj}$ für alle j. Die k-te Zeile von B ist also das c-fache Vielfache der k-ten Zeile von A, die anderen Zeilen von B stimmen mit den entsprechenden Zeilen von A überein. Offenbar ist $A_{k1} = B_{k1}$. Nach Induktionsvoraussetzung gilt ferner für alle $i \neq k$

$$\det(B_{i1}) = c\det(A_{i1}).$$

Also folgt

$$\begin{aligned}
\det B &= \sum_{i=1}^{n}(-1)^{i+1}b_{i1}\det B_{i1} \\
&= (-1)^{k+1}ca_{k1}\det A_{k1} + \sum_{i\neq k}(-1)^{i+1}a_{i1}c\det A_{i1} \\
&= c\det A.
\end{aligned}$$

Damit ist die Behauptung bewiesen. □

Beispiel:

i) Es ist

$$\det\begin{pmatrix}1 & 3 & 7\\ 4 & 8 & 2\\ 5 & 1 & 1\end{pmatrix} = 2\cdot\det\begin{pmatrix}1 & 3 & 7\\ 2 & 4 & 1\\ 5 & 1 & 1\end{pmatrix}$$

ii) Es ist

$$\det\begin{pmatrix}1 & 3 & 7\\ 4 & 8 & 2\\ 5 & 1 & 1\end{pmatrix} = \det\begin{pmatrix}1 & 3 & 7\\ 3 & 5 & -5\\ 5 & 1 & 1\end{pmatrix} + \det\begin{pmatrix}1 & 3 & 7\\ 1 & 3 & 7\\ 5 & 1 & 1\end{pmatrix}$$

Korollar 3.6 *Enthält A eine Nullzeile, so gilt* $\det A = 0$.

Beweis : Enthält A eine Nullzeile, so können wir diese Zeile mit 0 multiplizieren, ohne A zu ändern. Also gilt nach Satz 3.5

$$\det A = 0 \cdot \det A = 0.$$

□

Lemma 3.7 *Stimmen zwei benachbarte Zeilen einer Matrix überein, so ist* $\det A = 0$.

Beweis : mit Induktion nach n:

Induktionsanfang $(n = 1)$**:**

Ist $n = 1$, so stimmt die Behauptung, da es solche 1×1-Matrizen nicht gibt.

Induktionsanschluss: Die Behauptung gelte für alle $(n-1) \times (n-1)$-Matrizen. Es sei A eine $n \times n$-Matrix, deren k-te und $(k+1)$-te Zeile übereinstimmen.

Dann ist für alle $i \neq k, k+1$ die k-te Zeile von A_{i1} gleich der $(k+1)$-ten Zeile von A_{i1}, also gilt nach Induktionsvoraussetzung $\det A_{i1} = 0$.

Außerdem ist $A_{k1} = A_{k+1,1}$ und $a_{k1} = a_{k+1,1}$, woraus

$$\begin{aligned} \det A &= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1} \\ &= (-1)^{k+1} a_{k1} \det A_{k1} + (-1)^{k+2} a_{k+1,1} \det A_{k+1,1} \\ &= 0 \end{aligned}$$

folgt. □

Lemma 3.8 *Addiert man ein Vielfaches einer Zeile zu einer benachbarten Zeile hinzu, so ändert sich die Determinante nicht.*

Beweis : Es sei $A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ die Matrix mit den Zeilen $z_1, \ldots, z_n$, und $B = \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ z_{i+1} + cz_i \\ \vdots \\ z_n \end{pmatrix}$ für ein $i \in \{1, \ldots, n-1\}$ und ein $c \in \mathbb{R}$. Dann gilt nach Satz 3.5

$$\det B = \det \begin{pmatrix} | \\ z_i \\ z_{i+1} + cz_i \\ | \end{pmatrix} = \det \begin{pmatrix} | \\ z_i \\ z_{i+1} \\ | \end{pmatrix} + c \det \begin{pmatrix} | \\ z_i \\ z_i \\ | \end{pmatrix} = \det A,$$

da der zweite Term nach Lemma 3.7 verschwindet.

Genauso argumentiert man für die Matrix $B = \begin{pmatrix} z_1 \\ \vdots \\ z_i + cz_{i+1} \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix}$. □

Lemma 3.9 *Entsteht B aus A durch Vertauschen zweier benachbarter Zeilen, so ist* $\det B = -\det A$.

Beweis : Es sei $A = \begin{pmatrix} | \\ z_i \\ z_{i+1} \\ | \end{pmatrix}$ und $B = \begin{pmatrix} | \\ z_{i+1} \\ z_i \\ | \end{pmatrix}$. Dann gilt

$$\begin{aligned}
\det B &= \det \begin{pmatrix} | \\ z_{i+1} \\ z_i \\ | \end{pmatrix} \overset{3.8}{=} \det \begin{pmatrix} | \\ z_{i+1} \\ z_i - z_{i+1} \\ | \end{pmatrix} \overset{3.8}{=} \det \begin{pmatrix} | \\ z_{i+1} + (z_i - z_{i+1}) \\ z_i - z_{i+1} \\ | \end{pmatrix} \\
&= \det \begin{pmatrix} | \\ z_i \\ z_i - z_{i+1} \\ | \end{pmatrix} \overset{3.8}{=} \det \begin{pmatrix} | \\ z_i \\ z_i - z_{i+1} - z_i \\ | \end{pmatrix} = \det \begin{pmatrix} | \\ z_i \\ -z_{i+1} \\ | \end{pmatrix} \\
&\overset{3.5}{=} -\det \begin{pmatrix} | \\ z_i \\ z_{i+1} \\ | \end{pmatrix} = -\det A.
\end{aligned}$$

□

Damit können wir Lemma 3.7 verallgemeinern.

Satz 3.10 *Stimmen zwei beliebige Zeilen einer Matrix A überein, so ist* $\det A = 0$.

Beweis : Stimmen in A die i-te und die j-te Zeile mit $i < j$ überein, so vertauschen wir die j-te Zeile solange mit der nächsthöheren Zeile, bis sie in der $(i+1)$-ten Zeile landet. Bei jeder Vertauschung wechselt nach Lemma 3.9 das Vorzeichen der Determinante, also erhalten wir so eine Matrix B mit $\det B \pm \det A$. Nach Lemma 3.7 folgt $\det B = 0$, also auch $\det A = 0$. □

Jetzt können wir auch Lemma 3.8 und Lemma 3.9 verallgemeinern.

Satz 3.11 *Entsteht B aus $A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, indem für $i \neq j$ und ein $c \in \mathbb{R}$ die i-te Zeile z_i durch $z_i + cz_j$ ersetzt wird (also durch eine elementare Zeilenumformung vom Typ I), so ist* $\det B = \det A$.

Beweis : Mit Hilfe von Lemma 3.7 können wir das jetzt genauso beweisen wie Lemma 3.8:

$$\det B = \det \begin{pmatrix} | \\ z_i + cz_j \\ | \\ z_j \\ | \end{pmatrix} \overset{3.5}{=} \det \begin{pmatrix} | \\ z_i \\ | \\ z_j \\ | \end{pmatrix} + c \det \begin{pmatrix} | \\ z_j \\ | \\ z_j \\ | \end{pmatrix} \overset{3.7}{=} \det A.$$

□

Satz 3.12 *Entsteht B aus A durch Vertauschen zweier beliebiger Zeilen, so gilt* $\det B = -\det A$.

Beweis : Es sei

$$A = \begin{pmatrix} | \\ z_i \\ | \\ z_j \\ | \end{pmatrix} \text{ und } B = \begin{pmatrix} | \\ z_j \\ | \\ z_i \\ | \end{pmatrix}$$

Dann können wir genauso argumentieren wie im Beweis von Lemma 3.9:

$$\begin{aligned} \det \begin{pmatrix} | \\ z_i \\ | \\ z_j \\ | \end{pmatrix} \quad &\overset{3.11}{=} \quad \det \begin{pmatrix} | \\ z_i + (z_j - z_i) \\ | \\ z_j - z_i \\ | \end{pmatrix} \overset{3.11}{=} \det \begin{pmatrix} | \\ z_i + (z_j - z_i) \\ | \\ z_j - z_i \\ | \end{pmatrix} \\ &\overset{3.11}{=} \quad \det \begin{pmatrix} | \\ z_j \\ | \\ z_j - z_i - z_j \\ | \end{pmatrix} \overset{3.5}{=} -\det \begin{pmatrix} | \\ z_j \\ | \\ z_i \\ | \end{pmatrix}. \end{aligned}$$

□

Wir haben in Satz 3.5, Satz 3.11 und Satz 3.12 gesehen, wie sich die Determinante unter elementaren Zeilenumformungen verhält. Damit können wir jetzt die Determinante von Elementarmatrizen ausrechnen.

Proposition 3.13 *Es sei L eine $n \times n$-Elementarmatrix.*

i) Ist $L = E + ce_{ij}$ *vom Typ I, so ist* $\det L = 1$.

ii) Ist $L = E_n - e_{ii} - e_{jj} + e_{ij} + e_{ji}$ *vom Typ II, so ist* $\det L = -1$.

iii) Ist $L = E_n + (c-1)e_{ii}$ *vom Typ III, so ist* $\det L = c$.

Beweis : Nach Lemma 2.2 entsteht $L = LE_n$ aus der Einheitsmatrix durch Anwenden einer elementaren Zeilenumformung vom Typ I, II oder III. Daher ergibt sich

i) für Typ I: $\det L = 1$ nach Satz 3.11

ii) für Typ II: $\det L = -1$ nach Satz 3.12

iii) für Typ III: $\det L = c$ nach Satz 3.5

□

Lemma 3.14 *Ist L eine Elementarmatarix und A eine beliebige Matrix, so gilt*

$$\det(LA) = \det L \det A.$$

Beweis : LA entsteht aus A durch Anwenden einer elementaren Zeilenoperation. Also gilt: Ist L vom Typ I, so ist $\det LA = \det A$ nach Satz 3.11. Nach Proposition 3.13 ist $\det L = 1$, also $\det LA = \det L \det A$. Ist L vom Typ II, so ist $\det LA = -\det A$ nach Satz 3.12. Nach Proposition 3.13 ist $\det L = -1$, also $\det LA = \det L \det A$. Ist L vom Typ III, so ist $\det LA = c \det A$ nach Satz 3.5. Nach Proposition 3.13 ist $\det L = c$, also $\det LA = \det L \det A$. □

Nun lässt sich jede quadratische Matrix A mit elementaren Zeilenumformungen entweder in die Einheitsmatrix E_n oder in eine Matrix A' überführen, deren letzte Zeile eine Nullzeile ist.

Im ersten Fall ist

$$L_r \cdots L_1 A = E_n,$$

im zweiten Fall ist

$$L_r \cdots L_1 A = A'$$

für Elementarmatrizen $L_1, \ldots, L_r$.

Also folgt im ersten Fall mit Lemma 3.14

$$1 = \det(E_n) = \det(L_r \cdots L_1 A) = \det(L_r) \cdots \det(L_1) \det A.$$

Da alle $\det L_i \neq 0$ sind, folgt $\det A \neq 0$.

Im zweiten Fall gilt

$$0 = \det A' = \det(L_r \cdots L_1 A) = \det(L_r) \cdots \det(L_1) \det A.$$

Da alle $\det L_i \neq 0$ sind, folgt hieraus $\det A = 0$.

Korollar 3.15 *Eine quadratische Matrix A ist genau dann invertierbar, wenn $\det A \neq 0$ ist.*

Beweis : Ist A invertierbar, so lässt sich A nach Satz 2.9 durch elementare Zeilenumformungen in die Einheitsmatrix überführen. Wie wir oben nachgerechnet haben, folgt daraus $\det(A) \neq 0$. Ist umgekehrt $\det A \neq 0$, so überführen wir A durch elementare Zeilenumformungen in eine Matrix in Zeilenstufenform. Wie wir oben ausgerechnet haben, kann die so entstehende Matrix keine Nullzeile haben, sie ist also die Einheitsmatrix. Daher ist A nach Satz 2.9 invertierbar. □

Satz 3.16 (Axiomatische Charakterisierung der Determinante) *Es sei*

$$d : \mathbb{R}^{n \times n} \to \mathbb{R}$$

eine Funktion mit folgenden Eigenschaften:

i) $d(E_n) = 1$

ii) d *ist linear in den Zeilen*

iii) Stimmen zwei benachbarte Zeilen einer Matrix A überein, so ist $d(A) = 0$.

Dann stimmt d mit der Determinantenfunktion überein, d.h. es gilt $d = \det$. Die Determinante ist also durch die Regeln i) - iii) eindeutig bestimmt.

Beweis : Analysiert man die Beweise für Korollar 3.6, Lemma 3.8, Lemma 3.9, Satz 3.10, Satz 3.11, Satz 3.12, Proposition 3.13 und Lemma 3.14, so sieht man, dass all diese Resultate alleine aus den Eigenschaften i) - iii) für die Determinantenfunktion folgen. Sie gelten also auch für jede andere Funktion d, die i) - iii) erfüllt. Es sei A eine beliebige $n \times n$-Matrix, die wir mit elementaren Zeilenumformungen in Zeilenstufenform überführen. Es ist also

$$L_r \cdots L_1 A = A'$$

für Elementarmatrizen $L_1, \ldots, L_r$ und eine Matrix A', die entweder die Einheitsmatrix ist oder eine Nullzeile hat. Es gilt $\det L_i = d(L_i)$ nach Proposition 3.13, $\det A' = d(A')$ nach i) bzw. Korollar 3.6. Also folgt aus Lemma 3.14 $\det A = d(A)$. □

Satz 3.17 *Für $n \times n$-Matrizen A und B gilt* $\det(AB) = \det A \ \det B$.

Beweis : Ist A eine Elementarmatrix, so folgt dies aus Lemma 3.14. Für ein beliebiges A unterscheiden wir zwei Fälle.

1. Fall: A ist invertierbar. Dann ist A nach Satz 2.9 Produkt von Elementarmatrizen, d.h. $A = L_r \cdots L_1$. Also folgt mit Induktion nach r aus Lemma 3.14

$$\begin{aligned} \det(AB) &= \det(L_r) \cdots \det(L_1) \det B \\ &= \det(A) \det(B). \end{aligned}$$

2. Fall: A ist nicht invertierbar. Dann ist nach Korollar 3.15 $\det A = 0$. Wir müssen also zeigen, dass auch $\det(AB) = 0$ gilt. Durch elementare Zeilenumformungen kann A in eine Matrix A' überführt werden, deren letzte Zeile eine Nullzeile ist. Es ist $A' = L_r \cdots L_1 A$ für gewisse Elementarmatrizen L_i. Nun ist auch die letzte Zeile von $A'B$ eine Nullzeile, nach Korollar 3.6 ist also $\det(A'B) = 0$. Aus Lemma 3.14 folgt dann

$$\begin{aligned} \det(AB) &= \det(L_1^{-1}) \cdots \det(L_r^{-1}) \det(A'B) \\ &= 0. \end{aligned}$$

□

Korollar 3.18 *Ist A invertierbar, so gilt*

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

Beweis : Nach Satz 3.17 gilt

$$\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(E_n) = 1.$$

□

Satz 3.19 *Es sei A eine quadratische Matrix und A^t ihre Transponierte. Dann ist*

$$\det A = \det A^t.$$

Beweis : Wir unterscheiden zwei Fälle:

i) A ist invertierbar. Dann gibt es nach Satz 2.9 Elementarmatrizen $L_1, \ldots, L_r$ mit $A = L_r \cdots L_1$. Also ist nach Lemma 2.15 $A^t = L_1^t \cdots L_r^t$. Da nach Satz 3.17 $\det A = \prod_{i=1}^{r} \det L_i$ und $\det A^t = \prod_{i=1}^{r} \det L_i^t$ ist, genügt es zu zeigen, dass für jede Elementarmatrix L gilt $\det L = \det L^t$.

   Ist $L = E_n + ce_{ij}$ vom Typ I, so ist $L^t = E_n + ce_{ji}$ und $\det L = \det L^t = 1$ nach Proposition 3.13. Ist $L = E_n - e_{ii} - e_{jj} + e_{ij} + e_{ji}$ vom Typ II, so ist $L^t = L$, also auch $\det L = \det L^t$. Ist $L = E_n + (c-1)e_{ii}$ vom Typ III, so ist ebenfalls $L^t = L$, also auch hier $\det L = \det L^t$.

ii) A ist nicht invertierbar. Dann lässt sich A mit elementaren Zeilenumformungen in eine Matrix in Zeilenstufenform überführen, deren letzte Zeile eine Nullzeile ist. Nach einer Zeilenvertauschung erhalten wir eine Matrix B, deren erste Zeile eine Nullzeile ist. Es gibt also Elementarmatrizen $L_1, \ldots, L_r$ mit $B = L_r \cdots L_1 A$. Nach Lemma 2.15 ist $B^t = A^t L_1^t \cdots L_r^t$. Also ist $\det B = \prod_{i=1}^{r} \det L_i \cdot \det A$ und $\det B^t = \prod_{i=1}^{r} \det L_i^t \cdot \det A$. Da nach Korollar 3.15 alle $\det L_i \neq 0$ sind und nach i) $\det L_i = \det L_i^t$ gilt, müssen wir nur $\det B = \det B^t$ zeigen. Nach Korollar 3.6 ist $\det B = 0$. Berechnen wir $\det B^t$ mit der Formel aus Lemma 3.1, so steht in jedem Summanden ein Element aus der ersten Spalte von B^t, d.h. aus der ersten Zeile von B. Also ist auch $\det B^t = 0$.

□

Korollar 3.20 *Die Determinantenfunktion*

$$\det : \mathbb{R}^{n \times n} \to \mathbb{R}$$

hat folgende Eigenschaften:

i) $\det$ *ist linear in den Spalten der Matrix.*

ii) Stimmen zwei Spalten in A überein, so ist $\det A = 0$.

iii) Enthält A eine Nullspalte, so ist $\det A = 0$.

iv) Addiert man ein Vielfaches einer Spalte zu einer anderen hinzu, so bleibt die Determinante unverändert.

v) Vertauscht man zwei Spalten, so ändert die Determinante ihr Vorzeichen.

Beweis : Die Linearität in den Spalten ist hier genauso definiert wie in Definition 3.4, wenn man überall „Zeile“ durch „Spalte“ ersetzt.

Die Aussagen i) - v) folgen aus Satz 3.5, Satz 3.10, Korollar 3.6, Satz 3.11 und Satz 3.12, wenn man jeweils Satz 3.19 anwendet. □

Wir haben die Determinante in Lemma 3.1 durch die „Entwicklung nach der ersten Spalte“ kennengelernt, d.h. durch die Formel

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det(A_{i1})$$

Diese Formel lässt sich folgendermaßen verallgemeinern:

Satz 3.21 (Entwicklungsformeln für die Determinante)

i) *Für alle $j = 1, \ldots, n$ gilt*

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

Das nennt man „Entwicklung nach der j-ten Spalte“.

ii) *Für alle $i = 1, \ldots, n$ gilt*

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

Das nennt man „Entwicklung nach der i-ten Zeile“.

Beweis :

i) Wir ziehen mit $(j-1)$-Spaltenvertauschungen die j-te Spalte nacheinander an der $(j-1)$-ten, der $(j-2)$-ten, $\ldots$ Spalte vorbei, bis sie in der ersten Spalte landet. Die entstehende Matrix nennen wir A'. Die Reihenfolge der Spalten $s_1, \ldots, s_n$ von A ist also in A' $s_j, s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_n$. Nach Korollar 3.20 gilt $\det A = (-1)^{j-1} \det A'$. Ist $A = (a_{ij})$ und $A' = (a'_{ij})$, so gilt $a'_{i1} = a_{ij}$ für $i = 1, \ldots, n$. Ferner ist $A'_{i1} = A_{ij}$ für alle i. Wir berechnen $\det A'$ mit der Formel aus Lemma 3.1:

$$\begin{aligned} \det A' &= \sum_{i=1}^{n} (-1)^{i+1} a'_{i1} \det A'_{i1} \\ &= \sum_{i=1}^{n} (-1)^{i+1} a_{ij} \det A_{ij}, \end{aligned}$$

woraus

$$\begin{aligned}\det A &= (-1)^{j-1}\det A' \\ &= \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij},\end{aligned}$$

folgt.

ii) Nach Korollar 3.20 ist $\det A = \det A^t$. Da $(A^t)_{ij} = (A_{ji})^t$ gilt, folgt $\det A^t_{ij} = \det A_{ji}$. Also erhalten wir mit i) für alle $j = 1, \ldots, n$:

$$\begin{aligned}\det A &= \det A^t \\ &= \sum_{i=1}^{n}(-1)^{i+j}a_{ji}\det(A^t)_{ij} \\ &= \sum_{i=1}^{n}(-1)^{i+j}a_{ji}\det A_{ji},\end{aligned}$$

woraus nach Vertauschen der Indizes i und j die Behauptung ii) folgt.

$\square$

Zum Abschluss dieses Kapitels über Determinanten wollen wir noch die sogenannte Leibnizformel kennenlernen, die die Determinante einer Matrix vollständig durch die Matrixkoeffizienten ausdrückt.

Definition 3.22 *Eine $n \times n$-Matrix P heißt* ***Permutationsmatrix****, falls in jeder Zeile und in jeder Spalte von P genau eine 1 und ansonsten nur Nullen stehen.*

Beispiel:

i) E_n ist eine Permutationsmatrix.

ii) $\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$, $\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ sind alle 2×2-Permuationsmatrizen.

iii) $\begin{pmatrix}1 & 0 & 0\\0 & 0 & 1\\1 & 0 & 0\end{pmatrix}$ ist keine Permutationsmatrix.

Wir bezeichnen mit $e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ den n-dimensionalen Spaltenvektor, der in der j-ten Zeile den Eintrag 1 hat und ansonsten nur Nullen enthält. Wir nennen die Vektoren $e_1, \dots, e_n$ auch die **kanonischen Einheitsvektoren**.

Lemma 3.23 *Für jede $n \times n$-Matrix A ist Ae_j gerade die j-te Spalte von A.*

Beweis : Übungsaufgabe. □

Nach Definition ist jede Spalte einer Permutationsmatrix P einer der Vektoren $e_1, \dots, e_n$. Da P in jeder Zeile nur einen Eintrag $\neq 0$ enthält, kommt jeder Vektor $e_1, \dots, e_n$ höchstens einmal als Spalte in P vor. Wir definieren eine Funktion

$$p : \{1, \dots, n\} \to \{1, \dots, n\}$$

durch $p(j) = k$ genau dann, wenn die j-te Spalte von P gerade e_k ist.

Da P n verschiedene Spalten hat, muss jedes e_k als Spalte auftreten, d.h. p ist eine bijektive Abbildung.

Definition 3.24 *Sei $n \in \mathbb{N}$. Eine bijektive Abbildung $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ heißt* ***Permutation*** *von $\{1, \dots, n\}$. Die Menge aller Permutationen von $\{1, \dots, n\}$ nennt man auch „ symmetrische Gruppe“ und bezeichnet sie mit $\mathcal{S}_n$. Eine Permutation τ, die zwei Zahlen $i \neq j$ vertauscht und alle anderen festlässt, heißt* ***Transposition****.*

Man schreibt eine Permutation manchmal in der Form $\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$ In der unteren Zeile steht jede der Zahlen $1, \dots, n$ genau einmal.

Beispiel:

i) $\begin{pmatrix} 1 & 2 & \dots & n \\ n & n-1 & \dots & 1 \end{pmatrix}$ ist eine Permutation.

ii) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$ ist eine Transposition.

