

# Lineare Algebra zur Elektrotechnik

Dirk Lorenz und Thomas Sonar

Wintersemester 2021/22

## Inhaltsverzeichnis

<b>Vorbemerkung</b>	<b>3</b>
<b>Notation</b>	<b>5</b>
<b>1 Komplexe Zahlen</b>	<b>7</b>
<b>2 Die Fourier-Transformation</b>	<b>11</b>
<b>3 Vektoren</b>	<b>14</b>
<b>4 Normen und Skalarprodukte</b>	<b>17</b>
<b>5 Das Skalarprodukt im <math>\mathbb{R}^n</math></b>	<b>22</b>
<b>6 Kreuzprodukt und Spatprodukt</b>	<b>27</b>
<b>7 Geraden und Ebenen im <math>\mathbb{R}^3</math></b>	<b>31</b>
<b>8 Untervektorräume, Basen und Dimension</b>	<b>35</b>
<b>9 Lineare Gleichungssysteme und Matrizen</b>	<b>40</b>
<b>10 Orthogonale und unitäre Matrizen, Kern und Lösungsraum</b>	<b>44</b>
<b>11 Das Eliminationsverfahren</b>	<b>48</b>
<b>12 Der Rang einer Matrix</b>	<b>53</b>
<b>13 Inverse Matrizen</b>	<b>58</b>
<b>14 Die LR-Zerlegung</b>	<b>63</b>
<b>15 Permutationen und Determinanten</b>	<b>68</b>
<b>16 Eigenschaften von Determinanten</b>	<b>73</b>
<b>17 Lineare Abbildungen</b>	<b>78</b>

<b>18</b>	<b>Basiswechsel</b>	<b>83</b>
<b>19</b>	<b>Orthogonalität</b>	<b>88</b>
<b>20</b>	<b>Orthogonale Projektionen</b>	<b>93</b>
<b>21</b>	<b>Bestapproximation</b>	<b>97</b>
<b>22</b>	<b>Konstruktion von Drehungen und Spiegelungen und die QR-Zerlegung</b>	<b>102</b>
<b>23</b>	<b>Lineare Ausgleichsprobleme</b>	<b>107</b>
<b>24</b>	<b>Eigenwerte und Eigenvektoren</b>	<b>113</b>
<b>25</b>	<b>Diagonalisierbarkeit</b>	<b>118</b>
<b>26</b>	<b>Symmetrische/hermitesche Matrizen und die Hauptachsentransformation</b>	<b>123</b>
<b>27</b>	<b>Die Jordansche Normalform</b>	<b>127</b>

---

## Vorbemerkung

Dies ist das Skript zur Vorlesung „Lineare Algebra zur Elektrotechnik“, gehalten an der TU Braunschweig im Wintersemester 2021/22. Es erhebt keinerlei Anspruch, auch ohne Besuch der Vorlesung verständlich zu sein. Das Skript geht auf die Vorlesung von Thomas Sonar im Wintersemester 2020/21 zurück. Diese Version enthält mit an Sicherheit grenzender Wahrscheinlichkeit zahlreiche Fehler; seien Sie also aufmerksam beim Lesen und melden Sie sich bei mir, wenn Sie Fehler gefunden haben, damit ich die Fehler korrigieren kann; alle, die nach Ihnen das Skript lesen, werden es Ihnen danken.

Zur Mathematik für die Elektrotechnik gibt es zahlreiche Bücher (Sie finden sie zum Beispiel unter Namen wie „Mathematik für Ingenieure“, „Höhere Mathematik“). Im ersten Studienjahr sind die beiden Themen „Analysis“ und „Lineare Algebra“ wichtig. Dabei umfasst lineare Algebra die Themen Vektorrechnung, Geometrie in  $n$  Dimensionen, lineare Gleichungssysteme, Matrizenrechnung und Vektorräume. In der Analysis geht es um die Begriffe Grenzwerte, Funktionen, Stetigkeit, Ableitung und Integral. In vielen Büchern wird zuerst die Analysis behandelt und dann die lineare Algebra, oder beide Themengebiete werden vermischt. Bücher, die zu erst die lineare Algebra behandeln sind z.B. [Ansorge u. a. \[2010\]](#); [Neher \[2018\]](#); [Papula \[2009\]](#). Das Buch [Burg u. a. \[2013\]](#) widmet sich ausschließlich der lineare Algebra.

Die Aufzeichnungen der Vorlesungen zu diesem Skript finden Sie entweder im Stud.IP-Eintrag zur Vorlesung oder auf dem YouTube Kanal

<https://www.youtube.com/channel/UCn9kJUL71fuWXXxX-o6090g>.

Alle weiteren Materialien finden Sie in Stud.IP.

Braunschweig, den 24. Februar 2022

Dirk Lorenz  
d.lorenz@tu-braunschweig.de

- 
- [Ansorge u. a. 2010] ANSORGE, Rainer ; OBERLE, Hans J. ; ROTH, Kai ; SONAR, Thomas: *Mathematik für Ingenieure, Band 1: Lineare Algebra und analytische Geometrie, Differential- und Integralrechnung einer Variablen*. 4., erweiterte Auflage. WILEY-VCH, 2010
- [Burg u. a. 2013] BURG, Klemens ; HAF, Herbert ; WILLE, Friedrich: *Höhere Mathematik für Ingenieure: Bd. 2: Lineare Algebra*. Springer-Verlag, 2013. – URL <https://link.springer.com/book/10.1007/978-3-8351-9230-0>
- [Neher 2018] NEHER, Markus: *Anschauliche höhere Mathematik für Ingenieure und Naturwissenschaftler*. Springer, 2018. – URL <https://link.springer.com/book/10.1007/978-3-658-19420-8xs>
- [Papula 2009] PAPULA, Lothar: *Mathematik für Ingenieure und Naturwissenschaftler Band 1*. Springer, 2009. – URL <https://link.springer.com/book/10.1007/978-3-658-21746-4>

---

## Notation

Hier eine (ggf. noch unvollständige) Liste mit mathematischer Schreibweisen, die wir benutzen werden:

Summen- und Produktzeichen:

$$\sum_{k=1}^n a_k = a_1 + a_2 + \cdots + a_n$$

$$\prod_{k=1}^n a_k = a_1 \cdot a_2 \cdots a_n.$$

Für  $x > 0$  bezeichnet  $\sqrt{x}$  die *positive* Lösung von  $y^2 = x$ .

Wir geben Mengen entweder als Listen an

$$\{a, b, c\}$$

oder definieren sie über Eigenschaften der Elemente

$$\{x \mid x \text{ erfüllt Eigenschaft } E\}.$$

Lies: Die Menge aller  $x$ , für die gilt “ $x$  erfüllt die Eigenschaft  $E$ ”.

Für  $a < b$  schreiben wir Intervalle als

$$[a, b] = \{x \mid a \leq x \leq b\}$$

$$]a, b[ = \{x \mid a < x < b\}$$

$$[a, b[ = \{x \mid a \leq x < b\}$$

$$]a, b] = \{x \mid a < x \leq b\}.$$

Für zwei Mengen  $X, Y$  bezeichnen wir mit

$$f : X \rightarrow Y, \quad x \mapsto f(x)$$

Lies:  $f$  ist eine Abbildung von  $X$  nach  $Y$  die  $x$  auf  $f(x)$  abbildet.

eine Funktion von  $X$  nach  $Y$ , d.h. für jedes  $x \in X$  ist mit  $f(x)$  genau ein Element von  $Y$  gegeben.

Wir werden neben den lateinischen Buchstaben auch griechische benutzen. Manche davon haben zwei Schreibweisen:

$\alpha$	alpha	$\theta, \vartheta$	theta	$\pi$	pi
$\beta$	beta	$\iota$	iota	$\rho, \varrho$	rho
$\gamma$	gamma	$\kappa$	kappa	$\sigma, \varsigma$	sigma
$\delta$	delta	$\lambda$	lambda	$\tau$	tau
$\epsilon, \varepsilon$	epsilon	$\mu$	mu	$\upsilon$	upsilon
$\zeta$	zeta	$\nu$	nu	$\phi, \varphi$	phi
$\eta$	eta	$\xi$	xi	$\psi$	psi
		$\chi$	chi	$\omega$	omega

Von den griechischen Großbuchstaben benutzen wir nur diejenigen, die sich von den lateinischen unterscheiden:

$\Gamma$	Gamma	$\Xi$	Xi	$\Phi$	Phi
$\Delta$	Delta	$\Pi$	Pi	$\Psi$	Psi
$\Theta$	Theta	$\Sigma$	Sigma	$\Omega$	Omega
$\Lambda$	Lambda	$\Upsilon$	Upsilon		



# 1 Komplexe Zahlen

Wir kennen schon folgende Mengen von Zahlen:

$$\begin{aligned}\mathbb{N} &= \{0, 1, 2, \dots\} && \text{(Natürliche Zahlen)} \\ \mathbb{N}_+ &= \mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\} \\ \mathbb{Z} &= \{\dots, -2, -1, 0, 1, 2, \dots\} && \text{(Ganze Zahlen)} \\ \mathbb{Q} &= \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}_+ \right\} && \text{(Rationale Zahlen)} \\ \mathbb{R} &= \mathbb{Q} \cup \{x \mid x \text{ irrational}\} && \text{(Reelle Zahlen)}\end{aligned}$$

$\mathbb{Q}$  steht für "Quotienten".

In gewissem Sinne gibt es viel mehr reelle Zahlen als rationale Zahlen. Irrationale Zahlen sind z.B.  $\sqrt{2}$ ,  $\log(5)$  oder auch  $\pi$ . Der wichtigste Unterschied zwischen den rationalen und den reellen Zahlen ist, dass die letzteren *vollständig* sind, aber dies wird erst in der Analysis behandelt. Weiterhin bezeichnen wir:

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}$$

dabei ist  $\mathbb{R} \times \mathbb{R}$  das *cartesische Produkt* von  $\mathbb{R}$  mit sich selbst. Die Punkte in  $\mathbb{R}^2$  bilden die *euklidische Ebene* und  $(x, y)$  sind die cartesianischen Koordinaten dieser Punkte. Ebenso ist

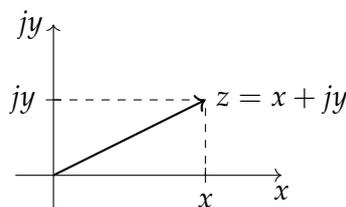
$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{n\text{-mal}}$$

der  $n$ -dimensionale euklidische Raum. Die Elemente von  $\mathbb{R}^n$  sind *Vektoren*.

Obwohl  $\mathbb{R}$  in gewissem Sinne vollständig ist, gibt es immer noch einfache Gleichungen, die keine Lösung in  $\mathbb{R}$  haben, z.B. die Gleichung  $x^2 + 1 = 0$ . Abhilfe bringen hier die *komplexen Zahlen*

$$\mathbb{C} := \{x + jy \mid x, y \in \mathbb{R}, j^2 = -1\}.$$

Die komplexen Zahlen visualisieren wir in der *Gauß'schen Zahlenebene*:



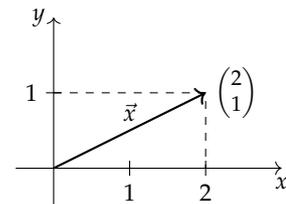
Für eine komplexe Zahl  $z = x + jy$  ist  $x = \text{Re}(z)$  der *Realteil* und  $y = \text{Im}(z)$  der *Imaginärteil*. Die Zahl  $j$  heißt *imaginäre Einheit*.

Mit komplexen Zahlen rechnen wir genau wie mit reellen Zahlen, nur dass wir  $j^2 = -1$  beachten. Für  $z_1 = x_1 + jy_1$ ,  $z_2 =$

Im Allgemeinen ist das cartesische Produkt von zwei Mengen  $A$  und  $B$  die Menge  $A \times B := \{(x, y) \mid x \in A, y \in B\}$  der geordneten Paare von Elementen aus  $A$  und  $B$ .

Im  $\mathbb{R}^2$  visualisieren wir wie folgt: Der

Vektor  $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$  ist



In Physik und Mathematik wird die imaginäre Einheit meist  $i$  genannt.

$x_2 + jy_2$  ist:

$$\begin{aligned} \text{Addition: } z_1 + z_2 &= (x_1 + jy_1) + (x_2 + jy_2) \\ &= \underbrace{(x_1 + x_2)}_{\text{Re}(z_1+z_2)} + j \underbrace{(y_1 + y_2)}_{\text{Im}(z_1+z_2)} \end{aligned}$$

$$\begin{aligned} \text{Multiplikation: } z_1 \cdot z_2 &= (x_1 + jy_1) \cdot (x_2 + jy_2) \\ &= x_1x_2 + jx_1y_2 + jx_2y_1 + j^2y_1y_2 \\ &= \underbrace{(x_1x_2 - y_1y_2)}_{\text{Re}(z_1z_2)} + j \underbrace{(x_1y_2 + x_2y_1)}_{\text{Im}(z_1z_2)} \end{aligned}$$

Um die Division

$$\frac{z_1}{z_2} = \frac{x_1 + jy_1}{x_2 + jy_2}$$

zu verstehen, benötigen wir die *komplex konjugierte Zahl*  $z^*$  :=  $x - jy$  zu  $z = x + jy$ . Wir erweitern  $z_1/z_2$  mit  $z_2^*$ :

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*}.$$

Wir bemerken:

$$\begin{aligned} z_2 z_2^* &= (x_2 + jy_2)(x_2 - jy_2) = x_2^2 - jx_2y_2 + jx_2y_2 - j^2y_2^2 \\ &= x_2^2 + y_2^2 \end{aligned}$$

und bekommen immer eine positive reelle Zahl!

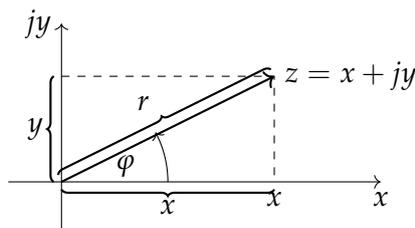
Wir bemerken weiter, dass  $z_2 z_2^* = x_2^2 + y_2^2$  nach dem Satz des Pythagoras das Quadrat der Länge von  $z$  ist, also definieren wir für  $z = x + jy$  den *Betrag* von  $z$  durch

$$|z| = \sqrt{x^2 + y^2} = \sqrt{z z^*}.$$

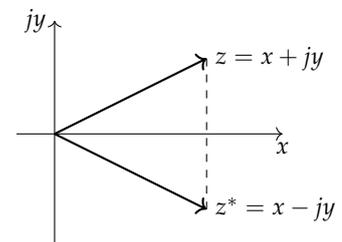
Damit ist

$$\begin{aligned} \text{Division: } \frac{z_1}{z_2} &= \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{(x_1 + jy_1)(x_2 - jy_2)}{x_2^2 + y_2^2} \\ &= \frac{(x_1x_2 + y_1y_2) + j(x_2y_1 - x_1y_2)}{x_2^2 + y_2^2} \\ &= \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + j \frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2} \end{aligned}$$

Damit haben wir in  $\mathbb{C}$  die Grundrechenarten und eine "Länge" von Zahlen bestimmt. Es geht aber noch mehr. Neben der cartesischen Beschreibung über Real- und Imaginärteil haben wir auch die *Polarform* über Winkel und Länge:



Konjugation ist die Spiegelung an der reellen Achse:



Beachte: Die komplex konjugierte wird auch mit  $\bar{z}$  statt  $z^*$  bezeichnet.

Elementargeometrisch erkennen wir  $x = r \cos(\varphi)$  und  $y = r \sin(\varphi)$ , also ist die Polarform

$$z = r(\cos(\varphi) + j \sin(\varphi)), \quad r = |z|.$$

Nach Leonard Euler sind Sinus und Kosinus durch ihre Potenzreihen gegeben:

$$\begin{aligned} \sin(x) &= \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} \mp + \dots \\ \cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \pm + \dots \end{aligned}$$

mit der *Fakultät*

$$n! = \begin{cases} 1 & ; n = 0 \\ 1 \cdot 2 \cdot 3 \cdots n & ; n > 0. \end{cases}$$

Die sogenannte *Eulersche e-Funktion* hat die Potenzreihe

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Wir schreiben auch  $\exp(x) = e^x$  und nennen die Funktion auch *Exponentialfunktion*.

Damit berechnen wir  $e^{j\varphi}$ :

$$\begin{aligned} e^{j\varphi} &= 1 + \frac{j\varphi}{1!} + \frac{(j\varphi)^2}{2!} + \frac{(j\varphi)^3}{3!} + \frac{(j\varphi)^4}{4!} + \frac{(j\varphi)^5}{5!} \dots \\ &= 1 + j\varphi - \frac{\varphi^2}{2!} - j\frac{\varphi^3}{3!} + \frac{\varphi^4}{4!} + j\frac{\varphi^5}{5!} \mp \dots \\ &= 1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} \mp \dots + j \left( \varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} \mp \dots \right) \\ &= \cos(\varphi) + j \sin(\varphi). \end{aligned}$$

Wir nutzen  $j^2 = -1, j^3 = j^2 j = -j$   
 $j^4 = j^3 j = -j j = 1, j^5 = j^4 j = j$  usw.

Damit haben wir die *Eulersche Formel*

$$e^{j\varphi} = \cos(\varphi) + j \sin(\varphi)$$

gezeigt. Wir sehen insbesondere, dass  $e^{j\varphi}$   $2\pi$ -periodisch ist, d.h. es gilt für  $k \in \mathbb{Z}$

$$\begin{aligned} e^{j(\varphi+k \cdot 2\pi)} &= \cos(\varphi + k \cdot 2\pi) + j \sin(\varphi + k \cdot 2\pi) = \cos(\varphi) + j \sin(\varphi) \\ &= e^{j\varphi} \quad (*) \end{aligned}$$

Wir erhalten folgende Darstellung der Polarform

$$z = r \cdot e^{j\varphi} = r(\cos(\varphi) + j \sin(\varphi)).$$

Mit  $\varphi = \pi$  folgt die "schönste Formel der Mathematik"  $e^{j\pi} = -1$ .

Die Polarform mit der *e-Funktion* macht das Rechnen oft einfacher:

**Potenzieren:** Für natürliche Exponenten  $n \in \mathbb{N}$  haben wir für

$$z = r \cdot e^{j\varphi}$$

$$\begin{aligned} z^n &= (r \cdot e^{j\varphi})^n = r^n \cdot e^{jn\varphi} \\ &= r^n (\cos(n\varphi) + j \sin(n\varphi)) \quad (\text{Formel von Moivre}) \end{aligned}$$

**Wurzeln:** Für  $n \in \mathbb{N}_+$  und  $a = a_0 \cdot e^{j\alpha}$  mit  $a_0 > 0$  betrachten wir die Gleichung  $z^n = a$ . Wir benutzen die Polarform  $z = r \cdot e^{j\varphi}$  und bekommen

$$z^n = r^n \cdot e^{jn\varphi} = a_0 \cdot e^{j\alpha}.$$

Vergleich von Länge und Exponent gibt

$$a_0 = r^n, \quad \alpha = n\varphi.$$

Aber Achtung! Da  $e^{j\varphi}$  nach (\*)  $2\pi$ -periodisch ist haben wir

$$n\varphi = \alpha + k \cdot 2\pi \text{ also } \varphi = \frac{\alpha + k \cdot 2\pi}{n}.$$

Es gibt also genau  $n$  verschiedene Wurzeln von  $a = a_0 \cdot e^{j\alpha}$ , nämlich

$$z_k = \sqrt[n]{a_0} e^{j \frac{\alpha + k \cdot 2\pi}{n}}, \quad k = 0, 1, 2, \dots, n-1.$$

Die Polarform ist auch hilfreich zum Multiplizieren und Dividieren. Für  $z_i = r_i e^{j\varphi_i}$ ,  $i = 1, 2$  ist

$$\begin{aligned} z_1 z_2 &= r_1 \cdot r_2 \cdot e^{j(\varphi_1 + \varphi_2)} \\ \frac{z_1}{z_2} &= \frac{r_1}{r_2} \cdot e^{j(\varphi_1 - \varphi_2)} \end{aligned}$$

Im Allgemeinen gilt sogar:

**Satz 1.1** (Hauptsatz der Algebra). *Jedes Polynom  $z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$  mit  $a_0, \dots, a_{n-1} \in \mathbb{C}$  hat in  $\mathbb{C}$  genau  $n$  Nullstellen.*

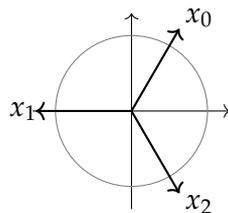
**Beispiel.** Die Gleichung  $x^3 + 2 = 0$  hat in  $\mathbb{R}$  nur die eine Nullstelle  $x_1 = -\sqrt[3]{2}$ . In  $\mathbb{C}$  bekommen wir über  $-2 = 2 \cdot e^{j\pi}$  die Nullstellen  $x_k = \sqrt[3]{2} \cdot e^{j \frac{\pi + k \cdot 2\pi}{3}}$ ,  $k = 0, 1, 2$ , also

$$\begin{aligned} x_0 &= \sqrt[3]{2} \cdot e^{j \frac{\pi}{3}} \\ x_1 &= \sqrt[3]{2} \cdot e^{j \frac{3\pi}{3}} = -\sqrt[3]{2} \\ x_2 &= \sqrt[3]{2} \cdot e^{j \frac{5\pi}{3}}. \end{aligned}$$

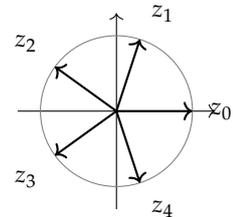
Da  $360^\circ \hat{=} 2\pi$ , gilt  $1^\circ \hat{=} \frac{\pi}{180}$  und also

$$\begin{aligned} \frac{\pi}{3} &\hat{=} \frac{180^\circ}{3} = 60^\circ \\ \frac{5\pi}{3} &\hat{=} 5 \cdot 60^\circ = 300^\circ \end{aligned}$$

und wir finden die drei Lösungen in der Gauß'schen Zahlenebene



Dies sind die 5 Lösungen der Gleichung  $z^5 = 1$ :



Wobei mehrfache Nullstellen wie z.B. die dreifache Nullstelle von  $x^3 = 0$  bei 0 entsprechend gezählt werden.

## 2 Die Fourier-Transformation

Zur Untersuchung von Signalen (also Funktionen  $s(t)$  die von einer Zeitvariablen  $t$  abhängen) werden diese oft in komplexe Frequenzen zerlegt.

**Definition 2.1** (Fourier-Transformation). Für eine Funktion  $s(t)$  ist

$$\hat{s}(\omega) := \int_{-\infty}^{\infty} s(\tau) e^{-j\omega\tau} d\tau$$

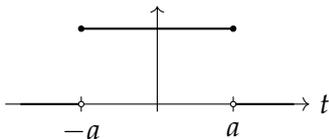
die *Fourier-Transformierte* von  $s$ .

Man kann zeigen, dass in vielen Fällen gilt

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{s}(\omega) e^{j\omega t} d\omega.$$

Die Funktion  $\hat{s}(\omega)$  heißt *kontinuierliches Spektrum* von  $f$ .

*Beispiel* (Rechteckimpuls). Für  $a > 0$  und

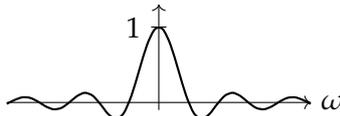
$$s(t) = \begin{cases} 1 & ; -a \leq t \leq a \\ 0 & ; |t| > a \end{cases}$$


ist die Fourier-Transformierte für  $\omega \neq 0$

$$\begin{aligned} \hat{s}(\omega) &= \int_{-\infty}^{\infty} s(\tau) e^{-j\omega\tau} d\tau \\ &= \int_{-a}^a e^{-j\omega\tau} d\tau \\ &= \left[ -\frac{1}{j\omega} e^{-j\omega\tau} \right]_{\tau=-a}^a \\ &= -\frac{1}{j\omega} (e^{-j\omega a} - e^{j\omega a}) \\ &= \frac{2}{\omega} \sin(\omega a). \end{aligned}$$

Für  $\omega = 0$  ergibt sich wegen  $\exp(0) = 1$  sofort  $\hat{s}(0) = \int_{-a}^a 1 d\tau = 2a$ .

Die Funktion

$$\text{sinc}(x) := \begin{cases} \frac{\sin(x)}{x} & ; x \neq 0 \\ 1 & ; x = 0 \end{cases}$$


heißt *Sinus cardinalis* und damit ist

$$\hat{s}(\omega) = 2a \text{sinc}(\omega a).$$

Natürlich muss das uneigentliche Integral existieren!

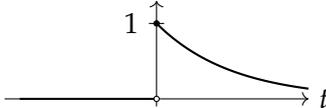
Andere Schreibweisen für  $\hat{s}(\omega)$  sind  $S(\omega)$  oder  $\mathcal{F}\{s\}(\omega)$ .

Die transformierte Variable  $\omega$  ist die *Kreisfrequenz* der Spektraldarstellung. Man kann die Fourier-Transformation auch mit Hilfe der *Frequenz*  $f = \omega/(2\pi)$  definieren, nämlich als  $S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi f t} dt$ . In diesem Fall ändert sich die Inversionsformel zu  $s(t) = \int_{-\infty}^{\infty} S(f) e^{j2\pi f t} df$ .

Aus der Euler-Formel  $e^{jx} = \cos(x) + j \sin(x)$  bekommen wir durch Addition bzw. Subtraktion von  $e^{jx}$  und  $e^{-jx}$  die Identitäten  $\cos(x) = \frac{1}{2}(e^{jx} + e^{-jx})$  und  $\sin(x) = \frac{1}{2j}(e^{jx} - e^{-jx})$ .

Die sinc-Funktion findet Anwendungen im Filter-Design und bei der Theorie des Abtastens von Signalen.

Beispiel (Spektrum einer Kondensatorentladung). Wir betrachten zu  $a > 0$

$$s(t) = \begin{cases} e^{-at} & ; t \geq 0 \\ 0 & ; t < 0. \end{cases}$$


Die Fourier-Transformierte ist

$$\begin{aligned} \hat{s}(\omega) &= \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt \\ &= \int_0^{\infty} e^{-at} e^{-j\omega t} dt \\ &= \int_0^{\infty} e^{-(a+j\omega)t} dt \\ &= \left[ -\frac{1}{a+j\omega} e^{-(a+j\omega)t} \right]_{t=0}^{\infty} \\ &= 0 - \left( -\frac{1}{a+j\omega} \right) = \frac{1}{a+j\omega}. \end{aligned}$$

Die Abbildung, die  $f$  auf die Fourier-Transformierte  $\hat{f}$  abbildet, heißt *Fourier-Transformation* und ist *linear*, d.h., es gilt

(a)  $\widehat{(f + g)}(\omega) = \hat{f}(\omega) + \hat{g}(\omega)$  (Superpositionsprinzip)

(b) Für  $\alpha \in \mathbb{C}$  ist  $\widehat{(\alpha f)}(\omega) = \alpha \hat{f}(\omega)$  (Skalierung/Homogenität)

Wir werden in dieser Vorlesung lineare Abbildung im Allgemeinen studieren. Die Fourier-Transformation taucht hier schon einmal als Beispiel auf, da sie in vielen Bereichen, wie z.B. der Signalverarbeitung oder auch der Wahrscheinlichkeitstheorie, fundamental ist.

Weitere Rechenregeln sind:

(i) **Streckung:** Für  $c \neq 0$  gilt

$$\mathcal{F}\{s(ct)\}(\omega) = \frac{1}{|c|} \mathcal{F}\{s\}\left(\frac{\omega}{c}\right).$$

(ii) **Verschiebungssätze:** Für  $a \in \mathbb{R}$  gilt

$$\begin{aligned} \mathcal{F}\{s(t-a)\}(\omega) &= e^{-j\omega a} \mathcal{F}\{s\}(\omega) \\ \mathcal{F}\{e^{jat}s(t)\}(\omega) &= \mathcal{F}\{s\}(\omega - a). \end{aligned}$$

(iii) **Ableitung:**

$$\mathcal{F}\{s'\}(\omega) = j\omega \mathcal{F}\{s\}(\omega).$$

In der digitalen Welt ist die *diskrete Fourier-Transformation* besonders wichtig:

Folgt beides aus den entsprechenden Regeln für Integrale, nämlich  $\int f + g dt = \int f dt + \int g dt$  und  $\int \alpha f dt = \alpha \int f dt$ .

**Definition 2.2.** Für eine Funktion  $f$  seien die Werte

$$f_k := f(x_k) \quad \text{an den Stellen} \quad x_k = \frac{2\pi k}{N}, \quad k = 0, 1, \dots, N-1$$

bekannt. Die Abbildung  $\mathcal{F}_N : \mathbb{C}^N \rightarrow \mathbb{C}^N$ , die einen Vektor  $\vec{f} = \begin{pmatrix} f_0 \\ \vdots \\ f_{N-1} \end{pmatrix} \in \mathbb{C}^N$  mittels

$$\hat{c}_l = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi k l}{N}}, \quad l = 0, 1, \dots, N-1$$

auf den Vektor  $\hat{c} = \begin{pmatrix} \hat{c}_0 \\ \vdots \\ \hat{c}_{N-1} \end{pmatrix}$  abbildet, heißt *diskrete Fourier-Transformation*.

Die diskrete Fourier-Transformation ist ebenfalls eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen, und lässt sich daher (wie wir später sehen werden) als *Matrix-Vektor-Produkt* schreiben, nämlich

$$\hat{c} = \frac{1}{N} F_N \vec{f}$$

mit der komplexen  $N \times N$  Fourier-Matrix

$$F_N := \left( e^{-j \frac{2\pi k l}{N}} \right)_{\substack{l=0, \dots, N-1 \\ k=0, \dots, N-1}}.$$

Für  $N = 2$  bekommen wir zum Beispiel

$$F_2 = \begin{pmatrix} e^0 & e^0 \\ e^0 & e^{-j\pi} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

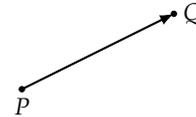
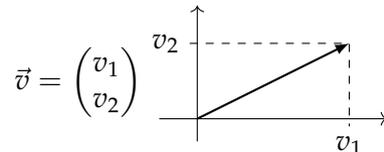
und für  $N = 4$  ergibt sich

$$F_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{pmatrix}.$$

Die diskrete Fourier-Transformation wird uns später in der linearen Algebra wieder begegnen. Die (kontinuierliche) Fourier-Transformation sehen wir erst in einer späteren Vorlesung wieder.  
der.xs

### 3 Vektoren

Wir motivieren Vektoren zuerst aus der geometrischen Anschauung: Für zwei Punkte  $P$  und  $Q$  in der Ebene ist  $\vec{v} = \overline{PQ}$  der Vektor mit Anfangspunkt  $P$  und Endpunkt  $Q$ ; dabei sehen wir alle Vektoren, die durch Parallelverschiebung aus  $\vec{v}$  hervorgehen als gleich an (manchmal spricht man auch von "freien Vektoren"). Wir geben Vektoren immer so an, dass ihr Anfangspunkt im Ursprung des Koordinatensystems liegt. In der Ebene ist das also



Ein Vektor im  $\mathbb{R}^n$  ist

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

und die  $v_i$  heißen *Komponenten* oder *Koordinaten* von  $\vec{v}$ . Im  $\mathbb{C}^n$  sind die Bezeichnungen analog.

Für Vektoren sind die Vektorraumoperationen der Vektoraddition und der Skalierung erklärt:

**Vektoraddition:** Für  $\vec{v}, \vec{w} \in \mathbb{R}^n$  (analog für  $\mathbb{C}^n$ ) ist

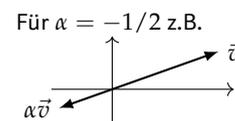
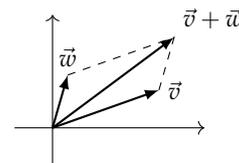
$$\vec{v} + \vec{w} := \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} \in \mathbb{R}^n.$$

**Skalierung:** Für  $\alpha \in \mathbb{R}$  und  $\vec{v} \in \mathbb{R}^n$  (bzw. auch in  $\mathbb{C}$ , bzw.  $\mathbb{C}^n$ )

$$\alpha \vec{v} := \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{pmatrix}.$$

*Bemerkung.* Für die Gauß'sche Zahlenebene haben wir komplexe Zahlen  $z = x + jy$  als Vektoren  $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$  geschrieben. Allerdings sind komplexe Zahlen spezieller als allgemeine Vektoren, da man für komplexe Zahlen Multiplikation und Division erklären kann, was für allgemeine Vektoren nicht geht. ■

Die Zahlen mit denen ein Vektor skaliert werden kann, nennen wir *Skalare*. Man nennt daher die Skalierung auch *skalare Multiplikation*.



Wir werden in Zukunft auch die Vektorräume  $\mathbb{R}^n$  und  $\mathbb{C}^n$  verlassen und abstraktere Vektoren betrachten. In diesem Kontext können z.B. auch Funktionen Vektoren sein! Wir legen klare Regeln zu Grunde, nämlich die *Vektorraumaxiome*: Diese gelten zum Beispiel auf  $\mathbb{R}^n$  und  $\mathbb{C}^n$  und sie lauten:

$$\begin{array}{l}
 (\text{VR1}) \left\{ \begin{array}{ll}
 (i) & \text{Für alle Vektoren } \vec{v}, \vec{w} \text{ gilt} & \vec{v} + \vec{w} = \vec{w} + \vec{v} \\
 (ii) & \text{Für alle Vektoren } \vec{v}, \vec{w}, \vec{u} \text{ gilt} & \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w} \\
 (iii) & \text{Es gibt } \vec{0}, \text{ so dass für alle Vektoren } \vec{v} \text{ gilt} & \vec{v} + \vec{0} = \vec{v} \\
 (iv) & \text{Für alle Vektoren } \vec{v}, \text{ gibt es } -\vec{v}, \text{ so dass} & \vec{v} + (-\vec{v}) = \vec{0}
 \end{array} \right. \\
 (\text{VR2}) \left\{ \begin{array}{ll}
 (v) & \text{Für alle Vektoren } \vec{v} \text{ gilt} & 1 \cdot \vec{v} = \vec{v} \\
 (vi) & \text{Für alle Skalare } \alpha, \beta \text{ und alle Vektoren } \vec{v} \text{ gilt} & \alpha(\beta\vec{v}) = (\alpha\beta)\vec{v} \\
 (vii) & \text{Für alle Skalare } \alpha, \beta \text{ und alle Vektoren } \vec{v} \text{ gilt} & (\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v} \\
 (viii) & \text{Für alle Skalare } \alpha \text{ und alle Vektoren } \vec{v}, \vec{w} \text{ gilt} & \alpha(\vec{v} + \vec{w}) = \alpha\vec{v} + \alpha\vec{w}
 \end{array} \right.
 \end{array}$$

Im  $\mathbb{R}^n$  und  $\mathbb{C}^n$  ist der Nullvektor  $\vec{0}$

$$\vec{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

(ein Vektor mit  $n$  Einträgen, die alle Null sind) und der Vektor  $-\vec{v}$  ist

$$-\vec{v} = (-1)\vec{v}.$$

**Definition 3.1** (Vektorraum). Eine Menge  $V$  auf der eine Addition und eine Skalierung (mit einer bestimmten Menge als Skalaren) erklärt ist, heißt *Vektorraum*, wenn die Axiome (VR1) und (VR2) erfüllt sind. Ist die Menge der Skalare  $\mathbb{R}$  (bzw.  $\mathbb{C}$ ), spricht man von einem *reellen* (bzw. *komplexen*) Vektorraum.

*Beispiel.* (a)  $\mathbb{R}^n$  und  $\mathbb{C}^n$  sind Vektorräume.

(b) Die Menge der Polynome (mit reellen bzw. komplexen Koeffizienten) sind

$$\begin{aligned}
 \Pi_n(\mathbb{R}) &= \left\{ p(x) = \sum_{k=0}^n a_k x^k \mid a_k \in \mathbb{R} \right\} \\
 \Pi_n(\mathbb{C}) &= \left\{ p(x) = \sum_{k=0}^n a_k x^k \mid a_k \in \mathbb{C} \right\}.
 \end{aligned}$$

Sie sind Vektorräume, wenn wir erklären

$$\begin{aligned}
 (p + q)(x) &:= p(x) + q(x) = \sum_{k=0}^n a_k x^k + \sum_{k=0}^n b_k x^k \\
 &= \sum_{k=0}^n (a_k + b_k) x^k \\
 (\alpha p)(x) &:= \alpha \cdot p(x) = \alpha \sum_{k=0}^n a_k x^k = \sum_{k=0}^n (\alpha a_k) x^k.
 \end{aligned}$$

- (c) Die Menge aller stetigen Funktionen auf einem abgeschlossenen Intervall  $[a, b]$  bezeichnen wir mit

$$C[a, b] := \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig}\}.$$

Vorsehen mit den Operationen

$$(f + g)(x) := f(x) + g(x)$$

$$(\alpha f)(x) := \alpha \cdot f(x)$$

bildet diese Menge einen reellen Vektorraum.

Analog bilden komplexwertige stetige Funktionen, also Funktionen  $f : [a, b] \rightarrow \mathbb{C}$  einen komplexen Vektorraum.

## 4 Normen und Skalarprodukte

Im  $\mathbb{R}^2$  und  $\mathbb{R}^3$  können wir die Länge eines Vektors mit dem Satz des Pythagoras bestimmen: Wir bezeichnen sie mit  $\|\vec{v}\|$ , nennen sie die *Euklidische Norm* von  $\vec{v}$  und sie ist für  $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$  bzw.  $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3$

$$\|\vec{v}\|_2 := \sqrt{v_1^2 + v_2^2} \text{ bzw. } := \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

Für den allgemeinen Fall definieren wir:

**Definition 4.1.** Die *Euklidische Norm* eines Vektors  $\vec{v} \in \mathbb{R}^n$  ist

$$\|\vec{v}\|_2 := \sqrt{\sum_{k=1}^n v_k^2}.$$

Für komplexe Vektoren  $\vec{v} \in \mathbb{C}^n$  ist die *unitäre Norm* definiert als

$$\|\vec{v}\|_2 := \sqrt{\sum_{k=1}^n |v_k|^2}.$$

Neben der Länge im geometrischen Sinn gibt es andere Möglichkeiten, Entfernungen zu messen! Auf einer Landkarte entspricht der Euklidische Abstand zweier Punkte dem Luftlinien-Abstand. Man könnte aber auch den Abstand entlang des Straßennetzes messen (wie es z.B. das Navigationsgerät angibt).

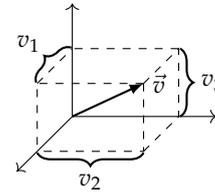
Wir definieren den allgemeinen Begriff einer Norm:

**Definition 4.2.** Ein Norm ist eine Abbildung, die einem Vektor  $\vec{v}$  eine reelle Zahl  $\|\vec{v}\|$  zuordnet und welche die folgenden Eigenschaften hat:

- (i) **Positivität:** Für alle  $\vec{v}$  ist  $\|\vec{v}\| \geq 0$ .
- (ii) **Definitheit:** Ist  $\|\vec{v}\| = 0$ , so ist  $\vec{v} = \vec{0}$ .
- (iii) **Absolute Homogenität:** Für jeden (reellen bzw. komplexen) Skalar  $\alpha$  und jeden Vektor  $\vec{v}$  gilt  $\|\alpha\vec{v}\| = |\alpha| \|\vec{v}\|$ .
- (iv) **Dreiecksungleichung:** Für je zwei Vektoren  $\vec{v}$  und  $\vec{w}$  gilt die *Dreiecksungleichung*  $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ .

*Bemerkung.* (a) Wir zeigen, dass die Euklidische Norm die Norm-Axiome erfüllt:

- (i) Klar, da die Wurzel per Definition nicht-negative Ergebnisse liefert.
- (ii) Klar, da in der Summe  $\sum_{k=1}^n v_k^2$  alle Summanden  $\geq 0$  sind, und daher  $v_k^2 = 0$  für alle  $k$  gelten muss, wenn  $\|\vec{v}\|_2 = 0$  gilt.



Die Bedingungen (i) bis (iv) heißen *Norm-Axiome*.

(iii) Es ist

$$\begin{aligned}\|\alpha \vec{v}\|_2 &= \sqrt{\sum_{k=1}^n \alpha^2 v_k^2} = \sqrt{\alpha^2 \sum_{k=1}^n v_k^2} \\ &= |\alpha| \sqrt{\sum_{k=1}^n v_k^2} = |\alpha| \|\vec{v}\|_2.\end{aligned}$$

(iv) Zuerst sehen wir, dass  $\|\vec{v} + \vec{w}\|_2 \leq \|\vec{v}\|_2 + \|\vec{w}\|_2$  äquivalent zur Bedingung  $\|\vec{v} + \vec{w}\|_2^2 \leq \|\vec{v}\|_2^2 + 2\|\vec{v}\|_2\|\vec{w}\|_2 + \|\vec{w}\|_2^2$  ist. Schreiben wir die Summen aus, so ist dies

$$\sum_{k=1}^n (v_k + w_k)^2 \leq \sum_{k=1}^n v_k^2 + 2\sqrt{\sum_{k=1}^n v_k^2} \sqrt{\sum_{k=1}^n w_k^2} + \sum_{k=1}^n w_k^2.$$

Multiplizieren wir unter dem ersten Quadrat aus, sehen wir, dass dies äquivalent ist zu

$$\sum_{k=1}^n v_k w_k \leq \sqrt{\sum_{k=1}^n v_k^2} \sqrt{\sum_{k=1}^n w_k^2}.$$

Haben wir also gezeigt, dass die letzte Ungleichung stets gilt, so haben wir gezeigt, dass auch die Dreiecksungleichung gilt. Diese Ungleichung wird *Cauchy-Schwarz-Ungleichung* genannt und wir beweisen sie später.

- (b) Einen (reellen oder komplexen) Vektorraum  $V$ , auf dem eine Norm  $\vec{v} \mapsto \|\vec{v}\|$  definiert ist, nennen wir *normierten Vektorraum*.
- (c) In einem normierten Vektorraum nennen wir  $\|\vec{v} - \vec{w}\|$  den *Abstand* zwischen  $\vec{v}$  und  $\vec{w}$ .

■

*Beispiel.* • Auf dem  $\mathbb{R}^n$  können wir auch folgende Normen definieren:

$$\|\vec{v}\|_\infty := \max\{|v_1|, \dots, |v_n|\}, \quad \text{Maximumsnorm oder } L^\infty\text{-Norm}$$

$$\|\vec{v}\|_1 := \sum_{k=1}^n |v_k|, \quad L^1\text{-Norm.}$$

(Für jedes  $p \geq 1$  ist  $\left(\sum_{k=1}^n |v_k|^p\right)^{1/p}$  eine Norm, und daher bezeichnen wir die Euklidische Norm mit  $\|\vec{v}\|_2$ ).

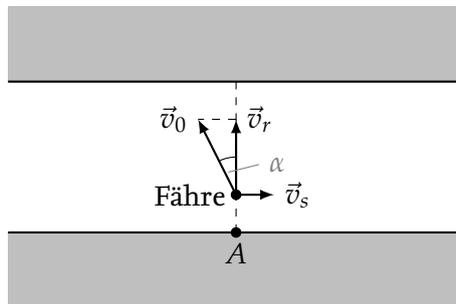
- Auf dem Vektorraum  $C[a, b]$  der stetigen Funktionen auf dem Intervall  $[a, b]$  definieren wir für  $f : [a, b] \rightarrow \mathbb{R}$

$$\|f\|_\infty := \max\{|f(t)| \mid a \leq t \leq b\}, \quad \text{Maximumsnorm oder } L^\infty\text{-Norm}$$

$$\|f\|_2 := \sqrt{\int_a^b (f(t))^2 dt}, \quad L^2\text{-Norm}$$

$$\|f\|_1 := \int_a^b |f(t)| dt, \quad L^1\text{-Norm.}$$

**Beispiel.** Wir betrachten folgende Situation: Eine Fähre fährt mit  $\|\vec{v}_0\|_2 = 4 \frac{\text{m}}{\text{s}}$  relativ zum Fluß vom Uferpunkt  $A$  auf kürzestem Weg zum gegenüberliegenden Ufer.



