Kapitel II

Lineare Algebra

§1 Vektorräume und Lineare Abbildungen

Es sei K ein Körper. Meistens interessieren uns nur die Fälle $K = \mathbb{R}$ und $K = \mathbb{C}$.

Definition.

Ein *Vektorraum (über K)* ist eine Menge V mit einem ausgezeichneten Element $\mathbf{0}$, so daß gilt:

1. Es gibt eine Verknüpfung $+$ auf V, so daß $(V, +, \mathbf{0})$ eine kommutative Gruppe ist.
2. Es gibt zusätzlich eine Verknüpfung $\cdot$ zwischen den Elementen von K (den „Skalaren") und den Elementen von V (den „Vektoren"), so daß für $\alpha, \beta \in K$ und $\mathbf{v}$, $\mathbf{w} \in V$ gilt:
   (a) $(\alpha + \beta) \cdot \mathbf{v} = (\alpha \cdot \mathbf{v}) + (\beta \cdot \mathbf{v})$,
   (b) $\alpha \cdot (\mathbf{v} + \mathbf{w}) = (\alpha \cdot \mathbf{v}) + (\beta \cdot \mathbf{w})$,
   (c) $\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha\beta) \cdot \mathbf{v}$,
   (d) $1 \cdot \mathbf{v} = \mathbf{v}$.

Bemerkung : Wenn keine Klammern gesetzt werden, gilt wie üblich „Punkt vor Strich".

Beispiele :

1. Natürlich sind die Mengen $\mathcal{V}_2$ und $\mathcal{V}_3$ der „anschaulichen" ebenen und räumlichen Vektoren auch Vektorräume (über $\mathbb{R}$). Sie dienten ja gerade als Modell für den abstrakten Vektorraum-Begriff.

2. Jeder Vektorraum enthält eine „Null". Weitere Elemente braucht es aber laut Definition nicht zu geben. Deshalb ist $\{\mathbf{0}\}$ ist das einfachste Beispiel eines Vektorraumes.

3. Der n–dimensionale Zahlenraum $K^n := \{\mathbf{x} = (x_1, \ldots, x_n) \mid x_i \in K \text{ für alle } i\}$ ist ein Vektorraum über K, wenn man die Verknüpfungen komponentenweise erklärt:

$$\begin{aligned}
(v_1, \ldots, v_n) + (w_1, \ldots, w_n) &:= (v_1 + w_1, \ldots, v_n + w_n), \\
-(v_1, \ldots, v_n) &:= (-v_1, \ldots, -v_n), \\
\mathbf{0} &:= (0, \ldots, 0), \\
\alpha \cdot (v_1, \ldots, v_n) &:= (\alpha \cdot v_1, \ldots, \alpha \cdot v_n).
\end{aligned}$$

Da es ja letztlich egal ist, ob man die Komponenten eines n–Tupels auf dem Papier nebeneinander oder übereinander schreibt, können wir ein Element $\mathbf{v} = (v_1, \ldots, v_n)$ des K^n auch als Spaltenvektor

$$\vec{v} := \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

schreiben. Normalerweise bräuchte man nicht zwischen Zeilen- und Spaltenvektoren zu unterscheiden. Es gibt aber Rechenverfahren, die sich besser einprägen, wenn man die eine oder die andere Schreibweise verwendet.

4. Ist $n = 1$, so wird aus K^n der Körper K selbst. Der ist also auch ein K–Vektorraum. Die Multiplikation mit Skalaren ist in diesem Falle nichts anderes als die interne Multiplikation in dem Körper K.

5. Ist $N = n \cdot m$, so kann man die N–Tupel des K^N auch in einem $n \times m$–reihigen rechteckigen Muster anordnen. Mit anderen Worten: Die Menge $M_{n,m}(K)$ der Matrizen mit n Zeilen und m Spalten und Koeffizienten aus K bildet einen Vektorraum über K, wenn man Addition und skalare Multiplikation komponentenweise erklärt.

6. Es sei $\Pi(K)$ die Menge der Polynome mit Koeffizienten in K. Hier wollen wir uns ausdrücklich auf die Fälle $K = \mathbb{R}$ und $K = \mathbb{C}$ beschränken. Die Summe zweier Polynome ist wieder ein Polynom, das Produkt eines Polynoms mit einem Skalar ist ebenfalls wieder ein Polynom. Alle anderen Regeln sind leicht nachzurechnen. Also ist $\Pi(K)$ ein Vektorraum über K, auch wenn dabei die geometrische Vorstellung von Vektoren gänzlich untergeht. Aber das macht ja gerade die Stärke des abstrakten Vektoraum-Begriffes aus.

Weitere Beispiele bekommen wir auf folgendem Wege:

Definition.

Sei V ein Vektorraum über K. Eine Teilmenge $U \subset V$ heißt ein *Untervektorraum*, falls gilt:

1. $\mathbf{0} \in U$,
2. Sind $\mathbf{v}, \mathbf{w} \in U$, so ist auch $\mathbf{v} + \mathbf{w} \in U$.
3. Ist $\alpha \in K$ und $\mathbf{v} \in U$, so ist auch $\alpha \cdot \mathbf{v} \in U$.

II.1.1 Satz. *Ist $U \subset V$ ein Untervektorraum, so ist U auch selbst ein Vektorraum.*

BEWEIS: Die Null und zwei Verknüpfungen gibt es nach Voraussetzung. Daß alle Gesetze gelten, folgt daraus, daß sie ja in dem größeren Raum V schon gelten. □

Beispiele :

1. Der *Nullraum* $\{\mathbf{0}\}$ ist Untervektorraum jedes anderen Vektorraumes.

2. Ist V ein Vektorraum über K und $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$, so ist

$$K\mathbf{v} := \{\alpha \cdot \mathbf{v} \mid \alpha \in \mathbb{R}\}$$

ein Untervektorraum.

(a) $\mathbf{0} = 0 \cdot \mathbf{v}$ liegt offensichtlich in $K\mathbf{v}$.

(b) Sind $\mathbf{w}_1 = \alpha \cdot \mathbf{v}$ und $\mathbf{w}_2 = \beta \cdot \mathbf{v}$ Elemente von $K\mathbf{v}$, so ist auch $\mathbf{w}_1 + \mathbf{w}_2 = (\alpha+\beta) \cdot \mathbf{v}$ ein Element von $K\mathbf{v}$.

(c) Ist $\mathbf{w} = \alpha \cdot \mathbf{v} \in K\mathbf{v}$ und $\gamma \in K$, so liegt auch $\gamma \cdot \mathbf{w} = (\gamma\alpha) \cdot \mathbf{v}$ in $K\mathbf{v}$.

3. $\mathbb{C}$ ist ein $\mathbb{R}$–Vektorraum, und $\mathbb{R} \subset \mathbb{C}$ ist ein Untervektorraum.

$\mathbb{C}$ ist natürlich auch ein $\mathbb{C}$–Vektorraum, aber $\mathbb{R}$ ist kein $\mathbb{C}$-Vektorraum und daher auch — bezogen auf den Skalarenkörper $\mathbb{C}$ — kein Untervektorraum von $\mathbb{C}$.

4. Die Menge $\Delta_n(K)$ aller Diagonal–Matrizen

$$\begin{pmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix}$$

bildet einen Untervektorraum von $M_{n,n}(K)$.

5. Jedes Element $c \in K$ kann man als konstantes Polynom auffassen. In diesem Sinne bildet K einen Untervektorraum des Raumes $\Pi(K)$ aller Polynome.

6. Die Kugel $B := \{\vec{v} \in \mathcal{V}_3 \mid \|\vec{v}\| \leq 1\}$ stellt *keinen* Untervektorraum von $\mathcal{V}_3$ dar!

Warum nicht? Ist etwa $\vec{v} \in B$, $\vec{v} \neq \vec{0}$, so ist $\varepsilon := \|\vec{v}\|$ eine positive reelle Zahl ≤ 1. Wäre B ein Untervektorraum, so müßte auch $\vec{w} := \frac{2}{\varepsilon} \cdot \vec{v}$ zu B gehören. Es ist aber $\|\vec{w}\| = \frac{2}{\varepsilon} \cdot \varepsilon = 2 > 1$.

Sind $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ und $\alpha_1, \ldots, \alpha_n \in K$, so nennt man den Ausdruck

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i := \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$$

eine *Linearkombination* der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Definition.

Es seien $\mathbf{v}_1, \ldots, \mathbf{v}_n$ Vektoren in V. Dann wird die Menge

$$< \mathbf{v}_1, \ldots, \mathbf{v}_n > := \{\sum_{i=1}^{n} \alpha_i \mathbf{v}_i \mid \alpha_1, \ldots, \alpha_n \in K\}$$

aller Linearkombinationen dieser Vektoren als *Spann von* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ (oder als *die von* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *erzeugte Menge* bezeichnet.

II.1.2 Satz. *Sind $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, so ist $< \mathbf{v}_1, \ldots, \mathbf{v}_n >$ ein Untervektorraum von V.*

BEWEIS:
1) $\mathbf{0} = 0 \cdot \mathbf{v}_1 + \cdots + 0 \cdot \mathbf{v}_n$ liegt in $< \mathbf{v}_1, \ldots, \mathbf{v}_n >$.

2) Sind

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \quad \text{und} \quad \mathbf{y} = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$$

Elemente von $< \mathbf{v}_1, \ldots, \mathbf{v}_n >$, sowie $\alpha \in K$, dann gilt:

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= \sum_{i=1}^{n} (\alpha_i + \beta_i) \mathbf{v}_i \quad \in < \mathbf{v}_1, \ldots, \mathbf{v}_n > \\ \text{und} \quad \alpha \cdot \mathbf{x} &= \sum_{i=1}^{n} (\alpha \cdot \alpha_i) \mathbf{v}_i \quad \in < \mathbf{v}_1, \ldots, \mathbf{v}_n > . \end{aligned}$$

Damit ist alles gezeigt. □

Definition.

Ein System $E := \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ von Vektoren aus V heißt

1. *ein Erzeugendensystem von V*, falls gilt:

$$< \mathbf{v}_1, \ldots, \mathbf{v}_n > = V,$$

2. *linear unabhängig*, falls gilt:

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0} \Longrightarrow \alpha_1 = \ldots = \alpha_n = 0,$$

3. eine *Basis von V*, falls E gleichzeitig linear unabhängig *und* ein Erzeugendensystem von V ist.

II.1.3 Satz. *Sei $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ eine Basis des Vektorraumes V. Dann läßt sich jeder Vektor $\mathbf{v} \in V$* **eindeutig** *als Linearkombination der $\mathbf{v}_i$ schreiben.*

BEWEIS: Sei $\mathbf{v} \in V$ beliebig vorgegeben. Da $V =< \mathbf{v}_1, \ldots, \mathbf{v}_n >$ ist, kann man $\mathbf{v}$ natürlich als Linearkombination der $\mathbf{v}_i$ schreiben. Wenn es nun zwei solche Darstellungen gibt,

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{n} \beta_i \mathbf{v}_i,$$

dann ist

$$\sum_{i=1}^{n} (\alpha_i - \beta_i) \mathbf{v}_i = \mathbf{0}.$$

Und weil die v_i linear unabhängig sind, muß dann $\alpha_i - \beta_i = 0$ sein, für $i = 1, \ldots, n$.

Also stimmen die beiden Darstellungen überein. □

Beispiele :

1. Die Einheitsvektoren
$$\vec{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots \ , \ \vec{e}_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$
bilden eine Basis von K^n.

2. Die 1 ist eine Basis von K (als K–Vektorraum aufgefaßt).

3. 1 und $\mathbf{j}$ bilden eine Basis von $\mathbb{C}$ über $\mathbb{R}$.

4. Die „Monome“ $1, x, x^2, x^3, \ldots$ ergeben ein Erzeugendensystem für den Vektorraum $\Pi(\mathbb{R})$ der reellen Polynome. Da dies unendlich viele Elemente sind, können wir gegenwärtig nicht entscheiden, ob sie auch eine Basis bilden.

Wie steht es mit der Existenz von Basen? Im K^n haben wir zumindest die Standardbasis.

Nun sei $W \subset K^n$ ein Untervektorraum, $W \neq \{\mathbf{0}\}$. Um zu zeigen, daß es in W eine Basis gibt, müssen wir eine Tatsache voraussetzen, die wir erst im nächsten Paragraphen beweisen können:[6]

Im K^n gibt es höchstens n linear unabhängige Vektoren.

II.1.4 Satz.
Jeder Untervektorraum $W \subset K^n$, $W \neq \{\mathbf{0}\}$, besitzt eine Basis.

BEWEIS: Da $W \neq \{\mathbf{0}\}$ ist, gibt es in W wenigstens einen Vektor $\mathbf{v}_1 \neq \mathbf{0}$, also wenigstens ein linear unabhängiges Element. Nun sei $k \leq n$ die maximale Anzahl linear unabhängiger Vektoren, die man in W finden kann, und es sei

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

ein solches System linear unabhängiger Vektoren.

Ist $\mathbf{w} \in W$ beliebig, so sind $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{w}$ linear abhängig. Es gibt also eine Darstellung der Null,

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k + \beta \mathbf{w},$$

wo nicht alle Koeffizienten $= 0$ sind. Wäre nun $\beta = 0$, so müßten — wegen der linearen Unabhängigkeit der $\mathbf{v}_i$ — auch $\alpha_1 = \ldots = \alpha_k = 0$ sein. Da das gerade ausgeschlossen wurde, muß $\beta \neq 0$ sein, und dann gilt:

$$\mathbf{w} = (-\frac{\alpha_1}{\beta}) \cdot \mathbf{v}_1 + \cdots + (-\frac{\alpha_k}{\beta}) \cdot \mathbf{v}_k.$$

Das bedeutet, daß $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ auch ein Erzeugendensystem ist. □

[6] Wir werden dort darauf achten, daß wir die Existenz einer Basis nicht benutzen!

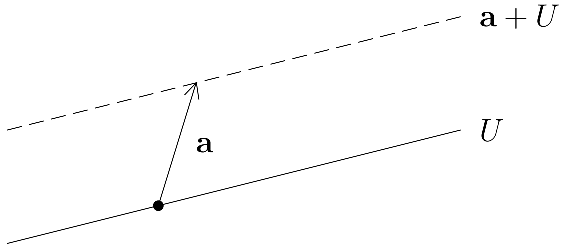
Definition.

Sei V ein Vektorraum, $\mathbf{a} \in V$ und $U \subset V$ ein Untervektorraum. Dann nennt man

$$\mathbf{a} + U := \{\mathbf{a} + \mathbf{u} \mid \mathbf{u} \in U\}$$

eine (zu U parallele) *lineare Mannigfaltigkeit*, oder auch einen *affinen Unterraum*.

Allgemeine Geraden, Ebenen usw. sind affine Unterräume in Vn.



Ist $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ eine Basis von U, so ist

$$\mathbf{a} + U = \{\mathbf{x} \in V \mid \exists\, \alpha_1, \ldots, \alpha_m \in K \text{ mit } \mathbf{x} = \mathbf{a} + \alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m\}.$$

Das ist dann eine Parameterdarstellung für $\mathbf{a} + U$.

Wir kommen nun zum zweiten fundamentalen Begriff dieses Abschnittes:

Definition.

Es seien V und W zwei (gleiche oder verschiedene) K–Vektorräume.

Eine Abbildung $f : V \to W$ heißt *linear*, falls für alle $\mathbf{u}, \mathbf{v} \in V$ und $\alpha \in K$ gilt:

1. $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.
2. $f(\alpha \cdot \mathbf{v}) = \alpha \cdot f(\mathbf{v})$.

Beispiele :

1. $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := mx$ ist linear, aber $g(x) := mx + b$ ist für $b \neq 0$ nicht linear, denn es ist z.B.

   $$g(u + v) = mu + mv + b, \text{ aber } g(u) + g(v) = mu + mv + 2b.$$

   Man nennt solche Funktionen zwar trotzdem in der Analysis gerne linear, sollte aber besser *affin-linear* sagen.

2. $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := x^2$ ist mit Sicherheit *nicht* linear!

3. Ist $c \in \mathbb{C}$ eine feste komplexe Zahl, so wird durch $z \mapsto c \cdot z$ eine lineare Abbildung des $\mathbb{R}$–Vektorraumes $\mathbb{C}$ auf sich definiert.

4. Ist V ein beliebiger $\mathbb{R}$–Vektorraum, $\mathbf{v} \in V$ fest gewählt, so ist $f : \mathbb{R} \to V$ mit $f(\lambda) := \lambda \cdot \mathbf{v}$ linear.

5. Sei $t \in \mathbb{R}$ und $R : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch

$$R : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos(t) - y\sin(t) \\ x\sin(t) + y\cos(t) \end{pmatrix}.$$

   Das ist eine lineare Abbildung, die mit $z \mapsto e^{\mathbf{j}t} \cdot z$ übereinstimmt.

Nun wollen wir versuchen, etwas Überblick zu bekommen:

Sei $F : K^m \to K^n$ eine beliebige lineare Abbildung.

Die Standardbasis von K^m sei $\{\vec{e}_1, \ldots, \vec{e}_m\}$, die von K^n sei $\{\vec{f}_1, \ldots, \vec{f}_n\}$

Ist $\vec{x} = \sum\limits_{j=1}^{m} x_j \, \vec{e}_j \in K^m$, so muß wegen der Linearität von F gelten:

$$F(\vec{x}) = \sum_{j=1}^{m} x_j \cdot F(\vec{e}_j).$$

Die Abbildung ist also durch die Werte $F(\vec{e}_1), \ldots, F(\vec{e}_m)$ schon völlig festgelegt. Und diese Bildvektoren im K^n sind ihrerseits durch ihre Komponenten festgelegt:

$$\vec{a}_j := F(\vec{e}_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = \sum_{i=1}^{n} a_{ij} \, \vec{f}_i \, .$$

Wenn wir also die $n \times m$ Zahlen a_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, m$, kennen, dann kennen wir auch die lineare Abbildung F. Die Zahlen können wir zu einer Matrix $A \in M_{n,m}(K)$ zusammenfassen, und wir haben gerade gesehen, daß jede solche Matrix A eine lineare Abbildung $F_A : K^m \to K^n$ bestimmt, und umgekehrt!

Definition.

Ist $\vec{x} \in K^m = M_{m,1}$ und $A \in M_{n,m}$, so wird $A \circ \vec{x} \in M_{n,1} = K^n$ definiert durch

$$A \circ \vec{x} := F_A(\vec{x}).$$

Das ergibt eine Verknüpfung $M_{n,m} \times M_{m,1} \to M_{n,1}$. Wie diese Verknüpfung genau aussieht, das wollen wir jetzt ausrechnen:

II.1.5 Satz. *Es ist*

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \circ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + x_2 a_{12} + \cdots + x_m a_{1m} \\ \vdots \\ x_1 a_{n1} + x_2 a_{n2} + \cdots + x_m a_{nm} \end{pmatrix}.$$

Legt man also die Spalte $\vec{x}$ auf die i–te Zeile von A, multipliziert dann jeweils die übereinander liegenden Terme und summiert alles auf, so erhält man den i–ten Eintrag im Bildvektor.

Beweis: Mit den oben gewählten Bezeichnungen gilt:

$$\begin{aligned} A\circ \vec{x} = F_A(\vec{x}) &= \sum_{j=1}^{m} x_j \cdot F(\vec{e}_j) \\ &= \sum_{j=1}^{m} x_j \cdot \left(\sum_{i=1}^{n} a_{ij}\ \vec{f}_i\right) \\ &= \sum_{j=1}^{m}\sum_{i=1}^{n} x_j a_{ij}\ \vec{f}_i \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} x_j a_{ij}\right) \vec{f}_i\ . \end{aligned}$$

Daraus kann man die Behauptung ablesen. □

Beispiel:

Sei $F : \mathbb{R}^3 \to \mathbb{R}^3$ definiert durch

$$F\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) := \begin{pmatrix} 3x_1 - 17x_2 \\ x_1 + 5x_2 + x_3 \\ 2(x_2 - x_1) \end{pmatrix}.$$

Dann kann man auch schreiben:

$$F\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 3 & -17 & 0 \\ 1 & 5 & 1 \\ -2 & 2 & 0 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Eine lineare Abbildung $F : K^n \to K$ nennt man auch eine *Linearform*. Wenn wir $a_1 := a_{11}, \ldots, a_n := a_{1n}$ setzen, dann können wir F in der Form

$$F : (x_1, \ldots, x_n)^\top \mapsto a_1x_1 + \cdots + a_nx_n$$

schreiben.

Ist $n = 3$ und $\vec{a} := (a_1, a_2, a_3)^\top$, so ist $F(\vec{x}) = \vec{a}^\top \circ \vec{x} = \mathbf{a} \circ \mathbf{x}^\top = \vec{a} \bullet \vec{x}$.

Definition.

Sei $f : V \to W$ eine lineare Abbildung. Dann heißt

$$\begin{aligned} &\mathrm{Im}\,(f) := \{\mathbf{w} \in W \mid \exists\, \mathbf{v} \in V \text{ mit } \mathbf{w} = f(\mathbf{v})\} \quad \text{das } \textit{Bild} \text{ von } f \\ \text{und}\quad &\mathrm{Ker}\,(f) := \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}\} \quad \text{der } \textit{Kern} \text{ von } f. \end{aligned}$$

II.1.6 Satz. *Ist $f : V \to W$ linear, so sind $\mathrm{Im}\,(f) \subset W$ und $\mathrm{Ker}\,(f) \subset V$ jeweils Untervektorräume.*

Beweis:

1) Ist f linear, so ist auf jeden Fall $f(\mathbf{0}) = f(0 \cdot \mathbf{0}) = 0 \cdot f(\mathbf{0}) = \mathbf{0}$. Also liegt die Null in $\mathrm{Im}\,(f)$.

Ist $\mathbf{w}_1 = f(\mathbf{v}_1)$ und $\mathbf{w}_2 = f(\mathbf{v}_2)$, so ist $\mathbf{w}_1 + \mathbf{w}_2 = f(\mathbf{v}_1) + f(\mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2)$, und $\alpha \cdot \mathbf{w}_1 = \alpha \cdot f(\mathbf{v}_1) = f(\alpha \cdot \mathbf{v}_1)$.

2) Da $f(\mathbf{0}) = \mathbf{0}$ ist, liegt die Null in $\mathrm{Ker}\,(f)$.

Ist $f(\mathbf{v}_1) = \mathbf{0}$ und $f(\mathbf{v}_2) = \mathbf{0}$, so ist auch $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) = \mathbf{0}$, sowie $f(\lambda \cdot \mathbf{v}_1) = \lambda \cdot f(\mathbf{v}_1) = \mathbf{0}$. □

Bemerkung: Ist $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ eine Basis von V, so ist

$$\mathrm{Im}\,(f) = < f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n) > .$$

II.1.7 Satz. *Sei $f : V \to W$ linear.*

1. *f ist genau dann surjektiv, wenn $\mathrm{Im}\,(f) = W$ ist.*
2. *f ist genau dann injektiv, wenn $\mathrm{Ker}\,(f) = \{0\}$ ist.*

BEWEIS: 1) ist trivial.

2) a) Sei f injektiv, $f(\mathbf{v}) = \mathbf{0}$. Da auch $f(\mathbf{0}) = \mathbf{0}$ ist, haben wir $f(\mathbf{v}) = f(\mathbf{0})$, und wegen der Injektivität muß dann $\mathbf{v} = \mathbf{0}$ sein.

b) Sei $\mathrm{Ker}\,(f) = \{\mathbf{0}\}$. Ist $f(\mathbf{v}_1) = f(\mathbf{v}_2)$, so ist $f(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$, also $\mathbf{v}_1 - \mathbf{v}_2 \in \mathrm{Ker}\,(f)$. Da dieser nur die Null enthält, muß $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$ sein, also $\mathbf{v}_1 = \mathbf{v}_2$. Damit ist f injektiv. □

Beispiele:

1. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x_1, x_2) := x_1 - x_2$.

   Da $x = x - 0 = f(x, 0)$ für jede reelle Zahl x gilt, ist $\mathrm{Im}\,(f) = \mathbb{R}$.

   Weiter ist $\mathrm{Ker}\,(f) = \{(x_1, x_2) \mid x_1 = x_2\} = \{(x, x) \mid x \in \mathbb{R}\} \neq \{\mathbf{0}\}$. Also ist f surjektiv, aber nicht injektiv.

2. Sei V ein beliebiger Vektorraum, $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ linear unabhängig, und $f : K^m \to V$ definiert durch
   $$f(\alpha_1, \ldots, \alpha_m) := \sum_{i=1}^{m} \alpha_i \mathbf{v}_i.$$
   Dann ist $\mathrm{Im}\,(f) = < \mathbf{v}_1, \ldots, \mathbf{v}_m >$ und $\mathrm{Ker}\,(f) = \{\mathbf{0}\}$.

Definition.

Eine **bijektive** lineare Abbildung $f : V \to W$ heißt ein *Isomorphismus*.

Zwei Vektorräume V und W heißen *isomorph* (in Zeichen $V \cong W$), wenn es zwischen ihnen einen Isomorphismus gibt.

II.1.8 Satz. *Ist $f : V \to W$ ein Isomorphismus, so ist $f^{-1} : W \to V$ linear, also auch ein Isomorphismus.*

Beweis: Seien $\mathbf{w}_1, \mathbf{w}_2 \in W$, $\alpha \in K$. Dann gilt:

$$\exists\, \mathbf{v}_1, \mathbf{v}_2 \in V \text{ mit } f(\mathbf{v}_1) = \mathbf{w}_1 \text{ und } f(\mathbf{v}_2) = \mathbf{w}_2.$$

Also ist

$$\begin{aligned} f^{-1}(\mathbf{w}_1 + \mathbf{w}_2) &= f^{-1}(f(\mathbf{v}_1) + f(\mathbf{v}_2)) \\ &= f^{-1}(f(\mathbf{v}_1 + \mathbf{v}_2)) \\ &= \mathbf{v}_1 + \mathbf{v}_2 \\ &= f^{-1}(\mathbf{w}_1) + f^{-1}(\mathbf{w}_2), \end{aligned}$$

und

$$\begin{aligned} f^{-1}(\alpha \cdot \mathbf{w}_1) &= f^{-1}(\alpha \cdot f(\mathbf{v}_1)) \\ &= f^{-1}(f(\alpha \cdot \mathbf{v}_1)) \\ &= \alpha \cdot \mathbf{v}_1 \\ &= \alpha \cdot f^{-1}(\mathbf{w}_1). \end{aligned}$$

□

II.1.9 Satz. *Sei $f : V \to W$ ein Isomorphismus und $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ eine Basis von V. Dann ist auch $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\}$ eine Basis von W.*

Beweis: Sei $\mathbf{w}_i := f(\mathbf{v}_i)$, für $i = 1, \ldots, n$.

1) Da $\mathrm{Im}\,(f) = < \mathbf{w}_1, \ldots, \mathbf{w}_n >$ und f surjektiv ist, bilden die $\mathbf{w}_i$ ein Erzeugendensystem.

2) Seien $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum_{i=1}^{n} \alpha_i \mathbf{w}_i = \mathbf{0}$. Dann ist

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i) = \sum_{i=1}^{n} f(\alpha_i \mathbf{v}_i) = f\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right).$$

Da $\mathrm{Ker}\,(f) = \{\mathbf{0}\}$ ist, muß schon $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$ sein. Da die $\mathbf{v}_i$ aber in V linear unabhängig sind, geht das nur, wenn $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ ist.

Also sind die $\mathbf{w}_i$ linear unabhängig. □

Da in einem Vektorraum jedes Element auf eindeutige Weise als Linearkombination der Basisvektoren geschrieben werden kann (sofern eine Basis existiert), überträgt ein Isomorphismus die Struktur eines Vektorraumes komplett auf einen anderen.

Die meisten Vektorräume, die uns interessieren, besitzen eine endliche Basis. Mit deren Hilfe kann man dann einen Isomorphismus auf einen der Räume K^n konstruieren. Daher werden wir uns im Folgenden vor allem mit dem Vektorraum K^n beschäftigen.

