Lineare Algebra für Informatiker und Statistiker

vorläufiges Vorlesungsskript WS 2009/2010

Michael Prähofer
Ludwig-Maximilians-Universität München

Stand:

12. Februar 2010

Inhaltsverzeichnis

0 Einführung **3**
0.1 Wozu Mathematik? . . . 3
0.2 Was macht (man mit) Mathematik? . . . 3
0.3 Beispiele für lineare Probleme . . . 4
0.4 Lineare Gleichungssysteme . . . 7
0.5 Gaußsches Eliminationsverfahren . . . 9
0.6 Rechnen mit Vektoren des $\mathbb{R}^n$ . . . 14
0.7 Rechnen mit Matrizen . . . 16

1 Algebraische Grundbegriffe **21**
1.1 Logik . . . 21
1.2 Mengen und Abbildungen . . . 22
1.3 Gruppen . . . 29
1.4 Ringe, Körper, Polynome . . . 32
1.5 Vektorräume . . . 40
1.6 Basis und Dimension . . . 46

2 Lineare Abbildungen **52**
2.1 Definition und Beispiele . . . 52
2.2 Lineare Abbildungen und Matrizen . . . 54
2.3 Koordinatentransformationen . . . 57

3 Determinanten **61**
3.1 Beispiele und Definition . . . 61
3.2 Cramersche Regel . . . 67
3.3 Determinante eines Endomorphismus . . . 68

4 Eigenwerte und Eigenvektoren **69**
4.1 Definition und Beispiele . . . 69
4.2 Das charakteristische Polynom . . . 70
4.3 Polynome von Matrizen und linearen Abbildungen . . . 74
4.4 Die Jordan-Normalform einer linearen Abbildung . . . 76

5 Euklidische Vektorräume **80**
5.1 Definitionen und Beispiele . . . 80
5.2 Orthogonale Transformationen . . . 83
5.3 Symmetrische lineare Abbildungen . . . 87

0 Einführung

20.10.09

0.1 Wozu Mathematik?

Angewandte Mathematik dient als Hilfswissenschaft zur Beschreibung, Vorhersage und Beeinflussung der (technischen, wirtschaftlichen, ...) Wirklichkeit.
Reine Mathematik untersucht abstrakte Strukturen und deren allgemeine Gesetzmäßigkeiten.
Kenntnis und Verständnis solcher Denkkonzepte ermöglichen es erst, reale Gegebenheiten in einer mathematischen Sprache zu formulieren.
Die **lineare Algebra**, und damit untrennbar verbunden die **Matrizenrechnung**, ist zusammen mit der Analysis ein wesentlicher Grundbestandteil auf dem fast alle modernen mathematischen Bereiche aufbauen.

0.2 Was macht (man mit) Mathematik?

- Ausrechnen: $\frac{1}{2} - \frac{1}{3} = ?$ $(\frac{1}{6})$.
- Gleichungen lösen (Algorithmen, Formeln).
  Gleichung: $ax^2 + bx + c = 0$, Lösung: $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, aber nur falls $a \neq 0$.
  Im Fall $a = 0$: Lineare Gleichung $bx + c = 0$.
  1. Fall $b \neq 0$: Lösung $x = -\frac{c}{b}$.
  2. Fall $b = 0$, $c \neq 0$: keine Lösung.
  3. Fall $b = 0$, $c = 0$: x ist beliebig, unendlich viele Lösungen.
- Klassifikation von möglichen Lösungsmengen. (z.B. Diskriminante D einer quadratischen Gleichung: für $D > 0$ gibt es zwei, für $D = 0$ eine und für $D < 0$ keine reelle Lösung.)
- Unterschiedliche Standardformen von Gleichungen

$$\underbrace{ax^2 + bx + c}_{\text{Normalform}} = \underbrace{a(x + \tfrac{b}{2a})^2 + c - \tfrac{b^2}{4a}}_{\text{Scheitelform}} = \underbrace{a(x - x_1)(x - x_2)}_{\text{Produktform}}$$

- Allgemeine Aussagen beweisen:
  Für jedes quadratische Polynom $p\,x^2 + qx + c$ gibt es eindeutige (komplexe) Zahlen x_1, x_2 mit $p\,x^2 + qx + c = (x - x_1)(x - x_2)$

0.3 Beispiele für lineare Probleme

Beispiel. Wie weit ist das Gewitter entfernt?
Zwischen Blitz und Donner liegen 3 s. Die Schallgeschwindigkeit ist etwa $333,32\,\frac{\mathrm{m}}{\mathrm{s}}$ (bei 3 °C). Der Blitz schlug also in $3\,\mathrm{s} \cdot 333,32\,\frac{\mathrm{m}}{\mathrm{s}} \approx 1\,\mathrm{km}$ Entfernung ein.

Beispiel. Wie weit ist das Ufer entfernt?
Der waagerechte Daumen (2 cm) am ausgestreckten Arm (60 cm) überdeckt gerade den Kirchturm am Ufer mit geschätzter Höhe von 30 m. Das Ufer ist also ca. $30\,\mathrm{m} \cdot \frac{60\,\mathrm{cm}}{2\,\mathrm{cm}} = 900\,\mathrm{m}$ entfernt.

Beispiel. Thomas Mann, Joseph und seine Brüder (adaptiert):
Kaufe ich einen Hektar Land hinzu, habe ich dreimal so viel Land wie mein Nachbar. Kauft hingegen er einen Hektar Land hinzu, habe ich nur doppelt soviel Land wie er. Wieviel Hektar Land haben wir? Durch Raten findet man leicht, dass ich 8 und der Nachbar 3 Hektar Land besitzen. Man zeige: es gibt keine weitere Möglichkeit. [HA]

Beispiel. Eine Parabel wird durch drei Punkte bestimmt.
Man bestimme die Koeffizienten der Parabel $y = ax^2 + bx + c$, die durch die Punkte (x_1, y_1), (x_2, y_2), (x_3, y_3) verläuft. Dabei sei zur Vereinfachung $x_2 - x_1 = x_3 - x_2$.