Wir stellen folgende Fragen:

- Unter welchem Winkel  $\alpha$  muss die Fähre gegen die Strömung fahren, wenn die Strömung die Geschwindigkeit  $\|\vec{v}_s\|_2 = 1 \frac{\text{m}}{\text{s}}$  hat?
- Wie groß ist die resultierende Geschwindigkeit  $\|\vec{v}_r\|_2$  der Fähre?

Zu (a): Es ist

$$\vec{v}_0 = \begin{pmatrix} -\|\vec{v}_0\|_2 \sin(\alpha) \\ \|\vec{v}_0\|_2 \cos(\alpha) \end{pmatrix}, \quad \vec{v}_s = \begin{pmatrix} \|\vec{v}_s\|_2 \\ 0 \end{pmatrix}, \quad \vec{v}_r = \begin{pmatrix} 0 \\ \|\vec{v}_r\|_2 \end{pmatrix}$$

Vektoraddition ergibt  $\vec{v}_r = \vec{v}_0 + \vec{v}_s$ , also

$$\begin{pmatrix} 0 \\ \|\vec{v}_r\|_2 \end{pmatrix} = \begin{pmatrix} -\|\vec{v}_0\|_2 \sin(\alpha) \\ \|\vec{v}_0\|_2 \cos(\alpha) \end{pmatrix} + \begin{pmatrix} \|\vec{v}_s\|_2 \\ 0 \end{pmatrix} = \begin{pmatrix} -\|\vec{v}_0\|_2 \sin(\alpha) + \|\vec{v}_s\|_2 \\ \|\vec{v}_0\|_2 \cos(\alpha) \end{pmatrix}.$$

Aus der ersten Zeile lesen wir ab

$$-\|\vec{v}_0\|_2 \sin(\alpha) + \|\vec{v}_s\|_2 = 0 \text{ also } \sin(\alpha) = \frac{\|\vec{v}_s\|_2}{\|\vec{v}_0\|_2} \text{ d.h. } \sin(\alpha) = \frac{1}{4}.$$

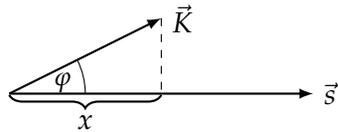
Die letzte Gleichung hat viele Lösungen, aber genau eine Lösung, die zwischen  $0$  und  $90^\circ$  liegt, nämlich

$$\alpha = \arcsin(1/4) \approx 14.48^\circ.$$

Zu (b): Für  $\|\vec{v}_r\|_2$  ergibt sich

$$\|\vec{v}_r\|_2 = \|\vec{v}_0\|_2 \cos(\alpha) \approx 4 \frac{\text{m}}{\text{s}} \cos(14.48^\circ) \approx 3.87 \frac{\text{m}}{\text{s}}.$$

**Beispiel.** Eine Kraft  $\vec{K}$  bewegt einen Massenpunkt in Richtung  $\vec{s}$ . Welche Arbeit wird dabei verrichtet? Wir kennen die Regel "Arbeit ist Kraft mal Weg", jedoch ist dabei nur der Anteil der Kraft  $\vec{K}$  relevant, der in Richtung von  $\vec{s}$  wirkt:



Wir sehen

$$\cos(\varphi) = \frac{x}{\|\vec{K}\|_2} \text{ also } x = \|\vec{K}\|_2 \cos(\varphi).$$

Die Arbeit ist also  $A = x\|\vec{s}\|_2 = \|\vec{K}\|_2\|\vec{s}\|_2 \cos(\varphi)$ .

**Definition 4.3.** Für Vektoren  $\vec{v}, \vec{w} \in \mathbb{R}^n \setminus \{\vec{0}\}$  ist

$$\langle \vec{v}, \vec{w} \rangle := \|\vec{v}\|_2 \|\vec{w}\|_2 \cos(\angle(\vec{v}, \vec{w}))$$

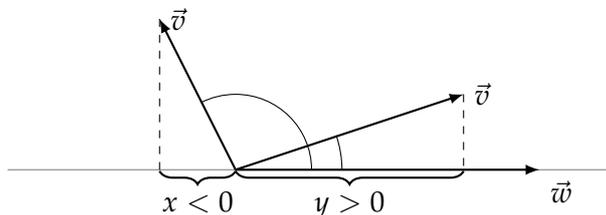
das (Euklidische) Skalarprodukt von  $\vec{v}$  und  $\vec{w}$ . Für  $\vec{v} = \vec{0}$  oder  $\vec{w} = \vec{0}$  ist  $\langle \vec{v}, \vec{w} \rangle := 0$ .

*Bemerkung.* (a) Mit dem Skalarprodukt lässt sich ein Vektor auf den anderen projizieren, denn

$$\frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|_2} = \|\vec{v}\|_2 \cos(\angle(\vec{v}, \vec{w}))$$

ist die Länge der Projektion von  $\vec{v}$  auf  $\vec{w}$ .

Beachte, dass diese Länge ein Vorzeichen hat!



(b) Mit dem Skalarprodukt lässt sich der Winkel zwischen zwei Vektoren bestimmen. So ist zum Beispiel  $\vec{v}$  senkrecht auf  $\vec{w}$ , wenn das Skalarprodukt der Vektoren Null ist, in Formeln

$$\vec{v} \perp \vec{w} \iff \langle \vec{v}, \vec{w} \rangle = 0.$$

■

Auch für Skalarprodukte gibt es eine Liste von Axiomen:

**Definition 4.4.** Ein Skalarprodukt ist eine Abbildung, die zwei Vektoren  $\vec{v}$  und  $\vec{w}$  eines reellen Vektorraums eine Zahl  $\langle \vec{v}, \vec{w} \rangle$  zuordnet, und die folgende Axiome erfüllt:

- (i) **Symmetrie:** Für alle Vektoren  $\vec{v}, \vec{w}$  gilt  $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ .
- (ii) **Linearität im ersten Argument:** Für alle Vektoren  $\vec{u}, \vec{v}, \vec{w}$  und Skalare  $\alpha$  gilt  $\langle \vec{u} + \alpha\vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \alpha\langle \vec{v}, \vec{w} \rangle$ .

Wegen Symmetrie gilt auch Linearität im zweiten Argument!

- (iii) **Positive Definitheit:** Für alle Vektoren  $\vec{v}$  gilt  $\langle \vec{v}, \vec{v} \rangle \geq 0$  und es gilt  $\langle \vec{v}, \vec{v} \rangle = 0$  nur für  $\vec{v} = \vec{0}$ .

Auch für komplexe Vektorräume gibt es den Begriff des Skalarproduktes. Hier fordert man etwas leicht anderes:

- (i) **Konjugierte Symmetrie:** Für alle Vektoren  $\vec{v}, \vec{w}$  gilt  $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle^*$ .
- (ii) **Konjugierte Linearität im ersten Argument:** Für alle Vektoren  $\vec{u}, \vec{v}, \vec{w}$  und Skalare  $\alpha$  gilt  $\langle \vec{u} + \alpha \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \alpha^* \langle \vec{v}, \vec{w} \rangle$ .

Die positive Definitheit bleibt unverändert, und insbesondere muss  $\langle \vec{v}, \vec{v} \rangle$  immer reell sein (auch für komplexe Vektoren). Wegen der konjugierten Symmetrie gilt übrigens Linearität im zweiten Argument:  $\langle \vec{u}, \alpha \vec{v} \rangle = \langle \alpha \vec{v}, \vec{u} \rangle^* = (\alpha^* \langle \vec{v}, \vec{u} \rangle)^* = \alpha \langle \vec{v}, \vec{u} \rangle^* = \alpha \langle \vec{u}, \vec{v} \rangle$ .

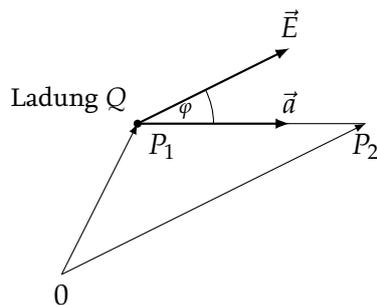
Beachte: Manchmal wird gefordert, dass ein komplexes Skalarprodukt im ersten Argument linear ist; dann bekommt man konjugierte Linearität im zweiten Argument.

## 5 Das Skalarprodukt im $\mathbb{R}^n$

*Beispiel.* Wir betrachten ein konstantes elektrisches Feld mit Feldstärkevektor

$$\vec{E} = \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} 10^6 \text{V m}^{-1}.$$

In diesem Feld soll die Punktladung  $Q = 10^{-7} \text{C}$  vom Punkt  $P_1 = (-2 \ 3 \ 4)^T \text{m}$  geradlinig längs des Richtungsvektors  $\vec{a} = (2 \ -1 \ 2)^T$  um 6 m in den Punkt  $P_2$  verschoben werden.



Um Platz zu sparen wurde hier die Schreibweise  $(-2 \ 3 \ 4)^T$  benutzt. Das  $^T$  bedeutet "transponieren" und macht aus Zeilenvektoren Spaltenvektoren (und umkehrt).

Wir wollen folgende Fragen beantworten:

- Welche Arbeit wird an der Punktladung verrichtet?
- Wie groß ist der Winkel  $\varphi$ ?

Zu (a): Nach Regeln der Elektrotechnik ergibt sich die Kraft  $\vec{F}$  aus

$$\begin{aligned} \vec{F} &= Q \cdot \vec{E} = 10^{-7} \text{C} \cdot \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} 10^6 \text{V m}^{-1} \\ &= \begin{pmatrix} 0.1 \\ -0.3 \\ 0.5 \end{pmatrix} \text{N} \end{aligned}$$

Es gilt  $\text{C} \times \text{V/m} = \text{N}$ .

Der Richtungsvektor  $\vec{a}$  hat die Länge

$$\|\vec{a}\| = \sqrt{2^2 + (-1)^2 + 2^2} \text{m} = 3 \text{m}.$$

Da der Vektor  $\vec{P_1P_2}$  die selbe Richtung wie  $\vec{a}$  hat, aber doppelt so lang sein soll, gilt

$$\vec{s} := \vec{P_1P_2} = 2\vec{a} = \begin{pmatrix} 4 \\ -2 \\ 4 \end{pmatrix} \text{m}.$$

Die Arbeit ist also

$$A = \langle \vec{F}, \vec{s} \rangle = \|\vec{F}\| \|\vec{s}\| \cos(\angle(\vec{F}, \vec{s})).$$

Es gilt

$$\begin{aligned}\|\vec{F}\| &= \sqrt{0.1^2 + 0.3^2 + 0.5^2} \text{N} \approx 0.592 \text{N} \\ \|\vec{s}\| &= 6 \text{m}.\end{aligned}$$

Also ist

$$A = 0.592 \text{N} \cdot 6 \text{m} \cos(\angle(\vec{F}, \vec{s})).$$

Wir müssen allerdings den Winkel  $\varphi$  zu Berechnung von  $A$  kennen! Und zur Berechnung des Winkels  $\varphi$  benötigen wir den Wert von  $A \dots$ ?

Das Beispiel zeigt: Wir brauchen eine Methode, um das Skalarprodukt aus den Komponenten der Vektoren zu berechnen.

Im  $\mathbb{R}^3$  heißen die drei Vektoren

$$\vec{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

die *standard Einheitsvektoren*. Jeden Vektor  $\vec{v} \in \mathbb{R}^3$  können wir mit Hilfe dieser Vektoren schreiben als

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3.$$

Oder auch "kanonische Einheitsvektoren".

Benutzen wir nun die Axiome für Skalarprodukte, so können wir rechnen

$$\begin{aligned}\langle \vec{v}, \vec{w} \rangle &= \langle v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3, w_1 \vec{e}_1 + w_2 \vec{e}_2 + w_3 \vec{e}_3 \rangle \\ &= \langle v_1 \vec{e}_1, w_1 \vec{e}_1 \rangle + \langle v_1 \vec{e}_1, w_2 \vec{e}_2 \rangle + \langle v_1 \vec{e}_1, w_3 \vec{e}_3 \rangle \\ &\quad + \langle v_2 \vec{e}_2, w_1 \vec{e}_1 \rangle + \langle v_2 \vec{e}_2, w_2 \vec{e}_2 \rangle + \langle v_2 \vec{e}_2, w_3 \vec{e}_3 \rangle \\ &\quad + \langle v_3 \vec{e}_3, w_1 \vec{e}_1 \rangle + \langle v_3 \vec{e}_3, w_2 \vec{e}_2 \rangle + \langle v_3 \vec{e}_3, w_3 \vec{e}_3 \rangle\end{aligned}$$

Jetzt ziehen wir die Skalare aus den Skalarprodukten und nutzen aus, dass die standard Einheitsvektoren normiert sind ( $\|\vec{e}_k\|_2 = 1$ ) und paarweise senkrecht aufeinander stehen (d.h.  $\vec{e}_j \perp \vec{e}_k$  wenn  $j \neq k$ ). Letzteres bedeutet

$$\langle \vec{e}_1, \vec{e}_2 \rangle = \langle \vec{e}_1, \vec{e}_3 \rangle = \langle \vec{e}_2, \vec{e}_3 \rangle = \langle \vec{e}_2, \vec{e}_1 \rangle = \langle \vec{e}_3, \vec{e}_1 \rangle = \langle \vec{e}_3, \vec{e}_2 \rangle = 0.$$

Außerdem folgt wegen  $\cos(0) = 1$

$$\langle \vec{e}_1, \vec{e}_1 \rangle = \langle \vec{e}_2, \vec{e}_2 \rangle = \langle \vec{e}_3, \vec{e}_3 \rangle = 1$$

und wir bekommen für das Euklidische Skalarprodukt

$$\langle \vec{v}, \vec{w} \rangle = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

Damit können wir unser Beispiel weiterführen: Die Arbeit ist

$$\begin{aligned} A &= \langle \vec{F}, \vec{s} \rangle = F_1 s_1 + F_2 s_2 + F_3 s_3 \\ &= 0.1 \cdot 4 \text{ Nm} + (-0.3) \cdot (-2 \text{ Nm}) + 0.5 \cdot 4 \text{ Nm} \\ &= 3 \text{ Nm} \end{aligned}$$

Zu (b): Auch können wir den Winkel  $\varphi = \angle(\vec{F}, \vec{s})$  berechnen:

$$\begin{aligned} 3 \text{ Nm} &= \langle \vec{F}, \vec{s} \rangle = \|\vec{F}\| \|\vec{s}\| \cos(\varphi) \\ &\approx 0.592 \text{ N} \cdot 6 \text{ m} \cdot \cos(\varphi). \end{aligned}$$

Es folgt

$$\cos(\varphi) \approx \frac{3}{6 \cdot 0.592} = 0.845.$$

Diese Gleichung hat viele Lösungen, und es gibt genau eine, die zwischen  $0^\circ$  und  $180^\circ$  liegt, nämlich

$$\varphi \approx \arccos(0.845) \approx 32.33^\circ$$

Halten wir noch einmal fest: Eine Abbildung  $\langle \cdot, \cdot \rangle$  die zwei Elementen eines reellen Vektorraumes  $V$  eine reelle Zahl zuordnet, und die die Axiome auf Definition 4.4 erfüllt, heißt *Skalarprodukt* auf  $V$ . Das Paar  $(V, \langle \cdot, \cdot \rangle)$  nennt man dann auch *Innenprodukt-Raum*.

*Beispiel.* (a) Auf  $V = \mathbb{R}^n$  ist

$$\langle \vec{v}, \vec{w} \rangle := \sum_{k=1}^n v_k w_k$$

ein Skalarprodukt, genannt das *Euklidische Skalarprodukt*.

(b) Auf  $V = C[a, b]$  definieren wir für  $f, g \in C[a, b]$

$$\langle f, g \rangle := \int_a^b f(x)g(x)dx.$$

Auch diese Abbildung ist ein Skalarprodukt auf  $V$ , genannt das  $L^2$ -Skalarprodukt.

**Satz 5.1** (Cauchy-Schwarz-Ungleichung). *Ist  $V$  ein reeller Vektorraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ , so gilt für alle  $\vec{v}, \vec{w} \in V$*

$$\langle \vec{v}, \vec{w} \rangle^2 \leq \langle \vec{v}, \vec{v} \rangle \langle \vec{w}, \vec{w} \rangle.$$

*Beweis.*

Für  $\vec{w} = 0$  ist die Behauptung klar, da auf beiden Seiten der Ungleichung 0 steht. Wir wissen, dass für beliebige Vektoren  $\vec{v}$  und  $\vec{w} \neq 0$  und  $\alpha \in \mathbb{R}$  immer gilt

$$0 \leq \langle \vec{v} - \alpha \vec{w}, \vec{v} - \alpha \vec{w} \rangle.$$

Nutzen wir die weiteren Axiome für Skalarprodukte, bekommen

Der Name kommt daher, dass das Skalarprodukt manchmal auch *inneres Produkt* genannt wird.

Auf dem  $\mathbb{C}^n$  ist  $\langle \vec{v}, \vec{w} \rangle := \sum_{k=1}^n v_k^* w_k$  ein Skalarprodukt, dazu später etwas mehr.

wir

$$0 \leq \langle \vec{v}, \vec{v} \rangle - 2\alpha \langle \vec{v}, \vec{w} \rangle + \alpha^2 \langle \vec{w}, \vec{w} \rangle.$$

Da dies für alle  $\alpha \in \mathbb{R}$  gilt, können wir versuchen, die rechte Seite so klein wie möglich zu machen (also über  $\alpha$  zu minimieren). Dies führt auf  $\alpha = \langle \vec{v}, \vec{w} \rangle / \langle \vec{w}, \vec{w} \rangle$  und auf die Ungleichung

$$0 \leq \langle \vec{v}, \vec{v} \rangle - 2 \frac{\langle \vec{v}, \vec{w} \rangle^2}{\langle \vec{w}, \vec{w} \rangle} + \frac{\langle \vec{v}, \vec{w} \rangle^2}{\langle \vec{w}, \vec{w} \rangle}$$

und durch Umstellen erhalten wir die Behauptung.  $\square$

**Satz 5.2.** Ist  $\langle \cdot, \cdot \rangle$  ein Skalarprodukt auf  $V$ , so ist durch  $\|\vec{v}\| := \sqrt{\langle \vec{v}, \vec{v} \rangle}$  eine Norm auf  $V$  definiert.

Man nennt diese Norm die von  $\langle \cdot, \cdot \rangle$  induzierte Norm.

*Beweis.*

Die Positivität von  $\|\cdot\|$  folgt, da wir (per Konvention) die positive Wurzel nehmen, und die Definitheit der Norm folgt aus der positiven Definitheit des Skalarproduktes. Die absolute Homogenität folgt aus

$$\|\alpha \vec{v}\| = \sqrt{\langle \alpha \vec{v}, \alpha \vec{v} \rangle} = \sqrt{\alpha^2 \langle \vec{v}, \vec{v} \rangle} = |\alpha| \|\vec{v}\|.$$

Für die Dreiecksungleichung der Norm rechnen wir zunächst

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle \\ &= \|\vec{v}\|^2 + 2\langle \vec{v}, \vec{w} \rangle + \|\vec{w}\|^2. \end{aligned}$$

Nach der Cauchy-Schwarz-Ungleichung ist

$$|\langle \vec{v}, \vec{w} \rangle| \leq \sqrt{\langle \vec{v}, \vec{v} \rangle} \sqrt{\langle \vec{w}, \vec{w} \rangle} = \|\vec{v}\| \|\vec{w}\|$$

und es folgt

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &\leq \|\vec{v}\|^2 + 2\|\vec{v}\| \|\vec{w}\| + \|\vec{w}\|^2 \\ &= (\|\vec{v}\| + \|\vec{w}\|)^2, \end{aligned}$$

woraus die Behauptung folgt.  $\square$

Ziehen wir die Wurzel, so können wir die Cauchy-Schwarz-Ungleichung äquivalent schreiben

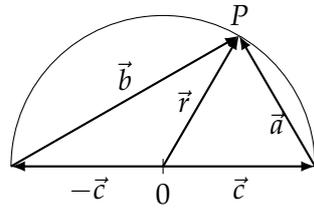
$$|\langle \vec{v}, \vec{w} \rangle| \leq \sqrt{\langle \vec{v}, \vec{v} \rangle} \sqrt{\langle \vec{w}, \vec{w} \rangle}.$$

Die Cauchy-Schwarz-Ungleichung im  $\mathbb{R}^n$  lautet

$$\left| \sum_{k=1}^n v_k w_k \right| \leq \sqrt{\sum_{k=1}^n v_k^2} \sqrt{\sum_{k=1}^n w_k^2}.$$

Hier sieht man die Stärke der Abstraktion: Die Cauchy-Schwarz-Ungleichung gilt für jedes Skalarprodukt, d.h. es folgt auch die Ungleichung  $|\int_a^b f(x)g(x)dx| \leq \sqrt{\int_a^b f(x)^2 dx} \sqrt{\int_a^b g(x)^2 dx}$  für alle stetigen Funktionen  $f$  und  $g$ .

*Beispiel.* Wir können den Satz von Thales folgern: Betrachten wir folgende Skizze:



Wir wollen zeigen, dass der Winkel im Punkt  $P$   $90^\circ$  beträgt: Wir lesen ab, dass  $\vec{r} = \vec{c} + \vec{a} = -\vec{c} + \vec{b}$  gilt. Damit bekommen wir

$$\begin{aligned} \langle \vec{b}, \vec{a} \rangle &= \langle \vec{r} + \vec{c}, \vec{r} - \vec{c} \rangle \\ &= \langle \vec{r}, \vec{r} \rangle - \langle \vec{r}, \vec{c} \rangle + \langle \vec{c}, \vec{r} \rangle - \langle \vec{c}, \vec{c} \rangle \\ &= \|\vec{r}\|^2 - \|\vec{c}\|^2. \end{aligned}$$

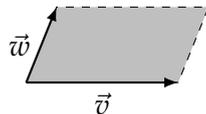
Da  $\vec{r}$  und  $\vec{c}$  gleich lang sind, folgt  $\langle \vec{a}, \vec{b} \rangle = 0$ , was  $\vec{a} \perp \vec{b}$  zeigt.

## 6 Kreuzprodukt und Spatprodukt

Im  $\mathbb{R}^3$  gibt es noch ein anderes wichtiges Produkt zwischen Vektoren. Während das Skalarprodukt aus zwei Vektoren einen Skalar macht, liefert das Kreuzprodukt zweier Vektoren im  $\mathbb{R}^3$  wieder einen Vektor im  $\mathbb{R}^3$ . Der Vektor  $\vec{v} \times \vec{w}$  ist der Vektor im  $\mathbb{R}^3$ , welche durch folgende Eigenschaften eindeutig bestimmt ist:

- (a)  $\|\vec{v} \times \vec{w}\|_2 = \|\vec{v}\|_2 \|\vec{w}\|_2 |\sin(\angle(\vec{v}, \vec{w}))|$ .
- (b)  $\vec{v} \times \vec{w}$  steht senkrecht auf  $\vec{v}$  und  $\vec{w}$ .
- (c) Die Vektoren  $(\vec{v}, \vec{w}, \vec{v} \times \vec{w})$  bilden in dieser Reihenfolge ein *Rechtssystem*, d.h. wenn  $\vec{v}$  in Richtung des Daumens der rechten Hand zeigt und  $\vec{w}$  in Richtung des Zeigefingers der rechten Hand zeigt, so zeigt  $\vec{v} \times \vec{w}$  in Richtung des Mittelfingers der rechten Hand.

Aus der Skizze



erkennen wir, dass das Parallelgramm die Höhe  $\|\vec{w}\|_2 \sin(\angle(\vec{v}, \vec{w}))$  hat. Daher ist  $\|\vec{v}\|_2 \|\vec{w}\|_2 |\sin(\angle(\vec{v}, \vec{w}))|$  die Fläche des Parallelogramms.

Das Vektorprodukt hat weitere wichtige Eigenschaften:

- (i)  $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$
- (ii)  $(\alpha \vec{v}) \times \vec{w} = \alpha(\vec{v} \times \vec{w})$
- (iii)  $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$ .

Suchen wir auch hier eine Möglichkeit, das Vektorprodukt aus den Komponenten der Vektoren zu berechnen: Für jeden Vektor  $\vec{v}$  gilt  $\vec{v} \times \vec{v} = \vec{0}$ . Für die Standard-Einheitsvektoren gilt natürlich ebenfalls  $\vec{e}_i \times \vec{e}_i = \vec{0}$ ,  $i = 1, 2, 3$  und außerdem gilt nach (a), (b) und (c):

$$\begin{aligned}\vec{e}_1 \times \vec{e}_2 &= \vec{e}_3 \\ \vec{e}_2 \times \vec{e}_3 &= \vec{e}_1 \\ \vec{e}_3 \times \vec{e}_1 &= \vec{e}_2.\end{aligned}$$

Rechtssysteme sind aus der Elektrotechnik bekannt: Fließt durch ein Leiterstück  $\vec{l}$  ein Strom  $I$  und befindet es sich in einem homogenen Magnetfeld  $\vec{B}$ , so erfährt es die Lorentzkraft  $\vec{K} = I(\vec{l} \times \vec{B})$ .

Demit bekommen wir

$$\begin{aligned}
 \vec{v} \times \vec{w} &= (v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3) \times (w_1\vec{e}_1 + w_2\vec{e}_2 + w_3\vec{e}_3) \\
 &= v_1w_1 \underbrace{(\vec{e}_1 \times \vec{e}_1)}_{=\vec{0}} + v_1w_2 \underbrace{(\vec{e}_1 \times \vec{e}_2)}_{=\vec{e}_3} + v_1w_3 \underbrace{(\vec{e}_1 \times \vec{e}_3)}_{=-\vec{e}_2} \\
 &\quad + v_2w_1 \underbrace{(\vec{e}_2 \times \vec{e}_1)}_{=-\vec{e}_3} + v_2w_2 \underbrace{(\vec{e}_2 \times \vec{e}_2)}_{=\vec{0}} + v_2w_3 \underbrace{(\vec{e}_2 \times \vec{e}_3)}_{=\vec{e}_1} \\
 &\quad + v_3w_1 \underbrace{(\vec{e}_3 \times \vec{e}_1)}_{=\vec{e}_2} + v_3w_2 \underbrace{(\vec{e}_3 \times \vec{e}_2)}_{=-\vec{e}_1} + v_3w_3 \underbrace{(\vec{e}_3 \times \vec{e}_3)}_{=\vec{0}} \\
 &= (v_2w_3 - v_3w_2)\vec{e}_1 + (v_3w_1 - v_1w_3)\vec{e}_2 + (v_1w_2 - v_2w_1)\vec{e}_3 \\
 &= \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix}.
 \end{aligned}$$

Wir überprüfen die Eigenschaften (i) – (iii):

(i)  $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$ :

$$\begin{aligned}
 \vec{v} \times \vec{w} &= \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix} \\
 &= - \begin{pmatrix} w_2v_3 - w_3v_2 \\ w_3v_1 - w_1v_3 \\ w_1v_2 - w_2v_1 \end{pmatrix} = -\vec{w} \times \vec{v}.
 \end{aligned}$$

(ii)  $(\alpha\vec{v}) \times \vec{w} = \alpha(\vec{v} \times \vec{w})$ :

$$\begin{aligned}
 (\alpha\vec{v}) \times \vec{w} &= \begin{pmatrix} \alpha v_2w_3 - \alpha v_3w_2 \\ \alpha v_3w_1 - \alpha v_1w_3 \\ \alpha v_1w_2 - \alpha v_2w_1 \end{pmatrix} \\
 &= \alpha(\vec{v} \times \vec{w}).
 \end{aligned}$$

(iii)  $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$ :

$$\begin{aligned}
 \vec{u} \times (\vec{v} + \vec{w}) &= \begin{pmatrix} u_2(v_3 + w_3) - u_3(v_2 + w_2) \\ u_3(v_1 + w_1) - u_1(v_3 + w_3) \\ u_1(v_2 + w_2) - u_2(v_1 + w_1) \end{pmatrix} \\
 &= \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix} + \begin{pmatrix} u_2w_3 - u_3w_2 \\ u_3w_1 - u_1w_3 \\ u_1w_2 - u_2w_1 \end{pmatrix} \\
 &= \vec{u} \times \vec{v} + \vec{u} \times \vec{w}
 \end{aligned}$$

Beispiel. (a) Für

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$$

ergibt sich

$$\vec{v} \times \vec{w} = \begin{pmatrix} 2 \cdot 1 - 4 \cdot 3 \\ 4 \cdot (-2) - 1 \cdot 1 \\ 1 \cdot 3 - 2 \cdot (-2) \end{pmatrix} = \begin{pmatrix} -10 \\ -9 \\ 7 \end{pmatrix}.$$

(b) Wir berechnen die Fläche  $F$  des Dreiecks mit den Eckpunkten

$$P_1 = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 6 \\ -1 \\ 1 \end{pmatrix}.$$

Mit den Verbindungsvektoren  $\vec{v} = P_1 P_2$  und  $\vec{w} = P_1 P_3$  gilt

$$F = \frac{1}{2} \|\vec{v} \times \vec{w}\|_2.$$

Es ist

$$\vec{v} = \begin{pmatrix} 1-4 \\ 0-2 \\ 5-3 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \\ 2 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 6-4 \\ -1-2 \\ 1-3 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \\ -2 \end{pmatrix}$$

und es folgt

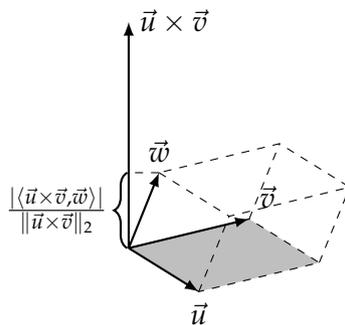
$$\vec{v} \times \vec{w} = \begin{pmatrix} (-2) \cdot (-2) - 2 \cdot (-3) \\ 2 \cdot 2 - (-3) \cdot (-2) \\ (-3) \cdot (-3) - (-2) \cdot 2 \end{pmatrix} = \begin{pmatrix} 10 \\ -2 \\ 13 \end{pmatrix}.$$

Also ergibt sich  $\|\vec{v} \times \vec{w}\| = \sqrt{10^2 + (-2)^2 + 13^2} = \sqrt{273}$  und daher ist  $F = \frac{1}{2} \sqrt{273}$ .

Ein weiteres spezielles Produkt im  $\mathbb{R}^3$  ist das *Spatprodukt* (von drei Vektoren): Für  $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$  ist

$$[\vec{u}, \vec{v}, \vec{w}] := \langle \vec{u} \times \vec{v}, \vec{w} \rangle.$$

Wir machen eine Skizze der Situation:



Der Flächeninhalt der grauen Fläche ist  $\|\vec{u} \times \vec{v}\|_2$  und da  $\vec{u} \times \vec{v}$  senkrecht aus  $\vec{u}$  und  $\vec{v}$  steht, ist das Volumen des Parallelotops welches von  $\vec{u}, \vec{v}$  und  $\vec{w}$  aufgespannt wird gleich  $|\langle \vec{u} \times \vec{v}, \vec{w} \rangle| = |[\vec{u}, \vec{v}, \vec{w}]|$ .

Außerdem gilt:

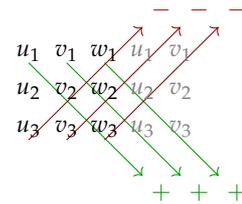
Statt Parallelotop sagt man auch Parallelepiped oder Spat (was den Namen des Produktes erklärt).

- $[\vec{u}, \vec{v}, \vec{w}]$  ist positiv, wenn  $(\vec{u}, \vec{v}, \vec{w})$  ein Rechtssystem bilden, sonst negativ.
- $[\vec{u}, \vec{v}, \vec{w}] = 0$ , wenn  $\vec{u}, \vec{v}, \vec{w}$  *koplanar* sind, d.h. wenn sie in einer Ebene liegen.

Das Spatprodukt lässt sich aus dem Komponenten der Vektoren berechnen als

$$\begin{aligned} [\vec{u}, \vec{v}, \vec{w}] &= \langle \vec{u} \times \vec{v}, \vec{w} \rangle \\ &= \left\langle \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right\rangle \\ &= w_1u_2v_3 - w_1u_3v_2 + w_2u_3v_1 - w_2u_1v_3 + w_3u_1v_2 - w_3u_2v_1 \\ &= u_1v_2w_3 + v_1w_2u_3 + w_1u_2v_3 \\ &\quad - u_3v_2w_1 - v_3w_2u_1 - w_3u_2v_1. \end{aligned}$$

Die Formel für das Spatprodukt kann man sich grafisch wie folgt merken:



**Beispiel.** (a) Was ist das Volumen  $V$  des Parallelotops, welches von

$$\vec{u} = \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} -2 \\ 8 \\ 7 \end{pmatrix}$$

aufgespannt wird?

**Lösung:**

$$\begin{aligned} V &= [\vec{u}, \vec{v}, \vec{w}] \\ &= 1 \cdot 2 \cdot 7 + 3 \cdot 8 \cdot 6 + (-2) \cdot 3 \cdot 2 - 6 \cdot 2 \cdot (-2) - 2 \cdot 8 \cdot 1 - 7 \cdot 3 \cdot 3 \\ &= 14 + 144 + (-12) - (-24) - 16 - 63 \\ &= 91. \end{aligned}$$

(b) Für welchen Wert von  $c$  liegen die vier Punkte

$$P_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 5 \\ 3 \\ 5 \end{pmatrix}, \quad P_4 = \begin{pmatrix} 6 \\ 1 \\ c \end{pmatrix}$$

in einer Ebene?

**Lösung:** Die Punkte liegen in einer Ebene, wenn das Volumen  $V$  des Parallelotops, welches von  $\vec{u} = \overrightarrow{P_1P_2}$ ,  $\vec{v} = \overrightarrow{P_1P_3}$  und  $\vec{w} = \overrightarrow{P_1P_4}$  aufgespannt wird gleich Null ist.

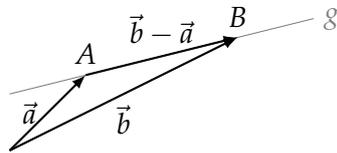
Es gilt also

$$\begin{aligned} V &= \left[ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 0 \\ c-1 \end{pmatrix} \right] \\ &= 0 + 0 + 5 \cdot 1 \cdot 4 - 1 \cdot 2 \cdot 5 - 0 - (c-1) \cdot 1 \cdot 4 \\ &= 20 - 10 - 4c + 4 = 14 - 4c. \end{aligned}$$

Dies ergibt Null, genau dann, wenn  $c = 7/2$  ist.

## 7 Geraden und Ebenen im $\mathbb{R}^3$

Mit Hilfe der Vektorrechnung beschreiben wir eine Gerade durch zwei Punkte  $A$  und  $B$  wie in der Skizze

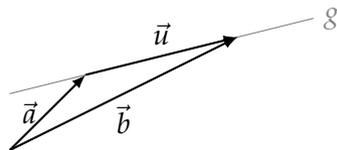


durch

$$g: \{ \vec{x} \mid \vec{x} = \vec{a} + \lambda(\vec{b} - \vec{a}), \lambda \in \mathbb{R} \}.$$

Diese Form nennt man *Zweipunktform* der Geraden (und analog ist die Zweipunktform einer Geraden im  $\mathbb{R}^n$  ebenfalls  $\{ \vec{x} \mid \vec{x} = \vec{a} + \lambda(\vec{b} - \vec{a}), \lambda \in \mathbb{R} \}$  für Vektoren  $\vec{a}, \vec{b} \in \mathbb{R}^n$ ).

Sind Punkt und Richtungsvektor gegeben, wie in dieser Skizze



so ist die Gerade

$$g: \{ \vec{x} \mid \vec{x} = \vec{a} + \lambda \vec{u}, \lambda \in \mathbb{R} \},$$

und diese Form nennt man *Punkt-Richtungsform*.

Im  $\mathbb{R}^2$  ist eine weitere Beschreibung die *Normalenform* durch einen Vektor  $\vec{n}$  und ein Skalar  $\gamma$ , nämlich

$$g: \{ \vec{x} \mid \langle \vec{n}, \vec{x} \rangle = \gamma \}$$

Man kann diese Form z.B. aus der Punkt-Richtungsform wie folgt herleiten: Es gilt

$$x_1 = a_1 + \lambda u_1$$

$$x_2 = a_2 + \lambda u_2.$$

Die erste Gleichung lösen wir nach  $\lambda$  und bekommen  $\lambda = (x_1 - a_1) / u_1$  (falls  $u_1 \neq 0$ ). Wir setzen dies in die zweite Gleichung ein und bekommen  $x_2 = a_2 + \frac{x_1 - a_1}{u_1} u_2$ , was wir umformen zu

$$u_2 x_1 - u_1 x_2 = a_1 u_2 - a_2 u_1.$$

Wir sehen also, dass wir wie folgt die Normalenform bekommen:

$$\vec{n} = \begin{pmatrix} u_2 \\ -u_1 \end{pmatrix}, \quad \gamma = a_1 u_2 - a_2 u_1,$$

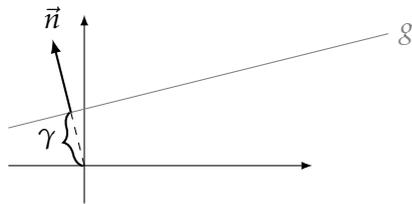
Ist  $u_1 = 0$  so argumentieren wir mit  $u_2$  statt  $u_1$ . Sind sowohl  $u_1$ , also auch  $u_2$  gleich Null, so ist der Richtungsvektor  $\vec{u}$  Null, und es gibt gar keine Gerade, sondern  $g$  ist nur ein Punkt.

denn dann gilt  $\langle \vec{n}, \vec{x} \rangle = \gamma$  für jedes  $\vec{x}$  der Form  $\vec{x} = \vec{a} + \lambda \vec{u}$ . Wir bemerken

$$\langle \vec{n}, \vec{u} \rangle = u_2 u_1 - u_1 u_2 = 0,$$

d.h. der Vektor  $\vec{n}$  steht senkrecht auf dem Richtungsvektor  $\vec{u}$ . Da man statt "steht senkrecht auf" auch "ist normal zu" sagt, erklärt dies den Namen Normalenform. Den Vektor  $\vec{n}$  nennt man auch „Normalevektor“ der Ebene.

Man kann die Normalenform weiter spezialisieren: In der Darstellung  $\langle \vec{n}, \vec{x} \rangle = \gamma$  können wir  $\vec{n}$  und  $\gamma$  beliebig skalieren: Ersetzen wir  $\vec{n}$  und  $\gamma$  durch  $\alpha \vec{n}$  und  $\alpha \gamma$ , so ändert sich die Menge der  $\vec{x}$  mit  $\langle \vec{n}, \vec{x} \rangle = \gamma$  nicht. Tut man dies, so dass  $\|\vec{n}\| = 1$  und  $\gamma \geq 0$  gilt, spricht von der *Hesse'schen Normalform*:

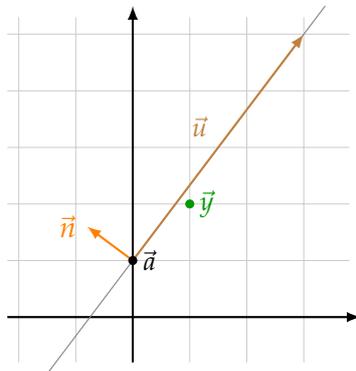


Genauer: Wir ersetzen  $\vec{n}$  durch  $\text{sign}(\gamma)\vec{n}/\|\vec{n}\|$  und  $\gamma$  durch  $\text{sign}(\gamma)\gamma/\|\vec{n}\|$ , wobei  $\text{sign}(\gamma)$  das Vorzeichen von  $\gamma$  ist.

- (a) Da  $\vec{n}$  normiert ist, ist  $\langle \vec{n}, \vec{x} \rangle$  die Länge der Projektion von  $\vec{x}$  auf  $\vec{n}$ , d.h.  $g$  ist die Menge aller  $\vec{x}$ , die Projektion auf  $\vec{n}$  die Länge  $\gamma$  haben.
- (b) Da  $\alpha \vec{n}$  genau dann auf  $g$  liegt, wenn  $\gamma = \langle \vec{n}, \alpha \vec{n} \rangle = \alpha \|\vec{n}\|_2 = \alpha$  gilt, ist  $\gamma$  der Abstand der Geraden vom Ursprung.
- (c) Für beliebiges  $\vec{x}$  gibt  $d = \langle \vec{n}, \vec{x} \rangle - \gamma$  gibt Information über den Abstand von  $\vec{x}$  zu  $g$ :
- Ist  $d > 0$ , so liegen  $\vec{x}$  und  $\vec{0}$  auf verschiedenen Seiten der Geraden.
  - Ist  $d < 0$ , so liegen  $\vec{x}$  und  $\vec{0}$  auf derselben Seite der Geraden.
  - Ist  $d = 0$ , so liegt  $\vec{x}$  auf der Geraden.

*Beispiel.* Es sei

$$g: \left\{ \vec{x} \mid \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \lambda \in \mathbb{R} \right\}, \text{ also } \vec{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \vec{u} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$



Wir bestimmen nun  $\vec{n}$  und  $\gamma$  aus der Normalenform und bekommen

$$\vec{n} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}, \quad \gamma = 0 \cdot 4 - 1 \cdot 3 = -3.$$

Die Gerade ist also bestimmt durch

$$g : \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid 4x_1 - 3x_2 = -3 \right\}.$$

Wir normieren noch zu

$$\vec{n} = -\frac{1}{\sqrt{3^2+4^2}} \begin{pmatrix} 4 \\ -3 \end{pmatrix} = \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix}, \quad \gamma = \frac{3}{5}$$

und bekommen die Hesse'sche Normalform

$$g : \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid -\frac{4}{5}x_1 + \frac{3}{5}x_2 = +\frac{3}{5} \right\}.$$

Damit lässt sich einfach prüfen, dass der Punkt  $\vec{y} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  wegen

$$d = \langle \vec{n}, \vec{y} \rangle - \gamma = \left\langle \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle - \frac{3}{5} = -\frac{4}{5} + \frac{6}{5} - \frac{3}{5} = -\frac{1}{5} < 0$$

auf der gleichen Seite wie der Ursprung liegt und den Abstand  $\frac{1}{5}$  zur Geraden hat.

