

# Mathematik für Informatiker I

Andreas Griewank

(griewank@math.hu-berlin.de)

Vertretung :

Jürgen Geiser (geiser@mathematik.hu-berlin.de)

Wiss. Mitarbeiter:

Hans-Dietrich Niepage (niepage@mathematik.hu-berlin.de)

Holger Heitsch (heitsch@mathematik.hu-berlin.de)

Lutz Lehmann (llehmann@mathematik.hu-berlin.de)

Stefan Körkel (stefan.koerkel@mathematik.hu-berlin.de)

Institut für Angewandte Mathematik  
Humboldt Universität zu Berlin

30. Januar 2007







# Teil B

## Lineare Algebra

### Vorläufige Gliederung

#### Teil A Algebraische Grundstrukturen

- (i) Algebraische Grundlagen
- (ii) Algebraische Teilstrukturen
- (iii) Algebraische Erweiterungen
- (iv) Äquivalenzrelationen und Quotientenstrukturen
- (v) Euklidische Ringe
- (vi) Polynome
- (vii) ...

#### Teil B Lineare Algebra

- (i) Vektorraum
- (ii) Analytische Geometrie im  $\mathbb{R}^3$
- (iii) Basis und Dimension

# B - 1 Einführung

Der Grundbegriff der linearen Algebra ist der des **Vektorraumes**, mit dessen Hilfe sich eine Vielzahl von mathematischen Objekten und Anwendungsmodellen beschreiben läßt.

## Definition B-1.1 (Vektorraum $\mathcal{V}$ bzw. linearer Raum)

Ein  $\mathcal{K}$ -Vektorraum ist eine Menge  $\mathcal{V}$

- (i) mit Addition  $+: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$  und
- (ii) skalarer Multiplikation  $\bullet: \mathcal{K} \times \mathcal{V} \rightarrow \mathcal{V}$  mit Elementen des Skalarkörpers  $\mathcal{K}$   
( $\mathcal{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_7, \dots$ )

Dabei ist  $(\mathcal{V}, +)$  eine additive Gruppe, d.h.  $\exists \mathbf{0} \in \mathcal{V}$  &  $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$ :

- ▶  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ , **Assoziativität**
- ▶  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ , **Kommutativität**
- ▶  $\mathbf{u} + \mathbf{0} = \mathbf{u}$ , **Neutrales Element**
- ▶  $\exists (-\mathbf{u}) \in \mathcal{V} : \mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ . **Inverses Element**

Die skalare Multiplikation muss mit der Multiplikation in  $\mathcal{K}$  verträglich sein,  
d.h.  $\forall \lambda, \gamma \in \mathcal{K}$ :

- ▶  $\lambda \bullet (\gamma \bullet \mathbf{u}) = (\lambda \gamma) \bullet \mathbf{u}$  sowie
- ▶  $1 \bullet \mathbf{u} = \mathbf{u}$ ,
- ▶  $\lambda \bullet (\mathbf{u} + \mathbf{v}) = \lambda \bullet \mathbf{u} + \lambda \bullet \mathbf{v}$ ,
- ▶  $(\lambda + \gamma) \bullet \mathbf{u} = \lambda \bullet \mathbf{u} + \gamma \bullet \mathbf{u}$ .

## Beispiel B - 1.2 (Spaltenvektorraum)

Für jedes  $n \in \mathbb{N}_+$  ist der Spaltenvektorenraum  $\mathbb{R}^{n \times 1}$  mit komponentenweisen Operationen auch ein  $\mathbb{R}$ -Vektorraum.

Ein Spaltenvektor  $\mathbf{v}$  aus  $n$  reellen Zahlen  $\nu_i$ ,  $i = 1, \dots, n$ , wird notiert als

$$\mathbf{v} = (\nu_1, \nu_2, \dots, \nu_n)^T \quad \text{oder} \quad \mathbf{v} = (\nu_i)_{i=1}^n .$$

## Beispiel B - 1.3 (Funktionsräume)

Für je zwei reellwertige Funktionen  $f, g : \mathcal{D} \rightarrow \mathbb{R}$  mit gemeinsamen Definitionsbereich  $\mathcal{D}$  kann man die Summe  $h = f + g$  als die Funktion mit den Werten

$$h(x) = f(x) + g(x) \quad \text{für} \quad x \in \mathcal{D}$$

definieren. Entsprechend erhält man  $h = \lambda \bullet f$  als die Funktion mit den Werten

$$h(x) = \lambda f(x) \quad \text{für} \quad x \in \mathcal{D} .$$



## Beispiel B - 1.4 (Zahlenfolgen)

Funktionen des Typs  $a : \mathbb{N} \rightarrow \mathbb{R}$  nennt man **Zahlenfolgen** und notiert diese als  $a = (a_n)_{n \in \mathbb{N}}$ ,  $b = (b_n)_{n \in \mathbb{N}}$ .

Addition und Multiplikation mit Skalaren werden gliedweise durchgeführt,

$$a + b = (a_n + b_n)_{n \in \mathbb{N}} \quad \text{und} \quad \lambda \bullet a = (\lambda a_n)_{n \in \mathbb{N}}.$$

Die Zahlenfolgen mit diesen Operationen bilden den  $\mathbb{R}$ -Vektorraum  $\ell(\mathbb{N}, \mathbb{R})$ .

Man kann für jedes  $k \in \mathbb{N}$  die eingliedrige Folge  $\delta^k = (\delta^k_n)_{n \in \mathbb{N}}$  mit

$$\delta^k_n = \begin{cases} 1 & \text{für } k = n \\ 0 & \text{sonst bei } k \neq n \end{cases}$$

definieren.

Mittels Addition und skalarer Multiplikation lassen sich aus diesen elementaren Folgen **immer nur** Folgen  $c = (c_n)_{n \in \mathbb{N}}$  endlicher Länge konstruieren, d.h. es gibt immer ein  $K \in \mathbb{N}$  mit  $c_n = 0$  für  $n > K$ .

Die meisten Untersuchungen und Ergebnisse der linearen Algebra beschäftigen sich mit Variationen der folgende Frage:

## Problem B-1.5

*Gegeben seien eine Familie von  $r$  Vektoren  $\mathbf{v}_i$  ( $i = 1, \dots, r$ ) und ein spezieller Vektor  $\mathbf{w}$  aus einem gemeinsamen Vektorraum  $\mathcal{V}$ .*

*Gibt es nun eine Familie von Skalaren  $\lambda_i$  ( $i = 1, \dots, r$ ), so dass*

$$\mathbf{w} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_r \mathbf{v}_r = \sum_{i=1}^r \lambda_i \mathbf{v}_i$$

*gilt? Wenn ja, wie kann man geeignete Koeffizienten  $\lambda_i$  berechnen? Sind diese dann eindeutig bestimmt?*

Unter anderem lassen sich Fragen nach Basisdarstellungen sowie die Suche nach den Lösungen linearer Gleichungssysteme

$$b_1 = a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n$$

$$b_2 = a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n$$

$$\vdots$$

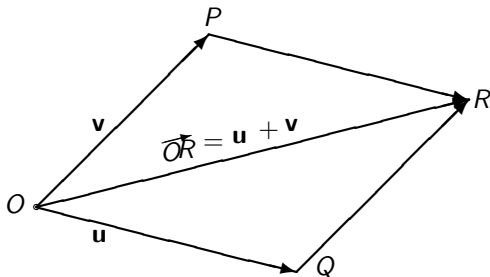
$$b_m = a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n$$

in dieser Art formulieren.

## B-2 Vektoren im Anschauungsraum

Man kann Verschiebungen im Raum als Vektoren auffassen. Vektoren  $\mathbf{v} = \overrightarrow{OP}$  und  $\mathbf{u} = \overrightarrow{OQ}$  sind dann diejenigen Verschiebungen, die den Ursprungspunkt  $O$  nach  $P$  bzw.  $Q$  verschieben. Die Summe zweier Verschiebungen ergibt sich als  $\mathbf{w} = \mathbf{u} + \mathbf{v} = \overrightarrow{OR}$ , wobei sich der Punkt  $R$  sowohl als Verschiebung von  $Q$  um  $\mathbf{v}$  als auch als Verschiebung von  $P$  um  $\mathbf{u}$  ergibt.

Bildlich liegt  $R$  an der Spitze des Pfeils  $\mathbf{v}$  (bzw.  $\mathbf{u}$ ), wenn dessen Anfang in den Punkt  $Q$  (bzw.  $P$ ) gelegt wird. Die Beliebigkeit in der Reihenfolge der Ausführung der Verschiebungen ist gleichbedeutend mit der Kommutativität der Vektoraddition.



Legt man den Ursprungspunkt  $O$  fest, so lassen sich alle **Raumpunkte**  $P$  mit ihren sogenannten **Ortsvektoren**  $\mathbf{v} = \overrightarrow{OP}$  identifizieren und man schreibt auch  $P = P(\mathbf{v})$ .

Vereinbart man weiterhin ein System von *drei rechtwinkligen Koordinatenachsen* mit geeigneter Skalierung, so lässt sich jeder Vektor  $\mathbf{v}$  mit Hilfe von drei Koordinaten  $\nu_1, \nu_2, \nu_3 \in \mathbb{R}$  wie folgt darstellen:

$$\mathbf{v} = \nu_1 \mathbf{e}_1 + \nu_2 \mathbf{e}_2 + \nu_3 \mathbf{e}_3$$

Hierbei verlaufen die drei **Einheitsvektoren**  $\mathbf{e}_1$ ,  $\mathbf{e}_2$  und  $\mathbf{e}_3$  entlang der  $x$ -,  $y$ - bzw.  $z$ -Achse. Sie bilden eine sogenannte **Basis** des Anschauungsraumes und werden zuweilen auch mit  $\vec{i}$ ,  $\vec{j}$  und  $\vec{k}$  bezeichnet. Hat man sich auf ein bestimmtes Koordinatensystem festgelegt, so kann man die Vektoren mit ihren entsprechenden **Koordinatentripeln** identifizieren und schreibt dann einfach

$$\mathbf{v} = (\nu_1, \nu_2, \nu_3)^T \in \mathbb{R}^3.$$

Inbesondere erhält man die **Basisvektoren** selbst als **kanonische Basisvektoren** des  $\mathbb{R}^3$ ,

$$\mathbf{e}_1 = (1, 0, 0)^T, \quad \mathbf{e}_2 = (0, 1, 0)^T, \quad \mathbf{e}_3 = (0, 0, 1)^T.$$

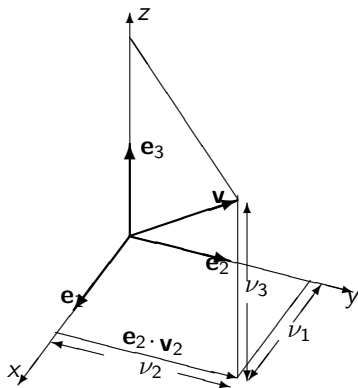
Addition, Subtraktion und Multiplikation erfolgen nun komponentenweise, z.B. für

$$\mathbf{u} = (3, -1, 2)^T \quad \text{und} \quad \mathbf{v} = (0, 2, 4)^T$$

ergibt sich

$$\mathbf{u} + \mathbf{v} = (3, 1, 6)^T, \quad \mathbf{u} - \mathbf{v} = (3, -3, -2)^T \quad \text{und} \quad 3\mathbf{u} = (9, -3, 6)^T,$$

wobei der Faktor 3 in der letzten Gleichung die Rolle eines Skalars spielt.



## Länge und Richtungskosinus

Wegen der vorausgesetzten Rechtwinkligkeit der Koordinatenachsen ergibt sich aus dem Satz des Pythagoras

### Definition B-2.1 (Länge eines Vektors, Euklidische Norm)

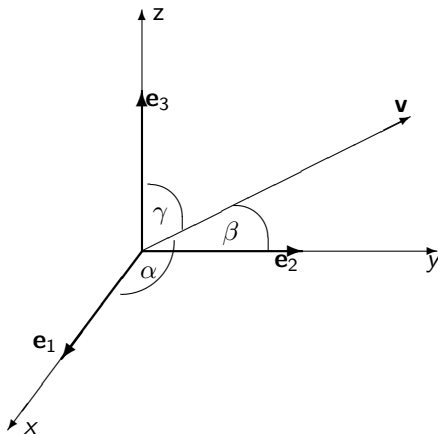
Der Vektor  $\mathbf{v} = (\nu_1, \nu_2, \nu_3)$  hat die **Länge**

$$\|\mathbf{v}\| = \sqrt{\nu_1^2 + \nu_2^2 + \nu_3^2} \quad \text{d.h.} \quad \|\mathbf{v}\|^2 = \mathbf{v}^T \mathbf{v}.$$

Diese nichtnegative reelle Zahl ist eine Verallgemeinerung des Betrages von reellen oder komplexen Zahlen und wird auch die **euklidische Norm** des Vektors  $\mathbf{v}$  genannt.

Dividiert man nun einen Vektor  $\mathbf{v} \neq 0$  durch seinen Betrag, so erhält man einen Vektor der Länge 1, dessen Komponenten als Kosinusse von drei Winkeln  $\alpha, \beta, \gamma \in [0, \pi]$  dargestellt werden können. Es gilt also

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \left( \frac{\nu_1}{\|\mathbf{v}\|}, \frac{\nu_2}{\|\mathbf{v}\|}, \frac{\nu_3}{\|\mathbf{v}\|} \right)^T = (\cos \alpha, \cos \beta, \cos \gamma)^T$$



Wie aus der Zeichnung ersichtlich ist, bilden  $\alpha$ ,  $\beta$  und  $\gamma$  die Winkel zwischen  $\mathbf{v}$  und den Basisvektoren  $\mathbf{e}_1$ ,  $\mathbf{e}_2$  und  $\mathbf{e}_3$ . Man kann also einen Vektor eindeutig durch diese drei Winkel und seine Länge definieren.



## Konstruktion des Lotfusspunktes

Gegeben sei ein Punkt  $P = O + \mathbf{v}$  und eine Gerade  $g = \{O + t\mathbf{u} : t \in \mathbb{R}\}$ . Der Punkt  $Q = O + t\mathbf{u}$  auf  $g$  mit kürzestem Abstand zu  $P$  wird Lotfusspunkt genannt.

$PQ = Q - P = t\mathbf{u} - \mathbf{v}$ . Abstand ist minimal  $\iff$  Quadrat des Abstandes ist minimal.

$$\begin{aligned}\|t\mathbf{u} - \mathbf{v}\|^2 &= (t\mathbf{u} - \mathbf{v})^T (t\mathbf{u} - \mathbf{v}) = t^2 \mathbf{u}^T \mathbf{u} - t(\mathbf{u}^T \mathbf{v} + \mathbf{v}^T \mathbf{u}) + \mathbf{v}^T \mathbf{v} \\ &= t^2 \|\mathbf{u}\|^2 - 2t\mathbf{u}^T \mathbf{v} + \|\mathbf{v}\|^2 \\ &= \left(t\|\mathbf{u}\| - \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\|}\right)^2 + \|\mathbf{v}\|^2 - \frac{(\mathbf{u}^T \mathbf{v})^2}{\|\mathbf{u}\|^2}\end{aligned}$$

Der Lotfusspunkt ergibt sich für  $t^* = \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\|^2}$  mit minimalem Abstand

$$\|t^* \mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{v}\|^2 - \frac{(\mathbf{u}^T \mathbf{v})^2}{\|\mathbf{u}\|^2} = \|\mathbf{v}\|^2 - (t^*)^2 \|\mathbf{u}\|^2.$$

## Geometrische Interpretation

Im Dreieck  $\triangle OPQ$  gilt der Satz des Pythagoras mit rechtem Winkel bei  $Q = O + t^* \mathbf{u}$ :

$$\|\mathbf{v}\|^2 = \|t^* \mathbf{u} - \mathbf{v}\|^2 + \|t^* \mathbf{u}\|^2, \quad \text{d.h.} \quad \|OP\|^2 = \|PQ\|^2 + \|OQ\|^2.$$

Deshalb kann durch  $\cos \varphi = \frac{\|OQ\|}{\|OP\|}$ ,  $\sin \varphi = \frac{\|PQ\|}{\|OP\|}$  der Winkel  $\varphi$  bei  $O$  zwischen  $\mathbf{u}$  und  $\mathbf{v}$  bestimmt werden, wenn  $\varphi \in [0, \frac{\pi}{2}] = [0^\circ, 90^\circ]$ .  
Allgemeiner gilt

$$\cos \varphi = \frac{t^* \|\mathbf{u}\|}{\|\mathbf{v}\|} = \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|},$$
$$\sin \varphi = \frac{\|t^* \mathbf{u} - \mathbf{v}\|}{\|\mathbf{v}\|} = \sqrt{1 - \frac{|\mathbf{u}^T \mathbf{v}|^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2}}$$

Insbesondere gilt  $|\mathbf{u}^T \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ .

## Skalar- oder inneres Produkt

Neben der Norm spielte die Größe  $\mathbf{u}^T \mathbf{v}$  eine besondere Rolle.

### Definition B-2.2 (Skalar- oder inneres Produkt)

Für zwei beliebige Vektoren  $\mathbf{u} = (\mu_1, \mu_2, \mu_3)^T$  und  $\mathbf{v} = (\nu_1, \nu_2, \nu_3)^T$  nennt man den Skalar

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \cdot \mathbf{v} = \mu_1 \nu_1 + \mu_2 \nu_2 + \mu_3 \nu_3$$

das Skalar- oder innere Produkt von  $\mathbf{u}$  und  $\mathbf{v}$ .

## Lemma B-2.3 (Eigenschaften Skalarprodukt)

*Es läßt sich nun leicht nachprüfen, daß für alle  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$  und  $\lambda \in \mathbb{R}$  gilt:*

$$\begin{aligned}\langle \mathbf{v}, \mathbf{u} \rangle &= \langle \mathbf{u}, \mathbf{v} \rangle \\ \lambda \langle \mathbf{u}, \mathbf{v} \rangle &= \langle \lambda \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \lambda \mathbf{v} \rangle \\ \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle \\ \langle \mathbf{u}, \mathbf{u} \rangle &= \|\mathbf{u}\|^2 \geq 0\end{aligned}$$

Weiter gilt die

## Lemma B-2.4 (Cauchy-Schwarzsche Ungleichung)

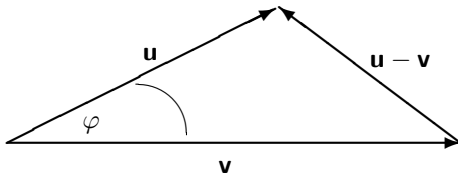
$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

### Bemerkung:

Die beiden Seiten sind nur dann genau gleich, wenn  $\varphi = 0$  oder  $\varphi = \pi = 180^\circ$ , d.h.  $\mathbf{v} = \lambda \mathbf{u}$  oder  $\mathbf{u} = \lambda \mathbf{v}$  sind für ein  $\lambda$ , das auch Null sein kann.

## Interpretation inneres Produkt im Anschauungsraum

Man betrachte das Dreieck  $\triangle ABC$  mit den Kanten  $\mathbf{u} = AB$ ,  $\mathbf{v} = AC$  und  $\mathbf{v} - \mathbf{u} = BC$ .



Für  $\|\mathbf{u} - \mathbf{v}\|^2$  gilt nach den oben aufgeführten Regeln

$$\begin{aligned}\|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} \rangle - \langle \mathbf{u} - \mathbf{v}, \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2 \langle \mathbf{u}, \mathbf{v} \rangle .\end{aligned}$$

Sei  $\varphi$  der von  $\mathbf{u}$  und  $\mathbf{v}$  eingeschlossene Winkel. Setzt man nun

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos \varphi$$

in diese Gleichung ein, so ergibt sich der

### Lemma B-2.5 (Kosinussatz)

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \|\mathbf{v}\| \cos \varphi \quad \text{bzw.} \\ |BC|^2 &= |AB|^2 + |AC|^2 - 2|AB| |AC| \cos(\sphericalangle BAC) \end{aligned}$$

Hierbei haben wir natürlich vorausgesetzt, dass weder  $\mathbf{u}$  noch  $\mathbf{v}$  gleich dem Nullvektor ist.

## Definition B-2.6 (Orthogonale Vektoren)

Man bezeichnet zwei Vektoren  $\mathbf{u}$  und  $\mathbf{v}$  als orthogonal zueinander, wenn der von ihnen eingeschlossene Winkel  $\pi/2 = 90^\circ$  ist, oder wenn einer von ihnen verschwindet, d.h. gleich Null ist. Formelmäßig schreibt man

$$\mathbf{u} \perp \mathbf{v}, \quad \text{falls} \quad \langle \mathbf{u}, \mathbf{v} \rangle = 0.$$

## Beispiel B-2.7

Die Einheitsvektoren  $\mathbf{e}_i$  bilden ein Orthogonalsystem in dem Sinne, dass

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0 \quad \text{falls} \quad i \neq j.$$

Es gibt aber auch noch andere Vektortripel mit dieser Eigenschaft. Das innere Produkt lässt sich entsprechend auf allen endlich dimensional Räumen definieren, es gibt dazu sogar mehrere Möglichkeiten.

