Kapitel 3

Lineare Abbildungen

3.1 Definition, Beispiele, Eigenschaften

Einleitung: [Koe97, §6, S. 35]: Seit Erwachen der Analysis im 17. Jahrhundert haben die Mathematiker immer wieder nach Funktionen gesucht, die sich durch besonders schöne Eigenschaften auszeichnen. Eine interessante Klasse solcher Eigenschaften fasst man unter dem Stichwort „Funktionalgleichungen" zusammen. Für Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ sind die folgenden Funktionalgleichungen denkbar:

$$
\begin{aligned}
f(x+1) &= f(x), \\
f(xy) &= f(x)f(y), \\
f(x+y) &= f(x)f(y), \\
f(x+y) &= f(x)+f(y).
\end{aligned}
\tag{3.1}
$$

Die erste Funktionalgleichung wird von allen *periodischen* Funktionen mit der Periode 1 erfüllt; die zweite Gleichung von Potenzfunktionen $f(x) = x^{\alpha}$; in der dritten Klasse liegen die Exponentialfunktionen $f(x) = a^x$ und schließlich wird die vierte Klasse von den linearen Funktionen $f(x) = cx$ erfüllt. Doch, sind dies alle Lösungen? Bereits Augustin Louis Cauchy (1789–1857) zeigte, dass alle *stetigen* Lösungen der Gleichung (3.1) die Form $f(x) = cx$ mit einer Konstanten $c \in \mathbb{R}$ haben. Wir werden sehen, dass dies nicht alle Lösungen sind.

Definition 3.1 Es seien V und W Vektorräume über $\mathbb{K}$. Eine $\mathbb{K}$*-lineare Abbildung* ist eine Abbildung $T\colon V \to W$, so dass für alle $v_1, v_2 \in V$ und $\lambda \in \mathbb{K}$ gilt

$$T(v_1+v_2) = T(v_1)+T(v_2) \quad \text{(Additivität)} \quad \text{und} \quad T(\lambda v_1) = \lambda T(v_1) \quad \text{(Homogenität)}. \tag{3.2}$$

Die beiden Bedingungen lassen sich zu einer Bedingung zusammenfassen, die für alle $v_1, v_2 \in V$ und alle $\lambda_1, \lambda_2 \in \mathbb{K}$ gelten muss: $T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2)$. Wir bezeichnen die Menge der $\mathbb{K}$-linearen Abbildungen von V nach W mit $\mathrm{L}_{\mathbb{K}}(V, W)$ oder $\mathrm{L}(V, W)$ oder $\mathrm{Hom}_{\mathbb{K}}(V, W)$.

Wenn der Grundkörper $\mathbb{K}$ eindeutig fixiert ist, sprechen wir einfach von *linearen Abbildungen.* Andere Bezeichnungen für eine lineare Abbildung sind *linearer Operator* oder *lineare Transformation* oder *Homomorphismus* von Vektorräumen.

Eine lineare Abbildung $T \in \mathrm{L}(V, W)$ heißt

> *Monomorphismus*, wenn sie injektiv ist, *Epimorphismus*, wenn sie surjektiv ist, *Isomorphismus*, falls sie bijektiv ist, *Endomorphismus*, wenn $V = W$ und *Automorphismus*, wenn sie bijektiv ist und $V = W$.

Die Menge der Endomorphismen bezeichnen wir mit $\mathrm{L}(V)$, die Menge der Automorphismen $T\colon V \to V$ mit $\mathrm{GL}(V)$.

Bemerkung 3.1 (a) $\mathrm{L}(V, W) \subset \mathrm{Abb}(V, W)$ ist ein Untervektorraum von $\mathrm{Abb}(V, W)$. In der Tat ist $\mathrm{L}(V, W)$ nichtleer, da die Nullabbildung linear ist $(0(\lambda_1 v_1 + \lambda_2 v_2) = 0 = \lambda_1 0(v_1) + \lambda_2 0(v_2))$ und stets zu $\mathrm{L}(V, W)$ gehört. Seien nun $S, T \in \mathrm{L}(V, W)$; wir müssen zeigen, dass dann auch $S + T$ und λS linear sind. In der Tat gilt für alle $v_1, v_2 \in V$, $\lambda_1, \lambda_2 \in \mathbb{K}$

$$\begin{aligned}(S+T)(\lambda_1 v_1 + \lambda_2 v_2) &= S(\lambda_1 v_1 + \lambda_2 v_2) + T(\lambda_1 v_1 + \lambda_2 v_2)\\ &= \lambda_1 S(v_1) + \lambda_2 S(v_2) + (\lambda_1 T(v_1) + \lambda_2 T(v_2)) = \lambda_1 (S+T)(v_1) + \lambda_2 (S+T)(v_2).\end{aligned}$$

Somit ist $S + T$ linear. Außerdem gilt wegen der Linearität von S

$$(\lambda S)(\lambda_1 v_1 + \lambda_2 v_2) = \lambda(S(\lambda_1 v_1 + \lambda_2 v_2)) = \lambda\lambda_1 S(v_1) + \lambda\lambda_2 S(v_2) = \lambda_1(\lambda S)(v_1) + \lambda_2(\lambda S)(v_2).$$

Also ist λS linear. Damit sind Linearkombinationen von linearen Abbildungen linear; $\mathrm{L}(V, W)$ ist ein linearer Teilraum.
(b) Sind $U \xrightarrow{f} V \xrightarrow{g} W$ lineare Abbildungen, so auch $g \circ f\colon U \to W$. In der Tat ist

$$g(f(\lambda v_1 + v_2)) = g(\lambda f(v_1) + f(v_2)) = \lambda g(f(v_1)) + g(f(v_2)) = \lambda g \circ f(v_1) + g \circ f(v_2).$$

Die Identität $\mathrm{id}_V\colon V \to V$ ist stets linear.

Beispiel 3.1 (a1) **Die lineare Abbildung einer Matrix.** Es sei $A \in \mathbb{K}^{m\times n}$, dann definieren wir $T_A \in \mathrm{L}(\mathbb{K}^n, \mathbb{K}^m)$ für $x = (x_1, \ldots, x_n)^\top$ über $T_A(x) := A \cdot x$. Das Ergebnis ist ein Spaltenvektor in $\mathbb{K}^m$. In der Tat ist T_A linear, denn wegen der Distributivität der Matrixmultiplikation ist $T_A(\lambda x + \mu y) = A(\lambda x + \mu y) = \lambda A x + \mu A y = \lambda T_A(x) + \mu T_A(y)$ für alle $x, y \in \mathbb{K}^n$, $\lambda, \mu \in \mathbb{K}$.

Die Bilder der Einheitsspaltenvektoren e_j, $j = 1, \ldots, n$ unter T_A sind genau die Spalten der Matrix A:

$$(A\mathrm{e}_j)_i = \sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{ij}, \qquad A\mathrm{e}_j = A \cdot \begin{pmatrix} 0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} = \begin{pmatrix} a_{1j}\\ a_{2j}\\ \vdots\\ a_{nj}\end{pmatrix} = \sum_{i=1}^{m} a_{ij}\mathrm{e}_i.$$

(a2) **Die Komposition.** Es seien $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ und $B = (b_{jk}) \in \mathbb{K}^{n\times p}$ verkettete Matrizen. Dann ist $AB \in \mathbb{K}^{m\times p}$ und es gilt

$$T_A \circ T_B = T_{AB}$$

als lineare Abbildungen von $\mathbb{K}^p \to \mathbb{K}^m$.
Beweis. Es sei $x = (x_1, \ldots, x_p)^\top \in \mathbb{K}^p$. Dann gilt

$$(T_A \circ T_B)(x) = T_A(T_B(x)) = T_A\left(\left(\sum_{k=1}^{p} b_{jk} x_k\right)_{j=1\cdots n}\right) = \left(\sum_{j=1}^{n} a_{ij} \sum_{k=1}^{p} b_{jk} x_k\right)_{i=1\cdots m}$$
$$= \left(\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) x_k\right)_{i=1\cdots m} = \left(\sum_{k=1}^{p} (AB)_{ik} x_k\right)_{i=1\cdots m} = f_{AB}(x).$$