In drei Dimensionen wird eine Ebene  $E$  durch drei Punkte  $\vec{a}$ ,  $\vec{b}$ ,  $\vec{c}$  festgelegt:

$$E : \{ \vec{x} = \vec{a} + \lambda(\vec{b} - \vec{a}) + \mu(\vec{c} - \vec{a}), \lambda, \mu \in \mathbb{R} \}.$$

Mit den Richtungsvektoren  $\vec{u} = \vec{b} - \vec{a}$  und  $\vec{v} = \vec{c} - \vec{a}$  ist dies

$$E : \{ \vec{x} = \vec{a} + \lambda\vec{u} + \mu\vec{v}, \lambda, \mu \in \mathbb{R} \}.$$

Um die Normalenform zu bestimmen suchen wir einen Vektor  $\vec{n}$  mit  $\vec{n} \perp \vec{u}$  und  $\vec{n} \perp \vec{v}$ . Nach den Eigenschaften des Kreuzproduktes wissen wir, dass wir  $\vec{n} = \vec{u} \times \vec{v}$  nehmen können. Damit gilt für  $\vec{x} \in E$

$$\langle \vec{x}, \vec{n} \rangle = \langle \vec{a}, \vec{n} \rangle + \lambda \underbrace{\langle \vec{u}, \vec{n} \rangle}_{=0} + \mu \underbrace{\langle \vec{v}, \vec{n} \rangle}_{=0},$$

also

$$\langle \vec{x}, \vec{n} \rangle = \gamma$$

mit  $\gamma = \langle \vec{a}, \vec{n} \rangle$ . Auch hier können wir  $\vec{n}$  und  $\gamma$  normieren (d.h. auf  $\|\vec{n}\|_2 = 1$  und  $\gamma > 0$  bringen) und haben damit die Hessesche Normalform einer Ebene im  $\mathbb{R}^3$

Um Geraden um  $\mathbb{R}^3$  zu beschreiben, nehmen wir die Punkt-Richtungsform  $\vec{x} = \vec{a} + \lambda \vec{u}$  und versuchen  $\lambda$  zu eliminieren:

$$\left. \begin{array}{l} x_1 = a_1 + \lambda u_1 \\ x_2 = a_2 + \lambda u_2 \\ x_3 = a_3 + \lambda u_3 \end{array} \right\} \implies \left\{ \begin{array}{l} \lambda = \frac{x_1 - a_1}{u_1} \\ \lambda = \frac{x_2 - a_2}{u_2} \\ \lambda = \frac{x_3 - a_3}{u_3} \end{array} \right.$$

(wenn alle  $u_i \neq 0$ ). Wir bekommen dadurch also zwei Gleichungen:

$$\frac{x_1 - a_1}{u_1} = \frac{x_2 - a_2}{u_2} = \frac{x_3 - a_3}{u_3}.$$

Wir sehen, dass die Gerade durch zwei Gleichungen beschrieben wird. Jede Gleichung beschreibt eine Ebene und daher bekommen wir die Gerade als Schnittmenge von zwei Ebenen.

Um auf eine Darstellung analog zur Normalform zu kommen, brauchen wir zwei Vektoren

$$\vec{m} = \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix}, \quad \vec{n} = \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix}$$

welche auf einem Richtungsvektor  $\vec{u}$  senkrecht stehen. Damit bekommen wir, wenn wir das Skalarprodukt der Geradengleichung mit  $\vec{m}$ , bzw.  $\vec{n}$  bilden

$$\langle \vec{x}, \vec{m} \rangle = \langle \vec{a}, \vec{m} \rangle + \lambda \underbrace{\langle \vec{u}, \vec{m} \rangle}_{=0 \text{ da } \vec{u} \perp \vec{m}}, \quad \langle \vec{x}, \vec{n} \rangle = \langle \vec{a}, \vec{n} \rangle + \lambda \underbrace{\langle \vec{u}, \vec{n} \rangle}_{=0 \text{ da } \vec{u} \perp \vec{n}}.$$

Die beiden Gleichungen sind also

$$\begin{aligned} \alpha_1 x_1 + \beta_1 x_2 + \gamma_1 x_3 &= \delta_1 := \langle \vec{a}, \vec{m} \rangle \\ \alpha_2 x_1 + \beta_2 x_2 + \gamma_2 x_3 &= \delta_2 := \langle \vec{a}, \vec{n} \rangle \end{aligned}$$

und beschreiben jeweils eine Ebene. Wir bekommen die Gerade wiederum als Schnittmenge von zwei Ebenen.

## 8 Untervektorräume, Basen und Dimension

Erinnere: Ein Vektorraum ist eine nicht-leere Menge, auf der eine Vektoraddition und eine Skalierung erklärt sind, für welche die Vektorraum-Axiome (VR1) und (VR2) aus Abschnitt 3 gelten. Wir unterscheiden bisher zwischen  $\mathbb{R}$ -Vektorräumen und  $\mathbb{C}$ -Vektorräumen. In den meisten Fällen gelten die Resultate für beide Fälle und wir schreiben  $\mathbb{K}$  für „ $\mathbb{R}$  oder  $\mathbb{C}$ “.

Folgende Beispiele kennen wir schon:

- (a)  $\mathbb{K}^n$  als reeller bzw. komplexer Vektorraum
- (b) Die Räume der Polynome vom Grad  $\leq n$ :

$$\Pi_n := \left\{ \sum_{k=0}^n a_k x^k \mid a_k \in \mathbb{K} \right\}.$$

- (c) Den Raum

$$C(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{K} \mid f \text{ stetig}\}$$

der stetigen Funktionen von  $\mathbb{R}$  nach  $\mathbb{K}$ .

Wir können auch den Raum *aller* Polynome bilden (d.h. von beliebigem Grad). Dieser ist

$$\Pi := \bigcup_{n \in \mathbb{N}} \Pi_n$$

und selbst ein Vektorraum.

Es gilt offenbar

$$\Pi_n \subset \Pi \subset C(\mathbb{R}).$$

**Definition 8.1** (Unterraum). Eine Teilmenge  $W \subset V$  heißt *Untervektorraum* (UVR, auch linearer Unterraum oder linearer Teilraum) eines Vektorraums  $V$ , wenn gilt

$$(UVR) \begin{cases} (i) & \vec{0} \in W \\ (ii) & \vec{v}, \vec{w} \in W \implies \vec{v} + \vec{w} \in W \\ (iii) & \vec{v} \in W, \alpha \in \mathbb{K} \implies \alpha \vec{v} \in W \end{cases}$$

Es folgt: Ein Untervektorraum  $W$  ist bezüglich der Vektorraumoperationen in  $V$  selbst ein Vektorraum.

*Beispiel.* (a) Die Kugel  $W := \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\|_2 \leq 1\}$  ist *kein* UVR des  $\mathbb{R}^n$ , denn es gilt  $\vec{e}_1 \in W$ , aber  $2\vec{e}_1 \notin W$ .

(b) Die Menge  $W = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid xy = 0 \right\}$  ist ebenfalls *kein* UVR des  $\mathbb{R}^2$ , denn es gilt  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in W$ , aber  $\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin W$ .

- (c) Für jeden Vektor  $\vec{v} \in \mathbb{R}^n$ ,  $\vec{v} \neq \vec{0}$  ist die von  $\vec{v}$  aufgespannte Gerade

$$G := \{\alpha \vec{v} \mid \alpha \in \mathbb{R}\}$$

ein UVR des  $\mathbb{R}^n$ . Für zwei Vektoren  $\vec{v}, \vec{w}$  ist

$$E := \{\alpha \vec{v} + \beta \vec{w} \mid \alpha, \beta \in \mathbb{R}\}$$

ebenfalls ein UVR. Sind  $\vec{v}$  und  $\vec{w}$  nicht parallel, so ist  $E$  eine Ebene.

- (d)  $\Pi_n$  ist UVR von  $\Pi$  und  $\Pi$  ist UVR von  $C(\mathbb{R})$ .

**Definition 8.2** (Linearkombination, Spann). Für  $\vec{v}_1, \dots, \vec{v}_m \in V$  heißt jeder Vektor

$$\vec{x} = \sum_{i=1}^m \alpha_i \vec{v}_i, \quad \alpha_i \in \mathbb{K}$$

eine *Linearkombination* der Vektoren  $\vec{v}_1, \dots, \vec{v}_m$  (und Koeffizienten  $\alpha_1, \dots, \alpha_m$ ).

Die Menge aller Linearkombinationen von  $\vec{v}_1, \dots, \vec{v}_m$  heißt der *Spann* (auch die lineare Hülle) von  $\vec{v}_1, \dots, \vec{v}_m$ , geschrieben

$$\text{Spann}(\vec{v}_1, \dots, \vec{v}_m) := \left\{ \sum_{i=1}^m \alpha_i \vec{v}_i \mid \alpha_i \in \mathbb{K} \right\}.$$

Der Spann ist immer ein UVR von  $V$ !

Hat man einen UVR  $W$ , so interessiert man sich für möglichst kleine Mengen von Vektoren, die  $W$  aufspannen (d.h.  $W$  ist der Spann der Menge). Eine Menge, die die minimal mögliche Anzahl von Vektoren enthält nennt man dann *Basis* von  $W$ . Ist  $B = \{\vec{v}_1, \dots, \vec{v}_m\}$  eine solche Menge, so kann kein  $\vec{v}_k$  eine Linearkombination der anderen  $\vec{v}_i$  sein (sonst könnte man  $\vec{v}_k$  gleich weglassen). Dieses Konzept wird im Begriff „lineare Unabhängigkeit“ formalisiert.

**Definition 8.3** (Lineare Unabhängigkeit, Basis). Eine Menge  $B = \{\vec{v}_1, \dots, \vec{v}_m\}$  von Vektoren heißt *linear unabhängig*, wenn für alle Skalare  $\alpha_1, \dots, \alpha_m$  gilt

$$\sum_{i=1}^m \alpha_i \vec{v}_i = \vec{0} \implies \alpha_1 = \dots = \alpha_m = 0.$$

Die Menge  $B$  heißt *Basis* von  $W$ , wenn sie linear unabhängig ist und  $W = \text{Spann}(B)$  gilt.

Eine Menge  $B$ , die nicht linear unabhängig ist, heißt *linear abhängig*.

Intuitiv soll lineare Unabhängigkeit von einer Menge bedeuten, dass die Vektoren der Menge „alle in verschiedene Richtung zeigen“.

Mit anderen Worten:  $\vec{v}_1, \dots, \vec{v}_m$  sind linear unabhängig, wenn man den Nullvektor nur durch die sogenannte *triviale Linearkombination* (d.h. mit lauter Nullkoeffizienten) darstellen kann.

*Beispiel.* Die beiden Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

bilden eine Basis des  $\mathbb{R}^2$ , denn:

1. Sie sind linear unabhängig, denn ist

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 = \begin{pmatrix} \alpha_1 + \alpha_2 \\ \alpha_1 + 2\alpha_2 \end{pmatrix} = \vec{0}$$

so folgt aus der ersten Zeile  $\alpha_1 = -\alpha_2$  und damit aus der zweiten Zeile  $-\alpha_2 + 2\alpha_2 = \alpha_2 = 0$ , was wiederum  $\alpha_1 = 0$  bringt.

2. Sie spannen den ganzen  $\mathbb{R}^2$  auf, d.h. jeder Vektor  $\vec{v} \in \mathbb{R}^2$  lässt sich als  $\vec{v} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2$  darstellen.

Man überprüft leicht, dass dies mit  $\alpha_1 = 2v_1 - v_2$  und  $\alpha_2 = -v_1 + v_2$  funktioniert.

Die Vektoren  $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  und  $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  bilden ebenfalls eine Basis des  $\mathbb{R}^2$ , die sogenannte *kanonische Basis*.

*Bemerkung.* • Zwei Vektoren  $\vec{v}, \vec{w} \neq \vec{0}$  sind genau dann linear unabhängig, wenn die nicht parallel sind.

- Ist  $\{\vec{v}_1, \dots, \vec{v}_m\}$  eine Basis von  $W$ , so ist die *Basisdarstellung* eines jeden Vektors  $\vec{v} \in W$

$$\vec{v} = \sum_{i=1}^m \alpha_i \vec{v}_i$$

eindeutig, d.h. die Koeffizienten  $\alpha_i$  sind eindeutig bestimmt.

- Die Vektoren

$$\vec{e}_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^n, i = 1, \dots, n$$

welche genau eine 1 an der  $i$ -ten Stelle haben, bilden die *kanonische Basis* des  $\mathbb{K}^n$ . ■

*Beispiel.* Die Menge

$$W := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_1 - x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$$

ist ein UVR des  $\mathbb{R}^3$ , denn:

Hat man zwei verschiedene Darstellungen  $\sum_i \alpha_i \vec{v}_i = \sum_i \beta_i \vec{v}_i$ , so folgt  $\sum_i (\alpha_i - \beta_i) \vec{v}_i = \vec{0}$  und wegen der linearen Unabhängigkeit muss  $\alpha_i = \beta_i$  für alle  $i$  gelten.

(i)  $\vec{0} \in W$  (für  $x_1 = x_2 = 0$ ),

(ii) Ist

$$\vec{u} = \begin{pmatrix} x_1 \\ x_2 \\ x_1 - x_2 \end{pmatrix}, \vec{v} = \begin{pmatrix} y_1 \\ y_2 \\ y_1 - y_2 \end{pmatrix} \in W$$

so ist mit  $z_1 = x_1 + y_1, z_2 = x_2 + y_2$

$$\vec{u} + \vec{v} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ (x_1 + y_1) - (x_2 + y_2) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_1 - z_2 \end{pmatrix} \in W.$$

(iii) Für gleiches  $\vec{u}$  und  $\alpha \in \mathbb{R}$  ist mit  $z_i = \alpha x_i$

$$\alpha \vec{u} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha(x_1 - x_2) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_1 - z_2 \end{pmatrix} \in W.$$

Die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

bilden eine Basis von  $W$ , denn:

(i) Die Vektoren  $\vec{v}_1$  und  $\vec{v}_2$  sind linear unabhängig: Ist

$$\vec{0} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 = \begin{pmatrix} \alpha_1 + \alpha_2 \\ \alpha_2 \\ \alpha_1 \end{pmatrix}$$

so folgt aus der zweiten Zeile  $\alpha_2 = 0$  und aus der dritten Zeile  $\alpha_1 = 0$ .

(ii) Die Vektoren  $\vec{v}_1$  und  $\vec{v}_2$  spannen ganz  $W$  auf: Ist  $\vec{v} \in W$ , so gibt es  $x_1, x_2$ , so dass

$$\vec{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_1 - x_2 \end{pmatrix}$$

Mit  $\alpha_1 = x_1 - x_2$  und  $\alpha_2 = x_2$  folgt dann

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 = \begin{pmatrix} \alpha_1 + \alpha_2 \\ \alpha_2 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 + x_2 \\ x_2 \\ x_1 - x_2 \end{pmatrix} = \vec{v}.$$

Fragen wir uns jetzt, wie man eine Basis eines Untervektorraums  $W$  von  $V$  konstruieren kann, wenn  $W$  durch eine endliche Menge ausgespannt ist, also wenn gilt

$$W = \text{Spann}(\vec{v}_1, \dots, \vec{v}_m).$$

Es soll nun eine Basis  $\{\vec{w}_1, \dots, \vec{w}_r\}$  konstruiert werden.

- (a) Falls  $\{\vec{v}_1, \dots, \vec{v}_m\}$  linear unabhängig sind, setze  $r := m$  und  $\vec{w}_i = \vec{v}_i, i = 1, \dots, m$  und wir sind fertig.
- (b) Andernfalls gibt es  $\alpha_1, \dots, \alpha_m \in \mathbb{R}$  mit  $\sum_{i=1}^m \alpha_i \vec{v}_i = \vec{0}$  wobei nicht alle  $\alpha_i$  Null sind. Durch eventuelles Umm Nummerieren können wir  $\alpha_m \neq 0$  erreichen. Dann können wir nach  $\vec{v}_m$  auflösen

$$\vec{v}_m = \sum_{i=1}^{m-1} \left(-\frac{\alpha_i}{\alpha_m}\right) \vec{v}_i.$$

Wir sehen also, dass  $W$  schon von  $\vec{v}_1, \dots, \vec{v}_{m-1}$  aufgespannt wird. Wir ersetzen also  $m$  durch  $m-1$  und gehen zu Schritt (a).

Das Verfahren muss terminieren, denn wir brechen spätestens ab, wenn kein  $\vec{v}_i$  mehr da ist, also für  $r = 0$  (dann war  $W = \{\vec{0}\}$ ).

Wir folgern:

*Jeder Vektorraum, der von einer endlichen Menge aufgespannt wird, hat eine Basis.*

**Satz 8.4.** Ist  $W \subset V$  ein UVR von  $V$  mit Basis  $\{\vec{v}_1, \dots, \vec{v}_m\}$ , dann sind je  $n$  Vektoren  $\vec{w}_1, \dots, \vec{w}_n \in W$  mit  $n > m$  linear abhängig.

**Korollar 8.5** (+Definition). Die Anzahl der Elemente einer Basis von  $W$  ist für alle Basen gleich. Diese Anzahl heißt Dimension von  $W$  und wird mit  $\dim(W)$  bezeichnet.

Haben wir eine Basis  $B$  mit  $n$  Elementen und eine weitere Basis  $C$  mit  $m$  Elementen, so bekommen wir sofort einen Widerspruch: Ist  $m > n$ , so ist  $C$  linear abhängig und ist  $n > m$ , so ist  $B$  linear abhängig.

**Beispiel.** (a) Es gilt  $\dim(\mathbb{R}^n) = n$ , denn eine Basis ist die Standardbasis  $\{\vec{e}_1, \dots, \vec{e}_n\}$ .

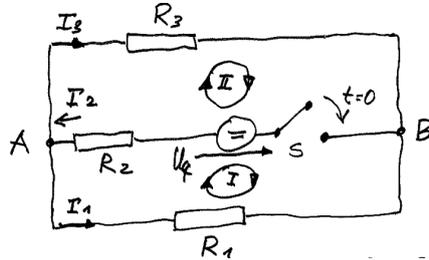
(b) Für  $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_1 - x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$  (vgl. voriges Beispiel) ist  $\dim(W) = 2$ .

(c) Für den Raum  $\Pi_n$  der Polynome vom Grad bis zu  $n$  gilt  $\dim(\Pi_n) = n + 1$ , denn die Monome  $\{1, x, x^2, \dots, x^n\}$  bilden eine Basis.

## 9 Lineare Gleichungssysteme und Matrizen

Lineare Gleichungssysteme sind fundamental für zahlreiche Anwendungen, aber auch für den Aufbau der linearen Algebra.

*Beispiel.* Wir betrachten den verzweigten Schaltkreis



mit ohmschen Widerständen  $R_1, R_2, R_3$ . Es wird eine Gleichspannung mit Quellenspannung  $U_q$  eingeschaltet. Wir wollen die Zweigströme  $I_1, I_2, I_3$  berechnen.

Die Knotenregel im Knoten A und die Maschenregel für die Maschen I und II ergeben die Gleichungen

$$\begin{aligned} -I_1 + I_2 - I_3 &= 0 \\ -R_1 I_1 - R_2 I_2 + U_q &= 0 \\ R_2 I_2 + R_3 I_3 - U_q &= 0. \end{aligned}$$

Wir haben also ein lineares Gleichungssystem mit drei Unbekannten und drei Gleichungen.

Man schreibt die übersichtlich in Matrix-Vektor Form:

$$\underbrace{\begin{pmatrix} -1 & +1 & -1 \\ -R_1 & -R_2 & 0 \\ 0 & R_2 & R_3 \end{pmatrix}}_{=:A} \underbrace{\begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix}}_{=:x} = \underbrace{\begin{pmatrix} 0 \\ -U_q \\ U_q \end{pmatrix}}_{=:b}.$$

Dies ist ein *inhomogenes lineares Gleichungssystem* (inhomogen, da die rechte Seite  $\vec{b}$  nicht der Nullvektor ist) mit *Koeffizientenmatrix*  $A$ .

Im allgemeinen sind Matrizen nichts anderes als rechteckige Felder von Zahlen. Wir bezeichnen die Menge der reellen Matrizen mit  $m$  Zeilen und  $n$  Spalten mit  $\mathbb{K}^{m \times n}$ . Für eine Matrix  $A$  schreibt man üblicherweise  $a_{ij}$  für den Eintrag in der  $i$ -ten Zeile und  $j$ -ten Spalte. Man schreibt auch

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

oder lässt die Erklärung für den Bereich der Indizes weg, wenn das aus dem Kontext klar ist.

Die Menge der Matrizen bildet einen Vektorraum, wenn wir Addition und Skalierung definieren als

$$\begin{aligned} A + B &:= (a_{ij} + b_{ij}) \\ \lambda A &:= (\lambda a_{ij}). \end{aligned}$$

Die Knotenregel besagt, dass die Summe der zu- und abfließenden Ströme sich in jedem Knoten zu Null summieren

Die Maschenregel besagt, dass die Summe der Spannungen in jeder Netzmasche gleich Null ist.

Um sich die Reihenfolge der Indizes zu merken bietet sich der Satz „Zeile zuerst, Spalte später“ an.

Beachte: Addition ist nur für Matrizen gleicher Größe erklärt!

Das *Matrix-Vektor Produkt* ist erklärt, wenn die Anzahl der Spalten einer Matrix  $A$  der Anzahl der Einträge in einem Vektor  $\vec{x}$  entspricht, also wenn  $A \in \mathbb{K}^{m \times n}$  ist, muss  $\vec{x} \in \mathbb{K}^n$  gelten. In diesem Fall ist

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

Es gibt verschiedene Möglichkeiten, sich das zu merken:

1. Der  $k$ te Eintrag von  $A\vec{x}$  ist (bis auf eine Konjugation im komplexen) das Skalarprodukt der  $k$ ten Zeile von  $A$  mit  $\vec{x}$ .

Im reellen schreiben wir  $A = (\vec{a}^{(1)} \dots \vec{a}^{(m)})^T$  mit  $\vec{a}^{(m)} \in \mathbb{R}^n$  (d.h.  $(\vec{a}^{(m)})^T$  ist die  $i$ -te Zeile von  $A$ ), so ist

$$A\vec{x} = \begin{pmatrix} \langle \vec{a}^{(1)}, \vec{x} \rangle \\ \dots \\ \langle \vec{a}^{(m)}, \vec{x} \rangle \end{pmatrix}$$

2. Der  $k$ te Eintrag von  $A\vec{x}$  ist  $\sum_{j=1}^n a_{kj}x_j$  (die Indizes, die nebeneinander stehen, müssen gleich sein, der übrige gibt den Eintrag).
3. Der Vektor  $A\vec{x}$  ist eine Linearkombination der Spalten von  $A$  und die Einträge von  $\vec{x}$  sind die Koeffizienten.

Schreiben wir, anders als in 1.,  $A = (\vec{a}^{(1)} \dots \vec{a}^{(n)})$ , wobei jetzt  $\vec{a}^{(j)} \in \mathbb{K}^m$  die  $j$ -te Spalte von  $A$  ist, so ist

$$A\vec{x} = x_1\vec{a}^{(1)} + \dots + x_n\vec{a}^{(n)} = \sum_{j=1}^n x_j\vec{a}^{(j)}$$

*Beispiel.*

$$\begin{pmatrix} 4 & 3 & 0 & 1 & 2 \\ 2 & 1 & 4 & 0 & 1 \\ 0 & 0 & 4 & 1 & 0 \\ 2 & 0 & 1 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4+0+0+1+0 \\ 2+0+12+0+0 \\ 0+0+12+1+0 \\ 2+0+3+0+0 \end{pmatrix} = \begin{pmatrix} 5 \\ 14 \\ 13 \\ 5 \end{pmatrix}.$$

Es gelten folgende Rechenregeln:

- (i)  $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$
- (ii)  $A(\lambda\vec{x}) = \lambda(A\vec{x}) = (\lambda A)\vec{x}$
- (iii)  $(A + B)\vec{x} = A\vec{x} + B\vec{x}$ .

Die Aussagen (i) und (ii) bedeuten, dass die Abbildung  $T(\vec{x}) := A\vec{x}$  eine *lineare Abbildung* ist.

Es ist auch eine Multiplikation von Matrizen definiert, jedenfalls, wenn die Anzahl der Spalten der ersten Matrix gleich der Anzahl der Zeilen der zweiten Matrix ist: Das Produkt von  $A \in \mathbb{R}^{l \times m}$  und  $B \in \mathbb{R}^{m \times n}$  ist  $C := AB$  und es wird so erklärt, dass für alle  $\vec{x} \in \mathbb{R}^n$  gilt

$$A(B\vec{x}) = (AB)\vec{x}.$$

Wir rechnen die linke Seite aus, und zwar zuerst  $B\vec{x}$ :

$$B\vec{x} = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n b_{1j}x_j \\ \vdots \\ \sum_{j=1}^n b_{mj}x_j \end{pmatrix}$$

also ist

$$\begin{aligned} A(B\vec{x}) &= \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{l1} & \dots & a_{lm} \end{pmatrix} \begin{pmatrix} \sum_{j=1}^n b_{1j}x_j \\ \vdots \\ \sum_{j=1}^n b_{mj}x_j \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^m [a_{1k} (\sum_{j=1}^n b_{kj}x_j)] \\ \vdots \\ \sum_{k=1}^m [a_{lk} (\sum_{j=1}^n b_{kj}x_j)] \end{pmatrix} = \begin{pmatrix} (\sum_{j=1}^n (\sum_{k=1}^m a_{1k}b_{kj})x_j) \\ \vdots \\ (\sum_{j=1}^n (\sum_{k=1}^m a_{lk}b_{kj})x_j) \end{pmatrix} \\ &=: \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{l1} & \dots & c_{ln} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

Wir sehen, dass  $C = AB$  eine  $l \times n$ -Matrix ist und die Einträge

$$c_{ij} = \sum_{k=1}^m a_{ik}b_{kj}$$

hat. Auch für die Matrix-Matrix-Multiplikation gibt es Merkgeln:

1. Der  $ij$ te Eintrag von  $AB$  ist das Skalarprodukt der  $i$ -ten Zeile von  $A$  und  $j$ ten Spalte von  $B$ . (In komplexen Vektorräumen stimmt das nicht ganz, da bei der Matrix-Matrix-Multiplikation nichts konjugiert wird.)
2. Die  $j$ te Spalte von  $AB$  ist eine Linearkombination der Spalten von  $A$  deren Koeffizienten die Einträge der  $j$ ten Spalte  $B$  sind.

*Beispiel.*

$$\begin{pmatrix} 4 & 3 & 0 & 1 & 2 \\ 2 & 1 & 4 & 0 & 1 \\ 0 & 0 & 4 & 1 & 0 \\ 2 & 0 & 1 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 21 & 13 \\ 14 & 8 & 8 \\ 13 & 1 & 7 \\ 5 & 4 & 13 \end{pmatrix}$$

Um sich die Größe von  $AB$  zu merken, merke man sich, dass die „mittleren Dimensionen“, welche gleich sein müssen, wegfallen und die äußeren stehen bleiben:  $(l \times n) \cdot (n \times m) \rightarrow (l \times m)$ .

Es gelten folgende Rechenregeln:

$$\begin{aligned} (i) \quad & (A + B)C = AC + BC \\ (ii) \quad & A(B + C) = AB + AC \\ (iii) \quad & A(BC) = (AB)C \end{aligned}$$

Eine besondere Matrix ist die  $k \times k$  Einheitsmatrix. Diese ist

$$I_k := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathbb{K}^{k \times k}$$

und hat die Eigenschaft, dass für  $A \in \mathbb{K}^{m \times n}$  und  $x \in \mathbb{K}^n$  gilt

$$AI_n = A, I_m A = A, \text{ und } I_n \vec{x} = \vec{x}.$$

Wichtig: Das Matrix-Matrix-Produkt ist im Allgemeinen *nicht* kommutativ:

- Einerseits muss  $BA$  nicht unbedingt erklärt sein, wenn  $AB$  erklärt ist. Ist zum Beispiel  $A \in \mathbb{K}^{l \times m}$  und  $B \in \mathbb{K}^{m \times n}$ , so ist  $AB$  erklärt, aber  $BA$  nicht, wenn  $l \neq n$  ist.
- Sogar wenn  $AB$  und  $BA$  beide erklärt sind, haben sie im Allgemeinen verschiedene Größe: Ist  $A \in \mathbb{K}^{l \times m}$  und  $B \in \mathbb{K}^{m \times l}$ , so ist  $AB \in \mathbb{K}^{l \times l}$ , aber  $BA \in \mathbb{K}^{m \times m}$ .
- Sogar, wenn  $AB$  und  $BA$  beide erklärt sind, und gleich groß sind, gilt im Allgemeinen

$$AB \neq BA.$$

- Weiterhin gilt, dass auch im Fall  $AB = 0$  (die Nullmatrix) immer noch  $BA \neq 0$  gelten kann.

*Beispiel.* Für

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

gilt

$$AB = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix},$$

aber

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

Übrigens können wir einen Vektor  $\vec{x} \in \mathbb{K}^n$  auch als  $n \times 1$ -Matrix interpretieren. Dann ist das Matrix-Vektor-Produkt ein Spezialfall des Matrix-Matrix-Produktes.

Für eine  $m \times n$  Matrix die nur Nullen enthält, schreiben wir  $0_{m \times n}$  oder auch nur 0, wenn die Größe von Kontext klar ist.

## 10 Orthogonale und unitäre Matrizen, Kern und Lösungsraum

Eine wichtige Operation für Matrizen ist das *Transponieren*: Für  $A \in \mathbb{K}^{m \times n}$  ist die *transponierte Matrix*  $A^T \in \mathbb{K}^{n \times m}$  und es gilt

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}, \text{ d.h. } (A^T)_{ij} = a_{ji}.$$

Durch Transposition wird aus einem Zeilenvektor  $\vec{x} \in \mathbb{K}^n$  ein Spaltenvektor  $\vec{x}^T \in \mathbb{K}^{1 \times n}$ .

Das Transponieren spiegelt eine Matrix an der Diagonalen.

Es gelten folgende Rechenregeln:

$$\begin{aligned} (i) \quad & (A + B)^T = A^T + B^T \\ (ii) \quad & (\lambda A)^T = \lambda A^T \\ (iii) \quad & (A^T)^T = A \\ (iv) \quad & (AB)^T = B^T A^T. \end{aligned}$$

Mit der Transposition können wir das reelle Skalarprodukt von  $\vec{x}, \vec{y} \in \mathbb{R}^n$  schreiben als

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{y} = (x_1 \ \dots \ x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i.$$

Man sieht am Beispiel von Vektoren noch einmal sehr deutlich, dass das Matrix-Matrix-Produkt nicht kommutativ ist:  $\vec{x}^T \vec{y}$  ist in  $\mathbb{K}$  (also ein Skalar), aber  $\vec{x} \vec{y}^T$  ist die  $n \times n$  Matrix

$$\vec{x} \vec{y}^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (y_1 \ \dots \ y_n) = \begin{pmatrix} x_1 y_1 & \dots & x_1 y_n \\ \vdots & & \vdots \\ x_n y_1 & \dots & x_n y_n \end{pmatrix}.$$

Man nennt das Produkt  $\vec{x} \vec{y}^T$  auch *dyadisches Produkt* und die Matrix  $\vec{x} \vec{y}^T$  *Dyade*.

**Definition 10.1.** Eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt

- *symmetrisch*, wenn  $A = A^T$  gilt,
- *orthonormal*, wenn  $AA^T = A^T A = I_n$  gilt (d.h. die Zeilen von  $A$  sind paarweise orthonormal zueinander, ebenso die Spalten),
- *positiv definit*, wenn  $A$  symmetrisch ist und für alle  $\vec{x} \in \mathbb{R}^n$  mit  $\vec{x} \neq \vec{0}$  gilt  $\vec{x}^T A \vec{x} > 0$ .

Hat  $A$  die Spalten  $\vec{a}^{(i)}$ , so ist der  $ij$ -te Eintrag von  $A^T A$  gerade  $\langle \vec{a}^{(i)}, \vec{a}^{(j)} \rangle$ . Orthonormalität zweier Vektoren bedeutet, dass beide Vektoren normiert sind (also Länge 1 haben) und orthogonal zueinander sind.

Die Identitätsmatrix ist positiv definit, denn es gilt  $\vec{x}^T I_n \vec{x} = \|\vec{x}\|_2^2 > 0$  für  $\vec{x} \neq \vec{0}$ .

Für komplexe  $m \times n$ -Matrizen  $A$  definieren wir noch die *konjugierte* Matrix  $\bar{A}$  als

$$\bar{A} := (\bar{a}_{ij}), \text{ d.h. } (A^*)_{ij} = \bar{a}_{ji}$$

und die *adjungierte* Matrix (auch *transponiert-konjugierte* Matrix)

$$A^* := \bar{A}^T.$$

Es gelten folgende Rechenregeln:

- (i)  $(A + B)^* = A^* + B^*$
- (ii)  $(\lambda A)^* = \bar{\lambda} A^*$
- (iii)  $(A^*)^* = A$
- (iv)  $(AB)^* = B^* A^*$ .

**Definition 10.2.** Eine quadratische Matrix  $A \in \mathbb{C}^{n \times n}$  heißt

- *hermitesch*, wenn  $A = A^*$  gilt,
- *unitär*, wenn  $AA^* = A^*A = I_n$  gilt,
- *positiv definit*, wenn  $A$  hermitesch ist und für alle  $\vec{x} \in \mathbb{C}^n$  mit  $\vec{x} \neq \vec{0}$  gilt  $\vec{x}^* A \vec{x} > 0$ .

Das komplexe Skalarprodukt lässt sich jetzt ausdrücken als

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y} = \sum_{i=1}^n x_i^* y_i$$

Wir werden nun lineare Gleichungssysteme in  $n$  Variablen mit  $m$  Gleichungen systematisch lösen. Es ist also gegeben

$$A\vec{x} = \vec{b}$$

mit Koeffizientenmatrix  $A \in \mathbb{K}^{m \times n}$  und rechter Seite  $\vec{b} \in \mathbb{K}^m$ .

**Definition 10.3.** Die Lösungsmenge  $L(A, \vec{b})$  eines linearem Gleichungssystems  $A\vec{x} = \vec{b}$  mit  $A \in \mathbb{K}^{m \times n}$  und  $\vec{b} \in \mathbb{K}^m$  ist

$$L(A, \vec{b}) := \left\{ x \in \mathbb{K}^n \mid A\vec{x} = \vec{b} \right\}.$$

**Definition 10.4** (Homogene und inhomogene Gleichungssysteme). Das lineare Gleichungssystem  $A\vec{x} = \vec{b}$  heißt *homogen*, wenn  $\vec{b} = \vec{0}$ , sonst *inhomogen*.

Ein homogenes Gleichungssystem hat stets die *triviale Lösung*  $\vec{x} = \vec{0}$ .

**Satz 10.5** (und Definition des Kerns). Die Lösungsmenge des homogenen Gleichungssystems  $A\vec{x} = \vec{0}$  ist ein Untervektorraum des  $\mathbb{K}^n$ , genannt *Kern der Matrix  $A$*  und geschrieben  $\text{Kern}(A)$ .

*Beweis.*

Wegen  $A\vec{0} = \vec{0}$  ist  $\vec{0} \in \text{Kern}(A)$ .

Sind nun  $\vec{v}, \vec{w} \in \text{Kern}(A)$ , so gilt  $A\vec{v} = \vec{0}$  und  $A\vec{w} = \vec{0}$ . Also gilt

$$A(\vec{v} + \vec{w}) = A\vec{v} + A\vec{w} = \vec{0} + \vec{0} = \vec{0},$$

also  $\vec{v} + \vec{w} \in \text{Kern}(A)$ . Analog folgt  $\alpha\vec{x} \in \text{Kern}(A)$  aus  $A(\alpha\vec{x}) = \alpha A\vec{x}$ .  $\square$

Der obige Satz sagt also nichts anderes als

$$L(A, \vec{0}) = \text{Kern}(A).$$

**Satz 10.6.** a) Ist  $(\vec{x}^{(1)}, \dots, \vec{x}^{(k)})$  eine Basis von  $\text{Kern}(A)$ , so ist jede Lösung des homogenen Systems  $A\vec{x} = \vec{0}$  von der Form

$$\vec{x}_h = \sum_{i=1}^k \alpha_i \vec{x}^{(i)}, \quad \alpha_i \in \mathbb{K}.$$

b) Ist  $\vec{x}_s$  eine Lösung des inhomogenen Systems  $A\vec{x} = \vec{b}$ , so ist jede Lösung des inhomogenen Systems von der Form

$$\vec{x} = \vec{x}_s + \vec{x}_h = \vec{x}_s + \sum_{i=1}^k \alpha_i \vec{x}^{(i)}.$$

Man nennt  $\vec{x}_h$  auch *allgemeine Lösung* des homogenen Systems, jedes  $\vec{x}_s$  *spezielle* (oder *partikuläre*) *Lösung* des inhomogenen Systems und  $\vec{x} = \vec{x}_s + \vec{x}_h$  *allgemeine Lösung* des inhomogenen Systems.

*Beweis.*

a) Folgt direkt, da  $\text{Kern}(A)$  ein UVR ist.

b) Es gilt

$$A(\vec{x}_s + \vec{x}_h) = A\vec{x}_s + A\vec{x}_h = \vec{b} + \vec{0} = \vec{b}.$$

Also lösen alle Vektoren der Form  $\vec{y} = \vec{x}_s + \vec{x}_h$  das inhomogene System.

Umgekehrt: Ist  $\vec{y}$  eine Lösung des inhomogenen Systems, dann ist

$$A(\vec{y} - \vec{x}_s) = A\vec{y} - A\vec{x}_s = \vec{b} - \vec{b} = \vec{0}.$$

Also ist  $\vec{y} - \vec{x}_s$  eine Lösung des homogenen Systems, d.h. es gilt

$$\text{nach a) } \vec{y} - \vec{x}_s = \sum_{i=1}^k \alpha_i \vec{x}^{(i)}.$$

$\square$

Die Aussage dieses Satzes in Kurzform ist also

$$L(A, b) = \vec{x}_s + \text{Kern}(A)$$

wobei  $\vec{x}_s$  eine beliebige spezielle Lösung des inhomogenen Systems ist.

*Beispiel.* Hier ein einfaches Beispiel, bei dem man alles einfach ablesen kann. Wir betrachten das Gleichungssystem

$$\begin{pmatrix} 4 & -2 & 1 & 4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}.$$

Suchen wir zuerst eine spezielle Lösung: Die dritte Zeile ist immer erfüllt (sie lautet  $0 = 0$ ) und ist daher irrelevant. In der zweiten Zeile  $x_2 - x_3 + 3x_4 = 2$  können wir  $x_3$  und  $x_4$  beliebig wählen und wählen  $x_3 = x_4 = 0$  und bekommen  $x_2 = 2$ . Gehen wir damit in die erste Zeile, wird diese zu  $4x_1 + (-2)2 + 0 + 0 = 0$  und wir bekommen  $x_1 = 1$ . Eine spezielle Lösung ist also

$$\vec{x}_s = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}.$$

Um homogene Lösungen zu bestimmen, betrachten wir das Gleichungssystem mit  $\vec{b} = \vec{0}$

$$\begin{pmatrix} 4 & -2 & 1 & 4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wieder liefert die dritte Zeile keine Information, und wir wählen in der zweiten Zeile zuerst  $x_3 = 1$  und  $x_4 = 0$ . Damit ist die zweite Zeile  $x_2 - 1 = 0$ , also  $x_2 = 1$ . Die erste Zeile gibt  $4x_1 + (-2)1 + 1 = 0$ , also  $x_1 = \frac{1}{4}$ . Dies gibt eine homogene Lösung

$$\vec{x}_h^{(1)} = \begin{pmatrix} \frac{1}{4} \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Jetzt wählen wir in der zweiten Zeile  $x_3 = 0$  und  $x_4 = 1$  und bekommen  $x_2 - 0 + 3 = 0$ , also  $x_2 = -3$ . Das gibt in der ersten Zeile  $4x_1 + (-2)(-3) + 0 + 4 = 0$ , also  $x_1 = -\frac{10}{4}$ . Das gibt eine zweite, von der ersten linear unabhängige, homogene Lösung

$$\vec{x}_h^{(2)} = \begin{pmatrix} -\frac{10}{4} \\ -3 \\ 0 \\ 1 \end{pmatrix}.$$

Eine weitere linear unabhängige homogene Lösung gibt es nicht und der ganze Lösungsraum ist also

$$L(A, b) = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + \alpha_1 \begin{pmatrix} \frac{1}{4} \\ 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -\frac{10}{4} \\ -3 \\ 0 \\ 1 \end{pmatrix} \mid \alpha_1, \alpha_2 \in \mathbb{K} \right\}.$$

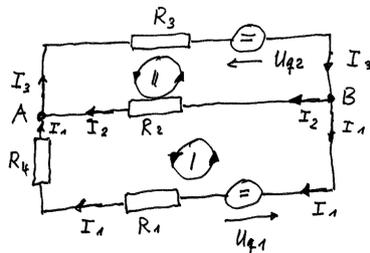
## 11 Das Eliminationsverfahren

Um bei Lösung von linearen Gleichungssysteme per Hand die Übersicht zu behalten, schreiben wir das Gleichungssystem  $A\vec{x} = \vec{b}$  in erweiterter Matrixform:

$$\left( \begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right)$$

Wir wenden nun sogenannte *elementare Zeilenumformungen* an, welche die Lösungsmenge des Gleichungssystems unverändert lassen, mit dem Ziel, unterhalb der Diagonale (also unterhalb der Einträge  $a_{ii}$ ) Nullen zu erzeugen. Wir gehen dabei systematisch vor und eliminieren zuerst die Nullen unterhalb von  $a_{11}$ , dann unterhalb von  $a_{22}$ , usw.

*Beispiel.* Wir betrachten noch einen Schaltkreis



mit den Daten  $R_1 = 1 \Omega$ ,  $R_2 = 2 \Omega$ ,  $R_3 = 3 \Omega$ ,  $R_4 = 5 \Omega$ ,  $U_{q1} = 10 \text{ V}$ ,  $U_{q2} = 20 \text{ V}$ . Gesucht sind die Ströme  $I_1$ ,  $I_2$  und  $I_3$ .

Die Knotenregel in Knoten A liefert die Gleichung

$$I_1 + I_2 - I_3 = 0.$$

Die Maschenregeln in den Maschen I und II gibt

$$\begin{aligned} -R_1 I_1 + R_2 I_2 - R_4 I_1 + U_{q1} &= 0 \\ -R_2 I_2 - R_3 I_3 + U_{q2} &= 0. \end{aligned}$$

Diese Gleichungen schreiben wir um als

$$\begin{aligned} -(R_1 + R_4) I_1 + R_2 I_2 &= -U_{q1} \\ +R_2 I_2 + R_3 I_3 &= U_{q2}. \end{aligned}$$

Wir haben also ein Gleichungssystem mit drei Unbekannten und drei Gleichungen:

$$\begin{aligned} I_1 + I_2 - I_3 &= 0 \\ -(R_1 + R_4) I_1 + R_2 I_2 &= -U_{q1} \\ +R_2 I_2 + R_3 I_3 &= U_{q2}. \end{aligned}$$

Die erweiterte Matrix des Gleichungssystems ist also

$$\left( \begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ -(R_1 + R_4) & R_2 & 0 & -U_{q1} \\ 0 & R_2 & R_3 & U_{q2} \end{array} \right)$$

Einsetzen der Zahlenwerte gibt

$$\left( \begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ -6 & 2 & 0 & -10 \\ 0 & 2 & 3 & 20 \end{array} \right)$$

Wir addieren nun das 6-fache der ersten Zeile zur zweiten und bekommen:

$$\left( \begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 8 & -6 & -10 \\ 0 & 2 & 3 & 20 \end{array} \right)$$

Dies ist eine der elementaren Zeilenumformungen!

und sind mit der Elimination unterhalb der Eintrags  $a_{11}$  schon fertig. Nun multiplizieren wir die letzte Zeile mit 4 und ziehen dann die zweite Zeile von der dritten ab und bekommen

$$\left( \begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 8 & -6 & -10 \\ 0 & 0 & 18 & 90 \end{array} \right).$$

Das multiplizieren einer Zeile mit einer Zahl ungleich Null ist eine weitere elementare Zeilenumformung.