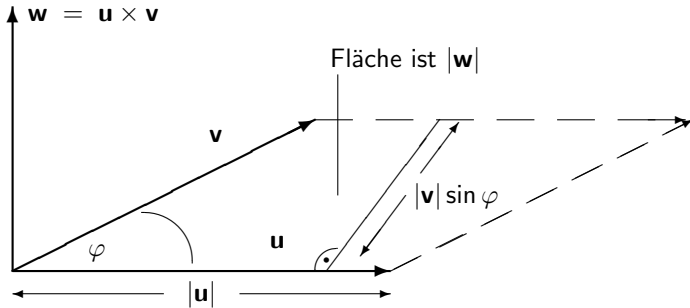
## Vektor- oder Kreuzprodukt

Dieses Produkt ist nur im dreidimensionalen Anschauungsraum eindeutig definiert.

### Definition B-2.8 (Vektor- oder Kreuzprodukt)

Zu je zwei nicht verschwindenden Vektoren  $\mathbf{u}$  und  $\mathbf{v}$  bezeichnet man als **Vektor- oder Kreuzprodukt** den Vektor  $\mathbf{w} = \mathbf{u} \times \mathbf{v}$ , dessen Richtung zu  $\mathbf{u}$  und  $\mathbf{v}$  orthogonal ist und dessen Länge  $|\mathbf{w}|$  gleich der Fläche des von  $\mathbf{u}$  und  $\mathbf{v}$  aufgespannten Parallelogramms ist. Weiter sei, von der Spitze von  $\mathbf{w}$  betrachtet,  $\mathbf{v}$  weniger als  $\pi = 180^\circ$  gegen den Uhrzeigersinn von  $\mathbf{u}$  weggedreht.





## Rechtssystem

Zeigt man mit dem Daumen und dem Zeigefinger längs der Vektoren  $\mathbf{u}$  und  $\mathbf{v}$ , so muss  $\mathbf{w}$  in die Richtung des nach innen abgeknickten Mittelfingers zeigen.

In diesem Sinne sind auch die drei Basisvektoren  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  rechtshändig orientiert.

Gemäß den oben genannten Anforderungen gilt nun insbesondere:

$$\begin{array}{ll} \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, & \mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2 \\ \mathbf{e}_2 \times \mathbf{e}_1 = -\mathbf{e}_3, & \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1 \\ \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2, & \mathbf{e}_3 \times \mathbf{e}_2 = -\mathbf{e}_1. \end{array}$$

Es soll also gelten:  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  bilden ein Rechtssystem,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{u}^T \mathbf{w} \\ \mathbf{v}^T \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{u}^T \\ \mathbf{v}^T \end{pmatrix} \mathbf{w} = \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix},$$

sowie

$$\begin{aligned} \|\mathbf{w}\| &= \|\mathbf{u}\| \|\mathbf{v}\| |\sin(\varphi)| \\ &= \|\mathbf{u}\| \|\mathbf{v}\| \sqrt{1 - \cos^2(\varphi)} \\ &= \sqrt{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \langle \mathbf{u}, \mathbf{v} \rangle^2} \\ &= \sqrt{\sum_{1 \leq j < k \leq 3} (\mu_j \nu_k - \mu_k \nu_j)^2}. \end{aligned}$$

### Bemerkung:

Man beachte, dass auf der letzten rechten Seite der Ausdruck unter der Wurzel nach der Cauchy-Schwarzschen Ungleichung nie negativ sein kann.

## Lemma B-2.9

*Dieses System von Bedingungen wird vom Vektor*

$$\mathbf{w} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} \mu_2\nu_3 - \mu_3\nu_2 \\ \mu_3\nu_1 - \mu_1\nu_3 \\ \mu_1\nu_2 - \mu_2\nu_1 \end{pmatrix}$$

*erfüllt.*

Diese Bildungsvorschrift merkt man sich am besten, indem man sie als die (formale) Determinante einer  $(3 \times 3)$  Matrix interpretiert. (Determinanten werden in einem der nächsten Abschnitte behandelt.) Und zwar gilt

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{vmatrix}.$$

### Bemerkung:

Hierbei handelt es sich allerdings nicht um eine gewöhnliche Matrix, da die drei Elemente in der ersten Zeile Vektoren, die Elemente der zweiten und dritten Zeile aber Skalare sind.

## Lemma B-2.10 (Vorzeichen des Vektorproduktes)

Werden die Vektoren  $\mathbf{u}$  und  $\mathbf{v}$  im Vektorprodukt vertauscht, dann ändert sich nur das Vorzeichen des Vektorproduktes:

$$\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$$

## Lemma B-2.11 (Bilinearität)

Für beliebige Vektoren  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  und Skalare  $\lambda$  gilt:

$$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$$

$$(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$$

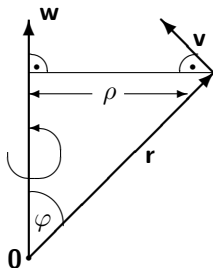
$$\lambda(\mathbf{u} \times \mathbf{v}) = (\lambda\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (\lambda\mathbf{v}) = -\lambda(\mathbf{v} \times \mathbf{u})$$

Eine wichtige Anwendung des Kreuzproduktes in der Mechanik ist die **Drehung eines Körpers um eine feste Achse mit der konstanten Winkelgeschwindigkeit**  $\omega$ . Man beschreibt diese Rotation durch einen Vektor  $\mathbf{w}$ , dessen Richtung  $\frac{\mathbf{w}}{\|\mathbf{w}\|}$  parallel zur Rotationsachse ist und dessen Länge die Winkelgeschwindigkeit repräsentiert, so dass  $\omega = \|\mathbf{w}\|$  ist.

Der Vektor  $\mathbf{w}$  ist so orientiert, dass die Drehung beim Blicken entlang seiner Richtung im Uhrzeigersinn erfolgt.

Ohne wesentliche Beschränkung der Allgemeinheit nehme man nun an, daß die Drehachse genau durch den Ursprung verläuft.

Dann erhält man den momentanen Geschwindigkeitsvektor  $\mathbf{v}$  eines Körperpunktes  $P$  mit derzeitigem Ortsvektor  $\mathbf{r} = \overrightarrow{OP}$  als  $\mathbf{v} = \mathbf{w} \times \mathbf{r}$ .



Diese Formel ergibt sich gemäß der Zeichnung aus der Beobachtung, daß die momentane Bewegungsrichtung  $\mathbf{v}/\|\mathbf{v}\|$  orthogonal zu  $\mathbf{w}$  und  $\mathbf{r}$  sein muß und dass der Geschwindigkeitsbetrag  $\|\mathbf{v}\|$  gleich  $\omega$  mal dem Abstand von der Achse, also  $\|\mathbf{r}\| \sin(\phi)$ , ist.

Hierbei ist  $\varphi$  der von den Vektoren  $\mathbf{w}$  und  $\mathbf{r}$  eingeschlossene Winkel und die Orientierung der resultierenden Geschwindigkeit  $\mathbf{v}$  ist so, daß  $\mathbf{w}$ ,  $\mathbf{r}$ ,  $\mathbf{v}$  ein rechtshändiges System bilden.

$$\mathbf{w} \cdot \mathbf{v} = 0 = \mathbf{r} \cdot \mathbf{v}$$

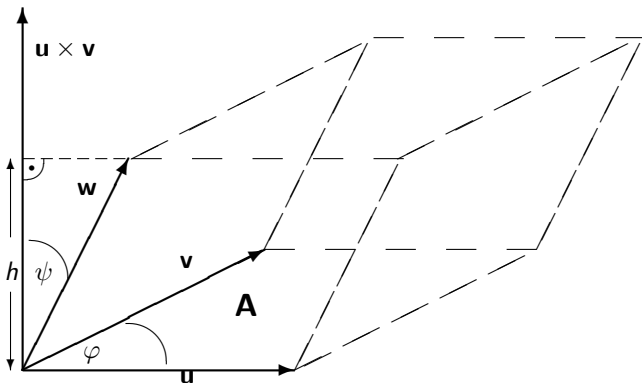
$$\|\mathbf{v}\| = \rho \omega = \|\mathbf{r}\| |\sin \varphi| \|\mathbf{w}\| = \|\mathbf{r} \times \mathbf{w}\|$$

# Spatprodukt

## Definition B-2.12 (Spatprodukt)

Bildet man das Skalarprodukt zwischen  $\mathbf{u} \times \mathbf{v}$  und einem dritten Vektor  $\mathbf{w}$ , so ergibt sich das sogenannte Spatprodukt:

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] \equiv (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} \in \mathbb{R}.$$





## Lemma B-2.13 (Betrag Spatprodukt)

*Gemäß der Zeichnung ergibt der Betrag*

$$|[\mathbf{u}, \mathbf{v}, \mathbf{w}]| = \underbrace{(\|\mathbf{u}\| \|\mathbf{v}\| \sin(\varphi))}_A \underbrace{\|\mathbf{w}\| \cos(\psi)}_h$$

*genau das Volumen des Parallelepipeds mit der Grundfläche  $A$  und der Höhe  $h$ .*

## Folgerung B-2.14

*Daraus sieht man unmittelbar, daß das Spatprodukt bis auf das Vorzeichen von der Reihenfolge der Vektoren  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$  unabhängig ist, da diese immer das gleiche Parallelepiped aufspannen.*

## Lemma B-2.15 (Vorzeichen Spatprodukt)

Für das **Vorzeichen** gilt die folgende Regel:

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] \quad \left\{ \begin{array}{lll} > 0 & \text{falls} & (\mathbf{u}, \mathbf{v}, \mathbf{w}) \text{ Rechtssystem} \\ < 0 & \text{falls} & (\mathbf{u}, \mathbf{v}, \mathbf{w}) \text{ Linkssystem} \\ = 0 & \text{falls} & (\mathbf{u}, \mathbf{v}, \mathbf{w}) \text{ linear abhängig} \end{array} \right.$$

Hierbei bezeichnet der Begriff **linear abhängig** den Zustand, dass die drei Vektoren in einer Ebene liegen und es deshalb nicht triviale Koeffizienten  $\alpha, \beta, \gamma$  gibt, für die

$$\alpha \mathbf{u} + \beta \mathbf{v} + \gamma \mathbf{w} = \mathbf{0} \quad \text{gilt}.$$

## Lemma B-2.16 (Identität im Anschauungsraum)

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \begin{vmatrix} \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \\ \omega_1 & \omega_2 & \omega_3 \end{vmatrix},$$

wobei  $\mathbf{w} = (\omega_1, \omega_2, \omega_3)$  ist.

## B-3 Abstandsnormen

Eine ganz zentrale Rolle in der linearen Algebra und (allgemeiner der sogenannten Funktionalanalysis) spielt der Begriff des **Abstandes** zwischen zwei Vektoren (z.B. auch Funktionen). Dadurch ergibt sich die Möglichkeit, „Kugeln“ und andere „Umgebungen“ von Vektoren zu betrachten die „nahe“ beieinander liegen.

### Definition B-3.1 (Norm und normierter Raum)

Ein linearer Vektorraum  $\mathcal{V}$  heisst **normiert**, wenn es zu jedem  $\mathbf{u} \in \mathcal{V}$  eine reelle Zahl  $\|\mathbf{u}\|$  gibt, so dass für beliebige  $\lambda \in \mathbb{R}$  und  $\mathbf{v} \in \mathcal{V}$  gilt:

- ▶  $\|\mathbf{u}\| \geq 0$  mit  $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$  **Definitheit**
- ▶  $\|\lambda \mathbf{u}\| = |\lambda| \|\mathbf{u}\|$  **Homogenität**
- ▶  $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$  **Dreiecksungleichung**

Hier ist  $|\lambda|$  der gewöhnliche Betrag reeller Zahlen.

Aus der Cauchy-Schwarz-Ungleichung folgt unmittelbar die Dreiecksungleichung für die euklidische Norm  $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\mathbf{u}^T \mathbf{u}}$ , da

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2 \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &\leq \|\mathbf{u}\|^2 + 2 |\langle \mathbf{u}, \mathbf{v} \rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2 \|\mathbf{u}\| \|\mathbf{v}\| + \|\mathbf{v}\|^2 \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

Auch die Homogenität ist gewährleistet, da

$$\|\lambda \mathbf{u}\| = \sqrt{\langle \lambda \mathbf{u}, \lambda \mathbf{u} \rangle} = |\lambda| \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

Dies gilt auch für verallgemeinerte Skalarprodukte, deren **Hilbert-Normen**  $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$  ebenfalls die Normeigenschaften erfüllen.

Man nennt den Vektorraum dann auch **Hilbert-Raum**.

In numerischen Anwendungen der Lineare Algebra werden neben der Euklidischen Norm häufig folgende anderen Normen benutzt:

- Für festes  $1 \leq p \leq \infty$  setze

$$\|\mathbf{v}\|_p = \|(\nu_1, \nu_2, \dots, \nu_n)^T\|_p = [|\nu_1|^p + |\nu_2|^p + \dots + |\nu_n|^p]^{1/p}$$

- Für  $p = 2$  erhält man wiederum die Euklidische Norm  $\|\mathbf{v}\|_2 = \|\mathbf{v}\|$ .  
Im Grenzfall  $p = \infty$  setzt man

$$\|\mathbf{v}\|_\infty = \|(\nu_1, \nu_2, \dots, \nu_n)^T\|_\infty = \max\{|\nu_1|, |\nu_2|, \dots, |\nu_n|\}$$

- Die Menge der Vektoren  $\mathbf{u}$  mit  $\|\mathbf{u}\|_1 \leq 1$  und  $\|\mathbf{u}\|_\infty \leq 1$  bilden für  $n = 2$  (d.h. in der Ebene) ein achsenparalleles bzw. diagonal orientiertes Quadrat.
- Bei den Zwischenwerten  $1 < p < \infty$  und insbesondere der Euklidischen Norm  $\|\mathbf{u}\|_2$  haben die verallgemeinerten Kugeln  $\{\mathbf{v} \in \mathcal{V} : \|\mathbf{u}\|_p \leq 1\}$  dagegen keine Ecken.
- Die beiden Grenzfälle  $p = 1$  und  $p = \infty$  haben den Vorteil, dass die entsprechenden Normen billig auswertbar sind.

## Lemma B-3.2 (Weitere Normeigenschaften)

- ▶ Per Induktion ergibt sich für die **Summe** endlich vieler Vektoren  $\mathbf{v}_i, i = 1 \dots m$ , die Ungleichung

$$\left\| \sum_{i=1}^m \mathbf{v}_i \right\| \leq \sum_{i=1}^m \|\mathbf{v}_i\|$$

- ▶ Aus der Dreiecksungleichung folgt für alle Normen die sogenannte **umgekehrte Dreiecksungleichung**

$$\|\mathbf{u} - \mathbf{v}\| \geq \left| \|\mathbf{u}\| - \|\mathbf{v}\| \right|$$

- ▶ Eine Norm  $\|\mathbf{v}\|$  ist genau dann eine **Hilbert-Norm**, wenn sie die folgende **Parallelogrammgleichung** erfüllt

$$\|\mathbf{u} - \mathbf{v}\|^2 + \|\mathbf{u} + \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2)$$

### Bemerkung:

Im letzteren Fall lässt sich die Identität

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4} [\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2]$$

auch als Definition des Inneren Produktes interpretieren.

## B-4 Basen und Unterräume

Im vorigen Abschnitt wurde festgestellt, dass im Anschauungsraum drei Vektoren **linear abhängig** sind (d.h. in einer Ebene liegen), wenn ihr Spatprodukt verschwindet.

Das Konzept der linearen Abhängigkeit bzw. Unabhängigkeit ist von zentraler Bedeutung für die Untersuchung beliebiger Räume und ihrer sogenannten Unterräume.

## Definition B-4.1 (Lineare Abhängigkeit und Unabhängigkeit)

Eine Familie (= Menge) von Vektoren  $\{\mathbf{v}_i\}_{i=1}^r \subset \mathcal{V}$  heißt **linear abhängig**, wenn es Skalare  $\{\lambda_i\}_{i=1}^r \subset \mathbb{R}$  gibt so daß gilt

$$\sum_{i=1}^n \lambda_i \mathbf{v}_i = \mathbf{0} \quad \text{und} \quad \sum_{i=1}^n |\lambda_i| \neq 0 .$$

Die zweite Bedingung schließt die Möglichkeit aus, dass alle  $\lambda_i$  verschwinden, in welchem Falle die erste Bedingung trivialerweise für jede Familie  $\{\mathbf{v}_i\}_{i=1}^r \subset \mathcal{V}$  zuträfe.

Umgekehrt heißt eine Familie  $\{\mathbf{v}_i\}_{i=1}^r \subset \mathcal{V}$  **linear unabhängig**, falls

$$\sum_{i=1}^n \lambda_i \mathbf{v}_i = \mathbf{0} \quad \implies \quad \sum_{i=1}^n |\lambda_i| = 0 .$$



## Folgerung B-4.2

*Man sieht leicht, daß eine Obermenge linear abhängiger Vektoren auch linear abhängig ist, während eine Untermenge linear unabhängiger Vektoren auch linear unabhängig ist.*

## Folgerung B-4.3

*Zwei Vektoren  $\mathbf{v}_1, \mathbf{v}_2$  sind genau dann linear abhängig, wenn sie parallel sind, da*

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 = \mathbf{0}, \lambda_1 \neq 0 \implies \mathbf{v}_1 = -(\lambda_2/\lambda_1) \mathbf{v}_2 .$$

## Bemerkung:

Hierbei haben wir ohne Beschränkung der Allgemeinheit vorausgesetzt, dass  $\lambda_1 \neq 0$ . Entsprechendes gilt, wenn  $\lambda_2 \neq 0$ , aber möglicherweise  $\lambda_1 = 0$ .

## Folgerung B - 4.4

*Zwei nicht verschwindende, zueinander orthogonale Vektoren  $\mathbf{v}_1 \perp \mathbf{v}_2$  sind auf jeden Fall linear unabhängig.*

### Beweis:

Um dies zu zeigen, bildet man das Skalarprodukt von  $\mathbf{v}_1$  mit beiden Seiten der Gleichung

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 = \mathbf{0}$$

und erhält

$$\lambda_1 \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + \lambda_2 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0 = \lambda_1 |\mathbf{v}_1|^2$$

und somit  $\lambda_1 = 0$ .

Entsprechend folgt aus dem Skalarprodukt mit  $\mathbf{v}_2$  die Gleichung  $\lambda_2 = 0$  und damit die behauptete lineare Unabhängigkeit von  $\mathbf{v}_1$  und  $\mathbf{v}_2$ . □

### Beobachtung:

Dieselbe Schlussfolgerung kann man leicht für eine Familie von beliebig vielen paarweise orthogonalen Vektoren  $\{\mathbf{v}_i\}_{i=1}^r \subset \mathcal{V}$  mit  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ ,  $\mathbf{v}_i \neq \mathbf{0}$ , für  $i \neq j$  durchführen. Deshalb sollte man Orthogonalität als eine besonders starke Form linearer Unabhängigkeit betrachten.

## Folgerung B-4.5 (Lineare Unabhängigkeit im $\mathbb{R}^3$ )

*Man kann zeigen, dass es im Anschauungsraum  $\mathbb{R}^3$  jeweils maximal drei linear unabhängige Vektoren ( wie z.B.  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ ) gibt.*

## Definition B-4.6 (Dimension eines Vektorraumes)

Die maximale Zahl linear unabhängiger Vektoren in einem Raum  $\mathcal{V}$  wird als dessen **Dimension**  $\dim(\mathcal{V})$  bezeichnet.

Falls es Familien linear unabhängiger Vektoren mit beliebig vielen Elementen in einem Raum  $\mathcal{V}$  gibt, so bezeichnet man ihn als **unendlichdimensional** und setzt  $\dim(\mathcal{V}) = \infty$ .

## Beispiel B-4.7

Der Raum aller Polynome ist unendlichdimensional, da die Familie der Monome (reine Potenzen von  $x$ )

$$x^j \quad j = 0, 1, \dots, n$$

für ein beliebiges  $n$  linear unabhängig ist.

## Bemerkung:

Im folgenden geht es darum nachzuweisen, dass der Dimensionsbegriff eindeutig ist und dass jede Menge linear unabhängiger Vektoren zu einer Basis (d.h. Menge von dim  $\mathcal{V}$  linear unabhängigen Vektoren) erweitert werden kann.

## Lemma B - 4.8 (Eindeutige Koeffizientendarstellung)

*Sei  $\{v_i\}_{i=1\dots m}$  eine Familie linear unabhängiger Vektoren irgendeines linearen Raumes  $\mathcal{V}$ . Dann besitzt jeder Vektor  $v \in \mathcal{V}$ , der zusammen mit den  $\{v_i\}_{i=1\dots m}$  keine linear unabhängige Menge bildet, eine eindeutige Darstellung*

$$\mathbf{v} = \sum_{i=1}^m \lambda_i \mathbf{v}_i$$

*Hierbei verschwinden alle  $\lambda_i$  genau dann wenn  $\mathbf{v} = 0$ .*

## Lemma B-4.9 (Austauschsatz)

Sei  $\{v_i\}_{i=1\dots m}$  eine Familie linear unabhängiger Vektoren irgendeines linearen Raumes  $\mathcal{V}$ .

