6 Lineare Abbildungen und Matrizen

6.1 Lineare Abbildungen Seien V, W Vektorräume über $\mathbb{K}$. Eine Abbildung $f : V \to W$ heißt *linear*, wenn sie die beiden linearen Operationen Addition und Skalarmultiplikation erhält, wenn also

$$f(u+v) = f(u) + f(v), \quad f(\alpha u) = \alpha f(u) \quad \text{für alle } u, v \in V, \alpha \in \mathbb{K}.$$

Das Grundprinzip der modernen Mathematik besteht darin, zunächst Strukturen und dann strukturerhaltende Abbildungen zu definieren. In diesem Fall ist die Struktur der lineare Vektorraum zusammen mit Addition und Skalarmultiplikation.

Man kann die Linearität einer Abbildung äquivalent mit der Bedingung $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$ definieren. Mehrfache Anwendung der Linearitätsbedingung liefert

$$f\Big(\sum_{i=1}^{k} \alpha_i v_i\Big) = \sum_{i=1}^{k} \alpha_i f(v_i).$$

Aus $f(0) = f(0 \cdot 0) = 0f(0)$ folgt $f(0) = 0$.

Der *Nullraum* oder *Kern* einer linearen Abbildung ist

$$\operatorname{Kern} f = \{v \in V : f(v) = 0\} \subset V.$$

Er ist Unterraum von V weil für alle $v, v' \in \operatorname{Kern} f$ und alle $\alpha \in \mathbb{K}$ gilt $f(v+v') = f(v) + f(v') = 0$, $f(\alpha v) = \alpha f(v) = 0$, womit das Unterraumkriterium aus Satz 5.2 erfüllt ist.

Eine einfache, aber oft verwendete Eigenschaft des Kerns: Eine lineare Abbildung f ist genau dann injektiv, wenn $\operatorname{Kern} f = \{0\}$. Enthält der Kern noch eine weiteres Element, so werden dieses und die Null auf die Null abgebildet. Enthält der Kern nur die Null, so kann es nicht sein, dass ein w zwei verschiedene Urbilder v, v' besitzt wegen $0 = f(v - v')$ und daher $v = v'$.

Der *Bildraum* oder das *Bild* einer linearen Abbildung ist

$$\operatorname{Bild} f = \{w \in W : \exists v \in V \text{ mit } f(v) = w\} \subset W.$$

Das Bild ist Unterraum von W, denn wenn $w, w' \in \operatorname{Bild} f$, so gibt es $v, v' \in V$ mit $f(v) = w$ und $f(v') = w'$. Damit ist $w + w' = f(v) + f(v') = f(v + v')$ und $w + w'$ ist aus dem Bild von f. Für die Skalarmultiplikation zeigt man das genauso.

Beispiele 6.1 (i) Die Abbildung in die Null $v \mapsto 0$ ist immer eine lineare Abbildung zwischen den Räumen V und W. In diesem Fall ist $\operatorname{Kern} f = V$ und $\operatorname{Bild} f = \{0\}$.

(ii) Die Identität $Id : V \to V$ ist linear mit $\operatorname{Kern} f = \{0\}$ und $\operatorname{Bild} f = V$.

(iii) Die *orthogonalen Transformationen* des $\mathbb{R}^2$, das sind Drehungen und Spiegelungen an einer Geraden, die durch den Nullpunkt läuft, sind linear. Auf die Konstruktion solcher Abbildungen werden wir später eingehen.

(iv) Dagegen ist die *Translation* um einen Vektor $v_0 \neq 0$, das ist $f(v) = v_0 + v$, keine lineare Selbstabbildung des $\mathbb{R}^2$. Man erkennt das schon daran, dass $f(0) = v_0 \neq 0$. □

Satz 6.2 (a) *Die Komposition linearer Abbildungen ist linear.*

(b) *Ist $f : V \to W$ linear und bijektiv, so ist auch die Inverse $f^{(-1)} : W \to V$ linear.*

(c) *Sind $f, g : V \to W$ linear, so ist die punktweise Summe $(f + g)(x) = f(x) + g(x)$ und die Multiplikation mit Skalaren $(\alpha f)(x) = \alpha f(x)$ linear.*

Beweis: (a) Ist $g : V \to W$ linear sowie $f : W \to X$ linear, so gilt

$$f\big(g(v + v')\big) = f\big(g(v) + g(v')\big) = f\big(g(v)\big) + f\big(g(v')\big).$$

Für die Skalarmultiplikation läuft das genauso.