§2 Matrizen und Lineare Gleichungssysteme

Ein allgemeines Lineares Gleichungssystem hat die folgende Gestalt:

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ \vdots & & & & & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array}$$

Die Koeffizienten a_{ij} sollen in einem Körper K liegen. Wir können sie dann zu einer Matrix $A \in M_{n,m}(K)$ zusammenfassen. Bilden wir außerdem aus $x_1, \ldots, x_m$ einen Spaltenvektor $\vec{x} \in K^m$, und aus $b_1, \ldots, b_n$ einen Spaltenvektor $\vec{b} \in K^n$, so bekommt das Lineare Gleichungssystem (kurz LGS) die übersichtliche Gestalt

$$A \circ \vec{x} = \vec{b} \qquad \text{oder} \qquad F_A(\vec{x}) = \vec{b}\ .$$

Die Spalten der Matrix A bilden m Spaltenvektoren $\vec{a}_1, \ldots, \vec{a}_m \in K^n$. Damit können wir das LGS auch wie folgt schreiben:

$$\sum_{j=1}^{m} x_j\ \vec{a}_j = \vec{b}\ .$$

Wir werden von Fall zu Fall die eine oder andere Darstellungsform wählen.

Definition.

Gegeben sei ein LGS $\quad A \circ \vec{x} = \vec{b}$.

1. Das LGS heißt *homogen*, falls $\vec{b} = \vec{0}$ ist. Andernfalls heißt es *inhomogen*.
2. Eine *Lösung* des LGS ist ein Vektor $\vec{x} \in K^m$, der die Gleichung erfüllt.

   Man kann A und $\vec{b}$ zur *erweiterten Matrix* $(A, \vec{b}) \in M_{n,m+1}(K)$ zusammenfassen. Mit $\text{Lös}(A, \vec{b})$ bezeichnet man die Menge aller Lösungen des LGS.

Offensichtlich gilt:

I.2.1 Satz.

$$\text{Lös}(A, \vec{0}) = \text{Ker}\,(F_A) \textit{ ist ein Unterverktorraum von } K^m.$$

Ein homogenes LGS $\quad A \circ \vec{x} = \vec{0}$ *besitzt immer die „triviale Lösung“* $\vec{x} = \vec{0}$. *Das ist genau dann die einzige Lösung, wenn F_A injektiv ist.*

Ist $A = (\vec{a}_1, \ldots, \vec{a}_m)$, so ist

$$<\vec{a}_1, \ldots, \vec{a}_m> = < F_A(\vec{e}_1), \ldots, F_A(\vec{e}_m) > = \text{Im}\,(F_A).$$

Daher gilt:

I.2.2 Satz. *Ein inhomogenes LGS* $\quad A \circ \vec{x} = \vec{b} \quad$ *ist genau dann lösbar, wenn* $\vec{b} \in <\vec{a}_1, \ldots, \vec{a}_m>$ *ist.*

Das liefert uns ein weiteres Lösungskriterium, das in der Praxis von besonderer Bedeutung ist:

Definition.

Ist A eine Matrix, so versteht man unter dem *Zeilenrang* (bzw. *Spaltenrang*) von A (in Zeichen: $\mathrm{rg}_z(A)$ bzw. $\mathrm{rg}_s(A)$) die Maximalzahl linear unabhängiger Zeilen (bzw. Spalten) von A.

I.2.3 Satz. *Das inhomogene LGS* $A\circ \vec{x}=\vec{b}$ *ist genau dann lösbar, wenn* $\mathrm{rg}_s(A,\vec{b}) = \mathrm{rg}_s(A)$ *ist.*

Beweis: Die Rang–Gleichung ist genau dann erfüllt, wenn $\vec{b}$ von den Spalten $\vec{a}_1, \ldots, \vec{a}_m$ linear abhängt, wenn also $\vec{b}\in<\vec{a}_1, \ldots, \vec{a}_m>$ ist. $\square$

Nehmen wir nun einmal an, das LGS $A\circ \vec{x}=\vec{b}$ sei lösbar. Wenn $\vec{x}_1$ und $\vec{x}_2$ zwei Lösungen sind, dann gilt:

$$A \circ (\vec{x}_1 - \vec{x}_2) = A\circ \vec{x}_1 - A\circ \vec{x}_2 = \vec{b} - \vec{b} = \vec{0}\ .$$

Das bedeutet: $\vec{x}_1$ und $\vec{x}_2$ unterscheiden sich um eine Lösung des *zugehörigen homogenen Systems*, also um ein Element von $\text{Lös}(A,\vec{0}) = \mathrm{Ker}\,(F_A)$.

Daraus folgt:

I.2.4 Satz. *Ist* $\vec{x}_0$ *eine spezielle Lösung des LGS* $A\circ \vec{x}=\vec{b}$ *, so ist*

$$\text{Lös}(A,\vec{b}) = \vec{x}_0 + \mathrm{Ker}\,(F_A).$$

Wenn also ein inhomogenes LGS überhaupt lösbar ist, dann ist die Lösungsmenge eine lineare Mannigfaltigkeit. Und das LGS ist genau dann eindeutig lösbar, wenn das zugehörige homogene System nur die triviale Lösung besitzt.

Beispiel :

Wir betrachten das LGS

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + x_3 &= -2 \end{aligned}$$

In Matrizenschreibweise sieht das folgendermaßen aus:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Wir bestimmen zunächst die Lösungsmenge des zugehörigen homogenen Systems:

$$\begin{aligned} \text{Lös}(A,\vec{0}) &= \{\vec{x} \mid x_1 + x_2 + x_3 = x_1 + x_3 = 0\} \\ &= \{\vec{x} \mid x_2 = 0 \text{ und } x_3 = -x_1\} \\ &= \{(x,0,-x)^\top \mid x \in \mathbb{R}\} \\ &= \mathbb{R}\cdot(1,0,-1)^\top. \end{aligned}$$

Nun brauchen wir noch eine spezielle Lösung des inhomogenen Systems: Setzen wir versuchsweise $x_1 = 1$, so bleiben uns die Gleichungen

$$x_2 + x_3 = 0 \text{ und } x_3 = -3.$$

Also ist $\vec{x}_0 := (1, 3, -3)^\top$ eine Lösung.

Die Gesamtlösungsmenge ist dann

$$\text{Lös}(A, \vec{b}) = \begin{pmatrix} 1 \\ 3 \\ -3 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \{ \begin{pmatrix} 1+x \\ 3 \\ -3-x \end{pmatrix} \mid x \in \mathbb{R}\}.$$

Im Allgemeinen geht es leider nicht so einfach. Zur Behandlung von großen Gleichungssystemen muß man systematisch vorgehen.

Wir betrachten eine Matrix

$$A = \left(a_{ij} \mid \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, m \end{array} \right),$$

mit den Spalten $\vec{a_1}, \dots, \vec{a_m} \in K^n$.

Wir wollen nun die Matrix verändern, indem wir Zeilen oder Spalten gewissen einfachen linearen Transformationen unterwerfen. Wenn wir das geschickt genug machen, dann bleibt der Lösungsraum erhalten oder wird nur unwesentlich verändert. Wir sprechen dann von *elementaren Umformungen.*

Typ (I):

Für $i \in \{1, \dots, n\}$ und $\lambda \in K$, $\lambda \neq 0$, sei

$$\varepsilon_{i,\lambda} : K^n \to K^n \text{ definiert durch } \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ \lambda \cdot x_i \\ \vdots \\ x_n \end{pmatrix}.$$

Das ist eine lineare Abbildung, ja sogar ein Isomorphismus mit Umkehrabbildung $\varepsilon_{i,1/\lambda}$.

Ist $A = (\vec{a}_1, \dots, \vec{a}_m)$, so wird bei der Transformation

$$A \mapsto \varepsilon_{i,\lambda}(A) := (\varepsilon_{i,\lambda}(\vec{a}_1), \dots, \varepsilon_{i,\lambda}(\vec{a}_m))$$

schlicht und ergreifend die i–te Zeile von A mit λ multipliziert.

Typ (II):

Für $i, k \in \{1, \ldots, n\}$ mit $i \neq k$ sei

$$\varepsilon_i^{+k} : K^n \to K^n \text{ definiert durch } \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_i + x_k \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix}.$$

Auch das ist eine lineare Abbildung, und auch ein Isomorphismus. Bei der Umkehrabbildung $\varepsilon_i^{-k} = \varepsilon_{k,-1} \circ \varepsilon_i^{+k} \circ \varepsilon_{k,-1}$ wird das k–te Element vom i–ten Element subtrahiert.

Bei der entsprechenden Matrizen–Operation

$$A \mapsto \varepsilon_i^{\pm k}(A) := (\varepsilon_i^{\pm k}(\vec{a}_1), \ldots, \varepsilon_i^{\pm k}(\vec{a}_m))$$

wird die k–te Zeile zur i–ten addiert oder von ihr subtrahiert.

Eine *elementare Zeilenumformung* ist eine beliebige Kombination von Transformationen vom Typ (I) oder (II). In diese Kategorie gehört zum Beispiel auch die Addition des λ–fachen der k–ten Zeile zur i–ten Zeile (mit $\lambda \neq 0$ und $i \neq k$). Das nennt man manchmal eine Umformung vom Typ (III).

Sogar Vertauschungen von Zeilen lassen sich so produzieren:

$$\begin{pmatrix} a_i \\ a_k \end{pmatrix} \mapsto \begin{pmatrix} a_i \\ -a_k \end{pmatrix} \mapsto \begin{pmatrix} a_i \\ a_i - a_k \end{pmatrix} \mapsto \begin{pmatrix} a_i - (a_i - a_k) \\ a_i - a_k \end{pmatrix} = \begin{pmatrix} a_k \\ a_i - a_k \end{pmatrix} \mapsto \begin{pmatrix} a_k \\ a_i \end{pmatrix}.$$

Zeilenvertauschungen nennen wir auch Umformungen vom Typ (IV).

Mit Spalten könnte man entsprechende Umformungen vornehmen, aber man darf beide Sorten nicht so ohne weiteres mischen. Deshalb werden wir bei den Spaltenoperationen nur *Vertauschungen von Spalten* zulassen!

Ist $A = (\vec{a}_1, \ldots, \vec{a}_m)$ eine Matrix und $\sigma \in S_m$ eine Permutation, so setzen wir $A_\sigma := (\vec{a}_{\sigma(1)}, \ldots, \vec{a}_{\sigma(m)})$. Das ist diejenige Matrix, die man aus A erhält, wenn man die Spalten gemäß σ permutiert.

Diese ganzen Umformungen sind natürlich nur dann sinnvoll, wenn dabei der Lösungsraum nicht oder nur unwesentlich verändert wird. Das wollen wir nun überprüfen.

I.2.5 Satz. *Wir betrachten ein inhomogenes LGS* $A \circ \vec{x} = \vec{b}$ *mit* $A \in M_{n,m}(K)$.

1. *Ist* ε *eine elementare Zeilenumformung, so ist*
$$\text{Lös}(\varepsilon(A), \varepsilon(\vec{b})) = \text{Lös}(A, \vec{b}).$$
2. *Ist* $\sigma \in S_m$ *und* $P_\sigma : K^m \to K^m$ *der Isomorphismus, der die Komponenten gemäß* σ *permutiert, so ist*
$$\text{Lös}(A_\sigma, \vec{b}) = P_\sigma(\text{Lös}(A, \vec{b})).$$

Fazit: Elementare Zeilenoperationen ändern am Lösungsraum überhaupt nichts, wenn man sie auf die erweiterte Matrix $(A, \vec{b})$ anwendet. Spaltenvertauschungen sorgen für eine Permutation der Komponenten der Lösungsvektoren. Darüber muß man sorgfältig Buch führen.

Beweis: Zu den Zeilenumformungen ε :

$$\begin{aligned} A\circ \vec{x}=\vec{b} &\iff x_1 \vec{a}_1 + \cdots + x_m \vec{a}_m = \vec{b} \\ &\iff \varepsilon(x_1 \vec{a}_1 + \cdots + x_m \vec{a}_m) = \varepsilon(\vec{b}) \\ &\iff x_1 \cdot \varepsilon(\vec{a}_1) + \cdots + x_m \cdot \varepsilon(\vec{a}_m) = \varepsilon(\vec{b}) \\ &\iff \varepsilon(A)\circ \vec{x} = \varepsilon(\vec{b}). \end{aligned}$$

Bei den Spalten–Vertauschungen müssen wir nur beachten, daß (wegen des Kommutativgesetzes) folgendes gilt:

$$x_1 \vec{a}_1 + \cdots + x_m \vec{a}_m = x_{\sigma(1)} \vec{a}_{\sigma(1)} + \cdots + x_{\sigma(m)} \vec{a}_{\sigma(m)} .$$

Daraus folgt:

$$A\circ \vec{x} = A_\sigma \circ P_\sigma(\vec{x}).$$

$\square$

Was wollen wir durch die Umformungen erreichen? Wir wollen die Matrix auf eine besondere Dreiecksgestalt bringen, so daß dann das LGS ganz einfach lösbar wird. Zu dem Zweck betrachten wir die folgende Klasse von Matrizen:

Definition.

Eine Matrix $A \in M_{n,m}(K)$ heißt r*–speziell*, wenn sie folgende Gestalt besitzt:

$$A = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \\ \hline & 0 & & C \end{array}\right),$$

mit $B \in M_{r,m-r}(K)$, $C \in M_{n-r,m-r}(K)$ und $a_{11}, \ldots, a_{rr} \neq 0$. Im Kästchen links oben stehen unterhalb der Diagonalen nur Nullen.

Beispiele :

1. Jede Matrix A ist 0–speziell! Die Teilmatrix B ist dann gar nicht vorhanden, es ist $A = C$.

2. Ist $r = m = n$, so ist A eine „obere Dreiecks–Matrix“.

3. Es muß natürlich $r \leq \min(n, m)$ sein.

   Ist $r = n \leq m$, so ist C nicht vorhanden.
   Ist $r = m \leq n$, so sind B und C beide nicht vorhanden.
   Ist $r < \min(n, m)$, so muß man B und C berücksichtigen.

I.2.6 Satz. *Sei A r–speziell, $C = 0$ oder nicht vorhanden. Dann gilt:*

1. *Die ersten r Spalten und die ersten r Zeilen sind jeweils linear unabhängig.*
2. *Es ist $\mathrm{rg}_s(A) = \mathrm{rg}_z(A) = r$.*
3. *Das Gleichungssystem $A \circ \vec{x} = \vec{b}$ ist genau dann lösbar, wenn $b_{r+1} = b_{r+2} = \ldots = b_n = 0$ ist. Insbesondere ist es immer lösbar, wenn $r = n$ ist.*

BEWEIS:
1) Sei $A = (\vec{a}_1, \ldots, \vec{a}_m)$, und $\sum_{j=1}^{r} \alpha_j \vec{a}_j = \vec{0}$, also

$$\alpha_1 \cdot \begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + \cdots + \alpha_r \cdot \begin{pmatrix} a_{1r} \\ \vdots \\ a_{rr} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}.$$

Dann ist $\alpha_r \cdot a_{rr} = 0$, also $\alpha_r = 0$. Aber dann ist auch $\alpha_{r-1} \cdot a_{r-1,r-1} = 0$, also $\alpha_{r-1} = 0$ usw. Die Spalten $\vec{a}_1, \ldots, \vec{a}_r$ sind linear unabhängig.

Die Zeilen von A seien mit $\mathbf{c}_1, \ldots, \mathbf{c}_n$ bezeichnet. Ist $\sum_{i=1}^{r} \beta_i \mathbf{c}_i = \mathbf{0}$, so ist insbesondere

$$\beta_1 \cdot (a_{11}, \ldots, a_{1r}) + \beta_2 \cdot (0, a_{22}, \ldots, a_{2r}) + \cdots + \beta_r \cdot (0, \ldots, 0, a_{rr}) = (0, \ldots, 0),$$

also

$$\begin{array}{ccccccccc} \beta_1 a_{11} & & & & & & & = & 0, \\ \beta_1 a_{12} & + & \beta_2 a_{22} & & & & & = & 0, \\ \vdots & & & & \ddots & & & & \vdots \\ \beta_1 a_{1r} & + & \beta_2 a_{2r} & + & \cdots & + & \beta_r a_{rr} & = & 0. \end{array}$$

Auch hier folgt offensichtlich, daß $\beta_1 = \beta_2 = \ldots = \beta_r = 0$ ist.

2) Es ist klar, daß $\mathrm{rg}_z(A) = r$ ist, denn es folgen höchstens noch Nullvektoren als Zeilen. Wir wissen schon, daß $\mathrm{rg}_s(A) \geq r$ ist. Daß es in A tatsächlich nicht mehr als r linear unabhängige Spalten geben kann, wird sich aus dem Beweis der Aussage (3) ergeben.

3) Wir müssen das folgende LGS betrachten:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \\ \hline & 0 & & 0 \end{array}\right) \circ \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ * \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ * \end{pmatrix}.$$

Ist das LGS lösbar, so muß offensichtlich $b_{r+1} = \ldots = b_n = 0$ sein.

Ist umgekehrt diese Bedingung erfüllt, so können wir $x_{r+1} = \ldots = x_m = 0$ setzen und brauchen dann nur noch das kleinere System

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{rr} \end{pmatrix} \circ \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}$$

zu lösen, also

$$\begin{array}{ccccccc} x_1 a_{11} & + & \cdots & + & x_r a_{1r} & = & b_1, \\ & & \ddots & & \vdots & & \vdots \\ & & & & x_r a_{rr} & = & b_r. \end{array}$$

Aber das ist mit „Rückwärtseinsetzen“ ganz einfach zu bewerkstelligen:
Es ist

$$\begin{aligned} x_r &= \frac{1}{a_{rr}} \cdot b_r, \\ x_{r-1} &= \frac{1}{a_{r-1,r-1}} \cdot (b_{r-1} - x_r a_{r-1,r}), \\ x_{r-2} &= \frac{1}{a_{r-2,r-2}} \cdot (b_{r-2} - x_r a_{r-2,r} - x_{r-1} a_{r-2,r-1}) \\ &\text{usw.} \end{aligned}$$

Jetzt können wir auch die Angelegenheit mit dem Spaltenrang erledigen: Das Gleichungssystem ist lösbar, wenn wir für $\vec{b}$ eine der Spalten $\vec{a}_{r+1}, \ldots, \vec{a}_m$ einsetzen. Aber das bedeutet, daß diese Spalte dann eine Linearkombination von $\vec{a}_1, \ldots, \vec{a}_r$ ist. Also kann $\text{rg}_s(A)$ nicht größer als r sein. $\square$

Definition.

Wir sagen, die Matrix A ist eine *Gauß–Matrix*, wenn A r–speziell ist, und $C = 0$ oder nicht vorhanden.

Bevor wir nun versuchen, ein LGS so umzuformen, daß es durch eine Gauß–Matrix beschrieben wird, wollen wir noch untersuchen, wie sich Spalten– und Zeilenrang bei den elementaren Umformungen verhalten.

I.2.7 Rang-Erhaltungssatz. *Bei elementaren Zeilenumformungen oder Vertauschungen von Spalten ändern sich Zeilenrang und Spaltenrang nicht.*

Dieser Satz ist **das** zentrale Ergebnis in unserer Theorie! Er liefert nicht nur den Schlüssel zur Auflösung Linearer Gleichungssysteme, er hat auch sehr wichtige Konsequenzen in der allgemeinen Vektorraum-Theorie.

Bevor wir zum Beweis schreiten, müssen wir noch einen Hilfssatz herleiten:

I.2.8 Austauschsatz. *Es seien $\mathbf{c}_1, \ldots, \mathbf{c}_n$ irgendwelche Vektoren im K^m. Die ersten r Vektoren seien linear unabhängig, und r sei auch die Maximalzahl linear unabhängiger Vektoren unter den $\mathbf{c}_i$. Weiter sei $\mathbf{d} \in \{\mathbf{c}_{r+1}, \ldots, \mathbf{c}_n\}$. Dann gilt:*

Es sind entweder $\mathbf{c}_1 + \mathbf{d}, \mathbf{c}_2, \ldots, \mathbf{c}_r$ *oder* $\mathbf{d}, \mathbf{c}_2, \ldots, \mathbf{c}_r$ *linear unabhängig.*

Beweis: Zunächst soll die Problematik an einem Beispiel erörtert werden:

Wir betrachten die Matrix $A := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}$ mit den Zeilen $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$. Offensichtlich sind $\mathbf{c}_1$ und $\mathbf{c}_2$ linear unabhängig, während $\mathbf{c}_3 = 1 \cdot \mathbf{c}_1 + (-1) \cdot \mathbf{c}_2$ ist.

Die Vektoren $\mathbf{c}_1$ und $\mathbf{c}_2 + \mathbf{c}_3 = \mathbf{c}_1$ sind linear abhängig. So etwas kann bei elementaren Zeilenumformungen passieren. Da jedoch auch $\mathbf{c}_1$ und $\mathbf{c}_3$ linear unabhängig sind, bleibt der Zeilenrang erhalten. Daß das immer funktioniert, zeigt der folgende Beweis:

Wir betrachten die im Satz beschriebene Situation. Offensichtlich ist

$$\mathbf{d} = \beta_1 \mathbf{c}_1 + \cdots + \beta_r \mathbf{c}_r,$$

mit irgendwelchen Koeffizienten β_i. Wir unterscheiden zwei Fälle:

1. Fall: $\beta_1 = 0$, also $\mathbf{d} = \beta_2 \mathbf{c}_2 + \cdots + \beta_r \mathbf{c}_r$.

Es sei $\alpha_1(\mathbf{c}_1 + \mathbf{d}) + \alpha_2 \mathbf{c}_2 + \cdots + \alpha_r \mathbf{c}_r = \mathbf{0}$. Die linke Seite können wir umformen zu

$$\alpha_1 \mathbf{c}_1 + (\alpha_1 \beta_2 + \alpha_2)\mathbf{c}_2 + \cdots + (\alpha_1 \beta_r + \alpha_r)\mathbf{c}_r.$$

Also muß gelten:

$$\alpha_1 = 0 \text{ und } \alpha_1 \beta_k + \alpha_k = 0 \text{ für } k = 2, \ldots, r.$$

Aber dann ist auch $\alpha_2 = \ldots = \alpha_r = 0$. Die Vektoren $\mathbf{c}_1 + \mathbf{d}$, $\mathbf{c}_2$, …, $\mathbf{c}_r$ sind in diesem Fall linear unabhängig.

2. Fall: $\beta_1 \neq 0$

Es sei $\alpha_1 \mathbf{d} + \alpha_2 \mathbf{c}_2 + \cdots + \alpha_r \mathbf{c}_r = \mathbf{0}$. Hier können wir die linke Seite umformen zu

$$\alpha_1 \beta_1 \mathbf{c}_1 + (\alpha_1 \beta_2 + \alpha_2)\mathbf{c}_2 + \cdots + (\alpha_1 \beta_r + \alpha_r)\mathbf{c}_r.$$

Dann muß $\alpha_1 = 0$ und damit auch $\alpha_2 = \ldots = \alpha_r = 0$ sein. Diesmal sind $\mathbf{d}$, $\mathbf{c}_2$, …, $\mathbf{c}_r$ linear unabhängig. $\square$

Nun zum Beweis des Rang-Erhaltungssatzes:

1) Sei $A = (\vec{a}_1, \ldots, \vec{a}_m)$. Es ist klar, daß sich der Spaltenrang bei einer Vertauschung der Spalten nicht ändert. Wir können deshalb o.B.d.A.[7] annehmen, daß die ersten r Spalten linear unabhängig sind, und daß $\text{rg}_s(A) = r$ ist.

Ist nun $\varepsilon : K^n \to K^n$ eine elementare Zeilen-Umformung, so ist ε insbesondere ein Isomorphismus. Das bedeutet, daß die Vektoren

$$\varepsilon(\vec{a}_1), \ldots, \varepsilon(\vec{a}_r)$$

ebenfalls linear unabhängig sind. Wenn es unter den Bildvektoren $\varepsilon(\vec{a}_j)$ sogar $r+1$ linear unabhängige gäbe, dann könnte man darauf den Isomorphismus ε^{-1} anwenden und hätte auch $r+1$ linear unabhängige Vektoren unter den $\vec{a}_j$. Da dem nicht so ist, muß auch $\text{rg}_s(\varepsilon(A)) = r$ sein.

[7] „Ohne Beschränkung der Allgemeinheit“, also ohne unzulässige Voraussetzungen.

2) Nun zum Zeilenrang: Es sei

$$A = \begin{pmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_n \end{pmatrix}$$

die Aufteilung von A in Zeilen. Ist $\sigma \in S_m$ eine Permutation, so ist

$$A_\sigma = (\vec{a}_{\sigma(1)}, \ldots, \vec{a}_{\sigma(m)}) = \begin{pmatrix} P_\sigma(\mathbf{c}_1) \\ \vdots \\ P_\sigma(\mathbf{c}_n) \end{pmatrix}.$$

Da $P_\sigma : K^m \to K^m$ ein Isomorphismus ist, ändert sich bei dieser Prozedur der Zeilenrang nicht.

Bei den elementaren Zeilenumformungen reicht es, wenn wir die vom Typ (I) und (II) betrachten, denn alle anderen ergeben sich ja aus diesen. Wir nehmen aber noch den Typ (IV) vorweg, die Vertauschung von Zeilen. Da ist es klar, daß sich der Zeilenrang nicht ändert. Also können wir o.B.d.A. annehmen, daß die ersten r Zeilen linear unabhängig sind, und daß $r = \mathrm{rg}_z(A)$ ist.

Beginnen wir mit der Transformation $\varepsilon_{i,\lambda}$. Ist $i > r$, so hat sie sicher keinen Einfluß auf $\mathrm{rg}_z(A)$. Ist $1 \le i \le r$ und

$$\alpha_1 \cdot \mathbf{c}_1 + \cdots + \alpha_i \cdot (\lambda \cdot \mathbf{c}_i) + \cdots + \alpha_r \cdot \mathbf{c}_r = \mathbf{0},$$

so ist $\alpha_1 = \ldots = \alpha_i\lambda = \ldots = \alpha_r = 0$, und das ist – da $\lambda \neq 0$ ist – nur möglich, wenn auch $\alpha_i = 0$ ist. Der Zeilenrang kann also nicht kleiner werden.

Nun schauen wir, was bei der Transformation ε_i^{+k} passiert: Ist $k \le r$ und

$$\alpha_1\mathbf{c}_1 + \cdots + \alpha_i(\mathbf{c}_i + \mathbf{c}_k) + \cdots + \alpha_r\mathbf{c}_r = \mathbf{0},$$

so muß

$$\alpha_1 = \ldots = \alpha_i = \ldots = \alpha_i + \alpha_k = \ldots = \alpha_r = 0$$

sein, wobei die Gleichung $\alpha_k = 0$ zunächst fehlt, aber sofort gefolgert werden kann. Also sind die Zeilenvektoren

$$\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{c}_i + \mathbf{c}_k, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_r$$

linear unabhängig, und der Zeilenrang kann höchstens größer werden.

Ist $\mathbf{d} := \mathbf{c}_k$ eine der Zeilen $\mathbf{c}_{r+1}, \ldots, \mathbf{c}_n$, so folgt aus dem Austauschsatz, daß der Zeilenrang bei Anwendung von ε_1^{+k} (und analog bei ε_i^{+k}) nicht kleiner wird.

Wie beim Spaltenrang können wir nun argumentieren, daß sich alle Zeilenumformungen umkehren lassen und daß der Zeilenrang deshalb auch nicht größer werden kann. $\square$

Nun gehen wir an die Lösung eines allgemeinen LGS:

I.2.9 Gauß'sches Eliminationsverfahren.
Gegeben sei ein LGS $A \circ \vec{x} = \vec{b}$, *mit* $A \in M_{n,m}(K)$ *und* $\mathrm{rg}_s(A) = r$. *Durch endlich viele elementare Zeilenumformungen und Spaltenvertauschungen kann man die Koeffizientenmatrix in eine* r*–spezielle Gauß–Matrix verwandeln.*

Beweis:
1) Jede Matrix, also auch A, ist 0–speziell.