Dann gilt für jeden nichtverschwindenden Vektor  $v \neq 0$

**entweder**

- ▶ Die Vereinigung  $\{v_i\}_{i=1\dots m+1}$  ist mit  $v_{m+1} \equiv v$  auch linear unabhängig.

**oder**

- ▶ Es gibt einen Index  $j \leq m$ , so dass

$$\{v_1, \dots, v_{j-1}, v, v_{j+1}, \dots, v_m\}$$

( $v_j$  wurde durch  $v$  ersetzt) weiterhin linear unabhängig ist.

## Bemerkung:

Wir nennen eine Familie von Vektoren **maximal linear unabhängig**, wenn die Hinzunahme irgendeines anderen Vektors die lineare Unabhängigkeit zerstört. Man nennt eine solche Menge dann auch **Basis des Raumes**. Der folgende Satz zeigt, dass alle Basen dieselbe Anzahl von Elementen haben: die Dimension des Raumes.

## Satz B-4.10 (Eindeutigkeit der Dimension)

*Seien  $\{v_i\}_{i=1\dots m}$  und  $\{w_i\}_{i=1\dots n}$  zwei maximal unabhängige Familien von Vektoren.*

*Dann gilt  $m = n = \dim(\mathcal{V})$  und für jeden Vektor  $u \in \mathcal{V}$  gibt es eindeutige Koeffizienten  $\{\alpha_i\}_{i=1}^n \subset \mathbb{R}$  und  $\{\beta_i\}_{i=1}^n \subset \mathbb{R}$  so dass*

$$\sum_{i=1}^n \alpha_i v_i = u = \sum_{i=1}^n \beta_i w_i .$$

## Unterräume und Linearkombinationen

Gerade in unendlichdimensionalen Räumen muss man oft praktische Untersuchungen auf einem endlichdimensionalen Unterraum beschränken (z.B. indem man den Grad von Polynomen mehr oder minder willkürlich beschränkt).

### Definition B-4.11 (Unterraum)

Ein **Unterraum** ist eine **Menge**  $\mathcal{U} \subset \mathcal{V}$ , die bezüglich der Addition von Vektoren und deren Multiplikation mit Skalaren **abgeschlossen** ist, d.h. es gilt für alle  $\mathbf{u}, \mathbf{v} \in \mathcal{U}$  und  $\lambda \in \mathbb{R}$  die Implikation

$$\mathbf{u}, \mathbf{v} \in \mathcal{U} \quad \implies \quad \mathbf{u} + \mathbf{v} \in \mathcal{U}, \quad \lambda \mathbf{u} \in \mathcal{U}.$$

### Beispiel B-4.12

Triviale Beispiele von Unterräumen sind  $\mathcal{V}$  selbst und der nur aus dem Nullvektor bestehende Raum  $\{\mathbf{0}\}$ , den man als nulldimensional betrachtet.

## Beispiel B-4.13 (Orthogonales Komplement)

Ein interessanteres Beispiel ist das **orthogonale Komplement**

$$v^\perp \equiv \mathcal{U} = \{\mathbf{u} \in \mathcal{V} \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0\}$$

eines fest vorgegebenen Vektors  $\mathbf{v}$ .

Die Abgeschlossenheit und damit Unterraumeigenschaft ersieht man aus der Tatsache, daß für alle  $\mathbf{u}, \mathbf{w} \in \mathcal{U}$  und  $\lambda \in \mathbb{R}$

$$\langle \mathbf{v}, \mathbf{u} \rangle = 0 = \langle \mathbf{v}, \mathbf{w} \rangle \quad \implies \quad \langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = 0 = \langle \mathbf{v}, \lambda \mathbf{u} \rangle .$$

Mit anderen Worten: Gehören  $\mathbf{u}$  und  $\mathbf{w}$  zum orthogonalen Komplement von  $\mathbf{v}$ , so gilt dies auch für die Summe  $\mathbf{u} + \mathbf{w}$  und das Produkt  $\lambda \mathbf{u}$ .



## Satz B-4.14 (Schnittprinzip, siehe auch Lemma ??)

*Für zwei Unterräume  $\mathcal{U}, \mathcal{W} \subset \mathcal{V}$  bildet deren Durchschnitt*

$$\mathcal{U} \cap \mathcal{W} = \{\mathbf{v} \in \mathcal{V} \mid \mathbf{v} \in \mathcal{U} \ \& \ \mathbf{v} \in \mathcal{W}\}$$

*einen Unterraum.*

## Satz B-4.15

*Der Schnitt mehrerer und sogar unendlich vieler Unterräume bildet einen Unterraum.*

## Beispiel B-4.16

Für eine beliebige Menge von Vektoren  $\mathcal{M} \subset \mathcal{V}$  ergibt sich das orthogonale Komplement als

$$\mathcal{M}^\perp \equiv \bigcap_{\mathbf{v} \in \mathcal{M}} \mathbf{v}^\perp = \{\mathbf{u} \in \mathcal{V} \mid \forall \mathbf{v} \in \mathcal{M} : \mathbf{u} \cdot \mathbf{v} = 0\}.$$

## Bemerkung:

Im Gegensatz zum Durchschnitt ist die mengentheoretische Vereinigung von zwei Unterräumen  $\mathcal{U}, \mathcal{W} \subset \mathcal{V}$  nur dann selbst ein Unterraum, wenn  $\mathcal{U}$  schon in  $\mathcal{W}$  oder  $\mathcal{W}$  schon in  $\mathcal{U}$  enthalten ist (siehe Warnung in **A-2**). Es gibt jedoch einen kleinsten Unterraum von  $\mathcal{V}$ , der sowohl  $\mathcal{U}$  als auch  $\mathcal{W}$  enthält und mit  $\mathcal{U} + \mathcal{W}$  bezeichnet wird.

Diese Bezeichnung ist sinnvoll, denn es gilt:

$$\mathcal{U} + \mathcal{W} = \{ \mathbf{u} + \mathbf{w} \mid \mathbf{u} \in \mathcal{U}, \mathbf{w} \in \mathcal{W} \}.$$

Man sagt dann auch, daß die Summe  $\mathcal{U} + \mathcal{W}$  von  $\mathcal{U}$  und  $\mathcal{W}$  aufgespannt wird.

Natürlich kann man auch die Summe mehrerer Unterräume bilden, was besonders dann von Interesse ist, wenn diese jeweils eindimensional sind.

## Definition B-4.17 (Linearkombination der Vektoren)

Für eine Familie  $\{\mathbf{v}_i\}_{i=1}^r \subset \mathcal{V}$  bezeichnet man jeden Vektor der Form

$$\mathbf{v} = \sum_{i=1}^r \lambda_i \mathbf{v}_i \quad \lambda_1, \dots, \lambda_r \in \mathbb{R}$$

als eine **Linearkombination der Vektoren  $\mathbf{v}_i$** .

## Definition B-4.18 (Lineare Hülle, vergleiche ??)

Die Menge aller möglichen Linearkombinationen von Vektoren  $\{\mathbf{v}_i\}_{i=1}^r = \mathcal{U}$  aus einer Teilmenge  $\mathcal{U} \subset \mathcal{V}$  bezeichnet man als deren **lineare Hülle**

$$\text{span}(\mathcal{U}) = \left\{ \mathbf{v} = \sum_{i=1}^r \lambda_i \mathbf{v}_i \mid \lambda_i \in \mathbb{R}, \mathbf{v}_i \in \mathcal{U} \right\}.$$

Die lineare Hülle ist abgeschlossen. Man bezeichnet sie deshalb auch als den von  $\{\mathbf{v}_i\}_{i=1}^r = \mathcal{U} \subset \mathcal{V}$  **aufgespannten Unterraum**.

## Definition B-4.19 (Basis eines Unterraumes)

Falls die Vektoren  $\{\mathbf{v}_i\}_{i=1}^r$  linear unabhängig sind, bezeichnet man sie als eine **Basis** des von ihnen aufgespannten Unterraumes.

## Folgerung B-4.20

*Aus der vorausgesetzten linearen Unabhängigkeit folgt die Eindeutigkeit der Darstellung eines beliebigen Vektors  $\mathbf{v}$  als Linearkombination.*

## Lemma B-4.21

Bezüglich einer bestimmten Basis  $\{\mathbf{v}_i\}_{i=1}^r$  hat jeder Vektor  $\mathbf{w} \in \mathcal{V}$  eine eindeutige Darstellung

$$\mathbf{w} = \sum_{i=1}^r \lambda_i \mathbf{v}_i$$

### Beweis.

Aus

$$\sum_{i=1}^r \lambda_i \mathbf{v}_i = \mathbf{w} = \sum_{i=1}^r \gamma_i \mathbf{v}_i$$

erhält man durch Abzug der rechten Seite von der linken

$$0 = \sum_{i=1}^r (\lambda_i \mathbf{v}_i - \gamma_i \mathbf{v}_i) = \sum_{i=1}^r (\lambda_i - \gamma_i) \mathbf{v}_i,$$

so dass wegen der linearen Unabhängigkeit der Basisvektoren notwendigerweise alle  $\lambda_i - \gamma_i = 0$  sind. Also sind die Koeffizienten  $\lambda_i = \gamma_i$  von  $\mathbf{w}$  bezüglich der gewählten Basis eindeutig bestimmt. □

## Beispiel B - 4.22

Es sei

$$\mathcal{P}_n \equiv \left\{ \sum_{i=0}^{n-1} \gamma_i x^i \mid \gamma_i \in \mathbb{R} \right\}$$

die Menge aller Polynome mit reellen Koeffizienten vom Grade kleiner  $n$ .  
Bezüglich der üblichen Addition von Polynomen und ihrer Multiplikation mit reellen Skalaren ist  $\mathcal{P}_n$  ein Vektorraum.

Die Monome  $\{\mathbf{v}_i = x^{i-1}\}_{i=1}^n$  bilden eine Basis des Vektorraumes  $\mathcal{P}_n$ , der deshalb  $n$ -dimensional ist.

### Beweisidee:

Zur linearen Unabhängigkeit: Es gebe eine nichttriviale Linearkombination, die das Nullpolynom ergibt,

$$P(x) = \sum_{i=1}^n \lambda_i \mathbf{v}_i = \sum_{i=1}^n \lambda_i x^{i-1} = 0.$$

Einerseits kann  $P(x)$  nur maximal  $n - 1$  Nullstellen haben, siehe Folgerung aus Korollar ?? . Im Widerspruch dazu ist jedes  $x \in \mathbb{R}$  eine Nullstelle von  $P(x)$ .

Also müssen alle Koeffizienten  $\lambda_i$  von  $P(x)$  gleich Null sein.



### Bemerkung:

Obwohl die Monombasis von  $\mathcal{P}_n$  sehr natürlich erscheint, ist sie keineswegs für alle im Zusammenhang mit Polynomen auftretenden mathematischen Aufgaben geeignet.

Allgemein kommen in linearen Räumen oftmals verschiedene Basen zur Anwendung, je nachdem, welche Art von Berechnung oder Untersuchung durchgeführt werden soll. Das Umrechnen der Koeffizienten eines Vektors von einer Basis auf eine andere nennt man Basistransformation. Diese verlangt normalerweise die Lösung eines linearen Gleichungssystems, wie sie im entsprechenden Abschnitt weiter unten behandelt wird.

## Bemerkung: Orthogonalitätsbedingung, orthonormale Basis

Rechnerisch besonders angenehm sind Basen, welche die  
**Orthogonalitätsbedingung**

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

erfüllen. Bei solchen **orthonormalen** Basen lassen sich die Koeffizienten  $\lambda_i$  eines beliebigen Vektors  $\mathbf{w}$  leicht berechnen:

Aus dem Ansatz 
$$\mathbf{w} = \sum_{j=1}^n \lambda_j \mathbf{v}_j$$

folgt durch die Bildung des inneren Produktes mit einem bestimmten Basisvektor  $\mathbf{v}_i$  sofort

$$\langle \mathbf{v}_i, \mathbf{w} \rangle = \sum_{j=1}^n \lambda_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \lambda_i,$$

da alle Summanden mit  $j \neq i$  verschwinden.



## Beispiel B - 4.23

In einem verallgemeinerten Sinne des Basisbegriffs bilden die trigonometrischen Funktionen

$$\mathbf{v}_{2j} \equiv \sin(jx) \quad \text{und} \quad \mathbf{v}_{2j+1} \equiv \cos(jx) \quad \text{für } j = 1, 2, \dots$$

zusammen mit der konstanten Funktion  $\mathbf{v}_1 \equiv 1$  eine Basis des unendlichdimensionalen Raumes aller Funktionen  $f(x)$ , die auf dem Intervall  $[-\pi, \pi]$  periodisch und quadratisch integrierbar sind.

Letzteres bedeutet, dass die Funktion  $f^2(x)$  ein endliches Integral auf  $[-\pi, \pi]$  hat. Das ist zum Beispiel dann der Fall, wenn  $f(x)$  bis auf endlich viele Sprungstellen **stetig** ist. Das innere Produkt, bezüglich dessen die trigonometrischen Funktionen eine orthogonale Basis bilden, ist nun das Integral

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) dx .$$

## Fortsetzung Beispiel

Die Orthogonalitätseigenschaften lassen sich hier mittels wiederholter partieller Integration oder mit Hilfe geeigneter trigonometrischer Umformungen leicht nachweisen. Allerdings müssen die Funktionen  $\mathbf{v}_j$  noch geeignet skaliert werden, so daß dann  $\|\mathbf{v}_j\| = 1$  gilt.

Dies vorausgesetzt lassen sich die Koeffizienten einer beliebigen Funktion  $f(x)$  bezüglich der Basisfunktion  $\sin(jx)$  als inneres Produkt

$$\int_{-\pi}^{\pi} f(x) \sin(jx) dx$$

berechnen.

## Warnung:

Da sich diese Integrale im allgemeinen nicht formelmäßig auswerten lassen, kommen hierbei in der Praxis oft Quadraturen, d.h. numerische Integrationsverfahren, zur Anwendung.

Streng genommen besteht der Vektorraum nicht aus den Funktionen selbst, sondern seine Elemente bilden Äquivalenzklassen von Funktionen, die sich nur an endlich vielen Punkten unterscheiden, so dass das Integral des Quadrates ihrer Differenz Null ergibt.

Die genauere Untersuchung und Beschreibung von Funktionenräumen und ihrer Basen ist der Ausgangspunkt der mathematischen Disziplin **Funktionalanalysis**.

## B-5 Lineare Abbildungen

### Definition B-5.1 (Lineare Abbildung)

Eine Abbildung  $\mathbf{F} : \mathcal{V} \rightarrow \mathcal{W}$  zwischen zwei reellen Vektorräumen  $\mathcal{V}$  und  $\mathcal{W}$  heißt linear, falls für alle  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  und  $\lambda \in \mathbb{R}$  gilt:

$$\mathbf{F}(\mathbf{u} + \mathbf{v}) = \mathbf{F}(\mathbf{u}) + \mathbf{F}(\mathbf{v})$$

**Additivität**

$$\mathbf{F}(\lambda \mathbf{u}) = \lambda \mathbf{F}(\mathbf{u})$$

**Homogenität**

### Bemerkung

Mit anderen Worten:  $F$  ist ein Vektorraumhomomorphismus im Sinne der auf Gruppen und Ringe zugeschnittenen algebraischen Definition A-6.4.

### Bemerkung

Die beiden definierenden Eigenschaften können auch zu der einen

$$F(\mathbf{u} + \lambda \mathbf{v}) = F(\mathbf{u}) + \lambda F(\mathbf{v}) \text{ für alle } \mathbf{u}, \mathbf{v} \in \mathcal{V}$$

$\lambda \in \mathbb{R}$  zusammengefasst werden.

## Assoziierte Unterräume

Entsprechend zum Lemma A-6.10 kann man zu einer linearen Abbildung  $F : \mathcal{U} \rightarrow \mathcal{V}$  als Gruppenhomomorphismus die Teilmengen

- ▶ **Kern**  $F = \{\mathbf{u} \in \mathcal{U} : F(\mathbf{u}) = \mathbf{0} \in \mathcal{V}\} = F^{-1}(\{\mathbf{0}\})$  und
- ▶ **Bild**  $F = \{F(\mathbf{u}) \in \mathcal{V} : \mathbf{u} \in \mathcal{U}\} = F(\mathcal{U})$

konstruieren.

Es ergeben sich folgende Aussage über Null, Bild und Kern.

### Lemma B-5.2

- (i) *Jede lineare Abbildung bildet die Null von  $\mathcal{U}$  in die Null von  $\mathcal{V}$  ab.*
- (ii) *Die Bildmenge  $F(\mathcal{W}) \subset \mathcal{V}$  jedes Unterraums  $\mathcal{W} \subset \mathcal{U}$  ist ein Unterraum von  $\mathcal{V}$ .*
- (iii) *Der Kern von  $F$  ist ein linearer Unterraum von  $\mathcal{U}$ .*

# Beispiele linearer Abbildungen

## Beispiel B-5.3 (Kreuzprodukt)

Im  $\mathbb{R}^3$  definiert auch das Kreuzprodukt eine lineare Abbildung  $K_{\mathbf{u}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$\begin{aligned} \mathbf{v} \mapsto K_{\mathbf{u}}(\mathbf{v}) &= \mathbf{u} \times \mathbf{v} = \mathbf{u} \times \sum_{k=1}^3 v_k \mathbf{e}_k \\ &= v_1(\mathbf{u} \times \mathbf{e}_1) + v_2(\mathbf{u} \times \mathbf{e}_2) + v_3(\mathbf{u} \times \mathbf{e}_3) \\ &= \sum_{j,k=1}^3 u_j v_k \mathbf{e}_j \times \mathbf{e}_k = \sum_{i=1}^3 \left( \sum_{j,k=1}^3 \epsilon_{ijk} u_j v_k \right) \mathbf{e}_i \end{aligned}$$

Die letzte Schreibweise benutzt den sog. „Epsilon-Tensor“ mit den Werten

$$\epsilon_{ijk} = \begin{cases} 1 & \text{für die geraden Permutationen } (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2); \\ -1 & \text{für die ungeraden Permutationen } (i, j, k) = (3, 2, 1), (2, 1, 3), (1, 3, 2); \\ 0 & \text{sonst für alle Nichtpermutationen.} \end{cases}$$

## Beispiel B - 5.4 (Ableitung im Polynomraum)

Betrachte den Raum  $\mathcal{U} = \mathcal{V} = \mathcal{P}_n$  der Polynome mit reellen Koeffizienten vom Grad kleiner  $n = \dim(\mathcal{P}_n)$  in einer Variablen  $x$ . Dann ist die

### Differentiation

$$\mathbf{w} = \mathbf{F}(\mathbf{v}) = \mathbf{v}' = \frac{d\mathbf{v}}{dx}$$

eine lineare Operation, deren Ergebnis wiederum ein Polynom  $\mathbf{w} \in \mathcal{U}$  ist. Mit den Koeffizientendarstellungen

$$\mathbf{v} = \sum_{i=1}^n \nu_i x^{i-1} \quad \text{und} \quad \mathbf{w} = \sum_{i=1}^n \omega_i x^{i-1}$$

gilt  $\mathbf{w} = \mathbf{F}(\mathbf{v})$  genau dann, wenn

$$\omega_i = i \nu_{i+1} \quad \text{für} \quad i = 1 \dots n-1$$

und  $\omega_n = 0$ . Ein beliebiges  $\mathbf{w} \in \mathcal{V}$  ist also genau dann das Bildelement  $\mathbf{F}(\mathbf{v})$  für ein geeignetes  $\mathbf{v}$ , wenn der höchste Koeffizient  $\omega_n$  verschwindet.

## Folgerung B - 5.5

*Wir erhalten im Wertevorrat  $\mathcal{V} = \mathcal{P}_n$  den  $(n - 1)$ -dimensionalen Bildbereich*

$$\text{Bild}(\mathbf{F}) = \left\{ \sum_{i=1}^{n-1} \omega_i x^{i-1} \mid \omega_i \in \mathbb{R} \right\} = \mathcal{P}_{n-1} .$$

*Umgekehrt fällt der Koeffizient  $\nu_1$  der konstanten Funktion  $x^0 = 1$  bei der Differentiation weg, und wir haben den eindimensionalen Kern*

$$\text{Kern}(\mathbf{F}) = \{ \nu_1 x^0 \mid \nu_1 \in \mathbb{R} \} = \mathcal{P}_1 .$$

Mit anderen Worten, die Differentiation bildet genau diejenigen Funktionen auf die Nullfunktion ab, die konstant sind. Es gilt also

$$\dim(\text{Bild}(\mathbf{F})) = n - 1 = \dim(\text{Dom}(\mathbf{F})) - \dim(\text{Kern}(\mathbf{F})) ,$$

wobei  $\text{Dom}(\mathbf{F}) = \mathcal{U}$  den **Definitionsbereich** von  $\mathbf{F}$  bezeichnet. Diese **Dimensionsformel** gilt auch im allgemeinen Fall linearer Abbildungen.



# Lineare Unabhängigkeit

## Lemma B - 5.6

*Linear abhängige Vektoren werden unter einer linearen Abbildung in linear abhängige Bildvektoren transformiert.*

Umkehrung:

Sind die Bildvektoren  $\mathbf{v}_1 = F(\mathbf{u}_1), \mathbf{v}_2 = F(\mathbf{u}_2), \dots, \mathbf{v}_k = F(\mathbf{u}_k)$  linear unabhängig, so sind es auch die Urbildvektoren  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ .

## Lemma B - 5.7

*Sei  $F : \mathcal{U} \rightarrow \mathcal{V}$  linear und  $\mathbf{u}_1, \dots, \mathbf{u}_n$  eine Basis von  $\mathcal{U}$ . Dann kann aus den Bildvektoren  $\mathbf{v}_1 = F(\mathbf{u}_1), \mathbf{v}_2 = F(\mathbf{u}_2), \dots, \mathbf{v}_n = F(\mathbf{u}_n)$  eine Basis des Bildunterraumes  $F(\mathcal{U}) = \text{Bild } F$  ausgewählt werden.*

Folgerung:

Ist  $\mathcal{U}$  endlichdimensional, so auch  $\text{Bild } F = F(\mathcal{U})$ .

## Satz B-5.8 (Dimensionsformel)

Seien  $\mathcal{U}, \mathcal{V}$  endlichdimensional und  $F : \mathcal{U} \rightarrow \mathcal{V}$  linear. Dann gilt

$$\dim \mathcal{U} = \dim \text{Kern } F + \dim \text{Bild } F$$

### Beweis

Seien  $\mathbf{b}_1, \dots, \mathbf{b}_r \in \mathcal{V}$  eine Basis von  $\text{Bild } F$  und  $\mathbf{n}_1, \dots, \mathbf{n}_s \in \mathcal{U}$  eine Basis von  $\text{Kern } F$ .

Dann gibt es  $\mathbf{a}_1, \dots, \mathbf{a}_r \in \mathcal{U}$  mit  $\mathbf{b}_1 = F(\mathbf{a}_1), \dots, \mathbf{b}_r = F(\mathbf{a}_r)$ .

Zu zeigen:

Das System  $\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{n}_1, \dots, \mathbf{n}_s$  ist eine Basis von  $\mathcal{U}$ , d.h.

- (i) es erzeugt  $\mathcal{U}$  und
- (ii) es ist linear unabhängig.

Ist dies gezeigt, so folgt die Behauptung aus

$$\dim \mathcal{U} = r + s = \dim \text{Bild } F + \dim \text{Kern } F .$$

zu (i): Das System  $\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{n}_1, \dots, \mathbf{n}_s$  erzeugt  $\mathcal{U}$ .

Sei  $\mathbf{u} \in \mathcal{U}$  beliebig. Dann gibt es  $x_1, \dots, x_r \in \mathbb{R}$  mit

$$F(\mathbf{u}) = x_1 \mathbf{b}_1 + \dots + x_r \mathbf{b}_r = F(x_1 \mathbf{a}_1 + \dots + x_r \mathbf{a}_r).$$

Daher ist  $\mathbf{u} - (x_1 \mathbf{a}_1 + \dots + x_r \mathbf{a}_r) \in \text{Kern } F$  und es gibt  $y_1, \dots, y_s \in \mathbb{R}$  mit

$$\begin{aligned} \mathbf{u} - (x_1 \mathbf{a}_1 + \dots + x_r \mathbf{a}_r) &= y_1 \mathbf{n}_1 + \dots + y_s \mathbf{n}_s \\ \iff \mathbf{u} &= x_1 \mathbf{a}_1 + \dots + x_r \mathbf{a}_r + y_1 \mathbf{n}_1 + \dots + y_s \mathbf{n}_s. \end{aligned}$$

zu(ii): Das System  $\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{n}_1, \dots, \mathbf{n}_s$  ist linear unabhängig

Für jede Linearkombination  $\mathbf{0} = x_1 \mathbf{a}_1 + \dots + x_r \mathbf{a}_r + y_1 \mathbf{n}_1 + \dots + y_s \mathbf{n}_s$  gilt

$$\mathbf{0} = F(\mathbf{0}) = x_1 \mathbf{b}_1 + \dots + x_r \mathbf{b}_r + \mathbf{0} \implies x_1 = \dots = x_r = 0.$$

Dann ist aber

$$\mathbf{0} = \mathbf{0} + y_1 \mathbf{n}_1 + \dots + y_s \mathbf{n}_s \implies y_1 = \dots = y_s = 0.$$

# Abbildungseigenschaften

## Folgerung B - 5.9

*Eine lineare Abbildung  $F : \mathcal{U} \rightarrow \mathcal{V}$  endlichdimensionaler Vektorräume ist genau dann injektiv, wenn  $\text{Kern } F = \{0\}$  gilt, was wiederum genau dann der Fall ist, wenn*

$$\dim \mathcal{U} = \dim \text{Bild } F$$

*gilt.*

*$F$  ist surjektiv, wenn  $\text{Bild } F = \mathcal{V}$  gilt, hier also genau dann, wenn*

$$\dim \text{Bild } F = \dim \mathcal{V}$$

*gilt.*

*$F$  ist also bijektiv genau dann, wenn*

$$\dim \mathcal{U} = \dim \text{Bild } F = \dim \mathcal{V}$$

*gilt.*

## Vektorgleichungen

Sei  $\mathbf{F} : \mathcal{U} \rightarrow \mathcal{V}$  linear und ein Vektor  $\mathbf{w} \in \mathcal{V}$  gegeben. Dann interessiert man sich für die Lösungen  $\mathbf{u} \in \mathcal{U}$  der **Vektorgleichung**

$$\mathbf{F}(\mathbf{u}) = \mathbf{w} .$$

Diese ist genau dann **lösbar**, wenn  $\mathbf{w} \in \text{Bild } \mathbf{F}$ . Für zwei Lösungen  $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$  gilt

$$\mathbf{0} = \mathbf{w} - \mathbf{w} = \mathbf{F}(\mathbf{u}_2) - \mathbf{F}(\mathbf{u}_1) = \mathbf{F}(\mathbf{u}_2 - \mathbf{u}_1) ,$$

d.h. die Differenz zweier Lösungen ist ein Kernvektor,  
 $\mathbf{v} = \mathbf{u}_2 - \mathbf{u}_1 \in \text{Kern } \mathbf{F}$ . Ist umgekehrt  $\mathbf{u}$  eine Lösung und  $\mathbf{v} \in \text{Kern } \mathbf{F}$ , so gilt

$$\mathbf{F}(\mathbf{u} + \mathbf{v}) = \mathbf{F}(\mathbf{u}) + \mathbf{F}(\mathbf{v}) = \mathbf{w} + \mathbf{0} = \mathbf{w} ,$$

die Lösungsmenge ist somit die um die **spezielle inhomogene Lösung**  $\mathbf{u}$  verschobene **homogene Lösungsmenge**, wie der Kernunterraum auch genannt wird.

## Lemma B-5.10 (Restklassen bezüglich eines Unterraums)

$\mathcal{W} \subset \mathcal{U}$  linearer Unterraum impliziert, dass

$$\mathbf{u} \sim \mathbf{w} \iff \mathbf{u} - \mathbf{w} \in \mathcal{W} \iff \exists \mathbf{v} \in \mathcal{W} : \mathbf{u} = \mathbf{w} + \mathbf{v}$$

eine Äquivalenzrelation ist.

Die entsprechenden Äquivalenzklassen

$$[\mathbf{u}] = \{\mathbf{w} \in \mathcal{U} : \mathbf{w} \sim \mathbf{u}\}$$

bilden einen Vektorraum  $\mathcal{U}/\mathcal{W}$  bezüglich der Operationen

$$[\mathbf{u}] + [\mathbf{w}] = [\mathbf{u} + \mathbf{w}] \quad \text{und} \quad \lambda[\mathbf{u}] = [\lambda\mathbf{u}].$$

Voneinander verschiedene Restklassen sind disjunkt. Elemente innerhalb einer Restklasse unterscheiden sich um ein Element von  $\mathcal{W}$ . D.h. die Restklassen bilden eine „Blätterung“ von  $\mathcal{U}$  in parallele Ebenen, die sich als Parallelverschiebungen des Unterraums  $\mathcal{W}$  ergeben.

## Folgerung B-5.11

*Sei  $F : \mathcal{U} \rightarrow \mathcal{V}$  linear. Die Restklassen von  $\mathcal{U} / \text{Kern } F$  entsprechen den Urbildmengen von  $F$ , denn*

$$\begin{aligned} [\mathbf{u}] &= \{\mathbf{w} \in \mathcal{U} : \mathbf{w} - \mathbf{u} \in \text{Kern } F\} \\ &= \{\mathbf{w} \in \mathcal{U} : F(\mathbf{w}) = F(\mathbf{u})\} = F^{-1}\left(\{F(\mathbf{u})\}\right) \end{aligned}$$

*Daraus folgt die Isomorphie*

$$\mathcal{U} / \text{Kern } F \simeq \text{Bild } F ,$$

*dabei ist die Restklasse  $[\mathbf{u}]$  eindeutig dem Vektor  $F(\mathbf{u})$  zugeordnet.*

Von besonderem Interesse sind lineare Abbildungen  $\mathbf{F} : \mathcal{U} \rightarrow \mathcal{V}$ , die **regulär** sind in dem Sinne, dass Lösungen  $\mathbf{u} \in \mathcal{U}$  von  $\mathbf{F}(\mathbf{u}) = \mathbf{w}$ , wenn sie überhaupt existieren, eindeutig sind.

Dies ist die Eigenschaft der Injektivität von  $\mathbf{F}$ , was äquivalent zu  $\text{Kern } \mathbf{F} = \{\mathbf{0}\}$  ist.

Mit anderen Worten: die Lösung der **inhomogenen** Vektorgleichung  $\mathbf{F}(\mathbf{v}) = \mathbf{w}$  ist eindeutig genau dann, wenn die entsprechende **homogene** Gleichung  $\mathbf{F}(\mathbf{u}) = \mathbf{0}$  nur die triviale Lösung  $\mathbf{u} = \mathbf{0}$  hat.

Im regulären Falle bezeichnet man die Zuordnung des Urbildes  $\mathbf{u} \in \mathcal{U}$  zum gegebenen Bildvektor  $\mathbf{w} = \mathbf{F}(\mathbf{v}) \in \text{Bild } \mathbf{F} \subset \mathcal{V}$  als die **Umkehrabbildung** oder die **inverse Abbildung**

$$\mathbf{F}^{-1} : \text{Bild}(\mathbf{F}) \rightarrow \text{Dom}(\mathbf{F}) .$$



Falls sie überhaupt existiert, ist die Inverse einer linearen Abbildung auch immer linear, d.h. es gilt für  $\mathbf{v}, \mathbf{w} \in \text{Bild } \mathbf{F} \subset \mathcal{V}$

$$\begin{aligned}\mathbf{F}^{-1}(\mathbf{v} + \mathbf{w}) &= \mathbf{F}^{-1}(\mathbf{v}) + \mathbf{F}^{-1}(\mathbf{w}) \\ \mathbf{F}^{-1}(\lambda \mathbf{w}) &= \lambda \mathbf{F}^{-1}(\mathbf{w}) .\end{aligned}$$

Das Auffinden von  $\mathbf{v} = \mathbf{F}^{-1}(\mathbf{w})$  für gegebenes  $\mathbf{w}$  bezeichnet man auch als **Lösen** der Vektorgleichung  $\mathbf{F}(\mathbf{v}) = \mathbf{w}$ . Stellt man eine Vektorgleichung bezüglich gewählter Basen in  $\mathcal{U}$  und  $\mathcal{V}$  komponentenweise dar, so ergibt sich ein **System skalarer Gleichungen**, ein **lineares Gleichungssystem**.

Das effektive und genaue Lösen von linearen Gleichungssystemen bei gleichzeitiger Untersuchung ihrer Regularität ist nach wie vor eine zentrale Aufgabe im „**Wissenschaftlichen Rechnen**“. Dabei werden Ergebnisse und Methoden der Informatik und Numerischen Mathematik eingesetzt, um Systeme mit Tausenden oder sogar Millionen von Unbekannten zumindest näherungsweise zu lösen.

Eine zweite, für Anwendungen sehr wichtige Aufgabe ist die Lösung von **Eigenwertproblemen**. Ist eine lineare Abbildung  $F : \mathcal{V} \rightarrow \mathcal{V}$  gegeben, so werden Vektoren  $\mathbf{v} \in \mathcal{V}$  gesucht, deren Bildvektor zu  $\mathbf{v}$  parallel ist. D.h. es sind Paare aus einem Vektor  $\mathbf{v}$  und einem Skalar  $\lambda$  zu bestimmen mit der Eigenschaft

$$\mathbf{F}(\mathbf{v}) = \lambda \mathbf{v} \quad \text{und} \quad \mathbf{v} \neq \mathbf{0}.$$

Gilt diese Gleichung, so nennt man  $\lambda$  einen **Eigenwert** und  $\mathbf{v}$  einen **Eigenvektor** der linearen Abbildung  $\mathbf{F}$ .

Die Lösung des Eigenwertproblems wird dadurch erschwert, dass die Eigenwerte komplex sein können und  $\mathbf{F}$  deswegen auf einer komplexen Erweiterung von  $\mathcal{V}$  definiert werden muss.

Die praktische Lösung von linearen Gleichungssystemen und Eigenwertproblemen verlangt die komponentenweise Darstellung linearer Abbildungen mittels Matrizen.

## B-6 Matrizen und ihre Algebra

Ein **Tupel**  $a = (a_1, a_2, \dots, a_m)$  kann als Abbildung  $a : J \rightarrow X$  der endlichen Indexmenge  $J = \{1, 2, \dots, m\}$  in irgendeine Bildmenge  $X$  definiert werden. Dabei werden die Funktionswerte im „alten Stil“ als  $a_j = a(j)$  notiert.

Man kann beliebige endliche Mengen als Indexmengen verwenden, z.B. auch  $J = \{3, 7, 19, 125, 765\}$ . Weiter kann man die Elemente der Indexmenge mit einer Struktur versehen, z.B. indem man geordnete Paare aus einem **kartesischen Produkt**  $J \times K$  zweier endlicher Mengen  $J$  und  $K$  als Indizes verwendet.

### Definition B-6.1 (Matrix)

Sei  $X$  eine beliebige Menge,  $m, n \in \mathbb{N}$ . Als  $(m \times n)$ -**Matrix mit Einträgen bzw. Komponenten aus**  $X$  wird dann jede Abbildung

$$A : \{1, 2, \dots, m\} \times \{1, 2, \dots, n\} \rightarrow X$$

bezeichnet. Die Menge aller  $(m \times n)$ -Matrizen wird mit  $X^{m \times n}$  notiert.

## Matrixnotation

Die Werte dieser als  $(m \times n)$ -Matrix bezeichneten Abbildung  $A : \{1, 2, \dots, m\} \times \{1, 2, \dots, n\} \rightarrow X$  werden **Matriceinträge** oder **-komponenten** genannt und als  $a_{jk} = A(j, k)$  notiert.

Die Matrix selbst wird schematisch als

$$A = (a_{jk})_{j=1\dots m}^{k=1\dots n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

dargestellt.

$\mathbb{R}^{m \times n}$  bezeichnet also den Raum der  $(m \times n)$ -Matrizen mit reellwertigen Einträgen (vgl. Definition ??). In der schematischen Darstellung hat eine solche Matrix  $m$  Zeilen und  $n$  Spalten. Der erste Index in einem Eintrag der Matrix gibt immer die Zeile, der zweite die Spalte des Eintrages an.

## Zeilen- und Spaltenvektoren

Matrizen, die nur aus einer Spalte bestehen, z.B. Elemente von  $\mathbb{R}^{n \times 1}$ , bezeichnet man als **Spaltenvektoren** der Länge  $n$ . Der Tupelraum  $\mathbb{R}^n$  wird üblicherweise mit dem Spaltenvektorraum  $\mathbb{R}^{n \times 1}$  identifiziert.

Analog sind Matrizen mit nur einer Zeile, z.B. aus  $\mathbb{R}^{1 \times n}$  **Zeilenvektoren** der Länge  $n$ .

Jede Matrix kann als Zeilenvektor aus Spaltenvektoren

$$A = (a_{1:m,1} \ a_{2:m,1} \ \cdots \ a_{1:m,n}) \quad \text{mit} \quad a_{1:m,k} = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}$$

oder umgekehrt als Spaltenvektor aus Zeilenvektoren

$$A = \begin{pmatrix} a_{1,1:n} \\ \vdots \\ a_{m,1:n} \end{pmatrix} \quad \text{mit} \quad a_{j,1:n} = (a_{j1} \ a_{j2} \ \cdots \ a_{jm})$$

verstanden werden.

## Lemma B - 6.2

*Für alle  $m, n \in \mathbb{N}$  ist der Matrixraum  $\mathbb{R}^{m \times n}$  ein reeller Vektorraum. Addition wie skalare Multiplikation sind komponentenweise erklärt.*

Aus einer reellen Zahl  $\lambda \in \mathbb{R}$  und Matrizen  $A = (a_{jk})_{j=1 \dots m}^{k=1 \dots n}$  und  $B = (b_{jk})_{j=1 \dots m}^{k=1 \dots n}$  werden Summe und Vielfaches gebildet als

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

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

## Matrixmultiplikation

Zwei Matrizen  $A \in \mathbb{R}^{m_1 \times n_1}$  und  $B \in \mathbb{R}^{m_2 \times n_2}$  heißen (bzgl. der Multiplikation) **kompatibel**, wenn die **inneren Dimensionen** übereinstimmen, d.h.  $n_1 = m_2$  gilt.

### Definition B-6.3

Es seien  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{n \times p}$  zwei kompatible Matrizen. Deren Matrixprodukt  $C = A \cdot B \in \mathbb{R}^{m \times p}$  ist dann komponentenweise definiert als

$$c_{j\ell} = \sum_{k=1}^n a_{jk} b_{k\ell} .$$

Es werden also die Elemente des  $j$ -ten **Zeilenvektors** von  $A$  mit denen des  $\ell$ -ten **Spaltenvektors** von  $B$  gliedweise multipliziert und dann die Produkte aufaddiert.

**Zeile mal Spalte**

## Eigenschaften des Matrixprodukts

- (i) Die Matrixmultiplikation ist **assoziativ**, d.h. für beliebige kompatible reelle Matrizen  $A, B, C$  gilt

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

- (ii) Die Matrixmultiplikation ist **distributiv**, d.h. mit geeignet kompatiblen Matrizen  $A, B, C, D$  gilt

$$(A + B)C = AC + BC$$

$$A(C + D) = AC + AD .$$

- (iii) Die Matrixmultiplikation ist **linear**, d.h. für beliebige kompatible reelle Matrizen  $A, B, C$  und reelle  $\lambda \in \mathbb{R}$  gilt

$$(A + \lambda B)C = AC + \lambda BC .$$

### Folgerung B - 6.4

Die reellen quadratischen  $(n \times n)$ -Matrizen bilden den **Matrixring**  $\mathbb{R}^{n \times n}$ .



## Definition B-6.5 (Identitätsmatrix)

Bezüglich der Multiplikation von Matrizen gibt es ein neutrales Element, nämlich die **Einheits-** oder **Identitätsmatrix**

$$I = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Der die Größe der Matrix angegebende Index  $n$  kann wegfallen, wenn er sich aus dem Zusammenhang ergibt. Es gilt nun insbesondere

$$I_m A = A = A I_n \quad \text{für} \quad A \in \mathbb{R}^{m \times n}.$$

## Beispiel B-6.6

Man betrachte die beiden  $(3 \times 3)$  Matrizen

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

wobei  $\sigma = 1/\sqrt{2}$  ist.

Bezüglich der kartesischen Basis des dreidimensionalen Anschauungsraumes

beschreibt  $A$  die **Spiegelung** aller Vektoren  $\mathbf{v} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$  an der diagonalen Fläche  $y = z$ .

$B$  beschreibt bezüglich der kartesischen Basis eine **Achtel-Drehung entgegen dem Uhrzeigersinn** um die  $z$ -Achse  $\mathbf{e}_3$  (Achtung: Rechtssystem!!).

## Fortsetzung Beispiel

Wird nun *zuerst rotiert und dann reflektiert*, so ergibt sich die Matrix

$$\begin{pmatrix} \sigma & \sigma & 0 \\ 0 & 0 & 1 \\ -\sigma & \sigma & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \sigma & \sigma & 0 \\ -\sigma & \sigma & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Hier ergab sich zum Beispiel das Element in der dritten Zeile und zweiten Spalte des Produktes als  $(0, 1, 0) \cdot (\sigma, \sigma, 0)^T = 0 \cdot \sigma + 1 \cdot \sigma + 0 \cdot 0 = \sigma$ .

Tauscht man jedoch die Reihenfolge der Faktoren aus, so erhält man die Matrix

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

Diese Matrix beschreibt die *Hintereinanderausführung der Spiegelung und dann der Drehung*, was zu unterschiedlichen Ergebnissen führt.

Wie das Beispiel zeigt, ist die Matrixmultiplikation **nicht kommutativ**.