(b) Das folgt aus

$$f(v) = w, \quad f(v') = w', \quad f(v + v') = f(v) + f(v'),$$

wenn man in der letzten Gleichung auf beiden Seiten $f^{(-1)}$ anwendet,

$$v + v' = f^{(-1)}\big(f(v) + f(v')\big) \;\Rightarrow\; f^{(-1)}(w) + f^{(-1)}(w') = f^{(-1)}(w + w').$$

Für die Skalarmultiplikation folgt das ebenso einfach.

(c) Dazu brauchen wir eine sehr einfache Rechnung

$$\begin{aligned}(f+g)(\beta v + \beta' v') &= f(\beta v + \beta' v') + g(\beta v + \beta' v') = \beta f(v) + \beta' f(v') + \beta g(v) + \beta' g(v') \\ &= \beta(f+g)(v) + \beta'(f+g)(v').\end{aligned}$$

Für αf beweist man das ganz analog. □

Eine bijektive lineare Abbildung $f : V \to W$ heißt *Isomorphismus*. In diesem Fall heißen die beiden Räume V und W *isomorph* und man schreibt $V \cong W$. Nach Satz 6.2(b) ist die Inverse eines Isomorphismusses selber linear, insbesondere ist $f^{(-1)} : W \to V$ ein Isomorphismus.

Die Komposition von Isomorphismen ist ein Isomorphismus, denn nach Satz 6.2(a) ist sie linear und bekanntlich ist die Komposition bijektiver Abbildungen bijektiv.

Damit ist die Isomorphie eine Äquivalenzrelation. Es gilt $Id : V \to V$ linear und damit $V \cong V$. Da die Umkehrung eines Isomorphismus ebenfalls ein Isomorphismus ist, gilt $V \cong W$ genau dann, wenn $W \cong V$. Aus „Komposition von Isomorphismen ist wieder Isomorphismus" folgt die Transitivität von $\cong$.

Damit können isomorphe Vektorräume als Vektorräume nicht voneinander unterschieden werden. Sofern eine Aussage nur aus den beiden linearen Operationen aufgebaut ist, gilt sie in V genau dann, wenn sie auch in W gilt.

Satz 6.3 *Alle endlich dimensionalen Vektorräume über $\mathbb{K}$ der Dimension n sind zueinander isomorph. Insbesondere ist jeder n-dimensionale Vektorraum isomorph zu $\mathbb{K}^n$.*

Beweis: Wir nehmen eine beliebige Basis von V, sagen wir $v_1, \dots, v_n$, und definieren die *Koordinatenabbbildung*

$$v = \sum_{i=1}^{n} \alpha_i v_i \in V \;\mapsto (\alpha_1, \dots, \alpha_n)^T \in \mathbb{K}^n.$$

Die so definierte Abbildung $f : V \to \mathbb{K}^n$, $f(v) = (\alpha_1, \dots, \alpha_n)^T$ ist linear. Denn wenn $v = \sum_i \alpha_i v_i$, $v' = \sum_i \alpha_i' v_i$, so folgt

$$f(v + v') = (\alpha_1 + \alpha_1', \dots, \alpha_n + \alpha_n')^T = (\alpha_1, \dots, \alpha_n)^T + (\alpha_1', \dots, \alpha_n')^T = f(v) + f(v').$$

Es gilt $f(v_i) = e_i$ mit den kanonischen Einheitsvektoren e_i des $\mathbb{K}^n$. Da die e_i eine Basis des $\mathbb{K}^n$ bilden, ist $f(v) = f(\sum_i \alpha_i v_i) = \sum_i \alpha_i e_i = 0$, genau dann, wenn alle $\alpha_i = 0$. Damit ist f injektiv. f surjektiv ist noch offensichtlicher, weil jeder Punkt $\sum_i \alpha_i e_i$ im Bild von f liegt.

Da Isomorphie eine Äquivalenzrelation ist, sind alle $\mathbb{K}$-Vektorräume der Dimension n zueinander isomorph. □

Beispiel 6.4 Wir können jetzt die Konstruktionen in den Beispielen 5.1 mathematisch präziser fassen.