■

In diesem Beweis wurde die Assoziativität der Matrixmultiplikation faktisch noch einmal für einen Spezialfall gezeigt. Wenn man die Assoziativität benutzt, so lautet der kurze Beweis:

$$(T_A \circ T_B)(x) = T_A(T_B(x)) = A \cdot (B \cdot x) = (A \cdot B) \cdot x = T_{AB}(x).$$

(b) **Differentiation.** Es sei $V = C^1(\mathbb{R})$ der Vektorraum der stetig differenzierbaren Funktionen und $W = C(\mathbb{R})$ die stetigen Funktionen auf $\mathbb{R}$. Dann ist $D\colon V \to W$, $f \mapsto f'$, linear, denn $(f+g)' = f' + g'$ und $(\lambda f)' = \lambda f'$.
(c) **Integration.** Es sei $V = C(\mathbb{R})$, $W = \mathbb{R}$ und $a < b$, $a, b \in \mathbb{R}$. Dann ist die Abbildung $V \to W$, $f \mapsto \int_a^b f(x)\,\mathrm{d}x$, linear.
(d) **Evaluation.** Es sei V ein reeller Vektorraum, X eine beliebige Menge, $\alpha \in X$ und $W = \mathrm{Abb}(X, V)$. Dann ist die Abbildung $W \to V$, $f \mapsto f(\alpha)$, linear.
(d) **Grenzwerte.** Es sei c der Vektorraum der konvergenten reellen Zahlenfolgen. Dann ist $\lim\colon c \to \mathbb{R}$, $(x_n) \mapsto \lim_{n\to\infty} x_n$, eine lineare Abbildung, denn für konvergente Folgen (x_n) und (y_n) gilt der Grenzwertsatz $\lim_{n\to\infty}(\lambda_1 x_n + \lambda_2 y_n) = \lambda_1 \lim_{n\to\infty} x_n + \lambda_2 \lim_{n\to\infty} y_n$.

Lemma 3.1 *Es sei $T\colon V \to W$ linear. Dann gilt:*
(a) $T(0) = 0$, $T(-v) = -T(v)$, $T(v - w) = T(v) - T(w)$ *für alle* $v, w \in V$.
(b) $T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 T(v_1) + \cdots + \lambda_n T(v_n)$ *für alle* $v_i \in V$, $\lambda_i \in \mathbb{K}$.
(c) *Ist für eine Familie* $\{v_i \mid i \in I\}$ *die Menge* $\{T(v_i) \mid i \in I\}$ *linear unabhängig in* W*, so ist* $\{v_i \mid i \in I\}$ *linear unabhängig in* V.
(d) *Für lineare Teilräume* $V' \subset V$ *und* $W' \subset W$ *sind Bild* $T(V') \subset W$ *und Urbild* $T^{-1}(W') \subset V$ *lineare Teilräume.*
(e) *Es gilt* $\dim T(V) \le \dim V$.
(f) *Ist T ein Isomorphismus, so ist auch T^{-1} linear (und ein Isomorphismus).*

Beweis. (a) Wegen $(-1)v = -v$ und der Linearität von T gilt, $T(v - w) = T(v + (-1)w) = T(v) + (-1)T(w) = T(v) - T(w)$. Setzt man hier $v = w = 0$ ein, so hat man $T(0) = T(0-0) = T(0) - T(0) = 0$; setzt man $v = 0$ ein, so hat man $T(-w) = T(0 - w) = T(0) - T(w) = 0 - T(w) = -T(w)$.
(b) folgt durch vollständige Induktion über n.
(c) Sei $\sum_{k=1}^n \alpha_k v_k = 0$. Nach $T(0) = 0$ und (b) folgt durch Anwenden von T,

$$T(0) = 0 = T\left(\sum_{k=1}^{n} \alpha_k v_k\right) = \sum_{k=1}^{n} \alpha_k T(v_k).$$

Wegen der linearen Unabhängigkeit von $\{T(v_k) \mid k = 1,\dots,n\}$ folgt $\alpha_1 = \cdots = \alpha_n = 0$; also waren die $\{v_i\}$ bereits linear unabhängig.
(d) Wegen $0 \in V'$, $0 \in W'$ und $T(0) = 0$ sind die Mengen $T(V')$ und $T^{-1}(W')$ nichtleer. Wir zeigen, dass das Unterraumkriterium Lemma 2.2 für $T(v')$ erfüllt ist. Seien dazu $w_1, w_2 \in T(V')$, etwa $w_1 = T(v_1)$ und $w_2 = T(v_2)$ mit $v_1, v_2 \in V'$. Dann ist

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 T(v_1) + \lambda_2 T(v_2) = T(\lambda_1 v_1 + \lambda_2 v_2) = T(v_3),$$

dabei ist $v_3 = \lambda_1 v_1 + \lambda_2 v_2 \in V'$ wegen der Abgeschlossenheit der linearen Operationen in V'. Also $\lambda_1 w_1 + \lambda_2 w_2 \in T(V')$. Der Beweis des Unterraumkriteriums für $T^{-1}(W')$ verläuft analog.
(e) Sind $T(v_1),\dots,T(v_n)$ linear unabhängig in W, so sind nach (c) $v_1,\dots,v_n$ linear unabhängig in V; also $n \le \dim V$. Dies gilt insbesondere für eine Basis $\{T(v_1),\dots,T(v_n)\}$ von $T(V)$.
(f) Es seien $w, w' \in W$. Da T bijektiv ist, existieren eindeutig bestimmte Vektoren $v, v' \in V$ mit $w = T(v)$, $w' = T(v')$, so dass $v = T^{-1}(w)$, $v' = T^{-1}(w')$. Dann gilt $T(\lambda v + \mu v') = \lambda w + \mu w'$. Wendet man darauf T^{-1} an, so hat man $\lambda T^{-1}(w) + \mu T^{-1}(w') = T^{-1}(\lambda w + \mu w')$; also ist T^{-1} linear. ■

Lemma 3.2 (Existenz- und Eindeutigkeitssatz) *Es seien V und W Vektorräume und $B \subset V$ eine Basis von V. Ferner sei eine Abbildung $T_0\colon B \to W$ gegeben.*
Dann lässt sich T_0 eindeutig zu einer linearen Abbildung $T\colon V \to W$ fortsetzen. Das heißt, es gibt eine lineare Abbildung $T\colon V \to W$ mit $T(b) = T_0(b)$ für alle $b \in B$; und wenn für eine lineare Abbildung $S\colon V \to W$ gilt, dass $S(b) = T_0(b)$ für alle $b \in B$, dann ist $S = T$.

Beweis. (a) Es sei $v \in V$. Da B eine Basis ist, existieren $n \in \mathbb{N}$, $\alpha_k \in \mathbb{K}$ und $b_k \in B$ für $k = 1,\dots,n$, sodass $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$. Wir definieren

$$T(v) := \sum_{k=1}^{n} \alpha_k T_0(b_k).$$

Wegen der Eindeutigkeit der Linearkombination für v, ist $T(v)$ wohldefiniert; andererseits ist T nach Konstruktion linear, denn sei $w = \sum_{k=1}^n \beta_k b_k$, so ist $v + w = \sum_{k=1}^n (\alpha_k + \beta_k) b_k$ und $T(w) = \sum_k \beta_k T_0(b_k)$ und daher

$$T(v+w) = \sum_{k=1}^{n} (\alpha_k + \beta_k) T_0(b_k) = \sum_k \alpha_k T_0(b_k) + \sum_k \beta_k T_0(b_k) = T(v) + T(w).$$

Außerdem ist $\lambda v = \sum_k \lambda \alpha_k b_k$ und damit

$$T(\lambda v) = \sum_k \lambda \alpha_k T_0(b_k) = \lambda \sum_k \alpha_k T_0(b_k) = \lambda T(v).$$

Somit ist T linear.
(b) Eindeutigkeit. Es sei auch $S\colon V \to W$ eine lineare Fortsetzung von T_0. Dann gilt für alle $v \in V$ nach Konstruktion von T und wegen der Linearität von S

$$\begin{aligned} S(v) - T(v) = S\left(\sum_k \alpha_k b_k\right) - T\left(\sum_k \alpha_k b_k\right) &= \\ = \sum_k \alpha_k S(b_k) - \sum_k \alpha_k T(b_k) = \sum_k \alpha_k T_0(b_k) - \sum_k \alpha_k T_0(b_k) &= 0. \end{aligned}$$

■

Beispiel 3.2 Es sei V ein $\mathbb{K}$-Vektorraum und $B=\{b_1,\dots,b_n\}$ eine Basis von V. Dann gibt es nach Lemma 3.2 genau eine lineare Abbildung

$$\Phi_B\colon V\to\mathbb{K}^n,\quad \Phi_B(b_k)=\mathrm{e}_k,\quad k=1,\dots,n.$$

Offenbar ordnet diese Abbildung jedem Vektor $v=\alpha_1b_1+\alpha_2b_2+\cdots+\alpha_nb_n$ seine Koordinaten $(\alpha_1,\dots,\alpha_n)$ bezüglich der gegebenen Basis B zu. Diese Abbildung heißt *Koordinatenabbildung* von V bezüglich B. Die Koordinatenabbildung Φ_B ist bijektiv. Ihre Umkehrabbildung $\Phi_B^{-1}\colon\mathbb{K}^n\to V$ ist eindeutig bestimmt durch $\Phi_B^{-1}(\mathrm{e}_k)=b_k$, $k=1,\dots,n$.