2) Reduktions–Schritt:

Es sei $0 \le k < r$. Wir nehmen an, A sei schon k–speziell:

$$A = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1k} & \\ \vdots & \ddots & \vdots & B_k \\ 0 & \cdots & a_{kk} & \\ \hline & 0 & & C_k \end{array}\right),$$

mit $B_k \in M_{k,m-k}(K)$, $C_k \in M_{n-k,m-k}(K)$ und $a_{11}, \ldots, a_{kk} \ne 0$. Wäre $C_k = 0$, so wäre $\mathrm{rg}_s(A) = k$. Das ist nicht der Fall, also muß $C_k \ne 0$ sein.

Das bedeutet, daß man ein a_{ij} in C_k finden kann, das $\ne 0$ ist. Wir nennen dieses Element das *Pivot–Element*. Es ist nicht eindeutig bestimmt, und es ist Inhalt umfangreicher numerischer Untersuchungen, dieses Pivot–Element möglichst geschickt zu wählen. Für uns hier reicht aber die Existenz eines solchen $a_{ij} \ne 0$. Durch Zeilenvertauschungen und Spaltenvertauschungen kann man erreichen, daß es das Element $a_{k+1,k+1}$ ist.

Nun subtrahiert man Vielfache der $(k+1)$–ten Zeile von den folgenden Zeilen, um die Elemente $a_{k+2,k+1}$, …, $a_{n,k+2}$ zum Verschwinden zu bringen:

$$A \longrightarrow \left(\begin{array}{ccc|cc} a_{11} & \cdots & a_{1k} & & \\ \vdots & \ddots & \vdots & B_k & \\ 0 & \cdots & a_{kk} & & \\ \hline 0 & \cdots & 0 & \multicolumn{1}{c|}{a_{k+1,k+1}} & \\ \cline{4-4} \vdots & & \vdots & \multicolumn{1}{c|}{0} & * \\ \vdots & & \vdots & \multicolumn{1}{c|}{\vdots} & \\ 0 & \cdots & 0 & \multicolumn{1}{c|}{0} & \end{array}\right).$$

Das Ergebnis ist eine $(k+1)$–spezielle Matrix.

3) Abschluß des Verfahrens:

Ist schließlich $k = r$ erreicht, so sind die hinteren $m - r$ Spalten von den ersten r Spalten linear abhängig (denn der Spaltenrang bleibt ja immer gleich, und die ersten r Spalten sind linear unabhängig!). Das geht nur, wenn die rechts unten verbliebene Matrix C_r die Null–Matrix ist. Aber das bedeutet, daß wir eine Gauß-Matrix erhalten haben. □

I.2.10 Folgerung 1. *Bei einer beliebigen Matrix A ist stets* $\mathrm{rg}_s(A) = \mathrm{rg}_z(A)$.

Beweis: Es stimmt für Gauß–Matrizen, und man kann A in eine solche umformen, ohne Spalten– und Zeilenrang zu verändern. □

Wir sprechen daher künftig einfach vom *Rang* einer Matrix (in Zeichen: $\mathrm{rg}(A)$).

I.2.11 Folgerung 2. *Es gibt im K^n höchstens n linear unabhängige Vektoren.*

Beweis: Sind $\vec{a}_1, \ldots, \vec{a}_m$ linear unabhängige (Spalten–)Vektoren im K^n, so kann man sie zu einer Matrix $A \in M_{n,m}(K)$ zusammenfassen. Dann gilt:

$$m = \mathrm{rg}_s(A) = \mathrm{rg}_z(A) \le n.$$

□

I.2.12 Folgerung 3. *Jeder Untervektorraum des K^n besitzt eine Basis.*

Der BEWEIS wurde schon früher gebracht, unter Vorwegnahme von Folgerung 2.

Bevor wir zu weiteren Folgerungen kommen, wollen wir ein Beispiel durchrechnen:

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 3, \\ x_1 & - & x_2 & - & x_3 & = & 4, \\ x_1 & + & 3x_2 & + & 3x_3 & = & 1. \end{array}$$

Wir müssen die folgende erweiterte Matrix betrachten:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -1 & -1 & 4 \\ 1 & 3 & 3 & 1 \end{array}\right)$$

Da $a_{11} = 1 \neq 0$ ist, braucht man beim ersten Schritt nichts zu vertauschen. Man kann bereits a_{11} als Pivot–Element benutzen. Subtraktion der 1. Zeile von der 2. und 3. Zeile ergibt:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array}\right)$$

Diese Matrix ist 1–speziell. Als neues Pivot–Element kann man $a_{22} = -2$ benutzen. Addiert man die 2. Zeile zur 3. Zeile, so erhält man:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & -1 \end{array}\right)$$

Der Rang der ursprünglichen Matrix ist also 2, der Rang der erweiterten Matrix aber 3. Das LGS besitzt keine Lösung!

Wie man hier deutlich sehen konnte, dient das Gauß'sche Eliminationsverfahren auch zur Bestimmung des Ranges. Man braucht den Rang nicht schon vorher zu kennen

Ändert man in der Ausgangsgleichung $A\circ \vec{x}=\vec{b}$ den Vektor $\vec{b}= (3,4,1)^\top$ zu $(3,4,2)^\top$, so ergibt sich nach den Umformungen folgendes System:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Dieses (überbestimmte) LGS ist lösbar. Um eine spezielle Lösung $\vec{x}_0$ zu erhalten, kann man x_3 beliebig wählen, etwa $x_3 = 0$. Dann ist

$$\begin{array}{ccccc} x_1 & + & x_2 & = & 3, \\ & & -2x_2 & = & 1, \end{array} \qquad \text{also } \vec{x}_0= \frac{1}{2}\cdot \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix}.$$

Wie bekommt man nun die allgemeine Lösung des zugehörigen homogenen Systems? Wir müssen folgendes System lösen:

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 0, \\ & - & 2x_2 & - & 2x_3 & = & 0. \end{array}$$

Setzt man für x_3 einen beliebigen Parameter λ ein, so wird daraus:

$$\begin{array}{rcrcl} x_1 & + & x_2 & = & -\lambda, \\ & - & 2x_2 & = & 2\lambda. \end{array}$$

Das ergibt $\vec{x}= \begin{pmatrix} 0 \\ -\lambda \\ \lambda \end{pmatrix}$. Also ist

$$\text{Lös}(A, \vec{b}) = \{\vec{x} \in K^3 \mid \exists\, \lambda \in \mathbb{R} \text{ mit } \vec{x} = \frac{1}{2} \cdot \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -\lambda \\ \lambda \end{pmatrix}\}.$$

Das hier angewandte Verfahren zur Bestimmung der allgemeinen Lösung der homogenen Gleichung können wir verallgemeinern:

Wir betrachten die Gleichung $A \circ \vec{x} = \vec{0}$, wobei wir aber schon voraussetzen können, daß $A \in M_{n,m}(K)$ eine Gauß–Matrix vom Rang r ist. Da außerdem die Zeilen $\mathbf{c}_{r+1}, \ldots, \mathbf{c}_n$ alle $= \mathbf{0}$ sind und nichts zum Problem beitragen, können wir uns auf den Fall $n = r$ beschränken:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \end{array}\right) \circ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Die obere Dreiecksmatrix auf der linken Seite von A bezeichnen wir mit D, und den Vektor $\vec{x}$ teilen wir auf in zwei Vektoren $\vec{x}^{\,*} \in K^r$ und $\vec{x}^{\,**} \in K^{m-r}$, so daß gilt:

$$A = (D, B) \quad \text{und} \quad \vec{x} = \begin{pmatrix} \vec{x}^{\,*} \\ \vec{x}^{\,**} \end{pmatrix}.$$

Dann folgt:

$$\begin{aligned} A \circ \vec{x} = \vec{0} \quad &\iff \quad D \circ \vec{x}^{\,*} + B \circ \vec{x}^{\,**} = \vec{0} \\ &\iff \quad D \circ \vec{x}^{\,*} = -B \circ \vec{x}^{\,**}. \end{aligned}$$

Das ursprüngliche Gleichungssystem ist überbestimmt, und wie bei dem obigen Beispiel fassen wir jetzt $\vec{x}^{\,**}$ als Parameter auf. Danach müssen die (in $\vec{x}^{\,*}$ zusammengefaßten) verbliebenen Komponenten $x_1, \ldots, x_r$ festgelegt sein.

Tatsächlich ist für *jeden* Vektor $\vec{y} \in K^{m-r}$ das LGS $D \circ \vec{x}^{\,*} = \vec{y}$ eindeutig lösbar. Es hat ja die Gestalt

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1r}x_r & = & y_1 \\ & & \ddots & & \vdots & & \vdots \\ & & & & a_{rr}x_r & = & y_r\,, \end{array}$$

mit $a_{ii} \neq 0$ für $i = 1, \ldots, r$.

Das gilt dann insbesondere für $\vec{y} := -B \circ \vec{x}^{\,**}$, wenn $\vec{x}^{\,**}$ irgendwie gewählt wird.

Durchläuft $\vec{x}^{**}$ ganz K^{m-r}, so erhalten wir auf diesem Wege die Lösungsgesamtheit. Ist $\{\vec{e}_1, \ldots, \vec{e}_{m-r}\}$ die Standard–Basis des K^{m-r} und $\vec{z}_j$ jeweils der eindeutig bestimmte Vektor mit $D\circ \vec{z}_j = -B\circ \vec{e}_j$, für $j = 1, \ldots, m-r$, dann bilden die Vektoren

$$\vec{x}_j := \begin{pmatrix} \vec{z}_j \\ \vec{e}_j \end{pmatrix}, \quad j = 1, \ldots, m-r\,,$$

eine Basis des Lösungsraumes $\text{Lös}(A, \vec{0}) = \text{Ker}\,(F_A)$.

I.2.13 Folgerung.
Ist $A \in M_{n,m}(K)$, so besitzt $\text{Lös}(A, \vec{0})$ *eine Basis von $m - \text{rg}(A)$ Elementen.*

Ist $\text{rg}(A) < m$ *(sind also die Spalten von A linear abhängig), so gibt es eine nicht–triviale Lösung von $A\circ \vec{x} = \vec{0}$.*

Beweis:
1) Da man A zu einer Gauß–Matrix vom Rang $r = \text{rg}(A)$ umformen kann, folgt die erste Behauptung unmittelbar aus den obigen Überlegungen.

2) Ist $\text{rg}(A) < m$, so gibt es $m - \text{rg}(A) > 0$ linear unabhängige Lösungsvektoren, also mindestens einen $\neq \vec{0}$. □

Das können wir für den folgenden grundlegenden Satz benutzen:

I.2.14 Satz. *Im K^n besitzt* **jede** *Basis genau n Elemente.*

Beweis: Wir wissen schon, daß eine Basis höchstens n Elemente besitzen kann. Nun nehmen wir an, es gibt eine Basis $\{\vec{a}_1, \ldots, \vec{a}_m\}$ von K^n mit $m < n$.

Die Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n$ müssen natürlich als Linearkombinationn der $\vec{a}_j$ geschrieben werden können:

$$\vec{e}_i = \sum_{j=1}^{m} \alpha_{ji}\, \vec{a}_j, \text{ für } i = 1, \ldots, n.$$

Wir fassen die α_{ji} zu einer Matrix

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \in M_{m,n}(K)$$

zusammen. Dann ist $\text{rg}(A) \leq \min(n, m) = m < n$. Das bedeutet, daß das homogene LGS $A\circ \vec{x} = \vec{0}$ mindestens eine nicht-triviale Lösung $\vec{x} \in K^n$ besitzt. Dann ist

$$\begin{aligned} \vec{x} = \sum_{i=1}^{n} x_i\, \vec{e}_i &= \sum_{i=1}^{n} x_i (\sum_{j=1}^{m} \alpha_{ji}\, \vec{a}_j) \\ &= \sum_{j=1}^{m} (\sum_{i=1}^{n} \alpha_{ji} x_i)\, \vec{a}_j = \vec{0}, \end{aligned}$$

denn $\sum_{i=1}^{n} \alpha_{ji} x_i$ ist die j–te Komponente von $A\circ \vec{x}$, also 0 =. Aber das ist unmöglich, denn $\vec{x}$ enthält Komponenten, die $\neq 0$ sind. Die Annahme war falsch. □

I.2.15 Folgerung. *Ist $K^n \cong K^m$, so ist $n = m$.*

BEWEIS: Sei $F : K^n \to K^m$ ein Isomorphismus. Dann ist $\{F(\vec{e}_1), \ldots, F(\vec{e}_n)\}$ eine Basis von K^m, und das ist nur möglich, wenn $n = m$ ist. □

I.2.16 Folgerung. *Sei V ein K–Vektorraum. Wenn V eine Basis mit n Elementen besitzt, dann besitzt auch jede andere Basis von V genau n Elemente.*

BEWEIS: Wenn V eine Basis mit n Elementen besitzt, dann gibt es einen Isomorphismus $F : V \to K^n$. Hätte V noch eine Basis mit m Elementen, so gäbe es auch einen Isomorphismus $G : V \to K^m$. Dann wäre $G \circ F^{-1} : K^n \to K^m$ ebenfalls ein Isomorphismus, und das ist nur möglich, wenn $n = m$ ist. □

Beispiel:

Wir wollen jetzt ein etwas größeres Gleichungssystem betrachten:

$$\begin{array}{rcrcrcrcr}
x_1 & + & 3x_2 & - & 4x_3 & + & 3x_4 & = & 9, \\
3x_1 & + & 9x_2 & - & 2x_3 & - & 11x_4 & = & -3, \\
4x_1 & + & 12x_2 & - & 6x_3 & - & 8x_4 & = & 6, \\
2x_1 & + & 6x_2 & + & 2x_3 & - & 14x_4 & = & -12.
\end{array}$$

Für die systematische Anwendung des Gauß'schen Eliminationsverfahrens benutzen wir folgendes Schema:

x_1	x_2	x_3	x_4	b_i	
1	3	−4	3	9	
3	9	−2	−11	−3	
4	12	−6	−8	6	
2	6	2	−14	−12	
1	3	−4	3	9	bleibt stehen
0	0	10	−20	−30	(−3 × 1. Zeile)
0	0	10	−20	−30	(−4 × 1. Zeile)
0	0	10	−20	−30	(−2 × 1. Zeile)
x_1	x_3	x_2	x_4		Spaltenvertauschung
1	−4	3	3	9	
0	10	0	−20	−30	
0	10	0	−20	−30	
0	10	0	−20	−30	
1	−4	3	3	9	
0	10	0	−20	−30	bleibt stehen
0	0	0	0	0	(− 2. Zeile)
0	0	0	0	0	(− 2. Zeile)

Das ergibt (z.B.) die folgende spezielle Lösung:

$$\begin{aligned}
x_4 &= 0 \text{ (willkürlich gewählt)}, \\
x_2 &= 0 \text{ (ebenfalls willkürlich)}, \\
x_3 &= -3, \\
\text{und } x_1 &= 9 - 12 = -3.
\end{aligned}$$

Eine Basis für den Lösungsraum des zugehörigen homogenen Systems erhält man, indem man für $(x_2, x_4)^\top$ die beiden möglichen Einheitsvektoren einsetzt und dann die zugehörigen Werte von x_1 und x_3 bestimmt:

$x_2 = 1$ und $x_4 = 0$ ergibt

$$\begin{aligned} x_1 - 4x_3 + 3 &= 0 \\ \text{und} \quad 10x_3 &= 0, \end{aligned}$$

also $x_1 = -3$ und $x_3 = 0$.

$x_2 = 0$ und $x_4 = 1$ ergibt

$$\begin{aligned} x_1 - 4x_3 + 3 &= 0 \\ \text{und} \quad 10x_3 - 20 &= 0, \end{aligned}$$

also $x_1 = 5$ und $x_3 = 2$.

Die allgemeine Lösung des inhomogenen Systems ist also gegeben durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ -3 \\ 0 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} 5 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 - 3\alpha + 5\beta \\ \alpha \\ -3 + 2\beta \\ \beta \end{pmatrix}, \text{ mit } \alpha, \beta \in \mathbb{R}.$$

Zum Schluß noch etwas Theorie:

Definition.

Sei V ein K-Vektorraum. Wenn V eine Basis mit endlich vielen Elementen besitzt, so nennt man die (eindeutig bestimmte) Anzahl der Elemente dieser Basis die *Dimension* von V (in Zeichen: $\dim(V)$ oder $\dim_K(V)$).

Wenn V nicht der Nullraum ist und keine Basis mit endlich vielen Elementen besitzt, dann nennt man V *unendlich–dimensional*.

Ist $\mathbf{a} \in V$ und $U \subset V$ ein Untervektorraum, so erklärt man die *Dimension* des affinen Raumes $\mathbf{a} + V$ durch

$$\dim(\mathbf{a} + V) := \dim(V).$$

I.2.17 Satz. $\dim_K(K^n) = n$, *insbesondere* $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Der Beweis ist schon weiter oben erbracht worden.

I.2.18 Satz. *Ist* $V \cong W$, *so ist* $\dim(V) = \dim(W)$.

Umgekehrt gilt auch: Ist $\dim(V) = \dim(W) < \infty$, *so ist* $V \cong W$.

Beweis: Die erste Aussage ist trivial.

Ist $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ eine Basis von V und $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ eine Basis von W, so wird durch $\mathbf{a}_i \mapsto \mathbf{b}_i$ (für $i = 1, \dots, n$) ein Isomorphismus von V nach W definiert. □

I.2.19 Satz. *Sei $A \in M_{n,m}(K)$. Dann ist*

$$\mathrm{rg}(A) + \dim_K(\mathrm{Lös}(A, \vec{0})) = m.$$

Auch hierfür haben wir den Beweis schon geführt.

I.2.20 Dimensions-Formel für lineare Abbildungen.
Sei $F = F_A : K^m \to K^n$ eine lineare Abbildung. Dann gilt:

$$\dim_K(\mathrm{Ker}\,(F)) + \dim_K(\mathrm{Im}\,(F)) = m.$$

Beweis:
Zunächst ist $\mathrm{Ker}\,(F_A) = \mathrm{Lös}(A, \vec{0})$

Wir müssen noch zeigen, daß $\mathrm{rg}(A) = \dim_K(\mathrm{Im}\,(F_A))$ ist. Es gilt aber:
Ist $A = (\vec{a}_1, \dots, \vec{a}_m)$, so ist

$$\mathrm{Im}\,(F_A) = < F_A(\vec{e}_1), \dots, F(\vec{e}_m) > = < \vec{a}_1, \dots, \vec{a}_m > .$$

Das bedeutet, daß die $\vec{a}_j$ ein Erzeugendensystem für $\mathrm{Im}\,(F_A)$ bilden. Wählt man aus ihnen ein maximales System linear unabhängiger Vektoren aus, so erhält man eine Basis für $\mathrm{Im}\,(F_A)$ mit $\mathrm{rg}(A)$ Elementen. □

I.2.21 Folgerung. *Sei $F : K^n \to K^n$ linear. Dann sind die folgenden Aussagen äquivalent:*

1. *F ist injektiv.*
2. *F ist surjektiv.*
3. *F ist ein Isomorphismus.*

Beweis: Es ist $\dim_K(\mathrm{Ker}\,(F)) + \dim_K(\mathrm{Im}\,(F)) = n$. Dann gilt:

$$\begin{aligned} F \text{ injektiv} &\iff \mathrm{Ker}\,(F) = \{0\} \\ &\iff \dim(\mathrm{Ker}\,(F)) = 0 \\ &\iff \dim(\mathrm{Im}\,(F)) = n \\ &\iff \mathrm{Im}\,(F) = K^n \\ &\iff F \text{ surjektiv} . \end{aligned}$$

Insbesondere ist F schon dann bijektiv, wenn F nur injektiv oder surjektiv ist. □

§3 Matrizen und Lineare Abbildungen

Eine Matrix $A \in M_{n,m}(K)$ definiert eine lineare Abbildung

$$F_A : K^m \to K^n \text{ mit } F_A(\vec{x}) := A \circ \vec{x}\ .$$

Sind $\vec{e}_1, \ldots, \vec{e}_m$ die Einheitsvektoren im K^m, so ist

$$\vec{a}_j := A \circ \vec{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \text{ die j-te Spalte von } A.$$

Offensichtlich gilt dann für $\vec{x} = \sum_{j=1}^{m} x_j \vec{e}_j$: $\qquad A \circ \vec{x} = \sum_{j=1}^{m} x_j \vec{a}_j$.

Wir wollen jetzt die Beziehung zwischen linearen Abbildungen und Matrizen noch etwas genauer untersuchen. Es sei $L(K^m, K^n)$ die Menge aller linearen Abbildungen von K^m nach K^n. Man kann (wie bei den Funktionen oder den Polynomen) zwei lineare Abbildungen addieren oder eine lineare Abbildung mit einem Skalar multiplizieren:

1. Sind $f,\ g \in L(K^m, K^n)$, so definiert man $f + g \in L(K^m, K^n)$ durch

$$(f + g)(\vec{x}) := f(\vec{x}) + g(\vec{x}).$$

2. Ist $f \in L(K^m, K^n)$ und $\alpha \in K$, so definiert man $\alpha \cdot f \in L(K^m, K^n)$ durch

$$(\alpha \cdot f)(\vec{x}) := \alpha \cdot f(\vec{x}).$$

Es ist sehr leicht zu sehen, daß $L(K^m, K^n)$ auf diese Weise zu einem Vektorraum wird.

Von der Menge $M_{n,m}(K)$ der K-wertigen Matrizen mit n Zeilen und m Spalten wissen wir schon, daß sie ebenfalls die Struktur eines Vektorraums trägt. Nun gilt:

> *Durch $A \mapsto F_A$ wird ein Isomorphismus zwischen den Vektorräumen $M_{n,m}(K)$ und $L(K^m, K^n)$ definiert.*

Zum Beweis muß die Linearität und die Bijektivität der Zuordnung nachgerechnet werden. Wir verzichten hier darauf, merken uns aber, daß strukturell zwischen den Matrizen aus $M_{n,m}(K) \cong K^{n \cdot m}$ und den linearen Abbildungen aus $L(K^m, K^n)$ kein Unterschied besteht.

Beispiele :

1. Die *identische Abbildung* $\mathrm{id}_{K^n} : K^n \to K^n$ mit $\mathrm{id}(\vec{x}) = \vec{x}$ wird durch die *Einheitsmatrix*

$$1_n := \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \in M_{n,n}(K)$$

beschrieben. Häufig schreibt man auch E oder E_n an Stelle von 1_n.

2. Sei $r < n$. Dann definiert man die *Einbettung* $i = i_{r,n} : K^r \to K^n$ durch

$$i((x_1, \ldots, x_r)^\top) := (x_1, \ldots, x_r, 0, \ldots, 0)^\top.$$

   Die fehlenden Komponenten werden einfach mit Nullen aufgefüllt. Diese Abbildung ist linear und injektiv, sie bildet K^r isomorph auf den Untervektorraum

$$\mathrm{Im}\,(i) = \{(x_1, \ldots, x_n)^\top \in K^n \mid x_{r+1} = \ldots = x_n = 0\}$$

   des K^n ab.

   Man identifiziert K^r auf diesem Wege mit $\mathrm{Im}\,(i)$ und tut damit so, als ob K^r schon selbst ein Unterraum des K^n wäre.

   In Matrizenschreibweise wird i gegeben durch die Matrix

$$A = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ \hline & 0 & \end{pmatrix} \in M_{n,r}(K).$$

3. Die *Projektion* von K^n auf K^r ist die Abbildung $p = p_{n,r} : K^n \to K^r$ mit

$$p((x_1, \ldots, x_n)^\top) := (x_1, \ldots, x_r)^\top.$$

   Es werden einfach die letzten Koordinaten $x_{r+1}, \ldots, x_n$ weggelassen. Die zugehörige Matrix $A \in M_{r,n}(K)$ ist gegeben durch

$$A = \left(\begin{array}{ccc|c} 1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & 0 \\ 0 & \cdots & 1 & \end{array}\right).$$

   Man beachte, daß $p \circ i = \mathrm{id}_{K^r}$ ist, daß die Abbildungen aber nicht invers zueinander sind, denn es gilt:

$$i \circ p((x_1, \ldots, x_n)^\top) = (x_1, \ldots, x_r, 0, \ldots, 0)^\top.$$

4. Sei $\sigma \in \mathcal{S}_n$ eine Permutation, $P_\sigma : K^n \to K^n$ definiert durch

$$P_\sigma(x_1, \ldots, x_n) := (x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

   Dann ist $P_\sigma(\vec{e}_j) = \vec{e}_{\sigma(j)}$, für $j = 1, \ldots, n$. Also wird P_σ durch die „Permutationsmatrix“

$$E_\sigma := (\vec{e}_{\sigma(1)}, \ldots, \vec{e}_{\sigma(n)}) = (\,1_n)_\sigma$$

   beschrieben.

Wir betrachten jetzt **zwei** lineare Abbildungen

$$g : K^l \to K^m \text{ und } f : K^m \to K^n.$$

Dann gibt es Matrizen $A \in M_{n,m}(K)$ und $B \in M_{m,l}(K)$, so daß $f = F_A$ und $g = F_B$ ist.

Wie steht es nun mit der zusammengesetzten Abbildung $f \circ g : K^l \to K^n$? Es muß eine Matrix $C \in M_{n,l}(K)$ mit $F_C = f \circ g$ geben. Dadurch erhalten wir eine Verknüpfung

$$\begin{array}{ccccc} M_{n,m}(K) & \times & M_{m,l}(K) & \to & M_{n,l}(K) \\ (\quad A & , & B \quad) & \mapsto & A \circ B := C, \end{array}$$

mit $F_{A\circ B} = F_A \circ F_B$.

I.3.1 Satz. *Sei* $A = \left(a_{ij} \Big| \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, m \end{array} \right)$ *und* $B = \left(b_{jk} \Big| \begin{array}{l} j = 1, \dots, m \\ k = 1, \dots, l \end{array} \right)$.

Dann ist $A \circ B = \left(c_{ik} \Big| \begin{array}{l} i = 1, \dots, n \\ k = 1, \dots, l \end{array} \right)$, *mit* $\boxed{c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}}$.

Beweis: c_{ik} ist der i-te Eintrag in der k-ten Spalte von $C = A \circ B$.

Wir benutzen die Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_l \in K^l$ und schreiben die Matrizen mit ihren Spaltenvektoren wie folgt:

$$A = (\vec{a}_1, \dots, \vec{a}_m) \text{ und } B = (\vec{b}_1, \dots, \vec{b}_l), \text{ mit } \vec{a}_j \in K^n,\ \vec{b}_k \in K^m.$$

Dann ist c_{ik} der i-te Eintrag in

$$\begin{aligned} (A \circ B) \circ \vec{e}_k &= F_{A\circ B}(\vec{e}_k) \\ &= F_A(F_B(\vec{e}_k)) \\ &= A \circ (B \circ \vec{e}_k) \\ &= A \circ \vec{b}_k \\ &= (\vec{a}_1, \dots, \vec{a}_m) \circ \begin{pmatrix} b_{1k} \\ \vdots \\ b_{mk} \end{pmatrix} = \sum_{j=1}^{m} b_{jk} \vec{a}_j, \end{aligned}$$

also $c_{ik} = \sum_{j=1}^{m} b_{jk} a_{ij}$. □

Das bedeutet:

> *Das Element in der i-ten Zeile und k-ten Spalte von $A \circ B$ erhält man, indem man die i-te Zeile von A auf die k-te Spalte von B legt, die dann übereinander liegenden Elemente multipliziert und schließlich alle Produkte aufsummiert.*

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ \boxed{a_{i1} \ \cdots \ a_{im}} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \circ \begin{pmatrix} b_{11} & \cdots & \boxed{\begin{matrix} b_{1k} \\ \vdots \\ b_{mk} \end{matrix}} & \cdots & b_{1l} \\ \vdots & & & & \vdots \\ b_{m1} & \cdots & & \cdots & b_{ml} \end{pmatrix} = \begin{array}{c} \begin{pmatrix} & \vdots & \\ \cdots & \boxed{\sum_{j=1}^{m} a_{ij} b_{jk}} & \cdots \\ & \vdots & \end{pmatrix} \leftarrow i \\ \uparrow k \end{array}$$

Es ist also wichtig, daß die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt.