Lösung. Wir erhalten drei Gleichungen mit den drei Unbekannten a, b, c.

$$y_1 = ax_1^2 + bx_1 + c, \tag{0.1}$$
$$y_2 = ax_2^2 + bx_2 + c, \tag{0.2}$$
$$y_3 = ax_3^2 + bx_3 + c \tag{0.3}$$

Wir eliminieren geschickt c:

$$y_2 - y_1 = a(x_2^2 - x_1^2) + b(x_2 - x_1) \qquad (2) - (1) = (1')$$

$$y_3 - y_2 = a(x_3^2 - x_2^2) + b(x_3 - x_2) \qquad (3) - (2) = (2')$$

Schließlich erhält man wegen $x_2 - x_1 = x_3 - x_2$

$$y_3 - 2y_2 + y_1 = a(x_3^2 - 2x_2^2 + x_1^2), \qquad (2') - (1') = (1'')$$

woraus man a ablesen kann. b ergibt sich dann aus $(1')$ und c aus (1). [HA]

Beispiel. Widerstandsnetzwerke
Ersatzwiderstand einer Reihenschaltung von Widerständen: $R = R_1 + R_2$.
Ersatzwiderstand einer Parallelschaltung von Widerständen: $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$.

Wie lautet der Ersatzwiderstand des Widerstandsnetzwerkes in Abb. 0.1? 22.10.09
Für $R_0 = \infty$ ist der Ersatzwiderstand

$$R = \frac{1}{\frac{1}{R_1+R_3} + \frac{1}{R_2+R_4}}.$$

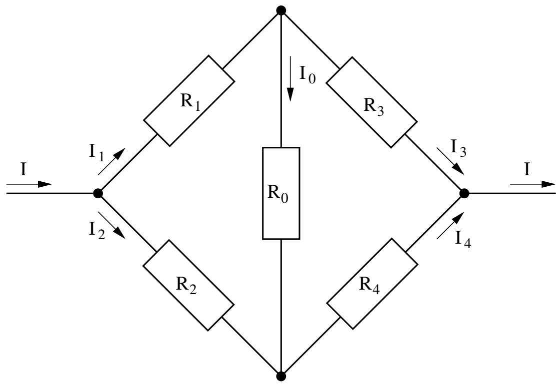


Abbildung 0.1: Eine Widerstandsbrücke

Für $R_0 = 0$ ist der Ersatzwiderstand

$$R = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}} + \frac{1}{\frac{1}{R_3} + \frac{1}{R_4}}.$$

Für beliebiges R_0 ist R gegeben durch $RI = R_1I_1 + R_3I_3$.
Sei I vorgegeben. Gesucht sind dann I_0, I_1, I_2, I_3, I_4.
Kirchhoffsche Gesetze:
1. Knotengleichungen: $I = I_1 + I_2$, $I_1 = I_0 + I_3$, $I_2 + I_0 = I_4$, $I_3 + I_4 = I$.
2. Maschengleichungen: $R_1I_1 + R_0I_0 - R_2I_2 = 0$, $R_3I_3 - R_4I_4 - R_0I_0 = 0$.
Lösung durch sukzessives Eliminieren von I_0, I_1, I_2, I_3, I_4. [HA]

Beispiel. Lineare Regression (Methode der kleinsten Quadrate)
Durch zwei verschiedene Punkte (x_1, y_1), (x_2, y_2) der Zeichenebene gibt es genau eine Gerade. Ist $x_1 \neq x_2$, so kann sie durch die Geradengleichung $y = mx + t$ beschrieben werden, wobei $m = \frac{y_2 - y_1}{x_2 - x_1}$ und $t = y_1 - mx_1$ ist.
Hat man mehr als zwei Punkte, $(x_1, y_1), \cdots, (x_N, y_N)$, die nur ungefähr auf einer Gerade liegen, sucht man diejenige „optimale" Gerade, $y = mx + t$, die, dadurch definiert ist, das die Summe der vertikalen Abstandsquadrate $\sum_{j=1}^{N}(mx_j + t - y_j)^2$ minimal ist. Als Ergebnis erhält man $m = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2}$ und $t = \overline{y} - m\overline{x}$.
Java applet: MatheVital von Jürgen Richter-Gebert.
`www-m10.ma.tum.de/bin/view/Lehrstuhl/Regression`
Anwendung: Regressionsgerade für den Zusammenhang zwischen Hausaufgabenpunkten und Klausurnote.
Fazit: **Hausaufgaben bearbeiten und korrigieren lassen!**

Beispiel. Die Google-Matrix: PageRank
wurde von Larry Page und Sergey Brin für die Stanford University patentiert (2001).
Relevanz einer Web-Seite i: $p(i)$, wobei $p(i) \geq 0$, $i = 1, \ldots, N$, N Anzahl der Webseiten.
Jeder Link von Seite j auf Seite i trägt zur Relevanz von Seite i bei, und zwar gewichtet

mit der Relevanz von j dividiert durch die Anzahl $n(j)$ der Links auf Seite j. Um Probleme mit Sackgassen und zyklischen Links zu vermeiden multipliziert man noch mit einem Dämpfungsfaktor $d \approx 0.85$ und addiert pauschal $1-d$ zu $p(i)$ dazu. Zeigt von den Seiten $j_1, \dots, j_k$ jeweils ein Link auf die Seite i, so berechnet sich der PageRank von i als

$$p(i) = \frac{d}{n(j_1)} p(j_1) + \cdots \frac{d}{n(j_k)} p(j_k) + 1 - d$$

Man erhält ein System von N linearen Gleichungen mit N Unbekannten, $p(1), \dots, p(N)$, wobei die Anzahl der Webseiten N z.B. schon 2004 in etwa 4 Milliarden war.