3.1.1 Die Matrix einer linearen Abbildung

Mit Hilfe von Koordinatenabbildungen lässt sich zu jeder linearen Abbildung T unter Auszeichnung von Basen im Ausgangsraum und im Bildraum eine Matrix zuordnen.

Definition 3.2 Gegeben sei eine lineare Abbildung $S\colon V\to W$ und Basen $B=\{b_1,\dots,b_n\}$ und $C=\{c_1,\dots,c_m\}$ von V bzw. von W. Dann gibt es genau eine Matrix $A\in\mathbb{K}^{m\times n}$, so dass das folgende Diagramm kommutativ ist

$$\begin{tikzcd} V \arrow[r, "S"] \arrow[d, "\Phi_B"'] & W \arrow[d, "\Phi_C"] \\ \mathbb{K}^n \arrow[r, "T_A"] & \mathbb{K}^m \end{tikzcd}$$

Hierbei ist T_A die gemäß Beispiel 3.1 (a1) der Matrix A zugeordnete lineare Abbildung $T_A\colon\mathbb{K}^n\to\mathbb{K}^m$. Es ist $T_A=\Phi_C\circ S\circ\Phi_B^{-1}$. Dabei gilt mit $A=(a_{ij})_{\substack{i=1\dots m\\ j=1\dots n}}$

$$\boxed{S(b_j)=\sum_{i=1}^m a_{ij}c_i,\quad j=1,\dots,n.}$$

In der Tat ist ja umgekehrt $S=\Phi_C^{-1}\circ T_A\circ\Phi_B$ und somit

$$S(b_j)=\Phi_C^{-1}(T_A(\Phi_B(b_j)))=\Phi_C^{-1}(A\cdot\mathrm{e}_j)=\Phi_C^{-1}\left(\sum_{i=1}^m a_{ij}\mathrm{e}_i\right)=\sum_{i=1}^m a_{ij}\Phi_C^{-1}(\mathrm{e}_i)=\sum_{i=1}^m a_{ij}c_i.$$

Wir bezeichnen die Matrix von S bezüglich B und C mit $A=M_{B,C}(S)$.

Beispiel 3.3 (a) Es sei $c\in\mathbb{K}$ fixiert und $S\colon V\to V$ gegeben durch $S(v)=cv$. Ferner sei $B=\{b_1,\dots,b_n\}$ eine Basis von V. Dann gilt S ist linear und die Matrix von S bezüglich der Basis B lautet:

$$M_{B,B}(S)=\begin{pmatrix} c & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & c\end{pmatrix}=cI_n.$$

Dies ist das c-fache der Einheitsmatrix I_n. Es gilt nämlich $S(b_i) = cb_i$, $i = 1, \dots, n$.
(b) Es sei B_n die Standardbasis von $\mathbb{K}^n$, $A \in \mathbb{K}^{m\times n}$. Dann gilt $A = M_{B_n,B_m}(T_A)$. Das heißt, dass die Matrix zur linearen Abbildung T_A bezüglich der Standardbasis genau die Matrix A selbst ist. Dies bedeutet, dass es eine bijektive Beziehung zwischen den Matrizen $\mathbb{K}^{m\times n}$ und den linearen Abbildungen $\mathrm{L}(\mathbb{K}^n, \mathbb{K}^m)$ gibt. In der einen Richtung ist sie gegeben durch die Zuordnung $A \mapsto T_A$ und in der anderen Richtung durch $S \mapsto M_{B_n,B_m}(S)$. Diese beiden Zuordnungen sind sogar zueinander inverse *lineare* Abbildungen zwischen Matrizen und linearen Abbildungen. Da bei linearen Isomorphismen Basen in Basen über gehen, siehe Lemma 3.5 unten, sind die Dimensionen beider Räume gleich groß

$$\dim \mathrm{L}(\mathbb{K}^n, \mathbb{K}^m) = \dim \mathbb{K}^{m\times n} = mn.$$

(c) Es sei $P\colon \mathbb{R}^3 \to \mathbb{R}^2$ gegeben durch $P(x_1, x_2, x_3) = (2x_1 - x_3, x_1 + 2x_2 + 3x_3)$. Dann gilt

$$A := M_{B_3,B_2}(P) = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$$

und $P = T_A$.
(d) Es sei $W = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ und $V = \mathbb{R}_2[x]$; $a \in \mathbb{R}$ sei fixiert. Die *Verschiebung* V_a (auch Translation oder Shift genannt), gegeben durch $V_a\colon W \to W$, $(V_a f)(x) = f(x-a)$, $f \in W$, ist eine bijektive lineare Abbildung. Wir betrachten die Einschränkung von V_a auf V und berechnen die Matrix von V_a bezüglich der Basen $B = \{1, x, x^2\}$ im Ausgangsraum und $C = \{1, 1+x, 1+x+x^2\}$ im Bildraum. Es ist

$$\begin{aligned}
V_a(1) &= 1,\\
V_a(x) &= x - a = (x+1) + (-a+1)1,\\
V_a(x^2) &= (x-a)^2 = x^2 - 2ax + a^2 = (x^2 + x + 1) + (-2a-1)(x+1) + (a^2 + 2a)
\end{aligned}$$

Hieraus kann man die Matrixeinträge von $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ spaltenweise sofort ablesen:

$$M_{B,C}(V_a) = A = \begin{pmatrix} 1 & 1-a & a^2+2a \\ 0 & 1 & -2a-1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Man beachte, dass $M_{B,C}(T)$ nicht nur von den Basen B und C abhängt, sondern auch von der Reihenfolge der Vektoren in der Basis. Wir betrachten also hier immer *geordnete* Basen.

3.2 Kern und Bild

3.2.1 Definition und Eigenschaften

Definition 3.3 Es sei $T\colon V \to W$ eine lineare Abbildung. Dann nennt man

$\operatorname{Im} T := T(V)$ das *Bild* von T,
$\operatorname{Ker} T := T^{-1}(\{0\})$ den *Kern* von T.

Als *Rang* von T bezeichnet man die Zahl $\operatorname{rg} T := \dim \operatorname{im} T = \dim T(V)$; als *Defekt* von T bezeichnet man die Zahl $\operatorname{def} T := \dim \operatorname{Ker} T$.

In der Tat sind nach Lemma 3.1 (d) $\operatorname{Im} T$ und $\operatorname{Ker} T$ lineare Teilräume von W bzw. von V, da V und $\{0\}$ Teilräume sind. Daher hat es Sinn, von deren Dimensionen zu reden.

Lemma 3.3 *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung. Dann sind äquivalent:*

(a) T ist surjektiv.
(b) Ist B eine Basis von V, so ist $T(B)$ ein Erzeugendensystem von W.

Ist darüber hinaus W endlichdimensional, so ist mit (a) und (b) gleichwertig

(c) $\operatorname{rg} T = \dim W$.