Wir hatten in 5.1(i) einem Polynom $p(x) = \sum_i \alpha_i x^i \in \mathbb{P}_n$ den Koeffizientenvektor $(\alpha_0, \dots, \alpha_n) \in \mathbb{R}^{n+1}$ zugeordnet. Dies definiert eine lineare Abbildung $I : \mathbb{P}_n \to \mathbb{R}^{n+1}$. Die Linearität hatten wir nachgewiesen, bijektiv ist I offenbar auch. Damit ist I ein Isomorphismus zwischen den angegebenen Räumen.

Genauso hatten wir in Beispiel 5.1(ii) den Matrizenraum $\mathbb{K}^{m\times n}$ linear und bijektiv auf den Raum $\mathbb{K}^{mn}$ abgebildet. Auch diese Räume sind daher isomorph. □

Mit $\mathcal{L}(V,W)$ bezeichnen wir die Menge der linearen Abbildungen zwischen den Vektorräumen V und W. Wie in Satz 6.2(c) gezeigt wurde, sind auf $\mathcal{L}(V,W)$ die punktweise Addition und Skalarmultiplikation, also $(f+g)(v) = f(v) + g(v)$ sowie $(\alpha f)(v) = \alpha f(v)$ erklärt, die $\mathcal{L}(V,W)$ ebenfalls zu einem $\mathbb{K}$-Vektorraum machen.

Sei $\dim V = n$ und $v_1, \ldots, v_n$ eine Basis von V. Dann ist wegen

$$f\Big(\sum_{i=1}^{n} \alpha_i v_i\Big) = \sum_{i=1}^{n} \alpha_i f(v_i) \tag{6.1}$$

jede lineare Abbildung durch die Werte auf einer Basis eindeutig bestimmt. Ferner wird das Bild durch eine Linearkombination dieser n Bildvektoren $f(v_i)$ aufgespannt. Daher ist das Bild eines n-dimensionalen Vektorraums endlich dimensional mit $\dim \operatorname{Bild} f \leq n$.

Die Dimension des Bildes einer linearen Abbildung f heißt *Rang* von f, geschrieben $\operatorname{rang} f$. Der folgende Satz wird auch *Rangformel* genannt:

Satz 6.5 *Sei V endlich dimensional und $f \in \mathcal{L}(V,W)$. Dann gilt*

$$\dim V = \dim \operatorname{Kern} f + \operatorname{rang} f.$$

Beweis: Da $\operatorname{Kern} f$ endlich dimensional ist, gibt es Basen $u_1, \ldots, u_r$ von $\operatorname{Kern} f$ und $w_1, \ldots, w_s$ von $\operatorname{Bild} f$ mit Urbildern $v_1, \ldots, v_s$. Wegen $\sum_i \alpha_i f(v_i) = \sum_i \alpha_i w_i$ spannt das Bild von $V_B = \operatorname{span}\{v_1, \ldots, v_s\}$ das gesamte Bild auf. Die Vektoren $\{v_i\}$ müssen daher l.u. sein und damit $\dim V_B = s = \operatorname{rang} f$. Für beliebiges $v \in V$ gilt $f(v) = \sum_i \alpha_i w_i$. Setze daher

$$v = v_B + v_K,\ v_B = \sum_{i=1}^{s} \alpha_i v_i \ \Rightarrow\ f(v - v_B) = 0 \ \Rightarrow\ f(v_K) = 0 \ \Rightarrow\ v_K = \sum_{i=1}^{r} \beta_i u_i.$$

Nach Konstruktion sind die Koeffizienten $\alpha_1, \ldots, \alpha_s, \beta_1, \ldots, \beta_r$ eindeutig bestimmt, daher die Vektoren $u_1, \ldots, u_r, v_1, \ldots, v_s$ linear unabhängig. Da sich jedes v nach diesen Vektoren entwickeln lässt, handelt es sich um eine Basis von V und es gilt $n = r + s$. □

Korollar 6.6 *Sind V, W endlich dimensional mit* $\dim V = \dim W$*, so gilt für jede lineare Abbildung $f \in \mathcal{L}(V,W)$*

$$f \text{ ist injektiv} \ \Leftrightarrow\ f \text{ ist surjektiv} \ \Leftrightarrow\ f \text{ ist bijektiv.}$$

Beweis: Nach der Rangformel impliziert jede dieser Bedingungen, dass $\dim \operatorname{Bild} f = \dim V$. □

Halten wir noch einmal die wichtigsten Aussagen dieses Abschnitts für ein n-dimensionales V und $f \in \mathcal{L}(V,W)$ fest:

Die Abbildung f ist bereits durch die Werte auf einer beliebigen Basis von V eindeutig bestimmt.