Beispiel. Für die 3D-Grafik von Google Earth muss für jeden Bildaufbau anhand der Position von darzustellenden Punkten im Raum (z.B. Eckpunkte eines Gebäudes) die zugehörige Position auf dem Computerbildschirm berechnet werden.
Einfaches Problem: Drehen eines Rechtecks mit den Ecken $O = (0,0)$, $A = (2,0)$, $B = (0,1)$ und $C = (2,1)$ in der Zeichenebene mit Winkel α um den Ursprung. Dann haben die gedrehten Punkte die Koordinaten $O' = O$, $A' = (2\cos\alpha, 2\sin\alpha)$, $B' = (-\sin\alpha, \cos\alpha)$ und schließlich

$$C' = A' + B' = (2 \cdot \cos\alpha - 1 \cdot \sin\alpha, 2 \cdot \sin\alpha + 1 \cdot \cos\alpha).$$

Für einen beliebigen Punkt $D = (x,y)$, $x, y \in \mathbb{R}$ lautet der gedrehte Punkt mit der gleichen Überlegung

$$D' = (\cos\alpha \cdot x - \sin\alpha \cdot y, \sin\alpha \cdot x + \cos\alpha \cdot y).$$

Schreibweise als Spaltenvektoren: $O = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $A = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $D = \begin{pmatrix} x \\ y \end{pmatrix}$.

Definition 1 (Anwenden einer Matrix auf einen Spaltenvektor)**.**

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

Setzt man

$$R = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

so gilt $D' = RD$ und entsprechend $O' = RO$, $A' = RA$, $B' = RB$, $C' = RC$.

Definition 2 (Matrixmultiplikation)**.**

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} := \begin{pmatrix} ax_1 + by_1 & ax_2 + by_2 & \cdots & ax_n + by_n \\ cx_1 + dy_1 & cx_2 + dy_2 & \cdots & cx_n + dy_n \end{pmatrix}$$

Die Koordinaten des um $\alpha = 45°$ gedrehten Rechtecks $OACB$ berechnet man so:

$$R \begin{pmatrix} 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \approx \begin{pmatrix} 0 & 1.414 & 0.707 & -0.707 \\ 0 & 1.414 & 2.121 & 0.707 \end{pmatrix}.$$

0.4 Lineare Gleichungssysteme

27.10.09

System von m linearen Gleichungen mit n Variablen: $x_1, \dots, x_n$

$$\begin{array}{ccccccccc} A_{11}x_1 & + & A_{12}x_2 & + & \cdots & + & A_{1n}x_n & = & b_1 \\ A_{21}x_1 & + & A_{22}x_2 & + & \cdots & + & A_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ A_{m1}x_1 & + & A_{m2}x_2 & + & \cdots & + & A_{mn}x_n & = & b_m \end{array} \qquad (*)$$

wobei $A_{11}, \dots, A_{mn}$, $b_1, \dots, b_m$ feste reelle Zahlen sind.
Lösungsmenge: $L = \{(x_1, \dots, x_n) \,|\, x_1, \dots, x_n$ reelle Zahlen, für die $(*)$ gilt$\}$

Definition 3. Für die Menge der reellen Zahlen $\mathbb{R}$ definiert man

- $\mathbb{R}^n := \{(x_1, \dots, x_n) \,:\, x_1, \dots, x_n \in \mathbb{R}\}$, die Menge der reellen N-Tupel.
- $\mathbb{R}^{m\times n} := \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \,:\, a_{11}, \dots, a_{mn} \in \mathbb{R} \right\}$,
  die Menge der reellen $m \times n$-Matrizen.

Schreibweise: $x \in \mathbb{R}^n$ bedeutet $x = (x_1, \dots, x_n)$ mit $x_1, \dots, x_n \in \mathbb{R}$.
Das Tupel $(x_1, \dots, x_n)$ schreibt man auch in senkrechter Anordnung

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := (x_1, \dots, x_n)$$

und nennt x einen Spaltenvektor.
$A \in \mathbb{R}^{m\times n}$ bedeutet $A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix}$ mit $A_{11}, \dots, A_{mn} \in \mathbb{R}$

Gleichheit:
Für $x, y \in \mathbb{R}^n$ ist $x = y$ gleichbedeutend mit $x_1 = y_1, \dots, x_n = y_n$.
Für $A, B \in \mathbb{R}^{m\times n}$ ist $A = B$ gleichbedeutend mit $A_{11} = B_{11}, \dots, A_{mn} = B_{mn}$.

Definition 4 (Anwenden einer Matrix auf einen Spaltenvektor)**.** Für $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$ ist

$$Ax := \begin{pmatrix} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n \\ \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n \end{pmatrix} \in \mathbb{R}^n$$

Damit lautet das Gleichungssystem $(*)$ kurz

$$Ax = b, \quad \text{bzw.,} \quad \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

mit $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. A heißt die **Koeffizientenmatrix** von $(*)$.

$$(A|b) := \begin{pmatrix} A_{11} & \cdots & A_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ A_{m1} & \cdots & A_{mn} & b_m \end{pmatrix} \in \mathbb{R}^{m \times (n+1)}$$

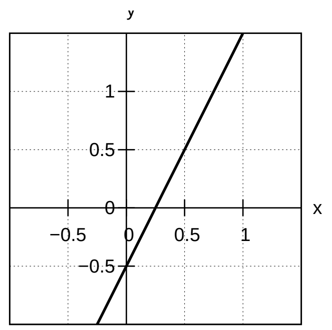
heißt die **erweiterte Koeffizientenmatrix** von $(*)$ und

$$L_{(A|b)} := \{x \in \mathbb{R}^n \,|\, Ax = b\}$$

ist die Lösungsmenge von $(*)$.

Beispiel. $m = 1$, $n = 2$: $4x - 2y = 1$
Lösungsmenge $L = \{(x, y) \in \mathbb{R}^2 \,|\, 4x - 2y = 1\}$
Auflösen nach y: $y = 2x - \frac{1}{2}$ (Geradengleichung)
Parametrisierung der Lösungsmenge: Zu jedem $t \in \mathbb{R}$ gibt es genau eine Lösung, nämlich $(t, 2t - \frac{1}{2}) \in \mathbb{R}^2$, d.h.