Beweis. (a) $\iff$ (b). Wegen der Linearität von T ist

$$\operatorname{span} T(B) = T(\operatorname{span} B) = T(V).$$

Ist T surjektiv, so ist $T(V) = W$ und $T(B)$ ist ein Erzeugendensystem für W. Ist umgekehrt $\operatorname{span} T(B) = W$, so ist T surjektiv. Ist nun $\dim W$ endlich, so folgt aus der Surjektivität von T, $\dim T(V) = \dim W$, also $\operatorname{rg} T = \dim W$. Ist umgekehrt $\dim W = \dim T(V)$, so folgt aus $T(V) \subset W$ und Lemma 2.8 (c) $T(V) = W$, also ist T surjektiv. ∎

Lemma 3.4 *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung. Dann sind äquivalent:*

(a) T ist injektiv.
(b) $\operatorname{Ker} T = \{0\}$.
(c) $\operatorname{def} T = 0$.
(d) Ist B eine Basis von V, so ist $T(B)$ linear unabhängig.

Ist V endlichdimensional, so ist außerdem gleichwertig

(e) $\operatorname{rg} T = \dim V$.

Beweis. Die Gleichwertigkeit von (b) und (c) ist klar, da der einzige nulldimensionale Raum der nur aus dem Nullvektor bestehende Raum $\{0\}$ ist.
(a) $\Longrightarrow$ (b): Wegen der Injektivität von T folgt aus $T(x) = T(0) = 0$ sofort $x = 0$; also ist $\operatorname{Ker} T = \{0\}$.
(b) $\Longrightarrow$ (d): Es sei

$$\alpha_1 T(b_1) + \cdots + \alpha_n T(b_n) = 0$$

eine Linearkombination von $T(B)$ in W. Dann gilt $T(\alpha_1 b_1 + \cdots + \alpha_n b_n) = 0$ und wegen $\operatorname{Ker} T = \{0\}$ folgt weiter $\alpha_1 b_1 + \cdots + \alpha_n b_n = 0$. Da aber B linear unabhängig in V war, ist $\alpha_1 = \cdots = \alpha_n = 0$. Somit ist $T(B)$ linear unabhängig.

(d) $\Longrightarrow$ (a): Es sei nun B eine Basis von V und $v, v' \in V$ mögen die Darstellungen $v = \sum_{i=1}^n \alpha_i b_i$ und $v' = \sum_{i=1}^n \beta_i b_i$ mit $b_i \in B$ besitzen. Wenn nun $T(v) = T(v')$ gilt, so folgt

$$T\left(\sum_i \alpha_i b_i\right) = T\left(\sum_i \beta_i b_i\right) \Longrightarrow \sum_i \alpha_i T(b_i) = \sum_i \beta_i T(b_i).$$

Nun ist aber nach der Annahme in (d) auch $T(B)$ eine Basis in W. Nach Lemma 2.4 (b) ist aber die Koordinatendarstellung bezüglich einer Basis eindeutig; folglich gilt $\alpha_i = \beta_i$ für alle $i = 1, \dots, n$.

(d) $\Longrightarrow$ (e): Sei nun V endlichdimensional mit einer Basis $B = \{b_1, \dots, b_n\}$. Nach (d) ist dann auch $\{T(b_1), \dots, T(b_n)\}$ linear unabhängig in W, also $\dim T(V) = \operatorname{rg} T \geq n$. Umgekehrt gilt nach Lemma 3.1 (e) $\operatorname{rg} T \leq n$. Folglich ist $\operatorname{rg} T = n$.

(e) $\Longrightarrow$ (b): Sei $\operatorname{rg} T = n = \dim V$ und $w_1 = T(b_1), \dots, w_n = T(b_n)$ sei eine Basis von $\operatorname{Im}(T)$. Nach Lemma 3.1 (c) ist dann $B := \{b_1, \dots, b_n\}$ linear unabhängig in V. Wegen $n = \dim V$ ist B sogar Basis in V. Sei nun $T(v) = 0$ für ein $v \in V$, etwa für $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$. Wendet man T darauf an und benutzt die Linearität von T und $T(0) = 0$, so folgt

$$\alpha_1 T(b_1) + \cdots + \alpha_n T(b_n) = 0.$$

Da aber $\{T(b_1), \dots, T(b_n)\}$ Basis in $T(V)$ ist, folgt $\alpha_1 = \cdots = \alpha_n = 0$. Somit ist $v = 0$ und $\operatorname{Ker} T = \{0\}$. ■

Wir fassen die Aussagen von Lemma 3.3 und Lemma 3.4 zusammen und erhalten eine Charakterisierung der bijektiven linearen Abbildungen (Isomorphismen).

Lemma 3.5 *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung. Dann sind die folgenden Aussagen äquivalent:*

(a) T ist ein Isomorphismus.
(b) Wenn B eine Basis von V ist, so ist $T(B)$ eine Basis von W.

Besitzen darüber hinaus V und W endliche Dimension, so ist außerdem äquivalent dazu

(c) $\dim V = \operatorname{rg} T = \dim W$.

Definition 3.4 Zwei Vektorräume V und W über $\mathbb{K}$ heißen *isomorph*, symbolisch $V \cong W$, wenn es einen Isomorphismus $T\colon V \to W$ gibt.

Folgerung 3.6 *Isomorphe Vektorräume besitzen dieselbe Dimension. Gilt umgekehrt* $\dim V = \dim W < \infty$*, so sind V und W isomorph. Insbesondere sind alle Vektorräume V mit* $\dim V = n$ *isomorph zu $\mathbb{K}^n$.*

Beweis. Der erste Teil folgt direkt aus dem obigen Lemma. Seien nun V und W endlichdimensionale Vektorräume mit derselben Dimension n und seien etwa $B = \{b_1, \dots, b_n\}$ und $C = \{c_1, \dots, c_n\}$ Basen in V bzw. in W. Dann lässt sich die Abbildung $T_0\colon B \to W$, $T(b_k) = c_k$ eindeutig zu einem Isomorphismus $T\colon V \to W$ fortsetzen. Mit Hilfe der Koordinatenabbildungen lässt sich T schreiben als $T = \Phi_C^{-1} \circ \Phi_B$. ■

Unendlichdimensionale Räume sind nicht notwendig isomorph.

3.2.2 Die Dimensionsformel für lineare Abbildungen

Ist $T \in \mathrm{L}(V, W)$, so kann man die Zahlen $\operatorname{def} T$ und $\operatorname{rg} T$ betrachten. Erstaunlicherweise hängt $\operatorname{def} T + \operatorname{rg} T$ gar nicht von T ab.

Satz 3.7 (Dimensionsformel für lineare Abbildungen) *Es sei V ein endlichdimensionaler Vektorraum. Für jede lineare Abbildung $T\colon V \to W$ gilt dann:*

$$\boxed{\operatorname{rg} T + \operatorname{def} T = \dim V.}$$

Beweis. Als Teilraum von V besitzt $\operatorname{Ker} T$ auch eine endliche Dimension; sei etwa $K = \{b_1, \dots, b_d\}$ eine Basis von $\operatorname{Ker} T$. Nach Lemma 2.8 kann man K zu einer Basis $\{b_1, \dots, b_d, b_{d+1}, \dots, b_n\}$ von V ergänzen. Wir werden zeigen, dass $E := \{T(b_{d+1}), \dots, T(b_n)\}$ eine Basis von $\operatorname{Im} T$ ist.
E ist erzeugend für $\operatorname{im} T$. Es sei $T(v) \in \operatorname{im} T$ mit $v = \sum_{i=1}^n \alpha_i b_i$. Wegen der Linearität von T und wegen $T(b_1) = T(b_2) = \dots = T(b_d) = 0$ folgt

$$T(v) = T\left(\sum_{i=1}^n \alpha_i b_i\right) = \sum_{i=1}^n \alpha_i T(b_i) = \sum_{i=d+1}^n \alpha_i T(b_i) \in \operatorname{span} E.$$

E ist linear unabhängig. Sei

$$\sum_{i=d+1}^n \alpha_i T(b_i) = 0 \Longrightarrow T\left(\sum_{i=d+1}^n \alpha_i b_i\right) = 0 \Longrightarrow \sum_{i=d+1}^n \alpha_i b_i \in \operatorname{Ker} T.$$

Folglich existieren $\alpha_i \in \mathbb{K}$, $i = 1, \dots, d$ mit

$$\sum_{i=d+1}^n \alpha_i b_i = \sum_{i=1}^d \alpha_i b_i.$$

Da aber $\{b_1, \dots, b_n\}$ eine Basis von V ist, folgt hieraus $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. Also ist E linear unabhängig und damit eine Basis von $\operatorname{Im} T$. Somit gilt

$$\operatorname{rg} T = \dim \operatorname{Im} T = \#\{b_{d+1}, \dots, b_n\} = n - d = \dim V - \operatorname{def} T.$$

∎

Satz 3.8 (Äquivalenzsatz) *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen gleicher Dimension,* $\dim V = \dim W < \infty$. *Dann sind äquivalent:*

(a) T ist injektiv.
(b) T ist surjektiv.
(c) T ist bijektiv.