Dies folgt aus (6.1), die gleichzeitig die wichtigste Formel der linearen Algebra ist. (6.1) zeigt auch, dass das Bild von f von den $f(v_i)$ aufgebaut wird. Demnach ist $\dim \operatorname{Bild} f \leq \dim V$. Eine lineare Abbildung kann also höchstens einen n-dimensionalen Bildraum aufspannen.

Für eine Basis $v_1, \ldots, v_n$ von V kann man $f(v_i) \in W$ beliebig vorgeben. Durch diese Vorgaben ist die lineare Abbildung f eindeutig bestimmt.

f ist demnach für $v = \sum_i \alpha_i v_i$ definiert durch $f(v) = \sum_i \alpha_i f(v_i)$. Dass ein so definiertes f linear ist, weist man ohne Mühe nach.

6.2 Darstellungsmatrizen linearer Abbildungen Eine $(m \times n)$-Matrix A hatten wir als rechteckiges Schema $A = (a_{ij})_{i=1,\dots,m,\ j=1,\dots n}$ definiert zusammen mit komponentenweiser Addition und Skalarmultiplikation. $\mathbb{K}^{m\times n}$ ist mit diesen Operationen ein linearer Vektorraum über $\mathbb{K}$ mit $\dim \mathbb{K}^{m\times n} = mn$.

Mit Hilfe von Matrizen lassen sich lineare Abbildungen $f : \mathbb{K}^n \to \mathbb{K}^m$ anschaulich beschreiben. Wir hatten im letzten Abschnitt eingesehen, dass f durch die Werte auf einer Basis von $\mathbb{K}^n$ eindeutig bestimmt ist. Naheliegend ist es, hier die kanonische Basis $\{e_j\}_{j=1,\dots,n}$ zu wählen. Es gilt also $f(e_j) = a_j \in \mathbb{K}^m$. Wir schreiben die Spaltenvektoren $(a_j)_{j=1,\dots,n}$ hintereinander und erhalten so eine $(m \times n)$-Matrix

$$A_f = (a_1|a_2|\dots|a_n) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{mit } a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

die als *Darstellungsmatrix* der linearen Abbildung f bezeichnet wird. Für $u = \sum_{j=1}^n x_j e_j$ gilt dann

$$f(u) = \sum_{j=1}^n x_j f(e_j) = \sum_{j=1}^n x_j a_j = \sum_{j=1}^n \sum_{i=1}^m x_j a_{ij} e_i = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

Man bestimmt also $f(u)$ nach der Regel „Zeile $\times$ Spalte“. Um die i-te Komponente von $f(u)$ zu bekommen, nimmt man die i-te Zeile der Matrix, das ist $a_{i1}, \dots, a_{in}$ und multipliziert sie mit den entsprechenden Einträgen des Spaltenvektors u, das ist $x_1, \dots, x_n$ und addiert schließlich das Ganze, also

$$(f(u)_i = a_{i1}x_1 + \dots + a_{in}x_n.$$

Beispiel 6.7 Sei $f : \mathbb{R}^3 \to \mathbb{R}^2$ mit $f(e_1) = (1,1)^T$, $f(e_2) = (2,1)^T$, $f(e_3) = (-1,3)^T$. Zu f gehört demnach die Darstellungsmatrix

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

Für $u = (1,2,-4)^T$ bestimmen wir $f(u)$ nach der Regel Zeile mal Spalte

$$f(u) = \begin{pmatrix} 1\cdot 1 + 2\cdot 2 + (-1)\cdot(-4) \\ 1\cdot 1 + 1\cdot 2 + 3\cdot(-4) \end{pmatrix} = \begin{pmatrix} 9 \\ -9 \end{pmatrix}.$$

Die Regel Zeile mal Spalte soll aber nicht vergessen lassen, dass wir eine Linearkombination der Spaltenvektoren bilden, in diesem Fall

$$f(u) = 1\cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 4\cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 9 \\ -9 \end{pmatrix}.$$