Lemma 3.25 *Es sei $|S_n|$ die Anzahl der Elemente in S_n. Dann gilt $|S_n| = n!$*

Beweis : mit Induktion nach n. Offenbar ist $\mathcal{S}_1 = \{\text{id}\}$, also $|\mathcal{S}_1| = 1! = 1$. Angenommen, $|\mathcal{S}_{n-1}| = (n-1)!$ für ein $n \geq 2$. Für jedes $i \in \{1, \ldots, n\}$ bildet die Menge aller Permutationen σ von $\{1, \ldots, n\}$ mit $\sigma(n) = i$ die Menge $\{1, \ldots, n-1\}$ bijektiv auf $\{1, \ldots, n\} \setminus \{i\}$ ab. Also gibt es nach Induktionsvoraussetzung $|\mathcal{S}_{n-1}| = (n-1)!$ viele Permutationen σ mit $\sigma(n) = i$. Da es n Wahlen für die Zahl i gibt, folgt $|\mathcal{S}_n| = n \cdot (n-1)! = n!$. $\square$

Sind $\sigma, \tau : \{1, \ldots, n\} \to \{1, \ldots, n\}$ Permutationen, so ist

$$\begin{aligned} \sigma\tau = \sigma \circ \tau : \{1, \ldots, n\} &\to \{1, \ldots, n\} \\ k &\mapsto \sigma(\tau(k)) \end{aligned}$$

ebenfalls eine Permutation. Wir nennen $\sigma\tau$ auch das Produkt von σ und τ.

Lemma 3.26 *Jedes $\sigma \in \mathcal{S}_n$ lässt sich als Produkt von Transpositionen schreiben.*

Beweis : Für jede Permutation σ sei $r(\sigma) \geq 0$ diejenige Zahl, für die $\sigma(i) = i$ für $i = 1, \ldots, r(\sigma)$ und $\sigma(r(\sigma) + 1) \neq r(\sigma) + 1$ ist. Schreiben wir

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

so ist $r(\sigma)$ die Stelle, bis zu der unten dasselbe wie oben steht. Ist $\sigma(1) \neq 1$, so ist $r(\sigma) = 0$. Wir führen für festes n eine absteigende Induktion nach $r(\sigma)$.

Ist $r(\sigma) = n$, so ist $\sigma = \text{id}$ und somit das leere Produkt von Transpositionen. Das ist der Induktionsanfang. Wir nehmen nun für den Induktionsschluss an, die Behauptung gelte für alle σ mit $r(\sigma) > r$ und zeigen sie für alle σ mit $r(\sigma) = r < n$. Aus $r(\sigma) = r$ folgt $\sigma(r+1) \neq r+1$. Da σ auf $\{1, \ldots, r\}$ die Identität ist, folgt $\sigma(r+1) > r+1$. Ist τ_1 die Vertauschung von $r+1$ und $\sigma(r+1)$, so lässt $\tau_1\sigma$ die Zahlen $1, 2, \ldots, r+1$ unverändert, d.h. es ist $r(\tau_1\sigma) \geq r+1$. Nach Induktionsvoraussetzung ist also $\tau_1\sigma = \tau_2 \cdots \tau_s$ das Produkt von Transpositionen $\tau_2, \ldots, \tau_s$. Da $\tau_1^2 = \text{id}$ ist, folgt hieraus $\sigma = \tau_1\tau_2 \cdots \tau_s$ durch Multiplikation mit τ_1 auf beiden Seiten, d.h. σ ist ebenfalls Produkt von Transpositionen. $\square$

Wir wir oben gesehen haben, können wir jeder Permutationsmatrix P die Permutation $\sigma : \{1 \ldots, n\} \to \{1, \ldots, n\}$ mit $Pe_j = e_{\sigma(j)}$ zuordnen. Von links nach rechts gelesen, besteht P also aus den Spalten $e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}$.

Mit anderen Worten, es gilt

$$P = e_{\sigma(1)1} + \cdots + e_{\sigma(n)n} = \sum_{j=1}^{n} e_{\sigma(j)j}.$$

Umgekehrt definiert jede Permutation $\sigma \in \mathcal{S}_n$ eine $(n \times n)$-Matrix

$$P_\sigma = \sum_{i=1}^{n} e_{\sigma(j)j}$$

mit den Koeffizienten

$$P_{ij} = \begin{cases} 0 & \text{für } \sigma(j) \neq i \\ 1 & \text{für } \sigma(j) = i \end{cases}$$

Offenbar steht hier in allen Spalten genau eine 1 (in Zeile $\sigma(j)$) und sonst Nullen.

Da σ bijektiv ist, existiert die Umkehrabbildung $\sigma^{-1} : \{1, \ldots, n\} \to \{1, \ldots, n\}$, die ebenfalls eine Permutation ist. In der i-ten Zeile von P_σ steht somit auch genau eine 1 (nämlich in der Spalte $\sigma^{-1}(i)$) und sonst Nullen. Also ist P_σ eine Permutationsmatrix. Ist σ eine Transposition, so ist P_σ eine Elementarmatrix vom Typ II (Übungsaufgabe).

Proposition 3.27

i) *Sind σ und τ die Permutationen zu den Permutationsmatrizen P und Q, so ist $\sigma\tau$ die Permutation zu PQ.*

ii) *Jede Permutationsmatrix P ist invertierbar mit $P^{-1} = P^t$.*

Beweis :

i) Es ist $(PQ)e_j = P(Qe_j) = P(e_{\tau(j)}) = e_{\sigma(\tau(j))} = e_{\sigma\tau(j)}$, woraus die Behauptung folgt.

ii) P ist die Matrix mit den Spalten $e_{\sigma(1)}, \ldots, e_{\sigma(n)}$, also ist P^t die Matrix mit den Zeilen $e^t_{\sigma(1)}, \ldots, e^t_{\sigma(n)}$. Da für alle i, j gilt

$$e_i e_j^t = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases},$$

folgt $PP^t = E_n$.

Nach Korollar 2.13 ist also $P^t = P^{-1}$.

□

Korollar 3.28 *Ist P eine Permutationsmatrix, so ist* $\det P = 1$ *oder* $\det P = -1$.

Beweis : Mit Satz 3.17 folgt aus Proposition 3.27 $\det P \det P^t = 1$, also wegen $\det P = \det P^t$ (siehe Satz 3.19) auch $(\det P)^2 = 1$, woraus die Behauptung folgt. □

Ist σ die Permutation zur Permutationsmatrix P, so nennt man die Zahl

$$\text{sgn}(\sigma) = \det P \in \{1, -1\}$$

auch das **Signum** der Permutation. Nach Satz 3.17 und Proposition 3.27 ist sgn $(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau)$.

Die Permutationsmatrix zu einer Transposition τ entsteht aus E_n durch Vertauschen zweier Spalten, also ist $\text{sgn}(\tau) = -1$ nach Korollar 3.20.

Satz 3.29 (Leibnizformel) *Für jede $n \times n$-Matriz $A = (a_{ij})$ ist*

$$\det A = \sum_{\substack{\sigma \text{ Permutation} \\ \text{von } \{1,\ldots,n\}}} sgn(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

und

$$\det A = \sum_{\substack{\sigma \text{ Permutation} \\ \text{von } \{1,\ldots,n\}}} sgn(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Beweis : Da $\det A = \det A^t$ ist, genügt es, eine der beiden Formeln zu beweisen.

Wir prüfen nach, dass die Funktion $d : \mathbb{R}^{n\times n} \to \mathbb{R}$, definiert durch $d(A) = \sum\limits_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ die drei Eigenschaften aus Satz 3.16 erfüllt:

i) Ist $A = E_n$, so ist $a_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$,

   woraus $a_{i\sigma(i)} = 0$ folgt, falls $\sigma(i) \neq i$ ist. Also bleibt nur der Summand für $\sigma = \text{id}$ übrig, und es folgt

$$d(E_n) = \text{sgn}(\text{id}) = 1.$$

ii) Ist $A = \begin{pmatrix} | \\ z_i \\ | \end{pmatrix}$, $B = \begin{pmatrix} | \\ z_i \\ | \end{pmatrix}$ und $C = \begin{pmatrix} | \\ z_i + \tilde{z}_i \\ | \end{pmatrix}$, so gilt $a_{k\sigma(k)} = b_{k\sigma(k)} = c_{k\sigma(k)}$ für alle $\sigma \in \mathcal{S}_n$ und alle $k \neq i$, sowie $a_{i\sigma(i)} + b_{i\sigma(i)} = c_{i\sigma(i)}$ für alle $\sigma \in \mathcal{S}_n$.

Somit folgt

$$\begin{aligned} d(C) &= \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) c_{1\sigma(1)} \cdots (a_{i\sigma(i)} + b_{i\sigma(i)}) \cdot \ldots \cdot c_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)} = \det A + \det B. \end{aligned}$$

Analog zeigt man $d \begin{pmatrix} | \\ az_i \\ | \end{pmatrix} = a \;\; d \begin{pmatrix} | \\ z_i \\ | \end{pmatrix}$.

Also ist d linear in den Zeilen.

iii) Sind zwei benachbarte Zeilen in $A = (a_{ij})$ gleich, so gilt für einen Zeilenindex i: $a_{ij} = a_{i+1j}$ für alle $j = 1, \ldots, n$. Es sei τ die Transposition, die i und $i+1$ vertauscht. Mit σ durchläuft auch $\sigma\tau$ die Menge $\mathcal{S}_n$. Also ist

$$\begin{aligned} d(A) &= \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma\tau) a_{1\sigma\tau(1)} \cdots a_{n\sigma\circ\tau(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} -\mathrm{sgn}(\sigma) a_{1\sigma\tau(1)} \cdots a_{n\sigma\tau(n)} \\ &\overset{\tau \text{ ausrechnen}}{=} \sum_{\sigma \in \mathcal{S}_n} -\mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i+1)} a_{i+1\sigma(i)} \cdots a_{n\sigma(n)} \\ &= -\sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{i+1\sigma(i+1)} a_{i\sigma(i)} \cdots a_{n\sigma(n)} \\ &= -d(A), \end{aligned}$$

woraus aus $d(A) = 0$ folgt.

Mit Satz 3.16 folgt also $d(A) = \det A$ für alle $A \in \mathbb{R}^{n \times n}$.

□

Die Leibniz-Formel eignet sich für großes n meist nicht zum Berechnen der Determinante, aber sie liefert die wichtige Information, dass die Determinante ein „ Polynom“ in den Koeffizienten ist. (Was ein Polynom ist, werden wir später definieren).

Sind etwa die Matrixeinträge $a_{ij} = a_{ij}(x)$ stetige oder differenzierbare Funktionen in x, so folgt sofort aus der Leibniz-Formel, dass auch $\det A$ eine stetige oder differenzierbare Funktion in x ist.

Wir wollen jetzt noch die sogenannten Cramerschen Regeln herleiten.

Definition 3.30 *Es sei $A = (a_{ij})$ eine $n \times n$-Matrix. Die Adjungierte Matrix AdjA zu A ist definiert als die $n \times n$-Matrix mit den Einträgen $(AdjA)_{ij} = (-1)^{i+j} \det A_{ji} =: \alpha_{ji}$, d.h.*

$$\begin{aligned} AdjA &= (\alpha_{ij})^t_{ij} \textit{ mit} \\ \alpha_{ij} &= (-1)^{i+j} \det A_{ij}. \end{aligned}$$

Beispiel:

i) $Adj \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

ii) $Adj \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 1 & -2 \\ -2 & 0 & 1 \\ -3 & -1 & 2 \end{pmatrix}^t = \begin{pmatrix} 4 & -2 & -3 \\ 1 & 0 & -1 \\ -2 & 1 & 2 \end{pmatrix}$

Satz 3.31 *Für jede $n \times n$-Matrix $A = (a_{ij})$ gilt*

$$\begin{aligned} (AdjA)A &= (\det A)E_n \textit{ und} \\ A(AdjA) &= (\det A)E_n, \end{aligned}$$

wobei $(\det A)E_n = \begin{pmatrix} \det A & & 0 \\ & \ddots & \\ 0 & & \det A \end{pmatrix}$

ist.

Beweis : Der Eintrag c_{ij} an der Stelle (i,j) im Matrixprodukt $(AdjA)A$ ist definitionsgemäß

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} (\mathrm{Adj}A)_{ik} a_{kj} = \sum_{k=1}^{n} \alpha_{ki} a_{kj} \\ &= \sum_{k=1}^{n} (-1)^{i+k} \det A_{ki} a_{kj} \end{aligned}$$

Für $i = j$ ist nach der Entwicklungsformel Satz 3.21 $c_{ij} = \det A$.

Für $i \neq j$ sei B die Matrix, die aus A entsteht, wenn man die i-te Spalte durch die j-te Spalte ersetzt. B hat also zwei übereinstimmende Spalten, woraus mit Korollar 3.20 $\det B = 0$ folgt.

Nun ist $B_{ki} = A_{ki}$ und $b_{ki} = a_{kj}$ für $k = 1, \ldots, n$.

Also ergibt die Entwicklung von $\det B$ nach der i-ten Spalte:

$$\begin{aligned} 0 = \det B &= \sum_{k=1}^{n} (-1)^{i+k} b_{ki} \det(B_{ki}) \\ &= \sum_{k=1}^{n} (-1)^{i+k} a_{kj} \det A_{ki}. \end{aligned}$$

Daher ist $c_{ij} = 0$ für $i \neq j$, woraus die Behauptung folgt. □

Korollar 3.32 (Cramer'sche Regel für die Matrixinversion)

Ist $\det A \neq 0$, *so gilt*

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

Beweis : Ist $\det A \neq 0$, so folgt aus Satz 3.31:

$$\left(\frac{1}{\det A}(\text{Adj}A)\right) A = E_n.$$

□

Beispiel:

i) Falls $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \neq 0$ ist, gilt

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

ii) Es ist $\det \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} = 1$, also folgt (siehe obiges Beispiel)

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}^{-1} = Adj \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & -2 & -3 \\ 1 & 0 & -1 \\ -2 & 1 & 2 \end{pmatrix}$$

Sei A eine $n \times n$-Matrix mit $\det A \neq 0$. Wir betrachten für einen n-dimensionalen Spaltenvektor b das lineare Gleichungssystem $Ax = b$. Da A nach Korollar 3.15 invertierbar ist, ist $x = A^{-1}b$ die einzige Lösung dieses Gleichungssystems.

Satz 3.33 (Cramer'sche Regel zur Lösung linearer Gleichungssysteme)

Ist $\det A \neq 0$*, so gilt für die eindeutig bestimmte Lösung* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *des linearen Gleichungssystems* $Ax = b$ *die Formel*

$$x_j = \frac{\det M_j}{\det A},$$

wobei M_j *die Matrix ist, die aus* A *entsteht, indem wir die* j*-te Spalte durch den Spaltenvektor* b *ersetzen.*

Beweis : Die einzige Lösung x von $Ax = b$ ist $x = A^{-1}b$. Nach Korollar 3.32 gilt also $x = \frac{1}{\det A}(\mathrm{Adj}A)b$, d.h. für $j = 1, \ldots, n$ ist

$$\begin{aligned} x_j &= \frac{1}{\det A} \sum_{k=1}^{n} (\mathrm{Adj}A)_{jk} b_k \\ &= \frac{1}{\det A} \sum_{k=1}^{n} \alpha_{kj} b_k = \frac{1}{\det A} \sum_{k=1}^{n} (-1)^{k+j} (\det A_{kj}) b_k. \end{aligned}$$

Entwickeln wir $\det M_j$ nach der j-ten Spalte, so folgt aus $(M_j)_{kj} = A_{kj}$ und der Tatsache, dass b_k der Eintrag von M_j an der Stelle (k, j) ist, dass $\det M_j = \sum_{k=1}^{n} (-1)^{k+j} \det A_{kj} b_k$ gilt. Daraus folgt $x_j = \frac{\det M_j}{\det A}$. □

Zur Berechnung der Lösung eines linearen Gleichungssystems gibt es im allgemeinen effektivere Verfahren als die Cramer'sche Regel. Trotzdem liefert uns Satz 3.33 eine wichtige Einsicht. Diese Formel beschreibt nämlich, wie die Lösung eines linearen Gleichungssystems $Ax = b$ von den Koeeffizienten von A und b abhängt: Die Koeffizienten der Lösung x sind nämlich jeweils ein Quotient von zwei „Polynomen" (was immer das ist) in den Koeffizienten von A und b.

4 Gruppen und Körper

Unter einer Verknüpfung auf einer Menge X verstehen wir eine Funktion

$$m : X \times X \to X.$$

Die Addition und die Multiplikation reeller Zahlen sind etwa Verknüpfungen auf $\mathbb{R}$. Die Schreibweise $m(x, y)$ ist für Verknüpfungen eher unpraktisch, wir werden stattdessen, je nachdem, um welche Verknüpfung es sich handelt, $a+b$, ab oder $a \circ b$ schreiben.

Definition 4.1 *Eine **Gruppe** ist eine Menge G zusammen mit einer Verknüpfung*

$$m : G \times G \to G,$$

die wir auch als $m(a, b) = ab$ schreiben, so dass gilt

i) Die Verknüpfung m ist assoziativ, d.h. es gilt $(ab)c = a(bc)$ für alle $a, b, c \in G$.

ii) G besitzt ein neutrales Element bezüglich m, d.h. es existiert ein $e \in G$ mit $ea = ae = a$ für alle $a \in G$.

iii) Jedes Element von G besitzt ein Inverses, d.h. für alle $a \in G$ existiert ein $b \in G$ mit $ab = ba = e$.

Ist G eine Gruppe und $a \in G$, so ist das Inverse zu a in G eindeutig bestimmt. Sind nämlich b und b' Elemente in G, für die $ab = ba = e$ und $ab' = b'a = e$ gilt, so folgt

$$b = eb = (b'a)b = b'(ab) = b'e = b'.$$

Also schreiben wir auch einfach a^{-1} für das Inverse von a.

Man beachte, dass in dieser Definition das Produkt ab nur eine Schreibweise für die Verknüpfung m ist. Würden wir die Bedingungen i) - iii) mit der Abbildung m schreiben, dann sähen sie recht unübersichtlich aus. Es muss aber für eine konkrete Gruppe G mit m nicht notwendig eine Multiplikation gemeint sein. Für manche Gruppen schreiben wir die Verknüpfung

$$m : G \times G \to G$$

als $m(a, b) = a + b$. Dann lauten die Bedingungen i) - iii) in Definition 4.1 wie folgt:

i) $(a + b) + c = a + (b + c)$ für alle $a, b, c \in G$.

ii) Es existiert ein $e \in G$ mit $e + a = a + e = a$ für alle $a \in G$. Wir nennen e auch Nullelement und bezeichnen es oft mit 0.

iii) Für alle $a \in G$ existiert ein Inverses b mit $a + b = b + a = e$. Dieses ist eindeutig bestimmt, und wir bezeichnen es mit $(-a)$.

Beispiele:

i) $\mathbb{Z}$ mit der Addition $+$, dem neutralen Element 0 und dem Inversen $(-a)$ zu a ist eine Gruppe.

ii) $\mathbb{N}$ mit der Addition $+$ bildet keine Gruppe, $\mathbb{N}_0$ auch nicht.

iii) $\mathbb{R}$ mit der Addition $+$ ist eine Gruppe, $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ mit der Multiplikation ist auch eine Gruppe.

iv) Die Menge $\mathrm{GL}_n(\mathbb{R})$ der invertierbaren Matrizen in $\mathbb{R}^{n\times n}$ ist eine Gruppe bezüglich der Multiplikation von Matrizen mit neutralem Element E_n und Inversem A^{-1}.

v) Die Menge S_n der Permutation von $\{1, \ldots, n\}$, also der bijektiven Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ ist eine Gruppe bezüglich der Hintereinanderausführung (Komposition) von Abbildungen. Das neutrale Element ist id (die Identität), definiert durch $\mathrm{id}(i) = i$. Das inverse Element zu einer Permutation σ ist die Permutation σ^{-1}, definiert durch

$$\sigma^{-1}(j) = i \text{ genau dann, wenn } \sigma(i) = j.$$

Wir wollen jetzt zeigen, dass eine Menge G mit einer Verknüpfung $m(a, b) = ab$ genau dann eine Gruppe ist, wenn gilt:

i) $(ab)c = a(bc)$ für alle $a, b, c \in G$

ii)′ Es gibt ein $e \in G$ mit $ea = a$ für alle $a \in G$

iii)′ Für alle $a \in G$ existiert ein Inverses $b \in G$ mit $ba = e$.

Mit anderen Worten, man kann die Bedingungen ii) bzw. iii) in Definition 4.1 durch die schwächeren Bedingungen ii)′ bzw. iii)′ ersetzen. Dies sieht man folgendermaßen:

Beweis : Für jedes $a \in G$ existiert nach iii)′ ein „Linksinverses“ b mit $ba = e$. Zu b existiert ebenfalls ein „Linksinverses“ c mit $cb = e$. Also folgt:

$$ab \overset{ii)'}{=} (ea)b = (cb)ab \overset{i)}{=} c(ba)b = c(eb) \overset{ii)'}{=} cb = e.$$

Also gilt $ab = ba = e$, d.h. es gilt iii) aus Definition 4.1.

Nun sei $a \in G$ und $b \in G$ ein Element mit $ab = ba = e$. Dann gilt:

$$ae = a(ba) \overset{i)}{=} (ab)a = ea \overset{ii)'}{=} a,$$

also ist $ae = ea = a$, d.h. ii) aus Definition 4.1 gilt. □

Definition 4.2 *Eine Gruppe G mit Verknüpfung $m(ab) = ab$ heißt **kommutativ oder abelsch**, falls gilt*

$$ab = ba \text{ für alle } a, b \in G.$$

Beispiel:

i) $(\mathbb{Z}, +), (\mathbb{R}, +)$ und $(\mathbb{R}^\times, \cdot)$ sind abelsche Gruppen.

ii) $GL_n(\mathbb{R})$ ist nicht abelsch für $n \geq 2$.

iii) S_n ist nicht abelsch für $n \geq 3$ (siehe Aufgabenzettel).

Definition 4.3 *Es sei G eine Gruppe mit der Verknüpfung $m(a, b) = ab$. Eine Teilmenge $H \subset G$ heißt **Untergruppe** von G, wenn gilt*

i) $a, b \in H \Rightarrow ab \in H$ (d.h. H ist abgeschlossen unter der Verknüpfung).

ii) $e \in H$ (d.h. H enthält das neutrale Element)

iii) $a \in H \Rightarrow a^{-1} \in H$ (d.h. für jedes Element in H liegt auch das Inverse in H).

Beispiele:

i) Jede Gruppe G hat die trivialen Untergruppen $\{e\}$ und G.

ii) $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{R}, +)$.

iii) Die Permutationsmatrizen bilden eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$.

Wir wollen nun als Beispiel für den Untergruppenbegriff alle Untergruppen der abelschen Gruppen $(\mathbb{Z}, +)$ bestimmen. Dazu brauchen wir den Begriff der Division mit Rest.

Lemma 4.4 *Es sei $q \in \mathbb{N}$. Jedes Element $a \in \mathbb{Z}$ lässt sich auf eindeutige Weise als*

$$a = kq + r$$

mit einem $k \in \mathbb{Z}$ und einem $r \in \{0, 1, \ldots, q-1\}$ darstellen.