Es folgen nun einige Rechenregeln für das Matrizenprodukt:

I.3.2 Satz.
Seien $A, A' \in M_{n,m}(K)$, $B, B' \in M_{m,l}(K)$ und $C \in M_{l,r}(K)$. Dann gilt:

1. $A \circ (B \circ C) = (A \circ B) \circ C$.
2. $A \circ (B + B') = A \circ B + A \circ B'$.
3. $(A + A') \circ B = A \circ B + A' \circ B$.
4. $\alpha \cdot (A \circ B) = (\alpha A) \circ B = A \circ (\alpha B)$ *für* $\alpha \in K$.
5. $1 \circ A = A$ *und* $A \circ 1 = A$.

Der BEWEIS erfolgt durch Nachrechnen.

Achtung!! Im allgemeinen ist $A \circ B \neq B \circ A$!!!
Hier ist ein einfaches Beispiel:

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 1\cdot 2 & 1\cdot 0 + 1\cdot 2 \\ 0\cdot 1 + 2\cdot 2 & 0\cdot 0 + 2\cdot 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 4 & 4 \end{pmatrix},$$

aber

$$\begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 0\cdot 0 & 1\cdot 1 + 0\cdot 2 \\ 2\cdot 1 + 2\cdot 0 & 2\cdot 1 + 2\cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 6 \end{pmatrix}.$$

Es gibt noch mehr seltsame Effekte: z.B. kann ein Matrizenprodukt die Null ergeben, auch wenn keiner der der beiden Faktoren Null ist:

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

Das bedeutet, daß man bei Matrizengleichungen i.a. nicht kürzen kann: Wenn $X \circ A = X \circ B$ ist, mit $X \neq 0$, so muß keineswegs $A = B$ gelten!

Es kann allerdings vorkommen, daß eine Matrix ein Inverses besitzt:

Definition.

Eine quadratische Matrix $A \in M_{n,n}(K)$ heißt *invertierbar*, falls es eine weitere Matrix $B \in M_{n,n}(K)$ gibt, so daß gilt:

$$A \circ B = B \circ A = 1_n.$$

Man schreibt dann auch: $B = A^{-1}$.

I.3.3 Satz. *Folgende Aussagen über eine Matrix $A \in M_{n,n}(K)$ sind äquivalent:*

1. *A ist invertierbar.*
2. *$F_A : K^n \to K^n$ ist ein Isomorphismus (mit $(F_A)^{-1} = F_{A^{-1}}$).*
3. $\mathrm{Ker}\,(F_A) = \{0\}$.
4. *Das LGS $A\circ \vec{x}=\vec{0}$ ist eindeutig lösbar.*
5. $\mathrm{rg}(A) = n$.

Beweis:
(1) $\Longrightarrow$ (2):
Sei $B := A^{-1}$. Dann ist

$$\mathrm{id}_{K^n} = F_{1} = F_{A\circ B} = F_A \circ F_B \text{ und genauso } \mathrm{id}_{K^n} = F_B \circ F_A.$$

Also ist F_A bijektiv (und damit ein Isomorphismus) mit $(F_A)^{-1} = F_B$.

(2) $\Longrightarrow$ (3):
Ist F_A ein Isomorphismus, so ist natürlich $\mathrm{Ker}\,(F_A) = \{0\}$.

(3) $\Longrightarrow$ (4):
Es ist $\mathrm{Lös}(A, \vec{0}) = \mathrm{Ker}\,(F_A)$, also $= \{0\}$. Damit ist das LGS eindeutig (durch $\vec{x}=\vec{0}$) lösbar.

(4) $\Longrightarrow$ (5):
Nach der Dimensionsformel ist $\mathrm{rg}(A) + \dim_K(\mathrm{Ker}\,(F_A)) = n$. Die eindeutige Lösbarkeit des LGS $A\circ \vec{x}=\vec{0}$ besagt, daß $\mathrm{Ker}\,(F_A) = \{0\}$ sein muß. Also ist $\mathrm{rg}(A) = n$.

(5) $\Longrightarrow$ (1):
Ist $A = (\vec{a}_1, \ldots, \vec{a}_n)$ und $\mathrm{rg}(A) = n$, so wird durch

$$g(\vec{a}_i) := \vec{e}_i \qquad \text{für } i = 1, \ldots, n$$

ein Isomorphismus $g : K^n \to K^n$ definiert, und zu dem gibt es eine Matrix $B \in M_{n,n}(K)$, so daß $g = F_B$ ist. Offensichtlich ist

$$B \circ A\circ \vec{e}_i=\vec{e}_i \text{ und } A \circ B\circ \vec{a}_i=\vec{a}_i, \text{ für } i = 1, \ldots, n,$$

also $B \circ A = A \circ B = 1$. □

Definition.

Eine Matrix $A \in M_{n,n}(K)$ heißt *regulär*, wenn sie eine der äquivalenten Eigenschaften von Satz 2 erfüllt.

Die Menge aller n-reihigen regulären Matrizen wird mit $\mathrm{GL}(n, K)$ bezeichnet („General Linear Group“).

„Regulär“ ist also nur ein anderes Wort für „invertierbar“. Zur Rechtfertigung des Namens der „Allgemeinen Linearen Gruppe“ brauchen wir noch den

I.3.4 Satz.

$$\mathrm{GL}(n,K);=\{A\in M_{n,n}(K)\mid \mathrm{rg}(A)=n\}\ \textit{ist eine Gruppe.}$$

BEWEIS:
1) 1_n liegt in $\mathrm{GL}(n,K)$ und spielt die Rolle des neutralen Elements.

2) Jedes $A\in \mathrm{GL}(n,K)$ besitzt ein Inverses.

3) Seien $A,B\in \mathrm{GL}(n,K)$. Dann gilt:

$$(A\circ B)\circ(B^{-1}\circ A^{-1})=A\circ(B\circ B^{-1})\circ A^{-1}=A\circ A^{-1}=1_n,$$

und genauso

$$(B^{-1}\circ A^{-1})\circ(A\circ B)=1_n.$$

Also ist $A\circ B$ invertierbar, mit $(A\circ B)^{-1}=B^{-1}\circ A^{-1}$. Man beachte die Reihenfolge!!

4) Offensichtlich ist $A\circ(B\circ C)=(A\circ B)\circ C$. □

Beispiele:

1. $\underline{n=1}$:

   $\mathrm{GL}(1,K)=\{a\in K\mid ax=0 \text{ eindeutig lösbar }\}=K^*=K\setminus\{0\}$.

2. $\underline{n=2}$:

   $A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ ist genau dann invertierbar, wenn $\begin{pmatrix} a\\ c\end{pmatrix}$ und $\begin{pmatrix} b\\ d\end{pmatrix}$ linear unabhängig sind, wenn also $\det(A)=ad-bc\neq 0$ ist.

3. Eine Diagonalmatrix

$$D=\begin{pmatrix} d_{11} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & d_{nn}\end{pmatrix}$$

   ist genau dann invertierbar, wenn $d_{11},\dots,d_{nn}\neq 0$ sind.

I.3.5 Satz. *Sei $A\in M_{n,m}(K)$ und $B\in M_{m,l}(K)$.*
Dann ist $\mathrm{rg}(A\circ B)\le \mathrm{rg}(B)$ und $\mathrm{rg}(A\circ B)\le\mathrm{rg}(A)$.

BEWEIS: Sei $f:=F_A:K^m\to K^n$ und $g:=F_B:K^l\to K^m$. Außerdem sei $V:=g(K^l)$. Dann ist $V\subset K^m$ ein Untervektorraum, und $\mathrm{rg}(B)=\dim_K(V)$.

Sei $k:=\mathrm{rg}(B)$ und $\{\mathbf{x}_1,\dots,\mathbf{x}_k\}$ eine Basis von V. Dann ist $\{f(\mathbf{x}_1),\dots,f(\mathbf{x}_k)\}$ ein Erzeugendensystem von $f(V)$, also $\dim_K(f(V))\le k$. Daraus folgt:

$$\begin{aligned}\mathrm{rg}(A\circ B) &= \dim_K(\mathrm{Im}\,(F_{A\circ B}))\\ &= \dim_K(f(V))\\ &\le k = \mathrm{rg}(B).\end{aligned}$$

Da $f(V)\subset f(K^m)=\mathrm{Im}\,(F_A)$ ist, ist auch $\mathrm{rg}(A\circ B)\le\mathrm{rg}(A)$. □

I.3.6 Folgerung. *Die Multiplikation mit einer regulären Matrix ändert den Rang nicht:*

Ist $A \in M_{n,m}(K)$, $P \in \mathrm{GL}(n,K)$ *und* $Q \in \mathrm{GL}(m,K)$, *so ist* $\mathrm{rg}(P \circ A \circ Q) = \mathrm{rg}(A)$.

BEWEIS: Nach dem Satz ist $\mathrm{rg}(P \circ A \circ Q) \le \mathrm{rg}(A)$.
Aber da $A = P^{-1} \circ (P \circ A \circ Q) \circ Q^{-1}$ ist, gilt auch die umgekehrte Ungleichung. □

Wir wollen jetzt ein Verfahren suchen, wie man A^{-1} (für eine reguläre Matrix A) berechnen kann.

Zunächst einige Vorbereitungen:

Sei $A = (\vec{a}_1, \ldots, \vec{a}_m) = \begin{pmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_n \end{pmatrix}$. Für $P \in M_{n,n}(K)$ und $Q \in M_{m,m}(K)$ gilt dann:

$$P \circ A = (P \circ \vec{a}_1, \ldots, P \circ \vec{a}_m) \text{ und } A \circ Q = \begin{pmatrix} \mathbf{c}_1 \circ Q \\ \vdots \\ \mathbf{c}_n \circ Q \end{pmatrix}.$$

Das wenden wir auf spezielle Matrizen P und Q an:

Ist ε eine Folge von elementaren Zeilenumformungen, so ist $\varepsilon : K^n \to K^n$ ein Isomorphismus, der durch eine Matrix P beschrieben werden kann. Damit gilt:

$$\varepsilon(A) = (\varepsilon(\vec{a}_1), \ldots, \varepsilon(\vec{a}_m)) = (P \circ \vec{a}_1, \ldots, P \circ \vec{a}_m) = P \circ A.$$

Ist σ eine Folge von Spaltenvertauschungen, so definieren wir $f_\sigma : K^m \to K^m$ durch

$$f_\sigma(x_1, \ldots, x_m) := (x_{\sigma(1)}, \ldots, x_{\sigma(m)}).$$

Diese Permutationsabbildung ist ein Isomorphismus, den wir auf die **Zeilenvektoren** $\mathbf{c}_i$ von A wirken lassen.
Ist $Q := E_\sigma = (\vec{e}_{\sigma(1)}, \ldots, \vec{e}_{\sigma(m)}) \in M_{m,m}(K)$ die zugehörige Matrix, so gilt:

$$\begin{aligned} A \circ Q &= \begin{pmatrix} \mathbf{c}_1 \circ Q \\ \vdots \\ \mathbf{c}_n \circ Q \end{pmatrix} = \begin{pmatrix} a_{1,\sigma(1)} & \cdots & a_{1,\sigma(m)} \\ \vdots & & \vdots \\ a_{n,\sigma(1)} & \cdots & a_{n,\sigma(m)} \end{pmatrix} \\ &= (\vec{a}_{\sigma(1)}, \ldots, \vec{a}_{\sigma(m)}) = A_\sigma. \end{aligned}$$

Da man A durch elementare Zeilenumformungen und Spaltenvertauschungen auf Gauß-Form bringen kann, bedeutet das:

> *Ist* $r = \mathrm{rg}(A)$, *so gibt es reguläre Matrizen* P *und* Q, *so daß* $P \circ A \circ Q$ *eine Gaußmatrix vom Rang* r *ist.*

Nun betrachten wir den Fall, daß $A \in \mathrm{GL}(n, K)$ ist. Allein mit Hilfe von Zeilenumformungen kann man erreichen, daß A r-speziell für ein $r \geq 0$ wird:

$$\varepsilon(A) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \\ \hline & & & \\ & 0 & & C \\ & & & \end{array}\right).$$

Wäre $r < n$ und die 1. Spalte von $C = 0$, so wäre die (r+1)-te Spalte von $\varepsilon(A)$ linear abhängig von den ersten r Spalten. Andererseits ist $\mathrm{rg}(\varepsilon(A)) = \mathrm{rg}(A) = n$, es sind also alle Spalten linear unabhängig.

Also gibt es in der ersten Spalte von C ein Element $\neq 0$, und mit Hilfe von Zeilenvertauschungen kann man es an die Stelle $(r+1, r+1)$ bringen. Das bedeutet, daß man keine Spaltenvertauschung braucht, um A auch $(r+1)$-speziell zu machen.

Dieses Verfahren läßt sich fortsetzen, solange $r < n$ ist. Das bedeutet: Es gibt eine Matrix $P \in \mathrm{GL}(n, K)$, die einer Folge von Zeilenumformungen entspricht, so daß $P \circ A$ eine obere Dreiecksmatrix vom Rang n ist. Durch weitere Zeilenumformungen kann man daraus schließlich sogar die Einheitsmatrix machen.

Ist aber $P \circ A = 1_n$, so ist $P = A^{-1}$. Um nun bei der Durchführung des Gaußverfahrens gleichzeitig auch die Matrix P zu erhalten, erweitern wir A zur Matrix $(A, 1)$. Dann gilt:

$$P \circ (A, 1) = (P \circ A, P \circ 1) = (1, A^{-1}).$$

I.3.7 Satz. *Sei $A \in \mathrm{GL}(n, K)$.*

1. *Es gibt eine Folge ε von Zeilenumformungen mit $\varepsilon(A) = 1$.*
2. *Ist $\varepsilon(A, 1) = (1, A^*)$, so ist $A^* = A^{-1}$.*

BEWEIS: Der erste Teil wurde gerade gezeigt.

Wenn ε durch die Matrix P verwirklicht wird, hat man die Gleichung

$$P \circ (A, 1) = (1, A^*),$$

also $P \circ A = 1$ und $P = P \circ 1 = A^*$. Damit ist $A^* = A^{-1}$. □

Beispiel:

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

Dann ist $\det(A) = \cos^2\varphi + \sin^2\varphi = 1$, also A regulär. Wir berechnen die inverse Matrix durch Umformen von $(A, \mathbb{1})$:

$\cos\varphi$	$-\sin\varphi$	1	0	$\times\cos\varphi$
$\sin\varphi$	$\cos\varphi$	0	1	$\times\sin\varphi$
$\cos^2\varphi$	$-\sin\varphi\cos\varphi$	$\cos\varphi$	0	$+$ 2. Zeile
$\sin^2\varphi$	$\sin\varphi\cos\varphi$	0	$\sin\varphi$	
1	0	$\cos\varphi$	$\sin\varphi$	
$\sin^2\varphi$	$\sin\varphi\cos\varphi$	0	$\sin\varphi$	$-\sin^2\varphi$(1. Zeile)
1	0	$\cos\varphi$	$\sin\varphi$	
0	$\sin\varphi\cos\varphi$	$-\sin^2\varphi\cos\varphi$	$\sin\varphi-\sin^3\varphi$	
1	0	$\cos\varphi$	$\sin\varphi$	
0	1	$-\sin\varphi$	$\cos\varphi$	

Also ist

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}^{-1} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Die Probe zeigt, daß es stimmt!

Das Invertieren von Matrizen hat viele Anwendungen:

Sei $A \in M_{n,n}(K)$ regulär. Dann ist jedes LGS $A\circ \vec{x}=\vec{b}$ eindeutig lösbar. Das wußten wir schon, aber jetzt kann man die Lösung auch sofort hinschreiben:

$$\vec{x} = A^{-1}\circ \vec{b}\ .$$

Eine andere Anwendung ist die Berechnung von Koordinatenwechseln. Das ist ein Thema, das erfahrungsgemäß zu Anfang einige Schwierigkeiten bereitet.

Sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von K^n,

$$B := (\vec{b}_1, \ldots, \vec{b}_n) \in \mathrm{GL}(n, K) \quad \text{die aus den Spalten } \vec{b}_i \text{ gebildete Matrix.}$$

Dann gibt es für jeden Vektor $\vec{x} \in K^n$ eine eindeutig bestimmte Darstellung

$$\vec{x} = \sum_{i=1}^{n} x_i^{(B)}\ \vec{b}_i\ .$$

Wir setzen jetzt

$$[\vec{x}]_{\mathcal{B}} := \begin{pmatrix} x_1^{(B)} \\ \vdots \\ x_n^{(B)} \end{pmatrix} \in K^n$$

und nennen das den *Koordinatenvektor* von $\vec{x}$ zur Basis $\mathcal{B}$. Dann ist

$$\begin{aligned} B\circ[\vec{x}]_{\mathcal{B}} &= (\vec{b}_1, \ldots, \vec{b}_n)\circ \begin{pmatrix} x_1^{(B)} \\ \vdots \\ x_n^{(B)} \end{pmatrix} \\ &= \sum_{i=1}^{n} x_i^{(B)}\ \vec{b}_i = \vec{x}, \end{aligned}$$

und weil B invertierbar ist, kann man auch schreiben:

$$\boxed{[\vec{x}]_{\mathcal{B}} = B^{-1} \circ \vec{x}\ .}$$

Beispiel :

Sei $B := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ mit den Spalten $\vec{b}_1, \vec{b}_2, \vec{b}_3$.

Dann ist $\mathcal{B} := \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ eine Basis des $\mathbb{R}^3$.

Um Koordinaten bezüglich der Basis $\mathcal{B}$ zu ermitteln, müssen wir zunächst B^{-1} berechnen. Das geschieht nach dem Standardverfahren:

1	1	0	1	0	0
1	0	1	0	1	0
0	1	1	0	0	1
1	1	0	1	0	0
0	-1	1	-1	1	0
0	1	1	0	0	1
1	0	1	0	1	0
0	-1	1	-1	1	0
0	0	2	-1	1	1
1	0	1	0	1	0
0	1	-1	1	-1	0
0	0	1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
1	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$
0	1	0	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$
0	0	1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Also ist $B^{-1} = \dfrac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

Der Vektor $\vec{x} := \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix}$ sei in kartesischen Koordinaten (also bezüglich der Standardbasis) gegeben. Bezüglich der Basis $\mathcal{B}$ besitzt er die Darstellung

$$[\vec{x}]_{\mathcal{B}} = B^{-1} \circ \vec{x} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = \begin{pmatrix} -6 \\ 18 \\ 12 \end{pmatrix}.$$

Tatsächlich ist

$$-6\,\vec{b}_1 + 18\,\vec{b}_2 + 12\,\vec{b}_3 = \begin{pmatrix} -6 \\ -6 \\ 0 \end{pmatrix} + \begin{pmatrix} 18 \\ 0 \\ 18 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \\ 12 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = \vec{x}\,.$$

Nun sei $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_n\}$ eine weitere Basis, B' die zugehörige Matrix. Ein Vektor $\vec{x} \in K^n$ kann dann auch in der Form

$$\vec{x} = B' \circ [\vec{x}]_{\mathcal{B}'}$$

geschrieben werden, und dann gilt:

$$\begin{aligned} B \circ [\vec{x}]_{\mathcal{B}} &= B' \circ [\vec{x}]_{\mathcal{B}'}, \\ \text{also} \quad [\vec{x}]_{\mathcal{B}} &= (B^{-1} \circ B') \circ [\vec{x}]_{\mathcal{B}'}, \quad \text{für alle } \vec{x} \in K^n. \end{aligned}$$

Die Matrix $A := B^{-1} \circ B'$ heißt *Übergangsmatrix.* Wir benötigen sie, wenn wir zwei verschiedene Koordinatenvektoren von $\vec{x}$ ineinander umrechnen wollen. Um ihre Bedeutung noch besser zu verstehen, schreiben wir sie in der Form

$$A = (\, \alpha_{ik} \mid i, k = 1, \ldots, n\,).$$

Da $B \circ A = B'$ ist, folgt:

$$\sum_{i=1}^{n} \alpha_{ik}\, \vec{b}_i = (\vec{b}_1, \ldots, \vec{b}_n) \circ \begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{nk} \end{pmatrix} = \vec{b}'_k\,.$$

A ist also die Koeffizientenmatrix, die man braucht, um die Basisvektoren von $\mathcal{B}'$ als Linearkombinationen der Basisvektoren von $\mathcal{B}$ zu schreiben.

Beispiel:

Wir führen das Beispiel von oben fort:

Es sei $B' := \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, und $\mathcal{B}'$ die aus den Spalten von B' gebildete Basis $\{\vec{b}'_1, \vec{b}'_2, \vec{b}'_3\}$ des $\mathbb{R}^3$. Die Übergangsmatrix zwischen den Basen $\mathcal{B}$ und $\mathcal{B}'$ ist die Matrix $A = B^{-1} \circ B' =$

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

Benutzt man etwa die zweite Spalte von A, so erhält man z.B.:

$$\frac{1}{2}(0\cdot \vec{b}_1 - 2\cdot \vec{b}_2 + 2\cdot \vec{b}_3) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \vec{b}'_2 .$$

Um den Koordinatenvektor von $\vec{x}$ bezüglich $\mathcal{B}'$ zu ermitteln, berechnen wir $(B')^{-1}$:

1	−1	−1	1	0	0
1	1	0	0	1	0
1	0	1	0	0	1
3	0	0	1	1	1
1	1	0	0	1	0
1	0	1	0	0	1
1	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
0	1	0	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$
0	0	1	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$

Demnach ist $(B')^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$, und

$$[\vec{x}]_{\mathcal{B}'} = (B')^{-1} \circ \vec{x} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \circ \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = \begin{pmatrix} 16 \\ -10 \\ 14 \end{pmatrix} .$$

Der Übergang von $[\vec{x}]_{\mathcal{B}'}$ nach $[\vec{x}]_{\mathcal{B}}$ wird durch Multiplikation mit der Übergangsmatrix A von links vollzogen:

$$A \circ [\vec{x}]_{\mathcal{B}'} = B^{-1} \circ B' \circ (B')^{-1} \circ \vec{x} = B^{-1} \circ \vec{x} = [\vec{x}]_{\mathcal{B}}.$$

Tatsächlich ist in unserem Beispiel

$$\frac{1}{2}\begin{pmatrix} 1 & 0 & -2 \\ 1 & -2 & 0 \\ 1 & 2 & 2 \end{pmatrix} \circ \begin{pmatrix} 16 \\ -10 \\ 14 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -12 \\ 36 \\ 24 \end{pmatrix} = \begin{pmatrix} -6 \\ 18 \\ 12 \end{pmatrix} .$$

Wir fassen die *Basiswechsel–Gleichungen* noch einmal zusammen:

$$\boxed{\begin{gathered} B' = B \circ A \\ A \circ [\vec{x}]_{\mathcal{B}'} = [\vec{x}]_{\mathcal{B}} \end{gathered}}$$

Man beachte die unterschiedliche Stellung von A in den beiden Formeln!

Sei jetzt V ein beliebiger n-dimensionaler Vektorraum über K, und $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ eine Basis von V. Der durch

$$\Phi_{\mathcal{A}}(\vec{e}_i) := \mathbf{a}_i \quad \text{für } i = 1, \ldots, n$$

gegebene Isomorphismus $\Phi_{\mathcal{A}} : K^n \to V$ heißt *Koordinatensystem* oder *Basisisomorphismus.*

Ist $\mathbf{x} \in V$, so besitzt $\mathbf{x}$ eine Darstellung $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{a}_i$, und es gilt:

$$(\Phi_{\mathcal{A}})^{-1}(\mathbf{x}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Auf diese Weise können wir die Elemente von V mit gewöhnlichen Spaltenvektoren aus K^n identifizieren.

Was passiert nun, wenn wir in V mit einer anderen Basis arbeiten wollen?

Ist $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine solche Basis, so gibt es eine Darstellung $\mathbf{x} = \sum_{i=1}^{n} x_i' \mathbf{b}_i$. Wir wollen die neuen Komponenten x_i' aus den alten x_k ausrechnen, in analoger Weise zu unserem Vorgehen beim Basiswechsel im K^n.

Dazu sei $\vec{b}'_i := (\Phi_{\mathcal{A}})^{-1}(\mathbf{b}_i)$, für $i = 1, \ldots, n$, und $B' := (\vec{b}'_1, \ldots, \vec{b}'_n)$. Offensichtlich gilt:

$$(\Phi_{\mathcal{A}})^{-1}(\mathbf{x}) = \sum_{i=1}^{n} x_i' (\Phi_{\mathcal{A}})^{-1}(\mathbf{b}_i) = \sum_{i=1}^{n} x_i' \, \vec{b}'_i \,.$$

Wir setzen

$$\begin{aligned} [\mathbf{x}]_{\mathcal{A}} &:= (\Phi_{\mathcal{A}})^{-1}(\mathbf{x}) \\ \text{und} \quad [\mathbf{x}]_{\mathcal{B}} &:= (B')^{-1} \circ [\mathbf{x}]_{\mathcal{A}} = [(\Phi_{\mathcal{A}})^{-1}(\mathbf{x})]_{\mathcal{B}'}. \end{aligned}$$

Beispiel:

Es sei $V := \{\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$. Das ist ein 2-dimensionaler Untervektorraum des $\mathbb{R}^3$, und offensichtlich bilden

$$\mathbf{a}_1 := (1, 0, -1) \quad \text{und} \quad \mathbf{a}_2 := (0, 1, -1)$$

eine Basis $\mathcal{A}$ von V. Wir benutzen diese Basis als Standardbasis und erhalten z.B. für $\mathbf{x} := (3, -2, -1) = 3\mathbf{a}_1 - 2\mathbf{a}_2$ die Darstellung

$$[\mathbf{x}]_{\mathcal{A}} = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \in \mathbb{R}^2.$$

Die Vektoren

$$\mathbf{b}_1 := (1, 1, -2) \quad \text{und} \quad \mathbf{b}_2 := (1, -1, 0)$$

bilden ebenfalls eine Basis $\mathcal{B}$ von V, wie man leicht nachrechnet. Nun ist

$$\begin{aligned} & (\Phi_{\mathcal{A}})^{-1}(\mathbf{b}_1) &=& (\Phi_{\mathcal{A}})^{-1}(\mathbf{a}_1+\mathbf{a}_2) = \begin{pmatrix}1\\1\end{pmatrix} \\ \text{und} \quad & (\Phi_{\mathcal{A}})^{-1}(\mathbf{b}_2) &=& (\Phi_{\mathcal{A}})^{-1}(\mathbf{a}_1-\mathbf{a}_2) = \begin{pmatrix}1\\-1\end{pmatrix}, \end{aligned}$$

also $B' = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ und $(B')^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Damit gilt:

$$\begin{aligned} [\mathbf{x}]_{\mathcal{B}} &= (B')^{-1} \circ [\mathbf{x}]_{\mathcal{A}} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \circ \begin{pmatrix} 3 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} \\ \frac{5}{2} \end{pmatrix}. \end{aligned}$$

Tatsächlich ist

$$\frac{1}{2}\mathbf{b}_1 + \frac{5}{2}\mathbf{b}_2 = (\frac{1}{2}, \frac{1}{2}, -1) + (\frac{5}{2}, -\frac{5}{2}, 0) = (3, -2, -1) = \mathbf{x}.$$

Sei $F : K^m \to K^n$ eine lineare Abbildung, mit $F(\vec{x}) = A\circ \vec{x}$. Wir wollen sehen, wie F mit Hilfe anderer Basen beschrieben werden kann (bisher haben wir immer nur die Standardbasen benutzt).