Damit sind wir fertig mit der Elimination und können von unten nach oben lösen:

- 3-te Zeile:  $18I_3 = 90$ , also  $I_3 = 5$ .
- 2-te Zeile:  $8I_2 - 6I_3 = -10$ , also  $8I_2 - 30 = -10$ , d.h.  $8I_2 = 20$ , also  $I_2 = 2.5$ .
- 1-te Zeile:  $I_1 + I_2 - I_3 = 0$ , also  $I_1 + 2.5 - 5 = 0$ , also  $I_1 = 2.5$ .

Die elementaren Zeilenumformungen, die wir benutzt haben sind

- Multiplikation einer beliebigen Zeile mit einer Zahl  $\neq 0$ ,
- Addition eines Vielfachen einer Zeile zu einer anderen.

Diese Operationen ändern nichts an der Menge der Lösungen!  
Eine weitere elementare Zeilenumformung ist:

- Tauschen zweier Zeilen.

Wir halten nun das systematische Vorgehen des Eliminationsverfahrens fest:

I. Ist  $a_{11} \neq 0$ ?

- (a) Falls  $a_{11} = 0$ , suche in ersten Spalte von  $A$  ein Element  $a_{k1} \neq 0$  und vertausche die  $k$ -te und die erste Zeile.

Das Verfahren wird oft Gaußsche Eliminationsverfahren genannt - es ist aber schon wesentlich älter und wurde schon um 200 n.Chr. in China eingesetzt.

- (b) Sind alle Elemente der 1-ten Spalte = 0, dann suche in der Restmatrix ein  $a_{ij} \neq 0$  und vertausche die  $i$ -te mit der ersten Spalte.

*Achtung: Vertauschen von **Spalten** ist eine Vertauschung von Variablen. Hierüber muss also Buch geführt werden!*

- (c) sind alle Elemente der Matrix = 0, dann ist der Prozess beendet.

II. Ist  $a_{11} \neq 0$ , dann heißt  $a_{11}$  das *Pivot-Element* für den ersten Eliminationsschritt: Wir rechnen für  $i = 1, \dots, m$

- Addiere das  $(-\frac{a_{i1}}{a_{11}})$ -fache der ersten Zeile zur  $i$ -ten Zeile.

Damit sind unterhalb von  $a_{11}$  Nullen erzeugt.

III. Wende das Verfahren ab Schritt I. auf die Restmatrix

$$\left( \begin{array}{ccc|c} a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & & \vdots & \vdots \\ a_{m2}^{(2)} & \cdots & a_{mn}^{(2)} & b_m^{(2)} \end{array} \right)$$

Wir kommen dadurch auf eine Matrix in *Zeilen-Stufen-Form*, nämlich in der Form

$$\left( \begin{array}{cccc|cc} a_{11} & a_{12} & \cdots & \cdots & a_{1n} & b_1 \\ & a_{22}^{(2)} & \cdots & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & & \vdots & \vdots \\ & & & a_{rr}^{(r)} & \cdots & a_{rn}^{(r)} & b_r^{(r)} \\ & & & 0 & \cdots & 0 & b_{r+1}^{(r)} \\ & & & \vdots & & \vdots & \vdots \\ & & & 0 & \cdots & 0 & b_m^{(r)} \end{array} \right)$$

IV. Wir folgern:

- Gilt  $b_k^{(r)} \neq 0$  für ein  $k \geq r + 1$ , so hat das Gleichungssystem keine Lösung.
- Gilt  $b_k^{(r)} = 0$  für  $k = r + 1, \dots, m$ , so kann man die *freien Variablen*  $x_{r+1}, \dots, x_n$  beliebig wählen und dann den Rest der  $x_k$  durch *Rückwärtssubstitution* lösen:

Für  $i = r, r - 1, \dots, 1$

$$- x_i = \left( b_i^{(r)} - \sum_{j=i+1}^n a_{ij}^{(r)} x_j \right) / a_{ii}.$$

- Wählt man  $x_{r+1} = \dots = x_n = 0$ , so bekommt man eine spezielle Lösung  $\vec{x}_s$ .
- Setzt man alle  $b_i^{(r)} = 0$  und wählt für  $j = 1, \dots, n - r$  die freien Variablen

$$x_{r+i}^{(j)} = \begin{cases} 0 & : i = 1, \dots, n - r, \quad i \neq j \\ 1 & : i = j \end{cases}$$

und rechnet per Rücksubstitution die  $x_1^{(j)}, \dots, x_r^{(j)}$  aus, so sind die Vektoren  $\vec{x}^{(j)}, j = 1, \dots, n - r$  linear unabhängig und bilden eine Basis des Kerns von  $A$ .

- Die allgemeine Lösung des inhomogenen Systems ist dann

$$\vec{x} = \vec{x}_s + \alpha_1 \vec{x}^{(1)} + \dots + \alpha_{n-r} \vec{x}^{(n-r)}, \quad \alpha_j \in \mathbb{K}.$$

- Ist ein  $b_k^{(r)}$  für  $k = r + 1, \dots, m$  ungleich Null, dann hat das Gleichungssystem keine Lösung.

*Beispiel.* Das Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 - x_2 - x_3 &= 4 \\ x_1 + 3x_2 + 3x_3 &= 1 \end{aligned}$$

hat die erweiterte Matrix

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

Addition des  $(-1)$ -fachen der ersten Zeile zur 2. und 3. Zeile gibt

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

Addieren der 2. Zeile zur 3. Zeile gibt dann

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

Wir sehen, dass dieses Gleichungssystem keine Lösung hat. Ändern wir die rechte Seite ab zu

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 - x_2 - x_3 &= 4 \\ x_1 + 3x_2 + 3x_3 &= 2 \end{aligned}$$

so kommen wir nach dem ersten Eliminationsschritt zu

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

und nach dem zweiten zu

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

Für eine spezielle Lösung setzen wir  $x_3 = 0$  und bekommen per Rückwärtssubstitution  $x_2 = -\frac{1}{2}$  und  $x_1 = \frac{7}{2}$ . Für die allgemeine Lösung des homogenen Systems betrachten wir

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

setzen  $x_3 = 1$  und bekommen durch Rückwärtssubstitution  $x_2 = -1$  und  $x_1 = 0$ . Damit ist der Lösungsraum

$$L(A, b) = \left\{ \left( \begin{array}{c} \frac{7}{2} \\ -\frac{1}{2} \\ 0 \end{array} \right) + \alpha \left( \begin{array}{c} 0 \\ -1 \\ 1 \end{array} \right) \mid \alpha \in \mathbb{K} \right\}$$

## 12 Der Rang einer Matrix

Der Eliminationsalgorithmus ist nicht nur eine praktische Rechenmethode (die, mit Anpassungen und Tricks, auch in Computern implementiert ist), sondern auch theoretisch hilfreich.

**Definition 12.1.** Die maximale Anzahl linear unabhängiger Zeilen einer Matrix  $A$  heißt (Zeilen-)Rang von  $A$ , geschrieben  $\text{Rang}(A)$ .

Da sich der Zeilen-Rang während des Eliminationsalgorithmus nicht ändert, bekommen wir sofort: Endet der Eliminationsalgorithmus in einer Matrix mit  $r$  nicht-null Zeilen, so gilt  $\text{Rang}(A) = r$ .

Außerdem sehen wir:

**Satz 12.2.** Das lineare Gleichungssystem  $A\vec{x} = \vec{b}$  hat genau dann eine Lösung, wenn gilt

$$\text{Rang}(A) = \text{Rang}(A|b).$$

**Satz 12.3.** (a) Für jede  $m \times n$ -Matrix  $A$  gilt die Dimensionsformel

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

(b) Ein unterbestimmtes, homogenes, lineares Gleichungssystem  $A\vec{x} = \vec{0}$ ,  $A \in \mathbb{K}^{m \times n}$  mit  $m < n$  hat stets einen nicht-trivialen Lösungsraum der Dimension  $n - \text{Rang}(A)$ .

(c) Ein lineares Gleichungssystem  $A\vec{x} = \vec{b}$  mit quadratischem  $A \in \mathbb{K}^{n \times n}$  ist genau dann eindeutig lösbar, wenn  $\text{Rang}(A) = n$  gilt.

Speziell für quadratische Matrizen bekommen wir außerdem:

**Satz 12.4.** Für  $A \in \mathbb{K}^{n \times n}$  sind folgende Aussagen äquivalent:

1. Die Zeilen von  $A$  sind linear unabhängig.
2.  $\text{Rang}(A) = n$
3.  $A\vec{x} = \vec{b}$  hat für jedes  $\vec{b} \in \mathbb{K}^n$  eine Lösung
4.  $A\vec{x} = \vec{b}$  hat für jedes  $\vec{b} \in \mathbb{K}^n$  genau eine Lösung
5.  $A\vec{x} = \vec{0}$  hat nur die triviale Lösung  $\vec{x} = \vec{0}$

Man kann sich lineare Gleichungssystem und speziell den Eliminationsalgorithmus aus geometrisch veranschaulichen:

(A) Eindeutig lösbarer Fall: Wir betrachten

$$\begin{pmatrix} 3 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 11 \\ 1 \end{pmatrix}.$$

Der Eliminationsalgorithmus geht hier wie folgt:

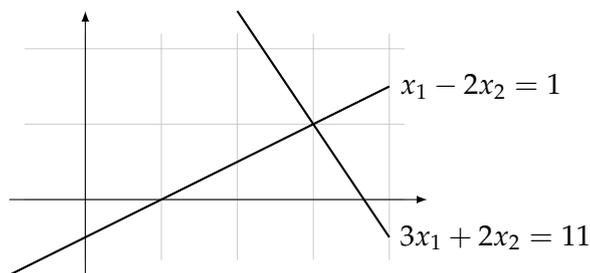
$$(A|\vec{b}) = \left( \begin{array}{cc|c} 3 & 2 & 11 \\ 1 & -2 & 1 \end{array} \right)$$

Zeilentausch  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 3 & 2 & 11 \end{array} \right)$

3faches der 1. von 2. abziehen  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 8 & 8 \end{array} \right)$

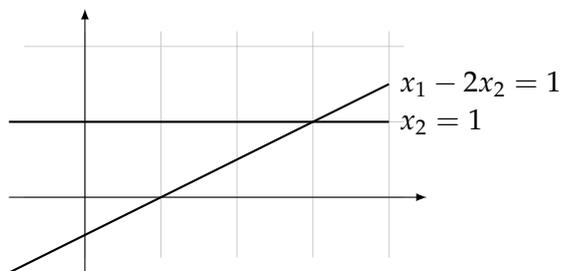
Rückwärtssubstitution: Aus  $8x_2 = 8$  folgt  $x_2 = 1$  und in der ersten Zeile kriegen wir  $x_1 - 2 = 1$ , also  $x_1 = 3$ .

Jede Zeile des Ausgangsgleichungssystems entspricht einer Geradengleichung:



Wir sehen, dass  $(3,1)$  als Schnittpunkt beider Geraden die einzige Lösung des Systems ist.

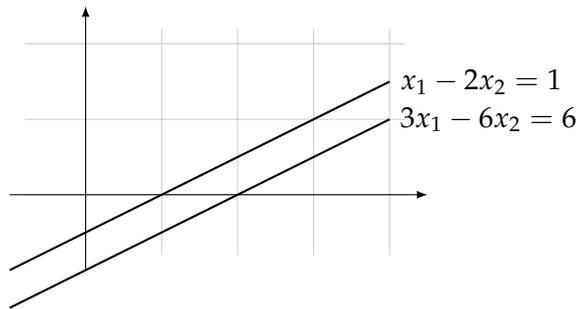
Die Elimination macht aus der Geradengleichung  $3x_1 + 2x_2 = 11$  die Gleichung  $8x_2 = 8$ , also  $x_2 = 1$ :



(B) Nicht lösbarer Fall: Ein Gleichungssystem ohne Lösung ist

$$\begin{pmatrix} 1 & -2 \\ 3 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \end{pmatrix}.$$

Das sieht man z.B. daran, dass die beiden zugehörigen Geraden parallel sind:



Das Eliminationsverfahren gibt:

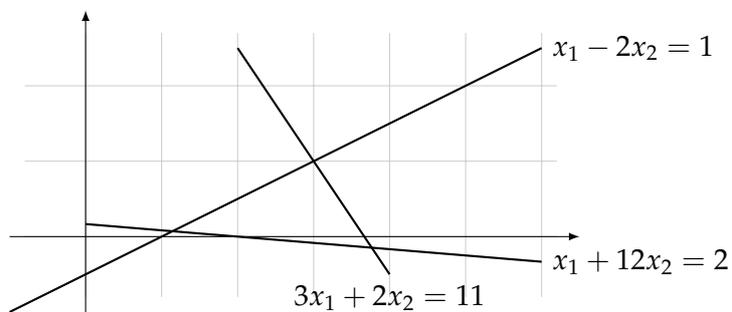
$$(A|\vec{b}) = \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 3 & -6 & 6 \end{array} \right)$$

$$3\text{faches der 1. von 2. abziehen} \rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 0 & 3 \end{array} \right)$$

Wir bekommen ebenso einen unlösbaren Fall, mit folgendem  $3 \times 2$  System:

$$\begin{pmatrix} 3 & 2 \\ 1 & -2 \\ 1 & 12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 11 \\ 1 \\ 2 \end{pmatrix}.$$

In diesem Fall schneiden sich die drei Geraden nicht in einem Punkt:



Das Eliminationsverfahren gibt:

$$(A|\vec{b}) = \left( \begin{array}{cc|c} 3 & 2 & 11 \\ 1 & -2 & -2 \\ 1 & 12 & 2 \end{array} \right)$$

1. und 2. Zeilen tauschen  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & -2 \\ 3 & 2 & 11 \\ 1 & 12 & 2 \end{array} \right)$

3faches der 1. von 2. abziehen  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 8 & 8 \\ 1 & 12 & 2 \end{array} \right)$

1. von 3. abziehen  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 8 & 8 \\ 0 & 14 & 1 \end{array} \right)$

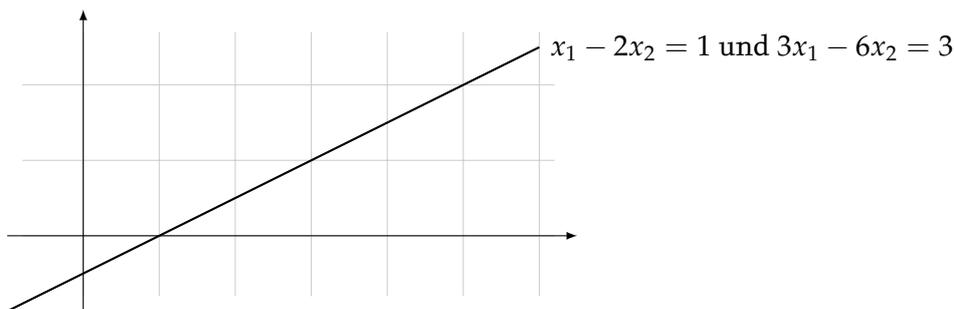
2. durch 8 teilen  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 14 & 1 \end{array} \right)$

14faches von 2. von 3. abziehen  $\rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -13 \end{array} \right)$

(C) Fall mit unendlich vielen Lösungen:

$$\left( \begin{array}{cc|c} 1 & -2 & 1 \\ 3 & -6 & 3 \end{array} \right) \rightarrow \left( \begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 0 & 0 \end{array} \right)$$

Hier ist  $x_2$  frei wählbar. In diesem Fall sind die beiden Geraden identisch:



Auf dem Computer muss man aufpassen, was bei Rundungsfehlern passiert. Um dies zu illustrieren betrachten wir

$$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Wir lösen dieses System zuerst exakt:

$$\left( \begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 1 & 1 & 2 \end{array} \right) \rightarrow \left( \begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 0 & 1 - 10^4 & 2 - 10^4 \end{array} \right).$$

Das führt auf

$$x_2 = \frac{9998}{9999} = 1 - \frac{1}{9999} = 0.999\,899\,98\dots$$

$$x_1 = 1 + \frac{1}{9999} = 1.000\,100\,01\dots$$

Allerdings ist das Rückwärtseinsetzen instabil: Rechnen wir mit 3-stelliger Arithmetik, so ist

$$1 - 10^4 = -9999 \stackrel{\circ}{=} -10^4$$

$$2 - 10^4 = -9998 \stackrel{\circ}{=} -10^4$$

$$10^{-4} = 0.0001 \stackrel{\circ}{=} 0.$$

Mit  $\stackrel{\circ}{=}$  ist hier „gleich bis auf Rundung“ gemeint.

Das Gleichungssystem nach Elimination ist also

$$\left( \begin{array}{cc|c} 0 & 1 & 1 \\ 0 & -10^4 & -10^4 \end{array} \right)$$

und hat unendlich viele Lösungen, nämlich  $x_2 = 1$  und  $x_1$  frei wählbar.

Auch wenn wir den Wert  $10^{-4}$  oben links stehen lassen, kommen wir auf

$$\left( \begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 0 & -10^4 & -10^4 \end{array} \right)$$

und also  $x_2 = 1$ , aber  $10^{-4}x_1 + 1 = 1$ , also  $x_1 = 0$ .

Die Situation wird besser, wenn *Pivot-Suche* betrieben wird, d.h. wenn Zeilen (und/oder Spalten) vertauscht werden, um als Element  $a_{11}$  eine möglichst große Zahl zu bekommen. Tauschen wir hier die Zeilen und eliminieren dann, bekommen wir

$$\left( \begin{array}{cc|c} 1 & 1 & 2 \\ 10^{-4} & 1 & 1 \end{array} \right) \rightarrow \left( \begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 - 10^{-4} & 1 - 2 \cdot 10^{-4} \end{array} \right)$$

was in 3-stelliger Arithmetik dem System

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

entspricht. Wir bekommen als Lösung also  $x_2 = 1$  und  $x_1 = 1$  was nah an der wahren Lösung ist.

Im Allgemeinen erhöht sich die Stabilität der Elimination, wenn man wie folgt *Pivot-Suche* betreibt:

- Wähle das betragsgrößte Element der Matrix und tausche Zeilen und Spalten, so dass dies zum Element  $a_{11}$  wird. Eliminiere wie gehabt und fahre entsprechend mit der Restmatrix fort.

### 13 Inverse Matrizen

**Definition 13.1.** Sei  $A \in \mathbb{K}^{n \times n}$  eine quadratische Matrix. Eine Matrix  $X \in \mathbb{K}^{n \times n}$  heißt *Inverse* zu  $A$ , wenn

$$AX = XA = I_n$$

gilt. Hat  $A$  eine Inverse, so heißt  $A$  *invertierbar*. Invertierbare Matrizen nennt man auch *regulär* und nicht-invertierbare *singulär*.

**Satz 13.2.** Ist eine  $n \times n$  Matrix  $A$  invertierbar, so gilt:

- (a) Das Gleichungssystem  $A\vec{x} = \vec{b}$  hat für jedes  $\vec{b}$  genau eine Lösung.  
 (b)  $\text{Rang}(A) = n$ .

*Beweis.*

- (a) Hat  $A$  eine Inverse  $X$ , so setzen wir  $\vec{x} = X\vec{b}$  und sehen  $A\vec{x} = AX\vec{b} = I_n\vec{b} = \vec{b}$ , was zeigt, dass dieses  $\vec{x}$  eine Lösung ist. Satz 12.4 zeigt auch schon die Eindeutigkeit.  
 (b) Nach Satz 12.4 wissen wir, dass aus Existenz von Lösungen auch  $\text{Rang}(A) = n$  folgt.

□

Es gilt aber auch:

**Satz 13.3.** Eine quadratische Matrix  $A \in \mathbb{K}^{n \times n}$  hat genau dann Zeilenrang  $n$ , wenn die Spalten von  $A$  linear unabhängig sind.

*Beweis.*

Wir stellen die Matrix spaltenweise dar, d.h wir schreiben

$$A = (\vec{a}^{(1)} \quad \dots \quad \vec{a}^{(n)}) \quad \text{mit} \quad \vec{a}^{(1)}, \dots, \vec{a}^{(n)} \in \mathbb{K}^n.$$

Dann gilt  $A\vec{x} = \vec{0}$  genau dann, wenn  $\sum_{i=1}^n x_i \vec{a}^{(i)} = \vec{0}$ . Aus Satz 12.4 schließen wir die Äquivalenzen

$$\begin{aligned} \text{Rang}(A) = n &\iff A\vec{x} = \vec{0} \text{ hat nur die triviale Lösung } \vec{x} = \vec{0} \\ &\iff \forall \vec{x} \in \mathbb{K}^n : \left( \sum_{i=1}^n x_i \vec{a}^{(i)} = \vec{0} \Rightarrow x_1 = \dots = x_n = 0 \right) \\ &\iff (\vec{a}^{(1)}, \dots, \vec{a}^{(n)}) \text{ linear unabhängig.} \end{aligned}$$

□

Mit etwas anderer Technik können wir eine ähnliche Aussage auch für rechteckige Matrizen treffen:

**Satz 13.4.** Für jede Matrix gilt

$$\text{Zeilenrang} = \text{Spaltenrang}$$

Hier ist Spaltenrang natürlich die maximale Anzahl linear unabhängiger Spalten.

*Beweis.*

Die Dimensionsformel aus Satz 12.3 sagt

$$\dim \left\{ \vec{x} \in \mathbb{K}^n \mid A\vec{x} = \vec{0} \right\} = \dim(\text{Kern}(A)) = n - \text{Rang}(A).$$

Andererseits ist

$$\text{Kern}(A) = \left\{ \vec{x} \in \mathbb{K}^n \mid \sum_{i=1}^n x_i \vec{a}^{(i)} = \vec{0} \right\}.$$

Sei nun  $l$  die maximale Anzahl von linear unabhängigen Spalten von  $A$ . Wie nummerieren diese so, dass  $(\vec{a}^{(1)}, \dots, \vec{a}^{(l)})$  linear unabhängig ist. Das bedeutet aber, dass alle weiteren Spalten  $\vec{a}^{(k)}$ ,  $k = l+1, \dots, n$  in  $\text{Spann}(\vec{a}^{(1)}, \dots, \vec{a}^{(l)})$  liegen. Wir können also für  $k = l+1, \dots, n$  schreiben

$$\vec{a}^{(k)} = \sum_{i=1}^l \alpha_{ki} \vec{a}^{(i)}.$$

Dann folgt

$$\begin{aligned} \vec{x} \in \text{Kern}(A) &\iff \sum_{i=1}^l x_i \vec{a}^{(i)} = \vec{0} \\ &\iff \sum_{i=1}^l \left( x_i + \sum_{k=l+1}^n \alpha_{ki} x_k \right) \vec{a}^{(i)} = \vec{0} \\ &\iff \text{für alle } i = 1, \dots, l \text{ gilt } x_i = - \sum_{k=l+1}^n \alpha_{ki} x_k \end{aligned}$$

Das bedeutet, dass der Kern von  $A$  die Dimension  $n - l$  hat, und das heißt  $l = \text{Rang}(A)$ .  $\square$

Da Zeilenrang gleich Spaltenrang ist, spricht man auch nur vom *Rang* einer Matrix. Wir sagen, dass eine  $m \times n$  Matrix *Vollrang* hat, wenn

$$\text{Rang}(A) = \min(m, n)$$

gilt (denn größer kann der Rang nicht sein).

**Satz 13.5.** *Eine quadratische Matrix ist genau dann invertierbar, wenn sie vollen Zeilenrang hat.*

*Beweis.*

Es bleibt nur noch die „Rückrichtung“ zu zeigen, d.h., dass eine Matrix von vollem Zeilenrang invertierbar ist. Es sei  $X = (\vec{x}^{(1)}, \dots, \vec{x}^{(n)})$ . Dann ist die Matrix-Gleichung  $AX = I_n$  äquivalent zu den  $n$  Gleichungssystemen  $A\vec{x}^{(i)} = \vec{e}_i$ . Wenn  $A$  vollen Zeilenrang hat, so haben diese Gleichungssysteme alle jeweils eine eindeutige Lösung. D.h. es gibt genau eine Matrix  $X$ , welche die Gleichung  $AX = I_n$  erfüllt.

Wir zeigen, dass  $X$  selbst Vollrang hat: Dazu nehmen wir an, dass  $\sum_{i=1}^n \alpha_i \vec{x}^{(i)} = \vec{0}$  gilt. Dann folgt

$$\vec{0} = \sum_{i=1}^n \alpha_i A \vec{x}^{(i)} = \sum_{i=1}^n \alpha_i \vec{e}_i$$

und daraus folgt  $\alpha_1 = \dots = \alpha_n = 0$ , da die Vektoren  $\vec{e}_1, \dots, \vec{e}_n$  linear unabhängig sind.

Wir können also genau wie am Anfang des Beweises schließen, dass die Gleichung  $XY = I_n$  eine eindeutige Lösung  $Y$  hat. Damit folgt aber

$$XA = XAI_n = (XA)(XY) = X(AX)Y = XI_nY = XY = I_n$$

und wir haben gezeigt, dass  $XA = AX = I_n$  und also ist  $X$  eine Inverse von  $A$ .  $\square$

Wir machen eine Reihe von Beobachtungen:

- Der obige Beweis zeigt auch, dass die zu  $A$  inverse Matrix eindeutig ist. Sie wird mit  $A^{-1}$  bezeichnet.
- Gilt  $AB = I_n$  für zwei quadratische Matrizen  $A$  und  $B$ , so sind  $A$  und  $B$  beide invertierbar und es gilt  $A = B^{-1}$  und  $B = A^{-1}$ . Insbesondere gilt  $(A^{-1})^{-1} = A$ .
- Der vorige Punkt zeigt, dass wir für die Definition der Inversen einer quadratischen Matrix  $A$  nur  $A^{-1}A = I_n$  hätten verlangen müssen, da daraus auch  $AA^{-1} = I_n$  folgt.
- Für zwei invertierbare Matrizen  $A, B$  gilt

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

$$\text{Es gilt } (AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

- Ist  $A$  invertierbar, so ist auch die transponierte  $A^T$  invertierbar, und es gilt

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

$$\text{Es gilt nach den Rechenregeln für Transposition } A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n.$$

- Man kann die inverse Matrix im Prinzip benutzen, um lineare Gleichungssysteme zu lösen, denn aus  $A\vec{x} = \vec{b}$  folgt  $A^{-1}\vec{b} = A^{-1}A\vec{x} = I_n\vec{x} = \vec{x}$ . In der Praxis macht man das allerdings *niemals*, weder per Hand, noch mit dem Computer. Der Aufwand zur Berechnung von inversen Matrizen ist viel höher, als die Anwendung des Eliminationsverfahrens und

Man spricht auch von „Links-Inversen“ und „Rechts-Inversen“, und sagt, dass für quadratische Matrizen jede Links-Inverse auch eine Rechts-Inverse ist. Für rechteckige Matrizen kann man auch von Links- und Rechts-Inversen sprechen. Das Beispiel  $A = \vec{e}_1 \in \mathbb{R}^{n \times 1}$  zeigt allerdings, dass dann Links-Inverse im Allgemeinen keine Rechts-Inversen sind: Mit  $X = \vec{e}_1^T \in \mathbb{R}^{1 \times n}$  gilt  $XA = 1 = I_1$ , aber  $AX \neq I_n$ .

Das nennt man auch die Schuhe-Socken-Identität. Ist  $B$  „Socken anziehen“ und  $A$  „Schuhe anziehen“, so ist das Ergebnis von  $AB$ , dass man Schuhe und Socken an hat. Um dies rückgängig zu machen, muss *erst* die Schuhe ausziehen, daher kommt  $A^{-1}$  zuerst.

auch die Genauigkeit beim Rechnen mit Rundungsfehlern leidet. Auch wenn man die Gleichung  $A\vec{x} = \vec{b}$  häufig für verschiedene rechte Seiten  $\vec{b}$ , aber mit gleichem  $A$  lösen muss, lohnt es sich nicht, die inverse Matrix zu benutzen, denn wir sehen im nächsten Abschnitt ein besseres Verfahren.

Man kann inverse Matrizen mit Hilfe eines Eliminationsverfahrens berechnen. Man startet dabei mit einer erweiterten Matrix der Form  $(A \mid I_n)$  und wendet so lange elementare Zeilenumformungen an, bis man auf  $(I_n \mid B)$  angekommen ist und liest dann  $A^{-1} = B$  ab. Dieses Verfahren nennt man das Gauß-Jordan-Verfahren.

*Beispiel.* Wir betrachten ein einfaches Zahlenbeispiel:

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

Wir führen das beschriebene Verfahren durch:

$$\begin{aligned} & \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \\ \text{2. Zeile von 3. abziehen} & \rightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array} \right) \\ \text{3. Zeile von 1. Zeile abziehen} & \rightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array} \right). \end{aligned}$$

Damit haben wir die inverse Matrix gefunden:

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

Nun noch ein einfaches Beispiel mit Parametern: Wir berechnen die Inverse einer allgemeinen  $2 \times 2$  Matrix (und bekommen nebenbei heraus, wann eine solche Matrix invertierbar ist).

*Beispiel.* Wir betrachten

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Wir rechnen

$$\begin{aligned}
 & \left( \begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right) \text{ Ziehe } \frac{c}{a}\text{-faches der 1. von 2. ab} \\
 \rightarrow & \left( \begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{array} \right) \text{ Fasse zusammen} \\
 \rightarrow & \left( \begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \frac{ad-bc}{a} & -\frac{c}{a} & 1 \end{array} \right) \text{ Multipliziere 2. mit } \frac{a}{ad-bc} \\
 \rightarrow & \left( \begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right) \text{ Ziehe } b\text{-faches der 2. von 1. ab} \\
 \rightarrow & \left( \begin{array}{cc|cc} a & 0 & 1 + \frac{bc}{ad-bc} & -\frac{ab}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right) \text{ Fasse zusammen} \\
 \rightarrow & \left( \begin{array}{cc|cc} a & 0 & \frac{ad}{ad-bc} & -\frac{ab}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right) \text{ Multipliziere 1. mit } \frac{1}{a} \\
 \rightarrow & \left( \begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right)
 \end{aligned}$$

Wir sehen also: Ist  $ad - bc \neq 0$ , so ist  $A$  invertierbar, und es gilt

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

## 14 Die LR-Zerlegung

Muss man ein System  $A\vec{x} = \vec{b}$  häufig und für verschiedene rechte Seiten  $\vec{b}$  lösen, so erscheint es unnützlich, jedes Mal wieder das ganze Eliminationsverfahren durchzuführen. Wir werden sehen, dass das auch nicht nötig ist, wenn man sich die richtigen Sachen merkt.

Schauen wir uns noch einmal den  $k$ -ten Schritt im Eliminationsverfahren an: Wir haben

$$A^{(k)}\vec{x} = \vec{b}^{(k)} : \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ & \ddots & & & \vdots \\ & & a_{kk} & \cdots & a_{kn} \\ & & \vdots & & \vdots \\ & & a_{nk} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_k^{(k)} \\ \vdots \\ b_n^{(k)} \end{pmatrix}$$

Der Einfachheit halber nehmen wir an, dass  $A$  quadratisch ist, und dass das Eliminationsverfahren keinen Zeilen- oder Spaltentausch braucht.

Wenn wir  $a_{kk}$  als Pivotelement nehmen, bekommen wir nach Elimination

$$A^{(k+1)}\vec{x} = \vec{b}^{(k+1)} : \begin{pmatrix} a_{11} & \cdots & \cdots & \cdots & \cdots & a_{1n} \\ & \ddots & & & & \vdots \\ & & a_{kk} & \cdots & \cdots & a_{kn} \\ & & 0 & \widetilde{a_{k+1,k+1}} & \cdots & * \\ & & \vdots & \vdots & & \vdots \\ & & 0 & * & \cdots & \widetilde{a_{nn}} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_k^{(k)} \\ b_{k+1}^{(k+1)} \\ \vdots \\ b_n^{(k+1)} \end{pmatrix}$$

Wie können wir den Übergang von  $(A^{(k)} \mid \vec{b}^{(k)})$  zu  $(A^{(k+1)} \mid \vec{b}^{(k+1)})$  beschreiben?

*Beispiel.* Wir machen uns diesen Schritt an einem kleinen  $2 \times 2$  Beispiel klar:

$$\begin{pmatrix} 1 & 2 \\ 5 & 11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 16 \end{pmatrix}, \quad \left( \begin{array}{cc|c} 1 & 2 & 3 \\ 5 & 11 & 16 \end{array} \right).$$

Wir haben  $k = 1$  und  $a_{11} = 1$ . Im Eliminationsschritt ziehen wir das 5-fache der ersten Zeile von der zweiten ab. Das können wir durch eine Multiplikation von links mit folgender Matrix machen:

$$M_1 = \begin{pmatrix} 1 & 0 \\ -5 & 1 \end{pmatrix}.$$

Wir rechnen

$$M_1 A = \begin{pmatrix} 1 & 0 \\ -5 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 5 & 11 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad M_1 \vec{b} = \begin{pmatrix} 1 & 0 \\ -5 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 16 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Im allgemeinen Fall können wir den  $k$ -ten Eliminationsschritt durch die Multiplikation von links mit der Matrix

$$M_k = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & -l_{k+1,k} & 1 & \\ & & \vdots & & \ddots \\ & & -l_{n,k} & & & 1 \end{pmatrix} \quad \text{mit } l_{ik} = \frac{a_{ik}}{a_{kk}}$$

darstellen, denn dann gilt  $A^{(k+1)} = M_k A^{(k)}$  und  $\vec{b}^{(k+1)} = M_k \vec{b}^{(k)}$ .

Als Ergebnis des Eliminationsalgorithmus bekommen wir die eine rechte obere Dreiecksmatrix

$$R = A^{(n)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n)} \end{pmatrix}.$$

Die Matrix  $R$  bekommen wir (wenn wir keine Vertauschungen vornehmen mussten) mit Hilfe der  $M_k$  als

$$R = M_{n-1} \cdot \dots \cdot M_2 M_1 A.$$

Die  $M_k$  heißen *Eliminationsmatrizen* und sind, wie wir gleich sehen werden, immer regulär. Daher ist

$$A = M_1^{-1} \cdot \dots \cdot M_{n-1}^{-1} R = LR, \quad L := M_1^{-1} \cdot \dots \cdot M_{n-1}^{-1}.$$

Der nächste Satz zeigt, dass  $L^{-1}$  immer eine normierte, linke, untere Dreiecksmatrix ist:

Das Wort „normiert“ meint hier, dass auf der Diagonale von  $L$  Einsen stehen.

**Satz 14.1.** (a) Es gilt

$$M_k^{-1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & l_{k+1,k} & 1 & \\ & & \vdots & & \ddots \\ & & l_{n,k} & & & 1 \end{pmatrix}.$$

(b) Es gilt

$$L := M_1^{-1} \cdot \dots \cdot M_{n-1}^{-1} = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{pmatrix}$$

*Beweis.*

■ Dies kann man beides einfach nachrechnen.  $\square$

Wir haben also  $A$  als Produkt von (normierter) linker, unterer Dreiecksmatrix und rechter oberer Dreiecksmatrix faktorisiert:

$$A = LR = (\triangle) (\nabla),$$

eine solche Zerlegung nennt man *LR-Zerlegung*.

*Beispiel.* Wir illustrieren das an einem Beispiel: Wir betrachten

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

Im ersten Schritt haben wir  $l_{21} = \frac{2}{4} = \frac{1}{2}$ ,  $l_{31} = 0$ , und  $l_{41} = \frac{2}{4} = \frac{1}{2}$ , also

$$M_1 = \begin{pmatrix} 1 & & & \\ -\frac{1}{2} & 1 & & \\ 0 & 0 & 1 & \\ -\frac{1}{2} & 0 & 0 & 1 \end{pmatrix}, \quad A^{(2)} = \begin{pmatrix} 4 & 3 & 0 & 1 \\ 0 & -\frac{1}{2} & 4 & -\frac{1}{2} \\ 0 & 0 & 4 & 1 \\ 0 & -\frac{3}{2} & 1 & -\frac{1}{2} \end{pmatrix}.$$

Im zweiten Schritt ist  $l_{32} = 0$  und  $l_{42} = 3$ , also

$$M_2 = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & -3 & 0 & 1 \end{pmatrix}, \quad A^{(3)} = \begin{pmatrix} 4 & 3 & 0 & 1 \\ 0 & -\frac{1}{2} & 4 & -\frac{1}{2} \\ 0 & 0 & 4 & 1 \\ 0 & 0 & -11 & 1 \end{pmatrix}.$$

Im dritten Schritt haben wir  $l_{43} = -\frac{11}{4}$  und wir bekommen

$$M_3 = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & \frac{11}{4} & 1 \end{pmatrix}, \quad A^{(4)} = \begin{pmatrix} 4 & 3 & 0 & 1 \\ 0 & -\frac{1}{2} & 4 & -\frac{1}{2} \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & \frac{15}{4} \end{pmatrix} =: R.$$

Berechnen wir jetzt die Inversen:

$$M_1^{-1} = \begin{pmatrix} 1 & & & \\ \frac{1}{2} & 1 & & \\ 0 & 0 & 1 & \\ \frac{1}{2} & 0 & 0 & 1 \end{pmatrix}, \quad M_2^{-1} = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 3 & 0 & 1 \end{pmatrix}$$

$$M_3^{-1} = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & -\frac{11}{4} & 1 \end{pmatrix}.$$

Das gibt uns

$$L := M_1^{-1} M_2^{-1} M_3^{-1} = \begin{pmatrix} 1 & & & \\ \frac{1}{2} & 1 & & \\ 0 & 0 & 1 & \\ \frac{1}{2} & 3 & -\frac{11}{4} & 1 \end{pmatrix}.$$

Auf Englisch nennt man diese Zerlegung *LU decomposition*. Im Allgemeinen kann man nicht auf einen Zeilentausch verzichten. Wenn man allerdings über das Tauschen von Zeilen Buch führt, kann man aber immer eine Zerlegung der Form  $PA = LR$  mit einer *Permutationsmatrix*  $P$  erreichen, welche die Zeilen von  $A$  vertauscht.

Wir haben also eine LR-Zerlegung von  $A$  gefunden:

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

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & 3 & -\frac{11}{4} & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 & 0 & 1 \\ 0 & -\frac{1}{2} & 4 & -\frac{1}{2} \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & \frac{15}{4} \end{pmatrix}.$$

Wie nutzen wir dies nun als, um ein gegebenes Gleichungssystem  $A\vec{x} = \vec{b}$  schnell zu lösen? Wir schreiben

$$A\vec{x} = L(R\vec{x}) = \vec{b}$$

und setzen  $\vec{y} := R\vec{x}$ . Um das Gleichungssystem zu lösen benutzen wir also folgende Schritte

1. Löse das Gleichungssystem  $L\vec{y} = \vec{b}$  durch *Vorwärtseinsetzen*:

$$y_i = b_i - \sum_{j=1}^{i-1} l_{ij}y_j, \quad i = 1, \dots, n.$$

2. Löse das Gleichungssystem  $R\vec{x} = \vec{y}$  durch *Rückwärtseinsetzen*:

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

*Beispiel.* Wir nehmen eine schon zerlegte Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = LR \text{ mit } \vec{b} = \begin{pmatrix} 0 \\ -2 \\ -5 \end{pmatrix}.$$

Die Lösung von

$$L\vec{y} = \vec{b}: \left( \begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & -2 \\ 3 & 1 & 1 & -5 \end{array} \right)$$

ist mit Vorwärtseinsetzen

$$\begin{aligned} y_1 &= 0 \\ 2y_1 + y_2 &= -2 \implies y_2 = -2 \\ 3y_1 + y_2 + y_3 &= -5 \implies y_3 = -5 - 0 - (-2) = -3. \end{aligned}$$

Das Rückwärtseinsetzen für

$$R\vec{x} = \vec{y}: \left( \begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 1 & -3 \end{array} \right)$$

gibt

$$\begin{aligned} x_3 &= -3 \\ x_2 + 2x_3 &= -2 \implies x_2 = -2 + 6 = 4 \\ x_1 + x_2 + x_3 &= 0 \implies x_1 = -4 + 3 = -1 \end{aligned} \quad , \quad \text{also } \vec{x} = \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix} .$$

Überschlagen wir den Rechenaufwand für das Lösen eines Gleichungssystems per Eliminationsverfahren und nach LR-Zerlegung:

- **Eliminationsverfahren:** Im  $k$ -ten Eliminationsschritt brauchen wir  $n - k$  Operationen um die Faktoren  $a_{ik}/a_{kk}$  zu berechnen. In der Elimination brauchen wir dann für jedes Element in den letzten  $n - k$  Zeilen zwei Operationen, nämlich  $\widetilde{a}_{r,s} = a_{r,s} - l_{r,s}a_{k,s}$ . Da es  $(n - k)(n - k + 1)$  Elemente in den letzten  $n - k$  Zeilen gibt, sind das  $2(n - k)(n - k + 1)$  Operationen. Insgesamt benötigt der  $k$ -te Eliminationsschritt  $2(n - k)(n - k + 1) + (n - k)$  Operationen. Die Gesamtzahl der Operationen ist also

$$\sum_{k=1}^{n-1} [2(n - k)(n - k + 1) + (n - k)] = \frac{2n^3 + 3n^2 - 5n}{3} .$$

Der höchste Exponent an  $n$  dominiert für große  $n$  und daher sagt man, dass der Aufwand für das Eliminationsverfahren *kubisch in  $n$*  ist.

Der Aufwand für das Berechnen einer LR-Zerlegung ist genau so, da dafür nichts weiter berechnet werden muss.

- **Lösen nach LR-Zerlegung:** Beim Vorwärtseinsetzen brauchen wir  $2(k - 1)$  Operationen für die  $k$ -te Zeile. Insgesamt haben wir also ist die Anzahl der Operationen für das Vorwärtseinsetzen:

$$\sum_{k=1}^n 2(k - 1) = n^2 - n .$$

Beim Rückwärtseinsetzen kommen wir auf die gleiche Anzahl von Operationen plus die  $n$  Divisionen durch die Diagonalelemente. Es ergeben sich also  $n^2$  Operationen, so dass wir insgesamt auf

$$2n^2 - n$$

Operationen kommen. Der Aufwand für das Lösen nach LR-Zerlegung ist daher nur *quadratisch in  $n$*  und daher um eine Größenordnung kleiner.

Genauer: Das Lösen nach LR-Zerlegung geht grob  $\frac{3}{4}n$ -mal schneller als das Lösen per Eliminationsverfahren!

## 15 Permutationen und Determinanten

Mit Determinanten berechnet man  $n$ -dimensionale Volumina von sogenannten Parallelotopen, also den Verallgemeinerungen von Parallelogrammen. In zwei und drei Dimensionen können wir das schon mit dem Kreuzprodukt und dem Spatprodukt. Um dies in  $n$  Dimensionen zu können, müssen wir etwas über *Permutationen* (Vertauschungen) wissen. Vorher führen wir noch drei Begriffe ein:

**Definition 15.1.** Eine Abbildung  $f : X \rightarrow Y$  zwischen zwei Mengen  $X$  und  $Y$  heißt *surjektiv*, wenn die Gleichung  $f(x) = y$  für jedes  $y$  mindestens eine Lösung hat, und *injektiv*, wenn sie für jedes  $y$  höchstens eine Lösung hat. Die Abbildung  $f$  heißt *bijektiv*, wenn sie *injektiv* und *surjektiv* ist.