Beweis : Wir betrachten zu a und q die Menge $R = \{a - kq : k \in \mathbb{Z}\} \cap \mathbb{N}_0$. Offenbar ist $R \neq \emptyset$. Es sei r die kleinste Zahl in R. Diese ist in $\{0, 1, \ldots, q-1\}$ enthalten, denn sonst liegt $r - q$ auch in R. Also gibt es ein $k \in \mathbb{Z}$ und ein $r \in \{0, 1, \ldots, q-1\}$ mit $a = kq + r$. Ist $a = k'q + r'$ mit $k' \in \mathbb{Z}$ und $r' \in \{0, 1, \ldots, q-1\}$ eine zweite solche Darstellung von a, so folgt $r' = k'q - a \in R$, also $r' \geq r$. Außerdem gilt $0 = (k'q + r') - (kq + r) = (k' - k)q + (r' - r)$, das heißt $r' - r = (k - k')q$. Nun ist $r' - r \in \{0, \ldots, q-1\}$, woraus $r' - r = 0 = (k - k')$ folgt. Also ist $r = r'$ und $k = k'$, das heißt die Darstellung ist eindeutig. $\square$

Satz 4.5 *Für jede ganze Zahl b ist die Teilmenge $b\mathbb{Z} := \{bk : k \in \mathbb{Z}\}$ von $\mathbb{Z}$ mit der Verknüpfung $+$ eine Untergruppe von $\mathbb{Z}$. Jede Untergruppe von $(\mathbb{Z}, +)$ ist von der Form $b\mathbb{Z}$ für ein $b \in \mathbb{Z}$.*

Beweis : Wir prüfen zunächst, dass $b\mathbb{Z}$ eine Untergruppe ist. Ist bk_1 und $bk_2 \in b\mathbb{Z}$, so ist $bk_1 + bk_2 = b(k_1 + k_2) \in b\mathbb{Z}$. Außerdem ist $0 = b \cdot 0 \in b\mathbb{Z}$ und mit $bk \in b\mathbb{Z}$ liegt auch $-bk = b(-k)$ in $b\mathbb{Z}$. Nach Definition 4.3 ist $b\mathbb{Z}$ also eine Untergruppe von $\mathbb{Z}$.

Nun zeigen wir, dass jede Untergruppe H von dieser Form ist. Da H eine Untergruppe von $(\mathbb{Z}, +)$ ist, gilt $0 \in H$. Ist 0 das einzige Element in H, so ist $H = \{0\} = 0\mathbb{Z}$ von der gewünschten Gestalt.

Enthält H ein $a \neq 0$ aus $\mathbb{Z}$, so ist entweder $a \in \mathbb{N}$ oder $-a \in \mathbb{N}$. Da H mit a auch das Inverse $-a$ enthält, existiert also eine natürliche Zahl $n \in H$. Wir definieren b als die kleinste natürliche Zahl in H und behaupten $H = b\mathbb{Z}$.

„$\supset$“: Sei $k \in \mathbb{Z}$. Ist $k > 0$, so ist $kb = \underbrace{b + \cdots + b}_{k\text{-mal}} \in H$, da H eine Untergruppe ist. Ist $k = 0$, so ist $kb = 0 \in H$. Ist $k < 0$, so haben wir schon gesehen, dass $(-k)b \in H$ it, also folgt $kb = -(-k)b \in H$.

„$\subset$“: Sei m eine beliebige Zahl in H. Division mit Rest durch die Zahl $b \in H$ liefert $k \in \mathbb{Z}$ und $r \in \{0, 1, \ldots, b-1\}$ mit

$$m = kb + r.$$

Nun ist $kb \in b\mathbb{Z} \subset H$, wie wir oben gezeigt haben. Also ist $-kb \in H$ und daher ist auch

$$r = m - kb$$

in H. Nun ist $r < b$, wegen der Minimalität von b folgt also $r = 0$. Somit ist $m = kb \in H$. $\square$

Lemma 4.6 *Es ist $d\mathbb{Z} = d'\mathbb{Z}$ genau dann, wenn $d = d'$ oder $d = -d'$ ist.*

Beweis : Die Richtung „$\Leftarrow$“ ist klar. Wir zeigen „$\Rightarrow$“ das heißt es gelte $d\mathbb{Z} = d'\mathbb{Z}$. Dann ist $d \in d'\mathbb{Z}$, das heißt $d = d'k$ für ein $k \in \mathbb{Z}$. Ferner ist $d' \in d\mathbb{Z}$, das heißt $d' = dl$ für ein $l \in \mathbb{Z}$. Somit folgt $d = d'k = dlk$. Ist $d \neq 0$, so impliziert dies $1 = lk$, das heißt $l, k \in \{\pm 1\}$ und $d = \pm d'$. Ist $d = 0$, so ist auch $d' = dl = 0$, das heißt $d = d'$. $\square$

Die Elemente der Untergruppe $b\mathbb{Z}$ von $\mathbb{Z}$ sind offenbar genau diejenigen ganzen Zahlen, die durch b teilbar sind.

Nun betrachten wir zwei ganze Zahlen a und b, die nicht beide Null sind. Es sei

$$\begin{aligned} a\mathbb{Z} + b\mathbb{Z} &= \{z \in \mathbb{Z} : z = r + s \text{ für } r \in a\mathbb{Z} \text{ und } s \in b\mathbb{Z}\} \\ &= \{z \in \mathbb{Z} : z = ak + bl \text{ für } k, l \in \mathbb{Z}\}. \end{aligned}$$

Diese Menge ist eine Untergruppe von $(\mathbb{Z}, +)$ (Übungsaufgabe). Sie heißt die von a und b erzeugte Untergruppe von $\mathbb{Z}$.

Lemma 4.7 *$a\mathbb{Z} + b\mathbb{Z}$ ist die kleinste Untergruppe von $\mathbb{Z}$, die a und b enthält, das heißt ist $H \subset \mathbb{Z}$ eine Untergruppe mit $a \in H$ und $b \in H$, so folgt $a\mathbb{Z} + b\mathbb{Z} \subset H$.*

Beweis : Da H eine Untergruppe ist, die a enthält, liegen 0 sowie alle $na = \underbrace{a + \cdots + a}_{n\text{-mal}}$ für $n \in \mathbb{N}$ in H.

Da H mit jedem Element sein Inverses enthält, liegt auch $(-n)a = -na \in H$. Insgesamt gilt also $a\mathbb{Z} \subset H$. Genauso folgt $b\mathbb{Z} \subset H$. Da H abgeschlossen unter $+$ ist, folgt $a\mathbb{Z} + b\mathbb{Z} \subset H$. □

Nach Satz 4.5 gibt es also ein $d \in \mathbb{Z}$, so dass die Untergruppe $a\mathbb{Z} + b\mathbb{Z}$ von $\mathbb{Z}$ mit $d\mathbb{Z}$ übereinstimmt. Da a und b nicht beide Null sind, ist auch $d \neq 0$. Nach Lemma 4.6 gibt es genau ein $d \in \mathbb{N}$ mit

$$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}.$$

Wir wollen diese Zahl d jetzt untersuchen.

Satz 4.8 *Es seien $a, b \in \mathbb{Z}$ nicht beide Null und $d \in \mathbb{N}$ die Zahl mit*

$$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}.$$

Dann gilt

i) d lässt sich schreiben als $d = ax + by$ für geeignete $x, y \in \mathbb{Z}$

ii) d teilt a und b.

iii) Teilt eine ganze Zahl e beide Zahlen a und b, so teilt e auch d.

Beweis :

i) folgt aus $d \in d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$.

ii) $a \in a\mathbb{Z} \subset a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$, also ist $a = dk$ für ein $k \in \mathbb{Z}$. Daher ist d ein Teiler von a. Genauso zeigt man, dass d ein Teiler von b ist.

iii) Ist e ein Teiler von a und b, so gilt $a = ek$ und $b = el$ für $k, l \in \mathbb{Z}$. Also ist $a \in e\mathbb{Z}$ und $b \in e\mathbb{Z}$, das heißt $e\mathbb{Z}$ ist eine Untergruppe von $\mathbb{Z}$, die a und b enthält. Da $d\mathbb{Z}$ nach Lemma 4.7 die kleinste solche Untergruppe ist, folgt $d\mathbb{Z} \subset e\mathbb{Z}$. Also gilt $d = em$ für ein $m \in \mathbb{Z}$, das heißt e teilt d.

$\square$

Aufgrund der Eigenschaften ii) und iii) von Satz 4.8 nennen wir die eindeutig bestimmte natürliche Zahl d mit

$$d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$$

auch den **größten gemeinsamen Teiler** von a und b und schreiben $d = \mathrm{ggT}(a, b)$. Nach Satz 4.8 i) kann man den größten gemeinsamen Teiler von a und b „ linear aus a und b kombinieren", das heißt es gibt $x, y \in \mathbb{Z}$ mit $d = ax + by$.

Das Problem, zu gegebenen ganzen Zahlen a und b den $\mathrm{ggT}(a, b)$ und die Zahlen x und y zu berechnen, lässt sich mit dem Euklidischen Algorithmus lösen.

Jetzt lernen wir eine weitere Struktur kennen, die die reellen Zahlen verallgemeinert.

Definition 4.9 *Ein **Körper** ist eine Menge K mit zwei Abbildungen*

$$\begin{aligned} + : \quad K \times K &\to K \quad \textit{(Addition)} \\ (a, b) &\mapsto a + b \end{aligned}$$

und

$$\begin{aligned} \cdot : \quad K \times K &\to K \quad \textit{(Multiplikation)} \\ (a, b) &\mapsto ab \end{aligned}$$

so dass folgende Bedingungen erfüllt sind:

i) $(a + b) + c = a + (b + c)$ für alle $a, b, c \in K$

ii) Es existiert ein Element $0 \in K$, so dass $0 + a = a$ für alle $a \in K$.

iii) *Für jedes* $a \in K$ *existiert ein Element* $(-a) \in K$ *mit* $(-a) + a = 0$.

iv) $a + b = b + a$ *für alle* $a, b \in K$

v) $(ab)c = a(bc)$ *für alle* $a, b, c \in K$

vi) *Es existiert ein Element* $1 \in K$ *mit* $1a = a$ *für alle* $a \in K$.

vii) *Für jedes* $a \neq 0$ *in* K *existiert ein Element* $a^{-1} \in K$ *mit* $a^{-1}a = 1$.

viii) $ab = ba$ *für alle* $a, b \in K$

ix) $a(b + c) = ab + ac$ *und* $(a + b)c = ac + bc$ *für alle* $a, b, c \in K$

x) $1 \neq 0$.

Wenn wir hier wie in ix) keine Klammern setzen, so soll immer die Multiplikation Vorrang vor der Addition haben.

Etwas prägnanter können wir sagen:

Definition 4.9' Ein Körper ist eine Menge K mit zwei Verknüpfungen

$$\begin{aligned} + : \quad & K \times K && \to K \text{ und} \\ \cdot : \quad & K \times K && \to K, \end{aligned}$$

so dass gilt:

i) $(K, +)$ ist eine abelsche Gruppe.

ii) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

iii) Es gelten die Distributivgesetze ix) aus Definition 4.9.

Wir schreiben $K^\times = K \setminus \{0\}$ und nennen diese multiplikative Gruppe die Einheitengruppe des Körpers. Außerdem schreiben wir für $n \in \mathbb{Z}$:

$$na = \begin{cases} \underbrace{a + \cdots + a}_{n\text{-mal}}, & n > 0 \\ 0, & n = 0 \\ -\underbrace{(a + \cdots + a)}_{(-n)\text{-mal}}, & n < 0 \end{cases}$$

sowie

$$a^n = \begin{cases} \underbrace{a \cdot \ldots \cdot a}_{n\text{-mal}}, & n > 0 \\ 1, & n = 0 \\ \underbrace{(a \cdot \ldots \cdot a)}_{(-n)\text{-mal}}{}^{-1}, & n < 0. \end{cases}$$

Beispiel: Die reellen Zahlen $\mathbb{R}$, die rationalen Zahlen $\mathbb{Q}$ und die komplexen Zahlen $\mathbb{C}$ (konstruiert in der Analysis) bilden jeweils einen Körper.

Ähnlich wie bei Gruppen hat man auch bei Körpern den Begriff des Unter- (oder Teil-)körpers.

Definition 4.10 *Es sei K ein Körper. Eine Teilmenge $L \subset K$ heißt **Teilkörper**, wenn gilt:*

i) $a, b \in L \Rightarrow a + b \in L$ *und* $ab \in L$

ii) $0 \in L$ *und* $1 \in L$

iii) $a \in L \Rightarrow -a \in L$

iv) $a \in L, a \neq 0 \Rightarrow a^{-1} \in L.$

*Ist $L \neq K$, so heißt L **echter Teilkörper** von K.*

Eine Teilmenge $L \subset K$ ist also genau dann ein Körper, wenn die Addition und die Multiplikation in K sich zu Abbildungen

$$\begin{aligned} + : \quad L \times L &\to L \text{ und} \\ \cdot : \quad L \times L &\to L \end{aligned}$$

einschränken lassen, und wenn L mit diesen Verknüpfungen selbst ein Körper ist. (Prüfen Sie das!)

Beispiel:

i) $\mathbb{Q} \subset \mathbb{R}$ ist ein (echter) Teilkörper.

ii) $\mathbb{R} \subset \mathbb{C}$ ist ein (echter) Teilkörper.

In jedem Körper K gelten einige Rechenregeln, die wir von den reellen Zahlen her kennen. Für alle $a, b \in K$ gilt:

i) $0a = a0 = 0$

ii) $(-1)a = -a$

iii) Ist $ab = 0$, so ist $a = 0$ oder $b = 0$

iv) $(a+b)^n = \sum\limits_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

für alle $n \in \mathbb{N}_0$.

Hier ist $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot \ldots \cdot (n-k+1)}{k \cdot \ldots \cdot 1} \in \mathbb{N}$ und das Produkt der natürlichen Zahl $\binom{n}{k}$ mit dem Körperelement $a^k b^{n-k}$ ist wie oben definiert.

Es gibt aber einen Punkt, an dem man mit den gewohnten Rechenregeln aufpassen muss. Das Körperelement $na = \underbrace{(a + \cdots + a)}_{n\text{-mal}} (n \in \mathbb{N})$ kann auch mal Null sein.

Genauer gesagt: Die Abbildung

$$\begin{array}{rcl} \mathbb{N} & \to & K \\ n & \mapsto & n1 = \underbrace{(1 + \cdots + 1)}_{n\text{-mal}} \end{array}$$

ist im allgemeinen nicht injektiv.

Definition 4.11 *Die kleinste natürliche Zahl n mit $n1 = 0$ heißt die **Charakteristik** von K. Wir bezeichnen sie mit char(K). Ist für jedes $n \in \mathbb{N}$ die Zahl $n1 \neq 0$, so wird die Charakteristik von K gleich 0 gesetzt, und wir schreiben char$(K) = 0$.*

Beispiel: char$(\mathbb{Q}) =$ char$(\mathbb{R}) =$ char$(\mathbb{C}) = 0$.

Lemma 4.12 *Ist K ein Körper mit char$(K) \neq 0$, so ist char(K) eine Primzahl. (Eine Primzahl ist eine natürliche Zahl p, die als einzige Teiler in $\mathbb{N}$ die Zahlen 1 und p besitzt.)*

Beweis : Sei $p =$ char(K). Dann ist $p \neq 0$. Angenommen, $p = nm$ mit $n, m \in \mathbb{N}$. Dann ist

$$0 = p1 = \underbrace{1 + \cdots + 1}_{nm\text{-mal}} = \underbrace{n1 + \cdots + n1}_{m\text{-mal}} = \underbrace{(1 + \cdots + 1)}_{m\text{-mal}}(n1) = (m1)(n1)$$

Also folgt $m1 = 0$ oder $n1 = 0$. Da p die kleinste natürliche Zahl mit $p1 = 0$ ist, folgt $p = m$ oder $p = n$. Also ist p eine Primzahl. $\square$

Wir wollen jetzt Beispiele für Körper kennenlernen, deren Charakteristik eine Primzahl ist.

Wir bezeichnen für jede ganze Zahl a mit $\overline{a} \in \{0, 1, \dots, p-1\}$ den Rest von a bei Division durch p, das heißt a ist die eindeutig bestimmte Zahl aus $\{0, 1, \dots, p-1\}$ mit $a = kp + \overline{a}$ für ein $k \in \mathbb{Z}$ (siehe Lemma 4.4).

Es gilt $\overline{a} = \overline{b}$ genau dann, wenn $a - b$ ein Vielfaches von p ist:

Gilt nämlich $\overline{a} = \overline{b}$, so sei $a = kp + \overline{a}$ und $b = lp + \overline{b}$ für $k, l \in \mathbb{Z}$. Dann ist $a - b = (k-l)p + (\overline{a} - \overline{b}) = (k-l)p$ ein Vielfaches von p. Ist umgekehrt $a - b = mp$ für ein $m \in \mathbb{Z}$ und $a = kp + \overline{a}$, so folgt $b = a - mp = (k-m)p + \overline{a}$, also $\overline{a} = \overline{b}$, da der Rest nach Lemma 4.4 eindeutig bestimmt ist. Ferner rechnet man leicht nach, dass $\overline{a+b} = \overline{a} + \overline{b}$ und $\overline{ab} = \overline{a}\overline{b}$ gilt. (Prüfen Sie das!)

Es sei $\mathbb{F}_p := \{0, 1, \dots, p-1\}$ die Menge der Zahlen $0, 1, \dots, p-1$. Wir definieren jetzt zwei Verknüpfungen $+$ und $\cdot$ auf $\mathbb{F}_p$. Um diese von der Addition und der Multiplikation auf den ganzen Zahlen unterscheiden zu können, bezeichnen wir letztere für den Moment mit $+_{\mathbb{Z}}$ bzw. $\cdot_{\mathbb{Z}}$.

Für $a, b \in \mathbb{F}_p$ setzen wir

$$a + b = \overline{a +_{\mathbb{Z}} b}$$

und

$$ab = \overline{a \cdot_{\mathbb{Z}} b}$$

Beispiel: Es sei $p = 11$. Dann gilt in $\mathbb{F}_{11}$:

$$3 + 5 = 8,\ 4 + 9 = 2,\ 10 + 10 = 9$$

sowie

$$2 \cdot 3 = 6,\ 3 \cdot 4 = 1,\ 2^4 = 5$$

Lemma 4.13 *Für $a_1, \dots, a_n \in \mathbb{F}_p$ gilt*

$$(a_1 \cdot \ldots \cdot a_{n-1}) \cdot a_n = \overline{a_1 \cdot_{\mathbb{Z}} \cdots \cdot_{\mathbb{Z}} a_n}$$

und

$$(a_1 + \cdots + a_{n-1}) + a_n = \overline{a_1 +_{\mathbb{Z}} \cdots +_{\mathbb{Z}} a_n}$$

Beweis : Das folgt mit Induktion nach n aus der Definition der Addition und der Multiplikation in $\mathbb{F}_p$. (Führen Sie das Argument aus!) □

Korollar 4.14 *Die Verknüpfungen $+$ und $\cdot$ auf $\mathbb{F}_p$ sind kommutativ, assoziativ und erfüllen die Distributivgesetze ix) in Definition 4.9.*

Beweis : Das folgt sofort aus Lemma 4.13. $\square$

Satz 4.15 $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ *mit den Verknüpfungen $+$ und $\cdot$ ist ein Körper der Charakteristik p.*

Beweis : Wir zeigen zunächst, dass $(\mathbb{F}_p, +)$ eine abelsche Gruppe ist. Offenbar ist 0 ein neutrales Element. Ferner gilt $0 + 0 = 0$ sowie $a + (p - a) = 0$ für alle $a \in \{1, \ldots, p-1\}$. Also hat jedes Element ein additives Inverses. Mit Korollar 4.14 ist $(\mathbb{F}_p, +)$ also eine abelsche Gruppe. Jetzt untersuchen wir $(\mathbb{F}_p \setminus \{0\}, \cdot)$. Offenbar ist 1 ein neutrales Element. Ist $a \in \mathbb{F}_p^\times = \mathbb{F}_p \setminus \{0\}$, so ist $a \in \{1, \ldots, p-1\}$, also folgt $\mathrm{ggT}(a, p) = 1$. Nach Satz 4.8 gibt es ganze Zahlen r und s mit

$$1 = r \cdot_{\mathbb{Z}} a +_{\mathbb{Z}} s \cdot_{\mathbb{Z}} p$$

Es sei $b = \overline{r} \in \mathbb{F}_p$. Dann ist $a \cdot_{\mathbb{Z}} b - a \cdot_{\mathbb{Z}} r$ durch p teilbar, also folgt $ab = \overline{a \cdot_{\mathbb{Z}} b} = \overline{a \cdot_{\mathbb{Z}} r}$. Da $\overline{s \cdot_{\mathbb{Z}} p} = 0$ ist, folgt $1 = \overline{r \cdot_{\mathbb{Z}} a +_{\mathbb{Z}} s \cdot_{\mathbb{Z}} p} = ab$. Also ist b ein Inverses zu a.

Mit Korollar 4.14 ist $(\mathbb{F}_p^\times, \cdot)$ also eine multiplikative Gruppe. Da die Distributivgesetze gelten, ist $\mathbb{F}_p$ nach Definition 4.9′ ein Körper. In $\mathbb{F}_p$ gilt

$$p \cdot 1 = \underbrace{1 + \cdots + 1}_{p-\text{mal}} = \overline{\underbrace{1 +_{\mathbb{Z}} \cdots +_{\mathbb{Z}} 1}_{p\text{-mal}}} = \overline{p} = 0$$

also gilt $\mathrm{char}\,(\mathbb{F}_p) = p$. $\square$

Für $a \in \mathbb{F}_p$ schreiben wir auch $-a$ für das Inverse bezüglich der Addition.

Also ist in $\mathbb{F}_7$ etwa $-5 = 2$.

Ist $a \neq 0$, so schreiben wir auch a^{-1} für das Inverse bezüglich der Multiplikation. Also ist in $\mathbb{F}_5$ etwa $3^{-1} = 2$.

Für die Existenz eines multiplikativen Inversen ist es entscheidend, dass die Zahl p eine Primzahl ist. Betrachtet man etwa die Menge $\{0, 1, \cdots, 5\}$ mit den analogen Verknüpfungen, definiert durch den Rest bei Division durch 6, so folgt $2 \cdot 3 = 0$. Also kann keines dieser Elemente invertierbar sein!

Sei nun K ein beliebiger Körper. Wir bezeichnen mit $K^{n\times n}$ die Menge der $n \times n$-Matrizen mit Einträgen in K.

Gehen wir nun noch einmal alle Definitionen und Resultate aus §1 - §3 durch, so stellen wir fest, dass wir überall in $\mathbb{R}$ nur die Rechenregeln verwendet haben, die in jedem Körper gelten. Also gilt

Satz 4.16 *Alle Ergebnisse aus §1, 2 und 3 gelten für Matrizen mit Einträgen in einem beliebigen Körper K bzw. für lineare Gleichungssysteme mit Koeffizienten in einem beliebigen Körper K, wenn man überall $\mathbb{R}$ durch K ersetzt.*

Wir bezeichnen für einen beliebigen Körper K mit $GL_n(K)$ die Gruppe der invertierbaren $n \times n$-Matrizen mit Einträgen in K.