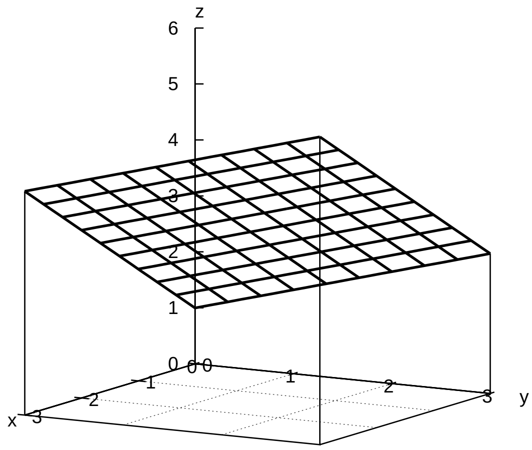
$$L = \{(t, 2t - \tfrac{1}{2}) \;:\; t \in \mathbb{R}\} \tag{0.4}$$



$$\boxed{\text{Eine lineare Gleichung mit zwei Unbekannten} \;\widehat{=}\; \text{Gerade im } \mathbb{R}^2}$$

Beispiel. $m = 1$, $n = 3$: $2x_1 - 2x_2 - x_3 = 2$
Lösungsmenge $L = \{x \in \mathbb{R}^2 \,|\, 2x_1 - 2x_2 - x_3 = 2\}$
Auflösen nach y: $x_1 = x_2 + \frac{1}{2}x_3 + 1$ (Ebenengleichung)
Parametrisierung der Lösungsmenge: Zu jedem $s, t \in \mathbb{R}$ gibt es genau eine Lösung, nämlich $(s + \frac{1}{2}t + 1, s, t) \in \mathbb{R}^3$, d.h.

$$L = \{(s + \tfrac{1}{2}t + 1, s, t) \;:\; s, t \in \mathbb{R}\} \tag{0.5}$$

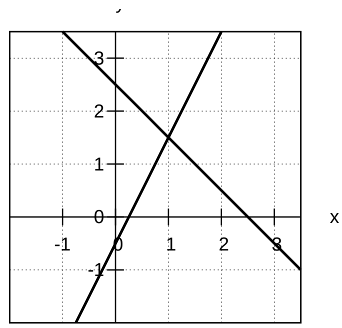


$$\boxed{\text{Eine lineare Gleichung mit drei Unbekannten} \mathrel{\widehat{=}} \text{Ebene im } \mathbb{R}^3}$$

Beispiel. $m = 2$, $n = 2$: $4x_1 - 2x_2 = 1$, $2x_1 + 2x_2 = 5$
Auflösen nach x_2,

$$x_2 = 2x_1 - \tfrac{1}{2}, \qquad x_2 = -x_1 + \tfrac{5}{2}$$

ergibt zwei Geradengleichungen. Die Lösung ist der Schnittpunkt der beiden Geraden.



Beispiel. $m = 2$, $n = 3$. Jede der zwei Gleichungen mit drei Unbekannten definiert eine Ebene im $\mathbb{R}^3$.
Die Lösungsmenge des Gleichungssystems entspricht der Schnittmenge der beiden Ebenen (i. Allg. eine Gerade).

Beispiel. $m = 3$, $n = 2$ [HA]

0.5 Gaußsches Eliminationsverfahren

Definition 5. $\mathbb{R}_n := \mathbb{R}^{1\times n}$ ist die Menge der reellen Zeilenvektoren.

Achtung: $a \in \mathbb{R}_n$ bedeutet $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$, dies ist *kein* n-Tupel von Zahlen,

$$\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \neq (a_1, \ldots, a_n) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n.$$

Für $a \in \mathbb{R}_n$ heißt der von links aus erste von 0 verschiedene Koeffizient der **führende Koeffizient**, oder Pivot.
$\begin{pmatrix} 0 & \cdots & 0 \end{pmatrix}$ heißt **Nullzeile**.

Definition 6 (Zeilen-Stufen-Form)**.**

- $A \in \mathbb{R}^{m \times n}$ heißt **in Zeilen-Stufen-Form**, falls gilt
  - (i) Alle Nullzeilen von A stehen unten in der Matrix.
  - (ii) Alle von Null verschiedenen Zeilen von A haben als führende Koeffizienten eine 1.
  - (iii) In jeder Zeile mit einer führenden Eins steht diese rechts von den darüberliegenden führenden Einsen.
- A heißt **in reduzierter Zeilen-Stufen-Form**, wenn jede Spalte von A, die eine führende Eins enthält, keine weiteren von Null verschiedenen Einträge hat.

Beispiel. $\begin{pmatrix} 1 & 2 & 4 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ sind in Zeilen-Stufen-Form.
$\begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ sind sogar in reduzierter Zeilen-Stufen-Form.

Bemerkung. Ist $(A|b) \in \mathbb{R}^{m \times (n+1)}$ in reduzierter Zeilen-Stufen-Form mit k führenden Einsen, so kann man eine Parametrisierung von $L_{(A|b)}$ durch $n - k$ Parameter direkt ablesen. Jeder Variablen, deren Zugehörige Spalte in A keine führende Eins enthält, wird einem Parameter gleichgesetzt. Die restlichen Variablen sind dann durch die Zeilen mit führender Eins eindeutig bestimmt.

Beispiel. $(A|b) = \begin{pmatrix} 1 & 2 & 0 & 1 & 3 \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $\begin{array}{ccccccccc} x_1 & + & 2 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 3 \\ & & & & 1 \cdot x_3 & + & 2 \cdot x_4 & = & 2 \\ & & & & & & 0 & = & 0 \end{array}.$
Ein Parameter für die 2. Spalte, $x_2 = s_1$ und einer für die 4. Spalte, $x_4 = s_2$. Somit ist

$$L_{(A|b)} = \{(3 - 2s_1 - s_2, s_1, 2 - 2s_2, s_2) \ : \ s_1, s_2 \in \mathbb{R}\}.$$

Bemerkung. Ist $(A|b)$ nur in Zeilen-Stufen-Form, so kann man eine Parametrisierung von $L_{(A|b)}$ durch sukzessives Einsetzen leicht bestimmen.