Sind $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ und $\mathcal{C} = (\vec{c}_1, \ldots, \vec{c}_m)$ Basen von K^n bzw. K^m und $B \in \mathrm{GL}(n, K)$ bzw. $C \in \mathrm{GL}(m, K)$ die zugehörigen Matrizen, so gilt für $\vec{x} \in K^m$ und $\vec{y} := F(\vec{x}) \in K^n$:

$$[\vec{x}]_{\mathcal{C}} = C^{-1}\circ \vec{x} \qquad \text{und} \qquad [\vec{y}]_{\mathcal{B}} = B^{-1}\circ \vec{y} = B^{-1} \circ A\circ \vec{x},$$

also

$$[\vec{y}]_{\mathcal{B}} = (B^{-1} \circ A \circ C) \circ [\vec{x}]_{\mathcal{C}}.$$

Das bedeutet: Wenn F bezüglich der Standardbasen von K^m und K^n durch die Matrix $A \in M_{n,m}(K)$ beschrieben wird, dann wird F bezüglich der Basen $\mathcal{C}$ und $\mathcal{B}$ durch die Matrix $B^{-1} \circ A \circ C$ beschrieben.

Ist $A' := B^{-1} \circ A \circ C = \left(a'_{\nu\mu} \mid \begin{array}{rcl} \nu &=& 1, \ldots, n \\ \mu &=& 1, \ldots, m \end{array} \right)$, so gilt:

$$\begin{aligned} F(\vec{c}_\mu) &= A \circ C\circ \vec{e}_\mu \\ &= B \circ (B^{-1} \circ A \circ C)\circ \vec{e}_\mu \\ &= B \circ (a'_{1\mu}, \ldots, a'_{n\mu})^\top \\ &= \sum_{\nu=1}^{n} a'_{\nu\mu} \vec{b}_\nu \, . \end{aligned}$$

Definition.

> Zwei Matrizen $A, A' \in M_{n,m}(K)$ heißen *äquivalent*, falls es reguläre Matrizen B und C gibt, so daß $A' = B^{-1} \circ A \circ C$ ist.

Äquivalente Matrizen beschreiben also die gleiche lineare Abbildung bezüglich verschiedener Basen. Unter dem *Normalformenproblem* versteht man die Aufgabe, die Basen B und C so zu wählen, daß A' eine möglichst einfache Gestalt annimmt. Wir wissen schon, daß man durch elementare Zeilenumformungen und Spaltenvertauschungen erreichen kann, daß A' eine Gaußmatrix wird. Läßt man auch noch beliebige Spaltenumformungen zu, so kann man zeigen:

Eine Matrix $A \in M_{n,m}(K)$ vom Rang r ist äquivalent zu der Matrix

$$A' = \left(\begin{array}{ccc|c} 1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & 0 \\ 0 & \cdots & 1 & \\ \hline & & & \\ & 0 & & 0 \\ & & & \end{array}\right),$$

mit r Einsen auf der Hauptdiagonale.

Bei Abbildungen von K^n auf K^n kann man versuchen, mit einer einzigen Basis auszukommen. Zwei Matrizen $A, A' \in M_{n,n}(K)$ heißen *ähnlich*, falls es eine reguläre Matrix B mit $A' = B^{-1} \circ A \circ B$ gibt. In diesem Fall ist die Lösung des Normalformenproblems sehr schwierig. Das Stichwort heißt hier „Jordansche Normalform“.

§4 Determinanten

In diesem Paragraphen ist immer $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Für 2-reihige und 3-reihige quadratische Matrizen haben wir jeweils schon Determinanten eingeführt:

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

$$\begin{aligned} \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} &= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1 \\ &= \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) a_{\sigma(1)} b_{\sigma(2)} c_{\sigma(3)}. \end{aligned}$$

Wir wollen nun für beliebiges n eine Funktion

$$\det : M_{n,n}(K) = K^n \times \ldots \times K^n \to K$$

suchen, die folgende Eigenschaften besitzt:

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

$$\begin{aligned} \det(\vec{a}_1, \ldots, \vec{a}_i + \vec{a}'_i, \ldots, \vec{a}_n) &= \det(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) \\ &\quad + \det(\vec{a}_1, \ldots, \vec{a}'_i, \ldots, \vec{a}_n), \\ \det(\vec{a}_1, \ldots, \alpha\, \vec{a}_i, \ldots, \vec{a}_n) &= \alpha \cdot \det(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n). \end{aligned}$$

D2) Sind zwei Spalten gleich (d.h. $\vec{a}_i = \vec{a}_j$ für Indizes $i \neq j$), so ist $\det(\vec{a}_1, \ldots, \vec{a}_n) = 0$.

D3) $\det(1_n) = 1$.

In Kapitel I wurde schon bewiesen, daß die Determinanten im Falle $n = 2$ und $n = 3$ die Eigenschaften D1, D2 und D3 besitzen.

I.4.1 Hilfssatz. *Sei V ein K-Vektorraum und $f : V \times V \to K$ eine Funktion, die in beiden Argumenten linear ist. Dann sind folgende Aussagen äquivalent:*

1. $f(\mathbf{x}, \mathbf{x}) = 0$ *für alle* $\mathbf{x} \in V$.
2. $f(\mathbf{y}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{y})$ *für* $\mathbf{x}, \mathbf{y} \in V$.

Beweis:
1) Ist $f(\mathbf{y}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{y})$, so ist $2 \cdot f(\mathbf{x}, \mathbf{x}) = 0$, also auch $f(\mathbf{x}, \mathbf{x}) = 0$.

2) Sei $f(\mathbf{x}, \mathbf{x}) = 0$ für alle $\mathbf{x}$. Dann gilt:

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

□

Man nennt solche Abbildungen *schiefsymmetrisch* oder *alternierend*. Die Eigenschaft, in jedem einzelnen Argument linear zu sein, nennt man *multilinear*. ACHTUNG! Wenn f zwei oder mehr Argumente hat und multilinear ist, dann ist f i.a. nicht linear! (Man überprüfe das an Hand der Abbildung $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ mit $(x, y) \mapsto x \cdot y$.)

Die Determinantenfunktion soll also eine auf n-Tupeln von Vektoren definierte multilineare alternierende Funktion werden, die außerdem durch Bedingung (D3) normiert ist.

Vertauscht man in

$$\det(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n)$$

die Argumente $\vec{a}_i$ und $\vec{a}_j$, so wechselt das Vorzeichen. Diesen Prozeß kann man iterieren. Daher muß schließlich für jede Permutation $\sigma \in S_n$ gelten:

$$\det(\vec{a}_{\sigma(1)}, \ldots, \vec{a}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot \det(\vec{a}_1, \ldots, \vec{a}_n).$$

Insbesondere folgt daraus für die Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n$:

$$\det(\vec{e}_{\sigma(1)}, \ldots, \vec{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma).$$

I.4.2 Satz von der Existenz und Eindeutigkeit der Determinante.

Es gibt genau eine Funktion $\det : K^n \times \ldots \times K^n \to K$ *mit den Eigenschaften (D1), (D2) und (D3).*

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

Die Spaltenvektoren $\vec{a}_1, \ldots, \vec{a}_n \in K^n$ fassen wir auf als Spalten einer Matrix $A = (a_{ij}) \in M_{n,n}(K)$. Dann ist

$$\vec{a}_j = \sum_{i=1}^{n} a_{ij}\, \vec{e}_i, \quad \text{für } j = 1, \ldots, n.$$

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

$$\vec{a}_1 = \sum_{i_1=1}^{n} a_{i_1 1}\, \vec{e}_{i_1}, \quad \ldots \quad, \vec{a}_n = \sum_{i_n=1}^{n} a_{i_n n}\, \vec{e}_{i_n}\ .$$

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

$$\begin{aligned} \det(\vec{a}_1, \ldots, \vec{a}_n) &= \det\Big(\sum_{i_1=1}^{n} a_{i_1 1}\, \vec{e}_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n n}\, \vec{e}_{i_n}\Big) \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1 1} \cdot \ldots \cdot a_{i_n n} \det(\vec{e}_{i_1}, \ldots, \vec{e}_{i_n}). \end{aligned}$$

Nun ist

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

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

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

Damit erhalten wir:

$$\boxed{\det(\vec{a}_1, \ldots, \vec{a}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n} \,.}$$

Die Determinante muß zwangläufig so aussehen! Das liefert uns die Eindeutigkeit, und auch die Existenz, wenn wir zeigen können, daß der gefundene Ausdruck (den wir mit $D(\vec{a}_1, \ldots, \vec{a}_n)$ bezeichnen wollen) die Bedingungen (D1), (D2) und (D3) erfüllt.

Das ist nicht ganz einfach! Wir beginnen mit der Multilinearität (D1) und beschränken uns dabei aus schreibtechnischen Gründen auf das erste Argument:

$$\begin{aligned}
&D(\vec{a}_1 + \vec{a}'_1, \vec{a}_2, \ldots, \vec{a}_n) = \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(a_{\sigma(1),1} + a'_{\sigma(1),1}) \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n} + a'_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n}) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a'_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= D(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) + D(\vec{a}'_1, \vec{a}_2, \ldots, \vec{a}_n).
\end{aligned}$$

Genauso geht es bei den anderen Argumenten. Weiter ist

$$\begin{aligned}
D(\alpha \cdot \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(\alpha \cdot a_{\sigma(1),1}) \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= \alpha \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= \alpha \cdot D(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n).
\end{aligned}$$

Nun nehmen wir an, daß $\vec{a}_1 = \vec{a}_2$ ist. $\tau \in S_n$ sei die Permutation, die lediglich 1 mit 2 vertauscht. Dann ist $\operatorname{sgn}(\sigma \circ \tau) = -\operatorname{sgn}(\sigma)$, für jedes $\sigma \in S_n$.

Wir definieren

$$\begin{aligned}
S'_n &:= \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\} \\
\text{und} \quad S''_n &:= \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = -1\}.
\end{aligned}$$

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

$$\begin{aligned}
&D(\vec{a}_1, \vec{a}_1, \ldots, \vec{a}_n) = \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= \sum_{\sigma \in S'_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&\quad + \sum_{\sigma \in S'_n} \operatorname{sgn}(\sigma \circ \tau) a_{\sigma\circ\tau(1),1} \cdot a_{\sigma\circ\tau(2),1} \cdot \ldots \cdot a_{\sigma\circ\tau(n),n} \\
&= \sum_{\sigma \in S'_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),1} \cdot \ldots \cdot a_{\sigma(n),n} - \sum_{\sigma \in S'_n} \operatorname{sgn}(\sigma) a_{\sigma(2),1} \cdot a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n} \\
&= 0.
\end{aligned}$$

Damit ist die Funktion D auch alternierend.

Wir kommen nun zu (D3): Das **Kronecker–Symbol** δ_{ij} wird definiert durch

$$\boxed{\delta_{ij} := \begin{cases} 0 & \text{falls } i \neq j \\ 1 & \text{falls } i = j \end{cases}}$$

Dann ist δ_{ij} die i-te Komponente des Einheitsvektors $\vec{e}_j$, und es gilt:

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

Daraus folgt:

$$\begin{aligned} D(\vec{e}_1, \ldots, \vec{e}_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \delta_{\sigma(1),1} \cdot \ldots \cdot \delta_{\sigma(n),n} \\ &= \operatorname{sgn}(\mathrm{id}) = 1. \end{aligned}$$

Damit ist tatsächlich alles bewiesen! □

Wir haben jetzt eine Formel für die Determinante, und für $n = 2$ und $n = 3$ kommt tatsächlich das heraus, was wir schon kennen. Man kann sich aber auch schnell davon überzeugen, daß so etwas wie die Regel von Sarrus für $n \geq 4$ nicht mehr zur Verfügung steht. Und die Original–Formel ist leider auch denkbar ungeeignet, Determinanten zu berechnen. Daher müssen wir noch etwas mehr Theorie betreiben.

Bemerkung : Man schreibt auch

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

Wir beginnen mit der Determinante der transponierten Matrix. Zur Erinnerung: Ist $A = (a_{ij}) \in M_{n,m}(K)$, so steht in der transponierten Matrix $A^\top \in M_{m,n}(K)$ an der Stelle (j, i) das Element a_{ij}.

Beispiel :

$$\begin{pmatrix} 3 & -1 & 7 \\ 5 & 0 & 2 \end{pmatrix}^\top = \begin{pmatrix} 3 & 5 \\ -1 & 0 \\ 7 & 2 \end{pmatrix}.$$

I.4.3 Satz. *Ist $A \in M_{n,m}(K)$ und $B \in M_{m,l}(K)$, so ist*

$$(A \circ B)^\top = B^\top \circ A^\top.$$

Beweis:

$$\begin{aligned}((A\circ B)^\top)_{ki} &= (A\circ B)_{ik}\\ &= \sum_{j=1}^m A_{ij}B_{jk}\\ &= \sum_{j=1}^m (A^\top)_{ji}(B^\top)_{kj}\\ &= (B^\top\circ A^\top)_{ki}.\end{aligned}$$

□

I.4.4 Satz. *Sei $A \in M_{n,n}(K)$. Dann gilt:*

$$\det(A^\top) = \det(A).$$

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

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

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

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

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

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

□

Damit ergeben sich folgende Regeln:

I.4.5 Satz. *Sei $A \in M_{n,n}(K)$.*

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

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

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

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

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

5. *Ist* $f : M_{n,n}(K) \to K$ *eine Funktion, die multilinear und alternierend in den Zeilen von* A *ist, mit* $f(\mathbb{1}_n) = 1$*, so ist* $f = \det$.

Beweis: Das meiste ist schon bekannt oder ergibt sich durch den Übergang $A \to A^\top$.

Ist $\text{rg}(A) < n$, so sind die Spalten $\vec{a}_1, \ldots, \vec{a}_n$ von A linear abhängig. Sei etwa

$$\vec{a}_1 = \sum_{j=2}^{n} \lambda_j \vec{a}_j .$$

Dann ist

$$\begin{aligned} \det(\vec{a}_1, \ldots, \vec{a}_n) &= \det(\sum_{j=2}^{n} \lambda_j \vec{a}_j, \vec{a}_2, \ldots, \vec{a}_n) \\ &= \sum_{j=2}^{n} \lambda_j \det(\vec{a}_j, \vec{a}_2, \ldots, \vec{a}_n) \\ &= 0, \end{aligned}$$

denn in der letzten Determinante kommt $\vec{a}_j$ zweimal als Argument vor. □

Durch elementare Umformungen kann man jede Matrix $A \in M_{n,n}(K)$ auf Dreiecksgestalt bringen. Daher ist der folgende Satz sehr nützlich:

I.4.6 Satz.

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

Beweis: Die Matrix sei mit A bezeichnet. Wir unterscheiden mehrere Fälle:

<u>1. Fall:</u> $a_{11} = 0$.

Dann hängt die erste Spalte von den anderen linear ab, und es ist $\det(A) = 0$.

<u>2. Fall:</u> $a_{11}, \ldots, a_{ii} \neq 0$ für ein i mit $1 \leq i < n$, aber $a_{i+1,i+1} = 0$.

Dann ist die $(i+1)$–te Spalte von den ersten i Spalten linear abhängig, und auch in diesem Fall ist $\det(A) = 0$.

<u>3. Fall:</u> $a_{11}, \ldots, a_{nn} \neq 0$.

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

$$\det(A) = a_{11} \cdot \ldots \cdot a_{nn}.$$

□

Der nächste Satz hat sowohl praktische als auch theoretische Bedeutung:

I.4.7 Determinanten–Produktsatz.
Es seien $A, B \in M_{n,n}(K)$. Dann gilt:

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

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

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

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

Nun sei $\delta : M_{n,n}(K) \to K$ definiert durch

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

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

Wir haben implizit mitbewiesen:

I.4.8 Satz. *$A \in M_{n,n}(K)$ ist genau dann regulär, wenn $\det(A) \neq 0$ ist.*

$$\textit{In diesem Falle ist} \quad \det(A^{-1}) = \frac{1}{\det(A)}.$$

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

Definition.

Sei $A = (\vec{a}_1, \ldots, \vec{a}_n)$. Dann nennt man

$$A_{ij} := \det(\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{e}_i, \vec{a}_{j+1}, \ldots, \vec{a}_n)$$

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

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

I.4.9 Satz. *Für alle i, j gilt:*

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

Beweis: Es ist

$$
\begin{array}{rcl}
A_{ij} & = & \left|\begin{array}{cccccc} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ & & & 1 & & & \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{array}\right| \leftarrow i \\
& = & (-1)^{(i-1)+(j-1)} \cdot \left|\begin{array}{ccccccc} 1 & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & & & & \vdots \\ 0 & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{array}\right| \\
& = & (-1)^{i+j} \cdot \det \left(\begin{array}{c|c} 1 & * \\ \hline 0 & \\ \vdots & S_{ij}(A) \\ 0 & \end{array}\right) \\
& = & (-1)^{i+j} \cdot \det S_{ij}(A). \qquad \text{(vgl. den folg. Hilfssatz!)}
\end{array}
$$

□

I.4.10 Hilfssatz. *Sind $A \in M_{r,r}(K)$, $B \in M_{r,n-r}(K)$ und $C \in M_{n-r,n-r}(K)$, so ist*

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

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

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

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

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

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

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

□

I.4.11 Laplace'scher Entwicklungssatz.
Sei $A = (a_{ij}) \in M_{n,n}(K)$. Dann gilt:

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

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

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

Sind $\vec{a}_1, \ldots, \vec{a}_n$ die Spalten von A, so gilt:

$$\begin{aligned}\det(\vec{a}_1, \ldots, \vec{a}_n) &= \det(\vec{a}_1, \ldots, \vec{a}_{j-1}, \sum_{i=1}^{n} a_{ij}\ \vec{e}_i, \vec{a}_{j+1}, \ldots, \vec{a}_n) \\ &= \sum_{i=1}^{n} a_{ij} \det(\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{e}_i, \vec{a}_{j+1}, \ldots, \vec{a}_n) \\ &= \sum_{i=1}^{n} a_{ij} A_{ij} \\ &= \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det S_{ij}(A).\end{aligned}$$

□

Die Vorzeichen sind wie bei einem Schachbrett verteilt:

$$\begin{array}{ccccc} + & - & + & - & \cdots \\ - & + & - & + & \\ + & - & + & - & \\ \vdots & & & & \end{array}$$

Beispiel:

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

Das war eine Entwicklung nach der ersten Zeile. Im allgemeinen wird man eine Zeile oder Spalte suchen, in der möglichst viele Nullen zu finden sind.

Eine Anwendung der Determinantentheorie ergibt sich für die LGS:

I.4.12 Cramersche Regel.
Sei $A = (\vec{a}_1, \ldots, \vec{a}_n) \in M_{n,n}(K)$. Dann gilt:

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

2. *Sei $\det(A) \neq 0$ und $\vec{x}= (x_1, \ldots, x_n)^\top$ der eindeutig bestimmte Lösungsvektor des LGS $A\circ \vec{x}=\vec{b}$. Dann ist*

$$x_i = \frac{1}{\det(A)} \cdot \det(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \ldots, \vec{a}_n).$$

Beweis:

1) Es gilt:

$$\begin{aligned} A\circ \vec{x}=\vec{b} \text{ ist eindeutig lösbar } &\iff \\ \iff\ & A\circ \vec{x}=\vec{0} \text{ eindeutig lösbar, und } \operatorname{rg}(A,\vec{b}) = \operatorname{rg}(A) \\ \iff\ & \operatorname{rg}(A) = \operatorname{rg}(A,\vec{b}) = n \\ \iff\ & \operatorname{rg}(A) = n \\ \iff\ & \det(A) \neq 0. \end{aligned}$$

2) Ist $\vec{x}$ der Lösungsvektor des LGS $A\circ \vec{x}=\vec{b}$, so ist

$$\vec{b}= \sum_{k=1}^{n} x_k \vec{a}_k .$$

Daraus folgt:

$$\begin{aligned} \det(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \ldots, \vec{a}_n) &= \\ &= \det(\vec{a}_1, \ldots, \vec{a}_{i-1}, \sum_{k=1}^{n} x_k \vec{a}_k, \vec{a}_{i+1}, \ldots, \vec{a}_n) \\ &= \sum_{k=1}^{n} x_k \det(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{a}_k, \vec{a}_{i+1}, \ldots, \vec{a}_n) \\ &= x_i \cdot \det(\vec{a}_1, \ldots, \vec{a}_n). \end{aligned}$$

□

Beispiel:

Sei $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$ und $\vec{b}= \begin{pmatrix} -13 \\ -7 \end{pmatrix}$. Dann gilt:

$$\det(A) = 2 + 15 = 17,$$

$$\det\begin{pmatrix} -13 & -3 \\ -7 & 1 \end{pmatrix} = -13 - 21 = -34 \text{ und } \det\begin{pmatrix} 2 & -13 \\ 5 & -7 \end{pmatrix} = -14 + 65 = 51.$$

Das LGS $A\circ \vec{x}=\vec{b}$ ist eindeutig lösbar, und für den Lösungsvektor $\vec{x}= (x_1, x_2)^\top$ gilt:

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

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

I.4.13 Satz. *Sei $A \in \mathrm{GL}(n, K)$. Dann ist $A^{-1} = (y_{ij})$ gegeben durch*

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

Man beachte die Reihenfolge der Indizes!

Beweis: Sei $\vec{y}_j := (y_{1j}, \ldots, y_{nj})^\top$ die j-te Spalte von A^{-1}. Da $A \circ A^{-1} = 1_n$ ist, gilt:

$$A \circ \vec{y}_j = \vec{e}_j \quad \text{für } j = 1, \ldots, n.$$

Aus der Cramerschen Regel folgt dann:

$$\begin{aligned} y_{ij} &= \frac{1}{\det(A)} \cdot \det(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{e}_j, \vec{a}_{i+1}, \ldots, \vec{a}_n) \\ &= \frac{1}{\det(A)} \cdot (-1)^{i+j} \det S_{ji}(A). \end{aligned}$$

□

Beispiel:

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

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

I.4.14 Satz. *Sei $A \in M_{n,n}(K)$ nicht die Null–Matrix. Dann ist $\mathrm{rg}(A)$ die größte natürliche Zahl r, zu der es eine r-reihige Unterdeterminante $\neq 0$ von A gibt.*

Beweis:
1) Es gebe eine r-reihige Untermatrix A' von A mit $\det(A') \neq 0$. Dann sind die Spalten von A' und damit auch r Spalten von A linear unabhängig. Also ist $\mathrm{rg}(A) \geq r$. Das gilt auch für das größtmögliche r, zu dem noch eine nicht verschwindende r-reihige Unterdeterminante existiert.

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

Beispiele:

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

2. Sei

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

Man kann nachrechnen, daß alle 3-reihigen Unterdeterminanten Null sind. Man kann aber auch leicht sehen, daß $B = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$ mit $2 \cdot \vec{b}_1 + \vec{b}_2 = \vec{b}_3$ ist. Also ist $\mathrm{rg}(B) < 3$. Rechts unten findet sich die Unterdeterminante $\mathbf{j} \cdot 5 - 3 \cdot \mathbf{j} = 2\mathbf{j} \neq 0$. Also ist $\mathrm{rg}(B) = 2$.

§5 Skalarprodukte

Im Folgenden müssen wir zwischen $\mathbb{R}$ und $\mathbb{C}$ unterscheiden.

Definition.

Sind $\vec{x}$ und $\vec{y}$ Vektoren im $\mathbb{R}^n$, so wird das *(Euklidische) Skalarprodukt* $\vec{x} \bullet \vec{y}$ definiert durch

$$\vec{x} \bullet \vec{y} := \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n.$$

I.5.1 Satz. *Für das Skalarprodukt gelten folgende Regeln:*

1. $\vec{x} \bullet \vec{y} = \vec{y} \bullet \vec{x}$. *(Symmetrie)*
2. $(\vec{x} + \vec{y}) \bullet \vec{z} = \vec{x} \bullet \vec{z} + \vec{y} \bullet \vec{z}$.
3. *Ist* $\lambda \in \mathbb{R}$, *so ist* $(\lambda \cdot \vec{x}) \bullet \vec{y} = \lambda \cdot (\vec{x} \bullet \vec{y}) = \vec{x} \bullet (\lambda \cdot \vec{y})$.
4. $\vec{x} \bullet \vec{x} \geq 0$.
5. $\vec{x} \bullet \vec{x} = 0 \iff \vec{x} = \vec{0}$.

Beweis: 1), 2) und 3) sind trivial
4) und 5): $\vec{x} \bullet \vec{x} = x_1^2 + \cdots + x_n^2$ ist natürlich immer ≥ 0, und $= 0$ kann nur herauskommen, wenn $x_1 = \ldots = x_n = 0$ ist. □

Bemerkung: Die Abbildung $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit $(\vec{v}, \vec{w}) \mapsto \vec{v} \bullet \vec{w}$ ist multilinear. In diesem Fall (bei *zwei* Argumenten) sagt man *bilinear* dazu. Die Eigenschaften 4) und 5) faßt man unter der Bezeichnung *Positive Definitheit* zusammen.

Das Skalarprodukt ist eine *positiv definite symmetrische Bilinearform.*

Definition.

Für $\vec{x} \in \mathbb{R}^n$ ist die *Euklidische Norm* (oder *Länge*) gegeben durch

$$\| \vec{x} \| = \sqrt{\vec{x} \bullet \vec{x}}.$$

Für kleines n entspricht die Norm genau derjenigen Länge, die sich aus dem Satz des Pythagoras ergibt.

I.5.2 Satz. *Für die Norm gelten folgende Regeln:*

1. $\| \vec{v} \| \geq 0$ *für alle Vektoren* $\vec{v}$.
2. $\| \vec{v} \| = 0 \iff \vec{v} = \vec{0}$.

3. $\| \alpha \cdot \vec{v} \| = |\alpha| \cdot \| \vec{v} \|$.

4. $| \vec{v} \bullet \vec{w} | \le \| \vec{v} \| \cdot \| \vec{w} \|$.
Gleichheit gilt genau dann, wenn $\vec{v}$ *und* $\vec{w}$ *linear abhängig sind.*

5. Es gilt die „Dreiecksungleichung“:
$\| \vec{v} + \vec{w} \| \le \| \vec{v} \| + \| \vec{w} \|$.

BEWEIS: 1) und 2) ergeben sich sofort aus den Eigenschaften des Skalarproduktes.
3) Es ist $\| \alpha \cdot \vec{v} \| = \sqrt{(\alpha \cdot \vec{v}) \bullet (\alpha \cdot \vec{v})} = \sqrt{\alpha^2 \cdot (\vec{v} \bullet \vec{v})} = |\alpha| \cdot \| \vec{v} \|$.
4) Quadriert bedeutet die Ungleichung:

$$\left(\sum_{i=1}^{n} v_i w_i \right)^2 \le \left(\sum_{i=1}^{n} v_i^2 \right) \cdot \left(\sum_{i=1}^{n} w_i^2 \right).$$

Aber das ist nichts anderes als die schon im 1. Kapitel bewiesene Schwarzsche Ungleichung.

Die zweite Aussage ergibt sich aus einer genauen Analyse des Beweises der Schwarzschen Ungleichung, wir wollen hier nicht weiter darauf eingehen.

5) Daß die Dreiecksungleichung aus der Schwarzschen Ungleichung folgt, haben wir schon im Falle $n = 3$ bewiesen. Mit beliebigem n geht's genauso. □

Definition.

Die Vektoren $\vec{v}$ und $\vec{w}$ im $\mathbb{R}^n$ seien beide $\neq 0$. Dann wird der *Winkel* $\varphi = \angle(\vec{v}, \vec{w})$ definiert durch

1. $\cos(\varphi) = \dfrac{\vec{v} \bullet \vec{w}}{\| \vec{v} \| \cdot \| \vec{w} \|}$.
2. $0 \le \varphi \le \pi$.

Definition.

Zwei Vektoren $\vec{v}$ und $\vec{w}$ im $\mathbb{R}^n$ sind *orthogonal* zueinander (in Zeichen $\vec{v} \perp \vec{w}$), wenn $\vec{v} \bullet \vec{w} = 0$ ist.