5 Vektorräume

Definition 5.1 *Es sei K ein Körper. Ein K-**Vektorraum** oder ein **Vektorraum über** K ist eine Menge V zusammen mit zwei Abbildungen.*

$$\begin{array}{rrcl} + : & V \times V & \to & V \qquad \textit{(Addition)} \\ & (v,w) & \mapsto & v+w \end{array}$$

und

$$\begin{array}{rcl} K \times V & \to & V \qquad \textit{(skalare Multiplikation)} \\ (a,v) & \mapsto & av \end{array}$$

so dass gilt:

i) $(V,+)$ *ist eine abelsche Gruppe.*

ii) $(ab)v = a(bv)$ *für alle* $a,b \in K$ *und* $v \in V$

iii) $1v = v$ *für alle* $v \in V$

iv) Für alle $a,b \in K$ *und* $v,w \in V$ *gelten die Distributivgesetze*

$$(a+b)v = av + bv$$

und

$$a(v+w) = av + aw.$$

Wir nennen die Elemente eines Vektorraums auch ***Vektoren.***

Beispiel:

i) $V = \{0\}$ mit $0 + 0 = 0$ und $a0 = 0$ für alle $a \in K$ ist ein Vektorraum über dem Körper K. Wir schreiben auch $V = 0$.

ii) Die Menge $K^n := K^{n\times 1} = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} : a_1, \ldots, a_n \in K \right\}$ der n-dimensionalen Spaltenvektoren mit Einträgen in K ist ein K-Vektorraum mit der in § 1 definierten Addition und skalaren Multiplikation:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

und

$$c \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} ca_1 \\ \vdots \\ ca_n \end{pmatrix}$$

Das Nullelement ist der Nullvektor $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, das inverse Element zu $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ ist $\begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix}$. Damit rechnet man leicht nach, dass $(K^n, +)$ eine abelsche Gruppe bildet. Die Rechenregeln ii), iii), iv) aus Definition 5.1 rechnet man ebenfalls mühelos nach.

Allgemeiner gilt, dass die Menge $K^{m\times n}$ der $m \times n$-Matrizen mit Einträgen in K ein K-Vektorraum ist ($\rightarrow$ Übungen).

iii) Ist $K \subset L$ ein Teilkörper von L, so wird L zusammen mit der Addition

$$+ : L \times L \rightarrow L$$

und der Einschränkung der Multiplikation

$$\cdot : K \times L \rightarrow L$$

ein K-Vektorraum. So ist etwa $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum.

iv) Jeder Körper K ist mit seiner Addition und Multiplikation ein K-Vektorraum. (Das ist ein Spezialfall von ii)).

Lemma 5.2 *In jedem K-Vektorraum V gilt:*

i) Ist 0_V das Nullelement von $(V,+)$, so ist $a0_V = 0_V$ für alle $a \in K$.

ii) Ist 0_K das Nullelement von $(K,+)$, so ist $0_K v = 0_V$ für alle $v \in V$.

iii) Für $a \in K$ und $v \in V$ gilt $(-a)v = a(-v) = -av$.

iv) Ist $av = 0_V$ für $a \in K$ und $v \in V$, so folgt $a = 0_K$ oder $v = 0_V$.

Beweis :

i) Es ist $a0_V + a0_V = a(0_V + 0_V) = a0_V$, nach Addieren von $-a0_V$ also $a0_V = 0$.

ii) Es gilt $0_K v + 0_K v = (0_K + 0_K)v = 0_K v, \Rightarrow 0_K v = 0_V$.

iii) Es gilt $av + (-a)v = (a + (-a))v = 0_K v \overset{\text{ii)}}{=} 0_V$, also ist $(-a)v = -av$.

Ferner ist $(-a)v = ((-1)a)v \overset{\text{5.1 ii)}}{=} a((-1)v)) = a(-v)$ nach dem gerade Gezeigten für $a = -1$.

iv) Angenommen, $av = 0_V$ und $a \neq 0$. Dann existiert a^{-1} im Körper K. Wir multiplizieren die Gleichung $0_V = av$ mit a^{-1} und erhalten $a^{-1}0_V = a^{-1}(av) = (a^{-1}a)v = v$. Nach i) ist $a^{-1}0_V = 0_V$, woraus $v = 0_V$ folgt.

□

In Zukunft werden wir in der Notation nicht mehr unterscheiden zwischen 0_K und 0_V; was jeweils gemeint ist, sollte aus dem Kontext hervorgehen.

Der Begriff des K-Vektorraums schließt den des gewählten Grundkörpers K mit ein. Es ist etwas anderes, ob wir $\mathbb{C}$ als $\mathbb{R}$-Vektorraum oder als $\mathbb{C}$-Vektorraum betrachten!

Definition 5.3 *Es sei V ein K-Vektorraum. Eine Teilmenge $U \subset V$ heißt **Untervektorraum** oder linearer Unterraum von V, falls gilt:*

i) $0 \in U$

ii) Für alle $u, v \in U$ ist $u + v \in U$

iii) Für alle $a \in K, u \in U$ ist $au \in U$.

Lemma 5.4 *Eine Teilmenge U eines K-Vektorraums V ist genau dann ein Untervektorraum von V, wenn U abgeschlossen unter $+$ und der skalaren Multiplikation ist und mit diesen Verknüpfungen selbst ein K-Vektorraum ist.*

Beweis : „ $\Rightarrow$“: Ist U ein Untervektorraum von V, so ist U abgeschlossen unter $+$ und skalarer Multiplikation nach Definition 5.3 ii) und iii). Da die Addition auf V kommutativ und assoziativ ist, gelten diese Regeln auch in U. Ferner ist $0 \in U$ und für jedes $u \in U$ auch $-u \overset{5.2}{=} (-1) \cdot u \in U$, da U abgeschlossen unter skalarer Multiplikation ist. Also ist $(U, +)$ eine abelsche Gruppe. Die Bedingungen ii), iii) und iv) aus Definition 5.1 gelten in V, also auch in $U \subset V$. Somit ist U ein Vektorraum.

„ $\Leftarrow$“ ist leicht. (Übungsaufgabe) □

Beispiel:

i) Jeder K-Vektorraum V enthält die Untervektorräume $\{0\}$ und V.

ii) Ist A eine $m \times n$-Matrix mit Einträgen in K, so ist die Menge der Lösungen $x \in K^n$ des linearen Gleichungssystems $Ax = 0$ ein Untervektorraum von K^n.

Definition 5.5 *Ein* ***Isomorphismus*** *ϕ von einem K-Vektorraum V ist eine bijektive Abbildung $\phi : V \to W$, für die gilt:*

i) $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$ *für alle* $v_1, v_2 \in V$

ii) $\phi(av) = a\phi(v)$ *für alle* $a \in K$ *und* $v \in V$.

Wir nennen dann V und W isomorph.

Beispiel:

i) Die Abbildung

$$\begin{aligned} \phi : \mathbb{R}^2 &\to \mathbb{C} \\ \begin{pmatrix} a \\ b \end{pmatrix} &\mapsto a + ib \end{aligned}$$

ist ein Isomorphismus von $\mathbb{R}$-Vektorräumen.

ii) Der K-Vektorraum der n-dimensionalen Zeilenvektoren $K^{1\times n}$ ist isomorph zum K-Vektorraum der n-dimensionalen Spaltenvektoren $K^n = K^{n\times 1}$ mit Hilfe des Isomorphismus

$$\begin{array}{rccl} \phi : & K^{1\times n} & \to & K^n \\ & (a_1, \ldots, a_n) & \mapsto & \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \end{array}$$

Wir wollen nun ein Verfahren zur Konstruktion von Untervektorräumen kennenlernen. Es sei K ein beliebiger Körper und V ein K-Vektorraum. Eine **Linearkombination** von Vektoren $v_1, \ldots, v_n \in V$ ist ein Vektor der Form

$$w = a_1 v_1 + \cdots + a_n v_n$$

für $a_1, \ldots, a_n \in K$.

Beispiel: Im $\mathbb{F}_3^3$ ist $w = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ eine Linearkombination von $v_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, denn es gilt

$$1v_1 + 2v_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Definition 5.6 *Für eine beliebige Teilmenge $\emptyset \neq M \subset V$ ist* ***die lineare Hülle*** *$\langle M \rangle$ von M definiert als*

$$\langle M \rangle = \left\{ \sum_{i=1}^{n} a_i v_i : n \in \mathbb{N}, a_i \in K, v_i \in M \right\}$$

$\langle M \rangle$ ist also die Menge aller Linearkombinationen, die man aus endlich vielen Elementen aus M bilden kann. Außerdem setzen wir $\langle \emptyset \rangle = 0$.

Beispiel:

i) Es ist $\langle 0 \rangle = 0$ und $\langle V \rangle = V$.

ii) Für jedes $v \in V$ ist $\langle \{v\} \rangle = \{av : a \in K\}$.

iii) Im K^3 gilt $\left\langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right\rangle = K^3$, da $\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

Satz 5.7 *Für jede Teilmenge $M \subset V$ ist $\langle M\rangle \subset V$ ein Untervektorraum. $\langle M\rangle$ ist der kleinste Untervektorraum, der M enthält, d.h.: Ist $H \subset V$ ein Untervektorraum mit $M \subset H$, so folgt $\langle M\rangle \subset H$.*

Beweis : Wir müssen zunächst die Eigenschaften i) - iii) aus Definition 5.3 prüfen. Ist $M = \emptyset$, so ist $\langle M\rangle = 0$ ein Untervektorraum von V. Wir können also $M \neq \emptyset$ annehmen. Dann gibt es ein $a \in M$, also liegt $0 = 0a \in \langle M\rangle$. Sind $\sum\limits_{i=1}^{n} a_i v_i$ und $\sum\limits_{j=1}^{m} b_j w_j \in \langle M\rangle$, so liegt auch

$$\sum_{i=1}^{n} a_i v_i + \sum_{j=1}^{m} b_j w_j$$

als Linearkombination der Vektoren $v_1, \ldots, v_n, w_1, \ldots, w_m$ in $\langle M\rangle$.

Außerdem ist für jedes $c \in K$ auch

$$c(\sum_{i=1}^{n} a_1 v_i) = \sum_{i=1}^{n} (ca_i) v_i$$

in $\langle M\rangle$. Also ist $\langle M\rangle$ ein Untervektorraum von V.

Ist H ein weiterer Unterraum mit $M \subset H$, so ist H abgeschlossen unter Addition und skalarer Multiplikation. Also sind alle Vektoren der Form $\sum\limits_{i=1}^{n} a_i v_i$ für $v_i \in M$ in H, d.h. $\langle M\rangle \subset H$. □

Manchmal ist es praktisch, mit endlichen geordneten Mengen von Vektoren zu arbeiten. Wir schreiben sie als $(v_1, \ldots, v_n)$. Die geordnete Menge (a, b) ist also verschieden von der geordneten Menge (b, a) und auch von (a, a, b), obwohl die gewöhnlichen (ungeordneten) Mengen $\{a, b\}, \{b, a\}, \{a, a, b\}$ gleich sind. Ist $(v_1, \ldots, v_n)$ eine geordnete Menge, so ist die lineare Hülle $\langle (v_1, \ldots, v_n)\rangle$ definiert als

$$\begin{aligned} \langle (v_1, \ldots, v_n)\rangle &= \langle \{v_1, \ldots, v_n\}\rangle \\ &= \left\{ \sum_{i=1}^{n} a_1 v_i : a_1, \ldots, a_n \in K \right\} \end{aligned}$$

Lemma 5.8 *Ist $M = (v_1, \ldots, v_n)$ eine endliche geordnete Menge von Vektoren in V und $v \in V$, so sei M' die geordnete Menge $M' = (v_1, \ldots, v_n, v)$. Dann gilt*

$$\langle M'\rangle = \langle M\rangle \Leftrightarrow v \in \langle M\rangle$$

Beweis : „ $\Rightarrow$ “: Offenbar liegt v in $\langle M' \rangle$. Also folgt aus $\langle M' \rangle = \langle M \rangle$ auch $v \in \langle M \rangle$.

„ $\Leftarrow$ “: Die Inklusion $\langle M \rangle \subset \langle M' \rangle$ ist klar.

Ist $v \in \langle M \rangle$, so ist $\langle M \rangle$ ein Untervektorraum, der alle Vektoren aus M' enthält. Nach Satz 5.7 folgt also $\langle M' \rangle \subset \langle M \rangle$. $\square$

Wir kommen nun zu einer sehr wichtigen Definition.

Definition 5.9

i) Eine endliche geordnete Menge $(v_1, \ldots, v_n)$ heißt ***linear abhängig****, wenn es $a_1, \ldots, a_n \in K$ gibt, die nicht alle Null sind, so dass*

$$a_1 v_1 + \cdots + a_n v_n = 0$$

gilt.

ii) Die Menge $(v_1, \ldots, v_n)$ heißt ***linear unabhängig****, wenn sie nicht linear abhängig ist. Mit anderen Worten: $(v_1, \ldots, v_n)$ ist genau dann linear unabhängig, wenn aus jeder Gleichung der Form $a_1 v_1 + \cdots + a_n v_n = 0$ mit $a_1 \ldots, a_n \in K$ bereits $a_1 = a_2 = \cdots = a_n = 0$ folgt.*

Wir nennen die Vektoren $v_1, \ldots, v_n$ linear abhängig bzw. unabhängig, wenn die geordnete Menge $(v_1, \ldots, v_n)$ linear abhängig bzw. unabhängig ist. Dieser Begriff hängt nicht von der Reihenfolge der v_i ab: Ist $(v_1, \ldots, v_n)$ linear abhängig bzw. unabhängig, so gilt dies auch für jede Umordnung dieser Menge.

Beispiel:

i) Die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$ und $\begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$ sind linear abhängig in $\mathbb{Q}^3$, denn

$$2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} = 0$$

ii) Ist $M = (v_1, \ldots, v_n)$ eine geordnete Menge mit $v_1 = 0$, so ist $1 v_1 + \sum\limits_{j \neq 1} 0 v_j = 0$, also ist M linear abhängig.

iii) Ist $M = (v_1, \ldots, v_n)$ eine geordnete Menge mit $v_i = v_j$ für $i \neq j$, so ist

$$1v_i + (-1)v_j = 0,$$

also ist M linear abhängig.

iv) Die Vektoren $\binom{2}{1}$ und $\binom{-1}{1}$ im $\mathbb{R}^2$ sind linear unabhängig, denn aus

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

folgt $2a - b = 0$ und $a + b = 0$, also $a = b = 0$.

Lemma 5.10 *Es seien $v_1, \ldots, v_n \in K^m$ und $A \in K^{m \times n}$ die Matrix mit den Spalten $v_1, \ldots, v_n$. Dann sind $v_1, \ldots, v_n$ genau dann linear abhängig, wenn das homogene lineare Gleichungssystem $Ax = 0$ eine nicht-triviale Lösung $x \neq 0$ besitzt.*

Beweis : Sind $v_1, \ldots, v_n$ linear abhängig, so gilt $a_1v_1 + \cdots + a_nv_n = 0$ für $a_1, \ldots, a_n \in K$, die nicht alle 0 sind. Also ist $x = \begin{pmatrix}a_1\\ \vdots \\ a_n\end{pmatrix} \neq 0$ eine nicht-triviale Lösung von $Ax = 0$. Ist umgekehrt $x = \begin{pmatrix}a_1\\ \vdots \\ a_n\end{pmatrix} \neq 0$ eine nicht-triviale Lösung von $Ax = 0$, so folgt $a_1v_1 + \cdots + a_nv_n = 0$. Da nicht alle a_1 Null sind, sind $v_1, \ldots, v_n$ linear abhängig. □

Da wir lineare Gleichungssysteme mit elementaren Zeilenumformungen lösen können (siehe § 2), gibt uns Lemma 5.10 ein Verfahren an die Hand, gegebene Vektoren auf lineare Abhängigkeit zu prüfen.

Proposition 5.11 *Es sei $M = (v_1, \ldots, v_n)$ eine linear unabhängige Menge in V und $v \in V$. Dann ist die Menge $M' = (v_1, \ldots, v_n, v)$ genau dann linear unabhängig, wenn $v \notin \langle M \rangle$ gilt.*

Beweis : Ist $v \in \langle M \rangle$, so gibt es $a_1, \ldots, a_n \in K$ mit $v = a_1v_1 + \cdots + a_nv_n$. Also ist

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n + (-1)v = 0$$

eine Linearkombination der 0. Da $-1 \neq 0$ als Koeffizient auftaucht, ist

$M' = \langle v_1, \ldots, v_n, v \rangle$ linear abhängig.

Umgekehrt nehmen wir an, dass M' linear abhängig ist. Dann gibt es $a_1, \ldots, a_n, a_{n+1}$ in K, die nicht alle Null sind, mit $a_1v_1 + \cdots + a_nv_n + a_{n+1}v' = 0$. Ist $a_{n+1} = 0$, so ist eines der $a_1, \ldots, a_n$ nicht 0 und $\sum\limits_{i=1}^{n} a_iv_i$ eine Linearkombination der 0. Das widerspricht der Tatsache, dass M linear unabhängig ist.

Also ist $a_{n+1} \neq 0$ und

$$v' = \frac{a_1}{a_{n+1}}v_1 + \cdots + \frac{a_n}{a_{n+1}}v_n$$

liegt in $\langle M \rangle$. □

Definition 5.12 *i) Ein K-Vektorraum V heißt **endlich erzeugt**, falls es eine endliche Menge M mit $\langle M \rangle = V$ gibt.*

*ii) Sei V ein endlich erzeugter K-Vektorraum. Eine endliche Menge $B = (v_1, \ldots, v_n)$ von Vektoren in V heißt **Basis** von V, falls B linear unabhängig ist und $V = \langle B \rangle$ gilt.*

Wir nehmen ab jetzt an, V sei ein endlich erzeugter K-Vektorraum und untersuchen Basen in V.

Satz 5.13 *Sei V ein endlich erzeugter K-Vektorraum. $B = (v_1, \ldots, v_n)$ ist genau dann eine Basis von V, wenn jedes $v \in V$ auf eindeutige Weise als Linearkombination der v_i geschrieben werden kann, d.h., wenn es für jedes $v \in V$ eindeutig bestimmte $a_1, \ldots, a_n \in K$ mit $v = a_1v_1 + \cdots + a_nv_n$ gibt.*

Beweis : Ist B eine Basis, so ist $V = \langle B \rangle$, also lässt sich jedes $v \in V$ als $v = a_1v_1 + \cdots + a_nv_n$ mit $a_i \in K$ schreiben. Wir müssen die Eindeutigkeit dieser Darstellung zeigen Gilt auch $v = b_1v_1 + \cdots + b_nv_n$ mit $b_i \in K$, so folgt $(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0$. Da $(v_1, \ldots, v_n)$ linear unabhängig sind, ist $a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$.

Lässt sich umgekehrt jedes $v \in V$ auf eindeutige Weise als Linearkombination der v_i schreiben, so gilt $V = \langle B \rangle$. Angenommen $a_1v_1 + \cdots + a_nv_n = 0$ ist eine Linearkombination der 0. Da $0 = 0a_1 + \cdots + 0a_n$ ist, folgt aus der Eindeutigkeit der Darstellung von 0, dass $a_1 = \cdots = a_n = 0$ ist. Somit ist B linear unabhängig, d.h. eine Basis von V. □

Beispiel:

i) Sei $V = K^n$ der Vektorraum der n-dimensionalen Spaltenvektoren und für alle $i = 1, \dots, n$, sei

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix} \qquad \in K^n.$$

Dann ist $e_1, \dots, e_n$ eine Basis von K^n. Jedes $v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n$ lässt sich nämlich auf eindeutige Weise als Linearkombination $v = a_1 e_1 + \cdots + a_n e_n$ der e_i schreiben.

ii) Die Vektoren $\binom{2}{1}, \binom{-1}{1}$ bilden eine Basis des $\mathbb{R}^2$. Wir haben oben gesehen, dass sie linear unabhängig sind. Die Matrix $A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ hat Determinante 3, ist also invertierbar in $\mathbb{R}^{2\times 2}$. Für jeden Vektor $\binom{x}{y} \in \mathbb{R}^2$ gibt es also einen Vektor $\binom{a}{b} \in \mathbb{R}^2$, nämlich $\binom{a}{b} = A^{-1}\binom{x}{y}$, mit $A\binom{a}{b} = \binom{x}{y}$. Also gilt $a\binom{2}{1} + b\binom{-1}{1} = \binom{x}{y}$, woraus $\langle \binom{2}{1}, \binom{-1}{1} \rangle = \mathbb{R}^2$ folgt. Also ist $\left(\binom{2}{1}, \binom{-1}{1}\right)$ auch ein Erzeugendensystem des $\mathbb{R}^2$.

Jetzt folgt ein wichtiger Satz:

Satz 5.14 *Es sei $V \neq 0$. Jedes endliche Erzeugendensystem von V enthält eine Basis. Insbesondere besitzt jeder endlich erzeugte Vektorraum eine Basis.*

Beweis : Es sei $M = (v_1, \dots, v_n)$ ein Erzeugendensystem von V. Ist M linear unabhängig, so ist M selbst eine Basis. Wir nehmen also an, dass M linear abhängig ist. Also gibt es $a_1, \dots, a_n \in K$, nicht alle Null, mit $a_1 v_1 + \cdots + a_n v_n = 0$. Ist a_i ein Koeffizient $\neq 0$, so folgt

$$v_i = -\sum_{j \neq i} \frac{a_j}{a_i} v_j,$$

d.h. für $\tilde{M} = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$ gilt $v_i \in \langle \tilde{M} \rangle$. Nach Lemma 5.8 folgt

$$\langle \tilde{M} \rangle = \langle M \rangle = V.$$

Wir haben also ein Erzeugendensystem von V mit $n-1$ Elementen gefunden.

Ist $\tilde{M}$ nicht linear unabhängig, so wenden wir dasselbe Verfahren erneut an, um einen Vektor zu streichen. Dies setzen wir so lange fort, bis wir an ein linear unabhängiges Erzeugendensystem von V, also eine Basis von V, gelangen. □

Der Beweis von Satz 5.14 gibt uns ein Verfahren an die Hand, wie wir aus gegebenen Vektoren $v_1, \ldots, v_n$, die ein Erzeugendensystem von V sind, eine Basis von V konstruieren können.

Was ist mit $V = 0$? Der Vektor 0 ist linear abhängig, so dass (0) keine Basis sein kann. Wir verabreden, dass die leere Menge linear unabhängig ist. Da $\langle \emptyset \rangle = 0$ gilt, ist also $\emptyset$ eine Basis des Nullraumes.

Satz 5.15 *Es sei V ein endlich erzeugter K-Vektorraum. Jede linear unabhängige Menge $(v_1, \ldots, v_r)$ in V kann zu einer Basis $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ von V ergänzt werden.*

Beweis : Es sei $M = (w_1, \ldots, w_m)$ eine endliche Menge mit $V = \langle M \rangle$.

Sind alle w_i in $\langle (v_1, \ldots, v_r) \rangle$ enthalten, so ist $V = \langle M \rangle \subset \langle (v_1, \ldots, v_r) \rangle \subset V$, also ist $(v_1, \ldots, v_r)$ ein Erzeugendensystem und damit eine Basis von V.

Ist das nicht der Fall, so wählen wir ein w_i mit $w_i \notin \langle (v_1, \ldots, v_n) \rangle$. Nach Proposition 5.11 ist dann $(v_1, \ldots, v_n, w_i)$ linear unabhängig. Wir wiederholen dieses Vorgehen so lange, bis wir ein linear unabhängiges Erzeugendensystem, d.h. eine Basis von V konstruiert haben. □

In Beweis von Satz 5.14 haben wir schon gesehen, dass man aus einem Erzeugendensystem Vektoren auswählen kann, die eine Basis bilden. Satz 5.15 sagt allgemeiner, dass wir durch geschicktes Auswählen von Vektoren aus einem Erzeugendensystem sogar vorgegebene linear unabhängige Vektoren zu einer Basis ergänzen können.

Wir wissen, dass jeder endlich erzeugte Vektorraum eine endliche Basis hat. Wir wollen jetzt untersuchen, was alle Basen gemeinsam haben.