Beispiel. $(A|b) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$, $\left.\begin{array}{rcrcl} x_1 & + & 2x_2 & = & 1 \\ & & x_2 & = & 3 \end{array}\right\}$, also $\left\{\begin{array}{rcl} x_2 & = & 3, \\ x_1 & = & 1 - 2x_2 = -5. \end{array}\right.$

Definition 7 (Elementare Zeilenumformungen)**.** Seien $A, B \in \mathbb{R}^{m \times n}$.

- B heißt **elementare Zeilenumformung** von A, wenn B aus A hervorgeht durch
  (i) Vertauschen zweier Zeilen,
  (ii) Multiplikation einer Zeile mit einer Zahl $\lambda \neq 0$,
  (iii) Addition des λ-fachen von Zeile i zu Zeile j, $i \neq j$.
- Geht B aus A durch eine endliche Folge von elementaren Zeilenumformungen hervor, so schreibt man
$$A \sim B$$
und sagt: „A ist äquivalent zu B".

29.10.09

Satz 1. *Aus $(A|b) \sim (A'|b')$ folgt $L_{(A|b)} = L_{(A'|b')}$.*

Beweis. Es genügt $L_{(A|b)} = L_{(A'|b')}$ zu zeigen, wenn $(A'|b')$ eine elementare Zeilenumformung von $(A|b)$ ist:

Typ (i) Vertauschen von zwei Gleichungen ändert nichts an der Lösungsmenge.

Typ (ii) $a_1x_1 + \cdots + a_nx_n = b$ gilt genau dann, wenn $\lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b$, falls $\lambda \neq 0$.

Typ (iii) Gilt $a_1x_1 + \cdots + a_nx_n = b$, so ist für beliebiges $\lambda \in \mathbb{R}$

$$c_1x_1 + \cdots + c_nx_n = d$$

gleichbedeutend mit

$$\underbrace{(c_1 + \lambda a_1)x_1 + \cdots + (c_n + \lambda a_n)x_n}_{=(c_1x_1+\cdots+c_nx_n)+\lambda(a_1x_1+\cdots+a_nx_n)} = d + \lambda b$$

□

Satz 2 (Gauß-Elimination)**.** *Jede Matrix ist äquivalent zu einer Matrix in Zeilen-Stufen-Form.*

Beweis. Sei $A \in \mathbb{R}^{m \times n}$. Rekursiver Algorithmus:

Fall (i) Ist $A_{11} \neq 0$, so multipliziert man die erste Zeile mit $\frac{1}{A_{11}}$ (Typ (ii)) und addiert dann zu Zeile j das $(-A_{j1})$-fache der neuen Zeile 1 hinzu (Typ (iii)), wobei $j = 2, \ldots, m$. Entweder die so erhaltene Matrix A' ist schon in Zeilen-Stufen-Form, oder man streicht die erste Zeile und erste Spalte und erhält $B = \begin{pmatrix} A'_{22} & \cdots & A'_{2n} \\ \vdots & & \vdots \\ A'_{m2} & \cdots & A'_{mn} \end{pmatrix}$. Von B bestimmt man eine Zeilen-Stufen-Form B'. Dann ist

$$A'' = \begin{pmatrix} 1 & A'_{12} & \cdots & A'_{1n} \\ 0 & B'_{11} & \cdots & B'_{1(n-1)} \\ \vdots & & \vdots & \\ 0 & B'_{(m-1)1} & \cdots & B'_{(m-1)(n-1)} \end{pmatrix}$$

eine Zeilen-Stufen-Form von A.

Fall (ii) Ist $A_{11} = 0$ und es gibt ein j mit $A_{j1} \neq 0$, so vertauscht man Zeile j mit Zeile 1 und bestimmt von der so erhaltenen Matrix eine Zeilen-Stufen-Form gemäß Fall (i).

Fall (iii) Ist die ganze erste Spalte Null, so bringt man $\begin{pmatrix} A_{12} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \in \mathbb{R}^{m\times(n-1)}$ auf Zeilen-Stufen-Form. Die Zeilen-Stufen-Form von A erhält man durch Wiederhinzufügen der ersten Spalte mit lauter Nullen.

□

Satz 3. *Jede Matrix in Zeilen-Stufen-Form ist äquivalent zu einer Matrix in reduzierter Zeilen-Stufen-Form.*

Beweis. (Skizze) Durch Umnummerieren der Variablen $x_1, \ldots, x_n$ (entsprecht dem Vertauschen von Spalten in der Koeffizientenmatrix) kann man erreichen, dass die führenden Einsen soweit links wie möglich in der Koeffizientenmatrix stehen.
Dann gilt (exemplarisch für $m = 4$, $n = 5$)

$$\begin{pmatrix} 1 & * & * & * & * \\ & 1 & * & * & * \\ & & 1 & * & * \\ & & & & \end{pmatrix} \sim \begin{pmatrix} 1 & * & 0 & * & * \\ & 1 & 0 & * & * \\ & & 1 & * & * \\ & & & & \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & * & * \\ & 1 & 0 & * & * \\ & & 1 & * & * \\ & & & & \end{pmatrix}$$

Hierbei bedeutet ein leerer Eintrag immer eine 0, $*$ bedeutet jeweils eine Zahl, deren Wert bestimmt für den Beweis aber nicht relevant ist.
Man addiert also geeignete Vielfache der letzten Zeile mit einer führenden Eins zu den vorhergehenden Zeilen, so dass die Koeffizienten über dieser Eins zu Null werden. Dies wiederholt man für die vorletzte, vorvorletzte, …Zeile mit führender Eins und erhält schließlich eine reduzierte Zeilen-Stufen-Form von der ursprünglichen Matrix. □

Beispiel.

$$\begin{aligned} x+y+2z &= 9 & &\boxed{1} & &\begin{pmatrix}1&1&2&9\\2&4&-3&1\\3&6&-5&0\end{pmatrix}\\ 2x+4y-3z &= 1 & &\boxed{2} & &\\ 3x+6y-5z &= 0 & &\boxed{3} & & \end{aligned}$$