## Lineare Abbildungen der Räume $\mathbb{R}^n$

Sei  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear. Dann kann  $F$  durch die Bilder

$$\mathbf{a}_1 = \mathbf{F}(\mathbf{e}_1), \mathbf{a}_2 = \mathbf{F}(\mathbf{e}_2), \dots, \mathbf{a}_n = \mathbf{F}(\mathbf{e}_n)$$

der kanonischen Basisvektoren eindeutig charakterisiert werden:

$$\begin{aligned}\mathbf{y} = \mathbf{F}(\mathbf{x}) &= \mathbf{F}(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n) \\ &= x_1\mathbf{F}(\mathbf{e}_1) + x_2\mathbf{F}(\mathbf{e}_2) + \dots + x_n\mathbf{F}(\mathbf{e}_n) \\ &= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n .\end{aligned}$$

Mit  $\mathbf{a}_k = a_{1:m,k} = (a_{1,k}, a_{2,k}, \dots, a_{m,k})$  ergibt sich für die  $j$ -te Komponente von  $\mathbf{y}$

$$y_j = a_{j,1}x_1 + a_{j,2}x_2 + \dots + a_{j,n}x_n .$$

Schreiben wir nun  $\mathbf{x} = (x_{1,1}, x_{2,1}, \dots, x_{n,1})^T \in \mathbb{R}^{n \times 1}$  und analog  $\mathbf{y} = (y_{1,1}, y_{2,1}, \dots, y_{m,1})^T \in \mathbb{R}^{m \times 1}$ , sowie  $A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \in \mathbb{R}^{m \times n}$ , so sind die Vektoren über das Matrix-Spaltenvektor-Produkt verbunden,  $\mathbf{y} = A\mathbf{x}$ .

## Lemma B - 6.7

*Es gibt eine eindeutige Beziehung zwischen der Menge  $\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$  der linearen Abbildungen von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  und*

*der Menge  $\mathbb{R}^{m \times n}$  der reellen  $(m \times n)$ -Matrizen.*

Jeder Matrix  $A \in \mathbb{R}^{m \times n}$  wird die durch das Matrix-Vektor-Produkt definierte lineare Abbildung  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  zugeordnet,  $A(x) = Ax$ .

Jeder Matrix  $A \in \mathbb{R}^{m \times n}$  kann also ein Kern und ein Bild zugeordnet werden, man nennt

$$\text{Rang}(A) = \dim \text{Bild}(A) \quad \text{und} \quad \text{Defekt}(A) = \dim \text{Kern}(A) .$$

es gilt die **Dimensionsformel**

$$n = \text{Rang}(A) + \text{Defekt}(A) .$$

## Verknüpfung linearer Abbildungen

Ist  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear mit Matrix  $A \in \mathbb{R}^{m \times n}$  und  $G : \mathbb{R}^p \rightarrow \mathbb{R}^n$  linear mit Matrix  $B \in \mathbb{R}^{n \times p}$ , so gilt für deren Verknüpfung und  $\mathbf{x} \in \mathbb{R}^p$

$$\begin{aligned}(F \circ G)(\mathbf{x}) &= \sum_{\ell=1}^p F(G(\mathbf{e}_\ell)) x_\ell = \sum_{\ell=1}^p F(\mathbf{b}_\ell) x_\ell \\&= \sum_{k=1}^n \sum_{\ell=1}^p F(\mathbf{e}_k) b_{k,\ell} x_\ell = \sum_{k=1}^n \sum_{\ell=1}^p \mathbf{a}_k b_{k,\ell} x_\ell \\&= \sum_{j=1}^m \sum_{k=1}^n \sum_{\ell=1}^p \mathbf{e}_j a_{j,k} b_{k,\ell} x_\ell = \sum_{j=1}^m \sum_{\ell=1}^p \mathbf{e}_j \left( \sum_{k=1}^n a_{j,k} b_{k,\ell} \right) x_\ell .\end{aligned}$$

Die Matrix  $C \in \mathbb{R}^{m \times p}$  der verknüpften Abbildung ergibt sich also als Matrixprodukt  $C = AB$ .

## Basisabbildung

Sei  $\mathcal{V}$  ein reeller Vektorraum,  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathcal{V}$  linear. Dann ist auch hier  $\mathbf{F}$  durch die Bilder der kanonischen Basisvektoren eindeutig bestimmt,

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{e}_1)x_1 + \cdots + \mathbf{F}(\mathbf{e}_n)x_n = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n .$$

Bilder von  $\mathbf{F}$  sind Linearkombinationen von  $\mathbf{v}_1, \dots, \mathbf{v}_n$ .  $\mathbf{F}$  ist genau dann

- ▶ injektiv, wenn  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängig sind,
- ▶ surjektiv, wenn  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  ein Erzeugendensystem ist und
- ▶ bijektiv, wenn  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  eine Basis ist.

Im letzten Fall wird  $\mathbf{F}$  Basisabbildung und die inverse Abbildung

$$\mathbf{F}^{-1} : \mathcal{V} \rightarrow \mathbb{R}^n$$

auch **Koordinatenabbildung** zur Basis  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  genannt.

## Matrixdarstellung allgemeiner linearer Abbildungen

Seien  $\mathcal{U}, \mathcal{V}$  Vektorräume mit Basen  $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathcal{U}$  und  $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_m \in \mathcal{V}$ .  
Die Basisabbildungen seien mit  $B : \mathbb{R}^n \rightarrow \mathcal{U}$  und  $\tilde{B} : \mathbb{R}^m \rightarrow \mathcal{V}$  bezeichnet.

Sei  $\mathbf{F} : \mathcal{U} \rightarrow \mathcal{V}$  linear. Dann ist  $\tilde{B}^{-1} \circ \mathbf{F} \circ B : \mathbb{R}^n \rightarrow \mathbb{R}^m$  ebenfalls linear  
und durch eine Matrix  $A$  darstellbar. Mit dieser gilt

$$\mathbf{F} = \tilde{B} \circ A \circ B^{-1} .$$

Für einen beliebigen Vektor  $\mathbf{u} = x_1 \mathbf{b}_1 + \dots + x_n \mathbf{b}_n \in \mathcal{U}$  gilt

$$\mathbf{v} = \mathbf{F}(\mathbf{u}) = x_1 \mathbf{F}(\mathbf{b}_1) + \dots + x_n \mathbf{F}(\mathbf{b}_n) = \sum_{j=1}^m \sum_{k=1}^n \tilde{b}_j a_{j,k} x_k$$

Die Koordinaten  $\mathbf{y} = (y_1, \dots, y_m)$  des Bildvektors  $\mathbf{v}$  ergeben sich als  
Matrix-Vektor-Produkt von  $A$  mit dem Koordinatenvektor

$\mathbf{x} = (x_1, \dots, x_n)$  von  $\mathbf{u}$ ,

$$\mathbf{y} = A\mathbf{x} .$$



## Beispiel B - 6.8

Bezüglich der monomialen Basis hat die schon im Beispiel B - 5.4 erwähnte Abbildung durch **Differentiation** in  $\mathcal{P}_n$  für  $n = 5$  die Matrix-Darstellung

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

Diese Matrix ergibt sich für  $i = 1, \dots, 5$  aus der Grundbeziehung

$$F(\mathbf{v}_i) = \mathbf{v}'_i = (i - 1) \mathbf{v}_{i-1} \quad \text{da} \quad \mathbf{v}_i = x^{i-1},$$

wobei hier  $\mathcal{W} = \mathcal{V}$  und deshalb  $\mathbf{w}_i = \mathbf{v}_i$ .

## Definition B-6.9 (Transposition)

Eine einfache, aber wichtige Operation auf Matrizen ist die

**Transposition**, die eine  $(m \times n)$  Matrix  $A$  in eine  $(n \times m)$  Matrix  $B = A^T$  überführt. Hierbei gilt  $\beta_{ij} = \alpha_{ji}$ , so dass in Matrixschreibweise

$$A^T = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{m2} \\ \dots\dots\dots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{mn} \end{pmatrix} = (\beta_{ij})_{i=1\dots n}^{j=1\dots m}.$$

### Bemerkung:

Nur die Diagonalelemente  $(\alpha_{ii})_{i=1\dots \min(m,n)}$  bleiben bei der Transposition unverändert, die anderen Elemente tauschen den Platz mit ihrem Gegenüber auf der anderen Seite der Diagonalen.

## Lemma B - 6.10 (Transpositionsregeln)

*Man kann sich leicht davon überzeugen, daß die folgenden Regeln für das Transponieren gelten:*

$$(A^T)^T = A$$

$$(A + B)^T = A^T + B^T$$

$$(\lambda A)^T = \lambda A^T$$

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

### Bemerkung:

Die Transposition ist also eine lineare Abbildung von  $\mathbb{R}^{m \times n}$  nach  $\mathbb{R}^{n \times m}$  und als solche sogar ihre eigene Inverse.

Die letzte Gleichung bedeutet, dass die Transponierte eines Produktes gleich dem Produkt der transponierten Faktoren in umgekehrter Reihenfolge ist. Hierbei müssen wir natürlich wieder davon ausgehen, dass die Faktoren bezüglich der Produktbildung kompatibel sind, woraus die Kompatibilität der Transponierten folgt.

## Dyadisches Produkt

Seien  $\mathbf{a} = (\alpha_j)_{j=1\dots n}$ ,  $\mathbf{b} = (\beta_j)_{j=1\dots n} \in \mathbb{R}^n = \mathbb{R}^{n \times 1}$  Spaltenvektoren. Dann ist  $\mathbf{a}^T = (\alpha_1, \alpha_2, \dots, \alpha_n)$  ein Zeilenvektor. Das Produkt dieses Zeilenvektors  $\mathbf{a}^T$  mit dem Spaltenvektor  $\mathbf{b}$  ergibt

$$\mathbf{a}^T \mathbf{b} = \langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{b}^T \mathbf{a} = \sum_{j=1}^n \alpha_j \beta_j \in \mathbb{R} = \mathbb{R}^{1 \times 1}.$$

Diese  $1 \times 1$  Matrix kann man also als Skalar mit dem inneren Produkt zwischen  $\mathbf{a}$  und  $\mathbf{b}$  identifizieren.

Wechselt man jedoch die Reihenfolge der Faktoren, so ergibt sich (auch für  $\mathbf{b} \in \mathbb{R}^m$ ,  $m \neq n$ ) die Matrix

$$\mathbf{b} \mathbf{a}^T = (b_i a_j)_{j=1\dots n}^{i=1\dots m} \in \mathbb{R}^{m \times n}$$

Diese nennt man das **dyadische Produkt** von  $\mathbf{a}$  und  $\mathbf{b}$ .

## Verbilligte Produkte

Normalerweise kostet für  $A \in \mathbb{R}^{m \times n}$  die Berechnung des Produktes  $A\mathbf{v}$  mit einem Vektor  $\mathbf{v} \in \mathbb{R}^n$  genau  $m \cdot n$  skalare Multiplikationen. Ist jedoch  $A = \mathbf{b}\mathbf{a}^T$  ein dyadisches Produkt, so berechnet man viel billiger

$$A\mathbf{v} = (\mathbf{b}\mathbf{a}^T)\mathbf{v} = \mathbf{b}(\mathbf{a}^T\mathbf{v}).$$

Beachte, dass  $\mathbf{b}(\mathbf{a}^T\mathbf{v})$  durch Bildung des inneren Produktes  $\mathbf{a}^T\mathbf{v} = \mathbf{a} \cdot \mathbf{v}$  und seine anschließende Multiplikation mit  $\mathbf{b}$  nur  $n + m$  skalare Multiplikationen verlangt. Demgegenüber kostet alleine die explizite Berechnung des dyadischen Produktes  $\mathbf{b}\mathbf{a}^T$  genau  $m \cdot n$  Multiplikationen.

Entsprechend berechnet man das Produkt mit einer Matrix  $V \in \mathbb{R}^{n \times p}$  als

$$(\mathbf{b}\mathbf{a}^T)V = \mathbf{b}(\mathbf{a}^TV) = \mathbf{b}(V^T\mathbf{a})^T$$

Die Produktbildung  $\mathbf{b}(V^T\mathbf{a})^T$  kostet nur  $(m + n) \cdot p$  skalare Multiplikationen während die Berechnung in der Form  $(\mathbf{b}\mathbf{a}^T)V$  mehr als  $m \cdot n \cdot p$  solche Operationen verlangt.

## Quadratische Matrix

$$A \in \mathbb{R}^{n \times n} \Rightarrow A^T \in \mathbb{R}^{n \times n}$$

Eine Matrix, deren Zeilenzahl gleich ihrer Spaltenzahl ist, heißt **quadratisch**. Alle linearen Abbildungen eines Raumes in sich selbst werden durch quadratische Matrizen beschrieben.

## Symmetrische Matrix

$$A^T = A \in \mathbb{R}^{n \times n}$$

Quadratische Matrizen, die bezüglich der Transposition invariant sind, heißen **symmetrisch**. Diese bilden einen Unterraum von  $\mathbb{R}^{n \times n}$ . Dieser Unterraum hat die Dimension  $n(n+1)/2$ , da man lediglich die  $n$  Elemente in der Diagonale und entweder die  $n(n-1)/2$  Elemente darüber oder die gleiche Zahl darunter frei wählen kann.

Matrizen der Form  $A^T A$  und  $A + A^T$  sind immer symmetrisch.

## Schiefsymmetrische Matrix

$$A^T = -A \in \mathbb{R}^{n \times n}$$

Quadratische Matrizen mit dieser Eigenschaft heißen **schiefsymmetrisch**.  
Wie wir später sehen werden, sind alle ihre Eigenwerte rein imaginär.

Für jede quadratische Matrix gilt

$$A = \underbrace{\frac{1}{2}(A + A^T)}_{\text{symmetrisch}} + \underbrace{\frac{1}{2}(A - A^T)}_{\text{schiefsymmetrisch}} .$$

Diese additive Zerlegung ist allerdings nicht sehr nützlich in Bezug auf die Eigenwerte, da diese in stark nichtlinearer Weise von der Matrix abhängen.

## Dreiecksmatrix

Falls für  $A = (\alpha_{ij}) \in \mathbb{R}^{n \times n}$

$$i > j \Rightarrow \alpha_{ij} = 0$$

gilt, so dass

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \alpha_{nn} \end{pmatrix},$$

dann nennt man  $A$  eine **obere Dreiecksmatrix**.

Analog definiert man auch die **untere Dreiecksmatrix**, deren oberhalb der Hauptdiagonale stehenden Elemente Null sind.



## Diagonale Matrizen

$A \in \mathbb{R}^{n \times n}$  heißt **diagonal**, wenn  $i \neq j \Rightarrow \alpha_{ij} = 0$  gilt, also

$$A = \begin{pmatrix} \alpha_{11} & 0 & \cdots & 0 \\ 0 & \alpha_{22} & \cdots & 0 \\ & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \alpha_{nn} \end{pmatrix}.$$

Man schreibt dann kurz  $A = \text{diag}(\alpha_i)_{i=1 \dots n}$ .

Insbesondere gilt

$$I = \text{diag}(1)_{i=1 \dots n}.$$

Summen und Produkte von diagonalen Matrizen sind wiederum diagonal:

$$A = \text{diag}(\alpha_i)_{i=1 \dots n}$$

$$B = \text{diag}(\beta_i)_{i=1 \dots n}$$

$$\Rightarrow$$

$$A + B = \text{diag}(\alpha_i + \beta_i)_{i=1 \dots n}$$

$$AB = \text{diag}(\alpha_i \beta_i)_{i=1 \dots n}$$

## Orthogonale Matrizen

$A \in \mathbb{R}^{n \times n}$  heißt **orthogonal**, falls

$$A^T A = I = A A^T$$

wobei sich zeigen läßt, daß die zweite Identität aus der ersten folgt.

Bezeichnet man mit  $\mathbf{a}_j = (\alpha_{ij})_{i=1 \dots n}$  den  $j$ -ten Spaltenvektor von  $A$ , so ist die Bedingung  $A^T A = I$  äquivalent zu

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \begin{cases} 0 & \text{falls } i \neq j \\ 1 & \text{falls } i = j \end{cases}$$

Das heißt: Die Matrix  $A$  ist genau dann orthogonal, wenn ihre Spaltenvektoren eine orthonormale Basis von  $\mathbb{R}^n$  bilden.

Da mit  $A$  auch  $A^T$  orthogonal ist, gilt dasselbe für die Zeilen von  $A$ , die ja die Spalten von  $A^T$  sind.

## Produkt orthogonaler Matrizen

Für zwei orthogonale Matrizen  $A$  und  $B$  ist jeweils auch deren Produkt orthogonal, da

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

Die Summe von orthogonalen Matrizen hat im allgemeinen nicht diese Eigenschaft. So ist zum Beispiel mit  $A$  auch  $-A$  orthogonal, aber deren Summe, die Nullmatrix  $A - A = 0$ , sicherlich nicht.

Orthogonale Matrizen bilden eine Gruppe  $O(n)$  mit der Einheitsmatrix als neutralem Element und inversen Matrizen  $A^{-1} = A^T$ .

## Beispiel B-6.11 (Drehungen in der Ebene)

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

$$A^T A = \begin{pmatrix} \cos(\varphi)^2 + \sin(\varphi)^2 & \cos(\varphi) \sin(\varphi) \cdot (1 - 1) \\ \sin(\varphi) \cos(\varphi) \cdot (1 - 1) & \cos(\varphi)^2 + \sin(\varphi)^2 \end{pmatrix} = I$$

## B-7 Lösung linearer Gleichungssysteme

### Lineare Systeme

Für eine lineare Abbildung

$$F : \mathcal{V} = \text{Span}\{\mathbf{v}_j\}_{j=1\dots n} \rightarrow \mathcal{W} = \text{Span}\{\mathbf{w}_i\}_{i=1\dots m}$$

und eine vorgegebene "Rechte Seite"  $\mathbf{w} = \sum_{i=1}^m b_i \mathbf{w}_i$  mit  $b_i \in \mathbb{R}$  findet man ein  $\mathbf{v} = \sum_{j=1\dots n} x_j \mathbf{v}_j$  mit  $F(\mathbf{v}) = \mathbf{w}$  durch Lösen des Gleichungssystems

$$\begin{array}{ccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \dots & + & \alpha_{1n}x_n & = & b_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + & \dots & + & \alpha_{2n}x_n & = & b_2 \\ & & & & \dots & & & & \\ \alpha_{i1}x_1 & + & \alpha_{i2}x_2 & + & \dots & + & \alpha_{in}x_n & = & b_i \\ & & & & \dots & & & & \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \dots & + & \alpha_{mn}x_n & = & b_m \end{array}$$

## Matrix–Vektor–Schreibweise

Äquivalenterweise ergibt sich in Matrix–Vektor–Schreibweise

$$\mathbf{A}\mathbf{x} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ \alpha_{21} & \dots & \alpha_{2j} & \dots & \alpha_{2n} \\ & \dots & \dots & \dots & \dots \\ \alpha_{m1} & \dots & \alpha_{mj} & \dots & \alpha_{mn} \end{pmatrix} \mathbf{x} = \mathbf{b}$$

wobei  $\mathbf{x} = (x_1, \dots, x_j, \dots, x_n)^T$  und  $\mathbf{b} = (b_1, \dots, b_i, \dots, b_m)^T$  sind.

Man bezeichnet das lineare System von  $m$  Gleichungen in  $n$  Unbekannten als

<b>unterbestimmt</b>	wenn $m < n$
<b>quadratisch</b>	wenn $m = n$
<b>überbestimmt</b>	wenn $m > n$

Im Allgemeinen haben **unterbestimmte** Gls. unendlich viele und **überbestimmte** Gls. keine Lösung. **Aber:** Für spezielle Gls. wird diese Regel verletzt, es gibt unterbestimmte Gls. ohne und überbestimmte Gls. mit Lösung.

## Definition B-7.1 (Regularität)

Eine Abbildung  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  und entsprechende Matrizen  $A$  heißen regulär, falls

$$A\mathbf{x} = F(\mathbf{x}) = 0 \quad \text{g.d.w.} \quad \mathbf{x} = 0,$$

andernfalls heißen sie singulär.

## Lemma B-7.2

*Falls  $A$  regulär ist, dann hat  $A\mathbf{x} = \mathbf{b}$  genau eine eindeutige Lösung für jedes  $\mathbf{b}$ .*

Ein Kriterium, ob eine Matrix regulär oder singulär ist, liefert die im Abschnitt **B-9** eingeführte **Determinante**  $\det(A)$ .

Wünschenswerte Lösungsalgorithmen prüfen die Regularität und liefern entweder die eindeutige Lösung oder Singularitätsbeschreibungen.

## Lösung Linearer Gleichungssysteme in Spezialfällen

Ist  $A$  eine Orthogonal-, Diagonal- oder Dreiecksmatrix (das sind diejenigen, deren Struktur sich auf das Produkt überträgt), so lassen sich die entsprechenden linearen Systeme  $A\mathbf{x} = \mathbf{b}$  relativ leicht lösen.