Satz 5.16 *Es seien M und L endliche Teilmengen von V. Es sei $\langle M \rangle = V$ und L sei linear unabhängig. Dann gilt $|M| \geq |L|$, wobei $|M|$ bzw.$|L|$ die Anzahl der Elemente in M bzw. L bezeichnet.*

Beweis : Es sei $M = (v_1, \dots, v_m)$ und $L = (w_1, \dots, w_n)$, d.h. es gilt $|M| = m$ und $|L| = n$. Da $\langle M \rangle = V$ ist, ist jedes w_j Linearkombination von $v_1, \dots, v_n$, d.h. es gibt $a_{1j}, \dots, a_{mj} \in K$ mit $w_j = a_{1j}v_1 + \cdots + a_{mj}v_m$. Es sei A die Matrix $(a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$

Für alle $c_1, \dots, c_n \in K^n$ gilt:

$$\sum_{j=1}^{n} c_j w_j = 0 \Leftrightarrow \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{m} a_{ij} v_i \right) = 0 \Leftrightarrow \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) v_i = 0.$$

Jede Lösung $x = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ des homogenen linearen Gleichungssystems $Ax = 0$ liefert also eine Gleichung der Form

$$c_1 w_1 + \cdots + c_n w_n = 0,$$

d.h. eine Linearkombination des Nullvektors durch L. Da L linear unabhängig ist, besitzt $Ax = 0$ nur die triviale Lösung. Aus Korollar 2.8 folgt daher, dass $Ax = 0$ nicht mehr Unbestimmte als Gleichungen haben darf. Also ist $m \geq n$, d.h. $|M| \geq |L|$. $\square$

Beispiel: Die Vektoren $e_1, \dots, e_n$ bilden eine Basis des K^n, insbesondere also ein Erzeugendensystem. Der K^n enthält also keine linear unabhängige Teilmenge mit $(n + 1)$ oder mehr Elementen.

Satz 5.17 *Es sei V ein endlich erzeugter Vektorraum und $M \subset V$ eine Teilmenge mit $\langle M \rangle = V$. Dann gibt es eine endliche Teilmenge $N \subset M$ mit $\langle N \rangle = V$.*

Beweis : Nach Voraussetzung gibt es eine endliche Teilmenge $\{v_1, \dots, v_n\} \subset V$ mit $\langle \{v_1, \dots, v_n\} \rangle = V$. Da $V = \langle M \rangle$ ist, liegt jedes $v_i \in \langle M \rangle$, ist also Linearkombination endlich vieler Elemente aus M. Wir sammeln alle Elemente aus M ein, die in der Linearkombination eines der v_i auftreten. Das gibt eine endliche Teilmenge $N \subset M$, für die $v_i \in \langle N \rangle$ gilt. Nach Satz 5.7 folgt $V = \langle \{v_1, \dots, v_n\} \rangle \subset \langle N \rangle$, woraus $\langle N \rangle = V$ folgt. $\square$

Wir wollen jetzt noch zeigen, dass alle linear unabhängigen Teilmengen von V endlich sind. Dazu müssen wir erst einmal definieren, was linear unabhängig für eine beliebige Teilmenge von V heißt.

Definition 5.18 *i) Eine beliebige Teilmenge $L \subset V$ heißt* ***linear unabhängig****, falls jede endliche Teilmenge von L linear unabhängig im Sinne von Definition 5.9 ist.*

ii) $L \subset V$ heißt ***linear abhängig****, falls L nicht linear unabhängig ist, d.h. falls es eine endliche Teilmenge von L gibt, die linear abhängig im Sinne von Definition 5.9 ist.*

iii) Eine beliebige Teilmenge $B \subset V$ heißt ***Basis*** *von V, falls B linear unabhängig ist und $\langle B \rangle = V$ gilt.*

Satz 5.19 *Es sei V ein endlich erzeugter K-Vektorraum.*

i) Jedes Erzeugendensystem $M \subset V$ enthält eine endliche Basis.

ii) Jede linear unabhängige Teilmenge $L \subset V$ ist endlich und lässt sich zu einer endlichen Basis von V ergänzen.

iii) Jede Basis von V ist endlich.

Beweis :

i) Nach Satz 5.17 gibt es eine endliche Teilmenge $N \subset M$ mit $\langle N \rangle = V$. Diese enthält nach Satz 5.14 eine endliche Basis.

ii) Es sei $L \subset V$ eine linear unabhängige Teilmenge und $M = \{v_1, \dots, v_n\}$ ein endliches Erzeugendensystem von V. Ist $L_1 \subset L$ eine beliebige endliche Teilmenge, so ist L_1 linear unabhängig. Aus Satz 5.16 folgt also $|L_1| \leq n$. Somit muss L eine endliche Teilmenge von V mit $|L| \leq n$ sein. Nach Satz 5.15 kann L zu einer endlichen Basis von V ergänzt werden.

iii) Da eine Basis insbesondere linear unabhängig ist, folgt das aus ii).

□

Wir wissen jetzt, dass jede Basis eines endlich erzeugten Vektorraums aus endlich vielen Elementen besteht. Jetzt wollen wir noch zeigen, dass die Anzahl der Elemente für jede Basis dieselbe ist.

Satz 5.20 *Sind B_1 und B_2 zwei Basen eines endlich erzeugten K-Vektorraums, so ist $|B_1 = |B_2|$, d.h. B_1 und B_2 haben gleich viele Elemente.*

Beweis : Nach Satz 5.19 sind B_1 und B_2 endlich. Wir können also Satz 5.16 auf $M = B_1$ und $L = B_2$ anwenden und erhalten $|B_1| \geq |B_2|$. Nochmalige Anwendung von Satz 5.16 auf $M = B_2$ und $L = B_1$ liefert $|B_2| \geq |B_1|$. □

Definition 5.21 *Die* ***Dimension*** *eines endlich erzeugten K-Vektorraums ist definiert als die Anzahl der Elemente einer Basis. Wir bezeichnen sie mit $\dim_K V$ oder auch mit $\dim V$.*

Nach Satz 5.20 hängt diese Definition nicht von der Wahl einer Basis von V ab.

Beispiel:

i) Ist $V = 0$, so ist $\dim_K V = 0$.

ii) Es ist $\dim_K K^n = n$ für alle $n \in \mathbb{N}$.

Wenn wir in Zukunft von einem Vektorraum V der Dimension n reden, so meinen wir, dass V eine Basis aus n Elementen besitzt. V ist also automatisch endlich erzeugt.

Satz 5.22 *Sei V ein endlich erzeugter K-Vektorraum.*

i) Ist $M \subset V$ mit $\langle M \rangle = V$, so folgt $|M| \geq \dim V$. Es gilt $|M| = \dim V$ genau dann, wenn M eine Basis von V ist.

ii) Ist $L \subset V$ eine linear unabhängige Teilmenge, so gilt $|L| \leq \dim V$. Es gilt $|L| = \dim V$ genau dann, wenn L eine Basis von V ist.

Beweis : Wir wissen aus Satz 5.14, dass V eine Basis B besitzt. Definitionsgemäß ist $\dim V = |B|$.

i) Ist $\langle M \rangle = V$, so folgt aus Satz 5.19 $|M| \geq \dim V$. Ist M sogar eine Basis von V, so gilt natürlich $\dim V = |M|$. Sei umgekehrt $|M| = \dim V$. Dann finden wir nach Satz 5.14 eine Basis $B' \subset M$ von V. Da $|B'| = \dim V = |M|$ gilt, folgt $B' = M$, d.h. M ist eine Basis von V.

ii) Ist $L \subset V$ linear unabhängig, so folgt aus Satz 5.19 $|L| \leq \dim V$. Ist L sogar eine Basis, so gilt definitionsgemäß $|L| = \dim V$. Sei umgekehrt $|L| = \dim V$. Dann können wir L nach Satz 5.15 zu einer Basis B' von V ergänzen. Da $|B'| = \dim V = |L|$ gilt, folgt $L = B'$, d.h. L ist eine Basis von V.

□

Satz 5.23 *Es sei V ein endlich erzeugter K-Vektorraum und $W \subset V$ ein Unterraum. Dann ist auch W endlich erzeugt, und es gilt $\dim W \leq \dim V$. Ferner ist genau dann $\dim W = \dim V$, wenn $W = V$ gilt.*

Beweis : Ohne Einschränkung ist $W \neq 0$. Wir überlegen uns zuerst, dass W endlich erzeugt ist. Wir starten mit einem beliebigen $0 \neq w_1 \in W$. Ist $\langle w_1 \rangle = W$, so ist W endlich erzeugt. Ansonsten gibt es ein $w_2 \in W$ mit $w_2 \notin \langle w_1 \rangle$. Nach Proposition 5.11 ist dann (w_1, w_2) linear unabhängig in V. Ist $\langle (w_1, w_2) \rangle = W$, so sind wir fertig. Ansonsten nehmen wir ein $w_3 \in W \setminus \langle (w_1, w_2) \rangle$ hinzu. Induktiv konstruieren wir so Systeme $(w_1, \ldots, w_k)$ von Vektoren in W, die linear unabhängig in V sind. Da in V nach Satz 5.22 jede $(\dim V + 1)$-elementige Teilmenge linear abhängig ist, bricht das Verfahren nach spätestens $\dim V$ Schritten ab und liefert ein endliches Erzeugendensystem von W.

Weil W endlich erzeugt ist, besitzt W eine Basis B mit $|B| = \dim W$. Da B linear unabhängig in V ist, folgt aus Satz 5.16 $\dim W = |B| \leq \dim V$. Ist $\dim W = \dim V$, dann ist B nach Satz 5.22 auch eine Basis von V, insbesondere ist $\langle B \rangle = W = V$. Aus $W = V$ folgt ferner trivialerweise $\dim W = \dim V$. □

Wir wollen nun noch kurz auf Vektorräume eingehen, die nicht endlich erzeugt sind.

Satz 5.24 *Ein K-Vektorraum V ist genau dann nicht endlich erzeugt, wenn er eine unendliche linear unabhängige Teilmenge L besitzt.*

Beweis : Angenommen, V besitzt eine unendliche linear unabhängige Teilmenge. Nach Satz 5.22 kann V dann nicht endlich erzeugt sein. Umgekehrt nehmen wir an, V sei nicht endlich erzeugt. Dann ist $V \neq 0$, d.h. es gibt ein $v_1 \neq 0$ in V. Die Menge (v_1) ist linear unabhängig. Da V nicht endlich erzeugt ist, gilt $V \neq \langle v_1 \rangle$, also gibt es ein $v_2 \in V \setminus \langle v_1 \rangle$. Nach Proposition 5.11 ist (v_1, v_2) linear unabhängig. Auf diese Weise konstruieren wir induktiv für alle n ein linear unabhängiges System $(v_1, \ldots, v_n)$ in V. Die Menge $\{v_k : k \in \mathbb{N}\}$ ist dann eine unendliche linear unabhängige Teilmenge von V. □

Auch Vektorräume, die nicht endlich erzeugt sind, besitzen eine Basis, wie wir jetzt zeigen wollen. Diese muss aus unendlich vielen Elementen bestehen. Wir benötigen hierfür das Zorn'sche Lemma.

Das Zorn'sche Lemma ist logisch äquivalent zum sogenannten Auswahlaxiom, das Teil der Mengenaxiome ist. Wir werden es also nicht beweisen. Zunächst erklären wir einige Begriffe.

Definition 5.25 *i) Es sei X eine Menge. Eine Relation auf X ist eine Teilmenge $R \subset X \times X$. Die Menge R besteht also aus gewissen Paaren (x, x') für $x \in X, x' \in X$.*

ii) Eine Relation R auf X heißt partielle Ordnung, und wir schreiben auch $x \leq x'$ statt $(x, x') \in R$, wenn folgende Bedingungen gelten:

a) $x \leq x$ für alle $x \in X$

b) Aus $x \leq x'$ und $x' \leq x''$ folgt $x \leq x''$.

c) Aus $x \leq x'$ und $x' \leq x$ folgt $x = x'$.

Eine partielle Ordnung heißt Totalordnung, falls zusätzlich gilt:

d) Für alle $x, x' \in X$ ist $x \leq x'$ oder $x' \leq x$.

In einer partiellen Ordnung müssen zwei Elemente also nicht unbedingt vergleichbar sein. So definiert zum Beispiel die Relation

$$A \leq B \text{ genau dann, wenn } A \subset B \text{ ist,}$$

eine partielle Ordnung auf der Menge aller Teilmengen von $\mathbb{R}$, aber keine Totalordnung. (Überlegen Sie sich das!) Die gewöhnliche Ordnung $\leq$ auf $\mathbb{R}$ ist hingegen eine Totalordnung.

Ist A eine Teilmenge einer Menge X und $\leq$ eine partielle Ordnung auf X, so heißt ein $s \in X$ obere Schranke von A, falls für alle $a \in A$ gilt: $a \leq s$.

Es muss also jedes Element von A

i) mit s vergleichbar und

ii) $\leq s$ sein.

Achtung: s muss nicht in A liegen.

Ein $m \in X$ heißt maximales Element von X, wenn für alle $x \in X$ aus $m \leq x$ schon $m = x$ folgt. Das bedeutet: Sind m und x vergleichbar, so ist $x \leq m$. Es kann aber natürlich Elemente x geben, die nicht mit m vergleichbar sind. Hat X ein maximales Element m, so muss m daher keine obere Schranke von X sein. Betrachtet man etwa die Menge X der Teilmengen von $\{1, \ldots, n\}$, die echt kleiner als $\{1, \ldots, n\}$ sind, versehen mit der partiellen Ordnung $A \leq B$ genau dann, wenn $A \subset B$ ist, so besitzt X die maximalen Elemente $\{2, 3, \ldots, n\}, \{1, 3, 4, \ldots, n\}, \{1, 2, 4, \ldots, n\}$ etc., aber keine obere Schranke (Prüfen Sie das!).

Jetzt können wir das Zorn'sche Lemma formulieren:

Satz 5.26 (Zorn'sches Lemma) *Es sei $(X, \leq)$ eine partiell geordnete Menge. Besitzt jede Teilmenge $A \subset X$, die bezüglich der Einschränkung von $\leq$ auf A total geordnet ist, eine obere Schranke in M, so hat X ein maximales Element.*

Satz 5.27 *Sei V ein beliebiger K-Vektorraum. Dann besitzt V eine Basis.*

Beweis : Wir wissen, dass in einem endlich erzeugten Vektorraum jede Basis B eine maximale linear unabhängige Teilmenge ist, das heißt, gilt $B \subset L$ für eine linear unabhängige Menge L, so folgt $B = L$. In einem beliebigen Vektorraum wollen wir ebenfalls eine Basis als „maximale linear unabhängige Teilmenge" definieren. Mit dem Zorn'schen Lemma zeigen wir, dass das geht.

Dazu sei M die Menge aller linear unabhängigen Teilmengen von V, d.h.

$$M = \{L \subset V : L \text{ linear unabhängig }\}.$$

M ist partiell geordnet bezüglich der Inklusion von Teilmengen, d.h. wir definieren

$$L_1 \leq L_2 \text{ genau dann, wenn } L_1 \subset L_2.$$

Es sei nun $\Sigma \subset M$ eine total geordnete Teilmenge. Für zwei beliebige linear unabhängige Teilmengen L_1, L_2 von V mit $L_1 \in \Sigma$ und $L_2 \in \Sigma$ gilt also $L_1 \subset L_2$ oder $L_2 \subset L_1$. Dann definieren wir S als Vereinigung aller $L \in \Sigma$, d.h.

$$S = \bigcup_{L \in \Sigma} L = \{v \in V : v \in L \text{ für ein } L \in \Sigma\}.$$

Wir behaupten, dass S eine linear unabhängige Teilmenge von V ist. Dazu müssen wir zeigen, dass jede endliche Teilmenge $\{v_1, \ldots, v_r\}$ von S linear unabhängig sind. Nach Konstruktion von S gibt es Mengen $L_1, \ldots, L_r \in \Sigma$ mit $v_1 \in L_1, \ldots, v_r \in L_r$. Wir zeigen nun, dass eine der Mengen L_i alle anderen enthält. Da Σ total geordnet ist, folgt aus $L_j \not\subset L_i$ bereits $L_i \subsetneq L_j$. Angenommen, keine der Mengen L_i enthält alle anderen, dann können wir also induktiv eine Kette $L_1 \subsetneq L_{i_1} \subsetneq L_{i_2} \subsetneq \cdots \subsetneq L_{i_k} \subsetneq \cdots$ definieren, wobei alle $L_{i_k} \in \{L_1, \ldots, L_r\}$ sind. An einer Stelle dieser Kette müsste sich dann ein Element aus $\{L_1, \ldots, L_r\}$ wiederholen, was zu einem Widerspruch führt. Also gibt es ein i mit $L_j \subset L_i$ für alle $j = 1, \ldots, r$. Dann gilt $\{v_1, \ldots, v_r\} \subset L_i$. Da L_i linear unabhängig ist, ist auch $\{v_1, \ldots, v_r\}$ linear unabhängig. Also haben wir gezeigt, dass S linear unabhängig ist. Das Element $S \in M$ ist somit eine obere Schranke von Σ.

Nach dem Zorn'schen Lemma besitzt M nun ein maximales Element, d.h. es gibt eine linear unabhängige Teilmenge $L \subset V$, so dass für jede linear unabhängige Teilmenge L' mit $L \subset L'$ schon $L = L'$ folgt.

Wir zeigen nun, dass dieses L eine Basis von V ist. Dazu bleibt zu zeigen, dass $< L >= V$ ist.

Angenommen, $v \in V$ ist ein Element mit $v \notin < L >$. Dann ist nach Proposition 5.11 auch $L \cup \{v\}$ linear unabhängig. Das widerspricht aber der Maximalität von L'. Also gilt $V =< L >$, d.h. L ist in der Tat eine Basis von V. $\square$

Der Beweis von Satz 5.27 beweist zwar die Existenz einer Basis, er gibt aber keine Information darüber, wie eine solche Basis aussieht!

Typische Beispiele für Vektorräume, die nicht endlich erzeugt sind, sind Folgen- oder Funktionenräume.

Beispiel: Sei $\mathbb{R}^\infty = \{(a_k)_{k \in \mathbb{N}_o} : a_k \in \mathbb{R}\}$ die Menge aller reellen Folgen. Wir definieren $(a_k)_{k\in\mathbb{N}_0} + (b_k)_{k\in\mathbb{N}_0} = (a_k + b_k)_{k\in\mathbb{N}_0}$ sowie für $c \in K$ eine skalare Multiplikation $c(a_k)_{k\in\mathbb{N}_0} = (ca_k)_{k\in\mathbb{N}_0}$. Mit diesen Verknüpfungen wird $\mathbb{R}^\infty$ ein $\mathbb{R}$-Vektorraum.

Es sei für alle $i \in \mathbb{N}_0$ $e^{(i)}$ die Folge

$$e^{(i)} = (0, \dots, 0, 1, 0, 0, \dots) \text{ mit 1 an der } i\text{-ten Stelle},$$

d.h. es ist $e_i^{(i)} = 1$ und $e_k^{(i)} = 0$ für $k \neq i$.

Dann ist die unendliche Menge

$$M = \{e^{(i)} : i \in \mathbb{N}_0\}$$

linear unabhängig in $\mathbb{R}^\infty$.

Ist nämlich $\{e^{(i_1)}, \dots, e^{(i_n)}\} \subset M$ eine endliche Teilmenge und gilt

$$a_1 e^{(i_1)} + \dots + a_n e^{(i_n)} = 0$$

für $a_1, \dots, a_n \in K$, so gilt für die einzelnen Folgenglieder, d.h. für alle $k \in \mathbb{N}_0$:

$$a_1 e_k^{(i_1)} + \dots + a_n e_k^{(i_n)} = 0.$$

Ist $k = i_1$, so gilt also $0 = a_1 e_{i_1}^{(i_1)} + a_2 e_{i_1}^{(i_2)} + \dots + a_n e_{i_1}^{(i_n)} = a_1$. Genauso zeigt man $a_2 = \dots = a_n = 0$.

Da $\mathbb{R}^\infty$ die unendliche, linear unabhängige Teilmenge M besitzt, kann $\mathbb{R}^\infty$ nicht endlich erzeugt sein.

Wir kommen zurück zu endlich-dimensionalen Vektorräumen. Wir werden nun sehen, dass Basen nützlich sind, um in beliebigen Vektorräumen mit Koordinaten zu rechnen.

Definition 5.28 *Es sei V ein endlich erzeugter K-Vektorraum und $B = (v_1, \ldots, v_n)$ eine Basis von V. Dann nennen wir für jedes $v \in V$ den Vektor*

$$\kappa_B(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n$$

mit $v = a_1 v_1 + \cdots + a_n v_n$ den ***Koordinatenvektor*** *von v bezüglich B. (Nach Satz 5.13 sind die $a_1, \ldots, a_n$ eindeutig bestimmt).*

Ist $V = K^n$, so bezeichnen wir für eine Basis $B = (v_1, \ldots, v_n)$ des K^n mit $[B]$ die $n \times n$-Matrix mit den Spalten $v_1, \ldots, v_n$.

Beispiel: $V = \mathbb{R}^2, B = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right)$. Dann ist $[B] = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$.

Man kann Koordinatenvektoren im K^n folgendermaßen berechnen:

Proposition 5.29 *Sei $B = (v_1, \ldots, v_n)$ eine Basis von K^n. Dann ist $[B]$ invertierbar und der Koordinatenvektor von $v \in K^n$ bezüglich B ist*

$$\kappa_B(v) = [B]^{-1} v$$

Beweis : Da die Spalten von $[B]$ linear unabhängig sind, ist $[B]$ invertierbar, (Übungsaufgabe). Definitionsgemäß ist $v = a_1 v_1 + \cdots + a_n v_n$ und $\kappa_B(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Also gilt $[B]\kappa_B(v) = a_1 v_1 + \cdots + a_n v_n = v$, woraus $\kappa_B(v) = [B]^{-1} v$ folgt. $\square$

In einem beliebigen K-Vektorraum V sind die Koordinatenvektoren nützlich, um V mit dem Vektorraum K^n zu identifizieren:

Satz 5.30 *Es sei V ein n-dimensionaler Vektorraum und B eine Basis von V. Dann ist*

$$\begin{aligned} \phi: \quad V &\to \quad K^n \\ v \quad &\mapsto \quad \kappa_B(v) \end{aligned}$$

ein Isomorphismus von K-Vektorräumen im Sinne von Definition 5.5.

Beweis : Offenbar ist $\kappa_B(v_1 + v_2) = \kappa_B(v_1) + \kappa_B(v_2)$ und $\kappa_B(av) = a\kappa_B(v)$ (Übungsaufgabe). Wir müssen also nur zeigen, dass ϕ bijektiv ist. Dazu definieren wir für $B = (v_1, \dots, v_n)$ die Abbildung $\psi : K^n \to V$ durch $\psi \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 v_1 + \dots + a_n v_n \in V$.

Dann ist $\kappa_B \left(\psi \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \right) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ und $\psi(\kappa_B(v)) = v$, d.h. ψ ist eine Umkehrabbildung von ϕ. Also ist ϕ bijektiv. $\square$

Hat man also eine Basis eines n-dimensionalen K-Vektorraums V gewählt, so kann man V mit Hilfe der Abbildung ϕ mit dem Vektorraum K^n identifizieren. Diese Identifizierung hängt allerdings von der Wahl der Basis ab. Es ist oft wichtig, dass man je nach Problem eine geeignete Basis wählt. Wir wollen daher jetzt untersuchen, was passiert, wenn man mit einer Basis von V zu einer anderen Basis übergeht.