**Definition 15.2.** (a) Eine *Permutation* ist eine bijektive Abbildung  $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ . Die Menge aller Permutationen von  $n$  Elementen bezeichnen wir mit  $S_n$ . Wir notieren  $\sigma_i := \sigma(i)$  und geben Permutationen auch als  $(\sigma_1, \dots, \sigma_n)$  an

(b) Für Permutationen ist eine Verknüpfung durch hinterausführen der Abbildungen definiert, d.h.

$$\mu \circ \sigma = (\mu_1, \dots, \mu_n) \circ (\sigma_1, \dots, \sigma_n) := (\mu_{\sigma_1}, \dots, \mu_{\sigma_n}).$$

*Beispiel.* Wir nehmen  $n = 4$  und  $\sigma = (2,4,1,3)$  und  $\mu = (1,3,2,4)$ . Dann ist

$$\mu \circ \sigma = (1,3,2,4) \circ (2,4,1,3) = (\mu_{\sigma_1}, \mu_{\sigma_2}, \mu_{\sigma_3}, \mu_{\sigma_4}).$$

Es ist

$$\begin{aligned} \mu_{\sigma_1} &= \mu_2 = 3, & \mu_{\sigma_2} &= \mu_4 = 4, \\ \mu_{\sigma_3} &= \mu_1 = 1, & \mu_{\sigma_4} &= \mu_3 = 2, \end{aligned}$$

also

$$(1,3,2,4) \circ (2,4,1,3) = (3,4,1,2).$$

*Bemerkung.* Die Menge  $S_n$  bildet mit der Verknüpfung eine mathematische Struktur, die *Gruppe* genannt wird. Sie heißt *symmetrische Gruppe*. Für  $n > 2$  ist die Verknüpfung von Permutationen nicht kommutativ, d.h. im Allgemeinen gilt  $\mu \circ \sigma \neq \sigma \circ \mu$ . ■

*Beispiel.*

$$\begin{aligned} (1,3,2,4) \circ (2,4,1,3) &= (3,4,1,2) \\ (2,4,1,3) \circ (1,3,2,4) &= (2,1,4,3) \end{aligned}$$

**Definition 15.3.** Eine Permutation die genau zwei Elemente vertauscht, nennen wir *Transposition*. Die Transposition, die die Elemente  $i$  und  $j$  vertauscht, bezeichnen wir mit  $\tau := [i, j]$ .

Also ist eine Abbildung  $f$  bijektiv, wenn die Gleichung  $f(x) = y$  für jedes  $y$  genau eine Lösung hat. Bijektive Abbildungen sind also genau die umkehrbaren Abbildungen.

**Satz 15.4.** (a) Jede Permutation lässt sich als Produkt von Transpositionen schreiben.

(b) Lässt sich eine Permutation auf zwei Arten als Produkt von Transpositionen schreiben, so ist die Anzahl der Transpositionen entweder in beiden Fällen gerade oder ungerade.

Für Transpositionen gelten folgende Rechenregeln: ( $m \neq n$ ,  $i \neq j$ )

$$\begin{aligned} (R1) \quad & [i, n] \circ [i, n] = \text{id} \\ (R2) \quad & [i, n] \circ [i, m] = [i, m] \circ [m, n] \\ (R3) \quad & [i, n] \circ [j, n] = [i, j] \circ [i, n] \\ (R4) \quad & [i, n] \circ [j, m] = [j, m] \circ [i, n]. \end{aligned}$$

*Beispiel.* Um  $\sigma = (5,3,4,1,2) \in S_5$  als Produkt von Transpositionen zu schreiben, rechnen wir:

$$\begin{aligned} \sigma_1 &= [1,5] \circ \sigma = (1,3,4,5,2) \\ \sigma_2 &= [2,3] \circ \sigma_1 = (1,2,4,5,3) \\ \sigma_3 &= [3,4] \circ \sigma_2 = (1,2,3,5,4) \\ \sigma_4 &= [4,5] \circ \sigma_3 = (1,2,3,4,5) = \text{id}. \end{aligned}$$

Also ist  $\text{id} = [4,5] \circ [3,4] \circ [2,3] \circ [1,5] \circ \sigma$  und daher

$$\sigma = [1,5] \circ [2,3] \circ [3,4] \circ [4,5]$$

**Definition 15.5.** Ist eine Permutation  $\sigma$  als Produkt von  $k$  Transpositionen dargestellt, so definieren wir das *Vorzeichen* (Signum) von  $\sigma$  als

$$\text{sign}(\sigma) := (-1)^k.$$

Ist  $\text{sign}(\sigma) = 1$ , so heißt  $\sigma$  *gerade*, sonst *ungerade*.

Es gilt:

- (a) Ist  $\tau$  eine Transposition, so ist  $\text{sign}(\tau) = -1$ .
- (b) Für beliebige Permutationen ist  $\text{sign}(\sigma \circ \mu) = \text{sign}(\sigma) \text{sign}(\mu)$ .
- (c) Für jede Permutation gilt  $\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)$ .

Kommen wir nun zur allgemeinen Determinante. Für die Intuition kann man sich folgendes merken:

Die Determinante einer Matrix  $A \in \mathbb{K}^{n \times n}$  mit Spalten  $\vec{a}^{(1)}, \dots, \vec{a}^{(n)}$  ist das (signierte)  $n$ -dimensionale Volumen des von den Spaltenvektoren aufgespannten Parallelotops.

Nach Satz 15.4 ist  $\text{sign}(\sigma)$  wohldefiniert, d.h. unabhängig von der Darstellung als Produkt von Transpositionen. Im Allgemeinen sagt man, dass etwas wohldefiniert ist, wenn die Definition von etwas anzuhängen scheint, aber trotzdem unabhängig davon ist (das Signum von Permutationen hängt eben nicht von der bestimmten Zerlegung in Transpositionen ab, sondern nur davon, ob die Anzahl von Transpositionen einer beliebigen Zerlegung gerade oder ungerade ist).

Beachte: Determinanten sind generell nur für quadratische Matrizen erklärt.

Hierbei bedeutet „signiert“, dass das Volumen ein Vorzeichen haben kann und das  $n$ -dimensionale Volumen erklärt sich im wesentlichen durch wiederholte Anwendung der Regel „Grundfläche mal Höhe“. So hat ein  $n$ -dimensionaler Quader mit den Seitenlängen  $a_1, \dots, a_n$  das  $n$ -dimensionale Volumen  $a_1 \cdot \dots \cdot a_n$ .

In den Dimensionen  $n = 2, 3$  kennen wir schon Formeln dafür, die wir mit Hilfe des Kreuz- und des Spatproduktes hergeleitet haben!

- $n = 2$ : Die Fläche des Parallelograms welches von den Vektoren

$$\vec{a}^{(1)} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \vec{a}^{(2)} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

aufgespannt wird, ist

$$a_{11}a_{22} - a_{12}a_{21}$$

(wobei der Flächeninhalt negativ ist, wenn, grob gesagt,  $\vec{a}^{(1)}$  „links von“  $\vec{a}^{(2)}$  liegt).

Wir haben hierbei in jedem Summanden die Faktoren so sortiert, dass der erste Index wächst. Die zweiten Indizes im ersten Summanden sind dann (1,2) und im zweiten Summanden (2,1), also jeweils eine Permutation von (1,2). Schauen wir die Vorzeichen an:

$$\begin{aligned} (1,2) &= \text{id}, & \rightsquigarrow \text{sign}(1,2) &= 1 \\ (2,1) &= [1,2], & \rightsquigarrow \text{sign}(2,1) &= -1. \end{aligned}$$

Das Signum der Permutation der zweiten Indizes gibt also das Vorzeichen des Summanden!

- $n = 3$ : Hier kennen wir für das Volumen des Parallelotops welches von den Vektoren

$$\vec{a}^{(1)} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad \vec{a}^{(2)} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \vec{a}^{(3)} = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

aufgespannt wird, die Formel

$$\begin{aligned} &a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Wieder haben wir in den Summanden die ersten Indizes aufsteigend geordnet. Untersuchen wir die Vorzeichen der Permutationen, die zu den zweiten Indizes gehören:

$$\left. \begin{aligned} (1,2,3) &= \text{id} \\ (2,3,1) &= [1,2] \circ [2,3] \\ (3,1,2) &= [1,3] \circ [2,3] \end{aligned} \right\} \text{Gerade Permutationen, also sign} = +1$$

$$\left. \begin{aligned} (3,2,1) &= [1,3] \\ (1,3,2) &= [2,3] \\ (2,1,3) &= [1,2] \end{aligned} \right\} \text{Ungerade Permutationen, also sign} = -1$$

In beiden Fällen gilt:

- Wir haben für jede Permutation einen Summanden mit je  $n$  Faktoren.
- Sind die ersten Indizes aufsteigend geordnet, so die zweiten Indizes eine Permutation.
- Das Vorzeichen des Summanden ist das Vorzeichen der Permutation.

Dies nehmen wir als Definition des allgemeinen Falls:

**Definition 15.6.** Für  $A \in \mathbb{K}^{n \times n}$  definieren wir die *Determinante* von  $A$  durch

$$\det(A) := \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Diese Formel für die Determinante ist nicht von großer praktischer Bedeutung. Nur in sehr wenigen Fällen kann man sie einfach anwenden um Determinanten zu berechnen. Hier eine Ausnahme dazu:

*Beispiel.* Es sei  $A$  eine obere Dreiecksmatrix, d.h.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}.$$

Jeder Summand in der Formel der Determinante enthält aus jeder Zeile und jeder Spalte genau einen Eintrag als Faktor. Also ist unter diesen Faktoren fast immer mindestens einer der „unterhalb der Diagonalen“ liegt. Die einzige Ausnahme ist der Eintrag der zu  $\sigma = \text{id}$  gehört und daher ist in diesem Fall

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix} = a_{11} \cdots a_{nn}.$$

Man kann die Formel aus Definition 15.6 aber gut benutzen, um Aussagen zu beweisen. Zum Beispiel lässt sich dieser Satz relativ einfach direkt mit der Formel zeigen.

**Satz 15.7.** Für  $A \in \mathbb{K}^{n \times n}$  gilt  $\det(A) = \det(A^T)$ .

Am folgenden Satz schauen wir uns an, wie ein solcher Beweis funktioniert.

**Satz 15.8.** Vertauscht man in einer Matrix zwei Zeilen, so ändert die Determinante nur das Vorzeichen.

*Beweis.*

Diese Formel für die Determinante wird auch *Leibniz-Formel* genannt.

Schon einmal vorweg: Auf diese Weise werden Determinanten eigentlich niemals ausgerechnet. Es gibt meist bessere und viel schnellere Wege!

Es sei  $\tilde{A}$  die Matrix, die aus  $A$  entsteht, wenn man die  $k$ -te und  $l$ -te Spalte tauscht (und wir nehmen  $k \leq l$  an). Dann ist Wir schauen uns an, was passiert, wenn wir die  $k$ -te und die  $l$ -te Zeile einer Matrix  $A$  tauschen:

$$\begin{aligned} \det(\tilde{A}) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \cdots a_{l\sigma(l)} \cdots a_{k\sigma(k)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma \circ [k, l]) a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{l\sigma(l)} \cdots a_{n\sigma(n)} \\ &= - \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{l\sigma(l)} \cdots a_{n\sigma(n)} = -\det(A). \end{aligned}$$

□

Wegen Satz 15.7 gilt auch, dass die Determinante ihr Vorzeichen wechselt, wenn man zwei Spalten vertauscht.

## 16 Eigenschaften von Determinanten

Die Determinante hat folgende wichtige Eigenschaften:

**Satz 16.1.** (a) Für die Einheitsmatrix gilt  $\det(I_n) = 1$ .

(b) Sind in einer Matrix  $A$  zwei Zeilen gleich, so gilt  $\det(A) = 0$  (ebenso für Spalten).

(c) Für eine Vertauschung der Zeilen mit einer Permutation  $\sigma$  gilt

$$\det \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} = \text{sign}(\sigma) \det \begin{pmatrix} \vec{a}_{\sigma(1)}^T \\ \vdots \\ \vec{a}_{\sigma(n)}^T \end{pmatrix}$$

(und für Spalten analog).

(d) Die Determinante ist eine  $n$ -Linearform, d.h. sie ist eine lineare Abbildung bzgl. jedes Zeilenvektors, genauer

$$\det \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_i^T + \vec{b}_i^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} = \det \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_i^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} + \det \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{b}_i^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix}$$

$$\det \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \alpha \vec{a}_i^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} = \alpha \det \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_i^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix}$$

Im Allgemeinen gilt  
 $\det(\alpha A) \neq \alpha \det(A)$  (sondern  
 $\det(\alpha A) = \alpha^n \det(A)$ ) und  
 $\det(A + B) \neq \det(A) + \det(B)$  (es  
gibt gar keine brauchbare Formel für  
 $\det(A + B)$ ).

(und ebenso ist sie in jeder Spalte linear).

Eine wichtige Folgerung dieses Satzes ist: Die Determinante einer Matrix ändert sich nicht bei der Anwendung von elementaren Zeilen- (oder Spalten-) Umformungen (nur bei Tauschungen, muss das Vorzeichen geändert werden). Zusammen mit der Erkenntnis, dass die Determinante einer oberen Dreiecksmatrix das Produkt der Diagonalelemente ist (letztes Beispiel n vorigem Abschnitt), ergibt sich eine elegante Methode zum Berechnen von Determinanten:

Bringe  $A$  mit dem Eliminationsalgorithmus auf obere Dreiecksform. Notiere dabei die Anzahl von Zeilen- und Spaltentausch und alle Faktoren, die man aus einzelnen Zeilen herauszieht. Am Ende bilde man das Produkt der Diagonalelemente.

*Beispiel.* Schauen wir uns ein  $4 \times 4$ -Beispiel an:

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

Da wir keinen Zeilen- oder Spaltentausch durchgeführt haben und keine Faktoren herausgezogen haben, ergibt sich als Determinante  $1 \cdot 2 \cdot (-2) \cdot 3 = -12$ .

Unser Wissen über den Eliminationsalgorithmus liefert uns eine weitere Charakterisierung für die Regularität einer Matrix. Für eine Matrix  $A \in \mathbb{K}^{n \times n}$  sind (unter anderem) äquivalent:

- (i)  $\det(A) \neq 0$
- (ii) die Zeilen von  $A$  sind linear unabhängig
- (iii) die Spalten von  $A$  sind linear unabhängig
- (iv)  $A$  ist invertierbar

**Satz 16.2** (Determinanten-Multiplikationssatz). Für  $A, B \in \mathbb{K}^{n \times n}$  gilt  $\det(AB) = \det(A) \det(B)$ .

**Korollar 16.3.** (a) Ist  $A \in \mathbb{K}^{n \times n}$  invertierbar, so gilt

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

(b) Das Produkt  $AB$  ist genau dann invertierbar, wenn sowohl  $A$  als auch  $B$  invertierbar sind.

- (a) Es gilt  $1 = \det(I_n) = \det(AA^{-1}) = \det(A) \det(A^{-1})$ .
- (b)  $\det(AB) \neq 0$  genau dann, wenn  $\det(A) \det(B) \neq 0$  was genau dann der Fall ist, wenn  $\det(A) \neq 0$  und  $\det(B) \neq 0$ .

Eine weitere Methode, um Determinanten auszurechnen, ist der Entwicklungssatz (von Laplace). Um diesem zu formulieren, führen wir folgende Schreibweise ein: Für eine Matrix  $A \in \mathbb{K}^{n \times n}$  und  $i, j \in \{1, \dots, n\}$  bezeichnen wir mit  $A_{ij}$  die  $(n-1) \times (n-1)$ -Matrix, die aus  $A$  entsteht, indem wir die  $i$ -te Zeile und  $j$ -te Spalte streichen.

**Satz 16.4** (Entwicklungssatz). Für alle  $i \in \{1, \dots, n\}$  gilt

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

und 
$$\det(A) = \sum_{k=1}^n (-1)^{i+k} a_{ki} \det(A_{ki}).$$

Am einfachsten versteht man das an einem kleinen Beispiel:

*Beispiel.* Wir entwickeln nach der ersten Spalte: Es gilt

$$\begin{aligned} \det \begin{pmatrix} 1 & 3 & 4 \\ 5 & 0 & 2 \\ 7 & 1 & 4 \end{pmatrix} &= 1 \cdot \det \begin{pmatrix} 0 & 2 \\ 1 & 4 \end{pmatrix} - 5 \cdot \det \begin{pmatrix} 3 & 4 \\ 1 & 4 \end{pmatrix} + 7 \cdot \det \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix} \\ &= 1 \cdot (0 - 2) - 5 \cdot (12 - 4) + 7 \cdot (6 - 0) \\ &= -2 - 40 + 42 = 0. \end{aligned}$$

Wir könnten auch nach der zweiten Zeile entwickeln:

$$\begin{aligned} \det \begin{pmatrix} 1 & 3 & 4 \\ 5 & 0 & 2 \\ 7 & 1 & 4 \end{pmatrix} &= -5 \cdot \det \begin{pmatrix} 3 & 4 \\ 1 & 4 \end{pmatrix} + 0 \cdot \det \begin{pmatrix} 1 & 4 \\ 7 & 4 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 1 & 3 \\ 7 & 1 \end{pmatrix} \\ &= -5 \cdot (12 - 4) + 0 - 2 \cdot (1 - 21) \\ &= -40 - (-40) = 0. \end{aligned}$$

Fasst man die Determinanten der Untermatrizen  $A_{ik}$  zu einer neuen Matrix wie folgt zusammen

$$\tilde{A} := \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}^T, \quad \text{mit } \alpha_{ik} := (-1)^{i+k} \det(A_{ik}),$$

so erhält man die zu  $A$  komplementäre Matrix. Für diese Matrix gilt

$$A\tilde{A} = \tilde{A}A = \det(A)I_n.$$

Für invertierbare  $A$  folgt daraus

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

*Beispiel.* Wir betrachten

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Es ist  $\det(A) = ad - bc$ . Die komplementäre Matrix ist

$$\tilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

und wir schließen

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

Im ersten Fall spricht man von „Entwicklung nach der  $i$ -ten Zeile“ und im zweiten von „Entwicklung nach der  $i$ -ten Spalte.“

Für die Vorzeichen im Entwicklungssatz merkt man sich am besten, dass diese nach einem Schachbrettmuster verteilt sind:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \ddots \\ + & - & + & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

In der Theorie kann man Determinanten benutzen, um lineare Gleichungssysteme zu lösen:

**Satz 16.5** (Cramersche Regel). Ist  $A = (\vec{a}^{(1)} \ \dots \ \vec{a}^{(n)}) \in \mathbb{K}^{n \times n}$  invertierbar und  $b \in \mathbb{K}^n$ , so ist die eindeutige Lösung  $\vec{x}$  des Gleichungssystem  $A\vec{x} = \vec{b}$  durch die Cramersche Regel gegeben:

$$x_k = \frac{\det \begin{pmatrix} \vec{a}^{(1)} & \dots & \vec{b} & \dots & \vec{a}^{(n)} \end{pmatrix}}{\det \begin{pmatrix} \vec{a}^{(1)} & \dots & \vec{a}^{(k)} & \dots & \vec{a}^{(n)} \end{pmatrix}}, \quad k = 1, \dots, n$$

*Beweis.*

Wir nutzen die Darstellung von  $\vec{b}$  als Linearkombination von Spalten von  $A$ , nämlich  $\vec{b} = \sum_{k=1}^n x_k \vec{a}^{(k)}$ . Damit rechnen wir die Determinante auf dem Bruchstrich aus indem wir die Linearität der Determinante benutzen. Beachte, dass der Vektor  $\vec{b}$  als  $k$ -te Spalte eingesetzt wurde:

$$\begin{aligned} \det \begin{pmatrix} \vec{a}^{(1)} & \dots & \vec{b} & \dots & \vec{a}^{(n)} \end{pmatrix} &= \det \begin{pmatrix} \vec{a}^{(1)} & \dots & \sum_{j=1}^n x_j \vec{a}^{(j)} & \dots & \vec{a}^{(n)} \end{pmatrix} \\ &= \sum_{j=1}^n x_j \det \underbrace{\begin{pmatrix} \vec{a}^{(1)} & \dots & \vec{a}^{(j)} & \dots & \vec{a}^{(n)} \end{pmatrix}}_{=0 \text{ für } j \neq k} \\ &= x_k \det \begin{pmatrix} \vec{a}^{(1)} & \dots & \vec{a}^{(k)} & \dots & \vec{a}^{(n)} \end{pmatrix}. \end{aligned}$$

□

Kommen wir nun zurück zur Interpretation der Determinante als Volumenfunktion: Zu jeder  $n \times n$  Matrix  $A$  mit Spalten  $\vec{a}^{(1)}, \dots, \vec{a}^{(n)}$  betrachten wir das von den Spaltenvektoren aufgespannte Parallelotop

$$P_A := \left\{ \sum_{i=1}^n \alpha_i \vec{a}^{(i)} \mid 0 \leq \alpha_i \leq 1 \right\}.$$

Dazu definieren wir:

**Definition 16.6.** Eine Abbildung  $V : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$  heißt (signierte) *Volumenfunktion*, wenn sie folgende Eigenschaften hat:

$$(V1) \quad V(\vec{a}^{(1)}, \dots, \vec{a}^{(i)}, \dots, \vec{a}^{(j)}, \dots, \vec{a}^{(i)}, \dots, \vec{a}^{(n)}) = -V(\vec{a}^{(1)}, \dots, \vec{a}^{(j)}, \dots, \vec{a}^{(i)}, \dots, \vec{a}^{(n)})$$

$$(V2) \quad V(\vec{a}^{(1)}, \dots, \alpha \vec{a}^{(i)}, \dots, \vec{a}^{(n)}) = \alpha V(\vec{a}^{(1)}, \dots, \vec{a}^{(i)}, \dots, \vec{a}^{(n)})$$

$$(V3) \quad V(\vec{a}^{(1)}, \dots, \vec{a}^{(i)} + \vec{b}^{(i)}, \dots, \vec{a}^{(n)}) = V(\vec{a}^{(1)}, \dots, \vec{a}^{(i)}, \dots, \vec{a}^{(n)}) + V(\vec{a}^{(1)}, \dots, \vec{b}^{(i)}, \dots, \vec{a}^{(n)})$$

$$(V4) \quad V(\vec{e}_1, \dots, \vec{e}_n) = 1$$

- (V4) entspricht geometrisch der Tatsache, dass der  $n$ -dimensionale Einheitswürfel das  $n$ -dimensionale Volumen 1 hat.

Es macht überhaupt keinen Sinn, diese Regel in der Praxis anzuwenden. Es müssen sehr viele Determinanten berechnet werden, was selbst mindestens so aufwändig ist, wie einmal das Eliminationsverfahren anzuwenden.

- (V<sub>3</sub>) besagt, dass das  $n$ -dimensionale Volumen sich additiv Verhält, wenn wir zu einer Seite einen Vektor addieren.
- (V<sub>2</sub>) besagt, dass die Skalierung einer Spalte einer Matrix  $A$  mit  $\alpha$  das  $n$ -dimensionale Volumen von  $P_A$  ebenfalls mit  $\alpha$  skaliert.
- (V<sub>1</sub>) gibt schließlich eine Orientierung vor (Vertauschen wir die Reihenfolge der aufspannenden Vektoren, kehrt sich das Vorzeichen um).

Die Determinante erfüllt alle Eigenschaften (V<sub>1</sub>) bis (V<sub>4</sub>) und ist daher eine Volumenfunktion! Es gilt sogar: Die Determinante ist die *einzig*e Funktion  $\mathbb{K}^{n \times n} \rightarrow \mathbb{K}$  die (V<sub>1</sub>) bis (V<sub>4</sub>) erfüllt!

Wir benutzen zur Formulierung des nächsten Satzes die Schreibweise  $\text{Vol}(A)$  für das  $n$ -dimensionale (signierte) Volumen des Parallelotops  $P_A$ .

**Satz 16.7.** *Es seien  $A, B \in \mathbb{K}^{n \times n}$  und  $BP_A$  das Parallelotop, welches entsteht, wenn wir alle Elemente  $x \in P_A$  mit  $B$  abbilden. Dann gilt*

$$\text{Vol}(BP_A) = \det(B) \text{Vol}(P_A).$$

*Beweis.*

Die Menge  $BP_A$  ist genau das Parallelotop, welches von den Vektoren  $B\vec{a}^{(1)}, \dots, B\vec{a}^{(n)}$  aufgespannt wird. Daher ist

$$\begin{aligned} \text{Vol}(BP_A) &= \det(B(\vec{a}^{(1)}, \dots, \vec{a}^{(n)})) \\ &= \det(B) \det(\vec{a}^{(1)}, \dots, \vec{a}^{(n)}) = \det(B) \text{Vol}(P_A). \end{aligned}$$

□

## 17 Lineare Abbildungen

Betrachten wir nun wieder etwas abstrakter zwei allgemeine  $\mathbb{K}$ -Vektorräume  $V$  und  $W$  und Abbildung  $T : V \rightarrow W$  zwischen ihnen.

**Definition 17.1.** Eine Abbildung  $T : V \rightarrow W$  zwischen zwei  $\mathbb{K}$ -Vektorräumen heißt *linear*, wenn für alle  $\vec{u}, \vec{v} \in V$  und  $\alpha \in \mathbb{K}$  gilt

$$\begin{aligned} T(\vec{u} + \vec{v}) &= T(\vec{u}) + T(\vec{v}) \\ T(\alpha\vec{v}) &= \alpha T(\vec{v}). \end{aligned}$$

Aus dem, was wir über Matrizen schon wissen, sehen wir sofort dass  $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$  definiert durch  $T(\vec{x}) = A\vec{x}$  mit  $A \in \mathbb{K}^{m \times n}$  eine lineare Abbildung ist.

**Definition 17.2.** Für eine lineare Abbildung  $T : V \rightarrow W$  sind der Kern und das Bild definiert als

$$\begin{aligned} \text{Kern}(T) &= \{ \vec{v} \in V \mid T(\vec{v}) = \vec{0} \} \subset V \\ \text{Bild}(T) &= \{ \vec{w} \in W \mid \exists \vec{v} \in V : T(\vec{v}) = \vec{w} \} \subset W. \end{aligned}$$

**Satz 17.3.** Kern und Bild einer linearen Abbildung sind Untervektorräume des Definitionsbereichs, bzw. des Bildraums.

*Beweis.*

Für den Kern haben wir das in Satz 10.5 gesehen und der Beweis ist geht in Fall von allgemeinen linearen Abbildungen genau so.

Sind  $\vec{w}^{(1)}, \vec{w}^{(2)} \in W$  im Bild von  $T$ , so gibt es  $\vec{v}^{(1)}, \vec{v}^{(2)} \in V$  mit  $T(\vec{v}^{(i)}) = \vec{w}^{(i)}$ . Dann ist aber auch  $T(\vec{v}^{(1)} + \vec{v}^{(2)}) = T(\vec{v}^{(1)}) + T(\vec{v}^{(2)}) = \vec{w}^{(1)} + \vec{w}^{(2)}$  und daher ist  $\vec{w}^{(1)} + \vec{w}^{(2)} \in \text{Bild}(T)$ . Ebenso ist  $T(\alpha\vec{v}^{(1)}) = \alpha T(\vec{v}^{(1)}) = \alpha\vec{w}^{(1)}$  und daher auch  $\alpha\vec{w}^{(1)} \in \text{Bild}(T)$  und wir haben gezeigt, dass  $\text{Bild}(T)$  ein Untervektorraum ist.  $\square$

Wir haben damit eine alternative Definition für den Rang, welche auch für abstrakte lineare Abbildungen funktioniert:

**Definition 17.4.** Der Rang einer linearen Abbildung  $T : V \rightarrow W$  ist  $\text{Rang}(T) := \dim(\text{Bild}(T))$ .

Sind  $V$  und  $W$  endlichdimensional, so ist jede lineare Abbildung durch Matrizen beschreibbar: Wir können dazu in  $V$  und  $W$  jeweils eine Basis wählen und bemerken, dass die Koeffizienten bezüglich dieser Basen durch eine bestimmte Matrix aufeinander abgebildet werden.

**Definition 17.5.** Es sei  $T : V \rightarrow W$  linear,  $B = \{ \vec{v}_1, \dots, \vec{v}_n \}$  eine Basis von  $V$  und  $C = \{ \vec{w}_1, \dots, \vec{w}_m \}$  eine Basis von  $W$ . Dann lässt sich jeder Vektor  $T(\vec{v}_j)$  in die Basis  $C$  entwickeln und wir nennen die Koeffizienten  $a_{ij}$ , d.h.

$$T(\vec{v}_j) = \sum_{i=1}^m a_{ij} \vec{w}_i.$$

Es reicht übrigens  $T(\alpha\vec{u} + \vec{v}) = \alpha T(\vec{u}) + T(\vec{v})$  zu fordern.

Der Kern ist also vollkommen analog wie für Matrizen definiert. Auch das Bild kann man für Matrizen genau so definieren.

Die Matrix  $A = (a_{ij})$  heißt *Darstellungsmatrix* (oder *Basisdarstellung*) von  $T$  bezüglich der Basen  $B$  und  $C$ .

**Satz 17.6.** Es seien  $V, W, T$  wie oben und  $A$  die Darstellungsmatrix von  $T$  bezüglich  $B$  und  $C$ . Dann gilt für  $\vec{w} = T(\vec{v})$ : Ist  $\vec{v} = \sum_{j=1}^n x_j \vec{v}_j$  und  $\vec{w} = \sum_{i=1}^m y_i \vec{w}_i$ , so gilt  $\vec{y} = A\vec{x}$ .

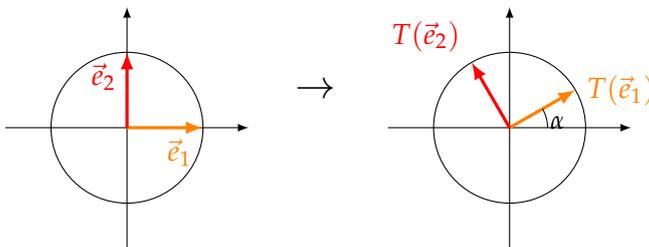
*Beweis.*

Wir benutzen die Linearität von  $T$

$$\begin{aligned} T(\vec{v}) &= T\left(\sum_{j=1}^n x_j \vec{v}_j\right) = \sum_{j=1}^n x_j T(\vec{v}_j) \\ &= \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} \vec{w}_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j\right) \vec{w}_i. \end{aligned}$$

Andererseits ist  $\vec{w} = T(\vec{v}) = \sum_{i=1}^m y_i \vec{w}_i$ . Die Behauptung folgt durch Vergleich der Koeffizienten.  $\square$

*Beispiel.* Wir betrachten  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  als eine Drehung der Ebene um den Ursprung und einen bestimmten Winkel  $\alpha$  und wollen diese Abbildung durch eine Matrix darstellen. Wir wählen nun im Definitionsbereich die Standard-Basis und bilden diese ab:



Geometrisch erkennen wir

$$T(\vec{e}_1) = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}, \quad T(\vec{e}_2) = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$$

Wir wählen nun im Bildraum ebenfalls die Standardbasis, d.h. wir bekommen

$$T(\vec{e}_1) = \cos(\alpha)\vec{e}_1 + \sin(\alpha)\vec{e}_2, \quad T(\vec{e}_2) = -\sin(\alpha)\vec{e}_1 + \cos(\alpha)\vec{e}_2.$$

Wir lesen jetzt die Darstellungsmatrix ab:

$$D_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Hier noch ein weiteres Beispiel:

*Beispiel.* Wir betrachten  $V = \Pi_2(\mathbb{R})$  und  $W = \Pi_1(\mathbb{R})$  und die lineare Abbildung definiert durch  $Tu = \frac{d}{dx}u = u'$ . Die Linearität

In Worten: Die Darstellungsmatrix bildet die Entwicklungskoeffizienten aufeinander ab.

der Abbildung  $T$  folgt aus den Ableitungsregeln:  $(\alpha u + v)' = \alpha u' + v'$ .

Als Basis von  $V$  wählen wir die Monome  $(\vec{v}_1, \vec{v}_2, \vec{v}_3) = (1, x, x^2)$ . Wir bilden diese mit  $T$  ab:

$$\begin{aligned} T(\vec{v}_1) &= \frac{d}{dx}(1) = 0, \\ T(\vec{v}_2) &= \frac{d}{dx}(x) = 1 = 1 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2, \\ T(\vec{v}_3) &= \frac{d}{dx}(x^2) = 2x = 0 \cdot \vec{v}_1 + 2 \cdot \vec{v}_2. \end{aligned}$$

In  $W$  wählen wir ebenfalls die Monom-Basis  $(\vec{w}_1, \vec{w}_2) = (1, x) = (\vec{v}_1, \vec{v}_2)$ . Damit können wir die Abbildungsmatrix von  $T$  ablesen:

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

Damit können wir tatsächlich Ableitungen berechnen: Wir betrachten  $p(x) = 4 + 5x + 7x^2$ . Dies entspricht in der Monom-Basis dem Vektor  $(4 \ 5 \ 7)^T$ . Wir wenden die Darstellungsmatrix an:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 14 \end{pmatrix}$$

und bekommen also  $T(p)(x) = 5 + 14x = p'(x)$ .

*Beispiel.* Betrachten wir nun die lineare Abbildung  $T(u) = \int_0^x u(t) dt$  als Abbildung von  $V = \Pi_1(\mathbb{R})$  nach  $W = \Pi_2(\mathbb{R})$ . Wir wählen wieder die Monom-Basen  $(\vec{v}_1, \vec{v}_2) = (1, t)$  in  $V$  und  $(\vec{w}_1, \vec{w}_2, \vec{w}_3) = (1, t, t^2)$  in  $W$ .

Es ist

$$\begin{aligned} T(\vec{v}_1) &= \int_0^1 1 dt = x = \vec{v}_2 \\ T(\vec{v}_2) &= \int_0^1 t dt = \frac{1}{2}x^2 = \frac{1}{2}\vec{w}_3. \end{aligned}$$

Für  $p(t) = a + bt = a\vec{v}_1 + b\vec{v}_2$  gilt  $T(p) = at + \frac{1}{2}bt^2 = a\vec{w}_2 + \frac{1}{2}b\vec{w}_3$ . Wir suchen also die Matrix  $A$  mit

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ a \\ \frac{1}{2}b \end{pmatrix}.$$

Man sieht schnell, dass dies durch die Matrix

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

geleistet wird. Damit können wir zum Beispiel  $p(t) = 7 - 5t$  integrieren, denn es ist

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 7 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 7 \\ -\frac{5}{2} \end{pmatrix}$$

und es stimmt  $\int_0^x p(t) dt = 7x - \frac{5}{2}x^2$ .

*Beispiel.* Eine lineare Abbildung  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  sei bzgl. der kanonischen Basen des  $\mathbb{R}^2$  und  $\mathbb{R}^3$  gegeben durch

$$T(\vec{x}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Wir wollen sie nun bezüglich der Basen

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \in \mathbb{R}^2 \text{ und} \\ \vec{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{w}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \vec{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Das Vorgehen ist wie folgt: Berechne die Bilder von  $\vec{v}_1$  und  $\vec{v}_2$  und stelle sie in der Basis  $(\vec{w}_1, \vec{w}_2, \vec{w}_3)$  dar:

$$\begin{aligned} T(\vec{v}_1) &= \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} \stackrel{!}{=} a_1 \vec{w}_1 + a_2 \vec{w}_2 + a_3 \vec{w}_3 \\ &= \begin{pmatrix} a_1 + a_2 + a_3 \\ a_2 + a_3 \\ a_3 \end{pmatrix}. \end{aligned}$$

Wir lösen durch Rückwärtssubstitution:  $a_3 = 9, a_2 = -2, a_1 = -2$ .  
Ebenso

$$\begin{aligned} T(\vec{v}_2) &= \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \stackrel{!}{=} a_1 \vec{w}_1 + a_2 \vec{w}_2 + a_3 \vec{w}_3 \\ &= \begin{pmatrix} a_1 + a_2 + a_3 \\ a_2 + a_3 \\ a_3 \end{pmatrix}. \end{aligned}$$

Wir bekommen  $a_3 = 3, a_2 = a_1 = 0$ . Die Darstellungsmatrix  $B$  bekommen wir nun, indem wir wie die Koeffizienten der Darstellung der Bilder in die Spalten schreiben:

$$B = \begin{pmatrix} -2 & 0 \\ -2 & 0 \\ 9 & 3 \end{pmatrix}.$$

Wir überprüfen das Ergebnis: Der Vektor  $\vec{v} = \vec{e}_1$  ist  $\vec{v} = \frac{1}{2}(\vec{v}_1 - \vec{v}_2)$ . Damit bekommen wir die Koordinaten von  $T(\vec{e}_1)$  bezüglich der Basis  $(\vec{w}_i)$  als

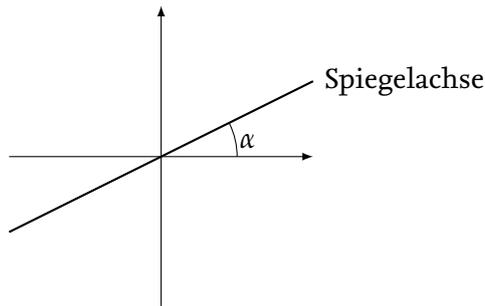
$$\vec{y} = \begin{pmatrix} -2 & 0 \\ -2 & 0 \\ 9 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix}.$$

Tatsächlich ist

$$T(\vec{e}_1) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = -1\vec{w}_1 - 1\vec{w}_2 + 3\vec{w}_3.$$

Hat man zwei Abbildungen  $T : V_1 \rightarrow V_2$  und  $S : V_2 \rightarrow V_3$  und Basen  $B_1, B_2, B_3$  in  $V_1, V_2$  bzw.  $V_3$  und die entsprechenden Darstellungsmatrizen  $A$  (für  $T$  bezüglich der Basen  $B_1$  und  $B_2$ ) und  $B$  (für  $S$  bezüglich  $B_1$  und  $B_2$ ), dann ist die Basisdarstellung von  $S \circ T$  bezüglich  $B_1$  und  $B_3$  gerade  $BA$ .

*Beispiel (Spiegelungen).* Es sei  $S_\alpha$  die lineare Abbildung des  $\mathbb{R}^2$  in sich selbst, welche an der folgenden Ursprungsgeraden spiegelt:



Wir suchen die Darstellungsmatrix dieser Abbildung bezüglich der Standardbasis. Diese Spiegelung ist die Komposition der drei Abbildungen

- Drehung um den Winkel  $-\alpha$
- Spiegelung an der  $x$ -Achse
- Drehung um den Winkel  $\alpha$ .

Die Spiegelung an der  $x$ -Achse bildet  $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$  auf  $\begin{pmatrix} v_1 \\ -v_2 \end{pmatrix}$  ab und hat daher die Darstellungsmatrix

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

Da die Drehung um  $\alpha$  die Abbildungsmatrix

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat, folgt (wegen  $\sin(-\alpha) = -\sin(\alpha)$  und  $\cos(-\alpha) = \cos(\alpha)$  und den Additionstheoremen für Sinus und Kosinus)

$$\begin{aligned} S_\alpha &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha)^2 - \sin(\alpha)^2 & 2\cos(\alpha)\sin(\alpha) \\ 2\cos(\alpha)\sin(\alpha) & -\cos(\alpha)^2 + \sin(\alpha)^2 \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}. \end{aligned}$$

## 18 Basiswechsel

Wir betrachten folgende Situation: Die Basisdarstellung  $A \in \mathbb{K}^{m \times n}$  einer Abbildung  $T : V \rightarrow W$  bezüglich zweier Basen  $(\vec{v}_1, \dots, \vec{v}_n)$  von  $V$  und  $(\vec{w}_1, \dots, \vec{w}_m)$  von  $W$  sei gegeben. Wir suchen nun die Basisdarstellung der gleichen Abbildung, aber bezüglich zweier anderer Basen  $(\vec{v}'_1, \dots, \vec{v}'_n)$  von  $V$  und  $(\vec{w}'_1, \dots, \vec{w}'_m)$  von  $W$ .

Dazu definieren wir  $S \in \mathbb{K}^{n \times n}$  für den *Basisübergang* von  $(\vec{v}'_1, \dots, \vec{v}'_n)$  zu  $(\vec{v}_1, \dots, \vec{v}_n)$  als

$$S := (s_{jk}), \quad \vec{v}'_k = \sum_{j=1}^n s_{jk} \vec{v}_j.$$

Die Matrix  $S$  nennen wir auch Basiswechselmatrix.

**Satz 18.1.** Die Matrix  $S$  ist regulär und die Inverse  $S^{-1}$  beschreibt den umgekehrten Basiswechsel von  $(\vec{v}_1, \dots, \vec{v}_n)$  zu  $(\vec{v}'_1, \dots, \vec{v}'_n)$ .

*Beweis.*

Es sei  $Q$  die Matrix des Basiswechsels von  $(\vec{v}_1, \dots, \vec{v}_n)$  zu  $(\vec{v}'_1, \dots, \vec{v}'_n)$ . Dann gilt nach Definition der Basiswechselmatrix

$$\vec{v}_j = \sum_{l=1}^n q_{lj} \vec{v}'_l.$$

Setzen wir dies in die Formel für die Koeffizienten  $s_{jk}$  ein, bekommen wir

$$\vec{v}'_k = \sum_{j=1}^n s_{jk} \sum_{l=1}^n q_{lj} \vec{v}'_l = \sum_{l=1}^n \left( \sum_{j=1}^n q_{lj} s_{jk} \right) \vec{v}'_l.$$

Daraus folgt per Koeffizientenvergleich

$$\sum_{j=1}^n q_{lj} s_{jk} = \delta_{lk} := \begin{cases} 0 & : l \neq k \\ 1 & : l = k \end{cases},$$

also gilt  $QS = I_n$ , woraus  $S^{-1} = Q$  folgt.  $\square$

Dieses  $\delta_{lk}$  nennt man *Kronecker-Delta*. Damit kann man z.B. schreiben  $I_n = (\delta_{jk})_{jk}$ .

Analog soll die Basiswechselmatrix für den Wechsel von der Basis  $(\vec{w}'_1, \dots, \vec{w}'_m)$  zu  $(\vec{w}_1, \dots, \vec{w}_m)$  in  $W$  durch  $R$  gegeben sein.

Um nun  $A$  bezüglich der neuen Basis darzustellen betrachten wir folgendes Schema:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \updownarrow & & \updownarrow \\ (\vec{v}_1, \dots, \vec{v}_n) & \xrightarrow{A} & (\vec{w}_1, \dots, \vec{w}_m) \\ S \uparrow & & \uparrow R \\ (\vec{v}'_1, \dots, \vec{v}'_n) & \xrightarrow{B} & (\vec{w}'_1, \dots, \vec{w}'_m) \end{array}$$

Die Matrizen  $A, B, S, R$  sind dabei festgelegt durch

$$\begin{aligned} S : \vec{v}'_k &= \sum_j s_{jk} \vec{v}_j & A : T(\vec{v}_j) &= \sum_i a_{ij} \vec{w}_i \\ R : \vec{w}'_l &= \sum_i r_{il} \vec{w}_i & B : T(\vec{v}'_k) &= \sum_l b_{lk} \vec{w}'_l. \end{aligned}$$

Dies setzen wir ineinander ein und bekommen

$$T(\vec{v}'_k) = T\left(\sum_j s_{jk} \vec{v}_j\right) = \sum_j s_{jk} T(\vec{v}_j) = \sum_j s_{jk} \sum_i a_{ij} \vec{w}_i = \sum_i \sum_j s_{jk} a_{ij} \vec{w}_i$$

und

$$T(\vec{v}'_k) = \sum_l b_{lk} \vec{w}'_l = \sum_l b_{lk} \sum_i r_{il} \vec{w}_i = \sum_i \sum_l b_{lk} r_{il} \vec{w}_i.$$

Durch Vergleich der Koeffizienten sehen wir, dass für alle  $i, k$  gelten muss

$$\sum_j a_{ij} s_{jk} = \sum_l r_{il} b_{lk}.$$

Das bedeutet aber nichts anderes, als dass

$$AS = RB.$$

Wir haben also folgenden Satz bewiesen:

**Satz 18.2.** Die Basisdarstellung  $B$  einer linearen Abbildung  $T : V \rightarrow W$  bezüglich der Basen  $(\vec{v}'_k)$  und  $(\vec{w}'_l)$  erhält man aus der Darstellung  $A$  bezüglich der Basen  $(\vec{v}_j)$  und  $(\vec{w}_i)$  durch

$$B = R^{-1}AS$$

mit  $R$  und  $S$  wie oben definiert.