Beweis. (a) $\rightarrow$ (b). Nach Lemma 3.4(e) gilt $\operatorname{rg} T = \dim V$, also nach Voraussetzung auch $\operatorname{rg} T = \dim W$. Nach Lemma 3.3 folgt dann T ist surjektiv.
(b) $\rightarrow$ (c). Nach den beiden Lemmas und nach Voraussetzung ist wieder $\operatorname{rg} T = \dim W = \dim V$, also ist T injektiv und damit bijektiv. Schließlich ist die Richtung (c) $\rightarrow$ (a) trivial. ∎

Der Satz gilt nicht mehr in unendlich-dimensionalen Räumen, wie das folgende Beispiel des *Verschiebungsoperators* $T\colon \omega \to \omega$ zeigt: $T(x) := T((x_1, x_2, \dots)) := (0, x_1, x_2, \dots)$. Offenbar ist T injektiv aber nicht surjektiv.

3.3 Matrizenrechnung

3.3.1 Rang und Defekt einer Matrix

Wir haben bereits in Kapitel 1.2 gesehen, wie Matrizen addiert, multipliziert und skalar vervielfacht werden. Unter dem Gesichtspunkt, dass wir Matrizen $A \in \mathbb{K}^{m\times n}$ mit ihren linearen Abbildungen $T_A \in \mathrm{L}(\mathbb{K}^n, \mathbb{K}^m)$ identifizieren, können wir die eben eingeführten Begriffe, Kern, Bild, Rang, Defekt aber auch Invertierbarkeit direkt auf Matrizen übertragen.

Definition 3.5 Für eine Matrix $A \in \mathbb{K}^{m\times n}$ definieren wir
(a) den *Kern* von A $\operatorname{Ker} A = \operatorname{Ker} T_A$, das *Bild* von A, $\operatorname{im} A = \operatorname{im} T_A$, den *Rang* $\operatorname{rg} A = \operatorname{rg} T_A$ und den Defekt von A über $\operatorname{def} A = \operatorname{def} T_A$.
(b) Eine Matrix heißt *quadratisch*, wenn ihre Zeilen- und Spaltenzahl überein stimmen.
(c) $\operatorname{Im} A$ heißt auch *Spaltenraum* von A und $\operatorname{im} A^\top$ nennt man den *Zeilenraum* von A.

Bemerkung 3.2 (a) Es ist klar, dass $\operatorname{Ker} A = \{x \in \mathbb{K}^n \mid Ax = 0\} = \text{Lös}(A, 0)$ genau die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ ist.
(b) Dagegen ist $\operatorname{Im} A = \{b \in \mathbb{K}^m \mid \exists x \in \mathbb{K}^n : Ax = b\}$ die Menge der Vektoren b in $\mathbb{K}^m$, für die das inhomogene lineare Gleichungssystem $Ax = b$ eine Lösung besitzt.
(c) Eine wichtige Rolle spielt die Bestimmung des Ranges einer Matrix. Nach unserer Definition ist der Spaltenraum $\operatorname{im} A$ der von den Spalten der Matrix A aufgespannte Teilraum im $\mathbb{K}^n$, $\operatorname{Im} A = \operatorname{span}\{A\mathrm{e}_1, \dots, A\mathrm{e}_n\}$. Somit gilt einerseits

$$\operatorname{rg} A = \dim \operatorname{span}\{A\mathrm{e}_1, \dots, A\mathrm{e}_n\} = \text{Dimension des Spaltenraumes von } A.$$

Da n Vektoren höchstens einen Raum der Dimension n aufspannen, gilt $\operatorname{rg} A \le n$. Andererseits ist $\operatorname{Im} A \subset \mathbb{K}^m$ und damit auch $\operatorname{rg} A \le m$. Folglich gilt $\operatorname{rg} A \le \min\{m, n\}$.

Ebenso könnte man aber den *Zeilenraum* von A betrachten, die lineare Hülle des von den Zeilen von A aufgespannten Unterraumes von $\mathbb{K}^n$. Natürlich ist der Zeilenraum von A gleich dem Spaltenraum von $A^\top$, *Zeilenraum* von $A = \operatorname{Im} A^\top$.

Satz 3.9 (Zeilenrang = Spaltenrang) *Für jede Matrix $A \in \mathbb{K}^{m\times n}$ gilt*

$$\operatorname{rg} A = \operatorname{rg} A^\top.$$

Beweis. Wir wollen eine Zeile oder Spalte von A als *linear überflüssig* bezeichnen, wenn sie die Linearkombination von anderen Zeilen bzw. Spalten ist. Verkleinert man eine Matrix durch Weglassen einer linear überflüssigen Spalte, so ändert sich der (Spalten)rang von A nicht.
(a) Wir werden zeigen, dass sich auch *der Zeilenrang dabei nicht ändert.* — Angenommen, die 1.Spalte der Matrix sei linear überflüssig, dann lässt sich die erste Komponente jeder Zeile linear kombinieren aus allen anderen Komponenten der Zeile. Dasselbe gilt für jede Linearkombination von Zeilen: auch hier lässt sich die erste Komponente linear aus den anderen Komponenten kombinieren (und zwar mit denselben Koeffizienten, wie bei allen Zeilen). Somit ist eine Linearkombination von Zeilen genau dann Null, wenn schon die entsprechende Linearkombination der Zeilen, die durch Weglassen der ersten Komponente entstehen, Null ist. Daher ändert sich der Zeilenrang nicht durch weglassen linear überflüssiger Spalten.
(b) Dasselbe Argument gilt natürlich, wenn man die Rolle von Zeilen und Spalten vertauscht: Durch Weglassen linear überflüssiger Zeilen verringert sich der Spaltenrang nicht.
(c) Nun verkleinern wir unsere Matrix A schrittweise durch Weglassen linear überflüssiger Zeilen und Spalten solange, bis es nicht weiter geht. Wir erhalten eine möglicherweise kleinere Matrix A', die denselben Zeilenrang und denselben Spaltenrang wie A hat. Also $\operatorname{rg} A = \operatorname{rg} A'$ und $\operatorname{rg} A^\top = \operatorname{rg} A'^\top$ und A' hat nur linear unabhängige Zeilen und linear unabhängige Spalten.
(d) *A' ist eine quadratische Matrix.* Angenommen, A' hat s Spalten und z Zeilen, dann gilt wegen der linearen Unabhängigkeit der Spalten $\operatorname{rg} A' = s \leq z$. Dasselbe Argument gilt aber für den Zeilenraum, da die Zeilen von A' linear unabhängig sind: $\operatorname{rg} A'^\top = z \leq s$. Also ist $\operatorname{rg} A = \operatorname{rg} A' = s = z =: r$ und es gilt Zeilenrang von A = Spaltenrang von $A = r$= Format der quadratischen Matrix A'. ∎

Bemerkung 3.3 Wir sahen bereits in Abschnitt 2.3.4, dass durch die 3 elementaren Zeilenoperationen des Gauß-Algorithmus, der Zeilenraum nicht verändert wird. Wir erhalten daher:

$$\boxed{r = \operatorname{rg} A = \text{Anzahl der führenden Einsen im Gauß-Algorithmus.}}$$

Der Gauß-Algorithmus (ohne Gauß-Jordan) ist daher bestens geeignet, den Rang einer Matrix zu ermitteln. Nach dem Dimensionssatz für lineare Abbildungen gilt ferner

$$\boxed{n - r = \operatorname{def} A = \text{Dimension des Lösungsraumes des homogenen Systems}}$$

Beispiel 3.4 (a) Es sei $c \neq 0$, dann ist

$$\operatorname{rg}\begin{pmatrix} c & \dots & c \\ c & \dots & c \\ \vdots & & \vdots \\ c & \dots & c \end{pmatrix} = 1,$$

denn die Matrix besitzt nur eine linear unabhängige Zeile. Für $c = 0$ wäre $\operatorname{rg} A = 0$.