Definition 5.31 *Es seien $B = (v_1, \dots, v_n)$ und $B' = (v'_1, \dots, v'_n)$ Basen von V. Die **Übergangsmatrix** $P = P^B_{B'}$ von B nach B' ist die Matrix mit den Spalten $\kappa_{B'}(v_1), \dots, \kappa_{B'}(v_n)$. Die i-te Spalte von P ist also der Koordinatenvektor von v_i bezüglich der Basis B'. Definitionsgemäß gilt für $P = (p_{ij})_{i,j}$:*

$$\sum_{i=1}^{n} p_{ij} v'_i = v_j.$$

Lemma 5.32 *Es seien B und B' Basen des endlich-dimensionalen Vektorraums V und $P = P^B_{B'}$ die Übergangsmatrix von B nach B'. Dann gilt für alle $v \in V$*

$$P\kappa_B(v) = \kappa_{B'}(v).$$

Beweis : Ist $B = (v_1, \dots, v_n)$ und $B' = (v'_1, \dots, v'_n)$, so schreiben wir $v = a_1 v_1 + \dots + a_n v_n$. Dann ist definitionsgemäß

$$\kappa_B(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Also gilt

$$\begin{aligned} P\kappa_B(v) &= P(a_1e_1 + \cdots + a_ne_n) \\ &= a_1Pe_1 + \cdots + a_nPe_n \quad (Pe_i = i\text{-te Spalte von } P) \\ &= a_1\kappa_{B'}(v_1) + \cdots + a_n\kappa_{B'}(v_n) \\ &= \kappa_{B'}(a_1v_1 + \cdots + a_nv_n) \\ &= \kappa_{B'}(v). \end{aligned}$$

□

Lemma 5.33 *Es seien B, B' und B'' Basen des endlich-dimensionalen K-Vektorraums V. Ferner sei $P = P^B_{B'}$ die Übergangsmatrix von B nach B' und $Q = P^{B'}_{B''}$ die Übergangsmatrix von B' nach B''.*

Dann ist QP die Übergangsmatrix von B nach B''. Mit anderen Worten, es gilt

$$P^{B'}_{B''}P^B_{B'} = P^B_{B''}.$$

Beweis : Für $B = (v_1, \ldots, v_n)$ hat $P = P^B_{B'}$ die Spalten $\kappa_{B'}(v_1), \ldots, \kappa_{B'}(v_n)$. QP ist also die Matrix mit den Spalten $Q\kappa_{B'}(v_1), \ldots, Q\kappa_{B'}(v_n)$.

Nach Lemma 5.32 ist $Q\kappa_{B'}(v_i) = \kappa_{B''}(v_i)$, also ist QP die Übergangsmatrix von B nach B''. □

Korollar 5.34 *Jede Übergangsmatrix zwischen zwei Basen eines endlich-dimensionalen Vektorraums ist invertierbar.*

Beweis : Ist P die Übergangsmatrix von B nach B', so gilt für die Übergangsmatrix Q von B' nach B nach Lemma 5.33, dass QP die Übergangsmatrix von B nach B ist. Also ist $QP = E_n$, d.h. P ist invertierbar. □

Beispiel: Es sei $B = (v_1, \ldots, v_n)$ eine Basis des K^n. Dann ist die Matrix $[B]$ mit den Spalten $v_1, \ldots, v_n$ die Übergangsmatrix von B nach $B' = (e_1, \ldots, e_n)$. Also ist $[B]^{-1}$ die Übergangsmatrix von $B' = (e_1, \ldots, e_n)$ nach B.

Zum Abschluss dieses Kapitels werden wir uns nun noch mit Summen von Vektorräumen befassen.

Sind $W_1, \ldots, W_r$ Unterräume des K-Vektorraums V, so definieren wir ihre Summe als

$$\sum_{i=1}^{r} W_i = W_1 + \cdots + W_r = \{v \in V : v = w_1 + \cdots + w_r : w_i \in W_i\}$$

$W_1 + \cdots + W_r$ ist ein Unterraum von V, der $W_1, \ldots, W_r$ enthält (Übungsaufgabe).

Definition 5.35 *i) $W_1, \ldots, W_r$ heißen* ***unabhängig****, falls für alle $w_1 \in W_1, \ldots, w_r \in W_r$ gilt: $w_1 + \cdots + w_r = 0$ genau dann, wenn $w_1 = \cdots = w_r = 0$ gilt.*

ii) V ist die ***direkte Summe*** *von $W_1, \ldots, W_r$, d.h. $V = W_1 \oplus \cdots \oplus W_r$ oder $V = \bigoplus_{i=1}^{r} W_i$ genau dann, wenn $V = W_1 + \cdots + W_r$ ist und $W_1, \ldots, W_r$ unabhängig sind.*

Beispiel:

i) Ein Unterraum W_1 ist stets unabhängig.

ii) W_1 und W_2 sind genau dann unabhängig, wenn $W_1 \cap W_2 = 0$ ist.

Satz 5.36 *Seien $W_1, \ldots, W_r$ Unterräume eines endlich erzeugten Vektorraums und sei B_i eine Basis von W_i.*

i) Es ist $(B_1, \ldots, B_n)$ genau dann eine Basis von V, wenn $V = \bigoplus_{i=1}^{r} W_i$ gilt.

ii) Es ist $\dim_K(W_1 + \cdots + W_r) \leq \dim_K W_1 + \cdots + \dim_K W_r$. Gleichheit steht genau dann, wenn $W_1, \ldots, W_r$ unabhängig sind.

Beweis :

i) Ist $(B_1, \ldots, B_n)$ eine Basis von V, so ist dieses System von Vektoren linear unabhängig. Es sei $w_1 + \cdots + w_r = 0$ für $w_i \in W_i$, d.h. $w_i \in \langle B_i \rangle$. Wir schreiben alle w_i als Linearkombinationen der Vektoren in B_i. Dann müssen alle in $w_1 + \cdots + w_r$ auftretenden Koeffizienten Null sein, also folgt $w_i = 0$. Ferner ist $V = \langle (B_1, \ldots, B_n) \rangle$, woraus $V = W_1 + \cdots + W_r$ folgt.

Ist umgekehrt $V = \bigoplus_{i=1}^{r} W_i$, so ist $V = W_1 + \cdots + W_r$. Also ist $(B_1, \ldots, B_r)$ ein Erzeugendensystem von V. Eine Linearkombination der 0 aus $(B_1, \ldots, B_r)$ lässt sich schreiben als

$$w_1 + \cdots + w_r = 0$$

für $w_1 \in \langle B_1 \rangle = W_1, \ldots, w_r \in \langle B_r \rangle = W_r$. Da $W_1, \cdots, W_r$ unabhängig sind, folgt $w_1 = \cdots = w_r = 0$. Da B_i linear unabhängig ist, sind die Koeffizienten der Basisvektoren, die in w_i auftreten, alle Null. Somit ist $(B_1, \dots, B_r)$ linear unabhängig.

ii) Offenbar ist $(B_1, \dots, B_r)$ ein Erzeugendensystem von $W_1 + \cdots + W_r$, daher gilt

$$\dim_K(W_1 + \cdots + W_r) \leq \sum_{i=1}^{r} |B_i| = \sum_{i=1}^{r} \dim_K W_i$$

Gilt hier Gleichheit, so ist $(B_1, \dots, B_r)$ nach Satz 5.22 eine Basis des Vektorraums $W_1 + \cdots + W_r$. Mit i) folgt also $W_1 + \cdots + W_r = W_1 \oplus \cdots \oplus W_r$, d.h. $W_1, \dots, W_r$ sind unabhängig.

Sind umgekehrt $W_1, \dots, W_r$ unabhängig, so ist $W_1 + \cdots + W_r = W_1 \oplus \cdots \oplus W_r$. Mit i) folgt

$$\dim_K(W_1 + \cdots + W_r) = \sum_{i=1}^{r} |Bi| = \sum_{i=1}^{r} \dim_K W_i.$$

$\square$

Korollar 5.37 *Sei W ein Unterraum eines endlich-dimensionalen Vektorraums V. Dann gibt es einen Unterraum W' von V mit $W \oplus W' = V$. Jeden solchen Unterraum W' nennt man ein **Komplement** von W in V.*

Beweis : Nach Satz 5.23 ist auch W endlich erzeugt. Sei $(w_1, \dots, w_r)$ eine Basis von W. Nach Satz 5.15 können wir dieses linear unabhängige System zu einer Basis $(w_1, \dots, w_r, w_{r+1}, \dots, w_n)$ von V ergänzen. Es sei $W' = \langle w_{r+1}, \dots, w_n \rangle$. Nach Satz 5.36 ist $V = W \oplus W'$. $\square$

Für gegebenes W gibt es im allgemeinen viele Komplemente W'.

Satz 5.38 (Dimensionsformel für Unterräume) *Es seien W_1 und W_2 Unterräume eines endlich erzeugten K-Vektorraums.*

Dann gilt

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2)$$

Beweis : Der Schnitt von zwei Unterräumen ist wieder ein Unterraum (Übungsaufgabe), der Ausdruck $\dim(W_1 \cap W_2)$ macht also Sinn. Es sei $(u_1, \dots, u_r)$ eine Basis von $W_1 \cap W_2$, d.h. $\dim(W_1 \cap W_2) = r$. Nun ist $W_1 \cap W_2 \subset W_1$ und $W_1 \cap W_2 \subset W_2$. Wir können also dieses linear unabhängige System von Vektoren nach Satz 5.15 zu einer Basis $(u_1, \ldots, u_r, x_1, \ldots, x_{m-r})$ von W_1 und zu einer Basis $(u_1, \ldots, u_r, y_1, \ldots, y_{n-r})$ von W_2 ergänzen. Dann ist $\dim W_1 = m$ und $\dim W_2 = n$.

Es genügt zu zeigen, dass $B = (u_1, \ldots, u_r, x_1, \ldots, x_{m-r}, y_1, \ldots, y_{n-r})$ eine Basis von $W_1 + W_2$ ist. Daraus folgt nämlich

$$\begin{aligned}\dim W_1 + W_2 &= r + (m-r) + (n-r) \\ &= m + n - r \\ &= \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).\end{aligned}$$

Also zeigen wir zunächst

$$W_1 + W_2 = \langle B \rangle.$$

Da alle u_i, x_i und y_i in dem Vektorraum $W_1 + W_2$ liegen, gilt offenbar „$\supset$ “. Um die andere Inklusion zu zeigen, betrachten wir ein $w_1 + w_2 \in W_1 + W_2$, d.h. $w_1 \in W_1$ und $w_2 \in W_2$. Dann lassen sich w_1 und w_2 durch die oben konstruierten Basen ausdrücken:

$$w_1 = a_1 u_1 + \cdots + a_r u_r + b_1 x_1 + \cdots + b_{m-r} x_{m-r}$$

und

$$w_2 = a_1' u_1 + \cdots + a_r' u_r + c_1 y_1 + \cdots + c_{n-r} y_{n-r}.$$

Rechnet man $w_1 + w_2$ aus, so folgt $w_1 + w_2 \in \langle B \rangle$.

Jetzt zeigen wir, dass B linear unabhängig ist. Es sei

$$a_1 u_1 + \cdots + a_r u_r + b_1 x_1 + \cdots + b_{m-r} x_{m-r} + c_1 y_1 + \cdots + c_{n-r} y_{n-r} = 0$$

eine Linearkombination der Null. Dann folgt

$$\begin{aligned}\sum_{i=1}^{n-r} c_i y_i &= -\sum_{i=1}^{r} a_i u_i - \sum_{i=1}^{m-r} b_i x_i \\ &\in \langle u_1, \ldots, u_r, x_1, \ldots, x_{m-r} \rangle = W_1\end{aligned}$$

Somit liegt $\sum_{i=1}^{n-r} c_i y_i \in W_1 \cap W_2$, lässt sich also als Linearkombination der $u_1, \ldots, u_r$ schreiben:

$$\sum_{i=1}^{n-r} c_i y_i = \sum_{j=1}^{r} d_j u_j \text{ für gewisse } d_j \in K.$$

Nun sind $(u_1, \dots, u_r, y_1, \dots, y_{n-r})$ aber linear unabhängig, also sind alle $d_j = 0$ und alle $c_i = 0$. Die obige Linearkombination wird also zu

$$\sum_{i=1}^{r} a_i u_i + \sum_{i=1}^{m-r} b_i x_i = 0.$$

Da $(u_1, \dots, u_r, x_1, \dots, x_{m-r})$ linear unabhängig sind, sind auch alle $a_i = 0$ und alle $b_i = 0$. Daher ist B auch linear unabhängig, also in der Tat eine Basis von $W_1 + W_2$.

$\square$

6 Lineare Abbildungen

Definition 6.1 *Es seien V und W K-Vektorräume. Eine Abbildung $f : V \to W$ heißt* ***linear*** *(oder auch Homomorphismus von K-Vektorräumen), falls gilt*

i) $f(v_1 + v_2) = f(v_1) + f(v_2)$ *für alle* $v_1, v_2 \in V$

ii) $f(av) = af(v)$ *für alle* $a \in K, v \in V$.

Eine lineare Abbildung vertauscht also mit der Addition und der skalaren Multiplikation.

Lemma 6.2 *Ist $f : V \to W$ eine lineare Abbildung, so gilt für alle $n \in \mathbb{N}$ und alle $a_1, \dots, a_n \in K$ und $v_1, \dots, v_n \in V$*

$$f(\sum_{i=1}^{n} a_1 v_i) = \sum_{i=1}^{n} a_i f(v_i)$$

Beweis : mit Induktion nach n folgt das sofort aus Definition 6.1. (Führen Sie das aus!) $\square$

Beispiel: Sei $A \in K^{m \times n}$ eine $m \times n$-Matrix mit Koeffizienten in K. Dann ist die Multiplikation mit A:

$$\begin{aligned} \lambda_A : K^n &\to K^m \\ v &\mapsto Av \end{aligned}$$

eine lineare Abbildung nach den Rechenregeln für die Matrizenmultiplikation.

Definition 6.3 *Für jede lineare Abbildung $f : V \to W$ definieren wir den Kern von f als*

$$Kern(f) = \{v \in V : f(v) = 0\}$$

und das Bild von f als

$$Bild(f) = \{w \in W : \text{ es gibt ein } v \in V \text{ mit } f(v) = w\}.$$

Beispiel: Ist $f = \lambda_A : K^n \to K^m$ für $A \in K^{m\times n}$, so ist Kern$(f) = \{v : Av = 0\}$ die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$.

Bild $(f) = \{w \in K^m$: es gibt ein $v \in K^n$ mit $Av = w\}$ ist die Menge aller Vektoren w im K^m, für die das inhomogene Gleichungssystem $Ax = w$ eine Lösung besitzt.

Lemma 6.4 *Kern(f) ist ein Unterraum von V und Bild(f) ist ein Unterraum von W.*

Beweis : Sind v_1 und v_2 in Kern$f$, so gilt $f(v_1 + v_2) \overset{6.1i)}{=} f(v_1) + f(v_2) = 0 + 0 = 0$ und für alle $a \in K$

$$f(av_1) \overset{6.1ii)}{=} af(v_1) = a0 = 0.$$

Somit sind $v_1 + v_2$ und av_1 in Kern$f$. Da $f(0_V) = f(0_K 0_V) \overset{6.1ii)}{=} 0_K f(0_V) = 0_W$ ist, ist ferner $0 \in$ Kern$f$.

Also ist Kern$f$ ein Unterraum von V. Da $f(0) = 0$ ist, liegt $0 \in$ Bild(f). Sind ferner $w_1 = f(v_1)$ und $w_2 = f(v_2)$ in Bild f, so folgt

$$w_1 + w_2 = f(v_1) + f(v_2) \overset{6.1i)}{=} f(v_1 + v_2)$$

und für alle $a \in K$

$$aw_1 = af(v_1) \overset{6.1ii)}{=} f(av_1).$$

Somit ist $w_1 + w_2$ und aw_1 in Bild(f). Also ist Bild(f) ein Unterraum von W. □

Wir wollen jetzt zeigen, dass wir auf einer linear unabhängigen Teilmenge von V die Bilder einer linearen Abbildung vorschreiben können.

Proposition 6.5 *Seien V und W endlich erzeugte K-Vektorräume.*

i) Ist $(v_1, \ldots, v_m)$ linear unabhängig in V, so gibt es für beliebige $w_1, \ldots, w_m \in W$ eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, m$.

ii) Ist $(v_1, \ldots, v_m)$ eine Basis von V, so ist die Abbildung f aus i) eindeutig bestimmt.

Beweis :

i) Wir ergänzen$(v_1, \ldots, v_m)$ zu einer Basis $(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ von V und setzen für alle $a_1, \ldots, a_n \in K$:

$$f\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{m} a_i w_i.$$

Insbesondere ist $f(v_i) = w_i$ für $i = 1, \ldots, m$ und $f(v_i) = 0$ für $i = m+1, \ldots, n$. Da es für jedes $v \in V$ eindeutig bestimmte $a_1, \ldots, a_n$ mit $v = \sum\limits_{i=1}^{n} a_i v_i$ gibt, ist f eine wohldefinierte Abbildung $f : V \to W$. (Prüfen Sie das!) Es ist für $v = \sum\limits_{i=1}^{n} a_i v_i$ und $v' = \sum\limits_{i=1}^{n} a_i' v_i$ in V:

$$f(v + v') = \sum_{i=1}^{n} (a_i + a_i') v_i = f(v) + f(v').$$

Genauso einfach zeigt man für alle $a \in K \quad f(av) = af(v)$. Also ist f eine lineare Abbildung.

ii) Sind f und g zwei Abbildungen $V \to W$ mit $f(v_i) = w_i$ und $g(v_i) = w_i$, so ist $(f - g)$ eine lineare Abbildung $V \to W$ mit $(f - g)(v_i) = 0$. Somit ist $v_i \in$ Kern$(f - g)$ für alle i. Nach Lemma 6.4 ist Kern$(f - g) \subset V$ ein Unterraum. Da dieser die Basis $(v_1, \ldots, v_n)$ enthält, muss Kern $(f - g) = V$ sein. Also ist $f = g$.

$\square$

Proposition 6.6 *Sei $f : V \to W$ eine lineare Abbildung. Ist $(v_1, \ldots, v_n)$ ein Erzeugendensystem von V, d.h. gilt $\langle (v_1, \ldots, v_n) \rangle = V$, so ist*

$$\textit{Bild}\ (f) = \langle f(v_1), \ldots, f(v_n) \rangle.$$

Beweis : Jedes $v \in V$ lässt sich schreiben als $v = \sum\limits_{i=1}^{n} a_i v_i$ für $a_i \in K$. Ist $w = f(v)$ ein Element in Bild(f), so folgt

$$w = f(v) = f\left(\sum_{i=1}^{n} a_i v_i\right) \overset{f \text{ linear}}{=} \sum_{i=1}^{n} a_i f(v_i),$$

d.h. $w \in \langle f(v_1), \ldots, f(v_n) \rangle$. $\square$

Satz 6.7 (Dimensionsformel für lineare Abbildungen) *Es sei $f : V \to W$ eine lineare Abbildung und es sei V endlich erzeugt. Dann sind auch Kern(f) und Bild(f) endlich erzeugte Unterräume, und es gilt*

$$\dim V = \dim(\textit{ Kern} f) + \dim(\textit{ Bild} f)$$

Beweis : Kern(f) ist nach Satz 5.23 als Unterraum eines endlich erzeugten Vektorraumes selbst endlich erzeugt. Bild(f) ist nach Proposition 6.6 endlich erzeugt.

Es sei $(u_1, \ldots, u_d)$ eine Basis von Kern$f$. Diese lässt sich nach Satz 5.15 zu einer Basis $(u_1, \ldots, u_d, v_1, \ldots, v_{n-d})$ von V fortsetzen. Also gilt dim(Kern$f$) $= d$ und $\dim V = n$.

Wir behaupten, dass $B = (f(v_1), \ldots, f(v_{n-d}))$ eine Basis von Bild(f) ist. Wir zeigen zuerst Bild$(f) = \langle B \rangle$.

Hier ist „$\supset$“ klar. Um die andere Inklusion zu zeigen, sei $w \in$ Bild(f), d.h. $w = f(v)$ für ein $v = \sum\limits_{i=1}^{d} a_i u_i + \sum\limits_{j=1}^{n-d} b_j v_j \in V$. Also ist

$$\begin{aligned} w = f(v) &= \sum_{i=1}^{d} a_i f(u_i) + \sum_{j=1}^{n-d} b_j f(v_j) \\ &= \sum_{j=1}^{n-d} b_j f(v_j) \in \langle B \rangle, \end{aligned}$$

da alle u_i in Kern(f) liegen.

Jetzt zeigen wir, dass B linear unabhängig ist. Angenommen, es gilt

$$\sum_{i=1}^{n-d} a_i f(v_i) = 0 \text{ für } a_i \in K.$$

Dann folgt $f\left(\sum\limits_{i=1}^{n-d} a_i v_i\right) = 0$, d.h. $\sum\limits_{i=1}^{n-d} a_i v_i \in$ Kern$(f) = \langle u_1, \ldots, u_d \rangle$. Also gibt es $b_1, \ldots, b_d \in K$ mit

$$\sum_{i=1}^{n-d} a_i v_i = \sum_{j=1}^{d} b_j u_j.$$

Da $(u_1, \ldots, u_d, v_1, \ldots, v_{n-d})$ eine Basis von V, also insbesondere linear unabhängig ist, sind alle $a_i = 0$ und alle $b_j = 0$. Also ist $B = (f(v_1), \ldots, f(v_{n-d}))$ eine Basis von Bild(f). Es folgt:

$$\begin{aligned} &\dim \text{ Kern}(f) + \dim \text{ Bild}(f) \\ &= d + n - d \\ &= n = \dim V. \end{aligned}$$

$\square$

Definition 6.8 *Ist $f : V \to W$ eine lineare Abbildung und V endlich erzeugt, so heißt* $\dim$ *Bild(f) auch **Rang von** f. Wir schreiben dafür rang(f).*

Wir wollen nun zeigen, dass sich alle linearen Abbildungen nach Wahl von Basen als Matrixmultiplikation ausdrücken können. Dazu betrachten wir zunächst den Spezialfall $V = K^n$.

Lemma 6.9 *Ist $f : K^n \to K^m$ eine lineare Abbildung, so ist $f = \lambda_A$ die Multiplikation mit einer Matrix $A \in K^{m\times n}$.*

Beweis : Es sei A die $(m \times n)$-Matrix mit den Spalten $f(e_1), \ldots, f(e_n)$ für die kanonische Basis $e_1, \ldots, e_n$ von K^n. Dann ist $Ae_i = i$-te Spalte von $A = f(e_i)$.

Für jedes $v = a_1e_1 + \cdots + a_ne_n \in K^n$ folgt also

$$Av = \sum_{i=1}^{n} a_i(Ae_i) = \sum_{i=1}^{n} a_i f(e_i) = f\left(\sum_{i=1}^{n} a_i e_i\right) = f(v).$$

$\square$

Definition 6.10 *Es sei $f : V \to W$ eine lineare Abbildung zwischen beliebig endlich erzeugten K-Vektorräumen. Sei $B = (v_1, \ldots, v_n)$ eine Basis von V und $C = (w_1, \ldots, w_m)$ eine Basis von W.*

*Die Matrix $A(f)^B_C \in K^{m\times n}$ mit den Spalten $\kappa_C(f(v_1)), \ldots, \kappa_C(f(v_n))$ heißt dann **Koordinatenmatrix** von f bezüglich der Basen B und C.*

Beispiel:

i) Ist $f = \mathrm{id} : V \to V$ und sind B und C Basen von V, so ist $A(\mathrm{id})^B_C = P^B_C$ die Übergangsmatrix von B nach C.

ii) Ist $V = K^n$ und $W = K^m$ sowie B und C die kanonischen Basen von V bzw. W, so gilt für jede lineare Abbildung $f : V \to W$, dass f gerade die Multiplikation mit $A(f)^B_C$ ist. (Rechnen Sie das nach!)