Eliminieren von x:

$$\begin{aligned} x+y+2z &= 9 & & & &\begin{pmatrix}1&1&2&9\\0&2&-7&-17\\0&3&-11&-27\end{pmatrix}\\ 2y-7z &= -17 & &\boxed{2}-2\cdot\boxed{1} & &\\ 3y-11z &= -27 & &\boxed{3}-3\cdot\boxed{1} & & \end{aligned}$$

Normieren der 2. Zeile:

$$\begin{aligned} x+y+2z &= 9 & & & &\begin{pmatrix}1&1&2&9\\0&1&-\frac{7}{2}&-\frac{17}{2}\\0&3&-11&-27\end{pmatrix}\\ y-\tfrac{7}{2}z &= -\tfrac{17}{2} & &\tfrac{1}{2}\cdot\boxed{2} & &\\ 3y-11z &= -27 & & & & \end{aligned}$$

Eliminieren von x:

$$\begin{aligned} x+y+2z &= 9 & & & &\begin{pmatrix}1&1&2&9\\0&1&-\frac{7}{2}&-\frac{17}{2}\\0&0&-\frac{1}{2}&-\frac{3}{2}\end{pmatrix}\\ y-\tfrac{7}{2}z &= -\tfrac{17}{2} & & & &\\ -\tfrac{1}{2}z &= -\tfrac{3}{2} & &\boxed{3}-3\cdot\boxed{2} & & \end{aligned}$$

Normieren der 3. Zeile:

$$\begin{aligned} x+y+2z &= 9 & & & &\begin{pmatrix}1&1&2&9\\0&1&-\frac{7}{2}&-\frac{17}{2}\\0&0&1&3\end{pmatrix} \quad \text{Zeilen-Stufen-Form}\\ y-\tfrac{7}{2}z &= -\tfrac{17}{2} & & & &\\ z &= 3 & &-\tfrac{1}{2}\cdot\boxed{3} & & \end{aligned}$$

Rücksubstitution von z:

$$\begin{aligned} x+y &= 3 & &\boxed{1}-2\cdot\boxed{3} & &\begin{pmatrix}1&1&0&3\\0&1&0&2\\0&0&1&3\end{pmatrix}\\ y &= 2 & &\boxed{2}+\tfrac{7}{2}\cdot\boxed{3} & &\\ z &= 3 & & & & \end{aligned}$$

Rücksubstitution von y:

$$\begin{aligned} x &= 1 & &\boxed{1}-\boxed{2} & &\begin{pmatrix}1&0&0&1\\0&1&0&2\\0&0&1&3\end{pmatrix} \quad \text{red. Zeilen-Stufen-Form}\\ y &= 2 & & & &\\ z &= 3 & & & & \end{aligned}$$

Korollar 1 (Gauß-Jordan-Elimination)**.** *Jede Matrix kann durch elementare Zeilenumformungen auf reduzierte Zeilen-Stufen-Form gebracht werden.*

Satz 4. *Die reduzierte Zeilen-Stufen-Form einer Matrix ist eindeutig.*

Beweis. in Kapitel 1. □

Definition 8 (Zeilenrang)**.** Der Zeilenrang einer Matrix A, $\mathrm{rg}(A)$ ist die Anzahl der führenden Einsen in der (reduzierten) Zeilen-Stufen-Form von A.

Bemerkung. Die Anzahl der führenden Einsen ändert sich nicht beim Übergang von Zeilen-Stufen-Form auf die reduzierte Zeilen-Stufen-Form.

Satz 5. *Sei $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$.*

- *Ist $rg(A) < rg(A|b)$, so besitzt $Ax = b$ keine Lösung.*

- *Ist* $rg(A) = rg(A|b)$*, so besitzt* $Ax = b$ *(mindestens) eine Lösung.*
- *Ist* $rg(A) = n$*, so besitzt* $Ax = b$ *höchstens eine Lösung.*

Beweis. (Skizze) Wir können annehmen dass $(A|b)$ und damit auch A bereits in reduzierter Zeilen-Stufen-Form vorliegen. Ohne Einschränkung seien die führenden Einsen links in der Matrix, d.h. sie liegen auf der Diagonalen, also (exemplarisch)

$$(A|b) = \begin{pmatrix} 1 & 0 & 0 & * & * & b_1 \\ & 1 & 0 & * & * & b_2 \\ & & 1 & * & * & b_3 \\ & & & & & b_4 \end{pmatrix}$$

.

- $\mathrm{rg}(A) < \mathrm{rg}(A|b)$ bedeutet $b_4 = 1$. D.h. die 4. Zeile entspricht der Gleichung $0 = 1$. Das Gleichungssystem besitzt also keine Lösung.
- $\mathrm{rg}(A) = \mathrm{rg}(A|b)$ bedeutet $b_4 = 0$. Dann ist $x_1 = b_1$, $x_2 = b_2$, $x_3 = b_3$ eine Lösung
- $\mathrm{rg}(A) = n$. Dann muss $m \geq n$ sein. Exemplarisch für $m = 5, n = 4$:

  $$(A|b) = \begin{pmatrix} 1 & & & & b_1 \\ & 1 & & & b_2 \\ & & 1 & & b_3 \\ & & & 1 & b_4 \\ & & & & b_5 \end{pmatrix}.$$

  Ist $b_5 = 1$, so gibt es keine Lösung, ist $b_5 = 0$, so ist die einzige Lösung $x_1 = b_1$, $x_2 = b_2$, $x_3 = b_3$, $x_4 = b_4$. Es gibt also höchstens eine Lösung.

□

Zusammenfassung: Jedes lineare Gleichungssystem kann durch elementare Zeilenumformungen auf reduzierte Zeilen-Stufen-Form gebracht werden, woraus eine Parametrisierung der Lösungsmenge einfach abzulesen ist.

0.6 Rechnen mit Vektoren des $\mathbb{R}^n$

Definition 9 (Addition und Skalarmultiplikation von Vektoren)**.** Sei $x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Dann ist

$$x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \in \mathbb{R}^n, \qquad \lambda x := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \in \mathbb{R}^n$$

Satz 6. *Für $x, y, z \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$ gilt*

$$\begin{aligned} x + y &= y + x, & (x+y)+z &= x+(y+z), \\ \lambda(x+y) &= \lambda x + \lambda y, & (\lambda+\mu)x &= \lambda x + \mu x, \\ \lambda(\mu x) &= (\lambda\mu)x, & 1x &= x. \end{aligned}$$

Beweis. Durch Einsetzen der Definition. □

Schreibweisen: $0 := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ heißt Nullvektor, $-x := (-1)x$, $x - y := x + (-y)$.