Sind $\vec{v}$ und $\vec{w}$ beide $\neq \vec{0}$, so sind sie genau dann zueinander orthogonal, wenn

$$\angle(\vec{v}, \vec{w}) = \pi/2$$

ist. „Orthogonal“ bedeutet also „senkrecht“.

I.5.3 Satz des Pythagoras. *Sind* $\vec{x}$ *und* $\vec{y}$ *Vektoren mit* $\vec{x} \bullet \vec{y} = 0$, *so ist*

$$\| \vec{x} + \vec{y} \|^2 = \| \vec{x} \|^2 + \| \vec{y} \|^2.$$

Der BEWEIS kann wörtlich aus dem 1. Kapitel übernommen werden.

Definition.

Ein *Orthogonalsystem* im $\mathbb{R}^n$ ist eine Menge $\{\vec{v}_1, \ldots, \vec{v}_n\}$ von paarweise zueinander orthogonalen Vektoren $\neq \vec{0}$. (Es ist also $\vec{v}_i \bullet \vec{v}_j = 0$ für $i \neq j$).

Ein *Orthonormalsystem* (kurz: *ON-System*) ist ein Orthogonalsystem, in dem die Vektoren $\vec{v}_i$ zusätzlich normiert sind (also $\| \vec{v}_i \| = 1$ für alle i).

Eine *Orthonormalbasis* (kurz: *ON-Basis*) ist eine Basis des $\mathbb{R}^n$, die zugleich ein ON-System ist.

Natürlich kann man alle diese Begriffe auch in Untervektorräumen des $\mathbb{R}^n$ definieren und benutzen.

Man beachte: Ein System $\{\vec{v}_1, \ldots, \vec{v}_n\}$ von Vektoren des $\mathbb{R}^n$ ist genau dann ein ON-System, wenn gilt:

$$\vec{v}_i \bullet \vec{v}_j = \delta_{ij} \qquad \text{(Kronecker-Symbol)} \quad , \text{für alle } i, j\,.$$

Beispiele :

1. $\{\vec{e}_1, \ldots, \vec{e}_n\}$ ist eine ON-Basis des $\mathbb{R}^n$.

2. Die Vektoren $\vec{a}_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\vec{a}_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ bilden ein Orthogonalsystem im $\mathbb{R}^2$, sie sind aber nicht normiert.

   $\vec{b}_1 := \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\vec{b}_2 := \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ bilden ein ON-System und natürlich auch eine ON-Basis, da $\det(\vec{b}_1, \vec{b}_2) \neq 0$ ist.

Daß das ON-System im 2. Beispiel schon eine ON-Basis war, war kein Zufall:

I.5.4 Satz. *Sei $U \subset \mathbb{R}^n$ ein Untervektorraum der Dimension k. Dann ist jedes Orthogonalsystem mit k Elementen in U auch eine Basis von U.*

BEWEIS:
Es sei ein solches Orthogonalsystem $\{\vec{a}_1, \ldots, \vec{a}_k\}$ gegeben. Nach Voraussetzung sind alle $\vec{a}_i \neq \vec{0}$. Ist $\sum_{i=1}^{k} \lambda_i \, \vec{a}_i = \vec{0}$, so ist $0 = \vec{0} \bullet \vec{a}_l = \lambda_l \cdot \| \vec{a}_l \|^2$, also $\lambda_l = 0$ für $l = 1, \ldots, k$. Damit sind die Vektoren linear unabhängig, und da ihre Anzahl mit der Dimension von U übereinstimmt, müssen sie eine Basis von U bilden. $\square$

Die Überprüfung, ob ein System von Vektoren eine Basis bildet, ist in solchen Fällen sehr einfach geworden. Besonders bequem ist auch die Ermittlung der Koordinaten eines Vektors bezüglich einer ON-Basis:

I.5.5 Satz. *Sei $U \subset \mathbb{R}^n$ ein Untervektorraum, $\{\vec{a}_1, \ldots, \vec{a}_k\}$ eine ON-Basis von U. Ist $\vec{x} \in U$ ein beliebiger Vektor, so findet man die Koeffizienten x_i in der Darstellung*

$$\vec{x} = \sum_{i=1}^{k} x_i \, \vec{a}_i$$

durch die Formel $\boxed{x_i = \vec{x} \bullet \vec{a}_i, \text{ für } i = 1, \ldots, k.}$

Der Beweis ist trivial, die Aussage sehr nützlich. Man sollte sie sich unbedingt merken, weil später etwas ähnliches im Zusammenhang mit orthogonalen Funktionensystemen (z.B. bei Fourierreihen) vorkommt.

Im $\mathbb{R}^n$ kennen wir schon ein Beispiel für eine ON-Basis. Bei einem beliebigen Unterraum des $\mathbb{R}^n$ stellt sich zunächst die Frage, ob es dort überhaupt ON-Basen gibt. Daß das in der Tat der Fall ist, zeigt der folgende Satz:

I.5.6 Schmidt'sches Orthogonalisierungsverfahren.
Sei $U \subset \mathbb{R}^n$ ein k-dimensionaler Untervektorraum und $\{\vec{x}_1, \ldots, \vec{x}_k\}$ eine Basis von U. Dann gibt es ein ON-System $\{\vec{a}_1, \ldots, \vec{a}_k\}$ in U, so daß gilt:

$$<\vec{a}_1, \ldots, \vec{a}_l> = <\vec{x}_1, \ldots, \vec{x}_l> \quad \text{für } l = 1, \ldots, k.$$

Beweis: Wir konstruieren die $\vec{a}_i$ aus den $\vec{x}_i$ und benutzen dabei vollständige Induktion:

Sei $\vec{a}_1 := \frac{1}{\| \vec{x}_1 \|} \cdot \vec{x}_1$. Dann ist $\| \vec{a}_1 \| = 1$ und $<\vec{a}_1> = <\vec{x}_1>$.

Nun nehmen wir an, wir hätten $\vec{a}_1, \ldots, \vec{a}_l$ schon mit den gewünschten Eigenschaften konstruiert. Dann sei

$$\vec{b}_{l+1} := \vec{x}_{l+1} - \sum_{i=1}^{l} (\vec{x}_{l+1} \bullet \vec{a}_i)\, \vec{a}_i \, .$$

Nun gilt:

1. $\vec{b}_{l+1} \in <\vec{x}_1, \ldots, \vec{x}_{l+1}>$.
2. $\vec{a}_1, \ldots, \vec{a}_l, \vec{b}_{l+1}$ sind linear unabhängig.
3. Für $j = 1, \ldots, l$ ist $\vec{b}_{l+1} \bullet \vec{a}_j = \vec{x}_{l+1} \bullet \vec{a}_j - \vec{x}_{l+1} \bullet \vec{a}_j = 0$.

Also bilden $\vec{a}_1, \ldots, \vec{a}_l, \vec{b}_{l+1}$ ein Orthogonalsystem mit $<\vec{a}_1, \ldots, \vec{a}_l, \vec{b}_{l+1}> = <\vec{x}_1, \ldots, \vec{x}_{l+1}>$.

Schließlich setzen wir $\vec{a}_{l+1} := \frac{1}{\| \vec{b}_{l+1} \|} \cdot \vec{b}_{l+1}$. □

I.5.7 Folgerung. *Jeder Untervektorraum des $\mathbb{R}^n$ besitzt eine ON-Basis.*

Beweis: Klar! □

Beispiel:

Die Vektoren $\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ und $\vec{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ bilden eine Basis des Untervektorraumes $U := \{\vec{x} \in \mathbb{R}^3 \mid x_3 = 0\}$. Das Schmidt'sche Verfahren liefert:

$$\vec{a}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

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

Definition.

Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so heißt

$$U^\perp := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \bullet \mathbf{u} = 0 \text{ für alle } \mathbf{u} \in U\}$$

das *orthogonale Komplement* von U.

I.5.8 Satz. *$U^\perp$ ist ein Untervektorraum.*

Beweis: 1) Offensichtlich ist $\mathbf{0} \in U^\perp$.
2) Sind $\mathbf{x}, \mathbf{y} \in U^\perp$, so ist $(\mathbf{x} + \mathbf{y}) \bullet \mathbf{u} = \mathbf{x} \bullet \mathbf{u} + \mathbf{y} \bullet \mathbf{u} = 0$ für alle $\mathbf{u} \in U$, also $\mathbf{x} + \mathbf{y} \in U^\perp$.
3) Ist $\mathbf{x} \in U^\perp$ und $\lambda \in \mathbb{R}$, so ist $(\lambda \cdot \mathbf{x}) \bullet \mathbf{u} = \lambda \cdot (\mathbf{x} \bullet \mathbf{u}) = 0$ für alle $\mathbf{u} \in U$,
also $\lambda \cdot \mathbf{x} \in U^\perp$. □

I.5.9 Satz. *Sei $U \subset \mathbb{R}^n$ ein Untervektorraum. Dann gilt:*

1. $\dim(U) + \dim(U^\perp) = n$.
2. *Jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ kann eindeutig zerlegt werden:*

$$\mathbf{x} = \mathbf{u} + \mathbf{v}, \quad \textit{mit } \mathbf{u} \in U \textit{ und } \mathbf{v} \in U^\perp.$$

3. *Es ist $U \cap U^\perp = \{\mathbf{0}\}$.*

Beweis: 1) Wir wählen eine ON-Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ von U.
Dann definieren wir eine lineare Abbildung $F : \mathbb{R}^n \to \mathbb{R}^k$ durch

$$F(\mathbf{x}) := (\mathbf{x} \bullet \mathbf{a}_1, \ldots, \mathbf{x} \bullet \mathbf{a}_k).$$

Dann ist $F(\mathbf{a}_i) = \mathbf{e}_i$ der i-te Einheitsvektor, $i = 1, \ldots, k$. Also ist F surjektiv.

$$\begin{aligned}
\text{Weiter ist} \quad F(\mathbf{x}) = 0 &\iff \mathbf{x} \bullet \mathbf{u} = 0 \text{ für alle } \mathbf{u} \in U \\
&\iff \mathbf{x} \in U^\perp,
\end{aligned}$$

also $U^\perp = \operatorname{Ker}(F)$. Nach der Dimensionsformel ist

$$\dim(U) + \dim(U^\perp) = k + \dim(\operatorname{Ker}(F)) = \dim(\operatorname{Im}(F)) + \dim(\operatorname{Ker}(F)) = n.$$

2) Es gibt eine ON-Basis $\{\mathbf{a}_{k+1}, \dots, \mathbf{a}_n\}$ von $U^\perp$. Da definitionsgemäß $\mathbf{a}_i \bullet \mathbf{a}_j = 0$ für $i = 1, \dots, k$ und $j = k+1, \dots, n$ ist, bildet

$$\{\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{a}_{k+1}, \dots, \mathbf{a}_n\}$$

eine ON-Basis des ganzen $\mathbb{R}^n$.

Ist $\mathbf{x} \in \mathbb{R}^n$ beliebig vorgegeben, so ist $\mathbf{x} = \sum\limits_{\nu=1}^{n} \lambda_\nu \mathbf{a}_\nu = \mathbf{u} + \mathbf{v}$ mit

$$\mathbf{u} := \sum_{i=1}^{k} \lambda_i \mathbf{a}_i \quad \text{und} \quad \mathbf{v} := \sum_{j=k+1}^{n} \lambda_j \mathbf{a}_j.$$

Das ist die gewünschte Zerlegung. Die Eindeutigkeit folgt aus (3).

3) Ist speziell $\mathbf{x} = \mathbf{u} + \mathbf{v} \in U$ so wie in (2) zerlegt, so ist $\mathbf{x} = \mathbf{u}$, also $\mathbf{v} = \mathbf{0}$ (wegen der Eindeutigkeit der Darstellung von $\mathbf{x}$ bezüglich einer Basis). Ist zusätzlich $\mathbf{x} \in U^\perp$, so ist $\mathbf{x} = \mathbf{v} = \mathbf{0}$. Damit haben wir (3) gezeigt. Und zusätzlich bekommen wir die Eindeutigkeit der Zerlegung:

Sind durch $\mathbf{x} = \mathbf{u} + \mathbf{v} = \mathbf{u}' + \mathbf{v}'$ zwei Zerlegungen gegeben, so ist $\mathbf{u} - \mathbf{u}' = \mathbf{v}' - \mathbf{v} \in U \cap U'$, also $= \mathbf{0}$. Aber das bedeutet, daß $\mathbf{u} = \mathbf{u}'$ und $\mathbf{v} = \mathbf{v}'$ ist. □

I.5.10 Folgerung. *Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so ist $U^{\perp\perp} = U$.*

Beweis: Ist $\mathbf{x} \in U$, so ist $\mathbf{x} \bullet \mathbf{v} = 0$ für alle $\mathbf{v} \in U^\perp$. Daher ist $U \subset U^{\perp\perp}$. Wegen der Dimensionsformel ist

$$\dim(U^{\perp\perp}) = n - \dim(U^\perp) = n - (n - \dim(U)) = \dim(U).$$

Also muß $U^{\perp\perp} = U$ sein. □

Von unseren rechtwinkligen Koordinatensystemen her kennen wir „orthogonale Projektionen“. Das soll nun verallgemeinert werden. Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so kann man eine Abbildung von $\mathbb{R}^n$ auf U wie folgt definieren:

Da jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ eine **eindeutige** Zerlegung $\mathbf{x} = \mathbf{u} + \mathbf{v}$ mit $\mathbf{u} \in U$ und $\mathbf{v} \in U^\perp$ besitzt, kann man

$$\mathrm{pr}_U : \mathbb{R}^n \to U$$

definieren durch

$$\mathbf{u} + \mathbf{v} \mapsto \mathbf{u}.$$

Diese Abbildung ist vernünftig definiert und - wie man leicht nachrechnen kann - auch linear. Wir nennen pr_U die *orthogonale Projektion auf U*. Sie hat folgende Eigenschaften:

1. $\mathrm{pr}_U \circ \mathrm{pr}_U(\mathbf{x}) = \mathrm{pr}_U(\mathbf{x})$ für alle $\mathbf{x} \in \mathbb{R}^n$.
2. $\mathrm{Ker}\,(\mathrm{pr}_U) = U^\perp$.
3. $\mathrm{pr}_U : \mathbb{R}^n \to U$ ist surjektiv.
4. $\|\, \mathbf{x} - \mathrm{pr}_U(\mathbf{x}) \,\| \le \|\, \mathbf{x} - \mathbf{u} \,\|$ für alle $\mathbf{u} \in U$.

Zum Beweis:

1) Ist $\mathbf{x} = \mathbf{u} + \mathbf{v}$, so ist $\mathrm{pr}_U(\mathbf{x}) = \mathbf{u} = \mathbf{u} + \mathbf{0}$, und daher

$$\mathrm{pr}_U \circ \mathrm{pr}_U(\mathbf{x}) = \mathrm{pr}_U(\mathbf{u}) = \mathbf{u} = \mathrm{pr}_U(\mathbf{x}).$$

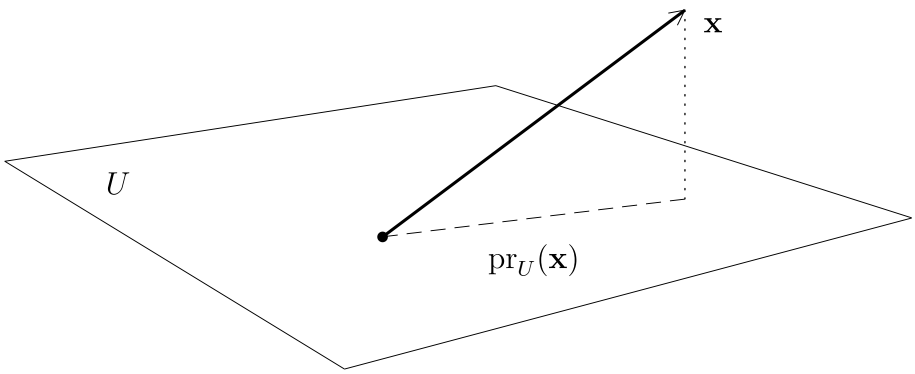
2) Es ist $\mathrm{pr}_U(\mathbf{u} + \mathbf{v}) = \mathbf{0} \iff \mathbf{u} = \mathbf{0} \iff \mathbf{u} + \mathbf{v} = \mathbf{v} \in U^\perp$.

3) ist klar.

4) Sei $\mathbf{x} = \mathbf{u}_0 + \mathbf{v}_0$ mit $\mathbf{u}_0 \in U$ und $\mathbf{v}_0 \in U^\perp$. Dann ist $\mathrm{pr}_U(\mathbf{x}) = \mathbf{u}_0$, und für $\mathbf{u} \in U$ ist $(\mathbf{x} - \mathbf{u}_0) \bullet (\mathbf{u}_0 - \mathbf{u}) = \mathbf{v}_0 \bullet (\mathbf{u}_0 - \mathbf{u}) = 0$. Daraus folgt:

$$\begin{aligned} \| \mathbf{x} - \mathbf{u} \|^2 &= \| (\mathbf{x} - \mathbf{u}_0) + (\mathbf{u}_0 - \mathbf{u}) \|^2 \\ &= \| \mathbf{x} - \mathbf{u}_0 \|^2 + \| \mathbf{u}_0 - \mathbf{u} \|^2 \\ &\geq \| \mathbf{x} - \mathbf{u}_0 \|^2. \end{aligned}$$

Die orthogonale Projektion von $\mathbf{x}$ auf U ist also derjenige Vektor $\mathbf{u} \in U$, der von $\mathbf{x}$ den kleinsten Abstand hat, der also $\mathbf{x}$ am besten approximiert.



Die Berechnung der orthogonalen Projektion eines Vektors ist recht einfach:

Wir wählen eine ON-Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ von U und ergänzen sie wie oben zu einer ON-Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ des ganzen $\mathbb{R}^n$. Dann gibt es eine Darstellung

$$\mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{a}_i, \text{ mit } \lambda_i = \mathbf{x} \bullet \mathbf{a}_i.$$

Offensichtlich ist

$$\mathrm{pr}_U(\mathbf{x}) = \sum_{i=1}^{k} (\mathbf{x} \bullet \mathbf{a}_i)\mathbf{a}_i,$$

und dafür braucht man die zusätzlichen Basisvektoren $\mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ gar nicht zu kennen!

Wir wenden uns jetzt den komplexen Vektorräumen zu:

Zunächst stellen wir fest, daß $\mathbb{R}^n \times \mathbb{R}^n$ folgendermaßen zu einem reellen Vektorraum gemacht werden kann:

$$\begin{aligned} (\vec{a}, \vec{b}) + (\vec{c}, \vec{d}) &:= (\vec{a} + \vec{c}, \vec{b} + \vec{d}), \\ \lambda \cdot (\vec{a}, \vec{b}) &:= (\lambda \cdot \vec{a}, \lambda \cdot \vec{b}). \end{aligned}$$

Natürlich ist dieser Vektorraum isomorph zu $\mathbb{R}^{2n}$, ein Isomorphismus ist etwa durch

$$(\vec{a}, \vec{b}) \quad \longleftrightarrow \quad \begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix}$$

gegeben.

Als nächstes definieren wir einen **reellen** Isomorphismus

$$\Phi : \mathbb{C}^n \to \mathbb{R}^n \times \mathbb{R}^n, \text{ durch } \vec{z} = \vec{x} + \mathbf{j}\,\vec{y} \mapsto (\vec{x}, \vec{y}).$$

Man sollte sich das so vorstellen, daß beim Übergang von links nach rechts einfach die „komplexe Struktur“ vergessen wird. Da sie in der umgekehrten Richtung aber leicht wieder rekonstruiert werden kann, lassen sich alle Eigentümlichkeiten der komplexen Struktur des $\mathbb{C}^n$ auch im $\mathbb{R}^n \times \mathbb{R}^n$ wiederfinden.

Multiplizieren wir etwa links den komplexen Vektor $\vec{z} = \vec{x} + \mathbf{j}\,\vec{y}$ mit $\mathbf{j}$, so müssen wir auf der rechten Seite $(\vec{x}, \vec{y})$ durch $(-\vec{y}, \vec{x})$ ersetzen. Das liefert dort einen Isomorphismus $J : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$.

Der Betrag einer komplexen Zahl $\lambda = \alpha + \mathbf{j}\beta$ ist bekanntlich die Zahl

$$|\,\lambda\,| := \sqrt{\lambda\overline{\lambda}} = \sqrt{\alpha^2 + \beta^2}.$$

Das motiviert die folgende Definition:

Definition.

1. Sind $\vec{z}$ und $\vec{w}$ Vektoren im $\mathbb{C}^n$, so wird das *Hermitesche Skalarprodukt* $<\vec{z}, \vec{w}>$ definiert durch

$$<\vec{z}, \vec{w}> := \sum_{i=1}^{n} z_i \overline{w}_i = z_1\overline{w}_1 + \cdots + z_n\overline{w}_n.$$

2. Für $\vec{z} \in \mathbb{C}^n$ ist die *Norm* definiert durch

$$\|\,\vec{z}\,\| := \sqrt{<\vec{z}, \vec{z}>}.$$

I.5.11 Satz. *Für das Hermitesche Skalarprodukt gelten folgende Regeln:*

1. $<\vec{z}, \vec{w}> = \overline{<\vec{w}, \vec{z}>}$.
2. $<\vec{z}_1 + \vec{z}_2, \vec{w}> = <\vec{z}_1, \vec{w}> + <\vec{z}_2, \vec{w}>$.
3. $<\alpha\cdot\vec{z}, \vec{w}> = \alpha\cdot <\vec{z}, \vec{w}>$ *und* $<\vec{z}, \alpha\cdot\vec{w}> = \overline{\alpha}\cdot <\vec{z}, \vec{w}>$.
4. $<\vec{z}, \vec{z}>$ *ist stets* **reell** *und* ≥ 0.
5. $<\vec{z}, \vec{z}> = 0 \iff z = 0$.

Beweis: 1) $\overline{w_1\overline{z}_1 + \cdots + w_n\overline{z}_n} = z_1\overline{w}_1 + \cdots + z_n\overline{w}_n$.

Die Punkte 2) bis 5) werden wie beim reellen Skalarprodukt bewiesen. Speziell ist

$$<\vec{z}, \vec{z}> = z_1\overline{z}_1 + \cdots + z_n\overline{z}_n = |\, z_1\,|^2 + \cdots + |\, z_n\,|^2.$$

□

Ist $\vec{z}=\vec{x} + \mathbf{j}\,\vec{y}$ und $\vec{w}=\vec{u} + \mathbf{j}\,\vec{v}$, so kann man auch die Real- und Imaginärteile $\vec{x}, \vec{y}, \vec{u}, \vec{v}$ als Elemente von $\mathbb{C}^n$ auffassen. Dann gilt:

$$\begin{aligned} <\vec{z}, \vec{w}> &= <\vec{x} + \mathbf{j}\,\vec{y}, \vec{u} + \mathbf{j}\,\vec{v}> \\ &= <\vec{x}, \vec{u}> + \mathbf{j} <\vec{y}, \vec{u}> - \mathbf{j} <\vec{x}, \vec{v}> + <\vec{y}, \vec{v}> \\ &= \vec{x} \bullet \vec{u} + \vec{y} \bullet \vec{v} + \mathbf{j}(\vec{y} \bullet \vec{u} - \vec{x} \bullet \vec{v}), \end{aligned}$$

also

$$\boxed{\mathrm{Re}\ <\vec{z}, \vec{w}> = \mathrm{Re}\,(\vec{z}) \bullet \mathrm{Re}\,(\vec{w}) + \mathrm{Im}\,(\vec{z}) \bullet \mathrm{Im}\,(\vec{w}) = \Phi(\vec{z}) \bullet \Phi(\vec{w}).}$$

Offensichtlich ist die Norm eines Vektors $\vec{z}$ im $\mathbb{C}^n$

$$\|\, \vec{z}\, \| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$$

nichts anderes als die euklidische Norm des entsprechenden Vektors $\Phi(\vec{z})$ im $\mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{R}^{2n}$. Daher gelten auch die gleichen Regeln wie im Reellen. Sogar die Schwarzsche Ungleichung läßt sich übertragen (was nicht ganz so selbstverständlich ist):

$$|<\vec{z}, \vec{w}>|^2 \le \|\, \vec{z}\,\|^2 \cdot \|\, \vec{w}\,\|^2.$$

Beim Skalarprodukt und bei der Orthogonalität ist die Situation etwas komplizierter, die Anschaulichkeit aus dem Reellen geht etwas verloren. Es gilt:

I.5.12 Satz.

$$<\vec{z}, \vec{w}> = 0 \iff \Phi(\vec{z}) \perp \Phi(\vec{w}) \ \ \textit{und}\ \ \Phi(\mathbf{j}\,\vec{z}) \perp \Phi(\vec{w}).$$

Beweis: Wir haben oben gesehen:

$$\begin{aligned} <\vec{z}, \vec{w}> &= \Phi(\vec{z}) \bullet \Phi(\vec{w}) + \mathbf{j} \cdot (-J(\Phi(\vec{z}))) \bullet \Phi(\vec{w}) \\ &= \Phi(\vec{z}) \bullet \Phi(\vec{w}) - \mathbf{j} \cdot \Phi(\mathbf{j}\,\vec{z}) \bullet \Phi(\vec{w}). \end{aligned}$$

Daraus folgt sofort der Satz. □

Orthogonal- und Orthonormalsysteme werden im Komplexen genauso wie im Reellen behandelt. Das Orthogonalisierungsverfahren läßt sich ebenfalls übertragen, und auch der Satz über das orthogonale Komplement.

Als nächstes wollen wir uns (für $K = \mathbb{R}$ und $K = \mathbb{C}$) mit dem Problem befassen, welche linearen Abbildungen Orthonormalbasen in Orthonormalbasen abbilden. Wir beginnen wieder mit dem reellen Fall:

Definition.

Eine lineare Abbildung $F : \mathbb{R}^n \to \mathbb{R}^n$ heißt eine *Isometrie*, falls gilt:

1. F ist surjektiv.
2. $\| F(\vec{x}) \| = \| \vec{x} \|$ für alle $\vec{x} \in \mathbb{R}^n$.

I.5.13 Satz. *Ist $F : \mathbb{R}^n \to \mathbb{R}^n$ eine Isometrie, so ist F ein Isomorphismus und $F(\vec{x}) \bullet F(\vec{y}) = \vec{x} \bullet \vec{y}$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$.*

Beweis:
1) Eine surjektive lineare Abbildung $F : \mathbb{R}^n \to \mathbb{R}^n$ ist automatisch ein Isomorphismus.

2) Es ist $\| F(\vec{x}) - F(\vec{y}) \|^2 = \| F(\vec{x} - \vec{y}) \|^2 = \| \vec{x} - \vec{y} \|^2$. Weiter gilt:

$$\begin{aligned} \| \vec{x} - \vec{y} \|^2 &= (\vec{x} - \vec{y}) \bullet (\vec{x} - \vec{y}) \\ &= \vec{x} \bullet \vec{x} + \vec{y} \bullet \vec{y} - 2\, \vec{x} \bullet \vec{y} \\ &= \| \vec{x} \|^2 + \| \vec{y} \|^2 - 2\, \vec{x} \bullet \vec{y}, \end{aligned}$$

also $\vec{x} \bullet \vec{y} = \dfrac{1}{2} \cdot (\| \vec{x} - \vec{y} \|^2 - \| \vec{x} \|^2 - \| \vec{y} \|^2)$.

Damit folgt die zweite Behauptung. □

Eine Isometrie ist also nicht nur längentreu, sondern sie respektiert auch das Skalarprodukt und ist daher winkeltreu.

I.5.14 Satz. *Sei $A \in M_{n,n}(\mathbb{R})$, $F = F_A : \mathbb{R}^n \to \mathbb{R}^n$. F ist genau dann eine Isometrie, wenn $A^\top \circ A = 1$ ist.*

Beweis: Wir benutzen, daß $\vec{x} \bullet \vec{y} = \vec{x}^\top \circ \vec{y}$ ist.