Das folgende Resultat zeigt, dass auch lineare Abbildungen zwischen beliebigen Vektorräumen durch Multiplikation mit der Koordinatenmatrix beschrieben werden können.

Proposition 6.11 *Es sei $f : V \to W$ eine lineare Abbildung und $n = \dim V < \infty$ sowie $m = \dim W < \infty$. Ferner sei B eine Basis von V und C eine Basis von W. Dann gilt für alle $v \in V$:*

$$\kappa_C(f(v)) = A(f)_C^B \cdot \kappa_B(v).$$

Beweis : Sei $B = (v_1, \ldots, v_n)$. Dann ist $\kappa_B(v_i) = e_i$, also

$$\begin{aligned} A(f)_C^B \cdot \kappa_B(v_i) &= i\text{-te Spalte von } A(f)_C^B \\ &= \kappa_C(f(v_i)). \end{aligned}$$

Ist $v = \sum\limits_{i=1}^{n} a_i v_i \in V$ ein beliebiger Vektor, so gilt

$$\begin{aligned} A(f)_C^B \cdot \kappa_B\left(\sum_{i=1}^{n} a_i v_i\right) &= A(f)_C^B \left(\sum_{i=1}^{n} a_i \kappa_B(v_i)\right) \\ &= \sum_{i=1}^{n} a_i (A(f)_C^B \cdot \kappa_B(v_i)) \\ &= \sum_{i=1}^{n} a_i \kappa_C(f(v_i)) \\ &= \kappa_C\left(\sum_{i=1}^{n} a_i f(v_i)\right) \\ &= \kappa_C\left(f\left(\sum_{i=1}^{n} a_i v_i\right)\right) = \kappa_C(f(v)). \end{aligned}$$

□

Man kann dieses Resultat auch so ausdrücken: Für die Isomorphismen

$$\begin{array}{rcl} V & \to & K^n \\ v & \mapsto & \kappa_B(v) \end{array} \quad \text{und} \quad \begin{array}{rcl} W & \to & K^m \\ w & \mapsto & \kappa_C(w) \end{array}$$

aus Satz 5.30 ist das Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\kappa_B"'] & W \arrow[d, "\kappa_C"] \\ K^n \arrow[r, "\lambda_{A(f)_C^B}"'] & K^m \end{tikzcd}$$

kommutativ, d.h. es ist

$$\lambda_{A(f)_C^B} \circ \kappa_B = \kappa_C \circ f.$$

Wir wollen jetzt noch untersuchen, wie sich die Koordinatenmatrix ändert, wenn man die Basen wechselt.

Proposition 6.12 *Sei $f \cdot V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen, und seien B und B' Basen von V sowie C und C' Basen von W. Dann gilt für $P = P^B_{B'}$ und $Q = P^C_{C'}$ die Basiswechselgleichung.*

$$A(f)^{B'}_{C'} = Q \cdot A(f)^B_C \cdot P^{-1}$$

Beweis : Es sei $B' = (v'_1, \ldots, v'_n)$. Nach Korollar 5.34 ist $P^{-1} = P^{B'}_B$ die Übergangsmatrix von B' nach B. Also ist P^{-1} die Matrix mit den Spalten $(\kappa_B(v'_1), \ldots, \kappa_B(v'_n))$. Nach Proposition 6.11 ist $A(f)^B_C \cdot P^{-1}$ die Matrix mit den Spalten $(\kappa_C(f(v'_1)), \ldots, \kappa_C(f(v'_n)))$. Also ist nach Lemma 5.32 $Q(A(f)^B_C \cdot P^{-1})$ die Matrix mit den Spalten

$$(\kappa_{C'}(f(v'_1)), \cdots, \kappa_{C'}(f(v'_n))),$$

mithin gleich $A(f)^{B'}_{C'}$. □

Lemma 6.13 *Sei V ein n-dimensionaler Vektorraum und B eine Basis von V. Für jede invertierbare $n \times n$-Matrix P über K (d.h. jedes $P \in GL_n(K)$) gibt es dann eine Basis B' von V mit $P = P^B_{B'}$.*

Beweis : Es sei $P^{-1} = (a_{ij})_{i,j=1,\ldots,n}$ und $B = (v_1, \ldots, v_n)$. Wir setzen $v'_k = \sum\limits_{i=1}^n a_{ik} v_i$ für alle $k = 1, \ldots, n$. Gilt $P = (b_{ij})_{i,j=1,\ldots,n}$, so folgt aus $P^{-1}P = E_n$, dass für alle j

$$v_j = \sum_{k=1}^n b_{kj} v'_k$$

gilt. (Rechnen Sie das nach!) Also ist $V = \langle v_1, \ldots, v_n \rangle \subset \langle v'_1, \ldots, v'_n \rangle \subset V$, d.h. $(v'_1, \ldots, v'_n)$ ist ein Erzeugendensystem von V. Gilt $\sum\limits_{k=1}^n c_k v'_k = 0$ für $c_k \in K$, so folgt $\sum\limits_{i=1}^n \left(\sum\limits_{k=1}^n a_{ik} c_k \right) v_i = 0.$

Da $(v_1, \ldots, v_n)$ linear unabhängig sind, sind alle $\sum\limits_{k=1}^n a_{ik} c_k = 0$. Somit ist $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ eine Lösung des homogenen Gleichungssystems $P^{-1}x = 0$. Da P^{-1} invertierbar ist, folgt $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = 0$, d.h. $B' = (v'_1, \ldots, v'_n)$ ist linear unabhängig und somit eine Basis von V.

Nach Konstruktion ist $\kappa_B(v'_k)$ die k-te Spalte von P^{-1}. Also ist $P^{-1} = P^{B'}_B$ und daher $P = P^B_{B'}$. □

Ist $f : V \to W$ eine lineare Abbildung, so ist es oft nützlich, Basen B von V und C von W zu finden, für die die Koordinatenmatrix $A(f)^B_C$ eine besonders einfache Gestalt hat.

Satz 6.14 *i) Sei $f : V \to W$ eine lineare Abbildung endlich erzeugter K-Vektorräume. Dann gibt es Basen B von V und C von W, so dass gilt:*

$$A(f)^B_C = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = \left(\begin{array}{cccccc} \multicolumn{3}{c}{\overbrace{\hphantom{1 \quad \ddots \quad 1}}^{r}} & & & \\ 1 & & & & 0 & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{array} \right) \quad \left.\vphantom{\begin{matrix}1\\ \ddots \\ 1\end{matrix}}\right\} r$$

wobei $r =$ rang(f) ist.

ii) Ist $A \in K^{m \times n}$ eine beliebige $m \times n$-Matrix, so gibt es Matrizen $Q \in GL_m(K)$ und $P \in GL_n(K)$, so dass

$$QAP^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

für $r =$ rang(λ_A) gilt.

Beweis :

i) Es sei $(u_1, \ldots, u_s)$ eine Basis von Kern$f$. Aufgrund der Dimensionsformel für lineare Abbildungen gilt für $r =$ rang$(f) = \dim$ Bild(f):

$$r + s = \dim V.$$

Wir können also $(u_1, \ldots, u_s)$ durch Hinzunahme von r Vektoren $v_1, \ldots, v_r$ zu einer Basis $B = (v_1, \ldots, v_r, u_1, \ldots, u_s)$ von V ergänzen. Wir setzen $w_i = f(v_i)$ für $i = 1, \ldots, r$. Da jedes $v \in V$ sich als $\sum\limits_{i=1}^{r} a_i v_i + \sum\limits_{j=1}^{s} b_j u_j$ schreiben lässt, sind alle Vektoren $f(v) \in W$ von der Form

$$\sum_{i=1}^{r} a_i f(v_i) + \sum b_j 0 = \sum_{i=1}^{r} a_i w_i \in \langle w_1, \ldots, w_r \rangle.$$

$(w_1, \ldots, w_r)$ ist also ein Erzeugendensystem von Bild(f) aus $r = \dim$ Bild(f) vielen Elementen, nach Satz 5.22 also eine Basis von Bild(f). Diese ergänzen wir zu einer Basis $(w_1, \ldots, w_r, w_{r+1}, \ldots, w_m) = C$ von W. Dann gilt für $i = 1, \ldots, r$

$$\kappa_C(f(v_i)) = \kappa_C(w_i) = e_i$$

und für $i = 1, \ldots, s$

$$\kappa_C(f(u_i)) = \kappa_C(0) = 0$$

Also ist $A(f)_C^B$ von der gewünschten Gestalt.

ii) A definiert eine lineare Abbildung

$$\begin{aligned} \lambda_A : K^n &\to K^m \\ v &\mapsto Av. \end{aligned}$$

Nach i) gibt es Basen B von K^n und C von K^m, so dass $A(\lambda_A)_C^B$ die gewünschte Gestalt hat. Es sei $P = [B]^{-1}$ die Übergangsmatrix von der kanonischen Basis von K^n nach B und $Q = [C]^{-1}$ die Übergangsmatrix von der kanonischen Basis von K^m nach C. Nach Korollar 5.34 sind beide invertierbar.

Da A die Koordinatenmatrix von λ_A bezüglich der kanonischen Basen ist, folgt aus Proposition 6.12 $A(\lambda_A)_C^B = QAP^{-1}$.

$\square$

Man nennt Matrizen A, A' aus $K^{m\times n}$ **äquivalent**, falls es $P \in \mathrm{GL}_n(K)$ und $Q \in \mathrm{GL}_m(K)$ mit $QAP^{-1} = A'$ gilt. Satz 6.14 ii) klassifiziert also Matrizen bis auf Äquivalenz. Sind A, A' quadratische Matrizen in $K^{n\times n}$, so nennt man A und A' **ähnlich**, falls es ein $P \in \mathrm{GL}_n(K)$ mit $PAP^{-1} = A'$ gilt.

Wie man zu A eine möglichst einfache Matrix A' findet, die ähnlich zu A ist, werden wir später studieren.

Definition 6.15 *Sei $A \in K^{m\times n}$ eine $m \times n$-Matrix. Der **Rang** von A (bezeichnet mit rang(A)) ist definiert als der Rang der linearen Abbildung $\lambda_A : K^n \to K^m$. Also ist rang$(A) = \dim$ Bild(λ_A).*

Lemma 6.16 *Ist $A \in K^{m\times n}$ und $P \in GL_n(K)$ sowie $Q \in GL_m(K)$, so ist rang$(A) =$ rang(QAP). Äquivalente Matrizen haben also denselben Rang.*

Beweis : Offenbar ist Bild $(\lambda_{AP}) = $ Bild$(\lambda_A \circ \lambda_P) \subset$ Bild(λ_A), d.h. rang$(AP) \leq$ rang(A). Ist $(v_1, \ldots, v_r)$ eine Basis von Bild$(\lambda_A \circ \lambda_P)$, so ist nach Proposition 6.6

$(Qv_1, \ldots, Qv_r)$ ein Erzeugendensystem von Bild $(\lambda_Q \circ \lambda_A \circ \lambda_P) = \text{Bild}(\lambda_{QAP})$. Also folgt $\text{rang}(QAP) \leq \text{rang} A$. Wir wenden dies auf QAP statt A und Q^{-1} statt Q sowie P^{-1} statt P an und erhalten

$$\text{rang}(A) = \text{rang}(Q^{-1}(QAP)P^{-1}) \leq \text{rang}(QAP).$$

Also folgt die Behauptung. $\square$

Mit Satz 6.14 ii) folgt also, dass zwei Matrizen genau dann äquivalent sind, wenn sie denselben Rang besitzen.

Satz 6.17 *Es sei $A \in K^{m\times n}$ eine Matrix mit den Zeilen $z_1, \ldots, z_m \in K^{1\times n}$ und den Spalten $s_1, \ldots, s_n \in K^{m\times 1}$. Dann ist*

$$\begin{aligned} \textit{rang}(A) &= \dim\langle(z_1, \ldots, z_m)\rangle \\ &= \dim\langle(s_1, \ldots, s_n)\rangle. \end{aligned}$$

Ferner stimmt $\text{rang}(A)$ sowohl mit der maximalen Anzahl linear unabhängiger Zeilen als auch mit der maximalen Anzahl linear unabhängiger Spalten von A überein.

Beweis : Da $(e_1, \ldots, e_n)$ ein Erzeugendensystem von K^n ist, ist nach Proposition 6.6 $(s_1, \ldots, s_n)$ ein Erzeugendensystem von $\text{Bild}(\lambda_A)$. Also gilt $\text{rang}(A) = \dim\langle(s_1, \ldots, s_n)\rangle$. Wenden wir dies auf die transponierte Matrix A^t an, so folgt rang $(A^t) = \dim\langle(z_1, \ldots, z_m)\rangle$. Für die erste Behauptung müssen wir also noch $\text{rang}(A) = \text{rang}(A^t)$ zeigen. Nach Satz 6.14 ii) gibt es $P \in GL_n(K)$ und $Q \in GL_m(K)$, so dass

$$QAP^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

für $r = \text{rang}(\lambda_A) = \text{rang}(A)$ gilt. Daraus folgt nach Transponieren:

$$(P^t)^{-1}A^tQ^t = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

Nach Lemma 6.16 ist also $\text{rang}(A^t) = \text{rang}(P^t)^{-1}A^tQ^t = r = \text{rang}(A)$, denn den Rang der Matrix auf der rechten Seite können wir sofort ablesen.

Nach Satz 5.14 enthält jedes endliche Erzeugendensystem M eines Vektorraums eine Basis. Diese ist nach Satz 5.22 eine maximale linear unabhängige Teilmenge von V. Daraus folgt der zweite Teil der Behauptung. $\square$

Korollar 6.18 *Es sei $A \in K^{n \times n}$ eine quadratische Matrix. Dann sind äquivalent*

i) *A ist invertierbar.*

ii) *$\text{rang}(A) = n$.*

iii) *$\text{Kern}(\lambda_A) = 0$.*

Beweis : i) $\Rightarrow$ ii): Ist A invertierbar, so gilt $E_n A A^{-1} = E_n$, also ist nach Lemma 6.16 $\text{rang}(A) = \text{rang}(E_n A A^{-1}) = \text{rang}(E_n) = n$.

ii) $\Rightarrow$ iii): Ist $\text{rang}(A) = n$, so ist nach der Dimensionsformel für lineare Abbildungen dim Kern $(\lambda_A) = 0$, also ist $\text{Kern}(\lambda_A) = 0$.

iii) $\Rightarrow$ i): Ist $\text{Kern}(\lambda_A) = 0$, so ist das homogene lineare Gleichungssystem nur trivial lösbar. Nach Satz 2.9 ist A invertierbar. □

7 Eigenwerte

Wir haben im letzten Abschnitt gezeigt, wie man für eine lineare Abbildung $f : V \to W$ besonders einfache Koordinatenmatrizen finden kann, indem man die Basen in V und W geeignet wählt. Jetzt wollen wir uns mit linearen Abbildungen $f : V \to V$ befassen. Solche linearen Abbildungen heißen auch **Endomorphismen** von V.

Definition 7.1 *i) Sei $f : V \to V$ ein Endomorphismus des K-Vektorraumes V. Ein $\alpha \in K$ heißt* ***Eigenwert*** *von f, falls es ein $v \neq 0$ in V gibt, so dass*

$$f(v) = \alpha v$$

gilt. Der Vektor v heißt dann ***Eigenvektor zu*** *α.*

ii) Ein $\alpha \in K$ heißt ***Eigenwert*** *einer Matrix $A \in K^{n \times n}$, falls α Eigenwert der Multiplikationsabbildung $\lambda_A : K^n \to K^n$ ist. Analog heißt $v \in K^n$ Eigenvektor von A, wenn $v \neq 0$ und Eigenvektor von λ_A ist.*

Also ist α genau dann ein Eigenwert von A, wenn $Av = \alpha v$ für ein $v \neq 0$ aus K^n gilt. Dieser Vektor v heißt ***Eigenvektor*** *zu α.*

Eigenvektoren müssen $\neq 0$ sein, da die Gleichung $f(v) = \alpha v$ für $v = 0$ für alle $\alpha \in K$ erfüllt ist und somit nichts aussagt.

Beispiel:

i) $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$ hat den Eigenvektor e_1 mit Eigenwert 3.

ii) Wir betrachten die Drehung $d_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ um einen Winkel $\theta \in [0, 2\pi[$, d.h. d_θ ist die Multiplikation mit $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Ist nämlich $v = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix}$, so ist $d_\theta(v) = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix}$ mit den Additionstheoremen für $\sin$ und $\cos$.

d_θ hat keinen reellen Eigenwert, wenn $\theta \neq 0$ und $\theta \neq \pi$ ist. (Prüfen Sie das!)

Lemma 7.2 *Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V und B eine beliebige Basis von V. Dann ist $\alpha \in K$ genau dann ein Eigenwert von f, wenn α ein Eigenwert der Koordinatenmatrix $A(f)^B_B$ ist.*

Beweis : Wir betrachten den Isomorphismus $\kappa_B : V \to K^n$. Sei α ein Eigenwert von f. Dann existiert ein $v \neq 0$ mit $f(v) = \alpha v$.

Nach Proposition 6.11 ist $\kappa_B(f(v)) = A(f)^B_B \kappa_B(v)$. Also folgt $A(f)^B_B \kappa_B(v) = \kappa_B(\alpha v) = \alpha\kappa_B(v)$. Da mit v auch $\kappa_B(v) \neq 0$ ist, ist α Eigenwert von $A(f)^B_B$. Sei umgekehrt α ein Eigenwert von $A(f)^B_B$, d.h. $A(f)^B_B w = \alpha w$ für ein $w \neq 0$ in K^n. Ferner sei $0 \neq v \in V$ der Vektor mit $\kappa_B(v) = w$. Es folgt

$$\kappa_B(f(v)) \overset{6.10}{=} A(f)^B_B \kappa_B(v) = A(f)^B_B w = \alpha w = \kappa_B(\alpha v),$$

also ist $f(v) = \alpha v$, d.h. α ist Eigenwert von f. $\square$

Korollar 7.3 *Ähnliche Matrizen in $K^{n\times n}$ haben dieselben Eigenwerte.*

Beweis : Ist $A' = S^{-1}AS$ für ein $S \in GL_n(K)$, so sind A und A' Koordinatenmatrizen derselben linearen Abbildung $K^n \to K^n$ für verschiedene Basen (Übungsaufgabe). Also haben sie nach Lemma 7.2 dieselben Eigenwerte. $\square$

Proposition 7.4 *Sei $f : V \to V$ ein Endomorphismus des endlich-dimensionalen Vektorraums V und B eine Basis von V. Dann ist die Koordinatenmatrix $A(f)_B^B$ von f bezüglich B genau dann ähnlich zu einer Diagonalmatrix $\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$, wenn f eine Basis aus Eigenvektoren besitzt. In diesem Fall sind $d_1, \dots, d_n$ die Eigenwerte von f.*

Beweis : Ist $A(f)_B^B$ ähnlich zu $D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$, so existiert ein $S \in \mathrm{GL}_n(K)$ mit

$$SA(f)_B^B S^{-1} = D.$$

Nach Lemma 6.13 ist $S = P_C^B$ die Übergangsmatrix von B zu einer Basis C von V. Also ist nach Proposition 6.12 $D = A(f)_C^C$ die Koordinatenmatrix von f bezüglich $C = (v_1, \dots, v_n)$. Somit ist $f(v_1) = d_1 v_1, \dots, f(v_n) = d_n v_n$, d.h. C ist eine Basis aus Eigenvektoren.

Ist umgekehrt $C = (v_1, \dots, v_n)$ eine Basis aus Eigenvektoren von f, so seien $d_1, \dots, d_n$ die zugehörigen Eigenwerte. Dann ist $A(f)_C^C = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$. Nach Proposition 6.12 ist diese Diagonalmatrix ähnlich zu $A(f)_B^B$. □

Eine Matrix, die ähnlich zu einer Diagonalmatrix ist, nennt man auch **diagonalisierbar**.

Besitzt f also eine Basis aus Eigenvektoren, so gibt es eine Basis von V, so dass die zugehörige Koordinatenmatrix von f Diagonalgestalt hat. Einen solchen Endomorphismus nennt man ebenfalls **diagonalisierbar**.

Lemma 7.5 *Es seien $v_1, \dots, v_s$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\alpha_1, \dots, \alpha_s$ von f. Dann ist $(v_1, \dots, v_s)$ linear unabhängig.*

Beweis : mit Induktion nach s.

Anfang: ($s = 1$) Das ist klar, da Eigenvektoren $\neq 0$.sind.

Induktionsschluss: Die Behauptung gelte für $s - 1$ Eigenvektoren. Angenommen $\sum_{i=1}^{s} c_i v_i = 0$ mit $c_i \in K$ ist eine Linearkombination der Null.

Dann folgt

$$0 = f(0) = \sum_{i=1}^{s} c_i f(v_i) = \sum_{i=1}^{s} \alpha_i c_i v_i.$$

Außerdem ist auch $\sum\limits_{i=1}^{s} \alpha_1 c_i v_i = \alpha_1 \sum\limits_{i=1}^{s} c_i v_i = 0.$

Also ist $\sum\limits_{i=2}^{s} (\alpha_i - \alpha_1) c_i v_i = 0.$

Nach Induktionsvoraussetzung folgt $(\alpha_i - \alpha_1)c_i = 0$ für alle $i = 2, \dots, n$. Da $\alpha_1, \dots, \alpha_s$ paarweise verschieden sind, folgt $c_2 = \dots = c_n = 0$, also wegen $\sum\limits_{i=1}^{s} c_i v_i = 0$ und $v_1 \neq 0$ auch $c_1 = 0$. □

Korollar 7.6 *Ein Endomorphismus eines n-dimensionalen Vektorraums hat höchstens n verschiedene Eigenwerte.*

Beweis : klar. □

Definition 7.7 *Sei α ein Eigenwert von $f : V \to V$. Dann heißt der Unterraum $V_\alpha =$ Kern$(f - \alpha id) = \{v \in V : f(v) = \alpha v\}$ von V der* ***Eigenraum*** *von f zu α. Er besteht aus allen Eigenvektoren von α und der Null.*

(Hier ist id : $V \to V$ die Abbildung id$(v) = v$.).

Korollar 7.8 *Es sei $f : V \to V$ ein Endomorphismus des endlich-dimensionalen K-Vektorraums V und $\alpha_1, \dots, \alpha_r$ die sämtlichen (paarweise verschiedenen) Eigenwerte von f. Dann sind $V_{\alpha_1}, \dots, V_{\alpha_r}$ unabhängig, d.h. $\sum\limits_{i=1}^{r} V_{\alpha_i} = \bigoplus\limits_{i=1}^{r} V_{\alpha_i} \cdot f$ ist genau dann diagonalisierbar, wenn $V = \bigoplus\limits_{i=1}^{r} V_{\alpha_i}$ gilt. Das ist äquivalent zu $\dim V = \sum\limits_{i=1}^{r} \dim V_{\alpha_i}$.*

Beweis : Ist $v_i \in V_{\alpha_i}$, so folgt aus $\sum\limits_{i=1}^{r} v_i = 0$ mit Lemma 7.5 $v_1 = \dots = v_r = 0$. Also sind $V_{\alpha_1}, \dots, V_{\alpha_r}$ unabhängig. Aus Proposition 7.4 folgt dann die zweite Behauptung. Die Dimensionsbedingung gilt nach Satz 5.36. □

Wir wollen jetzt noch ein Verfahren kennenlernen, mit dem man die Eigenwerte berechnen kann. Dazu betrachten wir zu einem Endomorphismus eines endlich-dimensionalen K-Vektorraums $f : V \to V$ und einem $\alpha \in K$ die lineare Abbildung

$$\begin{array}{rcl} \alpha\mathrm{id} - f : & = & V \\ v & \mapsto & \alpha v - f(v) \end{array}$$

Lemma 7.9 *Folgende Aussagen sind äquivalent:*

i) $\mathit{Kern}(\alpha \mathit{id} - f) \neq 0$

ii) $\mathit{Bild}(\alpha \mathit{id} - f) \neq V$

iii) Ist B eine beliebige Basis von V, so ist $\mathit{det}\big(A(\alpha id - f)_B^B\big) = 0$

iv) α ist ein Eigenwert von f.