$\square$

Nun untersuchen wir, wie die Komposition linearer Abbildungen mit den zugehörigen Darstellungsmatrizen realisiert werden kann. Seien $f : \mathbb{K}^m \to \mathbb{K}^l$ und $g : \mathbb{K}^n \to \mathbb{K}^m$ mit Matrixdarstellungen $A \in \mathbb{K}^{l\times m}$ von f und $B \in \mathbb{K}^{m\times n}$ von g. Mit $e_1^m, \dots, e_m^m$ bezeichnen wir die kanonische Basis von $\mathbb{K}^m$ und mit $e_1^l, \dots, e_l^l$ die kanonische Basis von $\mathbb{K}^l$. Dann gilt

$$f(e_k^m) = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{lk} \end{pmatrix} = \sum_i a_{ik} e_i^l, \quad g(x) = \begin{pmatrix} \sum_j b_{1j}x_j \\ \sum_j b_{2j}x_j \\ \vdots \\ \sum_j b_{mj}x_j \end{pmatrix} = \sum_{kj} b_{kj} x_j e_k^m,$$

und daher

$$f(g(x)) = f\Big(\sum_{kj} b_{kj} x_j e_k^m\Big) = \sum_{kj} b_{kj} x_j f(e_k^m) = \sum_{kji} b_{kj} x_j a_{ik} e_i^l$$

$$= \sum_{ji} \underbrace{\sum_{k=1}^{m} a_{ik} b_{kj}}_{\text{Zeile}\times\text{Spalte}=c_{ij}} \quad x_j e_i^l = \sum_{ij} c_{ij} x_j e_i^l = Cx, \quad C = (c_{ij}).$$

Um die Darstellungsmatrix für die Komposition zweier linearer Abbildungen zu bekommen gilt also wieder die Regel „Zeile × Spalte“. Ist $A = (a_{ik})$, $B = (b_{kj})$ und $C = (c_{ij})$ die Darstellungsmatrix für die Komposition, so gilt

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{im} b_{mj}.$$

Beispiel 6.8 Wir nehmen als f die gleiche lineare Abbildung wie in Beispiel 6.7 und für $g : \mathbb{R}^2 \to \mathbb{R}^3$ die Abbildung mit $g(e_1) = (2, 0, 1)^T$ und $g(e_2) = (2, 1, 0)^T$,

$$A_f = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 3 \end{pmatrix}, \quad B_g = \begin{pmatrix} 2 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Wir bestimmen das Produkt nach der Regel Zeile mal Spalte

$$C = \begin{pmatrix} 1 \cdot 2 + 2 \cdot 0 + (-1) \cdot 1 & 1 \cdot 2 + 2 \cdot 1 + (-1) \cdot 0 \\ 1 \cdot 2 + 1 \cdot 0 + 3 \cdot 1 & 1 \cdot 2 + 1 \cdot 1 + 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 5 & 3 \end{pmatrix}$$

□

6.3 Der Matrizenkalkül Der Raum $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ mit den Operationen der punktweisen Addition $(f + g)(x) = f(x) + g(x)$ und der punktweisen Skalarmultiplikation $(\alpha f)(x) = \alpha f(x)$ sind zum Matrizenraum $\mathbb{K}^{m\times n}$ isomorph mit dem Isomorphismus $I : f \mapsto A_f$. Der punktweisen Addition im Raum $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ entspricht genau die Matrizenaddition, bei der Skalarmultiplikation ist es genauso. Ferner können wir den Raum $\mathbb{K}^n$ mit dem Raum der Spaltenmatrizen $\mathbb{K}^{n\times 1}$ identifizieren. Im $\mathbb{K}^n$ haben wir damit zwei völlig äquivalente Begriffe, nämlich Vektoren und lineare Abblidungen auf der einen Seite und Matrizen auf der anderen Seite. Wir können daher die Herkunft aus der linearen Algebra vergessen und eine reine Matrizenrechnung betreiben, zumal wir jetzt nicht mehr zwischen Vektoren und Matrizen unterscheiden müssen. Letztere sind jetzt ebenfalls Matrizen und der Auswertung $f(u)$ entspricht das Matrizenprodukt Ax zwischen der Darstellungsmatrix A und der Matrix $x \in \mathbb{K}^{n,1}$. Dennoch werden wir zur besseren Unterscheidbarkeit von anderen Matrizen die $\mathbb{K}^{n\times 1}$-Matrizen als Vektoren bezeichnen.