(b) Sei $A=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$. Dann liefert der Gauß-Algorithmus

$$A\sim\begin{pmatrix}1&2&3\\0&-3&-6\\0&-6&-12\end{pmatrix}\sim\begin{pmatrix}1&2&3\\0&1&2\\0&0&0\end{pmatrix}$$

Folglich ist $\operatorname{rg} A=2$ und $\operatorname{def} A=1$.
(c) Es sei $x=(x_1,\dots,x_n)\in\mathbb{K}^{1\times n}$ ein Zeilenvektor und $y=(y_1,\dots,y_m)^\top\in\mathbb{K}^{m\times1}$ ein Spaltenvektor. Dann hat die Matrix

$$A=y\cdot x=\begin{pmatrix}y_1\\y_2\\\vdots\\y_m\end{pmatrix}\cdot(x_1,\dots,x_n)=(y_ix_j)_{\substack{i=1\dots m\\j=1\dots n}}$$

den Rang 1 oder 0. Dies folgt aus $\operatorname{rg} x,\operatorname{rg} y\le1$ und der Übungsaufgabe 6.5.

3.3.2 Die inverse Matrix

Wir wissen aus Abschnitt 1.1.3, dass nur die bijektiven Abbildungen invertierbar sind. Das gilt natürlich auch für lineare Abbildungen zwischen Vektorräumen. Nach Lemma 3.5 (c) kann T_A, $A\in\mathbb{K}^{m\times n}$, nur dann ein Isomorphismus sein, wenn $m=n$ ist und somit muss A quadratisch sein. Nur für quadratische Matrizen hat der Begriff der Invertierbarkeit Sinn. die zu $(T_A)^{-1}$ korrespondierende Matrix bezeichnen wir mit B (falls sie existiert). Schreibt man die Gleichung in linearen Abbildungen $T_A\circ(T_A)^{-1}=(T_A)^{-1}\circ T_A=\mathrm{id}_{\mathbb{K}^n}$ matriziell, so lautet sie $A\cdot B=B\cdot A=I_n$, wobei B die zu $(T_A)^{-1}$ gehörige Matrix ist. Wir definieren B als die zu A inverse Matrix.

Definition 3.6 (a) Eine Matrix $A\in\mathbb{K}^{n\times n}$ heißt *invertierbar*, wenn es eine Matrix $A^{-1}\in\mathbb{K}^{n\times n}$ gibt, für die

$$A\cdot A^{-1}=A^{-1}\cdot A=I_n.$$

Dabei ist I_n die Einheitsmatrix der Ordnung n. Wir nennen A^{-1} die zu A *inverse Matrix*.
(b) Die Menge der invertierbaren $n\times n$-Matrizen bezeichnen wir mit $\mathrm{GL}(n,\mathbb{K})$.

Invertierbare Matrizen nennt man auch *regulär* während die nichtinvertierbaren Matrizen *singulär* heißen.

Bemerkung 3.4 (a) In der Tat ist A^{-1} eindeutig durch A bestimmt, denn gäbe es ein $B\in\mathbb{K}^{n\times n}$ mit derselben Eigenschaft, so wäre $B=BI_n=B(AA^{-1})=(BA)A^{-1}=I_nA^{-1}=A^{-1}$.
(b) Aus der Gleichung für die entsprechenden linearen Abbildungen folgt (nach Beispiel 3.1 (a2))

$$T_A\circ T_{A^{-1}}=\mathrm{id}_{\mathbb{K}^n}=T_{A^{-1}}\circ T_A.$$

Das heißt, A ist invertierbar genau dann, wenn T_A invertierbar ist und in diesem Falle gilt $(T_A)^{-1}=T_{A^{-1}}$.

Lemma 3.10 *Es seien $A, B \in \mathrm{GL}(n, \mathbb{K})$.*
(a) Dann ist auch $AB \in \mathrm{GL}(n, \mathbb{K})$ und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.
(b) $A^{-1}, A^\top \in \mathrm{GL}(n.\mathbb{K})$ und $(A^{-1})^{-1} = A$ und $(A^\top)^{-1} = (A^{-1})^\top$.

Beweis. (a) Wegen $AB \cdot B^{-1}A^{-1} = A(BB^{-1})A^{-1} = A \cdot I_n \cdot A^{-1} = A \cdot A^{-1} = I_n$ und analog $B^{-1}A^{-1} \cdot AB = I_n$ folgt die Behauptung. (b) Folgt genauso wie $-(-x) = x$ und $\frac{1}{\frac{1}{x}} = x$. Ferner ist $(A^{-1})^\top \cdot A^\top = (AA^{-1})^\top = I_n^\top = I_n$. Hieraus folgt die Behauptung. ∎

Lemma 3.11 *Es sei $A \in \mathbb{K}^{n \times n}$. Dann sind äquivalent*
(a) A ist invertierbar.
(b) $\operatorname{rg} A = n$.
(c) $\operatorname{def} A = 0$.
(d) Die Spalten von A bilden eine Basis von $\mathbb{K}^n$.
(e) Die Zeilen von A bilden eine Basis von $\mathbb{K}^n$.
(f) T_A ist injektiv.
(g) T_A ist surjektiv.

Beweis. (a) $\longleftrightarrow$ (f) $\longleftrightarrow$ (g): Ist A invertierbar, so ist T_A invertierbar, also bijektiv. Nach dem Äquivalenzsatz ist dies gleichwertig mit (f) und (g). Nach Lemma 3.4 sind (c) und (f) gleichwertig und nach Lemma 3.3 sind (b) und (g) äquivalent. Ferner ist (g) äquivalent zu (d). Die Äquivalenz von (d) und (e) folgt schließlich aus Satz 3.9. ∎

3.3.3 Invertieren einer Matrix

Nach Lemma 3.11 sind genau die quadratischen Matrizen invertierbar, die den vollen Rang n haben. Anders ausgedrückt, der Gauß-Algorithmus, angewandt auf A liefert n führende Einsen; jedes lineare Gleichungssystem $A \cdot x = b$, $b \in \mathbb{R}^n$ ist eindeutig lösbar (universelle Lösbarkeit für alle b). Man kann diese Lösung sofort formal angeben, da ja A^{-1} existiert. Multipliziert man die Gleichung $A \cdot x = b$ von links mit A^{-1}, so erhält man

$$\begin{aligned} A^{-1} \mid \quad & A \cdot x = b \\ & A^{-1} \cdot (A \cdot x) = (A^{-1} \cdot A) \cdot x = I_n \cdot x = x = A^{-1} \cdot b. \end{aligned}$$

Die eindeutig bestimmte Lösung von $A \cdot x = b$ ist also $x = A^{-1} \cdot b$. Genau das nutzt man zur effektiven Berechnung der Matrix A^{-1}. Ist nämlich $b = \mathrm{e}_j$, $j = 1, \ldots, n$, und ist $x^{(j)}$ die Lösung des linearen Gleichungssystems $Ax = \mathrm{e}_j$, so gilt

$$Ax^{(j)} = \mathrm{e}_j \Leftrightarrow x^{(j)} = A^{-1}\mathrm{e}_j \Leftrightarrow x^{(j)} \text{ ist die } j\text{te Spalte der Matrix } A^{-1}.$$

Die letzte Feststellung folgt aus Beispiel 3.1 (a). Löst man also alle Gleichungssysteme $Ax = \mathrm{e}_j$, $j = 1, \ldots, n$, so erhält man die Spalten von A^{-1}. Somit ist das Verfahren zur Matrixinvertierung klar:

Verfahren zur Berechnung von A^{-1}. Gegeben sei eine Matrix $A \in \mathbb{K}^{n\times n}$. Man wendet den Gauß-Jordan-Algorithmus auf die erweiterte Koeffizientenmatrix an, wobei als rechte Seiten die n Spalteneinheitsvektoren genommen werden.
Entsteht dabei auf der Koeffizientenseite (linke Seite) eine halbe Nullzeile, so ist der Rang der Matrix A kleiner als n und die Matrix ist nicht invertierbar. Andernfalls entstehen n führende Einsen und die Koeffizientenmatrix A wird in die Einheitsmatrix I_n transformiert. Auf der rechten Seite kann man dann die inverse Matrix A^{-1} ablesen.