Nun ist die lineare Abbildung F_A genau dann eine Isometrie, wenn $\mathrm{rg}(A) = n$ ist, **und**

$$\vec{x}^\top \circ A^\top \circ A \circ \vec{y} = (A \circ \vec{x})^\top \circ (A \circ \vec{y}) = F_A(\vec{x})^\top \circ F_A(\vec{y}) = \vec{x}^\top \circ \vec{y}$$

für alle $\vec{x}, \vec{y}$.

Ist $B = (b_{ij})$ eine beliebige Matrix in $M_{n,n}(\mathbb{R})$, so ist $\vec{e}_i^{\,\top} \circ B \circ \vec{e}_j = b_{ij}$, für alle i, j.
Aus dem obigen Kriterium folgt, daß

$$\delta_{ij} = \vec{e}_i^{\,\top} \circ \vec{e}_j = \vec{e}_i^{\,\top} \circ (A^\top \circ A) \circ \vec{e}_j$$

ist, also $A^\top \circ A = 1$.

Sei umgekehrt $A^\top \circ A = 1$. Dann ist

$$1 = \det(1) = \det(A^\top \circ A) = \det(A)^2,$$

also $\det(A) \neq 0$ und damit $\mathrm{rg}(A) = n$. Insbesondere ist A dann invertierbar und $A^{-1} = A^\top$. Und natürlich ist dann

$$(A \circ \vec{x})^\top \circ (A \circ \vec{y}) = \vec{x}^\top \circ \vec{y}\,.$$

□

Definition.

$A \in M_{n,n}(\mathbb{R})$ heißt eine *orthogonale Matrix*, falls $A^\top \circ A = 1$ ist.

I.5.15 Satz. *Folgende Aussagen über eine Matrix $A \in M_{n,n}(\mathbb{R})$ sind äquivalent:*

1. *A ist orthogonal.*
2. *Die Spalten von A bilden ein ON-System.*
3. *Die Zeilen von A bilden ein ON-System.*
4. *F_A bildet ON-Basen auf ON-Basen ab.*
5. *F_A ist eine Isometrie.*

Beweis:
5) $\Longrightarrow$ 4): Ist F_A Isometrie, so ist F_A ein Isomorphismus und $F_A(\vec{x}) \bullet F_A(\vec{y}) = \vec{x} \bullet \vec{y}$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$. Dann ist aber klar, daß F_A ON-Basen auf ON-Basen abbildet.

4) $\Longrightarrow$ 2): F_A bilde ON-Basen auf ON-Basen ab. Da die Standardbasis eine ON-Basis ist und $F_A(\vec{e}_i) = A \circ \vec{e}_i$ die i-te Spalte von A ergibt, bilden die Spalten von A eine ON-Basis.

2) $\Longrightarrow$ 1): Ist $A = (\vec{a}_1, \ldots, \vec{a}_n)$, so sind die Einträge in $A^\top \circ A$ gerade die Produkte $\vec{a}_i^{\,\top} \circ \vec{a}_j = \vec{a}_i \bullet \vec{a}_j$, also $= \delta_{ij}$, wenn die Spalten von A ein ON-System bilden. Dann ist $A^\top \circ A = 1$.

1) $\Longrightarrow$ 5): Das haben wir schon gezeigt!

Es fehlt noch die Äquivalenz von 2) und 3): Aber die Gleichungen $A^\top \circ A = 1$ und $A \circ A^\top = 1$ sind gleichwertig, beide haben zur Folge, daß A invertierbar und $A^{-1} = A^\top$ ist, daß also jeweils auch die andere Gleichung gilt. □

I.5.16 Satz. *Ist A orthogonal, so ist $|\det(A)| = 1$.*

Beweis: Wir haben oben schon gesehen, daß $\det(A)^2 = 1$ ist. □

Im Falle $n = 2$ wollen wir alle orthogonalen Matrizen bestimmen:

Sei also $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ orthogonal. Da dann die Spaltenvektoren ein ON-System bilden müssen, ist

$$a^2 + c^2 = b^2 + d^2 = 1.$$

Außerdem ist $\varepsilon := ad - bc = \det(A) = \pm 1$. Wir müssen also zwei Fälle unterscheiden:

1) Ist $\varepsilon = +1$, so ist

$$\begin{aligned} (a-d)^2 + (b+c)^2 &= a^2 + b^2 + c^2 + d^2 + 2(bc - ad) \\ &= 1 + 1 - 2 \cdot 1 = 0, \end{aligned}$$

also $\quad a = d \quad$ und $\quad b = -c.$

Wegen $a^2 + c^2 = 1$ gibt es ein $\varphi \in [0, 2\pi)$, so daß $a = \cos(\varphi)$ und $c = \sin(\varphi)$ ist. Damit ist

$$A = R(\varphi) := \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

eine *Drehmatrix zum Winkel* φ.
$R(\varphi)$ bildet z.B. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ auf $\begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$ ab.

2) Ist $\varepsilon = -1$, so ist auch $A \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ orthogonal, und die neue Matrix besitzt die Determinante $+1$. Also gibt es ein φ, so daß

$$A = R(\varphi) \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$$

ist. Wir betrachten hier zunächst einmal den Spezialfall $\varphi = 0$. Dann ist $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ und $A \circ \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$. Das ist eine Spiegelung an der x–Achse.
Ist φ beliebig, so gilt:

$$\begin{aligned}
R(\frac{\varphi}{2}) \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \circ R(-\frac{\varphi}{2}) \; = \\
= \begin{pmatrix} \cos(\frac{\varphi}{2}) & -\sin(\frac{\varphi}{2}) \\ \sin(\frac{\varphi}{2}) & \cos(\frac{\varphi}{2}) \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \circ \begin{pmatrix} \cos(\frac{\varphi}{2}) & \sin(\frac{\varphi}{2}) \\ -\sin(\frac{\varphi}{2}) & \cos(\frac{\varphi}{2}) \end{pmatrix} \\
= \begin{pmatrix} \cos(\frac{\varphi}{2}) & -\sin(\frac{\varphi}{2}) \\ \sin(\frac{\varphi}{2}) & \cos(\frac{\varphi}{2}) \end{pmatrix} \circ \begin{pmatrix} \cos(\frac{\varphi}{2}) & \sin(\frac{\varphi}{2}) \\ \sin(\frac{\varphi}{2}) & -\cos(\frac{\varphi}{2}) \end{pmatrix} \\
= \begin{pmatrix} \cos^2(\frac{\varphi}{2}) - \sin^2(\frac{\varphi}{2}) & 2\sin(\frac{\varphi}{2})\cos(\frac{\varphi}{2}) \\ 2\sin(\frac{\varphi}{2})\cos(\frac{\varphi}{2}) & -\cos^2(\frac{\varphi}{2}) + \sin^2(\frac{\varphi}{2}) \end{pmatrix} \\
= \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix} \; = \; A.
\end{aligned}$$

Das zeigt, daß F_A in diesem Fall die Spiegelung an derjenigen Geraden ist, die mit der positiven x–Achse den Winkel $\frac{\varphi}{2}$ einschließt.

Wir benutzen die Ergebnisse in der Dimension 2 als Motivation für die folgende allgemeinere Definition:

Definition.

Eine orthogonale Matrix $A \in M_{n,n}(\mathbb{R})$ heißt *Drehung*, falls $\det(A) = 1$ ist, und *Drehspiegelung*, falls $\det(A) = -1$ ist.

Man kann zeigen, daß eine Drehspiegelung immer als Verknüpfung einer Drehung mit einer Spiegelung an einer Hyperebene[8] geschrieben werden kann.

Im Komplexen sehen die Isometrien etwas anders aus:
Da $<\vec{z},\vec{w}> = \vec{z}^\top \circ \overline{\vec{w}}$ ist, folgt:

I.5.17 Satz. *Sei $A \in M_{n,n}(\mathbb{C})$, $F = F_A : \mathbb{C}^n \to \mathbb{C}^n$. F ist genau dann eine Isometrie (bezüglich des hermiteschen Skalarproduktes), wenn $A^\top \circ \overline{A} = 1$ ist.*

Die Bedingung ist äquivalent zu der Bedingung $\overline{A}^\top \circ A = 1$.

Definition.

> $A \in M_{n,n}(\mathbb{C})$ heißt eine *unitäre Matrix*, falls $\overline{A}^\top \circ A = 1$ ist.

Die äquivalenten Bedingungen für die Orthogonalität einer Matrix lassen sich sinngemäß auf die unitären Matrizen übertragen. Weiter folgt:

$$A \text{ unitär} \implies |\det(A)|^2 = 1.$$

Aber VORSICHT! Eine komplexe Zahl vom Betrag 1 hat die Gestalt $e^{\mathbf{j}t}$. Daher kann man die Unterscheidung zwischen Drehungen und Drehspiegelungen **nicht** auf unitäre Matrizen übertragen!

Um die Notationen zu vereinfachen, setzt man gelegentlich auch

$$A^* := \begin{cases} A^\top & \text{falls } A \text{ reell,} \\ \overline{A}^\top & \text{falls } A \text{ komplex.} \end{cases}$$

Man nennt A^* die zu A *konjugierte Matrix*.

A ist also orthogonal (im reellen Fall) bzw. unitär (im komplexen Fall), wenn $A^*A = 1$ ist.

Ist $\{\vec{a}_1, \ldots, \vec{a}_n\}$ eine Basis des $\mathbb{R}^n$ bzw. $\mathbb{C}^n$, so kann man aus den Skalarprodukten der Basisvektoren eine Matrix $H := (h_{ij})$ aus $M_{n,n}(\mathbb{R})$ bzw. $M_{n,n}(\mathbb{C})$ bilden, durch

$$h_{ij} := <\vec{a}_i, \vec{a}_j> \qquad \text{für } i, j = 1, \ldots, n.$$

Offensichtlich ist $A^* = A$.

Definition.

> Eine Matrix A mit $A^* = A$ heißt *symmetrisch* (im reellen Fall) bzw. *hermitesch* (im komplexen Fall).

Den symmetrischen oder hermiteschen Matrizen entsprechen spezielle lineare Abbildungen:

---

[8]Eine *Hyperebene* im $\mathbb{R}^n$ ist ein (n-1)-dimensionaler Untervektorraum.

Definition.

Eine lineare Abbildung $f : \mathbb{C}^n \to \mathbb{C}^n$ mit $< f(\vec{z}), \vec{w} > = < \vec{z}, f(\vec{w}) >$ für alle $\vec{z}, \vec{w} \in \mathbb{C}^n$ heißt *selbstadjungiert.*

(Im Reellen wird „selbstadjungiert“ völlig analog definiert, aber natürlich mit Hilfe des Euklidischen Skalarproduktes!)

I.5.18 Satz.

1. *Eine Matrix* $A \in M_{n,n}(\mathbb{R})$ *ist genau dann symmetrisch, wenn* $F_A : \mathbb{R}^n \to \mathbb{R}^n$ *selbstadjungiert ist.*

2. *Eine Matrix* $A \in M_{n,n}(\mathbb{C})$ *ist genau dann hermitesch, wenn* $F_A : \mathbb{C}^n \to \mathbb{C}^n$ *selbstadjungiert ist.*

BEWEIS: Wir betrachten nur den komplexen Fall, der Reelle ist darin enthalten.

Es ist genau dann $< A \circ \vec{z}, \vec{w} > = < \vec{z}, A \circ \vec{w} >$ für alle $\vec{z}, \vec{w}$, wenn $A^\top = \overline{A}$ ist. $\square$

§6 Eigenwerte und Eigenvektoren

Definition.

Sei V ein K-Vektorraum, $F : V \to V$ linear.

Ein Vektor $\mathbf{x} \in V$ heißt *Eigenvektor* von F, falls gilt:

1. $\mathbf{x}$ ist nicht der Nullvektor.
2. Es gibt ein $\lambda \in K$, so daß $F(\mathbf{x}) = \lambda \cdot \mathbf{x}$ ist.

Der Skalar λ heißt *Eigenwert* von F zum Eigenvektor $\mathbf{x}$.

Ist $A \in M_{n,n}(K)$ eine quadratische Matrix, so versteht man unter einem Eigenvektor bzw. Eigenwert von A einen Eigenvektor bzw. Eigenwert von $F_A : K^n \to K^n$.

Beispiel:

Sei $A := \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ und $\vec{x} := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Dann ist

$$A \circ \vec{x} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 \\ 3 \cdot 2 + 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4 \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Also ist 4 ein Eigenwert von A zum Eigenvektor $\vec{x}$.

I.6.1 Satz. *Sei V ein K-Vektorraum, $F : V \to V$ linear und $\lambda \in K$. Dann ist*

$$E(\lambda) := \{\mathbf{x} \in V \mid F(\mathbf{x}) = \lambda \cdot \mathbf{x}\}$$

ein Untervektorraum von V.

Beweis:

1) Offensichtlich liegt $\mathbf{0}$ in $E(\lambda)$. Man beachte allerdings, daß $\mathbf{0}$ kein Eigenvektor ist!

2) Ist $\mathbf{x} \in E(\lambda)$, also $F(\mathbf{x}) = \lambda \cdot \mathbf{x}$, so ist auch

$$F(\alpha\mathbf{x}) = \alpha \cdot F(\mathbf{x}) = \alpha(\lambda\mathbf{x}) = \lambda(\alpha\mathbf{x}),$$

also $\alpha\mathbf{x} \in E(\lambda)$.

3) Sind $\mathbf{x}, \mathbf{y} \in E(\lambda)$, so ist $F(\mathbf{x}) = \lambda\mathbf{x}$ und $F(\mathbf{y}) = \lambda\mathbf{y}$. Also ist

$$F(\mathbf{x} + \mathbf{y}) = F(\mathbf{x}) + F(\mathbf{y}) = \lambda \cdot \mathbf{x} + \lambda \cdot \mathbf{y} = \lambda \cdot (\mathbf{x} + \mathbf{y}),$$

also $\mathbf{x} + \mathbf{y} \in E(\lambda)$. □

Definition.

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

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

Wozu sind Eigenvektoren gut?

Definition.

Sei V ein endlich-dimensionaler K-Vektorraum. Eine lineare Abbildung $F : V \to V$ heißt *diagonalisierbar*, wenn es eine Basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ von V gibt, bzgl. der F durch eine Diagonalmatrix beschrieben wird.

Eine Matrix $A \in M_{n,n}(K)$ heißt diagonalisierbar, wenn $F_A : K^n \to K^n$ diagonalisierbar ist.

Die Matrix $D = (d_{ij})$, die F bzgl. der Basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ beschreibt, gewinnt man bekanntlich aus der Gleichung

$$F(\mathbf{b}_j) = \sum_{i=1}^{n} d_{ij}\mathbf{b}_i, \quad j = 1, \dots, n.$$

D ist genau dann eine Diagonalmatrix, wenn $d_{ij} = 0$ für $i \neq j$ ist, also $F(\mathbf{b}_j) = d_{jj}\mathbf{b}_j$ für $j = 1, \dots, n$. Eine solche Matrix D kann man genau dann finden, wenn V eine Basis besitzt, die nur aus Eigenvektoren besteht.

Nun sei speziell $F = F_A : K^n \to K^n$, mit $A \in M_{n,n}(K)$.
Bezüglich der Basis $\{\vec{b}_1, \dots, \vec{b}_n\}$ werde F durch die Matrix D beschrieben. Fassen wir die Basisvektoren zu der regulären Matrix $B := (\vec{b}_1, \dots, \vec{b}_n)$ zusammen, so ist bekanntlich $D = B^{-1} \circ A \circ B$ (vgl. §3).
Damit haben wir:

I.6.2 Satz. *Folgende Aussagen über eine Matrix $A \in M_{n,n}(K)$ sind äquivalent:*

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

$$D = B^{-1} \circ A \circ B$$

ist.

Wie findet man Eigenvektoren zu einer Matrix A ?

Es ist einfacher, mit den Eigenwerten zu beginnen. Sei λ ein Eigenwert von A. Das bedeutet:

$$\begin{aligned} & \exists\ \vec{v} \neq \vec{0} \ \text{ mit } A \circ \vec{v} = \lambda \cdot \vec{v} \\ \iff & \exists\ \vec{v} \neq \vec{0} \ \text{ mit } A \circ \vec{v} - \lambda \cdot \vec{v} = \vec{0} \\ \iff & \exists\ \vec{v} \neq \vec{0} \ \text{ mit } (A - \lambda \cdot 1) \circ \vec{v} = \vec{0} \\ \iff & \operatorname{Ker}(A - \lambda \cdot 1) \neq \{\vec{0}\} \end{aligned}$$

$$\begin{aligned}
&\iff \operatorname{rg}(A - \lambda \cdot 1\!\!1) < n \\
&\iff \det(A - \lambda \cdot 1\!\!1) = 0 \\
&\iff \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0.
\end{aligned}$$

Wie sieht diese Determinante aus? Es ist

$$\begin{aligned}
\det(A - \lambda \cdot 1\!\!1) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \widetilde{a}_{\sigma(1),1} \cdot \ldots \cdot \widetilde{a}_{\sigma(n),n} \\
&\quad \left(\text{mit } \widetilde{a}_{ij} := \begin{cases} a_{ij} & \text{für } i \neq j \\ a_{ii} - \lambda & \text{für } i = j \end{cases} \right) \\
&= (a_{11} - \lambda) \cdot \ldots \cdot (a_{nn} - \lambda) + \text{Terme,} \\
&\qquad \text{bei denen } \lambda \text{ in höchstens } n-2 \text{ Faktoren vorkommt} \\
&= (-1)^n \lambda^n + (-1)^{n-1} (\sum_{i=1}^{n} a_{ii}) \lambda^{n-1} + \text{Polynom in } \lambda \text{ vom Grad} \leq n-2.
\end{aligned}$$

Definition.

$p_A(x) := \det(A - x \cdot 1\!\!1)$ heißt *charakteristisches Polynom* von A.

$p_A(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1} + c_n x^n$ ist ein Polynom vom Grad n über K, mit

$$\begin{aligned}
c_n &= (-1)^n, \\
c_{n-1} &= (-1)^{n-1} \cdot \sum_{i=1}^{n} a_{ii}, \\
&\vdots \\
c_0 &= \det(A).
\end{aligned}$$

Die Zahl $\operatorname{Spur}(A) := \sum_{i=1}^{n} a_{ii}$ heißt die *Spur von A*. Daß $c_0 = \det(A)$ ist, stellt man bei genauem Hinsehen fest:
Es ist nämlich

$$\widetilde{a}_{\sigma(1),1} \cdot \ldots \cdot \widetilde{a}_{\sigma(n),n} = a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n} + \text{Terme, die } \lambda \text{ enthalten.}$$

Das Entscheidende ist nun:

Die Eigenwerte von A sind die Nullstellen des charakteristischen Polynoms $p_A(x)$.

Das liefert eine praktische Berechnungsmöglichkeit.

Beispiele :

1. Wir betrachten noch einmal die Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Es gilt:

$$\det \begin{pmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4.$$

Also ist

$$p_A(\lambda) = 0 \iff \lambda = \frac{3 \pm \sqrt{9+16}}{2} = \begin{cases} 4 \\ -1 \end{cases}$$

Den einen dieser beiden Eigenwerte hatten wir schon kennengelernt. Nun suchen wir die zugehörigen Eigenvektoren. Dazu muß man bei gegebenem λ die Gleichung $A \circ \vec{x} = \lambda \cdot \vec{x}$ lösen:

$\lambda = -1$: $\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ hat z.B. $\vec{a}_1 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ als Lösung.

$\lambda = 4$: $\begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ hat $\vec{a}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ als Lösung.

$\vec{a}_1$ und $\vec{a}_2$ sind Eigenvektoren zu den Eigenwerten -1 bzw. 4. Da $\det \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = 5 \neq 0$ ist, bilden sie auch eine Basis.

Wir setzen $B := \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. Dann ist $B^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$ und

$$\begin{aligned} B^{-1} \circ A \circ B &= \frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \circ \begin{pmatrix} -1 & 8 \\ 1 & 12 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} -5 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Es kommt eine Diagonalmatrix heraus, und die Einträge darin sind gerade die Eigenwerte. Ob das Zufall ist oder ob es genau so sein muß, werden wir später untersuchen.

2. $R(\varphi)$ sei die Drehmatrix $R(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. Dann ist

$$\begin{aligned} \det(R(\varphi) - \lambda \cdot 1\!\!1) &= \det \begin{pmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{pmatrix} \\ &= (\cos\varphi - \lambda)^2 + \sin^2\varphi \\ &= \lambda^2 - 2\lambda\cos\varphi + 1. \end{aligned}$$

Dieser Ausdruck wird genau dann Null, wenn

$$\lambda = \frac{1}{2} \cdot (2\cos\varphi \pm \sqrt{4\cos^2\varphi - 4}) = \cos\varphi \pm \sqrt{-\sin^2\varphi}.$$

Das ist nur möglich, wenn $\varphi = 0, \pi, 2\pi$ ist, also $\sin\varphi = 0$. Das bedeutet, daß unter allen Drehmatrizen nur die speziellen Matrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \quad \text{und} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -1$$

Eigenwerte besitzen (und zwar $\lambda = 1$ bzw. $\lambda = -1$). Andere Drehungen besitzen keinen Eigenwert und daher auch keinen Eigenvektor.

Spaßeshalber kann man ja einmal nach *komplexen* Eigenwerten suchen. Auf Grund der obigen Berechnung ist sofort klar, daß $\lambda_{1,2} = e^{\pm \mathbf{j}\varphi}$ komplexe Eigenwerte sind.

Die Gleichung für einen Eigenvektor $\begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2$ zum Eigenwert $\lambda_1 = e^{\mathbf{j}\varphi}$ lautet:

$$\begin{pmatrix} -\mathbf{j}\sin\varphi & -\sin\varphi \\ \sin\varphi & -\mathbf{j}\sin\varphi \end{pmatrix} \circ \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Man sieht, daß $\begin{pmatrix} 1 \\ -\mathbf{j} \end{pmatrix}$ eine Lösung ist. Genauso ist $\begin{pmatrix} 1 \\ \mathbf{j} \end{pmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_2 = e^{-\mathbf{j}\varphi}$. Offensichtlich sind $\begin{pmatrix} 1 \\ \mathbf{j} \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -\mathbf{j} \end{pmatrix}$ linear unabhängig (auch über $\mathbb{C}$), es gibt also eine Basis von Eigenvektoren. Man sieht, daß eine Matrix über $\mathbb{R}$ nicht diagonalisierbar, aber über $\mathbb{C}$ diagonalisierbar sein kann.

I.6.3 Satz. *Sind $\vec{v}_1, \ldots, \vec{v}_k \in K^n$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$ der Matrix $A \in M_{n,n}(K)$, so sind sie linear unabhängig.*

BEWEIS: Dies ist Gelegenheit für einen Induktionsbeweis. Wir führen Induktion nach der Anzahl k.

$\underline{k = 1}$: *Ein* einzelner Eigenvektor $\vec{v}_1$ ist immer linear unabhängig, weil er ja $\neq \vec{0}$ ist.

$\underline{k - 1 \to k}$: Es seien Eigenvektoren $\vec{v}_1, \ldots, \vec{v}_k$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$ gegeben. Wir nehmen an, sie wären linear abhängig, und versuchen, daraus einen Widerspruch zu konstruieren:

Nach Induktionsvoraussetzung sind $\vec{v}_1, \ldots, \vec{v}_{k-1}$ linear unabhängig. Wenn sie durch Hinzunahme von $\vec{v}_k$ linear abhängig werden, muß es Koeffizienten $\alpha_j \in K$ geben, so daß gilt:

$$\vec{v}_k = \sum_{j=1}^{k-1} \alpha_j \, \vec{v}_j \, .$$

Daraus folgt:

$$\begin{aligned} \vec{0} = (A - \lambda_k 1) \circ \vec{v}_k &= (A - \lambda_k 1) \circ \sum_{j=1}^{k-1} \alpha_j \, \vec{v}_j \\ &= \sum_{j=1}^{k-1} \alpha_j \cdot (A - \lambda_k 1) \circ \vec{v}_j \\ &= \sum_{j=1}^{k-1} \alpha_j \cdot (\lambda_j - \lambda_k) \, \vec{v}_j \, . \end{aligned}$$

Da die $\lambda_j \neq \lambda_k$ sind, also $\lambda_j - \lambda_k \neq 0$ für $j = 1, \dots, k-1$, und da $\vec{v}_1, \dots, \vec{v}_{k-1}$ linear unabhängig sind, muß $\alpha_1 = \dots = \alpha_{k-1} = 0$ sein. Aber dann ist $\vec{v}_k = \vec{0}$, und das kann bei einem Eigenvektor nicht sein. WS! $\square$

I.6.4 Folgerung.
Hat $A \in M_{n,n}(K)$ n verschiedene Eigenwerte, so ist A diagonalisierbar.

Eine besondere Rolle spielt der Eigenwert 0:

I.6.5 Satz.
$A \in M_{n,n}(K)$ ist genau dann regulär, wenn 0 **kein** *Eigenwert von A ist.*

BEWEIS: Das charakteristische Polynom hat die Gestalt

$$p_A(x) = \sum_{i=0}^{n} c_i x^i.$$

Daher ist $p_A(0) = c_0 = \det(A)$, und es gilt:

$$\begin{aligned} A \quad \text{regulär} \quad &\iff \quad \det(A) \neq 0 \\ &\iff \quad p_A(0) \neq 0 \\ &\iff \quad 0 \text{ kein Eigenwert.} \end{aligned}$$

$\square$

Das charakteristische Polynom p_A einer Matrix $A \in M_{n,n}(K)$ hängt in Wirklichkeit nur von der linearen Abbildung $F = F_A$ ab, ganz egal, bezüglich welcher Basis man F beschreibt:

I.6.6 Satz. *Ist B invertierbar, so ist*

$$p_{B^{-1}AB}(x) = p_A(x) \text{ für alle } x \in \mathbb{R}.$$

BEWEIS: Es ist

$$\begin{aligned} p_{B^{-1}AB}(x) &= \det(B^{-1}AB - x\,1\,) \\ &= \det(B^{-1}AB - B^{-1}(x\,1\,)B) \\ &= \det(B^{-1} \circ (A - x\,1\,) \circ B) \\ &= (\det B)^{-1} \det(A - x\,1\,) \det B \\ &= \det(A - x\,1\,) \;=\; p_A(x). \end{aligned}$$

$\square$

I.6.7 Satz. *Besitzt $A \in M_{n,n}(K)$ n paarweise verschiedene Eigenwerte, so ist* $\dim(\text{Ker}\,(F_A)) \leq 1$, *also* $\text{rg}(A) \geq n - 1$.