### Lemma B-7.3 (Lösung orthogonaler Systeme)

*Falls  $A$  orthogonal ist, gilt:*

$$A\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad A^T A\mathbf{x} = \mathbf{x} = A^T \mathbf{b}$$

*In diesem Falle kann das Gleichungssystem also einfach durch die Multiplikation der rechten Seite  $\mathbf{b}$  mit der Transponierten  $A^T$  gelöst werden.*



## Lemma B-7.4 (Lösung diagonalen Systeme)

*Falls  $A = \text{diag}(\alpha_i)_{i=1\dots n}$  eine Diagonalmatrix ist, so reduziert sich das lineare System auf die Gleichungen  $\alpha_i x_i = b_i$ . Diese werden für beliebige  $b_i$  durch  $x_i = b_i / \alpha_i$  genau dann erfüllt, wenn keines der Diagonalelemente  $\alpha_i$  gleich Null ist.*

*Falls diese Regularitätsbedingung verletzt ist, muss **b** die Konsistenzbedingung*

$$\alpha_i = 0 \quad \Rightarrow \quad b_i = 0$$

*erfüllen. Die entsprechenden Lösungskomponenten  $x_i$  sind dann beliebig, so dass das Gleichungssystem  $A\mathbf{x} = \mathbf{b}$  mehrdeutig lösbar ist.*

## Lemma B - 7.5 (Lösung von Dreieckssystemen)

*Ist  $A$  eine untere Dreiecksmatrix, hat das entsprechende Gleichungssystem  $A\mathbf{x} = \mathbf{b}$  die folgende „gestaffelte“ Form:*

$$\begin{array}{rcl} \alpha_{11}x_1 & = & b_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 & = & b_2 \\ \vdots & & \vdots \\ \alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{ij}x_j & = & b_i \\ \vdots & & \vdots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{n,n-1}x_{n-1} + \alpha_{nn}x_n & = & b_n \end{array}$$

## Vorwärtssubstitution

Nun kann man zunächst aus der ersten Gleichung  $x_1$  bestimmen, dann diesen Wert in die Zweite einsetzen, um  $x_2$  zu erhalten, und so weiter. Unter der Regularitätsbedingung aus Lemma B-7.4, dass wiederum keines der diagonalen Elemente  $\alpha_{ii}$  verschwindet, hat man also

$$\begin{aligned}x_1 &= b_1/\alpha_{11} \\x_2 &= (b_2 - \alpha_{21}x_1)/\alpha_{22} \\x_3 &= (b_3 - \alpha_{31}x_1 - \alpha_{32}x_2)/\alpha_{33} \\&\vdots \\x_i &= (b_i - \alpha_{i1}x_1 - \cdots - \alpha_{i,i-1}x_{i-1})/\alpha_{ii} \\&\vdots \\x_n &= (b_n - \alpha_{n1}x_1 - \cdots - \alpha_{nj}x_j - \cdots - \alpha_{nn-1}x_{n-1})/\alpha_{nn}\end{aligned}$$

Man braucht  $n(n-1)/2$  Multiplikationen und Additionen sowie  $n$  Divisionen.

## Rückwärtssubstitution

Bei einer oberen Dreiecksmatrix  $A$  ergibt sich entsprechend das Verfahren der **Rückwärtssubstitution**, wobei jetzt die  $x_i$  für  $i = n, n-1, \dots, 1$  durch die Formel

$$x_i = \frac{1}{\alpha_{ii}} \left( b_i - \sum_{j=i+1}^n \alpha_{ij} x_j \right) \quad i = n, n-1, \dots, 1$$

bestimmt sind. Regularitätsbedingung ist wiederum, dass keines der Diagonalelemente verschwindet. Der Rechenaufwand ist auch hier von der Ordnung  $n^2/2$  arithmetischer Operationen.

Zur Lösung allgemeiner linearer Systeme kann man die Matrix  $A$  so modifizieren, daß sie eine der oben genannten speziellen Formen annimmt oder das Produkt solcher spezieller Matrizen wird. Das klassische Verfahren für eine solche Transformation ist die **Elimination nach Carl Friedrich Gauß** (1777 – 1855).

## B-8 Gauß-Elimination (1850)

Die Grundlage dieses Verfahrens ist die Beobachtung, dass für zwei Funktionen  $f(\mathbf{x})$  und  $g(\mathbf{x})$  eines Vektors  $\mathbf{x}$  und jeden beliebigen Skalar  $\lambda$  gilt:

$$\boxed{\begin{array}{l} f(\mathbf{x}) = 0 \\ g(\mathbf{x}) = 0 \end{array}} \iff \boxed{\begin{array}{l} f(\mathbf{x}) = 0 \\ g(\mathbf{x}) - \lambda f(\mathbf{x}) = 0 \end{array}}$$

$\underbrace{\hspace{10em}}_{=:\tilde{g}(\mathbf{x})}$

Mit anderen Worten: Die Menge  $\{\mathbf{x} | f(\mathbf{x}) = g(\mathbf{x}) = 0\}$  der Lösungen  $\mathbf{x}$  des Gleichungspaares  $f(\mathbf{x}) = 0$  und  $g(\mathbf{x}) = 0$  ist genau dieselbe wie die Lösungsmenge des Gleichungspaares  $f(\mathbf{x}) = 0$  und  $\tilde{g}(\mathbf{x}) = 0$ . Hierbei wurde die neue zweite Gleichung  $\tilde{g}(\mathbf{x}) = 0$  durch **Subtraktion eines Vielfachen der ersten von der alten zweiten Gleichung** erhalten.

Selbst wenn  $f(\mathbf{x})$  und  $g(\mathbf{x})$  nichtlinear sind, kann man gelegentlich durch solche Umformungen ein System von zwei oder mehreren Gleichungen sukzessive vereinfachen, bis eine explizite Lösung gelingt.

## Lineare Systeme in zwei Variablen

Zunächst betrachten wir hier den Fall von zwei linearen Gleichungen in zwei Unbekannten.

$$\alpha_{11}x_1 + \alpha_{12}x_2 = b_1$$

$$\alpha_{21}x_1 + \alpha_{22}x_2 = b_2$$

**Ausnahmefall:**  $\alpha_{11} = 0$

**Tauscht** man die beiden Gleichungen **aus**, so ergibt sich das gestaffelte Gleichungssystem

$$\begin{array}{l} \alpha_{21}x_1 + \alpha_{22}x_2 = b_2 \\ \alpha_{11}x_1 + \alpha_{12}x_2 = b_1 \end{array} \iff \tilde{A}\mathbf{x} = \begin{pmatrix} \tilde{\alpha}_{11} & \tilde{\alpha}_{12} \\ \tilde{\alpha}_{21} & \tilde{\alpha}_{22} \end{pmatrix} \mathbf{x} = \begin{pmatrix} \alpha_{21} & \alpha_{22} \\ 0 & \alpha_{12} \end{pmatrix} \mathbf{x} = \tilde{\mathbf{b}} = \begin{pmatrix} b_2 \\ b_1 \end{pmatrix},$$

wobei die Komponenten der rechten Seite auch vertauscht wurden.

Damit hat die Matrix  $\tilde{A}$  nun Dreiecksform. Vorausgesetzt, die beiden neuen Diagonalelemente  $\tilde{\alpha}_{11}$  und  $\tilde{\alpha}_{22}$  sind beide nicht Null, ergibt sich durch Rückwärtssubstitution

$$x_2 = \tilde{b}_2 / \tilde{\alpha}_{22} \quad \text{und} \quad x_1 = (\tilde{b}_1 - \tilde{\alpha}_{12}x_2) / \tilde{\alpha}_{11}.$$

## Normalfall: $\alpha_{11} \neq 0$

In diesem Fall läßt sich durch **Abziehen** des  $\boxed{\lambda_{21} \equiv \alpha_{21}/\alpha_{11}}$ -fachen **der ersten von der zweiten Gleichung** die Variable  $x_1$  aus Letzterer **eliminieren**.

Man erhält also

$$\underbrace{(\alpha_{21} - \lambda_{21}\alpha_{11})}_{\tilde{\alpha}_{21} = 0} x_1 + \underbrace{(\alpha_{22} - \lambda_{21}\alpha_{12})}_{\tilde{\alpha}_{22}} x_2 = \underbrace{(b_2 - \lambda_{21}b_1)}_{\tilde{b}_2}$$

Da  $\lambda_{21}$  gerade so gewählt wurde, dass  $\tilde{\alpha}_{21}$  verschwindet, hat das System nun wieder eine gestaffelte Form und die Lösungskomponenten  $x_2$  und  $x_1$  können durch **Rückwärtssubstitution** berechnet werden.

## Pivotierung

Das im Nenner von  $\lambda_{21}$  auftretende Diagonalelement  $a_{11}$  nennt man auch das **Pivotelement**.

Ist es ursprünglich gleich Null, so versucht man durch **Zeilenaustausch** (d.h. Umordnen der Gleichungen) ein nichtverschwindendes Pivotelement zu erhalten. Ist dies nicht möglich, so ist das Gleichungssystem singulär, d.h. nicht regulär. (Dieser Fall wird später betrachtet.)

Sind alle Diagonalelemente von  $A$  von Null verschieden, dann läßt sich  $A$  direkt durch  $n - 1$  sukzessive Eliminationsschritte **ohne** Zeilenaustausch auf Dreiecksform bringen.



# Lösung von Systemen beliebiger Dimension

Wir betrachten nun ein quadratisches Gleichungssystem von  $n$  Gleichungen mit  $n$  Unbekannten:

$$A\mathbf{x} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2j} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nj} & \dots & \alpha_{nn} \end{pmatrix} \mathbf{x} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}$$

## Erster Schritt

Eliminiere  $\alpha_{i1}$  mit Hilfe des nichtverschwindenden Diagonalelementes  $\alpha_{11}$ . Zu diesem Zwecke wird das  $\lambda_{i1}$ -fache der ersten Zeile mit

$$\lambda_{i1} = \alpha_{i1}/\alpha_{11} \qquad i = 2 \dots n$$

von allen anderen Zeilen abgezogen.

Dadurch erhalten die Elemente  $\alpha_{ij}$  mit  $i > 1$  und  $j > 1$  die neuen Werte

$$\alpha_{ij} \leftarrow \alpha_{ij} - \lambda_{i1}\alpha_{1j} \qquad i, j = 2 \dots n$$

Da die alten Werte nicht mehr gebraucht werden, kann man sie unmittelbar mit den Neuen überschreiben. (Deswegen haben wir hier nicht mehr wie im zweidimensionalen Fall die neuen Werte durch eine Tilde  $\sim$  von den Alten unterschieden.)

Entsprechend werden auch die Komponenten der rechten Seite nach der Formel

$$b_i \leftarrow b_i - \lambda_{i1} b_1 \quad i = 2 \dots n$$

"aufdatiert".

Anschließend hat das Gleichungssystem die Form

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2j} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & \alpha_{n2} & \dots & \alpha_{nj} & \dots & \alpha_{nn} \end{pmatrix} \mathbf{x} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}$$

## Zwischenergebnis nach k-1 Schritten

$$A\mathbf{x} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1,k-1} & \alpha_{1k} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2,k-1} & \alpha_{2k} & \dots & \alpha_{2n} \\ 0 & 0 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \alpha_{k-1,k-1} & \alpha_{k-1,k} & \dots & \alpha_{k-1,n} \\ \vdots & \vdots & & 0 & \alpha_{k,k} & \dots & \alpha_{k,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \alpha_{nk} & \dots & \alpha_{nn} \end{pmatrix} \mathbf{x} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \\ \vdots \\ \vdots \\ b_n \end{pmatrix}$$

## k-ter Schritt

Zur Elimination der letzten  $n - k$  Elemente in der  $k$ -ten Spalte subtrahiert man nun für  $i = k + 1, \dots, n$  das  $\lambda_{ik}$ -fache der  $k$ -ten Zeile mit

$$\lambda_{ik} = \alpha_{ik} / \alpha_{kk} \quad i = k + 1 \dots n$$

von der  $i$ -ten Zeile.

Es gilt also für  $j = k + 1, \dots, n$  die Aufdatierungsformel

$$\alpha_{ij} \leftarrow \alpha_{ij} - \lambda_{ik} \alpha_{kj} \quad i, j = k + 1 \dots n$$

und entsprechend für die rechte Seite

$$b_i \leftarrow b_i - \lambda_{ik} b_k \quad i = k + 1 \dots n.$$

## Spaltenpivotierung

Findet sich im  $k$ -ten Schritt in der Diagonale ein Element  $\alpha_{kk}$ , das gleich Null oder auch nur sehr klein ist, so sollte man einen Zeilenaustausch vornehmen.

Wenn die Matrix  $A$  regulär ist, dann muss mindestens eines der Elemente  $\alpha_{ik}$  mit  $i \geq k$  ungleich Null sein und kann dann durch Austausch der  $i$ -ten und  $k$ -ten Zeile in die Diagonale gebracht werden.

In Computerberechnungen wählt man im allgemeinen das  $\alpha_{ik}$  mit dem maximalen Betrag.

Bei Handrechnungen wählt man oft auch „glatte“ Zahlen, die die weitere Rechnung etwas erleichtern, auch wenn das ursprüngliche Diagonalelement nicht gleich Null ist.

## Aufwandsbetrachtung

Bei größeren Gleichungssystemen sind oft viele Elemente der gegebenen Matrix  $A$  gleich Null. Man kann dann bei der Pivotwahl darauf abzielen, möglichst viele von ihnen während der Aufdatierungen zu erhalten. Dadurch lassen sich Rechenaufwand und Speicherbedarf oft dramatisch reduzieren.

Sind alle Elemente von  $A$  ungleich Null, so beträgt der Rechenaufwand für die Gaußsche Elimination in etwa  $n^3/3$  Multiplikationen und Additionen.

Es ist bemerkenswert, dass dieser Aufwand nur einem Drittel des Aufwandes entspricht, der sich für die Multiplikation zweier quadratischer Matrizen im Standardverfahren ergibt.

## Zeilenoperationen mittels Elementarmatrizen

Wir betrachten das Gleichungssystem  $Ax = b$  bzw.  $[A, b]$ .

Die wichtigste Operation des Gaußschen Algorithmus ist es, das  $\lambda$ -fache der Zeile  $j$  von der Zeile  $k$  ( $\neq j$ ) abzuziehen. Das veränderte System in Matrixform ergibt sich durch Multiplikation beider Seiten von links mit der elementaren  $(m \times m)$ -Matrix

$$E_{kj}(\lambda) = I_m - \lambda e_k e_j^T, \quad \text{z.B.} \quad I_4 - 3 e_4 e_2^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 \end{pmatrix}.$$

Hat  $A$  in der ersten Spalte die Einträge  $(2, 1, 4, 3)^T$ , so wird durch Multiplikation mit der Beispielmatrix an der letzten Stelle eine Null erzeugt.

Diese Elementarmatrizen lassen sich leicht invertieren, es ist

$$(I_m + \lambda e_k e_j^T) \cdot (I_m - \lambda e_k e_j^T) = I_m + \lambda e_k e_j^T - \lambda e_k e_j^T - \lambda^2 e_k (e_j^T e_k) e_j^T = I_m.$$



Die Vertauschung zweier Zeilen  $j$  und  $k \neq j$  erfolgt mittels der Matrix

$$T_{j,k} = \sum_{i \neq j,k} e_i e_i^T + (e_j e_k^T + e_k e_j^T) = (I_m - e_j e_j^T - e_k e_k^T) + (e_j e_k^T + e_k e_j^T) .$$

$T_{j,k}$  ist zu sich selbst invers,  $T_{j,k}^2 = I_m$ .

Die Multiplikation der Zeile  $j$  mit einem Faktor  $\lambda \neq 0$  erfolgt mittels einer Diagonalmatrix

$$S_j(\lambda) = I_m + (\lambda - 1) e_j e_j^T = \text{Diag}(1, \dots, 1, \overbrace{\lambda}^j, 1, \dots, 1)$$

Es gilt  $S_j(\lambda) \cdot S_j(\lambda^{-1}) = I_m$ .

## Interpretation als LU-Faktorisierung

Angenommen, man hat den Gaußschen Algorithmus auf ein System  $[A, \mathbf{b}]$  angewandt und will nun ein System  $[A, \mathbf{c}]$  mit einer neuen rechten Seite lösen. Weiter sei angenommen, dass der Gaußsche Algorithmus ohne Zeilenvertauschungen und Skalierungen auskam, d.h. nur Vielfache oberer Zeilen zu unteren Zeilen hinzuaddiert wurden.

Dann kann man die erneute Reduktion von  $A$  auf Dreiecksform vermeiden, indem man sich die Elementaroperationen merkt und diese nur auf die neue rechte Seite  $\mathbf{c}$  wirken lässt.

Die Operation  $E_{j\,k}(\lambda_{j\,k})$  erzeugt eine Null an der Stelle  $(j, k)$  unterhalb der Diagonalen, d.h.  $j < k$ . An dieser nun freien Stelle kann der Multiplikator  $\lambda_{j\,k}$  nun statt des ursprünglichen  $\alpha_{j\,k}$  abgespeichert werden. Auf und oberhalb der Diagonalen verbleiben die Koeffizienten der vom Gauß-Algorithmus erzeugten oberen Dreiecksmatrix.

Vorausgesetzt, kein Zeilenaustausch war nötig, gilt für das ursprüngliche  $A$  und die aus der Gauß-Elimination resultierende obere (engl. **U**pper) Dreiecksmatrix  $U$

$$A = L U$$

wobei der linke Faktor  $L$  die folgende untere (engl. **L**ower) Dreiecksmatrix ist:

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda_{21} & 1 & 0 & \dots & 0 \\ \lambda_{31} & \lambda_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \lambda_{n3} & \dots & 1 \end{pmatrix}$$

Zu lösen bleibt

$$L(U\mathbf{x}) = L\mathbf{y} = \mathbf{c} \quad \text{mit} \quad U\mathbf{x} = \mathbf{y}.$$

Man löst also zunächst  $L\mathbf{y} = \mathbf{c}$  mittels Vorwärtssubstitution und dann  $U\mathbf{x} = \mathbf{y}$  mittels Rückwärtssubstitution. Der Gesamtaufwand entspricht recht genau  $n^2$  Operationen und damit einer Matrix-Vektor-Multiplikation.

## B-9 Determinante und Inverse

Für jede quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  läßt sich ein Skalarwert  $\det(A) \in \mathbb{R}$  berechnen, für den gilt:

$$\det(A) \neq 0 \iff A \text{ regulär}$$

Eine Dreiecksmatrix ist regulär, wenn alle ihre Diagonalelemente nicht Null sind. Man definiert also

$$\det(A) = \prod_{i=1}^n \alpha_{ii} \quad \text{für } A = \begin{array}{|c|} \hline \triangle \\ \hline \end{array} \quad \text{oder} \quad \begin{array}{|c|} \hline \triangle \\ \hline \end{array}.$$

Verlangt man nun noch

- (i) dass die Determinante konstant bleibt, wenn ein Vielfaches einer Zeile (Spalte) zu einer anderen Zeile (Spalte) addiert wird
- (ii) und daß sie lediglich das Vorzeichen wechselt, wenn zwei Zeilen (Spalten) ausgetauscht werden,

dann ist die Determinante schon eindeutig festgelegt.

## Berechnung durch Reduktion

Wie wir im Abschnitt **B-8 Gauss-Elimination** gesehen haben, läßt sich jede quadratische Matrix mittels elementarer Zeilen (Spalten)-Operationen und Zeilen (Spalten)-Vertauschungen in Dreiecksform überführen.

Dieses Vorgehen ist im allgemeinen auch der effizienteste Weg, eine Determinante zu berechnen.

### Beispiel B-9.1

$$\begin{aligned}\det \begin{pmatrix} -1 & 1 & 1 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{pmatrix} &= \det \begin{pmatrix} -1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & -1 \end{pmatrix} \\ &= -1 \cdot 2 \cdot (-1) = 2\end{aligned}$$

## Beispiel B-9.2

$$\det \begin{pmatrix} 1 & 3 & 2 & 4 \\ 2 & 6 & 4 & 12 \\ 4 & 15 & 7 & 11 \\ -2 & 3 & -6 & 1 \end{pmatrix} = -\det \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & 3 & -1 & -5 \\ 0 & 0 & 0 & 4 \\ 0 & 9 & -2 & 9 \end{pmatrix}$$

$$\det \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & 3 & -1 & -5 \\ 0 & 0 & 1 & 24 \\ 0 & 0 & 0 & 4 \end{pmatrix} = \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 3 & -1 & -5 \\ 0 & 0 & 1 & 24 \\ 0 & 0 & 0 & 4 \end{vmatrix} = 12$$

Der doppelte Vorzeichenwechsel resultiert aus den beiden Zeilenvertauschungen.

### Bemerkung

Die Betragstriche  $|A|$  stellen eine alternative Bezeichnung für  $\det(A)$  dar.

## Entwicklungssatz

Bei kleineren Matrizen läßt sich die Determinante auch rekursiv nach dem folgenden Entwicklungssatz berechnen.