03.11.09

Definition 10 (Skalarprodukt)**.** Für $x, y \in \mathbb{R}^n$ ist

$$x \cdot y := x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}$$

Satz 7. *Für $x, y, z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ gilt*

(i) $x \cdot y = y \cdot x$

(ii) $(x+y) \cdot z = x \cdot z + y \cdot z$, $x \cdot (y+z) = x \cdot y + x \cdot z$,

(iii) $(\lambda x) \cdot y = \lambda(x \cdot y) = x \cdot (\lambda y)$,

(iv) $x \cdot x \geq 0$, *und aus* $x \cdot x = 0$ *folgt* $x = 0$.

Beweis. (i), (ii), (iii) durch Einsetzen der Definition,
(iv): $x \cdot x = x_1^2 + \cdots + x_n^2 \geq 0$. Aus $x_1^2 + \cdots + x_n^2 \geq 0$ folgt $x_1, \ldots, x_n = 0$, also $x = 0$. □

Definition 11 (Länge eines Vektors)**.** Für $x \in \mathbb{R}^n$ ist

$$|x| := \sqrt{x \cdot x} \geq 0$$

Satz 8 (Cauchy-Schwarz-Ungleichung)**.** *Für $x, y \in \mathbb{R}^n$ ist $-|x| \cdot |y| \leq |x \cdot y| \leq |x| \cdot |y|$.*

Beweis. Zu zeigen ist $(x \cdot y)^2 \leq (x \cdot x)(y \cdot y)$.
1. Fall: $y = 0$ ist ok.
2. Fall: $y \neq 0$. Dann gilt für beliebiges $\lambda \in \mathbb{R}$, wegen Satz 7, (i) bis (iv):

$$\begin{aligned} 0 &\overset{\text{(iv)}}{\leq} (x - \lambda y) \cdot (x - \lambda y) \overset{\text{(ii),(iii)}}{=} x \cdot (x - \lambda y) - \lambda y \cdot (x - \lambda y) \\ &\overset{\text{(ii),(iii)}}{=} x \cdot x - \lambda(x \cdot y) - \lambda(y \cdot x) + \lambda^2 (y \cdot y) \\ &\overset{\text{(i)}}{=} (y \cdot y)\lambda^2 - 2(x \cdot y)\lambda + x \cdot x \\ &\overset{\text{Scheitelform},\ y \cdot y > 0}{=} (y \cdot y)\left(\lambda - \frac{x \cdot y}{y \cdot y}\right)^2 - \frac{(x \cdot y)^2}{y \cdot y} + x \cdot x \end{aligned}$$

Insbesondere gilt für $\lambda = \frac{x \cdot y}{y \cdot y}$

$$0 \leq -\frac{(x \cdot y)^2}{y \cdot y} + x \cdot x, \quad \text{bzw.} \quad (x \cdot y)^2 \leq (x \cdot x)(y \cdot y).$$

□

Definition 12 (Winkel zwischen zwei Vektoren)**.** Für $x, y \in \mathbb{R}^n$, $x, y \neq 0$ setzt man

$$\sphericalangle(x, y) := \arccos \underbrace{\frac{x \cdot y}{|x||y|}}_{\in[-1,1]} \in [0, \pi]$$

Bemerkung.

- Für $x \cdot y = 0$ schreibt man auch $x \perp y$ („x steht senkrecht auf y“, „x ist orthogonal zu y“). Dann ist nämlich, falls $x, y \neq 0$, $\sphericalangle(x, y) = \frac{\pi}{2} = 90°$
- Es gilt $x \cdot y = |x| \cdot |y| \cos \sphericalangle(x, y)$ nach obiger Definition.

Im $\mathbb{R}^3$ kann man ein Produkt mit Werten in $\mathbb{R}^3$ definieren:

Definition 13 (Vektorprodukt)**.** Für $x, y \in \mathbb{R}^3$ setzt man

$$x \times y := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

Satz 9. *Für $x, y, z \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$ gilt*

$$\begin{aligned} x \times y &= -(y \times x), \quad \textit{insbesondere} \quad x \times x = 0, \\ x \times (y + z) &= x \times y + x \times z, \\ (x + y) \times z &= x \times z + y \times z \\ \lambda(x \times y) &= (\lambda x) \times y = x \times (\lambda y) \end{aligned}$$

Beweis. Einsetzen der Definition. □

Satz 10. *Für $x, y \in \mathbb{R}^3$ gilt*

(i) $x \times y \perp x$, $x \times y \perp y$

(ii) $|x \times y| = |x||y| \sin \sphericalangle(x, y)$

Beweis. [HA] □

0.7 Rechnen mit Matrizen

Definition 14 (Addition und Skalarmultiplikation von Matrizen)**.** Seien $A, B \in \mathbb{R}^{m\times n}$, $\lambda \in \mathbb{R}$. Dann ist

$$\begin{aligned} A + B &\in \mathbb{R}^{m\times n}, \qquad (A + B)_{ij} := A_{ij} + B_{ij}, \quad i = 1, \ldots, m, j = 1, \ldots, n \\ \lambda A &\in \mathbb{R}^{m\times n}, \qquad (\lambda A)_{ij} := \lambda A_{ij}, \quad i = 1, \ldots, m, j = 1, \ldots, n \end{aligned}$$

Satz 11. *Für $A, B, C \in \mathbb{R}^{m\times n}$, $\lambda, \mu \in \mathbb{R}$ gilt*

$$\begin{aligned} A + B &= B + A, & (A+B)+C &= A+(B+C),\\ \lambda(A+B) &= \lambda A + \lambda B, & (\lambda+\mu)A &= \lambda A + \mu A,\\ \lambda(\mu A) &= (\lambda\mu)A, & 1A &= A. \end{aligned}$$

Beweis. Durch Einsetzen der Definition. □

Schreibweisen: $0 := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{m\times n}$ heißt Nullmatrix,
$-A := (-1)A$, $A - B := A + (-B)$.