**Definition 18.3.** Zwei Matrizen  $A, B \in \mathbb{K}^{m \times n}$  heißen *äquivalent*, wenn es reguläre Matrizen  $S \in \mathbb{K}^{n \times n}$  und  $R \in \mathbb{K}^{m \times m}$  gibt, so dass

$$B = R^{-1}AS$$

gilt.

*Beispiel.* Schauen wir uns wieder das Beispiel an, in dem  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  bezüglich der Standardbasen  $(\vec{e}_1, \vec{e}_2)$  und bzw.  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$  dargestellt ist als

$$T(\vec{x}) = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\vec{x}.$$

Also heißen zwei Matrizen äquivalent, wenn die im Prinzip die gleiche lineare Abbildung darstellen, nur bezüglich anderer Basen.

Berechnen wir noch einmal die Darstellungsmatrix  $B$  von  $T$  bezüglich der Basen

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \in \mathbb{R}^2 \text{ und}$$

$$\vec{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{w}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \vec{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Für den Basiswechsel  $(\vec{v}_k) \rightarrow (\vec{e}_j)$  betrachten wir

$$\vec{v}_k = \sum_{j=1}^2 s_{jk} \vec{e}_j, \quad k = 1, 2,$$

d.h.  $\vec{v}_1 = s_{11} \vec{e}_1 + s_{21} \vec{e}_2$ , also

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = s_{11} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + s_{21} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix},$$

d.h.  $s_{11} = 1$  und  $s_{21} = 1$ . Ebenso  $\vec{v}_2 = s_{12} \vec{e}_1 + s_{22} \vec{e}_2$ , also

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} = s_{12} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + s_{22} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} s_{12} \\ s_{22} \end{pmatrix},$$

d.h.  $s_{12} = -1$  und  $s_{22} = 1$ . Damit haben wir

$$S = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Für den Basiswechsel  $(\vec{w}_l) \rightarrow (\vec{e}_i)$  betrachten wir

$$\vec{w}_l = \sum_{j=1}^3 r_{jl} \vec{e}_j,$$

also

$$\vec{w}_1 = r_{11} \vec{e}_1 + r_{21} \vec{e}_2 + r_{31} \vec{e}_3,$$

d.h.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} r_{11} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ r_{21} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ r_{31} \end{pmatrix} = \begin{pmatrix} r_{11} \\ r_{21} \\ r_{31} \end{pmatrix}$$

und damit  $r_{11} = 1, r_{21} = 0, r_{31} = 0$ . Analog

$$\vec{w}_2 = r_{12} \vec{e}_1 + r_{22} \vec{e}_2 + r_{32} \vec{e}_3,$$

d.h.

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} r_{12} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ r_{22} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ r_{32} \end{pmatrix} = \begin{pmatrix} r_{12} \\ r_{22} \\ r_{32} \end{pmatrix}$$

und damit  $r_{12} = 1, r_{22} = 1, r_{32} = 0$  und

$$\vec{w}_3 = r_{13}\vec{e}_1 + r_{23}\vec{e}_2 + r_{33}\vec{e}_3,$$

d.h.

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} r_{13} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ r_{23} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ r_{33} \end{pmatrix} = \begin{pmatrix} r_{13} \\ r_{23} \\ r_{33} \end{pmatrix}$$

und damit  $r_{13} = 1, r_{23} = 1, r_{33} = 1$ . Wir bekommen

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wir berechnen die Inverse von  $R$  mit dem Eliminationsalgorithmus:

$$\begin{aligned} & \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ \Leftrightarrow & \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ \Leftrightarrow & \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right), \end{aligned}$$

also ist

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

Damit ist unsere Basiswechselmatrix

$$\begin{aligned} B = R^{-1}AS &= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 7 & 3 \\ 9 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ -2 & 0 \\ 9 & 3 \end{pmatrix}. \end{aligned}$$

Hier stellt sich eine naheliegende Frage: Ist eine lineare Abbildung  $T : V \rightarrow W$  gegeben, wie ist eine möglichst einfache Darstellung dieser Matrix bezüglich zweier Basen? Probleme dieser Art heißen *Normalformenprobleme*. Für Matrizen lautet ein solches Problem zum Beispiel:

Zu einer gegebenen Matrix  $A \in \mathbb{K}^{m \times n}$  finde man eine möglichst einfache äquivalente Matrix.

Dieses Normalformenproblem ist relativ einfach zu lösen:

**Satz 18.4.** Zwei Matrizen  $A, B \in \mathbb{K}^{m \times n}$  sind genau dann äquivalent, wenn sie den gleichen Rang haben. Hat  $A$  den Rang  $r$  so ist  $A$  äquivalent zur Matrix

$$D_r := \begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix}$$

*Beweis.*

Hat  $A$  den Rang  $r$ , hat der Kern von  $A$  nach der Dimensionsformel (Satz 12.3) die Dimension  $n - r$ . Wir wählen eine Basis  $\vec{v}_1, \dots, \vec{v}_n$  des  $\mathbb{K}^n$  so, dass die letzten  $r$  Vektoren  $\vec{v}_{r+1}, \dots, \vec{v}_n$  eine Basis von  $\text{Kern}(A)$  bilden. Nach Konstruktion gilt  $A\vec{v}_k = \vec{0}$  für  $k = r + 1, \dots, n$ . Wir definieren  $\vec{w}_k = A\vec{v}_k$  für  $k = 1, \dots, r$ . Diese  $\vec{w}_k$  sind linear unabhängig und bilden eine Basis von  $\text{Bild}(A)$ , welche wir beliebig zu einer Basis  $(\vec{w}_k)_{k=1, \dots, m}$  von  $\mathbb{K}^m$  ergänzen. Damit gilt  $A\vec{v}_k = \vec{w}_k$  für  $k = 1, \dots, r$  und  $A\vec{v}_k = \vec{0}$  für  $k = r + 1, \dots, n$ , was die Behauptung zeigt.  $\square$

Das Normalformenproblem ist wesentlich schwieriger, wenn wir quadratische Matrizen  $A \in \mathbb{K}^{n \times n}$  (bzw. lineare Selbstabbildungen  $T : V \rightarrow V$ ) betrachten und zusätzlich fordern, dass wir in Definitions- und Bildbereich die gleiche Basis wählen. In diesem Fall ist die Darstellungsmatrix nach Basiswechsel von der Form

$$B = S^{-1}AS.$$

**Definition 18.5.** Zwei Matrizen  $A, B \in \mathbb{K}^{n \times n}$  heißen *ähnlich*, wenn es eine invertierbare Matrix  $S \in \mathbb{K}^{n \times n}$  gibt, so dass  $B = S^{-1}AS$  gilt.

Auf dieses Normalformenproblem kommen wir zurück, wenn wir Eigenwerte und -vektoren von Matrizen thematisieren.

## 19 Orthogonalität

In diesem Abschnitt beschränken wir uns auf reelle euklidische Vektorräume  $V$  mit Skalarprodukt  $\langle \cdot, \cdot \rangle$  und der dadurch induzierten Norm  $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$ .

**Definition 19.1.** Eine Menge von Vektoren  $\{\vec{w}_1, \dots, \vec{w}_m\}$  bilden ein

- (a) *Orthogonalsystem*, falls sie paarweise aufeinander senkrecht stehen, d.h. falls

$$\langle \vec{w}_i, \vec{w}_j \rangle = \begin{cases} 0 & : i \neq j \\ > 0 & : i = j \end{cases}$$

für alle  $i, j$  gilt.

- (b) *Orthonormalsystem*, wenn sie ein Orthogonalsystem bilden und normiert sind, d.h. falls

$$\langle \vec{w}_i, \vec{w}_j \rangle = \delta_{ij} = \begin{cases} 0 & : i \neq j \\ 1 & : i = j \end{cases}$$

für alle  $i, j$  gilt.

**Satz 19.2.** Es sei  $(V, \langle \cdot, \cdot \rangle)$  ein  $n$ -dimensionaler, reeller euklidischer Vektorraum. Dann gilt:

- (a) Jedes Orthogonalsystem ist linear unabhängig und daher ist jedes Orthogonalsystem aus  $n$  Vektoren eine Basis von  $V$  (also eine Orthogonalbasis). Entsprechend ist jedes Orthonormalsystem aus  $n$  Vektoren eine Orthonormalbasis.
- (b) Ist  $(\vec{w}_1, \dots, \vec{w}_n)$  eine Orthonormalbasis, so hat jeder Vektor  $\vec{x} \in V$  die Fourier-Entwicklung

$$\vec{x} = \sum_{i=1}^n \langle \vec{w}_i, \vec{x} \rangle \vec{w}_i.$$

*Beweis.*

- (a) Ist  $\sum_{i=1}^m \alpha_i \vec{w}_i = \vec{0}$ , so bilden wir das Skalarprodukt der linken Seite mit  $\vec{w}_k$  und bekommen

$$0 = \left\langle \sum_{i=1}^m \alpha_i \vec{w}_i, \vec{w}_k \right\rangle = \sum_{i=1}^m \alpha_i \langle \vec{w}_i, \vec{w}_k \rangle = \alpha_k \langle \vec{w}_k, \vec{w}_k \rangle$$

woraus  $\alpha_k = 0$  folgt. Da dies für  $k = 1, \dots, m$  funktioniert, haben wir die lineare Unabhängigkeit gezeigt. Die weiteren Behauptungen sind direkte Konsequenzen.

- (b) Ist  $(\vec{w}_1, \dots, \vec{w}_n)$  eine Basis, so hat jedes  $\vec{x}$  eine eindeutige Darstellung  $\vec{x} = \sum_{i=1}^n \beta_i \vec{w}_i$  und wir müssen nur die Koeffizienten  $\beta_i$  bestimmen. Auch hier bilden wir das Skalarprodukt mit  $\vec{w}_k$

und bekommen

$$\langle \vec{w}_k, \vec{x} \rangle = \langle \vec{w}_k, \sum_{i=1}^n \beta_i \vec{w}_i \rangle = \sum_{i=1}^n \beta_i \langle \vec{w}_k, \vec{w}_i \rangle = \sum_{i=1}^n \beta_i \delta_{ik} = \beta_k.$$

□

Ist  $V = \mathbb{R}^n$  mit dem Standard-Skalarprodukt  $\langle \vec{v}, \vec{w} \rangle = \vec{v}^T \vec{w}$ , so ist die Fourier-Entwicklung

$$\vec{x} = \sum_{i=1}^n (\vec{w}_i^T \vec{x}) \vec{w}_i = \sum_{i=1}^n \vec{w}_i (\vec{w}_i^T \vec{x}) = \left( \sum_{i=1}^n \vec{w}_i \vec{w}_i^T \right) \vec{x}.$$

Es folgt also für Orthonormalbasen  $(\vec{w}_k)$  die Gleichung

$$I_n = \sum_{i=1}^n \vec{w}_i \vec{w}_i^T.$$

Kommen wir nun zur Konstruktion von Orthonormalbasen. Hierfür gibt das Orthogonalisierungsverfahren nach Erhard Schmidt (auch Gram-Schmidt-Verfahren genannt):

**Gegeben:** Vektorraum  $W$  mit Basis  $(\vec{v}_1, \dots, \vec{v}_m)$ .

**Gesucht:** Eine Orthogonalbasis  $(\vec{w}_1, \dots, \vec{w}_m)$  von  $W$ .

**Idee:** Konstruiere die  $\vec{w}_j$  rekursiv aus den  $\vec{v}_k$ .

Wir starten dabei einfach mit  $\vec{w}_1 = \vec{v}_1$  und konstruieren nun  $\vec{w}_2$  indem wir von  $\vec{v}_2$  ein gewisses Vielfaches von  $\vec{w}_1$  abziehen, d.h. wir machen den Ansatz

$$\vec{w}_2 = \vec{v}_2 - \alpha \vec{w}_1.$$

Wir bestimmen  $\alpha$  so, dass  $\vec{w}_1$  und  $\vec{w}_2$  orthogonal sind:

$$\begin{aligned} 0 &= \langle \vec{w}_1, \vec{w}_2 \rangle = \langle \vec{w}_1, \vec{v}_2 - \alpha \vec{w}_1 \rangle \\ &= \langle \vec{w}_1, \vec{v}_2 \rangle - \alpha \langle \vec{w}_1, \vec{w}_1 \rangle. \end{aligned}$$

Dies führt auf

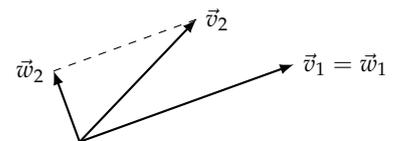
$$\alpha = \frac{\langle \vec{w}_1, \vec{v}_2 \rangle}{\|\vec{w}_1\|^2}$$

und damit auf

$$\vec{w}_2 = \vec{v}_2 - \frac{\langle \vec{w}_1, \vec{v}_2 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1.$$

Wir beschreiben noch den nächsten Schritt und dann wird das Prinzip klar: Um  $\vec{w}_3$  zu konstruieren, ziehen wir eine Linearkombination von  $\vec{w}_1$  und  $\vec{w}_2$  von  $\vec{v}_3$  ab, d.h. wir machen den Ansatz

$$\vec{w}_3 = \vec{v}_3 - \sum_{i=1}^2 \alpha_i \vec{w}_i = \vec{v}_3 - \alpha_1 \vec{w}_1 - \alpha_2 \vec{w}_2$$



und bestimmen  $\alpha_1, \alpha_2$  so, dass  $\vec{w}_3$  senkrecht auf  $\vec{w}_1$  und  $\vec{w}_2$  steht. Dies sind die beiden Gleichungen

$$0 = \langle \vec{w}_3, \vec{w}_1 \rangle = \langle \vec{v}_3 - \alpha_1 \vec{w}_1 - \alpha_2 \vec{w}_2, \vec{w}_1 \rangle = \langle \vec{v}_3, \vec{w}_1 \rangle - \alpha_1 \langle \vec{w}_1, \vec{w}_1 \rangle - \alpha_2 \underbrace{\langle \vec{w}_2, \vec{w}_1 \rangle}_{=0}$$

$$0 = \langle \vec{w}_3, \vec{w}_2 \rangle = \langle \vec{v}_3 - \alpha_1 \vec{w}_1 - \alpha_2 \vec{w}_2, \vec{w}_2 \rangle = \langle \vec{v}_3, \vec{w}_2 \rangle - \alpha_1 \underbrace{\langle \vec{w}_1, \vec{w}_2 \rangle}_{=0} - \alpha_2 \langle \vec{w}_2, \vec{w}_2 \rangle$$

wobei wir ausgenutzt haben, dass  $\vec{w}_1$  und  $\vec{w}_2$  bereits orthogonal sind. Wir bekommen die Koeffizienten einfach durch

$$\alpha_1 = \frac{\langle \vec{v}_3, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \quad \alpha_2 = \frac{\langle \vec{v}_3, \vec{w}_2 \rangle}{\|\vec{w}_2\|^2}.$$

Für den allgemeinen Schritt, nehmen wir an, dass wir  $\vec{w}_1, \dots, \vec{w}_{j-1}$  schon orthogonal bestimmt haben und machen den Ansatz

$$\vec{w}_j = \vec{v}_j - \sum_{i=1}^{j-1} \alpha_i \vec{w}_i.$$

Genau wie eben bekommen wir die Koeffizienten  $\alpha_i$  aus den Forderungen  $0 = \langle \vec{w}_j, \vec{w}_k \rangle, k = 1, \dots, j-1$  als

$$\alpha_k = \frac{\langle \vec{v}_j, \vec{w}_k \rangle}{\|\vec{w}_k\|^2}.$$

Wir halten das Vorgehen in einem Algorithmus fest:

**Eingabe:** Linear unabhängige Vektoren  $\vec{v}_1, \dots, \vec{v}_m$

**Ausgabe:** Orthogonalbasis  $\{\vec{w}_1, \dots, \vec{w}_m\}$  von  $\text{Spann}\{\vec{v}_1, \dots, \vec{v}_m\}$

```

 $\vec{w}_1 = \vec{v}_1$ 
for  $k = 2, \dots, m$  do
   $\vec{w}_k = \vec{v}_k - \sum_{i=1}^{k-1} \frac{\langle \vec{v}_k, \vec{w}_i \rangle}{\|\vec{w}_i\|^2} \vec{w}_i$ 
end for

```

Durch Normalisierung der Vektoren erhalten wir ebenso einfach eine Methode zur Konstruktion von Orthonormalbasen:

**Eingabe:** Linear unabhängige Vektoren  $\vec{v}_1, \dots, \vec{v}_m$

**Ausgabe:** Orthonormalbasis  $\{\vec{w}_1, \dots, \vec{w}_m\}$  von  $\text{Spann}\{\vec{v}_1, \dots, \vec{v}_m\}$

```

 $\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$ 
for  $k = 2, \dots, m$  do
   $\vec{u}_k = \vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{v}_k, \vec{w}_i \rangle \vec{w}_i$ 
   $\vec{w}_k = \frac{\vec{u}_k}{\|\vec{u}_k\|}$ 
end for

```

Wir haben das Verfahren in abstrakten Vektorräumen eingeführt und können es auch dort benutzen:

**Beispiel.** Wir orthogonalisieren die Monombasis  $(1, x, x^2, x^3)$  des Vektorraums der Polynome vom Grad  $\leq 3$  auf dem Intervall  $[-1, 1]$  bezüglich des Skalarproduktes

$$\langle f, g \rangle := \int_{-1}^1 f(x)g(x)dx$$

mit induzierter Norm  $\|f\| = \sqrt{\int_{-1}^1 f(x)^2 dx}$ . Wir setzen  $\vec{v}_1 := 1$ ,  $\vec{v}_2 := x$ ,  $\vec{v}_3 := x^2$  und  $\vec{v}_4 := x^3$ . Diese Monome sind nicht paarweise orthogonal, denn z.B. gilt

$$\langle \vec{v}_2, \vec{v}_4 \rangle = \int_{-1}^1 x \cdot x^3 dx = \int_{-1}^1 x^4 dx = \frac{1}{5}x^5 \Big|_{x=-1}^1 = \frac{2}{5}.$$

Wenden wir also den Orthogonalisierungsalgorithmus an:

- $\vec{w}_1 = \vec{v}_1 = 1$ .
- $\vec{w}_2 = \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1$  und wegen

$$\langle \vec{v}_2, \vec{w}_1 \rangle = \int_{-1}^1 x \cdot \frac{1}{2} dx = \frac{1}{4}x^2 \Big|_{x=-1}^1 = 0$$

ist  $\vec{w}_2 = x$ .

- $\vec{w}_3 = \vec{v}_3 - \frac{\langle \vec{v}_3, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 - \frac{\langle \vec{v}_3, \vec{w}_2 \rangle}{\|\vec{w}_2\|^2} \vec{w}_2$ . Wir berechnen

$$\langle \vec{v}_3, \vec{w}_1 \rangle = \int_{-1}^1 x^2 \cdot 1 dx = \frac{1}{3}x^3 \Big|_{x=-1}^1 = \frac{2}{3},$$

$$\langle \vec{v}_3, \vec{w}_2 \rangle = \int_{-1}^1 x^2 \cdot x dx = \frac{1}{4}x^4 \Big|_{x=-1}^1 = 0,$$

$$\|\vec{w}_1\|^2 = \int_{-1}^1 1 \cdot 1 dx = 2$$

und bekommen  $\vec{w}_3 = x^2 - \frac{1}{3}$ .

- $\vec{w}_4 = \vec{v}_4 - \frac{\langle \vec{v}_4, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 - \frac{\langle \vec{v}_4, \vec{w}_2 \rangle}{\|\vec{w}_2\|^2} \vec{w}_2 - \frac{\langle \vec{v}_4, \vec{w}_3 \rangle}{\|\vec{w}_3\|^2} \vec{w}_3$ . Wir berechnen

$$\langle \vec{v}_4, \vec{w}_1 \rangle = \int_{-1}^1 x^3 \cdot 1 dx = \frac{1}{4}x^4 \Big|_{x=-1}^1 = 0,$$

$$\langle \vec{v}_4, \vec{w}_2 \rangle = \int_{-1}^1 x^3 \cdot x dx = \frac{1}{5}x^5 \Big|_{x=-1}^1 = \frac{2}{5},$$

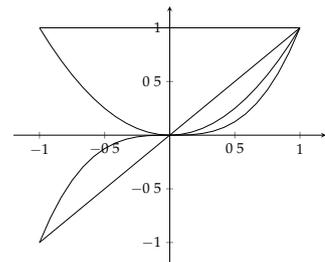
$$\langle \vec{v}_4, \vec{w}_3 \rangle = \int_{-1}^1 x^3 \cdot (x^2 - \frac{1}{3}) dx = \int_{-1}^1 x^5 - \frac{1}{3}x^3 dx = \frac{1}{6}x^6 - \frac{1}{12}x^4 \Big|_{x=-1}^1 = 0,$$

$$\|\vec{w}_2\|^2 = \int_{-1}^1 x \cdot x dx = \int_{-1}^1 x^2 dx = \frac{1}{3}x^3 \Big|_{x=-1}^1 = \frac{2}{3}$$

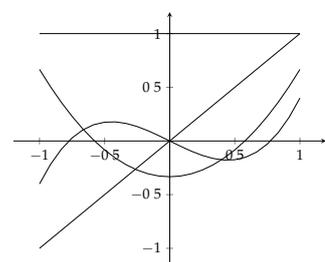
und bekommen  $\vec{w}_4 = x^3 - \frac{3}{5}x$ . Die Vektoren

$$\vec{w}_1 = 1, \quad \vec{w}_2 = x, \quad \vec{w}_3 = x^2 - \frac{1}{3}, \quad \vec{w}_4 = x^3 - \frac{3}{5}x$$

Hier die vier Funktionen der Monombasis:



Hier die vier Funktionen der orthogonalisierten Basis:



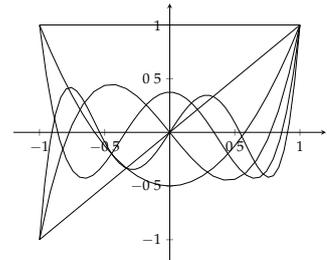
sind also orthogonal bezüglich des Skalarproduktes, was wir gewählt haben. Man rechnet noch aus, dass gilt

$$\|\vec{w}_1\|^2 = 2, \quad \|\vec{w}_2\|^2 = \frac{2}{3}, \quad \|\vec{w}_3\|^2 = \frac{8}{45}, \quad \|\vec{w}_4\|^2 = \frac{8}{175}.$$

Damit bekommt man die Orthonormalbasis

$$\frac{1}{\sqrt{2}}, \quad \sqrt{\frac{3}{2}}x, \quad \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right), \quad \sqrt{\frac{175}{8}}\left(x^3 - \frac{3}{5}x\right).$$

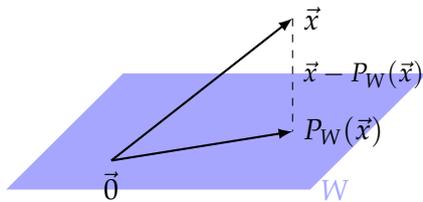
Dies sind (bis auf konstante Faktoren) die Legendre-Polynome. Die Legendre-Polynome werden üblicherweise so normiert, dass die an der Stelle 1 alle den Wert 1 haben. Hier die ersten sechs Legendre-Polynome:



Legt man andere Grundintervalle als  $[-1, 1]$  zu Grunde und nimmt man andere Skalarprodukte, z.B. gewichtete Produkte wie  $\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2/2}dx$  so bekommt man durch Orthogonalisierung der Monome andere Familien von orthogonalen Polynomen. Orthogonale Polynome spielen eine große Rolle in der numerischen Mathematik und der Optimierung.

## 20 Orthogonale Projektionen

Es sei  $W$  ein  $m$ -dimensionaler Unterraum eines euklidischen Vektorraums  $V$ . Wir wollen nun den Begriff der „orthogonalen Projektion“ auf  $W$  herleiten. Eine orthogonale Projektion auf einen Unterraum ist eine Abbildung  $P_W : V \rightarrow W$  mit der Eigenschaft, dass  $P_W(\vec{x})$  der „Fußpunkt des Lots von  $\vec{x}$  auf  $W$ “ ist. Geometrisch ist die Situation wie folgt:



Der Punkt  $P_W(\vec{x})$  erfüllt also insbesondere die Bedingung  $\langle P_W(\vec{x}), \vec{x} - P_W(\vec{x}) \rangle = 0$ . Wir sehen sogar, dass die Verbindungslinie  $\vec{x} - P_W(\vec{x})$  senkrecht auf ganz  $W$  steht, d.h. wir haben für jedes  $\vec{w} \in W$

$$\langle \vec{w}, \vec{x} - P_W(\vec{x}) \rangle = 0. \quad (\perp)$$

Die geometrische Situation lässt und folgendes vermuten:

Für gegebenes  $\vec{x} \in V$  suchen wir ein  $\vec{w} = P_W(\vec{x}) \in W$ , welches am nächsten an  $\vec{x}$  liegt, d.h. für welches der Wert  $\|\vec{x} - \vec{w}\|$  so klein wie möglich ist. Man nennt daher  $P_W(\vec{x})$  auch *Bestapproximation* von  $\vec{x}$  in  $W$ .

(Wir werden die im nächsten Abschnitt zeigen.)

Wie können wir  $P_W(\vec{x})$  berechnen? Sei dazu  $(\vec{v}_1, \dots, \vec{v}_m)$  eine Basis von  $W$ . Dann haben wir die Darstellung

$$P_W(\vec{x}) = \sum_{i=1}^m \alpha_i \vec{v}_i,$$

und suchen die Koeffizienten  $(\alpha_1, \dots, \alpha_m)$ . Da  $(\perp)$  genau das für alle  $w \in W$  gilt, wenn es für jeden Basisvektor  $\vec{v}_k$  gilt, betrachten wir für alle  $k$

$$\begin{aligned} 0 &= \langle \vec{v}_k, \vec{x} - \sum_{i=1}^m \alpha_i \vec{v}_i \rangle \\ &= \langle \vec{v}_k, \vec{x} \rangle - \sum_{i=1}^m \langle \vec{v}_k, \vec{v}_i \rangle \alpha_i. \end{aligned}$$

Diese Gleichungen sind äquivalent zum Gleichungssystem

$$\sum_{i=1}^m \langle \vec{v}_k, \vec{v}_i \rangle \alpha_i = \langle \vec{v}_k, \vec{x} \rangle, \quad k = 1, \dots, m,$$

welches wir mit Matrizen schreiben können:

$$\begin{pmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \cdots & \langle \vec{v}_1, \vec{v}_m \rangle \\ \vdots & & \vdots \\ \langle \vec{v}_m, \vec{v}_1 \rangle & \cdots & \langle \vec{v}_m, \vec{v}_m \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \langle \vec{v}_1, \vec{x} \rangle \\ \vdots \\ \langle \vec{v}_m, \vec{x} \rangle \end{pmatrix}. \quad (\text{G})$$

Die Matrix auf der linken Seite bezeichnet man auch mit  $G(\vec{v}_1, \dots, \vec{v}_m)$  und sie wird *Gramsche Matrix* der Basis  $(\vec{v}_1, \dots, \vec{v}_m)$  genannt.

Ist die Basis sogar eine Orthonormalbasis  $(\vec{w}_1, \dots, \vec{w}_m)$ , so ist die Gramsche Matrix wegen  $\langle \vec{w}_i, \vec{w}_j \rangle = \delta_{ij}$  einfach

$$G(\vec{w}_1, \dots, \vec{w}_m) = I_m$$

und wir haben für die Koeffizienten

$$\alpha_i = \langle \vec{w}_i, \vec{x} \rangle$$

und die orthogonale Projektion ist

$$P_W(\vec{x}) = \sum_{i=1}^m \langle \vec{w}_i, \vec{x} \rangle \vec{w}_i.$$

**Satz 20.1** (Projektionssatz). *Sei  $V$  ein euklidischer Vektorraum  $V$  und  $W$  ein endlichdimensionaler Unterraum von  $V$ . Dann existiert zu jedem  $\vec{x} \in V$  eine eindeutige Zerlegung  $\vec{x} = \vec{x}_1 + \vec{y}$  mit  $\vec{x}_1 \in W$  und  $\vec{y} \perp W$ .*

*Beweis.*

Die Existenz der Zerlegung haben wir schon gezeigt: Wir setzen einfach  $\vec{x}_1 = P_W(\vec{x})$  und  $\vec{y} = \vec{x} - P_W(\vec{x})$ .

Zur Eindeutigkeit: Wir nehmen an, es gäbe zwei solche Zerlegungen:

$$\vec{x} = \vec{x}_1 + \vec{y} = \vec{\tilde{x}}_1 + \vec{\tilde{y}}$$

$\vec{x}_1, \vec{\tilde{x}}_1 \in W, \vec{y}, \vec{\tilde{y}} \perp W$ . Dann gilt  $\vec{x}_1 - \vec{\tilde{x}}_1 = \vec{\tilde{y}} - \vec{y}$  und  $\vec{x}_1 - \vec{\tilde{x}}_1 \in W$  und  $\vec{\tilde{y}} - \vec{y} \perp W$ . Dies geht aber nur für  $\vec{x}_1 - \vec{\tilde{x}}_1 = \vec{\tilde{y}} - \vec{y} = \vec{0}$ .  $\square$

**Korollar 20.2.** *Die orthogonale Projektion  $P_W(\vec{x})$  ist unabhängig von der Wahl der Basis  $(\vec{v}_1, \dots, \vec{v}_m)$  eindeutig bestimmt durch die Bedingungen*

$$P_W(\vec{x}) \in W, \quad \vec{x} - P_W(\vec{x}) \perp W.$$

*Daher hat das lineare Gleichungssystem (G) stets eine eindeutige Lösung und die Gramsche Matrix ist stets invertierbar.*

Wir fassen noch einmal zusammen:

**Satz 20.3.** *Ist  $W$  ein endlichdimensionaler Unterraum eines euklidischen Vektorraumes  $V$  mit Basis  $(\vec{v}_1, \dots, \vec{v}_m)$ , dann ist die orthogonale Projektion von  $\vec{x} \in V$  auf  $W$  gegeben durch*

$$P_W(\vec{x}) = \sum_{i=1}^m \alpha_i \vec{v}_i,$$

Erweitert man im Fall eines endlichdimensionalen Raums  $V$  die ONB  $(\vec{w}_1, \dots, \vec{w}_m)$  von  $W$  zu einer ONB  $(\vec{w}_1, \dots, \vec{w}_n)$  von ganz  $V$ , so ist die Fourier-Darstellung von  $\vec{x}$

$$\vec{x} = \sum_{i=1}^n \langle \vec{w}_i, \vec{x} \rangle \vec{w}_i$$

und die Projektion

$$P_W(\vec{x}) = \sum_{i=1}^m \langle \vec{w}_i, \vec{x} \rangle \vec{w}_i$$

ist einfach die abgeschnittene Fourier-Darstellung.

wobei  $\vec{\alpha} \in \mathbb{R}^m$  die eindeutige Lösung des Gleichungssystems (G) ist.

Ist  $(\vec{w}_1, \dots, \vec{w}_m)$  ein Orthonormalsystem, so ist die orthogonale Projektion von  $\vec{x}$  auf  $W = \text{Spann}(\vec{w}_1, \dots, \vec{w}_m)$  gegeben durch

$$P_W(\vec{x}) = \sum_{i=1}^m \langle \vec{w}_i, \vec{x} \rangle \vec{w}_i.$$

Schauen wir uns den Fall  $V = \mathbb{R}^n$  mit dem Standard-Skalarprodukt genauer an: Hier können wir die Basisvektoren zu einer Matrix  $A = (\vec{v}_1, \dots, \vec{v}_m) \in \mathbb{R}^{n \times m}$  zusammenfassen. Die Gramsche Matrix ist dann  $G(\vec{v}_1, \dots, \vec{v}_m) = A^T A$  und das Gleichungssystem ist  $A^T A \vec{\alpha} = A^T \vec{x}$ , wobei  $\vec{\alpha}$  der Vektor mit den Einträgen  $\alpha_1, \dots, \alpha_m$  ist. Wir sehen insbesondere, dass die Gramsche Matrix  $G = A^T A$  stets invertierbar ist, wenn die Spalten von  $A$  linear unabhängig sind.

Haben wir ein Orthonormalsystem  $(\vec{w}_1, \dots, \vec{w}_m)$ , so ist die orthogonale Projektion auf  $W := \text{Spann}(\vec{w}_1, \dots, \vec{w}_m)$

$$\begin{aligned} P_W(\vec{x}) &= \sum_{i=1}^m \langle \vec{w}_i, \vec{x} \rangle \vec{w}_i \\ &= \sum_{i=1}^m (\vec{w}_i^T \vec{x}) \vec{w}_i \\ &= \left( \sum_{i=1}^m \vec{w}_i \vec{w}_i^T \right) \vec{x}. \end{aligned}$$

Wir haben also die Matrix-Darstellung der orthogonalen Projektion auf  $\text{Spann}(\vec{w}_1, \dots, \vec{w}_m)$  bezüglich der Standard-Basen

$$P_W = \sum_{i=1}^m \vec{w}_i \vec{w}_i^T.$$

*Beispiel.* Wir wollen die Orthogonalprojektion auf die durch die Gleichung  $3x_1 - 2x_2 + 5x_3 = 0$  gegebene Ebene  $E$  berechnen. Der Vektor

$$\vec{v}_3 = \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix}$$

steht senkrecht auf  $E$  und wir können diesen leicht zu einer Basis ergänzen durch

$$\vec{v}_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 5 \\ 0 \\ -3 \end{pmatrix}.$$

Dies sind zwei linear unabhängige Vektoren, die senkrecht auf  $\vec{v}_3$  stehen und daher eine Basis der Ebene.

Wir machen daraus mit dem Orthonormalisierungsverfahren eine Orthonormalbasis:

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|_2}, \quad \|\vec{v}_1\|_2 = \sqrt{2^2 + 3^2 + 0^2} = \sqrt{13} \rightsquigarrow \vec{w}_1 = \frac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

Für  $w_2$  berechnen wir das Skalarprodukt

$$\langle \vec{v}_2, \vec{w}_1 \rangle = \frac{1}{\sqrt{13}}(5 \cdot 2 + 0 \cdot 3 + (-3) \cdot 0) = \frac{10}{\sqrt{13}}$$

und damit

$$\begin{aligned} \vec{u}_2 &= \vec{v}_2 - \langle \vec{v}_2, \vec{w}_1 \rangle \vec{w}_1 \\ &= \begin{pmatrix} 5 \\ 0 \\ -3 \end{pmatrix} - \frac{10}{\sqrt{13}} \frac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \\ &= \frac{1}{13} \begin{pmatrix} 65 - 20 \\ 0 - 30 \\ -39 - 0 \end{pmatrix} \\ &= \frac{1}{13} \begin{pmatrix} 45 \\ -30 \\ -39 \end{pmatrix} \end{aligned}$$

und schließlich noch

$$\begin{aligned} \|\vec{u}_2\| &= \frac{1}{13} \sqrt{45^2 + 30^2 + 39^2} = \frac{1}{13} \sqrt{2025 + 900 + 1521} \\ &= \frac{1}{13} \sqrt{4446} = \frac{1}{13} \sqrt{9 \cdot 494} = \frac{3}{13} \sqrt{494}. \end{aligned}$$

Wir bekommen

$$\vec{w}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{13}{3\sqrt{494}} \frac{1}{13} \begin{pmatrix} 45 \\ -30 \\ -39 \end{pmatrix} = \frac{1}{\sqrt{494}} \begin{pmatrix} 15 \\ -10 \\ -13 \end{pmatrix}.$$

Damit bekommen wir die Matrixdarstellung der Orthogonalprojektion auf die Ebene als

$$\begin{aligned} P_E &= \vec{w}_1 \vec{w}_1^T + \vec{w}_2 \vec{w}_2^T \\ &= \frac{1}{13} \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \begin{pmatrix} 2 & 3 & 0 \end{pmatrix} + \frac{1}{494} \begin{pmatrix} 15 \\ -10 \\ -13 \end{pmatrix} \begin{pmatrix} 15 & -10 & -13 \end{pmatrix} \\ &= \frac{1}{13} \begin{pmatrix} 4 & 6 & 0 \\ 6 & 9 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{494} \begin{pmatrix} 225 & -150 & -195 \\ 150 & 100 & 130 \\ -195 & 130 & 169 \end{pmatrix} \\ &= \dots = \frac{1}{38} \begin{pmatrix} 29 & 6 & -15 \\ 6 & 34 & 10 \\ -15 & 10 & 13 \end{pmatrix}. \end{aligned}$$

Wir können auch eine Matrixdarstellung herleiten, wenn wir eine nicht-orthonormale Basis  $(\vec{v}_1, \dots, \vec{v}_m)$  von  $W$  haben: Der Vektor  $\vec{\alpha}$  der Koeffizienten der Basisdarstellung ist dann die Lösung von

$$A^T A \vec{\alpha} = A^T \vec{x}.$$

Da die Gramsche Matrix invertierbar ist, haben wir

$$\vec{\alpha} = (A^T A)^{-1} A^T \vec{x}.$$

Wegen  $P_W(\vec{x}) = \sum_{i=1}^m \alpha_i \vec{v}_i = A \vec{\alpha}$  folgt

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

## 21 Bestapproximation

Schauen wir uns den Zusammenhang von orthogonalen Projektionen und Approximation weiter an:

**Satz 21.1.** Sei  $V$  ein euklidischer Vektorraum und  $W$  ein endlichdimensionaler Unterraum von  $V$ . Zu  $\vec{x} \in V$  sei  $P_W$  die orthogonale Projektion auf  $W$ . Dann gilt für alle  $\vec{w} \in W$

$$\|\vec{x} - P_W(\vec{x})\| \leq \|\vec{x} - \vec{w}\|$$

und Gleichheit gilt nur für  $w = P_W(\vec{x})$ .

*Beweis.*

Aus der Gleichung

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + 2\langle \vec{u}, \vec{v} \rangle + \|\vec{v}\|^2$$

folgt der Satz des Pythagoras

$$\vec{u} \perp \vec{v} \iff \|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2.$$

Dann gilt für jedes  $w \in W$  auch  $P_W(\vec{x}) - \vec{w} \in W$  und daher  $P_W(\vec{x}) - \vec{w} \perp \vec{x} - P_W(\vec{x})$  und es folgt

$$\begin{aligned} \|\vec{x} - \vec{w}\|^2 &= \|\vec{x} - P_W(\vec{x}) + (P_W(\vec{x}) - \vec{w})\|^2 \\ &= \|\vec{x} - P_W(\vec{x})\|^2 + \|P_W(\vec{x}) - \vec{w}\|^2 \\ &\geq \|\vec{x} - P_W(\vec{x})\|^2 \end{aligned}$$

und Gleichheit gilt nur dann, wenn  $P_W(\vec{x}) - \vec{w} = 0$ .  $\square$

Beachte: Im obigen Satz muss nur  $W$  endlichdimensional sein. Der umgebende Raum  $V$  kann unendlichdimensional sein!

*Beispiel.* Es sei  $V$  der Vektorraum  $C[0, \frac{\pi}{2}]$  der stetigen, reellwertigen Funktionen auf dem Intervall  $[0, \frac{\pi}{2}]$ . Wir statten diesem mit dem  $L^2$ -Skalarprodukt aus:

$$\langle u, v \rangle = \int_0^{\pi/2} u(t)v(t)dt.$$

Wir suchen nun ein Polynom ersten Grades, welches die Funktion  $u(t) = \sin(t)$  im Sinne der induzierten  $L^2$ -Norm  $\|u\| = \sqrt{\langle u, u \rangle}$  am besten approximiert. Unser Unterraum ist also  $\Pi_1$  und wir suchen eine Funktion  $g(t) = \alpha_1 + \alpha_2 t$  für welche

$$\|u - g\| = \left( \int_0^{\pi/2} (g(t) - \sin(t))^2 dt \right)^{1/2}$$

so klein wie möglich ist.

Das heißt,  $P_W(\vec{x})$  ist tatsächlich das Element aus  $W$ , welches am nächsten an  $\vec{x}$  liegt, wie wir im letzten Abschnitt vermutet haben.

Wählen wir die Monombasis  $v_1(t) = 1$ ,  $v_2(t) = t$  und berechnen die Gramsche Matrix

$$\begin{aligned} \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{pmatrix} &= \begin{pmatrix} \int_0^{\pi/2} 1 dt & \int_0^{\pi/2} t dt \\ 0 & 0 \\ \int_0^{\pi/2} t dt & \int_0^{\pi/2} t^2 dt \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\pi}{2} & \frac{\pi^2}{8} \\ \frac{\pi^2}{8} & \frac{\pi^3}{24} \end{pmatrix}. \end{aligned}$$

Außerdem brauchen wir noch den Vektor

$$\begin{pmatrix} \langle v_1, u \rangle \\ \langle v_2, u \rangle \end{pmatrix} = \begin{pmatrix} \int_0^{\pi/2} \sin(t) dt \\ 0 \\ \int_0^{\pi/2} t \sin(t) dt \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Wir müssen also das Gleichungssystem

$$\begin{pmatrix} \frac{\pi}{2} & \frac{\pi^2}{8} \\ \frac{\pi^2}{8} & \frac{\pi^3}{24} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

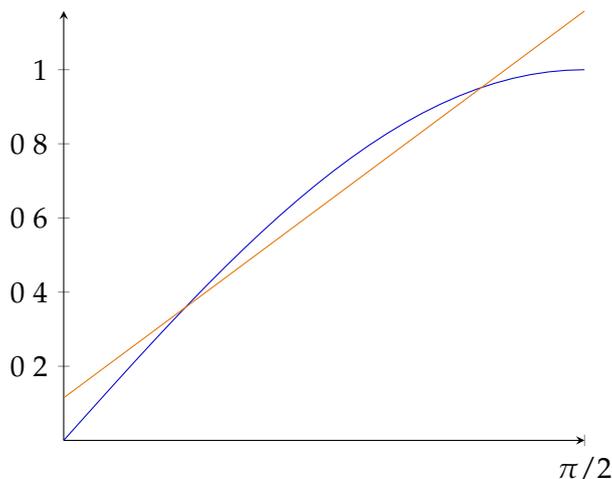
lösen. Die Inverse der  $2 \times 2$  Matrix ist

$$\frac{1}{\frac{\pi^4}{48} - \frac{\pi^4}{64}} \begin{pmatrix} \frac{\pi^3}{24} & -\frac{\pi^2}{8} \\ -\frac{\pi^2}{8} & \frac{\pi}{2} \end{pmatrix} = \frac{192}{\pi^3} \begin{pmatrix} \frac{\pi^2}{24} & -\frac{\pi}{8} \\ -\frac{\pi}{8} & \frac{1}{2} \end{pmatrix} = \frac{96}{\pi^3} \begin{pmatrix} \frac{\pi^2}{12} & -\frac{\pi}{4} \\ -\frac{\pi}{4} & 1 \end{pmatrix}$$

und wir bekommen die Lösung

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \frac{96}{\pi^3} \begin{pmatrix} \frac{\pi^2}{12} & -\frac{\pi}{4} \\ -\frac{\pi}{4} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{96}{\pi^3} \begin{pmatrix} \frac{\pi^2}{12} - \frac{\pi}{4} \\ -\frac{\pi}{4} + 1 \end{pmatrix} = \begin{pmatrix} \frac{8\pi-3}{24\frac{\pi^2}{\pi^3}} \end{pmatrix} \approx \begin{pmatrix} 0.1148 \\ 0.6644 \end{pmatrix}.$$

Das Ergebnis der Bestapproximation ist also  $g(t) \approx 0.1148 + 0.6644t$ :



*Beispiel.* Als weiteres Beispiel, betrachten wir die Approximation eines Sprunges mit trigonometrischen Funktionen. Wir betrachten den Vektorraum

$$V = \{f : [0, 2\pi] \rightarrow \mathbb{R} \mid f \text{ ist integrierbar und das Integral } \int_0^{2\pi} f(t)^2 dt \text{ existiert}\}$$

mit dem  $L^2$ -Skalarprodukt und der induzierten  $L^2$ -Norm

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt, \quad \|f\| = \sqrt{\int_0^{2\pi} f(t)^2 dt}.$$

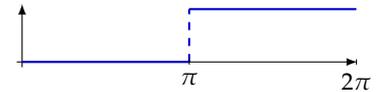
Weiterhin betrachten wir den Unterraum

$$W = \text{Spann}\{1, \sin(t), \cos(t), \sin(2t), \cos(2t), \sin(3t), \cos(3t)\}.$$

Als Funktion, die wir approximieren möchten, nehmen wir

$$u(t) = \begin{cases} 0 & : t \leq \pi \\ 1 & : t > \pi. \end{cases}$$

Den Nachweis, dass  $V$  wirklich ein Vektorraum ist, lassen wir weg.