### Satz B-9.3

*Bezeichnet man mit  $A_{ij}$  die  $(n-1) \times (n-1)$  Matrizen, die durch Weglassen der  $i$ -ten Zeile und  $j$ -ten Spalte aus  $A \in \mathbb{R}^{n \times n}$  hervorgegangen sind, so gilt für beliebiges (aber festes)  $i$  bzw.  $j$*

$$\begin{aligned}\det(A) &= \sum_{k=1}^n \alpha_{i k} \det(A_{i k}) (-1)^{i+k} \\ &= \sum_{k=1}^n \alpha_{k j} \det(A_{k j}) (-1)^{j+k} = \det(A^T)\end{aligned}$$

*Man sagt auch, die Determinante  $\det(A)$  wird **nach der  $i$ -ten Zeile** bzw.  **$j$ -ten Spalte** entwickelt.*

## HINWEISE

Der Satz ergibt sich ziemlich direkt aus einer auf Leibniz (Gottfried Wilhelm L., 1646 – 1716) zurückgehenden expliziten Formel für die Determinante. Es gilt nämlich

$$\det(A) = \sum \alpha_{1j_1} \alpha_{2j_2} \dots \alpha_{nj_n} \operatorname{sgn}(j_1, j_2, \dots, j_n)$$

wobei die Spaltenindizes  $(j_1, j_2, \dots, j_n)$  alle möglichen Permutationen der Zahlen  $(1, 2, \dots, n)$  durchlaufen. Das jeweilige  $\operatorname{sgn}(j_1, j_2, \dots, j_n)$  ist entweder  $+1$  oder  $-1$ , je nachdem, ob die Permutation durch eine gerade oder ungerade Zahl von Zweier-Vertauschungen aus der Grundpermutation  $(1, 2, \dots, n)$  gebildet werden kann.

Da die Gesamtzahl der Permutationen und damit der Summanden in der Leibnizschen Formel  $n!$  ist, wird diese in der Praxis selten angewandt.



## Beispiel B-9.4

Im folgenden wird der Entwicklungssatz zunächst auf die dritte Spalte, die zwei Nullen enthält (damit bleiben nur 2 Summanden übrig), angewendet.

$$\begin{aligned}
 \begin{vmatrix} 1 & 2 & 0 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 6 & 0 & 4 \\ 0 & 1 & 2 & 1 \end{vmatrix} &= -4 \begin{vmatrix} 1 & 2 & 2 \\ 3 & 6 & 4 \\ 0 & 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ 3 & 6 & 4 \end{vmatrix} \\
 &= -4 \begin{vmatrix} 6 & 4 \\ 1 & 1 \end{vmatrix} + 4 \cdot 3 \begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 1 & 3 \\ 6 & 4 \end{vmatrix} \\
 &\quad + 4 \begin{vmatrix} 2 & 2 \\ 6 & 4 \end{vmatrix} - 6 \begin{vmatrix} 2 & 2 \\ 1 & 3 \end{vmatrix} \\
 &= -4 \cdot 2 + 12 \cdot 0 - 2(-14) + 4(-4) - 6 \cdot 4 \\
 &= -8 + 28 - 16 - 24 = -20
 \end{aligned}$$

## Determinantenprodukt

Während sich die Determinante der Summe zweier Matrizen nicht leicht berechnen läßt, ergibt sich für Matrixprodukte die folgende multiplikative Regel:

### Satz B-9.5

*Sind  $A, B$  Matrizen vom Typ  $(n, n)$ , so gilt:*

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

### Beweis

$$A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \quad B = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \vdots & \vdots & & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{bmatrix}$$

$$C = AB = \left( \underbrace{\sum_{i=1}^n \beta_{i1} \mathbf{a}_i}_{\mathbf{c}_1}, \underbrace{\sum_{i=1}^n \beta_{i2} \mathbf{a}_i}_{\mathbf{c}_2}, \dots, \underbrace{\sum_{i=1}^n \beta_{in} \mathbf{a}_i}_{\mathbf{c}_n} \right)$$

## Fortsetzung: Beweis

Wir betrachten z.B.  $\tilde{B} = (\mathbf{b}_1 + \lambda \mathbf{b}_2, \mathbf{b}_2, \dots, \mathbf{b}_n) \implies \det(\tilde{B}) = \det(B)$

$$\tilde{C} = A\tilde{B} = (\mathbf{c}_1 + \lambda \mathbf{c}_2, \mathbf{c}_2, \dots, \mathbf{c}_n) \implies \det(A\tilde{B}) = \det(AB)$$

D.h. werden an  $B$  Spaltenoperationen  $B \longrightarrow \tilde{B}$  durchgeführt, die  $\det(\tilde{B})$  nicht ändern, dann bleibt auch  $\det(A\tilde{B}) = \det(AB)$  unverändert.

Man kann also  $B$  schrittweise in eine Dreiecks- bzw. sogar Diagonalmatrix

$$D = \text{diag}(\delta_1, \dots, \delta_n)$$

umformen, wobei

$$\det(AB) = \det(AD), \quad \det(D) = \det(B)$$

$$AD = (\delta_1 \mathbf{a}_1, \delta_2 \mathbf{a}_2, \dots, \delta_n \mathbf{a}_n)$$

$$\det(AD) = \det(A) \cdot \prod_{i=1}^n \delta_i = \det(A) \det(D) = \det(A) \det(B)$$



# Cramersche Regel

Gabriel Cramer (1704 – 1752)

Vorbetrachtung:

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

Entwicklung nach der j-ten Spalte

$$\sum_{i=1}^n (-1)^{i+j} \alpha_{ik} \det(A_{ij}) = 0$$

für  $k \neq j$

An die Stelle des Spaltenvektors  
 $\tilde{\mathbf{a}}_j$  wird  $\mathbf{a}_k$  gesetzt,  
 $\tilde{A}$  enthält zwei gleiche Spalten!!

## Lineares Gleichungssystem: $A\mathbf{x} = \mathbf{b}$

$$\left. \begin{array}{l} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n = b_1 \\ \vdots \\ \alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n = b_i \\ \vdots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{nn}x_n = b_n \end{array} \right\} \begin{array}{l} | (-1)^{1+j} \det(A_{1j}) \\ \\ | (-1)^{i+j} \det(A_{ij}) \\ \\ | (-1)^{n+j} \det(A_{nj}) \end{array} \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{geeignete} \\ \text{Multiplikationen,} \\ \text{anschließend} \\ \text{Summation} \end{array}$$

$$\begin{aligned} x_1 \underbrace{\sum_{i=1}^n (-1)^{i+j} \alpha_{i1} \det(A_{ij})}_0 + \cdots + x_j \underbrace{\sum_{i=1}^n (-1)^{i+j} \alpha_{ij} \det(A_{ij})}_{\det(A)} + \cdots \\ = \sum_{i=1}^n (-1)^{i+j} b_i \det(A_{ij}) = \det(A_j | \mathbf{b}) \end{aligned}$$

$A_j | \mathbf{b}$  bedeutet Ersetzung des Spaltenvektors  $\mathbf{a}_j$  durch Vektor  $\mathbf{b}$

## Satz B-9.6 (Cramersche Regel, 1850)

Falls  $\det(A) \neq 0$ , dann kann die eindeutige Lösung des Gleichungssystems  $Ax = b$  nach der **Cramerschen Regel** bestimmt werden:

$$x_j = \frac{1}{\det(A)} \det(A_j | \mathbf{b})$$

Dabei bedeutet  $A_j | \mathbf{b}$ , daß in  $A$  die  $j$ -te Spalte durch  $\mathbf{b}$  ersetzt wird.

### Bemerkung

Die Cramersche Regel ist rechnerisch sehr aufwendig und deshalb vorrangig von theoretischem Interesse. Sie wird angewandt in Fällen, wo die Koeffizienten  $\alpha_{ij}$  z. B. funktionelle Ausdrücke sind oder wenn eventuell nur eine der Unbekannten  $x_j$  benötigt wird.

## Beispiel B - 9.7

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \implies \det(A) = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2$$

$$\left. \begin{aligned} x_1 &= \frac{1}{2} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = \frac{1}{2} \\ x_2 &= \frac{1}{2} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -\frac{1}{2} \\ x_3 &= \frac{1}{2} \begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = \frac{1}{2} \end{aligned} \right\} \boxed{x = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}$$

# Anwendung der Cramerschen Regel auf die Bestimmung der inversen Matrix $A^{-1}$

**Bezeichnung**  $A^{-1} = B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ ,  $AB = I$

Die gesuchte Matrix  $B = A^{-1}$  läßt sich schrittweise aus den Gleichungssystemen

$$A\mathbf{b}_k = \mathbf{e}_k = (0, \dots, 0, 1, 0, \dots, 0)^T \quad k = 1, \dots, n$$

└─  $k$ -te Position

berechnen:

$$b_{jk} = \frac{1}{\det(A)} \det(A_j | \mathbf{e}_k)$$

$$= \frac{(-1)^{j+k}}{\det(A)} \det(A_{kj})$$

mit  $A_j | \mathbf{e}_k =$

$$\begin{bmatrix} \alpha_{11} & \dots & 0 & \dots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \textcircled{1} & & \vdots \\ \alpha_{n1} & \dots & 0 & \dots & \alpha_{nn} \end{bmatrix} \quad \leftarrow k$$

$\mathbf{a}_1$

$\mathbf{a}_j$

└─  $\mathbf{e}_k$



## Man beachte:

Vertauschung von Zeilen- bzw. Spaltenindex in  $b_{j\,k} = \frac{(-1)^{j+k}}{\det(A)} \det(A_{k\,j})$

Zweckmäßigerweise ergibt sich Darstellung:

$$\begin{aligned}
 A^{-1} &= \frac{1}{\det(A)} \left( ((-1)^{i+j} \det(A_{i\,j}))_{i=1\dots n}^{j=1\dots n} \right)^T \\
 &= \frac{1}{\det(A)} \begin{pmatrix} +\det(A_{11}) & -\det(A_{21}) & +\det(A_{31}) & \dots \\ -\det(A_{12}) & +\det(A_{22}) & -\det(A_{32}) & \dots \\ +\det(A_{13}) & -\det(A_{23}) & +\det(A_{33}) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}
 \end{aligned}$$

## Warnung:

Für rechteckige Matrizen läßt sich die Determinante nicht definieren.

## Beispiel B-9.8

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 2 & 0 \\ 3 & 1 & 4 \end{bmatrix} \implies \det(A) = \begin{vmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 1 & -2 \end{vmatrix} = \begin{vmatrix} 2 & 2 \\ 1 & -2 \end{vmatrix} = -6$$

$$A_{11} = \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} = 8, \quad A_{12} = \begin{vmatrix} -1 & 0 \\ 3 & 4 \end{vmatrix} = -4, \quad A_{13} = \begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} = -7$$

$$A_{21} = \begin{vmatrix} 0 & 2 \\ 1 & 4 \end{vmatrix} = -2, \quad A_{22} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2, \quad A_{23} = \begin{vmatrix} 1 & 0 \\ 3 & 1 \end{vmatrix} = 1$$

$$A_{31} = \begin{vmatrix} 0 & 2 \\ 2 & 0 \end{vmatrix} = -4, \quad A_{32} = \begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} = 2, \quad A_{33} = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2$$

$$A^{-1} = \frac{-1}{6} \begin{bmatrix} 8 & 4 & -7 \\ 2 & -2 & -1 \\ -4 & -2 & 2 \end{bmatrix}^T = \frac{-1}{6} \begin{bmatrix} 8 & 2 & -4 \\ 4 & -2 & -2 \\ -7 & -1 & 2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -8 & -2 & 4 \\ -4 & 2 & 2 \\ 7 & 1 & -2 \end{bmatrix}$$

**Probe:**  $A \cdot A^{-1} = I$

## Lösung allgemeiner linearer Systeme ( $m \neq n$ möglich)

Die Lösbarkeit eines linearen Systems

$$A\mathbf{x} = \mathbf{b} \in \mathbb{R}^m, \quad \mathbf{x} \in \mathbb{R}^n, \quad A \in \mathbb{R}^{m \times n}$$

hängt sehr stark vom **Rang** der  $m \times n$  Matrix  $A$  ab. Die natürliche Zahl  $\text{Rang}(A)$  ist gleichzeitig

- ▶ die maximale Zahl linear unabhängiger Spalten
- ▶ die maximale Zahl linear unabhängiger Zeilen
- ▶ die Dimension des Bildraumes  $\text{Bild}(A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$
- ▶  $n - \dim(\text{kern}(A))$  mit  $\text{kern}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$
- ▶ die maximale Größe einer quadratischen Untermatrix  $\tilde{A}$  (entsteht durch Streichen von geeigneten Zeilen und Spalten aus  $A$ ) mit nichtverschwindender Determinante  $\det(\tilde{A}) \neq 0$ .

Mittels der **Gauß-Elimination** läßt sich ein beliebiges, d.h. nicht notwendigerweise quadratisches, lineares System  $A\mathbf{x} = \mathbf{b}$  immer zu einem äquivalenten System  $\tilde{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$  umformen. Hierbei unterscheiden sich die Lösungsvektoren  $\tilde{\mathbf{x}}$  und  $\mathbf{x}$  nur in der Reihenfolge ihrer Elemente,  $\mathbf{x}$  kann also als Permutation von  $\tilde{\mathbf{x}}$  erhalten werden, wenn überhaupt Spaltenvertauschungen durchgeführt wurden. Die Matrix  $\tilde{A}$  hat Trapezform, so daß

$$\begin{array}{c} m \\ \left\{ \begin{array}{c} \overbrace{\begin{bmatrix} \times & \times & \dots & \times & \times & \times & \dots & \times \\ 0 & \times & \dots & \times & \times & \times & \dots & \times \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \times & \times & \times & & \times \\ 0 & 0 & \dots & 0 & \times & \times & & \times \\ 0 & 0 & \dots & 0 & 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & & 0 \end{bmatrix}}^n \\ \underbrace{\hspace{10em}}_r \quad \underbrace{\hspace{10em}}_{n-r} \end{array} \right\} \end{array} \begin{array}{c} \left[ \begin{array}{c} \times \\ \times \\ \vdots \\ \times \\ \times \\ \times \\ \times \\ \vdots \\ \times \end{array} \right] \end{array} \left. \vphantom{\begin{array}{c} \times \\ \times \\ \vdots \\ \times \\ \times \\ \times \\ \times \\ \vdots \\ \times \end{array}} \right\} \begin{array}{c} r \\ n-r \end{array} = \begin{array}{c} \left[ \begin{array}{c} \times \\ \times \\ \vdots \\ \times \\ \times \\ 0 \\ \vdots \\ 0 \end{array} \right] \end{array} \left. \vphantom{\begin{array}{c} \times \\ \times \\ \vdots \\ \times \\ \times \\ 0 \\ \vdots \\ 0 \end{array}} \right\} \begin{array}{c} r \\ m-r \end{array}$$

wobei die ersten  $r = \text{rang}(A)$  Elemente der Diagonale nicht Null sind.

## Eigenschaften der Lösung $\mathbf{x}$

Aus der Trapezform des linearen Systems lässt sich unmittelbar ablesen:

- ▶ die letzten  $m - r$  Komponenten von  $\tilde{\mathbf{b}}$  müssen verschwinden, sonst gibt es keine Lösungen,
- ▶ gibt es überhaupt Lösungen, so sind die letzten  $n - r$  Komponenten von  $\tilde{\mathbf{x}}$  beliebig.

Aus diesen beiden Beobachtungen ergeben sich für die entsprechenden Größen  $\mathbf{b}$  und  $\mathbf{x}$  die folgenden möglichen Situationen ( $r \leq \min(m, n)$  gilt immer):

**$n = r$ :** Es gilt  $\dim \text{Kern } A = n - r = 0$ . Wenn es eine Lösung gibt, so ist sie eindeutig.

**$n = r = m$ :**  $A$  ist invertierbar, es gibt immer eine Lösung.

**$n = r < m$ :** Es gibt  $\mathbf{b}$ , für welche das System unlösbar ist.

**$r < n$ :** Es gilt  $\dim \text{Kern } A = n - r > 0$ . Wenn es eine Lösung  $\mathbf{x}$  gibt, so unendlich viele, die einen affinen Unterraum (Gerade, Ebene, ...) bilden.

**$m, n < r$ :** Es gibt  $\mathbf{b}$ , für welche das System unlösbar ist.

**$n < m = r$ :** Es gibt immer mindestens eine Lösung  $\mathbf{x}$ .

## Fundamentale Eigenschaften der Determinante

- (i) Die Determinante ist *multilinear*. Sind  $\mathbf{a}_j, \tilde{\mathbf{a}}_j \in \mathbb{R}^n$ ,  $j = 1, \dots, n$  und ist  $\alpha \in \mathbb{R}$ , so gilt

$$\begin{aligned}\det(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_j + \alpha \tilde{\mathbf{a}}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) \\&= \det(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) \\&\quad + \alpha \det(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \tilde{\mathbf{a}}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) .\end{aligned}$$

- (ii) Die Determinante ist invariant unter Spalten-(Zeilen-)Operationen, ist  $\tilde{\mathbf{a}}_j = \mathbf{a}_k$  für ein  $k \neq j$ , so gilt

$$\det(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_j + \alpha \mathbf{a}_k, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) = \det(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n)$$

- (iii) Die Determinante ist invariant unter Transposition,  $\det(A) = \det(A^T)$ .

- (iv) Die Determinante ist multiplikativ,  $\det(AB) = \det(A) \det(B)$ .

## B-10 Eigenwerte und Eigenvektoren

### Komplexe Lineare Algebra

Das charakteristische Polynom  $\det(A - \lambda I_n)$  einer reellen  $(n \times n)$ -Matrix wird im allgemeinen Fall komplexe Nullstellen besitzen. Um die zugehörige Eigenvektorgleichung  $A\mathbf{v} = \lambda\mathbf{v}$  zu lösen, müssen die Komponenten von  $\mathbf{v}$  auch in den komplexen Zahlen gesucht werden. Es entsteht die Notwendigkeit, die bisher im Reellen betrachtete lineare Algebra auf die Räume  $\mathbb{C}^n$  der  $n$ -elementigen komplexen Vektoren  $\mathbf{a} \in \mathbb{C}^n$  mit

$$\mathbf{a} = (\alpha_i)_{i=1\dots n}, \quad \alpha_i = \operatorname{Re}(\alpha_i) + \mathbf{i}\operatorname{Im}(\alpha_i) \in \mathbb{C}$$

zu erweitern. Zu jedem solchen Vektor  $\mathbf{a}$  existiert der konjugiert-komplexe Vektor

$$\bar{\mathbf{a}} = (\bar{\alpha}_i)_{i=1\dots n}, \quad \bar{\alpha}_i = \operatorname{Re}(\alpha_i) - \mathbf{i}\operatorname{Im}(\alpha_i) \in \mathbb{C}.$$

Offensichtlich gilt, wie schon im skalaren Falle, für jeden **reellen Vektor**  $\mathbf{a} \in \mathbb{R}^n$ :  $\bar{\mathbf{a}} = \mathbf{a}$ .

## Sesquilinearform und Norm

Durch

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \bar{\mathbf{b}} = \bar{\mathbf{b}}^T \mathbf{a} = \sum_{i=1}^n \alpha_i \bar{\beta}_i \in \mathbb{C}$$

wird nun ein inneres Produkt (oder Skalarprodukt) definiert, welches im Gegensatz zum reellen Falle nicht symmetrisch ist, d.h. es ist von der Reihenfolge der Faktoren abhängig:

$$\langle \mathbf{b}, \mathbf{a} \rangle = \mathbf{b}^T \bar{\mathbf{a}} = \overline{\bar{\mathbf{a}}^T \mathbf{b}} = \overline{\langle \mathbf{a}, \mathbf{b} \rangle}$$

Da nun  $\langle \mathbf{a}, \mathbf{a} \rangle$  immer reell und nichtnegativ ist, läßt sich damit eine Norm

$$\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle} = \left[ \sum_{i=1}^n |\alpha_i|^2 \right]^{\frac{1}{2}}$$

definieren. Unter der Isomorphie  $\mathbb{C}^n \simeq \mathbb{R}^{2n}$  (als reelle VR) ist dies die euklidische Norm.



Es gelten die üblichen Normeigenschaften

$$\|\mathbf{a}\| \geq 0, \quad \|\mathbf{a}\| = 0 \iff \mathbf{a} = \mathbf{0}, \quad \|\gamma\mathbf{a}\| = |\gamma|\|\mathbf{a}\|$$

für beliebiges  $\gamma \in \mathbb{C}$ , sowie die Dreiecksungleichungen

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|, \quad \|\mathbf{a} - \mathbf{b}\| \geq \left| \|\mathbf{a}\| - \|\mathbf{b}\| \right|.$$

Erweitert man den Körper der Skalare von  $\mathbb{R}$  auf  $\mathbb{C}$ , so bleiben fast alle Definitionen und Sätze gültig. Das gilt insbesondere für die Begriffe

- Linearkombination
- Linearer Unterraum
- Dimension
- Lineare Unabhängigkeit
- Basis
- Lineare Abbildung

sowie Matrizen und ihre speziellen Formen.