Beispiel 3.5 Gegeben sei die reelle 4×4- Matrix

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

Wir berechnen A^{-1} durch simultanes Lösen von 4 linearen Gleichungssystemen:

1	0	1	1	1	0	0	0	$\cdot(-1)$ Add. zu Zeile 2 und 4
1	1	2	1	0	1	0	0	
0	−1	0	1	0	0	1	0	
1	0	0	2	0	0	0	1	
1	0	1	1	1	0	0	0	
0	1	1	0	−1	1	0	0	Addtion zu Zeile 3
0	−1	0	1	0	0	1	0	
0	0	−1	1	−1	0	0	1	
1	0	1	1	1	0	0	0	
0	1	1	0	−1	1	0	0	
0	0	1	1	−1	1	1	0	
0	0	−1	1	−1	0	0	1	
1	0	1	1	1	0	0	0	
0	1	1	0	−1	1	0	0	
0	0	1	1	−1	1	1	0	(-1) Add zu Zeile 2 und 1
0	0	0	2	−2	1	1	1	
1	0	0	0	2	−1	−1	0	
0	1	0	−1	0	0	−1	0	
0	0	1	1	−1	1	1	0	
0	0	0	1	−1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	Add zur 2., Subtr. von der 3.Zeile
1	0	0	0	2	−1	−1	0	
0	1	0	0	−1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	
0	0	1	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	
0	0	0	1	−1	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	

Somit gilt:

$$A^{-1} = \begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 4 & -2 & -2 & 0 \\ -2 & 1 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ -2 & 1 & 1 & 1 \end{pmatrix}$$

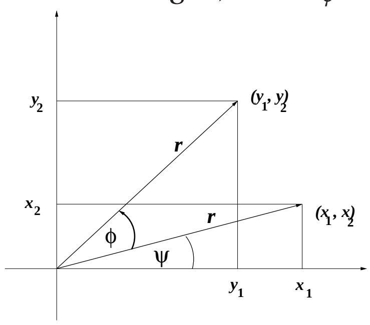
3.3.4 Spezielle Klassen von Matrizen

Es sei $\mathbb{K} = \mathbb{R}$ oder oder $\mathbb{K} = \mathbb{C}$.

Definition 3.7 Eine quadratische Matrix $A \in \mathbb{K}^{n\times n}$ heißt

symmetrisch, falls $A = A^\top$, also $a_{ij} = a_{ji}$ für alle $i, j = 1, \ldots, n$.
hermitesch, falls $A = A^*$, dabei ist $A^* = (\overline{A})^\top$, also $a_{ij} = \overline{a_{ji}}$ (komplexe Konjugation) für alle $i, j = 1, \ldots, n$.
unitär, falls $A^* A = A A^* = I_n$.
Eine unitäre Matrix mit nur reellen Einträgen heißt *orthogonal*. Für orthogonale Matrizen gilt also $A A^\top = A^\top A = I_n$.

Beispiel 3.6 (Drehung um φ) Es sei D_φ die Drehung des $\mathbb{R}^2$ um den Ursprung um den Winkel φ. Wir werden sehen, dass dies in der Tat eine lineare Abbildung ist, deren Matrix bestimmen und zeigen, dass D_φ orthogonal ist.



Die Koordinate y_1 erhält man als $y_1 = r\cos(\varphi + \psi) = r\cos\varphi\sin\psi - r\sin\varphi\cos\psi$. Nun ist aber $r\cos\psi = x_1$ und $r\sin\psi = x_2$. Also hat man $y_1 = \cos\varphi x_1 - \sin\varphi x_2$. Analog berechnet man y_2

$$y_1 = r\cos(\varphi + \psi) = x_1\cos\varphi - x_2\sin\varphi,$$
$$y_2 = r\sin(\varphi + \psi) = x_1\sin\varphi + x_2\cos\varphi.$$

Man erkennt, dass die Abbildung $D_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, $D_\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ linear ist mit der Matrix

$$T_{D_\varphi} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

Zur Vereinfachung der Symbolik identifizieren wir D_φ und T_{D_φ}. Man nennt T_{D_φ} die Drehmatrix um den Winkel φ. Es gilt

$$D_\varphi \cdot D_\varphi^\top = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} \cos^2\varphi + \sin^2\varphi & 0 \\ 0 & \cos^2\varphi + \sin^2\varphi \end{pmatrix} = I_2.$$

Das heißt, D_φ ist eine orthogonale 2×2-Matrix. In der Mechanik werden Drehbewegungen im Raum , das heißt, orthogonale Transformationen $T \in \mathrm{GL}(3, \mathbb{R})$ durch die *Eulerschen Winkel* beschrieben, siehe auch `http://de.wikipedia.org/wiki/Eulersche_Winkel`.

Jede Drehung im dreidimensionalen Raum $\mathbb{R}^3$ um den Ursprung ist die Hintereinanderausführung von drei ebenen Drehungen:eine um den Winkel ψ in der xy-Ebene, eine um den Winkel θ in der yz-Ebene und schließlich erneut eine Drehung um in der xy-Ebene um den Winkel φ.

Später werden wir die in Definition 3.7 eingeführten Begriffe auch für lineare Abbildungen definieren.

3.3.5 Koordinatenransformationen und Basistransformationen

Problem: wie verändern sich die Koordinaten eines Vektors, wenn man zu einer anderen Basis über geht?
Es sei V ein Vektorraum mit der Basis $E = \{e_1, \dots, e_n\}$. Dann gibt es für jeden Vektor $v \in V$ eindeutig bestimmte Zahlen, die Koordinaten von v bezüglich E, $(x_1, \dots, x_n)$ mit $v = x_1v_1 + \dots + x_nv_n$. Sei nun $F = \{f_1, \dots, f_n\}$ eine weitere Basis von V und $v = y_1f_1 + \dots + y_nf_n$.
Problem: Wie berechnen wir die neuen Koordinaten $(y_1, \dots, y_n)$ aus den alten Koordinaten $(x_1, \dots, x_n)$? Die Basistransformation sei gegeben durch eine Matrix $A \in \mathbb{K}^{n\times n}$ in der folgenden Weise:

$$f_j = \sum_{i=1}^{n} a_{ij} e_i, \quad j = 1, \dots, n; \tag{3.3}$$

dabei sei $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ die *Matrix der Basistransformation*. Für festes j schreiben wir also die Koordinaten a_{ij} von f_j bezüglich E in die $j$te *Spalte* der Matrix A. Da F eine Basis ist, gibt es auch eine Matrix $B = (b_{ij}) \in \mathbb{K}^{n\times n}$ mit

$$e_i = \sum_{k=1}^{n} b_{ki} f_k, \quad i = 1, \dots, n, \tag{3.4}$$

welche die umgekehrte Basistransformation beschreibt. Setzt man (3.4) in (3.3) ein, so hat man:

$$f_j = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ij} b_{ki}\, f_k, \quad j = 1, \dots, n.$$

Vertauscht man die Summation und beachtet, dass F eine Basis in V ist, so hat man

$$f_j = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} b_{ki} a_{ij}\right) f_k = \sum_{k=1}^{n} (BA)_{kj} f_k \longrightarrow (BA)_{kj} = \delta_{kj}, \quad k, j = 1, \dots, n.$$

Somit gilt $BA = I_n$ — die Matrizen A und B sind zueinander invers (und insbesondere regulär); $B = A^{-1}$. Durch Einsetzen der Basistransformation (3.3) in die Gleichung $v = y_1f_1 + \dots + y_nf_n$ erhalten wir eine Beziehung zwischen den Koordinaten x_i und y_j:

$$v = \sum_{j=1}^{n} y_j f_j = \sum_{j=1}^{n} y_j \sum_{i=1}^{n} a_{ij} e_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij} y_j\right) e_i = \sum_{i=1}^{n} x_i e_i.$$

Da die Koordinaten bezüglich der Basis $\{e_i\}$ eindeutig bestimmt sind, folgt $x_i = \sum_{j=1}^{n} a_{ij} y_j$, $i = 1, \dots, n$, oder matriziell $x = Ay$. Da A regulär ist, gilt auch $y = A^{-1}x$. Man beachte, dass in der Definition der Matrix A über (3.3) die *transponierte* Matrix einging. 'Formal' gilt $f = A^\top e$ und $e = (A^{-1})^\top f$.