Beweis : i) $\Leftrightarrow$ ii) folgt aus der Dimensionsformel $\dim V = \dim(\,\mathrm{Kern}(\alpha\mathrm{id} - f)) + \dim(\,\mathrm{Bild}(\alpha\mathrm{id} - f))$.

i) $\Rightarrow$ iv): Ist $\mathrm{Kern}(\alpha\mathrm{id} - f) \neq 0$, so existiert ein $v \neq 0$ mit $\alpha v = f(v)$. Also ist α ein Eigenwert von f.

iv) $\Rightarrow$ iii): Ist α ein Eigenwert von f, so gibt es ein $v \neq 0$ mit $f(v) = \alpha v$. Also ist $\alpha v - f(v) = 0$, d.h. 0 ist ein Eigenwert von $\alpha\mathrm{id} - f$. Nach Lemma 7.2 ist 0 dann ein Eigenwert von $A(\alpha\mathrm{id} - f)_B^B$, d.h. der Kern der Multiplikation $A(\alpha id - f)_B^B$ ist $\neq 0$. Also ist nach Korollar 6.18 $\det\big(A(\alpha id - f)_B^B\big) = 0$.

iii) $\Rightarrow$ i): Ist $\det\big(A(\alpha id - f)_B^B\big) = 0$, so ist nach Korollar 6.18 der Kern der Multiplikation mit dieser Matrix $\neq 0$. Also existiert ein $w \neq 0$ in K^n mit $A(\alpha id - f)_B^B w = 0$, woraus $(\alpha\mathrm{id} - f)(\kappa_B(w)) = 0$ folgt. Somit ist auch $\mathrm{Kern}(\alpha\mathrm{id} - f) \neq 0$. $\square$

Mit diesem Lemma können wir Eigenwerte berechnen.

Beispiel: Sei $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$ und $\lambda_A : \mathbb{Q}^2 \to \mathbb{Q}^2$ wie immer die lineare Abbildung $\lambda_A(v) = Av$. Dann ist A die Koordinatenmatrix von λ_A bezüglich der kanonischen Basis auf $\mathbb{Q}^2$. Also ist $\alpha E - A$ die Koordinatenmatrix von $\alpha\mathrm{id} - f$ bezüglich der kanonischen Basis. Nach Lemma 7.9 ist α genau dann ein Eigenwert von A, wenn $\det(\alpha E - A) = 0$ ist.

Es ist $\alpha E - A = \begin{pmatrix} \alpha - 1 & -4 \\ -1 & \alpha - 1 \end{pmatrix}$, also

$$\det(\alpha E - A) = (\alpha - 1)^2 - 4 = \alpha^2 - 2\alpha - 3 = (\alpha + 1)(\alpha - 3),$$

Die Eigenwerte von A sind also -1 und 3.

Um das zu formalisieren, benötigen wir einige Tatsachen über Polynomringe.

Es sei K ein Körper. Der Polynomring $K[X]$ in einer Variablen über K ist definiert als die Menge aller (formalen) Summen der Form

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$$

für $a_0, \ldots, a_n \in K$ und beliebiges $n \in \mathbb{N}_0$.

(„Formal“ bedeutet hier, dass wir das $+$ stur hinschreiben und uns über die Bedeutung keine Gedanken machen.)

Zwei solche formalen Summen $f(x) = a_0 + a_1X + \ldots + a_nX^n$ und $g(X) = b_0 + b_1X + \ldots + b_mX^m$ sind gleich, falls für alle $k \leq N = \max\{n, m\}$ gilt: $a_k = b_k$. Falls $n < N$, so setzen wir hier $a_{n+1} = \ldots = a_N = 0$ und analog, falls $m < N, b_{m+1} = \ldots = b_N = 0$.

Für $f(X)$ und $g(X)$ wie oben definieren wir ferner

$$f(X) + g(X) = (a_0 + b_0) + (a_1 + b_1)X + \ldots + (a_N + b_N)X^N,$$

wobei wir wieder die überzähligen Koeffizienten gleich Null gesetzt haben.

Für $c \in K$ definieren wir noch

$$cf(X) = ca_0 + ca_1X + \ldots + ca_nX^n.$$

Für $f(X) = \sum\limits_{i=0}^{n} a_iX^i$ und $g(X) = \sum\limits_{j=0}^{m} b_jX^j$ definieren wir nun das Produkt $fg \in K[X]$ durch

$$fg(X) = \sum_{k=0}^{m+n} [\sum_{l=0}^{k} (a_l + b_{k-l})]X^k.$$

Dieses Produkt entsteht, indem man $f(X)g(X)$ formal ausmultipliziert und dann nach Potenzen von X sortiert.

Dann rechnet man leicht nach, dass $fg = gf$ sowie $f(gh) = (fg)h$ und $1f = f1 = f$ für das Polynom $a_0 = 1$ ist. Mit ein bisschen Schreibarbeit zeigt man außerdem für $f, g, h \in K[X]$, dass man in bekannter Weise Klammern auflösen kann:

$$(f + g)h = fh + gh.$$

Der **Grad** des Polynoms $0 \neq f(X) = a_0 + a_1X + \cdots + a_nX^n$ ist die größte Zahl k mit $a_k \neq 0$. Wir bezeichnen den Grad von f mit grad(f) (in der englischen Literatur deg(f) für degree). Wir setzen grad$(0) = -\infty$.

Für $d = \operatorname{grad}(f)$ können wir f also schreiben als $f = a_0 + a_1X + \cdots + a_dX^d$. Der „höchste" Koeffizient a_d von f heißt auch **Leitkoeffizient**. Ist $a_d = 1$, so heißt f **normiert**.

Wir betrachten nun für $A = (a_{ij}) \in K^{n\times n}$ die Matrix

$$XE_n - A = \begin{pmatrix} X - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & X - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \dots & X - a_{nn} \end{pmatrix} \in (K[X])^{n\times n}$$

Obwohl diese Matrix als Koeffizienten Polynome und nicht Körperelemente hat, ist mit der Leibnizformel 3.29 ihre Determinante definiert. Diese ist ebenfalls ein Polynom in X. Es gelten alle Rechenregeln aus § 3, bei denen wir nicht durch Einträge der Matrix teilen.

Definition 7.10 *$\chi(X) = \chi_A(X) = \det(XE_n - A)$ heißt das **charakteristische Polynom** der Matrix A.*

Beispiel: $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix}$. *Dann ist*

$$XE_3 - A = \begin{pmatrix} X - 1 & -2 & -1 \\ 0 & X + 1 & 0 \\ -1 & -2 & X - 3 \end{pmatrix}$$

und

$$\begin{aligned} \chi_A(X) &= (X+1)\det\begin{pmatrix} X-1 & -1 \\ -1 & X-3 \end{pmatrix} \\ &= (X+1)\cdot((X-1)(X-3)-1) \\ &= (X+1)(X^2-4x+2) \\ &= (X+1)(X-(2+\sqrt{2}))(X-(2-\sqrt{2})) \end{aligned}$$

Lemma 7.11 *Für alle $\alpha \in K$ ist $\chi_A(\alpha) = \det(\alpha E_n - A)$.*

Beweis : Wir betrachten die Abbildung

$$\varphi : K[X] \to K,$$

die durch Einsetzen von α definiert ist, das heißt, $\varphi(\sum\limits_{k=0}^{n} a_k X^k) = \sum\limits_{k=0}^{n} a_k \alpha^k$. Dann gilt $\varphi(f+g) = \varphi(f) + \varphi(g)$ und $\varphi(f \cdot g) = \varphi(f)\varphi(g)$ für alle $f, g \in K[X]$. Es ist $\chi_A(X) = \det(XE_n - A) = \sum\limits_{\sigma} \operatorname{sgn}(\sigma) b_{1\sigma(1)}(X) \cdots b_{n\sigma(n)}(X)$ für $XE_n - A = (b_{ij}(X))_{i,j}$ mit $b_{ij}(X) \in K[X]$. Daher ist $\chi_A(\alpha) = \sum_\sigma \operatorname{sgn}(\sigma) b_{1\sigma(1)}(\alpha) \cdots b_{n\sigma(n)}(\alpha) = \det(\alpha E_n - A)$. □

Proposition 7.12 *α ist genau dann ein Eigenwert von A, wenn $\chi_A(\alpha) = 0$ gilt.*

Beweis : $\chi_A(\alpha) = 0 \overset{7.11}{\Leftrightarrow} \det(\alpha E_n - A) = 0. \overset{7.9}{\Leftrightarrow} \alpha$ ist ein Eigenwert von A. □

Die Eigenwerte von A sind also genau die Nullstellen des charakteristischen Polynoms $\chi_A(X)$.

Definition 7.13 *Für $A = (a_{ij}) \in K^{n \times n}$ sei Spur $(A) = a_{11} + \cdots + a_{nn}$. Spur (A) ist also die Summe über die Diagonalelemente.*

Proposition 7.14 *Sei $A \in K^{n \times n}$. Dann ist $\chi_A(X)$ von der Form*

$$\chi_A(X) = X^n - \textit{Spur}(A) X^{n-1} + a_{n-2} X^{n-2} + \cdots + a_1 X + (-1)^n \det A$$

mit Koeffizienten $a_1, \ldots, a_{n-2} \in K$.

Beweis : Berechnen wir $\det(XE_n - A)$ mit der Leibnizformel, so gilt

$$\begin{aligned} \det(XE_n - A) &= (X - a_n) \cdots (X - a_{nn}) + \text{ Terme vom grad } \leq n-2 \\ &= X^n - \sum_{i=1}^{n} a_{ii} X^{n-1} + \text{ Terme vom grad } \leq n-2 \\ &= X^n - \text{ Spur } (A) X^{n-1} + \text{ Terme vom grad } \leq n-2. \end{aligned}$$

Schreiben wir $\chi_A(X) = X^n - \text{ Spur}(A) X^{n-1} + a_{n-2} X^{n-2} + \cdots + a_1 X + a_0$, so ist $a_0 = \chi_A(0) = \det(0E_n - A) = \det(-A) = (-1)^n \det A$. □

Lemma 7.15 *Ähnliche Matrizen haben dasselbe charakteristische Polynom, also insbesondere dieselbe Spur und dieselbe Determinante.*

Beweis : Ist $A' = SAS^{-1}$ für ein $S \in \mathrm{GL}(n, K)$, so ist

$$\begin{aligned}
\chi_{A'}(X) &= \det(XE_n - A') \\
&= \det(XE_n - SAS^{-1}) \\
&= \det(S(XE_n)S^{-1} - SAS^{-1}) \\
&= \det(S[XE_n - A]S^{-1}) \\
&= \det(S)\det(XE_n - A)\det(S^{-1}) \\
&= \chi_A(X).
\end{aligned}$$

□

Definition 7.16 *Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V. Das charakteristische Polynom von f ist definiert als*

$$\chi_f(X) = \chi_{A(f)^B_B}(X),$$

wobei $A(f)^B_B$ die Koordinatenmatrix von f bezüglich einer beliebigen Basis B von V ist.

Analog sei $\mathrm{Spur}(f) := \mathrm{Spur}(A(f)^B_B)$. Nach Lemma 7.15 hängt diese Definition nicht von der Wahl der Basis B ab.

Satz 7.17 *Für $f : V \to V$ sind äquivalent:*

i) f ist diagonalisierbar.

ii) $\chi_f(X) = \prod\limits_{i=1}^{r}(X - \alpha_i)^{n_i}$ mit $\alpha_i \in K, n_i \geq 0$, (d.h. χ_f zerfällt über K vollständig in Linearfaktoren) und $n_i = \dim V_{\alpha_i}$.

Beweis : i) $\Rightarrow$ ii) Ist f diagonalisierbar, so gibt es eine Basis B von V mit $A(f)^B_B = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$.

Also ist $\chi_f(X) = \det\left(XE_n - A(f)^B_B\right) = \prod\limits_{i=1}^{n}(X - d_i)$. Wir fassen jeweils dieselben d_i zusammen und schreiben $\chi_f(X) = \prod\limits_{i=1}^{r}(X - \alpha_i)^{n_i}$, wobei r die Anzahl der paarweise verschiedenen Eigenwerte von f ist. Offenbar hat der Eigenraum von $A(f)^B_B$ bezüglich α_i die Dimension n_i. Daher folgt $\dim V_{\alpha_i} = n_i$.

ii) $\Rightarrow$ i) Da χ_f den Grad n hat, folgt $\sum\limits_{i=1}^{r} n_i = n$. Mit Satz 5.36 folgt $V = \bigoplus V_{\alpha_i}$. Nach Korollar 7.8 ist f dann diagonalisierbar. □

Wir betrachten nun ein Polynom $f(x) = a_0 + a_1 X + \dots a_m X^m$ in $K[X]$.

Ist $A \in K^{n \times n}$ eine quadratische Matrix, so definieren wir

$$f(A) = a_0 + a_1 A + \dots + a_m A^m$$

mit den Rechenregeln für Matrizen. Dann ist $f(A)$ ebenfalls in $K^{n \times n}$. Dies liefert eine Abbildung

$$\begin{array}{rcl} K[X] & \to & K^{n \times n} \\ f & \mapsto & f(A), \end{array}$$

die $(f+g)(A) = f(A) + g(A)$ und $(fg)(A) = f(A)g(A)$ erfüllt. (Man nennt sie Einsetzungshomomorphismus.)

Satz 7.18 (Satz von Cayley-Hamilton) *Es ist* $\chi_A(A) = 0$.

Beweis : $\chi_A(A)$ ist eine $n \times n$-Matrix über K. Es genügt zu zeigen, dass Kern $(\chi_A(A)) = K^n$ gilt. Sei $v \in K^n$ ein Vektor $\neq 0$. Wir betrachten $v, Av, A^2 v, \dots$ und wählen r so, dass $(v, Av, \dots, A^r v)$ linear unabhängig, aber $(v, Av, \dots, A^{r+1} v)$ linear abhängig ist. Dann ist $b_0 = v, b_1 = Av, \dots, b_r = A^r v$ eine Basis von $U = \langle v, Av, \dots, A^r v \rangle$. Diese ergänzen wir zu einer Basis $B = (b_0, \dots, b_r, b_{r+1}, \dots, b_{n-1})$ von K^n. Die lineare Abbildung

$$\begin{array}{rcl} \lambda_A =: \lambda : K^n & \mapsto & K^n \\ v & \mapsto & Av \end{array}$$

erfüllt offenbar $\lambda(U) \subset U$. Also hat die Koordinatenmatrix $A_{\lambda,B,B}$ von λ bezüglich B die Gestalt

$$A_{\lambda,B,B} = \begin{pmatrix} C & * \\ 0 & D \end{pmatrix}$$

mit der $(r+1) \times (r+1)$-Matrix

$$C = \begin{pmatrix} 0 & 0 & & 0 & c_0 \\ 1 & 0 & & \vdots & c_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & & & 1 & c_r \end{pmatrix}$$

wobei $A^{r+1} v = c_0 v + c_1 Av + \dots + c_r A^r v$ gilt, und einer $(n-r-1) \times (n-r-1)$-Matrix D.

Es ist
$$\chi_C(X) = \det(XE_n - C) = X^{r+1} - c_r X^r \cdots - c_1 X - c_0,$$
wie man durch Entwickeln nach der letzten Spalte errechnet. Also ist
$$\chi_C(A)v = A^{r+1}v - c_r A^r v \cdots - c_1 A v - c_0 = 0,$$
woraus
$$\chi_A(A)v = (\chi_D(A) \cdot \chi_C(A))v = 0$$
folgt. □

Definition 7.19 *Das Minimalpolynom $p_A(X) \in K[X]$ einer Matrix $A \in K^{n\times n}$ ist definiert als das normierte Polynom kleinsten Grades mit $p_A(A) = 0$.*

Ist $q \in K[X]$ ein beliebiges Polynom mit $q(A) = 0$, so können wir es mit Rest durch p_A teilen. Es gibt also Polynome $g, r \in K[X]$, so dass grad(r) <grad(p_A) ist mit
$$q(X) = g(X)p_A(X) + r(X).$$
(Diese Tatsache lässt sich genau wie Lemma 4.4 beweisen.)

Wir setzen A ein und erhalten
$$0 = q(A) = g(A)p_A(A) + r(A) = r(A).$$
Aufgrund der Minimalität von p_A muss also $r = 0$ sein. Somit ist q ein Vielfaches von p_A.

Insbesondere ist das charakteristische Polynom ein Vielfaches des Minimalpolynoms.

Beispiel:

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

Dann ist
$$\chi_A(X) = (X-2)^3.$$
Die Teiler dieses Polynoms sind
$$1, (X-2), (X-2)^2 \text{ und } (X-2)^3.$$
Da $(A-2) \neq 0$ und $(A-2)^2 \neq 0$ ist, folgt $p_A(X) = (X-2)^3 = \chi_A(X)$.

ii) $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ hat das Minimalpolynom $p_A(X) = (X-2)$.

Wir haben im Satz 7.17 gelernt, dass ein Endomorphismus $f : V \to V$ genau dann diagonalisierbar ist, wenn erstens das charakteristische Polynom von der Form

$$\chi_f(X) = (X-\alpha_1)^{m_1} \dots (X-\alpha_r)^{m_r}$$

mit $\alpha_i \in K$ ist und wenn zweitens

$$m_i = \dim V_{\alpha_i}$$

gilt.

Den Exponenten m_i nennt man auch die algebraische Vielfachheit von α_i, die Zahl $\dim V_{\alpha_i}$ (die Dimension des Eigenraums zu α_i) heißt auch geometrische Vielfachheit von α_i.

Ist $f : V \to V$ ein beliebiger Endomorphismus eines endlich-dimensionalen Vektorraums, so gilt

$$\dim V_{\alpha_i} \leq n_i,$$

das heißt es ist geometrische Vielfachheit $\leq$ algebraische Vielfachheit. (Beweisen Sie das !)

Wir wollen nun auf die Bedingung geometrische Vielfachheit = algebraische Vielfachheit verzichten und trotzdem noch eine möglichst einfache Koordinatenmatrix für f angeben.

Satz 7.20 *Sei $f : V \to V$ ein Endomorphismus des n-dimensionalen K-Vektorraums V. Wir nehmen an, dass das charakteristische Polynom χ_f vollständig in Linearfaktoren über K zerfällt, d.h.*

$$\chi_f(X) = (X-\alpha_1)^{m_1} \cdot \ldots \cdot (X-\alpha_r)^{m_1}$$

mit $\alpha_i \in K$. Dann gibt es eine Basis B von V, so dass die Koordinatenmatrix $A(f)_B^B$ von f bezüglich B eine obere Dreiecksmatrix ist. Auf der Diagonalen von $A(f)_B^B$ stehen die Eigenwerte $\alpha_1, \dots, \alpha_r$ mit ihren Vielfachheiten $m_1, \dots, m_r$ gezählt, d.h. jedes α_i kommt

m_i-mal vor:

$$A(f)_B^B = \begin{pmatrix} \underbrace{\begin{matrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_1 \end{matrix}}_{m_1} & & & & * \\ & & \ddots & & \\ & & & \alpha_r & \\ & 0 & & & \ddots \\ & & & & \underbrace{\begin{matrix} & & \\ & & \\ & & \alpha_r \end{matrix}}_{m_r} \end{pmatrix}$$

Beweis : mit Induktion nach $n = \dim V$. Für $n = 1$ ist nichts zu zeigen. Wir nehmen jetzt an, $n \geq 2$ und die Behauptung gelte für alle Vektorräume der Dimension $\leq n-1$. Es sei $b_1 \neq 0$ ein Eigenvektor zum Eigenwert α_1. Wir ergänzen b_1 zu einer Basis $C = (b_1, c_2, \ldots, c_n)$ von V. Die Koordinatenmatrix von f bezüglich C hat dann die Gestalt

$$A(f)_C^C = \begin{pmatrix} \alpha_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

mit einer Matrix $A' \in K^{(n-1)\times(n-1)}$. Es sei $W := \langle c_2, \ldots, c_n \rangle \subset V$ der von $c_2, \ldots, c_n$ erzeugte Unterraum von V. Offenbar ist $\dim W = n - 1$.

Es sei $g : W \to W$ die lineare Abbildung, deren Koordinatenmatrix bezüglich der Basis $(b_2, \ldots, b_n)$ gerade die Matrix A' ist. Dann gilt für alle $w \in W$

$$f(w) = \gamma(w)b_1 + g(w)$$

mit einem $\gamma(w) \in K$.

Nun gilt:

$$\begin{aligned} \chi_f(X) &= \det(XE_n - A(f)_C^C) \\ &= (X - \alpha_1)\det(XE_n - A') \\ &= (X - \alpha_1)\chi_g(X). \end{aligned}$$

Also ist $\chi_g(X) = (X-\alpha_1)^{m_1-1}(X-\alpha_2)^{m_2}\cdot\ldots\cdot(X-\alpha_r)^{m_r}$, d.h. χ_g zerfällt vollständig in Linearfaktoren über K. Nach Induktionsvoraussetzung gibt es also eine Basis

$B' = (b_2, \ldots, b_n)$ von W, so dass die Koordinatenmatrix von g bezüglich B' eine obere Dreiecksmatrix der Form

$$A_{g,B',B'} = \begin{pmatrix} \alpha_1 & & & & & & \\ & \ddots & & & & * & \\ & & \alpha_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r & & \\ & 0 & & & & \ddots & \\ & & & & & & \alpha_r \\ & \underbrace{\hphantom{\alpha_1 \ \ddots \ \alpha_1}}_{m_1 - 1} & & & & \underbrace{\hphantom{\alpha_r \ \ddots \ \alpha_r}}_{m_r} & \end{pmatrix}$$

ist. Wir setzen $B = (b_1, b_2, \ldots, b_n)$. Dann gilt

$$A(f)_B^B = \begin{pmatrix} \alpha_1 & * & \ldots & * \\ 0 & & & \\ \vdots & A_{g,B',B'} & & \\ 0 & & & \end{pmatrix},$$

woraus die Behauptung folgt. □

Einen Endomorphismus, der bezüglich einer geeigneten Basis eine Koordinatenmatrix besitzt, die eine obere Dreiecksmatrix ist, nennt man auch trigonalisierbar. Satz 7.20 sagt also, dass f trigonalisierbar ist, falls χ_f vollständig in Linearfaktoren über K zerfällt.

Satz 7.21 *Für eine Matrix $A \in K^{n \times n}$ sind äquivalent:*

i) χ_A zerfällt vollständig in Linearfaktoren über K.

ii) A ist ähnlich zu einer oberen Dreiecksmatrix.

Beweis : i) $\Rightarrow$ ii) folgt aus Satz 7.20, angewandt auf die lineare Abbildung $\lambda_A : K^n \to K^n$. (Führen Sie das aus!)

ii) $\Rightarrow$ i). Nach Lemma 7.15 haben ähnliche Matrizen dasselbe charakteristische Polynom. Ist $D = \begin{pmatrix} d_1 & & * \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ eine obere Dreiecksmatrix, so gilt offenbar

$$\chi_D(X) = \det(XE_n - D) = \prod_{i=1}^{n} (X - d_i),$$

d.h. χ_D zerfällt vollständig in Linearfaktoren über K. □