Unterschiede ergeben sich lediglich dort, wo das innere Produkt eine wesentliche Rolle spielt:

Eine **Familie von Vektoren**  $\mathbf{v}_i \in \mathbb{C}^n$ ,  $i = 1 \dots r$ , heißt **orthogonal**, wenn

$$\mathbf{v}_i \cdot \mathbf{v}_j = \bar{\mathbf{v}}_i^T \mathbf{v}_j = 0 \quad \text{für } i \neq j,$$

was weiterhin lineare Unabhängigkeit impliziert. Eine quadratische **Matrix**

$$A = (\alpha_{ij})_{i=1 \dots n}^{j=1 \dots n} \in \mathbb{C}^{n \times n}$$

heißt **unitär**, wenn ihre Spalten paarweise orthogonal sind und ihre Norm jeweils gleich 1 ist. Mittels der **konjugiert transponierten Matrix**

$$A^H = \bar{A}^T = \begin{pmatrix} \bar{\alpha}_{11} & \dots & \bar{\alpha}_{n1} \\ \vdots & & \vdots \\ \bar{\alpha}_{1n} & \dots & \bar{\alpha}_{nn} \end{pmatrix}$$

läßt sich die Unitarität von  $A$  durch die Beziehungen

$$A^H A = \bar{A}^T A = I = A \bar{A}^T = A A^H$$

beschreiben, wobei  $I$  weiterhin die Einheitsmatrix bezeichnet:

$$I = \text{diag}(1, 1, \dots, 1) \in \mathbb{R}^{n \times n} \subset \mathbb{C}^{n \times n}.$$

Abgesehen von der Orthogonalität erfährt auch der Begriff der **Symmetrie** bei der Erweiterung auf komplexe Matrizen eine Veränderung. Man kann zwar eine Matrix  $A = (\alpha_{ij}) \in \mathbb{C}^{n \times n}$  immer noch symmetrisch nennen, wenn  $\alpha_{ji} = \alpha_{ij}$  gilt, diese Eigenschaft ist aber wesentlich weniger interessant als die Erfüllung der Bedingung

$$\bar{\alpha}_{ji} = \alpha_{ij} \quad \Rightarrow \quad A^H = \bar{A}^T = A.$$

Solche Matrizen nennt man **hermitesch** und die entsprechenden linearen Abbildungen auch **selbstadjungiert**, da für beliebige Vektoren  $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$

$$\langle \mathbf{a}, A\mathbf{b} \rangle = \mathbf{a}^T \overline{A\mathbf{b}} = (\bar{A}^T \mathbf{a})^T \bar{\mathbf{b}} = \langle A^H \mathbf{a}, \mathbf{b} \rangle = \langle A\mathbf{a}, \mathbf{b} \rangle$$

gilt. Für uns wird nur von Bedeutung sein, dass hermitesche Matrizen (wie ihre Untermenge der reell symmetrischen Matrizen) nur reelle Eigenwerte haben.

## Definition B-10.1 (Eigenwerte und Eigenvektoren)

Eine komplexe Zahl  $\lambda \in \mathbb{C}$  heißt **Eigenwert** einer quadratischen Matrix  $A \in \mathbb{C}^{n \times n}$ , wenn es einen entsprechenden **Eigenvektor**  $\mathbf{v} \in \mathbb{C}^n$  gibt, so dass gilt:

$$A\mathbf{v} = \lambda\mathbf{v} \quad \text{mit} \quad \mathbf{v} \neq 0.$$

## Folgerung B-10.2

*Daraus folgt unmittelbar, dass  $\mathbf{v}$  eine nichttriviale Lösung des homogenen Systems*

$$(A - \lambda I_n)\mathbf{v} = 0$$

*ist. Eine solche existiert genau dann, wenn der Rang von  $(A - \lambda I_n)$  kleiner als  $n$  ist, und damit äquivalenterweise gilt*

$$P(\lambda) \equiv \det(A - \lambda I_n) = 0.$$

$P(\lambda)$  wird das **charakteristische Polynom von  $A$**  genannt.

## Satz B-10.3 (Polynomeigenschaft von $P(\lambda)$ )

$P(\lambda)$  ist ein Polynom  $n$ -ten Grades und hat die spezielle Form

$$P(\lambda) = (-\lambda)^n + \text{Tr}(A)(-\lambda)^{n-1} + \dots + \det(A),$$

wobei

$$\text{Tr}(A) = \sum_{i=1}^n \alpha_{ii}$$

die Spur (engl. trace) der Matrix  $A$  bezeichnet. Falls alle Elemente von  $A$  reell sind, so gilt dies auch für die Koeffizienten des charakteristischen Polynoms (allerdings nicht für die Wurzeln).

### Bemerkung:

Die spezielle Form des  $n$ -ten,  $(n-1)$ -ten und konstanten Koeffizienten wird hier nicht bewiesen.

## Beweis:

Durch Induktion nach  $n$  beweisen wir die etwas allgemeinere Behauptung:

$P(\lambda) \equiv \det(A - \lambda B)$  ist für jedes Paar von Matrizen  $A, B \in \mathbb{C}^{n \times n}$  mit  $B = (\beta_{ij})$  ein Polynom vom Grade kleiner oder gleich  $n$ .

**Induktionsanfang,  $n = 1$ :**  $\det(A - \lambda B) = \alpha_{11} - \lambda \beta_{11}$  ist offensichtlich ein Polynom vom Grade gleich 1.

**Induktionsvoraussetzung:**  $\deg(\det(A - \lambda B)) \leq n$  sei erfüllt für  $n$ .

**Induktionsschritt,  $n \rightarrow n + 1$ :**

Nach dem Entwicklungssatz gilt für zwei Matrizen  $A, B \in \mathbb{R}^{(n+1) \times (n+1)}$

$$\det(A - \lambda B) = \sum_{j=1}^{n+1} (-1)^{j+1} (\alpha_{1j} - \lambda \beta_{1j}) \det(A_{1j} - \lambda B_{1j})$$

wobei  $A_{1j}$  und  $B_{1j}$  die durch Weglassen der ersten Zeile und  $j$ -ten Spalte aus  $A$  bzw.  $B$  gebildeten  $n \times n$  Matrizen darstellen. Nach Induktionsvoraussetzung sind die Determinanten  $\det(A_{1j} - \lambda B_{1j})$  Polynome vom Grade höchstens  $n$ , so dass die Multiplikation mit den linearen Faktoren  $(\alpha_{1j} - \lambda \beta_{1j})$  den Grad höchstens auf  $n + 1$  erhöhen kann. □

## Algebraische Vielfachheit

Nach dem Fundamentalsatz der Algebra gibt es  $k \leq n$  verschiedene Nullstellen  $\lambda_i$  der Vielfachheit  $p_i$ , so dass das charakteristische Polynom als

$$P(\lambda) = (\lambda_1 - \lambda)^{p_1} (\lambda_2 - \lambda)^{p_2} \dots (\lambda_k - \lambda)^{p_k}$$

geschrieben werden kann, es gilt  $\sum_{j=1}^k p_j = n$ . Daraus folgt

$$\sum_{j=1}^k p_j \lambda_j = \text{Tr}(A), \quad \prod_{j=1}^k \lambda_j^{p_j} = \det(A).$$

Die Zahl  $p_j > 0$  heißt die **algebraische Vielfachheit** des Eigenwertes  $\lambda_j$ .

## Beispiel B - 10.4

Die Matrix

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

hat das charakteristische Polynom

$$\begin{aligned} P(\lambda) &= \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} \\ &= \lambda^2 + 1 = (\mathbf{i} - \lambda)(-\mathbf{i} - \lambda) \end{aligned}$$

Die Eigenwerte sind also  $\lambda_1 = \mathbf{i}$  und  $\lambda_2 = -\mathbf{i}$ , beide mit der algebraischen Vielfachheit  $p_1 = p_2 = 1$ . Man prüft leicht die Identitäten

$$\text{Tr}(A) = 0+0 = 0 = \mathbf{i}-\mathbf{i} \quad \text{und} \quad \det(A) = 1 = \mathbf{i}(-\mathbf{i}) = -\mathbf{i}^2 = -(-1)$$

Dabei ist  $\mathbf{i}$  die *imaginäre Einheit* der komplexen Zahlen.



## Berechnung der Eigenvektoren

Die zu einem Eigenwert  $\lambda_j$  gehörenden Eigenvektoren  $\mathbf{v}_j$  lassen sich als Lösungen des homogenen Gleichungssystems

$$(A - \lambda_j I) \mathbf{v} = 0$$

bestimmen. Sie bilden einen linearen Unterraum der Dimension

$$q_j \equiv \dim(\ker(A - \lambda_j I_n)) = n - \text{rang}(A - \lambda_j I_n)$$

Die Zahl  $q_j > 0$  heißt die **geometrische Vielfachheit** des Eigenwertes  $\lambda_j$  und ist immer kleiner oder gleich der algebraischen Vielfachheit  $p_j$  von  $\lambda_j$ :

$$q_j \leq p_j \quad \text{für } j = 1 \dots k.$$

Eigenwerte  $\lambda_i$  mit  $q_i < p_i$  heißen **defekt**.

Zum Eigenwert  $\lambda_i$  kann man immer  $q_i$  linear unabhängige Vektoren finden, die den **Eigenunterraum**  $\ker(A - I\lambda_i)$  aufspannen. Weiterhin gilt die folgende Aussage bezüglich verschiedener Eigenwerte.

## Lemma B - 10.5

Die zu  $r$  verschiedenen Eigenwerten  $\lambda_j, j = 1 \dots r$ , gehörenden Eigenvektoren  $(\mathbf{v}_j)_{j=1 \dots r}$  sind linear unabhängig.

### Beweis:

**Induktionsanfang,  $r=1$ :**  $\mathbf{v}_1$  ist wie jeder Eigenvektor definitionsgemäß ungleich Null und deshalb linear unabhängig, d.h.  $\gamma_1 \mathbf{v}_1 = 0$  kann nur mit  $\gamma_1 = 0$  erfüllt werden.

**Induktionsvoraussetzung,  $r$ :** Die Menge der Eigenvektoren  $(\mathbf{v}_j)_{j=1 \dots r}$  sei linear unabhängig, d.h.

$$\sum_{j=1}^r \gamma_j \mathbf{v}_j = 0 \quad \implies \quad \gamma_1 = \dots = \gamma_r = 0.$$

**Induktionsschritt,  $r \rightarrow r+1$ :** Es gelte für geeignete Koeffizienten  $\gamma_i$

$$\sum_{j=1}^{r+1} \gamma_j \mathbf{v}_j = 0.$$

## Fortsetzung: Beweis

Multiplikation mit der Matrix  $A$  bzw. dem Skalar  $\lambda_{r+1}$  liefert

$$0 = A \cdot 0 = A \cdot \sum_{j=1}^{r+1} \gamma_j \mathbf{v}_j = \sum_{j=1}^{r+1} \gamma_j A \mathbf{v}_j = \boxed{\sum_{j=1}^r \gamma_j \lambda_j \mathbf{v}_j} + \boxed{\gamma_{r+1} \lambda_{r+1} \mathbf{v}_{r+1}}$$

und

$$0 = \lambda_{r+1} \cdot 0 = \lambda_{r+1} \cdot \sum_{j=1}^{r+1} \gamma_j \mathbf{v}_j = \boxed{\sum_{j=1}^r \gamma_j \lambda_{r+1} \mathbf{v}_j} + \boxed{\gamma_{r+1} \lambda_{r+1} \mathbf{v}_{r+1}}$$

Aus der Differenz dieser beiden Gleichungen fällt der letzte Term mit  $\mathbf{v}_{r+1}$  ganz heraus:

$$0 = \sum_{j=1}^r \gamma_j (\lambda_j - \lambda_{r+1}) \mathbf{v}_j.$$

Laut Induktionsannahme sind die  $\mathbf{v}_i$ ,  $i = 1 \dots r$ , linear unabhängig, also müssen die zusammengesetzten Koeffizienten  $\gamma_j (\lambda_j - \lambda_{r+1})$  alle gleich Null sein. Da die  $\lambda_j$  verschieden sind, gilt  $\lambda_{r+1} - \lambda_j \neq 0$ ,  $i = 1 \dots r$ . Dies kann aber nur bedeuten, dass die Koeffizienten  $\gamma_j$  für  $j = 1 \dots r$  und damit auch  $\gamma_{r+1}$  gleich Null sind. Also ist auch die um einen Eigenvektor erweiterte Familie  $(\mathbf{v}_j)_{j=1 \dots r+1}$ , linear unabhängig.

## Eigenwertzerlegung von Matrizen

Sind nun alle Eigenwerte einfach oder zumindest nicht defekt, so kann man einen vollen Satz von  $n$  linear unabhängigen Eigenvektoren  $\mathbf{v}_i$  finden. Diese faßt man als Spalten zu einer quadratischen Matrix

$$V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \mathbb{C}^{n \times n}$$

zusammen, welche auf Grund der linearen Unabhängigkeit ihrer Spaltenvektoren eine Inverse  $V^{-1}$  besitzt. Nun kann man die  $n$  Vektorgleichungen  $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$  mit Hilfe der Diagonalmatrix  $\Lambda = \text{diag}(\lambda_i)_{i=1\dots n}$  zur Matrixgleichung

$$A V = V \Lambda$$

kombinieren. Multipliziert man von links bzw. von rechts mit  $V^{-1}$  so erhält man die Darstellung

$$\Lambda = V^{-1} A V \quad \text{bzw.} \quad A = V \Lambda V^{-1}.$$

## Eigenwerte: Operationen und spezielle Matrizen

### Definition B-10.6 (Ähnlichkeitstransformation)

Gibt es für zwei Matrizen  $A, B \in \mathbb{C}^{n \times n}$  eine reguläre Matrix  $T$ , so dass  
 $TA = TB$  und damit auch  $A = TBT^{-1}$  bzw.  $B = T^{-1}AT$   
gilt, so sagt man, die Matrizen  $A$  und  $B$  sind **ähnlich**. Die Überführung  
der Matrix  $B$  in  $A$  durch  $TBT^{-1}$  heisst **Ähnlichkeitstransformation**.

Daraus folgt mit  $\det(T) \det(T^{-1}) = 1$  unmittelbar:

### Folgerung B-10.7 (Eigenwerte ähnlicher Matrizen)

$$\begin{aligned}\det(A - \lambda I) &= \det(TBT^{-1} - \lambda TT^{-1}) \\ &= \det(T(B - \lambda I)T^{-1}) \\ &= \det(T) \det(B - \lambda I) \det(T^{-1}) \\ &= \det(B - \lambda I)\end{aligned}$$

*Also haben zueinander ähnliche Matrizen genau dasselbe  
charakteristische Polynom und damit auch dieselben Eigenwerte.*

## Verschiebung (Shift)

Addiert man zu einer Matrix  $A$  ein Vielfaches der Einheitsmatrix, so verschieben sich die Eigenwerte entsprechend, da

$$\det((A + \mu I) - \lambda I) = \det(A - (\lambda - \mu)I),$$

so dass  $\lambda$  genau dann ein Eigenwert von  $(A + \mu I)$  ist, wenn  $\lambda - \mu$  ein Eigenwert von  $A$  ist.

## Transponierung

Selbst bei komplexen Matrizen läßt die Transponierung den Determinantenwert unverändert, so daß

$$\det(A^T - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I).$$

Also haben  $A$  und  $A^T$  genau dieselben Eigenwerte, wobei die dazugehörigen Eigenvektoren im allgemeinen allerdings verschieden sind.

## Konjugierte und symmetrische Matrizen

Da die Konjugierung von komplexen Zahlen sich auf Faktoren und Summanden überträgt, gilt

$$\overline{\det(A - \lambda I)} = \det(\overline{A - \lambda I}) = \det(\bar{A} - \bar{\lambda} I) = \det(\bar{A} - \bar{\lambda} I),$$

so dass  $\bar{\lambda}$  genau dann ein Eigenwert von  $\bar{A}$  und damit auch  $\bar{A}^T$  ist, wenn  $\lambda$  ein Eigenwert von  $A$  ist.

Da für reelle Matrizen  $A = \bar{A}$  gilt, ist für diese mit jedem  $\lambda$  auch  $\bar{\lambda}$  ein Eigenwert. Das heißt:

### Folgerung B-10.8 (Eigenwerte reeller Matrizen)

**Alle Eigenwerte reeller Matrizen sind entweder reell oder treten als konjugiert-komplexe Paare  $(\lambda, \bar{\lambda})$  auf.**

## Hermiteische Matrizen

Für diese Klasse von Matrizen sind alle Eigenwerte reell und die Eigenvektoren können sogar orthogonal zueinander gewählt werden. Es gilt nämlich

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \mathbf{v}^H(\lambda \mathbf{v}) = \mathbf{v}^H(A\mathbf{v}) = \mathbf{v}^H A^H \mathbf{v} = (A\mathbf{v})^H \mathbf{v} = \bar{\lambda} \mathbf{v}^H \mathbf{v} = \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle,$$

wobei wir die Voraussetzung  $A^H = A$  benutzt haben. Da nun  $\mathbf{v}^H \mathbf{v} = |\mathbf{v}|^2$  nicht Null ist, gilt  $\lambda = \bar{\lambda}$ , und der Eigenwert  $\lambda$  muss deshalb reell sein.

Sind  $\lambda_1, \lambda_2 \in \mathbb{R}$  zwei verschiedene Eigenwerte mit Eigenvektoren  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^n$ , so gilt analog

$$\lambda_1 \mathbf{v}_2^H \mathbf{v}_1 = \mathbf{v}_2^H A \mathbf{v}_1 = \mathbf{v}_2^H A^H \mathbf{v}_1 = \lambda_2 \mathbf{v}_1^H \mathbf{v}_2.$$

Also muss  $0 = (\lambda_2 - \lambda_1) \mathbf{v}_2^H \mathbf{v}_1 = (\lambda_2 - \lambda_1) \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$  gelten, wegen  $\lambda_1 \neq \lambda_2$  muss  $\mathbf{v}_1$  orthogonal zu  $\mathbf{v}_2$  sein.



## Praktische Berechnung der Eigenwerte

Der offensichtliche Weg, Eigenwerte und die entsprechende Eigenvektoren zu berechnen, führt über das charakteristische Polynom  $P(\lambda) = \det(A - \lambda I)$ . Bei größeren Dimensionen ist jedoch schon die Berechnung der Koeffizienten von  $P(\lambda)$  sehr aufwendig. Rechnet man mit Standard-Datentypen, so erhält man außerdem die Eigenwerte als Nullstellen von  $P(\lambda)$  mit nur geringer Genauigkeit, da sie stark durch numerische Rundungsfehler beeinträchtigt werden.

Stattdessen führt man beim sogenannten **QR-Algorithmus** eine Folge von orthogonalen Ähnlichkeitstransformationen durch, die  $A$  sukzessive auf diagonale Form reduzieren (falls  $A$  wirklich diagonalisierbar ist). Der Gesamtaufwand dafür ist normalerweise mindestens  $5n^3$  skalare Multiplikationen und damit ein Vielfaches der Kosten für eine **LU-Faktorisierung**. Deswegen wird die Eigenwertzerlegung nur in ganz speziellen Fällen zur Lösung linearer Systeme  $Ax = b$  herangezogen.

## Nutzung der Diagonalisierung

Eigenwertzerlegungen spielen besonders bei der Analyse sogenannter dynamischer Systeme eine zentrale Rolle.

Betrachtet man zum Beispiel eine lineare Evolutionsgleichung

$$x_{neu} = Ax_{alt} + b ,$$

so ergibt deren wiederholte Anwendung den  $k$ -ten Zustand von  $x$  beginnend mit  $x = 0$  als einen Ausdruck der Form  $Q_k(A)b$  mit

$$Q_k(A) = \sum_{j=0}^{k-1} q_j A^j = V \sum_{j=0}^{k-1} q_j \Lambda^j V^{-1} = V Q_k(\Lambda) V^{-1} .$$

Hierbei gilt das Gleichheitszeichen unter der Voraussetzung, dass  $A = V \Lambda V^{-1}$  mit einer Diagonalmatrix  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ .

Insbesondere gilt dann  $A^j = V \Lambda^j V^{-1}$ .

Mit anderen Worten: Man kann  $V$  aus dem Matrixpolynom  $Q_k(A)$  herausziehen, so dass dessen Verhalten gerade für große  $k$  durch  $Q_k(\Lambda) = \text{diag}(Q_k(\lambda_j))_{j=1 \dots n}$  beschrieben ist.

## Kanonische Jordanform

Einige Matrizen (wie zum Beispiel  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ ) lassen sich nicht in die Form  $A = V\Lambda V^{-1}$  mit Diagonalmatrix  $\Lambda$  bringen. Vielmehr bleiben in  $\Lambda$  noch einige Elemente in der Superdiagonalen zurück, die einen sogenannten **Jordanblock** bilden. Die vollständige Diagonalisierung gelingt hier nicht, weil der Eigenwert ( $\lambda_1 = 1$ ) eine höhere algebraische ( $p_1 = 2$ ) als geometrische Vielfachheit ( $q_1 = 1$ ) besitzt, dh. defekt ist.

Dieser Effekt ist sehr speziell, da für ein beliebig kleines  $\varepsilon \neq 0$  die gestörte Matrix

$$\begin{bmatrix} 1 & 1 \\ 0 & 1+\varepsilon \end{bmatrix}$$

zwei unterschiedlichen Eigenwerte ( $1$  und  $1 + \varepsilon$ ) hat und deshalb diagonalisierbar ist mit  $V = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$ . Allerdings ist die entsprechende Matrix  $V^{-1} = \begin{bmatrix} 1 & -1/\varepsilon \\ 0 & 1/\varepsilon \end{bmatrix}$  dann sehr groß und explodiert, wenn man  $\varepsilon$  immer kleiner macht.

Zusammenfassend bleibt festzustellen, dass die Eigenwertzerlegung  $A = V\Lambda V^{-1}$  ein wesentlich kniffligeres Problem darstellt, als die Lösung linearer Gleichungssysteme  $Ax = b$ .