Beispiel 3.7 Wir betrachten das Beispiel aus der Übungsaufgabe 5.3 genauer: Die Vektoren $b_1 = (1,2,1)$, $b_2 = (2,9,0)$ und $b_3 = (3,3,4)$ bilden eine Basis B des $\mathbb{R}^3$. Bestimmen Sie die neuen Koordinaten (y_1, y_2, y_3) bezüglich B, wenn die alten Koordinaten (x_1, x_2, x_3) in der Standardbasis gegeben sind.
Wir lesen die Matrix $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ der Basistransformation spaltenweise ab:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix}$$

Mittels Gauß-Algorithmus bestimmt man die inverse Matrix

$$A^{-1} = \begin{pmatrix} -36 & 8 & 21 \\ 5 & -1 & -3 \\ 9 & -2 & -5 \end{pmatrix}$$

Somit gilt

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = y = A^{-1}x = \begin{pmatrix} -36x_1 + 8x_2 + 21x_3 \\ 5x_1 - x_2 - 3x_3 \\ 9x_1 - 2x_2 - 5x_3 \end{pmatrix}$$

Nun gilt $y_1 b_1 + y_2 b_2 + y_3 b_3 = x_1 \mathrm{e}_1 + x_2 \mathrm{e}_2 + x_3 \mathrm{e}_3$.

3.3.6 Verhalten von Matrizen bei Basistransformation

Problem: Gegeben sei eine lineare Abbildung $T \in \mathrm{L}(V, W)$. Man ermittle Basen B und C in V bzw. in W, so dass die zugehörige Matrix $M_{B,C}(T)$ eine *möglichst einfache* Gestalt hat. Wir werden sehen, dass dieses Problem eine sehr einfache Lösung besitzt (Normalformenproblem).
Wir haben gesehen, wie sich die Koordinaten ändern, wenn man zu einer neuen Basis über geht. Nun wollen wir uns überlegen, wie die neue Matrix M' einer linearen Abbildung $T \in \mathrm{L}(V;W)$ aussieht, wenn $M = M_{E,F}(T)$ die Darstellung von T bezüglich gegebener Basen E und F von V bzw. von W ist und wenn man in V und W zu neuen Basen E' bzw. F' über geht. Wir wollen also $M' = M_{E',F'}(T)$ aus M und den Transformationsmatrizen bestimmen.

A, B	Matrizen der Basistransformationen in V bzw. in W.
x, y	Koordinaten von $v \in V$ bzw. von $T(v) \in W$ bzgl. der Basen E und F.
x', y'	Koordinaten von $v \in V$ und $T(v) \in W$ bzgl. der Basen E' und F'.
M, M'	Matrizen von T bezüglich (E, F) bzw. (E', F').

Nach diesen Setzungen gilt mit Hilfe der Transformationsformeln

$$y = Mx \quad y' = M'x' \quad x' = A^{-1}x \quad y' = B^{-1}y.$$

Multipliziert man die erste Gleichung mit B^{-1} von links und setzt $x = Ax'$ ein, so hat man

$$B^{-1}y = B^{-1}Mx = B^{-1}MAx' \quad \Longrightarrow \quad y' = (B^{-1}MA)x',$$

sodass man unmittelbar die Transformationsformel erhält

$$M' = B^{-1}MA. \tag{3.5}$$

Besonders wichtig ist der Fall von Endomorphismen $T \in \mathrm{L}(V)$, wo man üblicherweise $E = F$ und $E' = F'$ wählt, sodass $A = B$ gilt. Hier lautet die Transformationsformel

$$M' = A^{-1}MA.$$

Definition 3.8 Zwei quadratische Matrizen $M, M' \in \mathbb{K}^{n\times n}$ heißen einander *ähnlich*, wenn es eine invertierbare Matrix $A \in \mathrm{GL}(n,\mathbb{K})$ gibt mit $M' = A^{-1}MA$.

Ähnliche Matrizen beschreiben ein und dieselbe lineare Abbildung $T\colon \mathbb{K}^n \to \mathbb{K}^n$ nur in verschiedenen Basen.

Beispiel 3.8 In Übungsaufgabe 5.3 war eine lineare Abbildung $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ gegeben durch

$$T(b_1) = (1,0), \quad T(b_2) = (-1,1), \quad T(b_3) = (0,1),$$

wobei die Basis $B = \{b_1, b_2, b_3\}$ gegeben ist durch $b_1 = (1,2,1)$, $b_2 = (2,9,0)$ und $b_3 = (3,3,4)$. Somit gilt

$$M = M_{B,B_2}(T) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

wobei B_2 die Standardbasis vom $\mathbb{R}^2$ ist. Gesucht ist die Darstellung von T bezüglich der Standardbasen B_3 und B_2. Die Transformationsmatrix A von B_3 nach B und deren Inverse A^{-1} sind dort angegeben. Da im Zielraum $\mathbb{R}^2$ gar keine Transformation stattfindet entfällt der Faktor $B = I_2$. Somit gilt:

$$M' = M_{B_3,B_2}(T) = MA^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -36 & 8 & 21 \\ 5 & -1 & -3 \\ 9 & -2 & -5 \end{pmatrix} = \begin{pmatrix} -41 & 9 & 24 \\ 14 & -3 & -8 \end{pmatrix}$$

3.3.7 Das Normalformproblem für $T \in \mathrm{L}(V,W)$, $V \neq W$

Es seien V und W endlichdimensionale Vektorräume und $T \in \mathrm{L}(V,W)$. Wir kommen nun zurück zur Frage, wie man Basen in V und W wählen kann, dass die Matrix M' von T möglichst einfache Gestalt hat. Es sei $r = \operatorname{rg} T$. Wie im Beweis von Satz 3.7 wählen wir zunächst eine Basis $\{b_{r+1}, b_{r+2}, \dots, b_n\}$ von $\operatorname{Ker} T$. Ergänzt man diese Basis zu einer Basis $\{b_1, \dots, b_r, b_{r+1}, \dots, b_n\}$ von V, so ist $\{T(b_1), \dots, T(b_r)\}$ eine Basis von $\operatorname{Im} T$. Ergänzt man wiederum diese r Vektoren beliebig zu einer Basis $C = \{T(b_1), \dots, T(b_r), w_{r+1}, \dots, w_m\}$ von W, so hat $M' = M_{B,C}(T)$ die Gestalt

$$M' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei I_r die Einheitsmatrix der Ordnung r ist. Dies folgt direkt aus der eingerahmten Formel in Definition 3.2.

Startet man mit einer Matrix $M \in \mathbb{K}^{m\times n}$ vom Rang $\operatorname{rg} M = r$, so gibt es nach den Überlegungen des vorigen Abschnitts reguläre Matrizen $A \in \mathrm{GL}(m, \mathbb{K})$ und $B \in \mathrm{GL}(n, \mathbb{K})$, sodass

$$M = AM'B = A \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} B.$$

Wir nehmen M' als Normalform von M.

Wesentlich schwieriger ist das Normalformenproblem für $T \in \mathrm{L}(V)$, wenn wir nämlich nur *eine* Basis in V wählen können, die für den Ausgangs- und Zielraum V gleichzeitig gültig ist. Wir kommen darauf im Kapitel Eigenwerttheorie zurück.

Literaturverzeichnis

[Ant98] H. Anton. *Lineare Algebra. Einführung, Grundlagen, Übungen.* Spectrum Lehrbuch. Spektrum Akademischer Verlag, Heidelberg, 1998.

[Jän04] K. Jänich. *Lineare Algebra.* Springer-Lehrbuch. Springer-Verlag, Berlin, 10 edition, 2004.

[Koe97] M. Koecher. *Lineare Algebra und analytische Geometrie.* Springer-Verlag, Berlin, 4 edition, 1997.

[Kow79] H.-J. Kowalsky. *Lineare Algebra.* De Gruyter Lehrbuch. Walter de Gruyter, Berlin, New York, 1979.

PD Dr. A. Schüler
Mathematisches Institut
Universität Leipzig
04009 Leipzig
Axel.Schueler@math.uni-leipzig.de