Als ersten schauen wir uns die Basis an und konstruieren eine Orthonormalbasis von  $W$ . Wir setzen

$$\begin{aligned} v_0 &\equiv 1, & v_1 &= \sin(t), & v_2 &= \cos(t) \\ v_3 &= \sin(2t), & v_4 &= \cos(2t), & v_5 &= \sin(3t), & v_6 &= \cos(3t). \end{aligned}$$

Man kann zeigen, dass diese Funktionen tatsächlich schon orthogonal bezüglich des  $L^2$ -Skalarproduktes sind. Um eine Orthonormalbasis von  $W$  zu erhalten, müssen wir die Funktionen nur normieren. Hierfür kann man berechnen

$$\|v_0\| = \sqrt{\int_0^{2\pi} 1 dt} = \sqrt{2\pi}$$

und für  $k = 1, 2, \dots$

$$\int_0^{2\pi} \sin(kt)^2 dt = \int_0^{2\pi} \cos(kt)^2 dt = \pi.$$

Wir bekommen  $\|v_j\| = \sqrt{\pi}$  für  $j = 1, \dots, 6$ . Also bilden die  $w_0, \dots, w_6$  definiert durch

$$\begin{aligned} w_0 &\equiv \frac{1}{\sqrt{2\pi}}, & w_1 &= \frac{1}{\sqrt{\pi}} \sin(t), & w_2 &= \frac{1}{\sqrt{\pi}} \cos(t) \\ w_3 &= \frac{1}{\sqrt{\pi}} \sin(2t), & w_4 &= \frac{1}{\sqrt{\pi}} \cos(2t), & w_5 &= \frac{1}{\sqrt{\pi}} \sin(3t), & w_6 &= \frac{1}{\sqrt{\pi}} \cos(3t). \end{aligned}$$

eine Orthonormalbasis von  $W$ . Die Bestapproximation von  $u$  in  $W$  ist

$$P_W(u) = \sum_{i=0}^6 \langle w_i, u \rangle w_i.$$

Wir berechnen die Skalarprodukte

Erinnere:  $u(t) = 0$  für  $0 \leq t \leq \pi$ ,  
 $u(t) = 1$  für  $\pi < t \leq 2\pi$ .

$$\langle w_0, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{2\pi}} dt = \frac{\pi}{\sqrt{2\pi}} = \sqrt{\frac{\pi}{2}},$$

$$\langle w_1, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{\pi}} \sin(t) dt = \left[ -\frac{1}{\sqrt{\pi}} \cos(t) \right]_{t=\pi}^{2\pi} = -\frac{2}{\sqrt{\pi}},$$

$$\langle w_2, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{\pi}} \cos(t) dt = \left[ \frac{1}{\sqrt{\pi}} \sin(t) \right]_{t=\pi}^{2\pi} = 0,$$

$$\langle w_3, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{\pi}} \sin(2t) dt = \left[ -\frac{1}{2\sqrt{\pi}} \cos(2t) \right]_{t=\pi}^{2\pi} = 0,$$

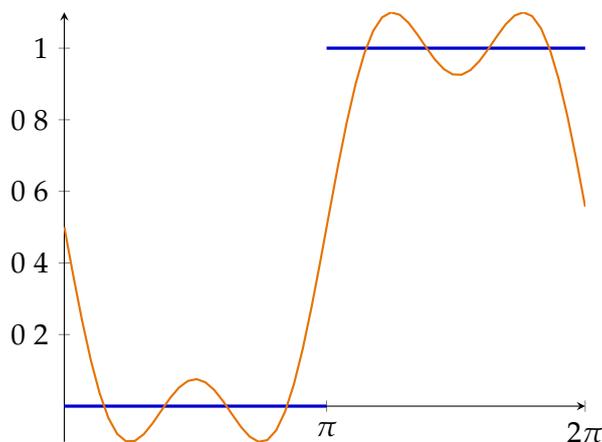
$$\langle w_4, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{\pi}} \cos(2t) dt = \left[ \frac{1}{2\sqrt{\pi}} \sin(2t) \right]_{t=\pi}^{2\pi} = 0,$$

$$\langle w_5, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{\pi}} \sin(3t) dt = \left[ -\frac{1}{3\sqrt{\pi}} \cos(3t) \right]_{t=\pi}^{2\pi} = -\frac{2}{3\sqrt{\pi}},$$

$$\langle w_6, u \rangle = \int_{\pi}^{2\pi} \frac{1}{\sqrt{\pi}} \cos(3t) dt = \left[ \frac{1}{3\sqrt{3\pi}} \sin(3t) \right]_{t=\pi}^{2\pi} = 0.$$

Die Bestapproximation ist also

$$\begin{aligned} P_W(u) &= \sqrt{\frac{\pi}{2}} w_0 - \frac{2}{\sqrt{\pi}} w_1 - \frac{2}{3\sqrt{\pi}} w_5 \\ &= \sqrt{\frac{\pi}{2}} \frac{1}{\sqrt{2\pi}} - \frac{2}{\sqrt{\pi}} \frac{1}{\sqrt{\pi}} \sin(t) - \frac{2}{3\sqrt{\pi}} \frac{1}{\sqrt{\pi}} \sin(3t) \\ &= \frac{1}{2} - \frac{2}{\pi} \sin(t) - \frac{2}{3\pi} \sin(3t) \end{aligned}$$



Beachte: Wir haben nebenbei auch die Bestapproximation in den Unterräumen  $W_0 = \text{Spann}\{w_0\}$ ,  $W_2 = \text{Spann}\{w_0, w_1, w_2\}$  und

$W_4 = \text{Spann} \{w_0, w_1, w_2, w_3, w_4\}$  berechnet; wir müssen die Summe nur entsprechend früher abbrechen:

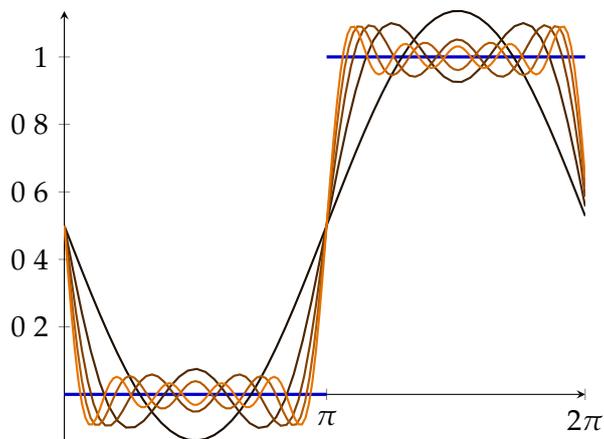
$$P_{W_0}(u) \equiv \frac{1}{2}, \quad P_{W_2}(u) = \frac{1}{2} - \frac{2}{\pi} \sin(t), \quad P_{W_4}(u) = \frac{1}{2} - \frac{2}{\pi} \sin(t).$$

Analog kann man ausrechnen, dass die Bestapproximation an  $u$  im Raum

$$\text{Spann} (\{\cos(kt) \mid 0 \leq k \leq K\} \cup \{\sin(kt) \mid 1 \leq k \leq K\})$$

gegeben ist durch

$$\frac{1}{2} - \frac{\pi}{2} \sum_{2k-1 \leq K} \frac{\sin((2k-1)t)}{2k-1}.$$





Da  $(\vec{w}_k)$  eine Orthonormalbasis ist, wissen wir  $\gamma_1 = \langle \vec{x}, \vec{w}_1 \rangle = \langle \vec{w}, \vec{x} \rangle$ . Wir bekommen

$$H(\vec{x}) = \vec{x} - 2\langle \vec{w}, \vec{x} \rangle \vec{w} = \vec{x} - 2\vec{w}^T \vec{x} \vec{w} = (I - 2\vec{w}\vec{w}^T)\vec{x}.$$

Wir definieren:

**Definition 22.1.** Die Spiegelung des  $\mathbb{R}^n$  an der Hyperebene  $\vec{w}^\perp$  mit  $\|\vec{w}\|_2 = 1$  wird durch die Matrix

$$H := I_n - 2\vec{w}\vec{w}^T$$

beschrieben, man nennt solche Matrizen auch *Householder-Matrizen*.

Die Matrix der Wpiegelung an  $\vec{v}^\perp$  für unnormiertes  $\vec{v}$  ist entsprechend  $H = I_n - \frac{2}{\|\vec{v}\|_2^2} \vec{v}\vec{v}^T$ .

**Satz 22.2.** Ist  $H$  eine Householder-Matrix, so gilt

- (a)  $H^2 = I_n$ ,
- (b)  $H^T = H$
- (c)  $H^T H = HH^T = I_n$

*Beweis.*

$$(a) \quad H^2 = HH = (I_n - 2\vec{w}\vec{w}^T)(I_n - 2\vec{w}\vec{w}^T) = I_n - 4\vec{w}\vec{w}^T + 4\vec{w} \underbrace{\vec{w}^T \vec{w}}_{=\|\vec{w}\|_2^2=1} \vec{w}^T = I_n.$$

$$(b) \quad H^T = (I_n - 2\vec{w}\vec{w}^T)^T = I_n - 2\vec{w}\vec{w}^T = H.$$

$$(c) \quad \text{Mit den ersten beiden Punkten folgt } H^T H = H^2 = I_n.$$

□

**Satz 22.3.** Sind  $\vec{x}, \vec{y} \in \mathbb{R}^n$  mit  $\vec{x} \neq \vec{y}$  und  $\|\vec{x}\|_2 = \|\vec{y}\|_2$ , so gibt es eine Householder-Matrix  $H$ , mit  $H\vec{x} = \vec{y}$  und dieses  $H$  ist von der Form  $H = I_n - 2\vec{w}\vec{w}^T$  mit  $\vec{w} = (\vec{x} - \vec{y}) / \|\vec{x} - \vec{y}\|_2$ .

*Beweis.*

Aus  $H = I_n - 2\vec{w}\vec{w}^T$  und  $H\vec{x} = \vec{y}$  folgt

$$\vec{y} = H\vec{x} = \vec{x} - 2\vec{w}(\vec{w}^T \vec{x}).$$

Wir formen um zu  $\vec{x} - \vec{y} = 2\langle \vec{w}, \vec{x} \rangle \vec{w}$  und sehen, dass  $\vec{w}$  ein Vielfaches von  $\vec{x} - \vec{y}$  sein muss, also  $\vec{w} = \gamma(\vec{x} - \vec{y})$ . Da  $\|\vec{w}\|_2 = 1$  gelten muss, folgt  $\gamma = \pm \frac{1}{\|\vec{x} - \vec{y}\|_2}$ . Wir setzen ein, um zu überprüfen, ob das positive Vorzeichen das richtige ist: Da  $\|\vec{x}\|_2^2 = \|\vec{y}\|_2^2$  gilt, folgt  $2\langle \vec{x} - \vec{y}, \vec{x} \rangle = 2\|\vec{x}\|_2^2 - 2\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\|_2^2 - 2\langle \vec{x}, \vec{y} \rangle + \|\vec{y}\|_2^2 = \|\vec{x} - \vec{y}\|_2^2$  und daher ist

$$\begin{aligned} 2\langle \vec{w}, \vec{x} \rangle \vec{w} &= \frac{2\langle \vec{x} - \vec{y}, \vec{x} \rangle}{\|\vec{x} - \vec{y}\|_2^2} (\vec{x} - \vec{y}) \\ &= \vec{x} - \vec{y}. \end{aligned}$$

□

Man kann orthonormale Matrizen benutzen, um lineare Gleichungen zu lösen. Eine Methode dafür ist QR-Zerlegung:

**Satz 22.4.** Es sei  $A \in \mathbb{R}^{m \times n}$  mit  $m \geq n$ . Dann existiert eine orthonormale Matrix  $Q \in \mathbb{R}^{m \times m}$  und eine obere Dreiecksmatrix  $R \in \mathbb{R}^{m \times n}$ , so dass  $A = QR$ .

*Beweis.*

Wir beweisen diesen Satz konstruktiv und geben einen Algorithmus zur Konstruktion von  $Q$  und  $R$  an: Es sei  $\vec{a}_1$  die erste Spalte von  $A$ . Wie definieren  $\alpha_1$  so, dass  $|\alpha_1| = \|\vec{a}_1\|_2$  und dass das Vorzeichen von  $\alpha_1$  umgekehrt zum ersten Eintrag in  $\vec{a}_1$  ist (hier und später im Algorithmus wenden wir die Konvention an, dass das Vorzeichen negativ ist, wenn der Eintrag in  $\vec{a}_1$  Null ist). Wir suchen eine Householder-Matrix  $H_1 \in \mathbb{R}^{m \times m}$  mit der Eigenschaft, dass

$$H_1 A = \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix}.$$

Die Matrix  $H_1$  muss also  $H_1 \vec{a}_1 = \alpha_1 \vec{e}_1$  erfüllen. Nach dem vorigen Satz und der Randbemerkung zu Definition 22.1 wird dies geleistet durch  $H_1 = I_m - 2 \frac{\vec{w}_1 \vec{w}_1^T}{\|\vec{w}_1\|_2^2}$  mit  $\vec{w}_1 = \vec{a}_1 - \alpha_1 \vec{e}_1$ .

Die Matrix  $A_2$  ist von der Größe  $(m-1) \times (n-1)$  und so lange  $m-1 > 1$  ist, fahren wir wie im ersten Schritt fort: Es sein  $\vec{a}_2 \in \mathbb{R}^{m-1}$  die erste Spalte von  $A_2$  und  $\alpha_2$  so, dass  $|\alpha_2| = \|\vec{a}_2\|_2$  und das Vorzeichen von  $\alpha_2$  ist umgekehrt zu dem des ersten Eintrags in  $\vec{a}_2$ . Setze  $\vec{w}_2 = \vec{a}_2 - \alpha_2 \vec{e}_1$  (beachte: Hier ist  $\vec{e}_1 \in \mathbb{R}^{m-1}$ ),  $H_2' = I_{m-1} - 2 \frac{\vec{w}_2 \vec{w}_2^T}{\|\vec{w}_2\|_2^2}$  und

$$H_2 = \begin{pmatrix} 1 & \\ & H_2' \end{pmatrix} \in \mathbb{R}^{m \times m}.$$

Dann ist

$$H_2 H_1 A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & H_2' & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & \alpha_2 & * & * \\ \vdots & 0 & & \\ 0 & 0 & & A_3 \end{pmatrix}.$$

So fahren wir weiter bis Schritt  $t = \min(m-1, n)$  fort, bis wir eine obere Dreiecksmatrix

$$R = H_t \cdots H_1 A$$

erreichen. Da die  $H_i$  alle orthonormal und symmetrisch sind, bekommen wir die QR-Zerlegung als

$$A = \underbrace{H_1 \cdots H_t}_:=Q R.$$

□

Einige Bemerkungen zur QR-Zerlegung:

- Mit Hilfe der QR-Zerlegung kann man konsistente überbestimmte Gleichungssysteme lösen: Hat man ein System  $A\vec{x} = \vec{b}$  und eine QR-Zerlegung, dann ist das System  $QR\vec{x} = \vec{b}$  und da  $Q$  orthonormal ist, folgt  $R\vec{x} = Q^T\vec{b}$ . Da  $R$  eine obere Dreiecksmatrix ist, lässt sich nun ablesen, ob es Lösungen gibt, und sie lassen sich durch Rückwärtseinsetzen bestimmen.
- Man kann den Speicheraufwand der QR-Zerlegung reduzieren: Da die obere Dreiecksmatrix von der Form

$$R = \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix} \text{ mit } \hat{R} \in \mathbb{R}^{n \times n}$$

ist, definiert man  $\hat{Q} \in \mathbb{R}^{m \times n}$  indem man die ersten  $n$  Spalten von  $Q$  in  $\hat{Q}$  schreibt. Dann gilt

$$QR = (\hat{Q} \quad Q') \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix} = \hat{Q}\hat{R}.$$

Da die Spalten von  $\hat{Q}$  orthonormal sind, gilt  $\hat{Q}^T\hat{Q} = I_n$  und  $A\vec{x} = \vec{b}$  ist äquivalent zu  $\hat{R}\vec{x} = \hat{Q}^T\vec{b}$ .

*Beispiel.* Wir betrachten die Matrix

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

Die Norm der ersten Spalte ist  $\|\vec{a}_1\|_2 = \sqrt{4 + 1 + 4} = 3$  und da der Eintrag 2 links oben positiv ist, ist  $\alpha_1 = -3$ . Daher nehmen wir

$$\vec{w}_1 = \vec{a}_1 + 3\vec{e}_1 = \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}$$

und wegen  $\|\vec{w}_1\|_2^2 = 25 + 1 + 4 = 30$  ist

$$H_1 = I_3 - 2 \frac{\vec{w}_1 \vec{w}_1^T}{\|\vec{w}_1\|_2^2} = I_3 - \frac{2}{30} \begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 25 & 5 & 4 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{14}{15} & -\frac{2}{15} \\ -\frac{2}{3} & -\frac{2}{15} & \frac{11}{15} \end{pmatrix}.$$

Das ergibt

$$H_1 A = \begin{pmatrix} -3 & -1 & 2 \\ 0 & 0 & \frac{16}{5} \\ 0 & -5 & \frac{12}{5} \end{pmatrix}.$$

Die QR-Zerlegung lässt sich auch anders ausrechnen, z.B. mit Hilfe des Gram-Schmidt-Verfahrens oder mit Hilfe von Givens-Rotationen.

Wir bekommen

$$A_2 = \begin{pmatrix} 0 & \frac{16}{5} \\ -5 & \frac{12}{5} \end{pmatrix}.$$

Hier ist die Norm der ersten Spalte 5 und der erste Eintrag ist 0, daher ist (nach unserer Konvention)

$$\vec{w}_2 = \begin{pmatrix} 0 \\ -5 \end{pmatrix} + 5\vec{e}_1 = \begin{pmatrix} 5 \\ -5 \end{pmatrix}.$$

Das gibt

$$H'_2 = I_2 - 2 \frac{\vec{w}_2 \vec{w}_2^T}{\|\vec{w}_2\|_2^2} = I_2 - \frac{2}{50} \begin{pmatrix} 25 & -25 \\ -25 & 25 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Also ist

$$H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

und wir bekommen

$$H_2 H_1 A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -3 & -1 & 2 \\ 0 & 0 & \frac{16}{5} \\ 0 & -5 & \frac{12}{5} \end{pmatrix} = \begin{pmatrix} -3 & -1 & 2 \\ 0 & -5 & \frac{12}{5} \\ 0 & 0 & \frac{16}{5} \end{pmatrix} =: R$$

Die Matrix  $Q$  bekommen wir als

$$Q := H_1 H_2 = \begin{pmatrix} -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{14}{15} & -\frac{2}{15} \\ -\frac{2}{3} & -\frac{2}{15} & \frac{11}{15} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{15} & \frac{14}{15} \\ -\frac{2}{3} & \frac{11}{15} & -\frac{2}{15} \end{pmatrix}.$$

Mit diesen Matrizen gilt  $A = QR$ .

## 23 Lineare Ausgleichsprobleme

Bei der Ausgleichsrechnung geht es um die Anpassung von Parametern an Messdaten. Anfang des Jahres 1801 beobachtete Giusseppe Piazzi einen neuen Himmelskörper und vermutete, dass es sich um einen neuen Planeten handeln könnte. Es konnte nur wenige Daten aufnehmen (Wolkennächte verhinderten mehr Aufnahmen) und publizierte diese. Carl-Friedrich Gauß bemerkte die Publikation und wollte herausfinden, ob es sich um einen Planeten oder einen Kometen handelte. Erstere haben eine elliptische Bahn, zweitere eine eher parabolische Bahn. Er nahm an, dass es eine elliptische Bahn war und wendete die *Methode der kleinsten Quadrate* an, um die Parameter der Bahn aus den Messdaten zu ermitteln. Tatsächlich konnte das Objekt gegen Ende des Jahres 1801 auf Grund seiner vorhergesagten wieder entdeckt werden.

Eine Ellipse in der Ebene ist die Menge aller Punkte  $(x, y)$ , die der Gleichung

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

für Parameter  $a, b$  genügen. Liegen zwei Messungen  $x_1 = x(t_1)$ ,  $y_1 = y(t_1)$  und  $x_2 = x(t_2)$ ,  $y_2 = y(t_2)$  zu verschiedenen Zeitpunkten vor, so könnte man die Parameter  $\alpha = 1/a^2$  und  $\beta = 1/b^2$  aus dem Gleichungssystem

$$\begin{pmatrix} x_1^2 & y_1^2 \\ x_2^2 & y_2^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

berechnen.

Aber Achtung: Jede Messung ist mit einem (ggf. kleinen) Fehler behaftet. Außerdem hat man häufig mehr Messungen als man für die Lösung des System benötigt (hier mehr als zwei). Welche soll man nehmen?

Wir betrachten eine (als bekannt angenommene) Funktion  $f$ , die von Parametern  $x_1, \dots, x_n$  abhängt und die einen Zusammenhang zwischen zwei Größen  $t$  und  $y$  beschreibt, also  $f(t; x_1, \dots, x_n) = y$ . Es seien  $(t_i, y_i)$ ,  $i = 1, \dots, m$  Messwerte bekannt, wobei  $m \geq n$  gelten soll (ggf. ist  $m$  sehr viel größer als  $n$ ). Wir fragen uns: Wie kann man die Parameter aus dem (ggf. nichtlinearen) System

$$f(t_i, x_1, \dots, x_n) = y_i, \quad i = 1, \dots, m$$

bestimmen oder wenigstens schätzen? Eine Lösung des Gleichungssystems können wir im allgemeinen (auf Grund von Messfehlern) nicht erwarten.

Stattdessen, versucht man, die *Residuen*

$$r_i := f(t_i, x_1, \dots, x_n) - y_i$$

so klein wie möglich zu machen. Hierfür gibt es viele Möglichkeiten, z.B.: Bestimme die Parameter  $x_1, \dots, x_n$  so, dass

Es handelte sich um einen „Planetoiden“ (Kleinplanet) der heute Ceres genannt wird.

- (a) die Summe der Quadrate, also  $\|\vec{r}\|_2^2 = \sum_{i=1}^m r_i^2$  möglichst klein wird. (Methode der kleinsten Quadrate - Method of least squares)
- (b) die Summe der absoluten Abweichungen  $\|\vec{r}\|_1 = \sum_{i=1}^m |r_i|$  möglichst klein wird. (Least absolute deviations)
- (c) die größte Abweichung  $\|\vec{r}\|_\infty = \max_{1 \leq i \leq m} |r_i|$  möglichst klein wird. (Tschebyscheff-Ausgleichsrechnung)

Wir untersuchen nur die Methode der kleinsten Quadrate und nehmen an, dass die Funktion  $f$  linear bezüglich der Parameter  $x_1, \dots, x_n$  ist.

*Beispiel.* Gegeben seien Messwerte  $(t_i, y_i)$  und wir möchte den Verlauf durch eine Ausgleichsparabel  $y = c_0 + c_1t + c_2t^2$  beschreiben. Unsere Funktion ist also

$$f(t; c_0, c_1, c_2) = c_0 + c_1t + c_2t^2.$$

In diesem Fall ist die Funktion linear in den Variablen  $c_0, c_1, c_2$ .

Ist  $f(t; x_1, \dots, x_n)$  linear in den Parametern, so gilt

$$f(t; x_1, \dots, x_n) = \sum_{j=1}^n x_j f_j(t)$$

mit Funktionen  $f_j$ . Die Residuen sind also

$$r_i = \sum_{j=1}^n f_j(t_i) x_j - y_i.$$

In Matrix-Schreibweise ist das

$$\underbrace{\begin{pmatrix} f_1(t_1) & \cdots & f_n(t_1) \\ \vdots & & \vdots \\ f_1(t_m) & \cdots & f_n(t_m) \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \approx \vec{0}.$$

**Definition 23.1.** Es sei  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$  mit  $\text{Rang}(A) = n$  und  $b \in \mathbb{R}^m$ . Das *lineare Ausgleichsproblem* sucht einen Vektor  $\vec{x} \in \mathbb{R}^n$ , so dass  $\|A\vec{x} - \vec{b}\|_2$  (bzw. äquivalent  $\|A\vec{x} - \vec{b}\|_2^2$ ) möglichst klein wird. Der Vektor  $\vec{r}(\vec{x}) := A\vec{x} - \vec{b}$  heißt *Residuum* und  $\|\vec{r}(\vec{x})\|_2^2$  heißt *Zielfunktion*.

Quadrieren der Zielfunktionen ändert die Minimierer nicht.

Um die Zielfunktion zu minimieren, suchen wir einen Punkt  $\vec{x}$  in dem sich die Zielfunktion unter kleinen Änderungen nur vergrößert. Wir betrachten also eine kleine Änderung  $\Delta\vec{x}$ , d.h. wir betrachten

$$\vec{r}(\vec{x} + \Delta\vec{x}) = A(\vec{x} + \Delta\vec{x}) - \vec{b} = \vec{r}(\vec{x}) + A\Delta\vec{x}.$$

Für die quadrierte Zielfunktion gilt dann

$$\begin{aligned}\|\vec{r}(\vec{x} + \Delta\vec{x})\|_2^2 &= \langle \vec{r}(\vec{x}) + A\Delta\vec{x}, \vec{r}(\vec{x}) + A\Delta\vec{x} \rangle \\ &= \langle \vec{r}(\vec{x}), \vec{r}(\vec{x}) \rangle + 2\langle \vec{r}(\vec{x}), A\Delta\vec{x} \rangle + \langle A\Delta\vec{x}, A\Delta\vec{x} \rangle \\ &= \|\vec{r}(\vec{x})\|_2^2 + 2\vec{r}(\vec{x})^T A\Delta\vec{x} + \|A\Delta\vec{x}\|_2^2 \\ &= \|\vec{r}(\vec{x})\|_2^2 + 2\Delta\vec{x}^T (A^T \vec{r}(\vec{x})) + \|A\Delta\vec{x}\|_2^2.\end{aligned}$$

wir folgern:

- (a) Gilt  $A^T \vec{r}(\vec{x}) = \vec{0}$ , so ist  $\|\vec{r}(\vec{x} + \Delta\vec{x})\|_2^2 \geq \|\vec{r}(\vec{x})\|_2^2$  für alle  $\Delta\vec{x}$  und das heißt,  $\vec{x}$  ist eine Lösung des Ausgleichsproblems.
- (b) Ist  $\text{Rang}(A) = n$ , so liefert  $A^T \vec{r}(\vec{x}) = \vec{0}$  sogar  $\|\vec{r}(\vec{x} + \Delta\vec{x})\|_2^2 > \|\vec{r}(\vec{x})\|_2^2$  für alle  $\Delta\vec{x} \neq 0$ , d.h.  $\vec{x}$  ist die *eindeutige* Lösung des Ausgleichsproblems.

Die Bedingung  $A^T \vec{r}(\vec{x})$  ist ausgeschrieben

$$A^T \vec{r}(\vec{x}) = \vec{0} \iff A^T(A\vec{x} - \vec{b}) = \vec{0} \iff A^T A\vec{x} = A^T \vec{b}.$$

**Definition 23.2.** Das Gleichungssystem  $A^T A\vec{x} = A^T \vec{b}$  heißt *Normalengleichung* des Ausgleichsproblems  $\min \|A\vec{x} - \vec{b}\|_2$ .

Die Matrix  $A^T A$  ist eine symmetrische  $n \times n$ -Matrix und hat  $A$  Rang  $n$ , so ist  $A^T A$  invertierbar. Das heißt, die Normalengleichung hat auch für überbestimmte und inkonsistente Gleichungssystem  $A\vec{x} = \vec{b}$  eine eindeutige Lösung. Wir halten fest:

**Satz 23.3.** Sei  $m \geq n$ ,  $A \in \mathbb{R}^{m \times n}$  mit  $\text{Rang}(A) = n$  und  $\vec{b} \in \mathbb{R}^m$ . Dann hat das lineare Ausgleichsproblem  $\min \|A\vec{x} - \vec{b}\|_2$  eine eindeutig bestimmte Lösung  $\vec{x}$ , welche die Lösung der Normalengleichung

$$A^T A\vec{x} = A^T \vec{b}$$

ist.

Bei der Lösung mit dem Computer wird die Normalengleichung meist nicht eingesetzt, da ihre Lösung oft instabil ist. Stabiler lassen sich lineare Ausgleichsprobleme auf dem Computer mit der QR-Zerlegung lösen: Dazu nutzen wir folgende Eigenschaft von orthonormalen Matrizen aus: Ist  $Q \in \mathbb{R}^{m \times m}$ , so gilt für alle  $\vec{x}$ :  $\|Q\vec{x}\|_2 = \|\vec{x}\|_2 = \|Q^T \vec{x}\|$ . Haben wir also  $A = QR$  zerlegt, so gilt für die Zielfunktion

$$\|A\vec{x} - \vec{b}\|_2 = \|QR\vec{x} - \vec{b}\|_2 = \|Q^T QR\vec{x} - Q^T \vec{b}\|_2 = \|R\vec{x} - Q^T \vec{b}\|_2.$$

Da hier  $m \geq n$  gilt, hat  $R$  die Form

$$R = \begin{pmatrix} \tilde{R} \\ 0_{m-n \times n} \end{pmatrix}, \quad \text{mit oberer Dreiecksmatrix } \tilde{R} \in \mathbb{R}^{n \times n}.$$

Zerlegen wir den Vektor  $Q^T \vec{b}$  als

$$Q^T \vec{b} = \begin{pmatrix} \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} \quad \text{mit} \quad \vec{c}_1 \in \mathbb{R}^n, \quad \vec{c}_2 \in \mathbb{R}^{m-n},$$

so bekommen wir für das quadrierte Zielfunktional

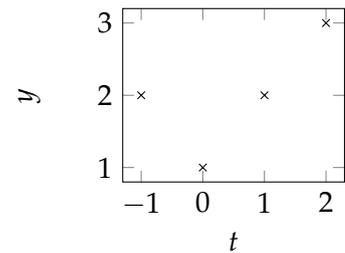
$$\|A\vec{x} - \vec{b}\|_2^2 = \left\| \begin{pmatrix} \tilde{R} \\ 0_{m-n \times n} \end{pmatrix} \vec{x} - \begin{pmatrix} \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} \right\|_2^2 = \|\tilde{R}\vec{x} - \vec{c}_1\|_2^2 + \|\vec{c}_2\|_2^2$$

und der optimale quadratische Fehler ist also  $\|\vec{c}_2\|_2^2$ . Die Lösung des Ausgleichsproblems erhält man also durch die Lösung per Rückwärtseinsetzen in  $\tilde{R}\vec{x} = \vec{c}_1$ .

Rechnen wir noch ein Beispiel per Hand:

*Beispiel.* Wir wollen eine Ausgleichsparabel durch die folgenden Punkte legen:

$$\begin{array}{c|cccc} t_i & -1 & 0 & 1 & 2 \\ \hline y_i & 2 & 1 & 2 & 3 \end{array}$$



Unsere Ausgleichsparabel hat die Form  $y = c_0 + c_1 t + c_2 t^2$  und wir bekommen die Gleichungen

$$\begin{aligned} y_1 &= c_0 + c_1 t_1 + c_2 t_1^2 = c_0 - c_1 + c_2 & = 2 \\ y_2 &= c_0 + c_1 t_2 + c_2 t_2^2 = c_0 & = 1 \\ y_3 &= c_0 + c_1 t_3 + c_2 t_3^2 = c_0 + c_1 + c_2 & = 2 \\ y_4 &= c_0 + c_1 t_4 + c_2 t_4^2 = c_0 + 2c_1 + 4c_2 & = 3 \end{aligned}$$

In Matrix-Schreibweise ist das Residuum also

$$\vec{r} = \underbrace{\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix}}_{=\vec{x}} - \underbrace{\begin{pmatrix} 2 \\ 1 \\ 2 \\ 3 \end{pmatrix}}_{=\vec{b}}.$$

Die Normalengleichung ist also

$$A^T A \vec{x} = A^T \vec{b}$$

und wir berechnen

$$\begin{aligned} A^T A &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{pmatrix} \\ A^T \vec{b} &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \\ 16 \end{pmatrix} \end{aligned}$$

Wir müssen also lösen:

$$\begin{pmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \\ 16 \end{pmatrix}$$

Wenden wir das Eliminationsverfahren auf das System

$$\left( \begin{array}{ccc|c} 4 & 2 & 6 & 8 \\ 2 & 6 & 8 & 6 \\ 6 & 8 & 18 & 16 \end{array} \right)$$

an. Wir dividieren zuerst die erste Zeile durch zwei:

$$\left( \begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ 2 & 6 & 8 & 6 \\ 6 & 8 & 18 & 16 \end{array} \right).$$

Jetzt ziehen wir die erste von der zweiten Zeile ab und das dreifache der ersten von der dritten:

$$\left( \begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ 0 & 5 & 5 & 2 \\ 0 & 5 & 9 & 4 \end{array} \right).$$

Jetzt ziehen wir die zweite Zeile von der dritten Zeile ab und bekommen

$$\left( \begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ 0 & 5 & 5 & 2 \\ 0 & 0 & 4 & 2 \end{array} \right).$$

Wir lösen durch Rückwärtseinsetzen:

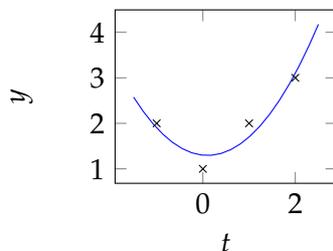
$$4c_2 = 2 \implies c_2 = \frac{1}{2}$$

$$5c_1 + 5c_2 = 2 \implies 5c_1 + \frac{5}{2} = 2 \implies c_1 = -\frac{1}{10}$$

$$2c_0 + c_1 + 3c_2 = 4 \implies 2c_0 - \frac{1}{10} + \frac{3}{2} = 4 \implies c_0 = \frac{13}{10}.$$

Unsere Ausgleichsparabel ist also

$$y = 1.3 - 0.1t + 0.5t^2.$$



Hier noch ein Beispiel, das zeigt, dass die Normalgleichungen instabil sein können: Wir betrachten die Matrix

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

für ein kleines  $\epsilon$ . Die Matrix hat Rang 2 (die Spalten sind linear unabhängig). Es ist aber

$$A^T A = \begin{pmatrix} 1 & \epsilon & 0 \\ 1 & 0 & \epsilon \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} = \begin{pmatrix} 1 + \epsilon^2 & 1 \\ 1 & 1 + \epsilon^2 \end{pmatrix}$$

Diese Matrix kann numerisch singular sein, nämlich dann, wenn  $\epsilon$  noch positiv ist,  $\epsilon^2$  jedoch so klein, dass es in Fließkommadarstellung schon  $= 0$  ist.

## 24 Eigenwerte und Eigenvektoren

*Beispiel.* Betrachten wir ein *lineares dynamisches System*: Zu einer gegebenen Matrix  $A \in \mathbb{R}^{n \times n}$  und gegebenen Anfangswert  $\vec{x}^{(0)}$  definieren wir rekursiv eine Folge

$$\vec{x}^{(k+1)} = A\vec{x}^{(k)}, k = 0, 1, \dots$$

und fragen uns, wie wir das Verhalten der Folge  $\vec{x}^{(k)}$  bestimmen können.

Konkret nehmen wir

$$A = \frac{1}{4} \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}.$$

Wir betrachten verschiedene Anfangswerte:

$$\begin{array}{lll} \vec{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & \vec{x}^{(1)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} & \vec{x}^{(2)} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} \\ \vec{x}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, & \vec{x}^{(1)} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} & \vec{x}^{(2)} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \end{pmatrix} \\ \vec{x}^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \vec{x}^{(1)} = \begin{pmatrix} \frac{5}{4} \\ \frac{3}{4} \end{pmatrix} & \vec{x}^{(2)} = \begin{pmatrix} 2.125 \\ 1.875 \end{pmatrix} \end{array}$$

Wir beobachten, dass, abhängig vom Anfangswert, ganz verschiedenes Verhalten auftritt. Die ersten beiden Anfangswerte werden auf Vielfache von sich selbst abgebildet. Der letzte Anfangswert liefert ein nicht so einfaches Bild, scheint sich aber einem Vielfachen des Vektors  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  anzunähern. Um dieses Phänomen zu verstehen untersuchen wir genauer, unter welchen Umständen Vektoren auf Vielfache von sich selbst abgebildet werden.

**Definition 24.1.** Eine Zahl  $\lambda \in \mathbb{R}/\mathbb{C}$  heißt *Eigenwert* einer quadratischen Matrix  $A \in \mathbb{K}^{n \times n}$ , wenn es einen Vektor  $\vec{v} \in \mathbb{K}^n \setminus \{\vec{0}\}$  gibt, so dass  $A\vec{v} = \lambda\vec{v}$  gilt. Der Vektor  $\vec{v} \neq 0$  heißt *Eigenvektor* von  $A$  zum Eigenwert  $\lambda$ .

Die Menge

$$E_\lambda(A) := \{\vec{v} \in \mathbb{K}^n \mid A\vec{v} = \lambda\vec{v}\}$$

heißt *Eigenraum* von  $A$  zum Eigenwert  $\lambda$ .

**Satz 24.2.** Für  $A \in \mathbb{K}^{n \times n}$  und  $\lambda \in \mathbb{K}$  gilt  $E_\lambda(A) = \text{Kern}(A - \lambda I_n)$ .

*Beweis.*

Eigenräume sind übrigens immer Unterräume!

Es gilt

$$\begin{aligned}\vec{v} \in E_\lambda(A) &\iff A\vec{v} = \lambda\vec{v} \\ &\iff (A - \lambda I_n)\vec{v} = \vec{0} \\ &\iff \vec{v} \in \text{Kern}(A - \lambda I_n).\end{aligned}$$

□

Das Bestimmen von Eigenvektoren und Eigenwerten ist *nicht-lineares Problem*, da die bestimmende Gleichung den nichtlinearen Term  $\lambda\vec{v}$  enthält. Ist jedoch ein Eigenwert bekannt, so ist das Problem, einen zugehörigen Eigenvektor zu finden linear (wie der vorige Satz zeigt). Doch wie kann man die Eigenwerte einer Matrix bestimmen? (Und hat überhaupt jede Matrix Eigenwerte?)

**Satz 24.3.** Es sei  $A \in \mathbb{K}^{n \times n}$ . Dann gilt:  $\lambda \in \mathbb{C}$  ist genau dann ein Eigenwert von  $A$ , wenn gilt  $\det(A - \lambda I_n) = 0$ .

Beweis.

Es gilt:

$$\begin{aligned}\lambda \text{ ist EW von } A &\iff \exists \vec{v} \neq \vec{0} : (A - \lambda I_n)\vec{v} = \vec{0} \\ &\iff \dim(\text{Kern}(A - \lambda I_n)) \geq 1 \\ &\iff A - \lambda I_n \text{ nicht invertierbar} \\ &\iff \det(A - \lambda I_n) = 0.\end{aligned}$$

□

**Definition 24.4.** Die Abbildung  $p_A(\lambda) := \det(A - \lambda I_n)$  heißt *charakteristisches Polynom* von  $A$ .

Das charakteristische Polynom ist tatsächlich ein Polynom:

$$p_A(\lambda) = \det \begin{pmatrix} (a_{11} - \lambda) & a_{12} & \cdots & a_{1n} \\ a_{21} & (a_{22} - \lambda) & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & (a_{nn} - \lambda) \end{pmatrix} = \sum_{\sigma \in S_n} \text{sign}(\sigma) \tilde{a}_{1\sigma_1} \cdots \tilde{a}_{n\sigma_n},$$

wobei

$$\tilde{a}_{jk} := \begin{cases} a_{jk} & : j \neq k \\ a_{jk} - \lambda & : j = k. \end{cases}$$

Wir sehen, dass jeder Summand ein Polynom in  $\lambda$  ist und daher ist auch  $p_A$  ein Polynom. Der Grad des Polynoms ist  $n$ , denn der Summand zur Permutation  $\sigma = \text{id}$  lautet  $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$  und hat den Grad  $n$ . Alle anderen Summanden haben mindestens zwei Faktoren mit der Form  $\tilde{a}_{j\sigma_j} = a_{j\sigma_j}$  und haben daher höchstens den Grad  $n - 2$ .

Wir wissen daher also:

$$\begin{aligned}p_A(\lambda) &= (a_{11} - \lambda) \cdots (a_{nn} - \lambda) + \text{Polynom vom Grad } \leq n - 2 \\ &= (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} (a_{11} + \cdots + a_{nn}) + \text{Polynom vom Grad } \leq n - 2.\end{aligned}$$

Mit dem Fundamentalsatz der Algebra bekommen wir:

**Satz 24.5.** Eine Matrix  $A \in \mathbb{K}^{n \times n}$  hat höchstens  $n$  reelle Eigenwerte und genau  $n$  komplexe Eigenwerte (wenn die Vielfachheit der Nullstelle gezählt wird).

Beachte: Reelle Matrizen können sowohl reelle, als auch komplexe Eigenwerte haben!

**Beispiel.** 1. Für die Matrix  $I_n$  gilt

$$P_{I_n}(\lambda) = \det(I_n - \lambda I_N) = \det((1 - \lambda)I_n) = (1 - \lambda)^n \det(I_n) = (1 - \lambda)^n.$$

Die Matrix hat also nur den Eigenwert 1, diesen mit der Vielfachheit  $n$ . Außerdem ist jeder Vektor  $\vec{v} \neq 0$  ein Eigenvektor und es gilt  $E_1(I_n) = \mathbb{K}^n$ .

2. Die Diagonalmatrix

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

hat das charakteristische Polynom  $p_D(\lambda) = (d_1 - \lambda) \cdots (d_n - \lambda)$  und somit die Eigenwerte  $d_1, \dots, d_n$ . Die Eigenvektoren zu  $d_i$  sind die Vektoren  $ce_i$  mit Faktoren  $c \neq 0$ .