BEWEIS: Sind $\lambda_1, \dots, \lambda_n$ die n verschiedenen Eigenwerte von A, so gibt es eine invertierbare Matrix B, so daß $D := B^{-1} \circ A \circ B$ eine Diagonalmatrix mit den λ_i auf der Diagonalen ist. Aber dann ist

$$p_A(x) = p_{B^{-1}AB}(x) = (-1)^n \cdot (x - \lambda_1) \cdot \ldots \cdot (x - \lambda_n).$$

Da höchstens ein Eigenwert $= 0$ sein kann, ist $\mathrm{rg}(A) = \mathrm{rg}(D) \geq n-1$, und daher auch $\dim(\mathrm{Ker}\,(F_A)) \leq 1$. $\square$

Leider kann man nicht immer erwarten, daß das charakteristische Polynom in lauter verschiedene Linearfaktoren zerfällt. Wir betrachten noch ein weiteres Beispiel:

Beispiel:

Sei $A := \begin{pmatrix} 6 & 4 & 4 \\ -1 & 0 & -1 \\ -2 & 0 & 0 \end{pmatrix}$. Dann gilt:

$$\begin{aligned} \det(A - \lambda 1) &= \det \begin{pmatrix} 6-\lambda & 4 & 4 \\ -1 & -\lambda & -1 \\ -2 & 0 & -\lambda \end{pmatrix} \\ &= (6-\lambda)\lambda^2 + 8 + 4 \cdot 0 - 8\lambda - 0 - 4\lambda \\ &= -\lambda^3 + 6\lambda^2 - 12\lambda + 8 \\ &= (2-\lambda)^3. \end{aligned}$$

Also ist $\lambda = 2$ die einzige Nullstelle von $p_A(x)$. Sie hat die Vielfachheit 3. Mehr Nullstellen kann es aus Gradgründen selbst im Komplexen nicht geben. Wie steht es nun mit den Eigenvektoren? Die zugehörige Gleichung hat die Form

$$\begin{pmatrix} 4 & 4 & 4 \\ -1 & -2 & -1 \\ -2 & 0 & -2 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir verwenden das schematisch Verfahren zum Lösen von Linearen Gleichungssystemen:

$$\begin{array}{ccc|cl} 4 & 4 & 4 & 0 & \\ -1 & -2 & -1 & 0 & \\ -2 & 0 & -2 & 0 & (+(-2)\times(2.\ \text{Zeile})\ ) \\ \hline 4 & 4 & 4 & 0 & \\ -1 & -2 & -1 & 0 & (+\frac{1}{4}\times(1.\ \text{Zeile})\ ) \\ 0 & 4 & 0 & 0 & \\ \hline 4 & 4 & 4 & 0 & \\ 0 & -1 & 0 & 0 & \\ 0 & 4 & 0 & 0 & (+4\times(2.\ \text{Zeile})\ ) \\ \hline 4 & 4 & 4 & 0 & \\ 0 & -1 & 0 & 0 & \\ 0 & 0 & 0 & 0. & \end{array}$$

Das resultierende Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 0, \\ x_2 &= 0 \end{aligned}$$

liefert als Lösungsraum den Eigenraum

$$E(2) = \{\alpha \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R}\}.$$

Der ist nur 1-dimensional.

Definition.

Sei $\lambda \in K$ ein Eigenwert der Matrix $A \in M_{n,n}(K)$.

$a(\lambda) :=$ Vielfachheit der Nullstelle λ des charakteristischen Polynoms $p_A(x)$ heißt *algebraische Vielfachheit von* λ.

$g(\lambda) := \dim_K(E(\lambda))$ heißt *geometrische Vielfachheit von* λ.

Das Verhalten im vorigen Beispiel war kein Zufall:

I.6.8 Satz. *Ist λ Eigenwert von $A \in M_{n,n}(K)$, so ist*

$$1 \le g(\lambda) \le a(\lambda) \le n.$$

Bevor wir diesen Satz beweisen können, müssen wir noch ein anderes Ergebnis nachtragen, den „Basisergänzungssatz“:

I.6.9 Satz. *Sei $V \subset K^n$ ein Untervektorraum, $d := \dim_K(V)$ und $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ ein System linear unabhängiger Vektoren von V. Dann kann dieses System zu einer Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{a}_{k+1}, \dots, \mathbf{a}_d\}$ ergänzt werden.*

Beweis: Ist $k < d$, so gibt es einen Vektor $\mathbf{a}_{k+1} \in V$, so daß $\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{a}_{k+1}$ linear unabhängig sind, denn andernfalls wäre $< \mathbf{a}_1, \dots, \mathbf{a}_k > = V$.

Mit dem um einen Vektor erweiterten System fährt man fort. Da es in V höchstens d linear unabhängige Elemente geben kann, bricht das Verfahren bei d Vektoren ab. Dann hat man das Ausgangssystem zu einer Basis ergänzt. $\square$

Nun können wir uns dem Vergleich der algebraischen und geometrischen Vielfachheit widmen:

Beweis: Sei $k := g(\lambda)$.
Dann gibt es eine Basis $\{\vec{v}_1, \dots, \vec{v}_k\}$ von $E(\lambda)$. Wir ergänzen diese Basis zu einer Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ des ganzen K^n. Dann ist $B := (\vec{v}_1, \dots, \vec{v}_n)$ eine reguläre Matrix, und es gilt:

$$\begin{aligned}
B^{-1} \circ A \circ B &= B^{-1} \circ (A \circ (\vec{v}_1, \dots, \vec{v}_k), A \circ (\vec{v}_{k+1}, \dots, \vec{v}_n)) \\
&= B^{-1} \circ (\lambda\, \vec{v}_1, \dots, \lambda\, \vec{v}_k, A \circ (\vec{v}_{k+1}, \dots, \vec{v}_n)) \\
&= (\lambda \cdot B^{-1} \circ (\vec{v}_1, \dots, \vec{v}_k), B^{-1} \circ A \circ (\vec{v}_{k+1}, \dots, \vec{v}_n)) \\
&= \left(\begin{array}{ccc|c} \lambda & \cdots & 0 & \\ \vdots & \ddots & \vdots & X \\ 0 & \cdots & \lambda & \\ \hline & 0 & & Y \end{array}\right),
\end{aligned}$$

mit irgendwelchen Matrizen $X \in M_{k,n-k}(K)$ und $Y \in M_{n-k,n-k}(K)$. Daraus folgt:

$$\begin{aligned} p_A(x) &= p_{B^{-1}AB}(x) \\ &= \det \left(\begin{array}{ccc|c} \lambda - x & \cdots & 0 & \\ \vdots & \ddots & \vdots & X \\ 0 & \cdots & \lambda - x & \\ \hline & 0 & & Y - x\,\mathbb{1} \end{array} \right) \\ &= (\lambda - x)^k \cdot \det(Y - x \cdot \mathbb{1}). \end{aligned}$$

Also ist λ eine Nullstelle von mindestens k-ter Ordnung, es ist $a(\lambda) \geq g(\lambda)$.
Da $p_a(x)$ höchstens n Nullstellen haben kann, ist $a(\lambda) \leq n$, und da λ nur dann Eigenwert ist, wenn es dazu mindestens einen Eigenvektor gibt, ist $g(\lambda) \geq 1$. □

I.6.10 Folgerung. *Sei $A \in M_{n,n}(K)$. A ist genau dann diagonalisierbar, wenn gilt:*

1. *$p_A(x)$ zerfällt vollständig in Linearfaktoren.*
2. *Für jeden Eigenwert λ ist $a(\lambda) = g(\lambda)$.*

Beweis: 1) Ist A diagonalisierbar, so gibt es eine Basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ von Eigenvektoren. Wir können die Basisvektoren so anordnen, daß gilt:
$\vec{v}_1, \ldots, \vec{v}_{k_1}$ sind Eigenvektoren zum Eigenwert λ_1,
$\vec{v}_{k_1+1}, \ldots, \vec{v}_{k_1+k_2}$ sind Eigenvektoren zum Eigenwert λ_2,
$\vdots$
$\vec{v}_{k_1+\cdots+k_{r-1}+1}, \ldots, \vec{v}_n$ sind Eigenvektoren zum Eigenwert λ_r.
Die Eigenwerte $\lambda_1, \ldots, \lambda_r$ seien paarweise verschieden. Wir bilden aus ihnen die Diagonalmatrix

$$\Delta := \begin{pmatrix} \lambda_1 & 0 & & & & & & \cdots & & 0 \\ 0 & \ddots & & & & & & & & \\ & & \lambda_1 & & & & & & & \\ & & & \lambda_2 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \lambda_2 & & & & \vdots \\ \vdots & & & & & & \ddots & & & \\ & & & & & & & \lambda_r & & \\ & & & & & & & & \ddots & 0 \\ 0 & & & & \cdots & & & & 0 & \lambda_r \end{pmatrix}.$$

Es gibt eine reguläre Matrix B, so daß $B^{-1}AB = \Delta$ ist. Daher gilt:

$$p_A(x) = p_\Delta(x) = (\lambda_1 - x)^{k_1} \cdot \ldots \cdot (\lambda_r - x)^{k_r}.$$

Das charakteristische Polynom zerfällt in Linearfaktoren, und es ist $a(\lambda_i) = k_i = g(\lambda_i)$ für $i = 1, \ldots, r$.

2) Sei umgekehrt das Kriterium erfüllt, $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenwerte mit den Vielfachheiten $k_i = a(\lambda_i) = g(\lambda_i)$ für $i = 1, \ldots, r$. Dann folgt:

$$\begin{aligned} \dim_K(E(\lambda_1)) + \cdots + \dim_K(E(\lambda_r)) &= g(\lambda_1) + \cdots + g(\lambda_r) \\ &= a(\lambda_1) + \cdots + a(\lambda_r) \\ &= \deg(p_A(x)) = n. \end{aligned}$$

Da Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, folgt jetzt die Existenz einer Basis von Eigenvektoren. □

Bemerkungen :

1. Über $\mathbb{C}$ zerfällt jedes Polynom in Linearfaktoren. Da ist die erste Bedingung überflüssig.

2. Mit diagonalisierbaren Matrizen kann man gut rechnen:

   Ist $B^{-1}AB = D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix, so ist z.B.

   $$B^{-1}A^pB = \underbrace{(B^{-1}AB) \circ \ldots \circ (B^{-1}AB)}_{\text{p-mal}} = D^p.$$

   Aber die Potenzen einer Diagonalmatrix kann man leicht ausrechnen. Daraus ergibt sich:

   $$A^p = B \circ \begin{pmatrix} \lambda_1^p & & 0 \\ & \ddots & \\ 0 & & \lambda_n^p \end{pmatrix} \circ B^{-1}.$$

   Auf diese Weise kann man Polynome von (diagonalisierbaren) Matrizen bilden. Ist $f(x) = \sum_{i=1}^{n} a_i x^i$, so setzt man

   $$f(A) := a_0 + a_1 \cdot A + a_2 \cdot A^2 + \cdots + a_n \cdot A^n.$$

   Insbesondere ist

   $$p_A(A) = B \circ p_A(D) \circ B^{-1} = B \circ \begin{pmatrix} p_A(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & p_A(\lambda_n) \end{pmatrix} \circ B^{-1} = 0.$$

   Das ist ein Spezialfall des Satzes von Cayley-Hamilton, in Wirklichkeit gilt die Gleichung $p_A(A) = 0$ für jede quadratische Matrix. Der Beweis dafür ist schwierig, er benutzt die Jordansche Normalform. Und man muß sich vor folgendem „Kurzschluß“ hüten:

   $$p_A(A) = \det(A - A \cdot \mathbf{1}) = \det(A - A) = \det(0) = 0.$$

   Das ist natürlich völliger Quatsch!! (Warum?)

Wir wollen uns jetzt mit den Eigenwerten spezieller Matrizen befassen:

I.6.11 Satz. *Sei $A \in M_{n,n}(K)$ orthogonal (im Falle $K = \mathbb{R}$) oder unitär (im Falle $K = \mathbb{C}$). Dann gilt:*

1. *Ist $\lambda \in K$ ein Eigenwert von A, so ist $|\lambda| = 1$.*
2. *Sind λ, μ zwei verschiedene Eigenwerte von A und $\vec{x}$ bzw. $\vec{y}$ jeweils Eigenvektoren dazu, so sind $\vec{x}$ und $\vec{y}$ zueinander orthogonal.*

BEWEIS: Wir betrachten nur den komplexen Fall, und wir arbeiten mit der linearen Abbildung $F = F_A$ statt mit der Matrix A.

1) Sei $\vec{x} \neq \vec{0}$, $F(\vec{x}) = \lambda \cdot \vec{x}$. Da F eine Isometrie ist, gilt:

$$\lambda\overline{\lambda} <\vec{x}, \vec{x}> = < \lambda \vec{x}, \lambda \vec{x}> = < F(\vec{x}), F(\vec{x}) > = <\vec{x}, \vec{x}> .$$

Daraus folgt, daß $|\lambda|^2 = 1$ ist.

2) Ist $F(\vec{x}) = \lambda \vec{x}$ und $F(\vec{y}) = \mu \vec{y}$, mit $\lambda \neq \mu$, so folgt:

$$\lambda\overline{\mu} <\vec{x}, \vec{y}> = < F(\vec{x}), F(\vec{y}) > = <\vec{x}, \vec{y}> .$$

Also ist $(\lambda\overline{\mu} - 1) \cdot <\vec{x}, \vec{y}> = 0$. Wäre $\lambda\overline{\mu} = 1$, so wäre $\overline{\mu} = \lambda^{-1} = \overline{\lambda}$ (weil $|\lambda| = 1$ ist). Aber dann wäre auch $\lambda = \mu$, im Widerspruch zur Voraussetzung. Also muß $<\vec{x}, \vec{y}> = 0$ sein. □

Bevor wir die Frage der Diagonalisierbarkeit erörtern, wollen wir noch den zweiten im Zusammenhang mit Skalarprodukten interessanten Typ von Matrizen betrachten:

I.6.12 Satz. *Sei $A \in M_{n,n}(K)$ symmetrisch (im Falle $K = \mathbb{R}$) oder hermitesch (im Falle $K = \mathbb{C}$). Dann gilt:*

1. *Das charakteristische Polynom von A zerfällt in Linearfaktoren.*
2. *Alle Eigenwerte von A sind reell.*
3. *Eigenvektoren zu verschiedenen Eigenwerten von A sind zueinander orthogonal.*

BEWEIS: Eine reelle symmetrische Matrix ist natürlich auch hermitesch, und über $\mathbb{C}$ zerfällt das charakteristische Polynom in Linearfaktoren. Wir brauchen also nur zu zeigen, daß alle Eigenwerte einer hermiteschen Matrix A reell sind.

Wieder arbeiten wir mit der zugehörigen linearen Abbildung $F = F_A$. Sie ist in diesem Fall selbstadjungiert.

Ist $\lambda \in \mathbb{C}$ Eigenwert von F, so gibt es einen Vektor $\vec{z} \in \mathbb{C}^n$ mit $\vec{z} \neq \vec{0}$, so daß $F(\vec{z}) = \lambda \cdot \vec{z}$ ist. Dann ist $<\vec{z}, \vec{z}>$ reell und > 0. Außerdem gilt:

$$\begin{aligned} \lambda <\vec{z}, \vec{z}> &= < \lambda \vec{z}, \vec{z}> = < F(\vec{z}), \vec{z}> \\ &= <\vec{z}, F(\vec{z}) > \\ &= <\vec{z}, \lambda \cdot \vec{z}> = \overline{\lambda} <\vec{z}, \vec{z}> . \end{aligned}$$

Es ist klar, daß dann $\lambda = \overline{\lambda}$ sein muß, also $\lambda \in \mathbb{R}$.

Zur Orthogonalität der Eigenvektoren:
Seien $\lambda \neq \mu$ zwei Eigenwerte, $\vec{z}$ bzw. $\vec{w}$ zugehörige Eigenvektoren. Dann gilt:

$$\begin{aligned} \lambda \cdot <\vec{z}, \vec{w}> &= <\lambda\, \vec{z}, \vec{w}> \\ &= <F(\vec{z}), \vec{w}> \\ &= <\vec{z}, F(\vec{w})> \\ &= <\vec{z}, \mu\, \vec{w}> \\ &= \mu \cdot <\vec{z}, \vec{w}> \text{ (weil } \mu \text{ reell).} \end{aligned}$$

Also ist $(\lambda - \mu) \cdot <\vec{z}, \vec{w}> = 0$, d.h. $<\vec{z}, \vec{w}> = 0$. □

Jetzt untersuchen wir die Diagonalisierbarkeit. Dabei wollen wir induktiv oder rekursiv vorgehen. Ist $F : K^n \to K^n$ linear, so wird ein Untervektorraum $U \subset K^n$ *F-invariant* genannt, wenn $F(U) \subset U$ ist.

Wählt man nun eine Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ von U und ergänzt sie zu einer Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n\}$ von K^n, so läßt sich F folgendermaßen beschreiben:

$$F(\mathbf{a}_j) = \sum_{i=1}^{n} \alpha_{ij} \mathbf{a}_i,\ j = 1, \ldots, n,$$

mit $\alpha_{ij} = 0$ für $j \leq k$ und $i > k$. Dann hat die Matrix $A := (\alpha_{ij})$ folgende Gestalt:

$$A = \left(\begin{array}{c|c} A' & A'' \\ \hline 0 & A''' \end{array} \right).$$

Es leuchtet also ein, daß sukzessives Aufsuchen von invarianten Unterräumen zumindest auf eine Matrix in oberer Dreiecksgestalt führt. Bei den speziellen linearen Abbildungen, die wir hier untersuchen wollen, gilt aber noch mehr:

I.6.13 Hilfssatz. *Sei $F : K^n \to K^n$ isometrisch oder selbstadjungiert. Ist $U \subset K^n$ ein F–invarianter Untervektorraum, so ist auch $U^\perp$ F–invariant.*

Beweis: Es sei $F(U) \subset U$. Wir müssen zeigen, daß auch $F(U^\perp) \subset U^\perp$ gilt.

Den Beweis dazu müssen wir in zwei Teilen ausführen:

a) Zunächst sei F selbstadjungiert. Dann ist

$$<F(\vec{z}), \vec{w}> = <\vec{z}, F(\vec{w})> \qquad \text{für alle } \vec{z}, \vec{w} \in K^n.$$

Ist nun $\vec{z} \in U$ und $\vec{w} \in U^\perp$, so gilt:

$$<\vec{z}, F(\vec{w})> = <F(\vec{z}), \vec{w}> = 0.$$

b) Ist F isometrisch, so ist F ein Isomorphismus und $<F(\vec{z}), F(\vec{w})> = <\vec{z}, \vec{w}>$ für alle $\vec{z}, \vec{w} \in K^n$.

Ist $\vec{z} \in U$, so gibt es ein $\vec{z}\,' \in U$ mit $F(\vec{z}\,') = \vec{z}$. Ist weiter $\vec{w} \in U^\perp$, so gilt:

$$<\vec{z}, F(\vec{w})> = <F(\vec{z}\,'), F(\vec{w})> = <\vec{z}\,', \vec{w}> = 0.$$

Damit ist der Hilfssatz bewiesen. □

I.6.14 Satz. *Sei $A \in M_{n,n}(K)$ symmetrisch, hermitesch, orthogonal oder unitär. Im orthogonalen Fall sei zusätzlich vorausgesetzt, daß $p_A(x)$ in Linearfaktoren zerfällt.*

Dann gibt es in K^n eine ON-Basis von Eigenvektoren von A.

BEWEIS: Sei $F := F_A : K^n \to K^n$. Wir behandeln den reellen und den komplexen Fall simultan und führen Induktion nach n:

$\underline{n = 1}$: Ist $a := F(1) \in K$, so ist $F(1) = a \cdot 1$, also $\{1\}$ eine ON-Basis aus Eigenvektoren, zum Eigenwert a.

$\underline{n - 1 \to n}$: Auf Grund der Voraussetzungen zerfällt $p_A(x)$ in Linearfaktoren. Es gibt also einen Eigenwert λ und dazu einen Eigenvektor $\vec{v}$. O.B.d.A. sei $\| \vec{v} \| = 1$. Wir setzen $U := K\cdot \vec{v}$. Offensichtlich gilt $F(U) \subset U$.

Nach dem Hilfssatz ist $F(U^\perp) \subset U^\perp$.

Die Abbildung $F' := F \big|_{U^\perp}: U^\perp \to U^\perp$ bleibt selbstadjungiert bzw. isometrisch. Da es in $U^\perp$ eine ON-Basis gibt, können wir $U^\perp$ mit dem K^{n-1} identifizieren und darauf die Induktionsvoraussetzung anwenden. Es gibt eine ON-Basis von Eigenvektoren von F' (die dann auch Eigenvektoren von F sind) in $U^\perp$. Zusammen mit dem Vektor $\vec{v}$ ergibt das eine ON-Basis von Eigenvektoren von F in K^n. $\square$

Als Folgerung haben wir:

I.6.15 Satz von der Hauptachsentransformation.
Ist $A \in M_{n,n}(K)$ eine symmetrische (bzw. hermitesche) Matrix, so gibt es eine orthogonale (bzw. unitäre) Matrix $S \in \mathrm{GL}(n, K)$, so daß $S^{-1} \circ A \circ S$ eine (reelle) Diagonalmatrix ist. Die Einträge in der Diagonalmatrix sind die Eigenwerte von A.

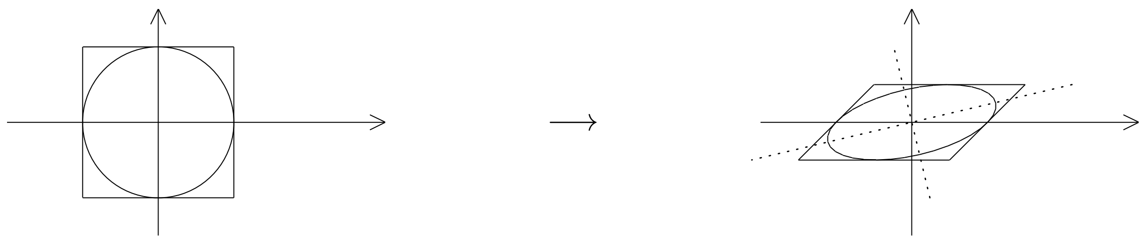
BEWEIS: Es gibt eine ON-Basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ aus Eigenvektoren von A. Dann ist $S := (\vec{v}_1, \ldots, \vec{v}_n)$ unitär, und bezüglich der ON-Basis wird $F := F_A$ durch die Matrix $S^{-1}\circ A\circ S$ beschrieben. Es ist klar, daß dies eine Diagonalmatrix ist, in der die Eigenwerte von A stehen. $\square$

Um den Begriff „Hauptachsentransformation“ zu erläutern, sei noch ein Beispiel angegeben:

Beispiel:

Sei $M := \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}^{\top} \circ \vec{x} = 1\} = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid (x_1)^2 + (x_2)^2 = 1\}$ der Einheitskreis im $\mathbb{R}^2$. Wir wenden die folgende lineare Transformation auf den $\mathbb{R}^2$ an:

Es sei $A := \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$ und $F(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) := A \circ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + \frac{1}{2}x_2 \\ \frac{1}{2}x_2 \end{pmatrix}$. Diese Abbildung verwandelt das Quadrat mit den Ecken $(1, 1)$, $(-1, 1)$, $(-1, -1)$ und $(1, -1)$ in ein Parallelogramm mit den Ecken $(\frac{3}{2}, \frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2})$, $(-\frac{3}{2}, -\frac{1}{2})$ und $(\frac{1}{2}, -\frac{1}{2})$. Dadurch werden Figuren in der Ebene in einer einfachen Parallelperspektive dargestellt. Insbesondere wird aus dem Einheitskreis eine Ellipse:



Wir können diese Ellipse genau angeben. Es ist

$$\begin{aligned} F(M) &= \{\vec{y} = F(\vec{x}) \mid \vec{x}^{\top} \circ \vec{x} = 1\} \\ &= \{\vec{y} \mid F^{-1}(\vec{y})^{\top} \circ F^{-1}(\vec{y}) = 1\} \\ &= \{\vec{y} \mid \vec{y}^{\top} \circ ((A^{-1})^{\top} \circ A^{-1}) \circ \vec{y} = 1\}. \end{aligned}$$

Da $\det(A) = \frac{1}{2}$ ist, gilt:

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

und damit

$$(A^{-1})^{\top} \circ A^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}.$$

Für $\vec{y} = (y_1, y_2)^{\top}$ folgt daraus:

$$\begin{aligned} \vec{y}^{\top} \circ ((A^{-1})^{\top} \circ A^{-1}) \circ \vec{y} &= (y_1, y_2) \circ \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix} \circ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= (y_1, y_2) \circ \begin{pmatrix} y_1 - y_2 \\ 5y_2 - y_1 \end{pmatrix} \\ &= y_1^2 - 2y_1y_2 + 5y_2^2. \end{aligned}$$

Mit Hilfe der symmetrischen Matrix $B := (A^{-1})^{\top} \circ A^{-1}$ wird die „quadratische Form“

$$q(\vec{y}) := \vec{y}^{\top} \circ B \circ \vec{y} = y_1^2 - 2y_1y_2 + 5y_2^2$$

beschrieben. Das Bild des Einheitskreises unter F ist die Ellipse

$$E := \{\vec{y} \mid q(\vec{y}) = 1\} = \{\vec{y} \mid y_1^2 - 2y_1y_2 + 5y_2^2 = 1\}.$$

Will man eine solche Ellipse zeichnen, so ist es nützlich, ihre Hauptachsen zu kennen (Manche Computer-Programme können Ellipsen mit Hilfe ihrer Hauptachsen zeichnen). Also führen wir eine Hauptachsentransformation durch:

Es ist

$$\begin{aligned} \det(B - \lambda \mathbf{1}) &= \det \begin{pmatrix} 1-\lambda & -1 \\ -1 & 5-\lambda \end{pmatrix} \\ &= (1-\lambda)(5-\lambda) - 1 \\ &= \lambda^2 - 6\lambda + 4. \end{aligned}$$

$$\text{und} \qquad \lambda^2 - 6\lambda + 4 = 0 \iff \lambda = \frac{6 \pm \sqrt{36-16}}{2} = 3 \pm \sqrt{5}.$$

Es gibt also zwei verschiedene (natürlich reelle) Eigenwerte

$$\lambda_1 = 3 + \sqrt{5} \quad \text{und} \quad \lambda_2 = 3 - \sqrt{5}.$$

Dazu müssen wir noch Eigenvektoren bestimmen. Das geschieht – für $\lambda \in \{\lambda_1, \lambda_2\}$ – durch Lösen der Gleichung $(B - \lambda \mathbb{1}) \circ \vec{x} = \vec{0}$.

1. $\lambda_1 = 3 + \sqrt{5}$.

$$\begin{pmatrix} -2-\sqrt{5} & -1 \\ -1 & 2-\sqrt{5} \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

liefert die Bedingung $x_1 = (2-\sqrt{5})x_2$. Also ist

$$\vec{v}_1 = \begin{pmatrix} 2-\sqrt{5} \\ 1 \end{pmatrix} \approx \begin{pmatrix} -0.24 \\ 1 \end{pmatrix}$$

ein passender Eigenvektor.

2. $\lambda_2 = 3 - \sqrt{5}$.

$$\begin{pmatrix} -2+\sqrt{5} & -1 \\ -1 & 2+\sqrt{5} \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

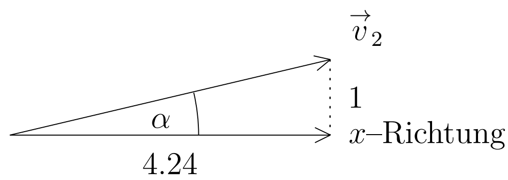
führt auf die Bedingung $x_1 = (2+\sqrt{5})x_2$ und damit auf den Eigenvektor

$$\vec{v}_2 = \begin{pmatrix} 2+\sqrt{5} \\ 1 \end{pmatrix} \approx \begin{pmatrix} 4.24 \\ 1 \end{pmatrix}.$$

Damit stehen die Richtungen der Hauptachsen fest. Offensichtlich stehen sie aufeinander senkrecht: Es ist $\vec{v}_1 \bullet \vec{v}_2 = (2-\sqrt{5})(2+\sqrt{5}) + 1 = 0$.

Eine Drehung des Standardkoordinatensystems um einen noch zu ermittelnden Winkel α führt die Einheitsvektoren in die Richtungen von $\vec{v}_2$ und $\vec{v}_1$ über. Dabei ist

$$\sin(\alpha) \approx \frac{1}{\sqrt{1^2 + (4.24)^2}} \approx 0.23,$$



also $\alpha \approx 13°$.

Als Gleichung in den neuen Koordinaten (bezogen auf $\{\vec{v}_2, \vec{v}_1\}$) erhält man für E:

$$\lambda_2 x_1^2 + \lambda_1 x_2^2 = 1.$$

Normalerweise hat die Gleichung einer Ellipse (mit dem Mittelpunkt im Nullpunkt) die Form

$$\left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 = 1.$$

Also erhalten wir

$$a = \frac{1}{\sqrt{\lambda_2}} \approx 1.14 \quad \text{und} \quad b = \frac{1}{\sqrt{\lambda_1}} \approx 0.44.$$