Auf dem Matrizenraum $\mathbb{K}^{m\times n}$ sind daher Addition und Skalarmultiplikation sowie das Matrizenprodukt zwischen Elementen aus $\mathbb{K}^{l\times m}$ und $\mathbb{K}^{m\times n}$ definiert mit Ergebnis im Raum $\mathbb{K}^{l\times n}$.

Satz 6.9 *Sofern die Operationen definiert sind, gelten die folgenden Rechenregeln:*

(a) *Das Matrizenprodukt ist assoziativ*

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

(b) *Es gelten die Distibutivgesetze*

$$A(B + C) = AB + AC, \quad (A + B)C = AC + BC.$$

(c) *Matrix- und skalare Multiplikation sind homogen*

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

Beweis: (a) Hier greift man besser auf die Herkunft des Matrizenprodukts als Komposition linearer Abbildungen zurück. Letztere ist bekanntlich assoziativ.

(b) Das gilt ebenfalls für lineare Abbildungen, folgt aber auch direkt aus der Definition „Zeile $\times$ Spalte“.

(c) Das ist trivial, ist doch $\alpha A = (\alpha a_{ij})$ □

Im Fall $m = n$ sprechen wir von *quadratischen Matrizen* $A \in \mathbb{K}^{n \times n}$. Die Matrizenprodukte AB und BA sind hier zwar definiert, stimmen in der Regel aber nicht überein wie das folgende Beispiel zeigt:

Beispiel 6.10 Es ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

□

Die *Einheitsmatrix* von $\mathbb{K}^{n \times n}$ ist

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

und entspricht der Identität im Raum $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^n)$. Es gilt $AE_n = E_n A = A$.

Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt *regulär* oder *invertierbar*, wenn es eine Matrix $A^{-1} \in \mathbb{K}^{n \times n}$ gibt mit $AA^{-1} = E_n$. Eine reguläre Matrix ist die Darstellungsmatrix eines Isomorphismusses des $\mathbb{K}^n$, A^{-1} ist demnach die Darstellungsmatrix des inversen Isomorphismusses. Daher gilt im Falle einer regulären Matrix auch $A^{-1}A = E_n$. Die regulären Matrizen bilden mit der Matrix-Multiplikation eine Gruppe mit neutralem Element E_n, die mit $GL(n, \mathbb{K})$ notiert und allgemeine lineare Gruppe (engl: general linear group) genannt wird.

Wir können viele Begriffsbildungen für lineare Abbildungen auf Matrizen $A \in \mathbb{K}^{m \times n}$ übertragen: Das *Bild* von A ist die Menge der durch Ax erzeugten Elemente

$$\operatorname{Bild} A = \{y = Ax : x \in \mathbb{K}^n\} = \{y = \sum_{i=1}^{n} x_i a_i,\ x_i \in \mathbb{K}\},$$

wobei $a_1, \ldots, a_n \in \mathbb{K}^m$ die Spalten von A bezeichnet. Der *Rang* von A ist die Dimension des Bildraums, also die Anzahl der linear unabhängigen Spalten von A. Da wir auch von der Anzahl der linear unabhängigen Zeilen sprechen können, wird der Rang im Zusammenhang mit Matrizen auch als *Spaltenrang* bezeichnet.

Der *Kern* ist die Menge der Vektoren, die auf die Null abgebildet werden,

$$\operatorname{Kern} A = \{x : Ax = 0\}.$$

Die Rangformel für $A \in \mathbb{K}^{m \times n}$ ist dann

$$n = \dim \operatorname{Kern} A + \operatorname{rang} A. \tag{6.2}$$

Lemma 6.11 *Es gilt* $\operatorname{rang} A \leq \min\{m, n\}$.

Beweis: Für $m \geq n$ ist das richtig, wir haben ja nur n Spalten zur Verfügung. Andererseits ist der Bildraum ein Teilraum des $\mathbb{K}^m$. Dort kann es höchstens m linear unabhängige Vektoren geben. □

Eine quadratische Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann regulär, wenn $\operatorname{rang} A = n$. Nur in diesem Fall ist die zugehörige Selbstabbildung surjektiv und damit bijektiv.