3. Hier eine reelle Matrix, die keine reellen Eigenwerte hat: Wir betrachten die  $2 \times 2$  Drehmatrix

$$D_\beta = \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}.$$

Das charakteristische Polynom ist

$$\begin{aligned} p_{D_\beta}(\lambda) &= \det(D_\beta - \lambda I_2) = \det \begin{pmatrix} \cos(\beta) - \lambda & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) - \lambda \end{pmatrix} \\ &= (\cos(\beta) - \lambda)^2 + \sin(\beta)^2. \end{aligned}$$

Wir setzen dies gleich Null, und rechnen

$$\begin{aligned} (\cos(\beta) - \lambda)^2 + \sin(\beta)^2 &= 0 \\ \iff (\cos(\beta) - \lambda)^2 &= -\sin(\beta)^2 \\ \iff \cos(\beta) - \lambda &= \pm j \sin(\beta) \\ \iff \lambda &= \cos(\beta) \pm j \sin(\beta) = e^{\pm j\beta}. \end{aligned}$$

Die Eigenwerte sind also die beiden komplexen Zahlen  $\lambda_1 = e^{j\beta}$  und  $\lambda_2 = e^{-j\beta}$ .

Jetzt, wo wir die Eigenwerte kennen, können wir die Eigenvektoren berechnen: Dazu müssen wir Vektoren im Kern von  $D_\beta - \lambda_{1/2} I_2$  bestimmen. Für  $\lambda_1$  rechnen wir

$$\begin{aligned} (D_\beta - \lambda_1 I_2) \vec{v}^{(1)} &= \vec{0} \\ \iff \begin{pmatrix} \cos(\beta) - (\cos(\beta) + j \sin(\beta)) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) - (\cos(\beta) + j \sin(\beta)) \end{pmatrix} \vec{v}^{(1)} &= \vec{0} \\ \iff \begin{pmatrix} -j \sin(\beta) & -\sin(\beta) \\ \sin(\beta) & -j \sin(\beta) \end{pmatrix} \vec{v}^{(1)} &= \vec{0} \end{aligned}$$

Wir wissen, dass die Matrix links den Rang 1 hat (sonst wäre  $\lambda_1$  kein Eigenwert) und müssen also nur die erste Zeile betrachten:

$$-j \sin(\beta)v_1^{(1)} - \sin(\beta)v_2^{(1)} = 0.$$

Wir können einen Eintrag von  $\vec{v}^{(1)}$  frei wählen und entscheiden uns für  $v_2^{(1)} = -j$ . Dann folgt für  $v_1^{(1)}$

$$-j \sin(\beta)v_1^{(1)} + j \sin(\beta) = 0,$$

also  $v_1^{(1)} = 1$  und wir bekommen als einen Eigenvektor

$$\vec{v}^{(1)} = \begin{pmatrix} 1 \\ -j \end{pmatrix}$$

(der Eigenraum  $E_{\lambda_1}$  ist der von diesem Vektor aufgespannte Unterraum).

Analog berechnet mal als einen Eigenvektor zu  $\lambda_2$

$$\vec{v}^{(2)} = \begin{pmatrix} 1 \\ j \end{pmatrix}.$$

**Satz 24.6.** Es seien  $\vec{v}_1, \dots, \vec{v}_l$  Eigenvektoren von  $A$  zu verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_l$  (d.h.  $\lambda_k \neq \lambda_m$  für  $k \neq m$ ). Dann sind die Vektoren  $(\vec{v}_k)_k$  linear unabhängig.

*Beweis.*

Um dies zu beweisen, benutzen wir die Technik der vollständigen Induktion:

Für den Fall  $m = 1$  gilt die Behauptung, da ein Eigenvektor  $\vec{v}_1$  per Definition  $\neq 0$  ist und daher linear unabhängig. (Dies war der sogenannte *Induktionsanfang*.)

Zeigen wir jetzt, dass die Behauptung im Fall  $m + 1$  gilt, wenn sie für den Fall  $m$  gilt. (Dies ist der *Induktionsschritt*. Mit ihm können wir vom Induktionsanfang ( $m = 1$ ) auf dem Fall  $m + 1 = 2$  und von immer weiter auf die Fälle  $3, 4, \dots, l$  schließen.) Um etwas Schreibarbeit zu sparen, nehmen wir also an, dass die Behauptung für den Fall  $m - 1$  (anstatt gelte. Wir betrachten die Linearkombination

$$\vec{0} = \sum_{i=1}^m \alpha_i \vec{v}_i. \quad (*)$$

Wir multiplizieren diese Gleichung mit  $A$  und benutzen, dass die  $\vec{v}_i$  Eigenvektoren sind:

$$\vec{0} = \sum_{i=1}^m \alpha_i A \vec{v}_i = \sum_{i=1}^m \alpha_i \lambda_i \vec{v}_i$$

Wir nehmen die Gleichung (\*) mit  $\lambda_m$  mal, ziehen sie von dieser

Gleichung ab, und bekommen

$$\vec{0} = \sum_{i=1}^m \alpha_i (\lambda_i - \lambda_m) \vec{v}_i = \sum_{i=1}^{m-1} \alpha_i (\lambda_i - \lambda_m) \vec{v}_i.$$

Da nach unserer Annahme die Vektoren  $\vec{v}_1, \dots, \vec{v}_{m-1}$  linear unabhängig sind, folgt

$$\alpha_i (\lambda_i - \lambda_m) = 0, \text{ für alle } i = 1, \dots, m-1$$

und da die  $\lambda_i$  alle verschieden sind, folgt  $\alpha_i = 0$  für  $i = 1, \dots, m-1$ . Damit wird (\*) zu  $\vec{0} = \alpha_m \vec{v}_m$  und wegen  $\vec{v}_m \neq \vec{0}$  folgt auch  $\alpha_m = 0$ .  $\square$

Schauen wir uns noch einmal das Beispiel vom Anfang des Abschnitts an: Die Matrix

$$A = \frac{1}{4} \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

hat das charakteristische Polynom  $p_A(\lambda) = (\frac{5}{4} - \lambda)^2 - (\frac{3}{4})^2$  und damit die Nullstelle  $\lambda = \frac{5}{4} \pm \frac{3}{4}$ , also  $\lambda_1 = 2, \lambda_2 = \frac{1}{2}$ . Also zugehörige Eigenvektoren können wir zum Beispiel die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

nehmen. Jetzt können wir verstehen, wie sich die Iteration verhält: Da unsere Eigenvektoren eine Basis bilden, können wir jeden Anfangswert  $\vec{x}^{(0)}$  darstellen als

$$\vec{x}^{(0)} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2.$$

Es folgt

$$\begin{aligned} \vec{x}^{(k)} &= A \vec{x}^{(k-1)} = AA \vec{x}^{(k-2)} = A \cdots A \vec{x}^{(0)} \\ &= A \cdots A (\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2) \\ &= A \cdots A (\alpha_1 A \vec{v}_1 + \alpha_2 A \vec{v}_2) \\ &= A \cdots A (\alpha_1 \lambda_1 \vec{v}_1 + \alpha_2 \lambda_2 \vec{v}_2) \\ &= \alpha_1 \lambda_1^k \vec{v}_1 + \alpha_2 \lambda_2^k \vec{v}_2. \end{aligned}$$

Da  $\lambda_1 = 2$  und  $\lambda_2 = \frac{1}{2}$  gilt, sehen wir, dass der Anteil von  $\vec{x}^{(0)}$ , welcher zu  $\vec{v}_2$  gehört, gegen Null geht und die Folge  $\vec{x}^{(k)}$  sich tatsächlich einem Vielfachen von  $\vec{v}_1$  nähert.

## 25 Diagonalisierbarkeit

Eigenwerte haben etwas mit dem Normalformenproblem für Selbstabbildungen zu tun! Das Normalformenproblem für quadratische Matrizen  $A \in \mathbb{K}^{n \times n}$  ist: Finde eine Basis  $\vec{v}_1, \dots, \vec{v}_n$  des  $\mathbb{K}^n$ , so dass die Matrixdarstellung der Abbildung  $T(\vec{x}) = A\vec{x}$  bezüglich dieser Basis möglichst einfach ist.

Betrachten wir nun den Fall, dass es zu  $A$  eine Basis  $(\vec{v}_1, \dots, \vec{v}_n)$  aus Eigenvektoren gibt. Dann gilt ja  $A\vec{v}_i = \lambda_i\vec{v}_i$  und wenn wir die  $\vec{v}_i$  spaltenweise in eine Matrix  $S \in \mathbb{K}^{n \times n}$  schreiben, bekommen wir

$$AS = A(\vec{v}_1 \ \cdots \ \vec{v}_n) = (A\vec{v}_1 \ \cdots \ A\vec{v}_n) = (\lambda_1\vec{v}_1 \ \cdots \ \lambda_n\vec{v}_n).$$

Die rechte Seite lässt sich mit Hilfe einer Diagonalmatrix auch schreiben als

$$(\lambda_1\vec{v}_1 \ \cdots \ \lambda_n\vec{v}_n) = (\vec{v}_1 \ \cdots \ \vec{v}_n) \underbrace{\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}{=:B}.$$

Wir haben in diesem Fall also  $AS = SB$  und da  $S$  invertierbar ist, haben wir

$$B = S^{-1}AS.$$

In diesem Fall gibt es also eine Basis (nämlich eine aus Eigenvektoren), so dass  $A$  bezüglich dieser Basis als Diagonalmatrix dargestellt wird!

**Definition 25.1.** Eine quadratische Matrix  $A \in \mathbb{K}^{n \times n}$  heißt *diagonalisierbar*, wenn es eine invertierbare Matrix  $S \in \mathbb{K}^{n \times n}$  gibt, so dass  $B = S^{-1}AS$  eine Diagonalmatrix ist.

Unsere Argumentation oben funktioniert auch andersherum und wir können festhalten: Ist  $A$  diagonalisierbar, so bilden die Spalten von  $S$  eine Basis aus Eigenvektoren und die zugehörigen Eigenwerte stehen auf der Diagonalen von  $B$ .

Aus Satz 24.6 können wir ein einfaches Kriterium herleiten, was Diagonalisierbarkeit garantiert:

**Satz 25.2.** *Hat eine Matrix  $A \in \mathbb{K}^{n \times n}$   $n$  verschiedenen Eigenwerte, so ist sie diagonalisierbar.*

Die Umkehrung gilt nicht: Die Einheitsmatrix  $I_n$  ist diagonalisierbar (selbst diagonal), hat aber nur den einen Eigenwert 1.

Hier ein einfaches Beispiel einer nicht-diagonalisierbaren Matrix:

Wir haben übrigens gezeigt: Eine Matrix ist genau dann diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.

*Beispiel.* Wir betrachten die Matrix

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

Das charakteristische Polynom ist

$$p_A(\lambda) = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2$$

und daher ist 1 der einzige Eigenwert. Bestimmen wir nun den zugehörigen Eigenraum, d.h., wir suchen  $\vec{v}$  so, dass  $(A - I_2)\vec{v} = \vec{0}$ , also den Kern von

$$(A - \lambda I_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Wir sehen, dass der Kern vom Vektor  $\vec{e}_1$  aufgespannt wird und die Dimension 1 hat. Da es keine weiteren Eigenwerte gibt, kann es keine Basis aus Eigenvektoren geben und  $A$  ist nicht diagonalisierbar.

**Definition 25.3.** Es sei  $\lambda$  ein Eigenwert von  $A$ .

- (a) Die Vielfachheit von  $\lambda$  als Nullstelle von des charakteristischen Polynoms  $p_A$  heißt *algebraische Vielfachheit* von  $\lambda$ , geschrieben als  $a(\lambda)$ .
- (b) Die Dimension des Eigenraumes  $E_\lambda(A)$  heißt *geometrische Vielfachheit* von  $\lambda$ , geschrieben  $g(\lambda)$ .

Wir erinnern daran, dass zwei Matrizen  $A$  und  $B$  ähnlich heißen, wenn es ein  $S$  gibt, so dass  $B = S^{-1}AS$ .

**Satz 25.4.** (a) Sind zwei Matrizen ähnlich, so haben sie das gleiche charakteristische Polynom, also auch die gleichen Eigenwerte mit den gleichen algebraischen Vielfachheiten.

(b) Sind zwei Matrizen ähnlich, so sind auch die geometrischen Vielfachheiten gleich. Genauer: Ist  $B = S^{-1}AS$  so gilt:

$$\vec{v} \text{ ist EV von } A \text{ zum EW } \lambda \iff S^{-1}\vec{v} \text{ ist EV von } B \text{ zu EW } \lambda.$$

**Beweis.**

(a) Wegen  $I_n = S^{-1}S$  folgt

$$\begin{aligned} p_B(\lambda) &= \det(B - \lambda I_n) = \det(S^{-1}AS - \lambda S^{-1}S) \\ &= \det(S^{-1}(A - \lambda I_n)S) = \underbrace{\det(S^{-1})}_{=1/\det(S)} \det(A - \lambda I_n) \det(S) \\ &= \det(A - \lambda I_n) = p_A(\lambda). \end{aligned}$$

(b) Wir multiplizieren die Eigenwertgleichung für  $\vec{v}$  mit  $S^{-1}$  und

Mit anderen Worten: Das charakteristische Polynom, die Eigenwerte und deren algebraische und geometrische Vielfachheit sind invariant gegenüber Koordinatentransformationen. Die Eigenvektoren hingegen nicht.

nutzen wieder  $SS^{-1} = I_n$ :

$$\begin{aligned} A\vec{v} = \lambda\vec{v} &\iff S^{-1}ASS^{-1}\vec{v} = \lambda S^{-1}\vec{v} \\ &\iff B(S^{-1}\vec{v}) = \lambda(S^{-1}\vec{v}). \end{aligned}$$

□

**Satz 25.5.** Für einen Eigenwert  $\lambda_0$  einer  $n \times n$ -Matrix  $A$  gilt immer

$$1 \leq g(\lambda_0) \leq a(\lambda_0) \leq n.$$

*Beweis.*

Die Ungleichung  $1 \leq g(\lambda_0)$  folgt, da die Matrix  $A - \lambda_0 I_n$  nicht invertierbar ist, und daher einen Kern mit Dimension  $\geq 1$  hat. Die Ungleichung  $a(\lambda_0) \leq n$  folgt daher, dass das charakteristische Polynom den Grad  $n$  hat. Für die mittlere Ungleichung betrachten wir eine Basis  $(\vec{v}_1, \dots, \vec{v}_k)$  des Eigenraums  $E_{\lambda_0}(A)$  (d.h. wir haben angenommen, dass  $g(\lambda_0) = k$  gilt). Wir ergänzen diese zu einer Basis  $(\vec{v}_1, \dots, \vec{v}_n)$  des ganzen  $\mathbb{K}^n$  und schreiben diese Basis spaltenweise in  $S$ . Mit dieser Matrix gilt dann für ein  $R$

$$\begin{aligned} AS &= (\lambda_0 \vec{v}_1 \quad \dots \quad \lambda_0 \vec{v}_k \quad A\vec{v}_{k+1} \quad \dots \quad A\vec{v}_n) \\ &= S \left( \begin{array}{ccc|ccc} \lambda_0 & & 0 & & & \\ & & \ddots & & & \\ & & & \lambda_0 & & \\ \hline & & & & & R \\ & & 0 & & & \end{array} \right) \end{aligned}$$

also gilt

$$B := S^{-1}AS = \left( \begin{array}{ccc|ccc} \lambda_0 & & 0 & & & \\ & & \ddots & & & \\ & & & \lambda_0 & & \\ \hline & & & & & R \\ & & 0 & & & \end{array} \right).$$

Die Matrix  $B$  hat nach dem vorigen Satz das gleiche charakteristische Polynom wie  $A$ , d.h. wir haben

$$p_A(\lambda) = p_B(\lambda) = (\lambda - \lambda_0)^k \det(R - \lambda I_{n-k}).$$

Damit ist  $\lambda_0$  eine Nullstelle mit Vielfachheit mindestens  $k$  was die Behauptung zeigt. □

**Satz 25.6.** (a) Im komplexen gilt: Eine Matrix  $A \in \mathbb{C}^{n \times n}$  ist genau dann diagonalisierbar, wenn für alle Eigenwerte die geometrische und die algebraische Vielfachheit übereinstimmt.

(b) *Im reellen Fall gilt: Ist  $A$  reell und diagonalisierbar und sind alle Eigenwerte reell, dann sind auch alle Eigenvektoren reell.*

*Beispiel.* Wir betrachten die Matrix

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

und berechnen das charakteristische Polynom durch Entwicklung der Determinante nach der letzten Spalte

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I_3) = \det \begin{pmatrix} 3 - \lambda & 2 & 0 \\ -3 & -2 - \lambda & 0 \\ -3 & -3 & 1 - \lambda \end{pmatrix} \\ &= (1 - \lambda) \det \begin{pmatrix} 3 - \lambda & 2 \\ -3 & -2 - \lambda \end{pmatrix} \\ &= (1 - \lambda) ((3 - \lambda)(-2 - \lambda) - 2(-3)) = (1 - \lambda) (-6 - \lambda + \lambda^2 + 6) \\ &= (1 - \lambda)(-\lambda + \lambda^2) = (1 - \lambda)\lambda(-1 + \lambda) = -(1 - \lambda)^2\lambda. \end{aligned}$$

Wir können die Nullstellen und die Vielfachheiten jetzt ablesen:

$$\begin{aligned} \lambda_1 &= 0, \quad a(\lambda_1) = 1 \\ \lambda_2 &= 1, \quad a(\lambda_2) = 2. \end{aligned}$$

Bestimmen wir jetzt Eigenräume:

$$E_0(A) = \text{Kern}(A - 0I_3) = \text{Kern}(A).$$

Mit dem Eliminationsalgorithmus bekommen wir:

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

und wir bekommen einen Vektor  $\vec{v}^1$  im Kern (d.h. einen Eigenvektor zu Eigenwert 0) in dem wir  $v_3^1 = \sigma$  setzen und aus der letzten Zeile  $v_2^1 = \sigma$  und aus der ersten Zeile  $v_1^1 = -\frac{2}{3}\sigma$  bekommen. Für einen ganzzahligen Eigenvektor wählen wir  $\sigma = 3$ , also

$$\vec{v}^1 = \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix}.$$

Für den Eigenwert  $\lambda = 1$  gehen wir entsprechend vor und bestimmen  $\vec{v}^2$  so, dass  $(A - I_3)\vec{v}^2 = \vec{0}$ :

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

Die Lösungsmenge ist zweidimensional, und besteht aus den Vektoren der Form

$$\begin{pmatrix} -\tau \\ \tau \\ \sigma \end{pmatrix}$$

und wir wählen

$$\vec{v}_1^2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Insgesamt ist  $(\vec{v}^1, \vec{v}_1^2, \vec{v}_2^2)$  eine Basis des  $\mathbb{R}^3$  aus Eigenvektoren und wir schreiben sie spaltenweise in die Transformationsmatrix

$$S = (\vec{v}^1, \vec{v}_1^2, \vec{v}_2^2) = \begin{pmatrix} -2 & -1 & 0 \\ 3 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

Die Inverse hierzu berechnet man als

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

Wir überprüfen noch

$$\begin{aligned} S^{-1}AS &= \begin{pmatrix} 1 & 1 & 0 \\ -3 & -2 & 0 \\ -3 & -3 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 0 \\ -3 & -2 & 0 \\ -3 & -3 & 1 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 3 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 \\ -3 & -2 & 0 \\ -3 & -3 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

## 26 Symmetrische/hermitesche Matrizen und die Hauptachsentransformation

Erinnere: Das komplexe Standardskalarprodukt und die zugehörige Norm sind

$$\langle \vec{u}, \vec{v} \rangle = \sum_{i=1}^n u_i^* v_i, \quad \|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{\sum_{i=1}^n |v_i|^2}.$$

Für  $\vec{v} \in \mathbb{C}^n$  bezeichnen wir (evtl. etwas missverständlich) mit  $\vec{v}^* \in \mathbb{C}^{1 \times n}$  den konjugiert-transponierten Vektor und mit  $A^*$  die konjugiert-transponierte Matrix. Wir nennen  $A$  *hermitesch*, wenn  $A^* = A$  gilt und *unitär*, wenn  $A^*A = AA^* = I_n$  gilt.

Aus der Definition des Skalarproduktes und der konjugiert-transponierten Matrix folgt

$$\begin{aligned} \langle \vec{u}, A\vec{v} \rangle &= \vec{u}^*(A\vec{v}) = (\vec{u}^*A)\vec{v} = (A^*\vec{u})^*\vec{v} \\ &= \langle A^*\vec{u}, \vec{v} \rangle. \end{aligned}$$

Ist  $A$  *reell* und hermitesch, so ist  $A$  symmetrisch.

**Satz 26.1.** Für eine hermitesche Matrix  $A \in \mathbb{C}^{n \times n}$  gilt:

- (a) Die Eigenwerte von  $A$  sind reell.  
 (b) Eigenvektoren von  $A$  zu verschiedenen Eigenwerten sind orthogonal.

*Beweis.*

- (a) Ist  $\vec{v} \in \mathbb{C}^n$  ein Eigenvektor zum Eigenwert  $\lambda \in \mathbb{C}$ , so gilt

$$\begin{aligned} \lambda \|\vec{v}\|^2 &= \lambda \langle \vec{v}, \vec{v} \rangle = \langle \vec{v}, \lambda \vec{v} \rangle = \langle \vec{v}, A\vec{v} \rangle \\ &= \langle A^*\vec{v}, \vec{v} \rangle = \langle A\vec{v}, \vec{v} \rangle = \langle \lambda \vec{v}, \vec{v} \rangle = \lambda^* \langle \vec{v}, \vec{v} \rangle \\ &= \lambda^* \|\vec{v}\|^2 \end{aligned}$$

und da  $\vec{v} \neq \vec{0}$  gilt, folgt  $\lambda = \lambda^*$ , also  $\lambda \in \mathbb{R}$ .

- (b) Es gelte  $A\vec{v} = \lambda\vec{v}$  und  $A\vec{w} = \mu\vec{w}$  mit  $\lambda \neq \mu$  und  $\vec{v}, \vec{w} \neq \vec{0}$ . Mit (a) folgt (da  $\lambda, \mu$  reell)

$$\begin{aligned} (\lambda - \mu) \langle \vec{v}, \vec{w} \rangle &= \langle \lambda \vec{v}, \vec{w} \rangle - \langle \vec{v}, \mu \vec{w} \rangle \\ &= \langle A\vec{v}, \vec{w} \rangle - \langle \vec{v}, A\vec{w} \rangle = \langle \vec{v}, A^*\vec{w} \rangle - \langle \vec{v}, A\vec{w} \rangle \\ &= \langle \vec{v}, (A^* - A)\vec{w} \rangle = 0. \end{aligned}$$

Da  $\lambda - \mu \neq 0$  gilt, folgt  $\langle \vec{v}, \vec{w} \rangle = 0$ .

□

**Satz 26.2** (Hauptachsentransformation). Jede hermitesche Matrix  $A \in \mathbb{C}^{n \times n}$  ist unitär diagonalisierbar, d.h. es existiert eine unitäre Matrix  $S$  (deren Spalten Eigenvektoren enthält), so dass  $S^{-1}AS = \Lambda$  mit reeller Diagonalmatrix  $\Lambda$  (auf deren Diagonale die Eigenwerte von  $A$  stehen).

Da  $S$  unitär ist, gilt  $S^{-1} = S^*$  und wir könnten auch  $S^*AS = \Lambda$  schreiben.

Ein paar Bemerkungen zu dem Satz:

1. Aus dem Satz zur Hauptachsentransformation folgt im reellen Fall: Ist  $A \in \mathbb{R}^{n \times n}$  symmetrisch, so ist  $A$  orthonormal diagonalisierbar, d.h. es existiert eine orthonormale Matrix  $S \in \mathbb{R}^{n \times n}$ , so dass  $S^T A S = \Lambda$  gilt mit reeller Diagonalmatrix  $\Lambda$ .

Da wir einen normierten Eigenvektor  $\vec{v}$  als Spalte von  $S$  auch durch  $-\vec{v}$  ersetzen können, können wir sogar  $\det(S) = 1$  fordern, d.h. die Matrix  $S$  kann immer als Drehung gewählt werden.

2. Der Satz zur Hauptachsentransformation bleibt gültig, wenn  $A$  nur  $A^* A = A A^*$  erfüllt (solche Matrizen nennt man *normal*). Allerdings müssen die Eigenwerte von  $A$  in diesem Fall nicht mehr reell sein.

Mit Hilfe der Hauptachsentransformation lassen sich quadratische Polynome in mehreren Variablen untersuchen:

**Definition 26.3.** Eine Funktion  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  der Form

$$q(\vec{x}) = \vec{x}^T A \vec{x} + \vec{b}^T \vec{x} + c = \sum_{i,j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c$$

mit  $A \in \mathbb{R}^{n \times n}$ ,  $b \in \mathbb{R}^n$  und  $c \in \mathbb{R}$  heißt *quadratisches Polynom* in den Variablen  $x_1, \dots, x_n$ .

Wegen der Symmetrie des Skalarproduktes gilt

$$\vec{x}^T A \vec{x} = \langle \vec{x}, A \vec{x} \rangle = \langle A^T \vec{x}, \vec{x} \rangle = \langle \vec{x}, A^T \vec{x} \rangle = \vec{x}^T A^T \vec{x}.$$

Es folgt

$$\vec{x}^T A \vec{x} = \vec{x}^T \left( \frac{A+A^T}{2} \right) \vec{x}$$

und daher können wir ohne Einschränkung annehmen, dass die Matrix  $A$  in einem quadratischen Polynom symmetrisch ist.

Die Nullstellenmenge eines quadratischen Polynoms, also die Menge

$$\{ \vec{x} \in \mathbb{R}^n \mid q(\vec{x}) = 0 \},$$

nennt man *Quadrik*.

*Beispiel.* Wir betrachten

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}, \quad \vec{b} = \vec{0}, \quad c = -1,$$

d.h. das quadratische Polynom ist

$$q(\vec{x}) = \alpha x_1^2 + 2\beta x_1 x_2 + \gamma x_2^2 - 1.$$

Man kann übrigens jede quadratische Matrix  $A$  in eine Summe aus einer symmetrischen und einer schiefsymmetrischen Matrix zerlegen (eine Matrix  $B$  heißt schiefsymmetrisch, oder auch antisymmetrisch, wenn  $B^T = -B$  gilt). Die Zerlegung ist einfach  $A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$ .

Wir untersuchen die Eigenwerte der Matrix. Diese hat die beiden Eigenwerte

$$\lambda_{1/2} = \frac{\alpha+\gamma}{2} \pm \sqrt{\frac{(\alpha-\gamma)^2}{4} + \beta^2}.$$

Die beiden zugehörigen Eigenvektoren  $\vec{w}^1$  und  $\vec{w}^2$  rechnen wir nicht aus, doch wir wissen, dass sie orthonormal gewählt werden können. Wir setzen  $S = (\vec{w}^1 \quad \vec{w}^2)$  und  $\vec{z} = S^{-1}\vec{x} = S^T\vec{x}$ , also gilt  $\vec{x} = S\vec{z}$ . Damit ist

$$q(\vec{x}) = \vec{x}^T \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \vec{x} - 1 = \vec{z}^T S^T A S \vec{z} - 1 = \vec{z}^T S^{-1} A S \vec{z} - 1$$

und wegen  $S^{-1}AS = \text{diag}(\lambda_1, \lambda_2)$ , folgt

$$q(\vec{x}) = \vec{z}^T \text{diag}(\lambda_1, \lambda_2) \vec{z} - 1 = \lambda_1 z_1^2 + \lambda_2 z_2^2 - 1.$$

Die Transformation von  $q$  durch die Substitution  $\vec{x} \mapsto \vec{z}$  nennt man Hauptachsenstransformation des Polynoms.

Wir interpretieren die Situation geometrisch, indem wir die Quadrik  $\{\vec{x} \in \mathbb{R}^2 \mid q(\vec{x}) = 0\}$  betrachten. Diese lautet  $ax_1^2 + bx_1x_2 + cx_2^2 - 1 = 0$  und es ist auf den ersten Blick nicht klar, welche geometrische Form sie hat. In den neuen Koordinaten  $\vec{z}$  wird diese zu

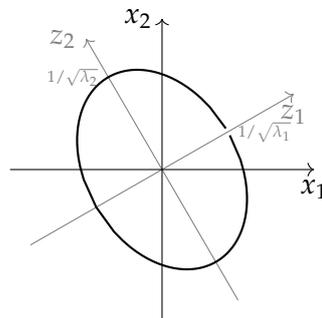
$$\lambda_1 z_1^2 + \lambda_2 z_2^2 - 1 = 0$$

und dies nennen wir die Normalform der Quadrik. Formales Auflösen nach  $z_2$  gibt

$$z_2 = \pm \sqrt{\frac{1}{\lambda_2} - \frac{\lambda_1}{\lambda_2} z_1^2}.$$

Die Form der Lösungsmenge hängt von den Vorzeichen von  $1/\lambda_2$  und  $\lambda_1/\lambda_2$  ab: Es gilt:

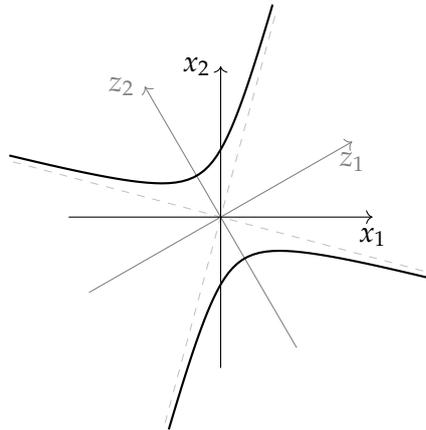
1.  $\lambda_1, \lambda_2 > 0$ : Hier ist der Ausdruck unter der Wurzel nur für  $|z_1| \leq 1/\sqrt{\lambda_1}$  nicht negativ und  $C$  ist eine Ellipse:



2.  $\lambda_1 < 0, \lambda_2 > 0$ : Hier ist der Ausdruck unter der Wurzel immer positiv und  $C$  ist eine Hyperbel

Wir können ein gegebenes Polynom  $p(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2 + c$  immer wie folgt in Matrix-Schreibweise ausdrücken:

$$p(x_1, x_2) = x^T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} x + c.$$



(Im Fall  $\lambda_1 > 0$  und  $\lambda_2 < 0$  ist die Hyperbel in die andere Richtung offen.)

3.  $\lambda_1 > 0, \lambda_2 = 0$ :  $C$  besteht aus den beiden Geraden  $z_1 = \pm\sqrt{\lambda_1}^{-1}$
4.  $\lambda_1, \lambda_2 < 0$ : Hier ist  $C$  leer.

Da eine symmetrische Matrix nur reelle Eigenwerte hat, können wir definieren:

**Definition 26.4.** Eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt:

- positiv definit :  $\iff$  alle Eigenwerte  $> 0$
- positiv semidefinit :  $\iff$  alle Eigenwerte  $\geq 0$
- negativ definit :  $\iff$  alle Eigenwerte  $< 0$
- negativ semidefinit :  $\iff$  alle Eigenwerte  $\leq 0$
- indefinit :  $\iff$  Es gibt Eigenwerte  $> 0$  und  $< 0$ .

**Satz 26.5.** Für eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  sind äquivalent:

- (a)  $A$  ist positiv definit.
- (b) Für alle  $\vec{x} \in \mathbb{R}^n, \vec{x} \neq 0$  gilt  $\vec{x}^T A \vec{x} > 0$ .
- (c) Für alle  $k = 1, \dots, n$  gilt

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0$$

- (d)  $A$  hat eine Cholesky-Zerlegung  $A = LDL^T$  mit einer normierten unteren Dreiecksmatrix  $L$  und eine Diagonalmatrix  $D$  mit positiven Einträgen auf der Diagonale.

## 27 Die Jordansche Normalform

Da nicht alle Matrizen diagonalisierbar sind, stellt sich die Frage nach einer allgemeinen Normalform für quadratische Matrizen. Was ist also die einfachst mögliche Form für  $S^{-1}AS$  für eine gegebene quadratische Matrix  $A$ ? Die Antwort auf diese Frage gibt die *Jordansche Normalform*.

**Definition 27.1.** Eine Matrix der Form

$$J = \begin{pmatrix} \lambda & 1 & & & 0 \\ 0 & \lambda & 1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}$$

heißt *Jordankasten* zum Wert  $\lambda$  (es ist also  $j_{ii} = \lambda$  für ein  $\lambda$ ,  $j_{i,i+1} = 1$  und  $j_{kl} = 0$  in allen anderen Fällen).

**Satz 27.2.** Zu jeder Matrix  $A \in \mathbb{K}^{n \times n}$  existiert eine invertierbare (ggf. komplexe) Matrix  $S$ , so dass

$$S^{-1}AS = J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_r \end{pmatrix}, \quad \text{wobei } J_i \text{ Jordankasten zum Wert } \lambda_j.$$

Die Matrix  $J$  heißt *Jordansche Normalform* von  $A$ . Sie ist, bis auf die Reihenfolge der Jordankästen, eindeutig bestimmt. Dabei können die  $J_i$  verschiedene Größen haben und es kann mehrere Kästen zum gleichen Wert  $\lambda$  geben.

Ein paar Bemerkungen:

- Da  $A$  und  $J$  ähnlich sind, haben sie die gleichen Eigenwerte und daher sind die Werte  $\lambda_1, \dots, \lambda_r$  die Eigenwerte von  $A$ .
- Die algebraische Vielfachheit des Eigenwertes  $\lambda_i$  ergibt sich als die Summen der Größen der Jordankästen zum Wert  $\lambda_i$ .
- Die geometrische Vielfachheit von  $\lambda_i$  hingegen ist die Anzahl der Jordankästen zum Wert  $\lambda_i$ .

Wie kann man die Jordansche Normalform  $J$  einer Matrix  $A$  bestimmen und die Transformationsmatrix  $S$  gleich dazu? Schauen wir uns dazu den Sonderfall an, dass  $J$  aus nur einem Jordankasten besteht, d.h. es gilt  $S^{-1}AS = J$  und ist  $S = (\vec{v}_1 \ \dots \ \vec{v}_n)$  so ist

$$\begin{aligned} AS &= (A\vec{v}_1 \ \dots \ A\vec{v}_n) \\ &= SJ = (\vec{v}_1 \ \dots \ \vec{v}_n) \begin{pmatrix} \lambda & 1 & & 0 \\ 0 & \ddots & \ddots & \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda \end{pmatrix} \\ &= (\lambda\vec{v}_1 \ \vec{v}_1 + \lambda\vec{v}_2 \ \dots \ \vec{v}_{n-1} + \lambda\vec{v}_n). \end{aligned}$$

Wir haben also

$$\left. \begin{array}{l} A\vec{v}_1 = \lambda\vec{v}_1 \\ A\vec{v}_2 = \vec{v}_1 + \lambda\vec{v}_2 \\ \vdots \\ A\vec{v}_n = \vec{v}_{n-1} + \lambda\vec{v}_n \end{array} \right\} \implies \left\{ \begin{array}{l} (A - \lambda I_n)\vec{v}_1 = \vec{0} \\ (A - \lambda I_n)\vec{v}_2 = \vec{v}_1 \\ \vdots \\ (A - \lambda I_n)\vec{v}_n = \vec{v}_{n-1} \end{array} \right.$$

Wir erkennen, dass  $\vec{v}_1$  ein Eigenvektor zum Eigenwert  $\lambda$  ist. Außerdem bekommen wir durch Einsetzen für  $k = 2, \dots, n$

$$(A - \lambda I_n)^k \vec{v}_k = \vec{0}, \quad \text{sowie} \quad (A - \lambda I_n)^{k-1} \vec{v}_k \neq \vec{0}.$$

Die motiviert folgende Definition:

**Definition 27.3.** Es sei  $A \in \mathbb{K}^{n \times n}$  und  $\lambda_i$  ein Eigenwert von  $A$ .

(a) Wir nennen  $\vec{v}_i^{(k)}$  einen *Hauptvektor  $k$ -ter Stufe* zum Eigenwert  $\lambda_i$ , wenn gilt

$$(A - \lambda_i I_n)^k \vec{v}_i^{(k)} = \vec{0}, \quad (A - \lambda_i I_n)^{k-1} \vec{v}_i^{(k-1)} \neq \vec{0}.$$

Wir können schließen, dass  $(A - \lambda I_n)^n \vec{v}_k = \vec{0}$  für alle  $k$  und da die  $\vec{v}_k$  eine Basis des  $\mathbb{K}^n$  bilden, folgt  $(A - \lambda I_n)^n = 0$ .

Ein Hauptvektor der ersten Stufe ist ein Eigenvektor!

(b) Eine endliche Folge  $\vec{v}_i^{(1)}, \dots, \vec{v}_i^{(m)}$  von Hauptvektoren heißt *Hauptvektorkette*, wenn gilt

$$(A - \lambda I_n)\vec{v}_i^{(1)} = \vec{0}, \quad (A - \lambda I_n)\vec{v}_i^{(2)} = \vec{v}_i^{(1)}, \quad \dots \quad (A - \lambda I_n)\vec{v}_i^{(m)} = \vec{v}_i^{(m-1)}$$

und wenn die Folge nicht verlängerbar ist, d.h. wenn  $(A - \lambda I_n)\vec{v} = \vec{v}_i^{(m)}$  keine Lösung hat. Wir nennen dann  $\vec{v}_i^{(m)}$  den Hauptvektor *höchster Stufe*.

Die Definition zeigt uns auch, wie wir im allgemeinen Fall der Jordanschen Normalform die Transformationsmatrix  $S$  bestimmen können: Es ist nämlich

$$S = \left( \underbrace{\vec{v}_1^{(1)} \dots \vec{v}_1^{(m_1)}}_{\text{Kette zu } \lambda_1}, \dots, \underbrace{\vec{v}_i^{(1)} \dots \vec{v}_i^{(m_i)}}_{\text{Kette zu } \lambda_i}, \dots, \underbrace{\vec{v}_r^{(1)} \dots \vec{v}_r^{(m_r)}}_{\text{Kette zu } \lambda_r} \right),$$

Hier das Verfahren zur Konstruktion der Jordanschen Normalform: Gegeben sei  $A \in \mathbb{K}^{n \times n}$ .

1. Bestimme die Eigenwerte  $\lambda_1, \dots, \lambda_s$  von  $A$  und die algebraischen Vielfachheiten  $a(\lambda_1), \dots, a(\lambda_s)$ .
2. Berechne zu jedem Eigenwert eine Basis  $(\vec{v}_{i,1}, \dots, \vec{v}_{i,g(\lambda_i)})$  des zugehörigen Eigenraumes (d.h. bestimme eine Basis von  $\text{Kern}(A - \lambda_i I_n)$ ). (Die Anzahl der Basisvektoren ist die geometrische Vielfachheit.)

3. Ist  $a(\lambda_i) = g(\lambda_i)$ , so ist für dieses  $i$  nichts mehr zu tun. Gilt  $>$ , so konstruiere man zu jedem  $\vec{v}_{i,k}$  ( $k = 1, \dots, g(\lambda_i)$ ) die zugehörige Hauptvektorenkette  $\vec{v}_{i,k}^{(l)}$ .
4. Alle die Vektoren aus den Hauptvektorenketten bilden dann die Spalten der Matrix  $S$ .

*Beispiel.* Wir betrachten die Matrix

$$A = \begin{pmatrix} -1 & 1 & 1 & -3 & -3 & -4 & 0 \\ 1 & -2 & -1 & 3 & 4 & 4 & 0 \\ -1 & 1 & 0 & -6 & -8 & -8 & 0 \\ 0 & 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 3 & 4 & 0 \\ 0 & 0 & 0 & -3 & -1 & -2 & 0 \\ 0 & 0 & 0 & 3 & 1 & 4 & 2 \end{pmatrix}.$$

Mit etwas Aufwand lässt sich (unter Ausnutzung der Null-Blöcke) das charakteristische Polynom berechnen:

$$p_A(\lambda) = -(\lambda - 2)^3(\lambda + 1)^4.$$

Wir haben also die beiden Eigenwerte  $\lambda_1 = 2$  und  $\lambda_2 = -1$ . Berechnen wir Basen der Eigenräume. Es gilt

$$\text{Kern}(A - 2I_7) = \text{Kern} \begin{pmatrix} -3 & 1 & 1 & -3 & -3 & -4 & 0 \\ 1 & -4 & -1 & 3 & 4 & 4 & 0 \\ -1 & 1 & -2 & -6 & -8 & -8 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 & 4 & 0 \\ 0 & 0 & 0 & -3 & -1 & -4 & 0 \\ 0 & 0 & 0 & 3 & 1 & 4 & 0 \end{pmatrix}.$$

Der Rang dieser Matrix ist 5, d.h. es gibt zwei linear unabhängige Eigenvektoren und wir nehmen

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ -2 \\ 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Jetzt berechnen wir die Hauptvektorketten: Für  $\vec{v}_1$  bekommen wir als eine mögliche Lösung von  $(A - 2I_7)\vec{v}_{1,2} = \vec{v}_1$  den Vektor

$$\vec{v}_{1,2} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Die Hauptvektorkette zu  $\vec{v}_2$  bricht sofort ab, da  $(A - 2I_7)\vec{w} = \vec{v}_2$  keine Lösung hat.

Der Eigenraum zu  $\lambda_2 = -1$  ist

$$\text{Kern}(A + I_7) = \text{Kern} \begin{pmatrix} 0 & 1 & 1 & -3 & -3 & -4 & 0 \\ 1 & -1 & -1 & 3 & 4 & 4 & 0 \\ -1 & 1 & 1 & -6 & -8 & -8 & 0 \\ 0 & 0 & 0 & 3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 4 & 4 & 0 \\ 0 & 0 & 0 & -3 & -1 & -1 & 0 \\ 0 & 0 & 0 & 3 & 1 & 4 & 3 \end{pmatrix}.$$

Wieder ist der Rang 5 und wir bekommen die beiden linear unabhängigen Eigenvektoren

$$\vec{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_4 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

Berechnen wir die Hauptraumketten: Zu  $\vec{v}_3$  bekommen wir aus  $(A + I_7)\vec{v}_{3,2} = \vec{v}_3$  zum Beispiel

$$\vec{v}_{3,2} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

und aus  $(A + I_7)\vec{v}_{3,3} = \vec{v}_{3,2}$  auch noch

$$\vec{v}_{3,3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

wonach die Hauptvektorkette abbricht. Die letzte Hauptvektorkette zu  $\vec{v}_4$  müssen wir nicht mehr ausrechnen, da wir mit

$$\{\vec{v}_1, \vec{v}_{1,2}, \vec{v}_2, \vec{v}_3, \vec{v}_{3,2}, \vec{v}_{3,3}, \vec{v}_4\}$$

schon eine Basis des  $\mathbb{R}^7$  gefunden haben. Wir setzen

$$S = (\vec{v}_1 \quad \vec{v}_{1,2} \quad \vec{v}_2 \quad \vec{v}_3 \quad \vec{v}_{3,2} \quad \vec{v}_{3,3} \quad \vec{v}_4)$$

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

und man kann nachrechnen

$$S^{-1}AS = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$

was die gesuchte Jordansche Normalform ist.

Wir lesen noch die algebraischen und geometrischen Vielfachheiten ab:

- Es gibt zwei Jordankästen zum Eigenwert 2 und daher ist die geometrische Vielfachheit  $g(2) = 2$ .
- Es gibt zwei Jordankästen zum Eigenwert  $-1$  und daher ist die geometrische Vielfachheit  $g(-1) = 2$ .
- Die Summe der Größen der Jordankästen zum Eigenwert 2 ist 3 und daher ist die algebraische Vielfachheit  $a(2) = 3$ .
- Der Summe der Größen der Jordankästen zum Eigenwert  $-1$  ist 4 und daher ist die algebraische Vielfachheit  $a(-1) = 4$ .