Beispiel 6.12 Wir untersuchen die folgenden Matrizen in $\mathbb{K} = \mathbb{C}$,

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & i & 1 \\ 0 & 1 & -i \\ 1 & 0 & 0 \end{pmatrix}.$$

In der Matrix A sind die Spaltenvektoren a_1 und a_3 l.u.. Ferner ist $a_4 = a_1 - a_3$. Damit gilt $\operatorname{Bild} A = \operatorname{span}\{a_1, a_3\}$. Nach der Rangformel ist $\dim \operatorname{Kern} A = 2$. Als Basis des Kerns können wir die Vektoren $e_2, (1, 0, -1, -1)^T \in \mathbb{C}^4$ nehmen.

In der Matrix B gilt $b_2 = ib_3$, damit ist $\operatorname{Bild} A = \operatorname{span}\{b_1, b_2\}$. Der Kern ist daher eindimensional, offenbar ist $(0, 1, -i)^T \in \operatorname{Kern} B$. □

Satz 6.13 (a) *Für Matrizen* $A : \mathbb{K}^{l \times m}$, $B : \mathbb{K}^{m \times n}$ *gilt*

$$\operatorname{rang} AB \leq \min\{\operatorname{rang} A, \operatorname{rang} B\}.$$

(b) *Sei* $A : \mathbb{K}^{m \times n}$, $B : \mathbb{K}^{m \times m}$, $C : \mathbb{K}^{n \times n}$ *mit* $\operatorname{rang} B = m$, $\operatorname{rang} C = n$. *Dann gilt*

$$\operatorname{rang} BA = \operatorname{rang} A, \quad \operatorname{rang} AC = \operatorname{rang} A.$$

Beweis: (a) Ist $\operatorname{rang} A \leq \operatorname{rang} B$, so gilt $\operatorname{Bild} AB \subset \operatorname{Bild} A$. Ist umgekehrt $\operatorname{rang} B \leq \operatorname{rang} A$, so bildet A das Bild von B auf einen Unterraum der Dimension $\leq \operatorname{rang} B$ ab.

(b) Wir können die Matrizen als Darstellungsmatrizen der zugehörigen linearen Abbildungen interpretieren. Dann sind B und C Isomorphismen zwischen den angegebenen Räumen. B bildet das Bild von A auf einen Unterraum gleicher Dimension ab. Das Bild von C spannt den ganzen $\mathbb{K}^n$ auf. In diesem Fall gilt sogar $\operatorname{Bild} A = \operatorname{Bild} AC$. □

Beispiel 6.10 zeigt, dass man in (a) nicht mehr zeigen kann. In diesem Beispiel haben die Matrizen A und B jeweils den Rang 1 und für die Produkte gilt $\operatorname{rang} AB = 1$, aber $\operatorname{rang} BA = 0$.

Für eine Matrix $A = (a_{ij}) \in \mathbb{K}^{m \times n}$ ist die *transponierte Matrix* $A^T = (a^T_{ij}) \in \mathbb{K}^{n \times m}$ definiert durch $a^T_{ij} = a_{ji}$ für alle $1 \leq i \leq m$, $1 \leq j \leq n$. Anschaulich klappt man die Matrix A von links unten nach rechts oben:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}.$$

Sind A, B Matrizen über $\mathbb{K}$, für die das Produkt AB erklärt ist, so gilt

$$(AB)^T = B^T A^T.$$

Mit $C = AB$ haben wir nämlich

$$c_{ij} = \sum_k a_{ik} b_{kj} \;\Rightarrow\; c^T_{ij} = c_{ji} = \sum_k a_{jk} b_{ki} = \sum_k b^T_{ik} a^T_{kj} = (B^T A^T)_{ij}.$$

Eine ähnliche Formel gilt für die Inverse des Produkts zweier regulärer Matrizen $A, B \in \mathbb{K}^{n \times n}$:

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

wegen

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

Für eine reguläre Matrix $A \in \mathbb{K}^{n \times n}$ ist es gleichgültig, ob man zuerst transponiert und dann invertiert oder umgekehrt:

$$(A^{-1})^T = (A^T)^{-1}$$

wegen

$$(A^{-1})^T A^T = (A A^{-1})^T = E_n,$$

also ist A^T die Inverse von $(A^{-1})^T$. Wir schreiben daher auch A^{-T} anstatt $(A^{-1})^T$ oder $(A^T)^{-1}$.