Definition 15 (Matrix-Multiplikation)**.** Für $A \in \mathbb{R}^{m\times r}$, $B \in \mathbb{R}^{r\times n}$ ist

$$AB \in \mathbb{R}^{m\times n}, \qquad (AB)_{ij} := A_{i1}B_{1j} + \cdots + A_{ir}B_{rj}, \quad i = 1,\ldots,m, j = 1,\ldots,n$$

Beispiel. $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix}$,

$$AB = \underbrace{\begin{pmatrix} 1 & 2 & 4 \\ \mathbf{2} & \mathbf{6} & \mathbf{0} \end{pmatrix}}_{2\times 3} \underbrace{\begin{pmatrix} 4 & 1 & \mathbf{4} & 3 \\ 0 & -1 & \mathbf{3} & 1 \\ 2 & 7 & \mathbf{5} & 2 \end{pmatrix}}_{3\times 4} = \underbrace{\begin{pmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & \mathbf{26} & 12 \end{pmatrix}}_{2\times 4}$$

Satz 12 (Eigenschaften der Matrixmultiplikation)**.** *Für $a, B \in \mathbb{R}^{m\times r}$, $C, D \in \mathbb{R}^{r\times n}$, $F \in \mathbb{R}^{n\times s}$, $\lambda \in \mathbb{R}$ gilt*

(i) $(A + B)C = AC + BC$, $A(C + D) = AC + BC$,

(ii) $(\lambda A)C = \lambda(AC) = A(\lambda C)$,

(iii) $(AC)F = A(CF)$.

Beweis. Durch Einsetzen der Definition. Wir führen den Beweis für (iii) aus.
Man vergleicht die einzelnen Einträge der Matrix auf der linken und rechten Seite der Gleichung: Für $i = 1,\ldots,m$, $j = 1,\ldots,s$ ist

$$\begin{aligned} ((AC)F)_{ij} &= (AC)_{i_1}F_{1j} + \cdots + (AC)_{i_n}F_{nj} \\ &= (A_{i1}C_{11} + \cdots + A_{ir}C_{r1})F_{1j} \\ &\quad + \cdots + \\ &\quad +(A_{i1}C_{1n} + \cdots + A_{ir}C_{rn})F_{nj} \end{aligned}$$

und

$$\begin{aligned}(A(CF))_{ij} &= A_{i1}(CF)_{1j} + \cdots + A_{ir}(CF)_{rj} \\ &= A_{i1}(C_{11}F_{1j} + \cdots + C_{1n}F_{nj}) \\ &\quad + \cdots + \\ &\quad + A_{ir}(C_{r1}F_{1j} + \cdots + C_{rn}F_{nj}).\end{aligned}$$

Nach Ausmultiplizieren sieht man, dass beide Ausdrücke als Summanden genau die rn Terme $A_{ik}C_{kl}F_{lj}$, $k = 1, \ldots, r$, $l = 1, \ldots, n$ haben und somit gleich sind. □

Beispiel.

- $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^{n\times 1}(\widehat{=}\mathbb{R}^n)$, $b \in \mathbb{R}^{m\times 1}$. Dann kann im linearen Gleichungssystem $Ax = b$ der Term Ax als Matrixprodukt einer $m \times n$ mit einer $n \times 1$-Matrix aufgefasst werden.
- Sei A' elementare Zeilenumformung von $A \in \mathbb{R}^{4\times 5}$
  (i) durch Vertauschen von Zeile 1 und 3, dann ist
  $$A' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A,$$
  (ii) durch Multiplikation von Zeile 1 mit $\lambda \neq 0$, dann ist
  $$A' = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A,$$
  (iii) durch Addition von λ mal Zeile 1 zu Zeile 2, dann ist
  $$A' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \lambda & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A.$$

Definition 16 (Transponierte einer Matrix)**.** Sei $A \in \mathbb{R}^{m\times n}$. Dann heißt

$$A^T \in \mathbb{R}^{n\times m}, \quad (A^T)_{ij} := A_{ji}, i = 1, \ldots, n,\ j = 1, \ldots, m,$$

die Transponierte von A.

Beispiel.

- $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$
- $x, y \in \mathbb{R}^n (\widehat{=}\mathbb{R}^{n\times 1})$, dann ist $x \cdot y = x^T y$, wenn man $\mathbb{R}^{1\times 1}$ mit $\mathbb{R}$ identifiziert.

Satz 13. *Für $A, B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times s}$, $\lambda \in \mathbb{R}$*

(i) $(A^T)^T = A$, *(ii)* $(A+B)^T = A^T + B^T$, *(iii)* $(\lambda A)^T = \lambda A^T$,

(iv) $(AC)^T = C^T A^T$.

Beweis. [HA] □

Quadratische Matrizen

05.11.09

Bemerkung. Die Matrixmultiplikation hat überraschende Eigenschaften:
Es gibt $A, B \in \mathbb{R}^{n\times n}$ mit

- $AB \neq BA$, z.B. $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.
- $AB = 0$, aber $A \neq 0$ und $B \neq 0$, z.B. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Für $A \in \mathbb{R}^{n\times n}$ heißt $(A_{11}, \ldots, A_{nn})$ die **(Haupt-)Diagonale** von A.

$$E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}$$

heißt $n \times n$-Einheitsmatrix. Sie hat lauter Einsen auf der Diagonalen und sonst nur Nulleinträge. Für beliebiges $B \in \mathbb{R}^{m\times n}$ gilt

$$E_m B = B = B E_n.$$

Wenn klar ist welche Einheitsmatrix gemeint ist, schreibt man auch nur E statt E_n.

Beispiel. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Satz 14. *Gilt für $A \in \mathbb{R}^{n\times n}$, dass $rg(A) = n$, so ist die reduzierte Zeilen-Stufen-Form von A gleich E_n.*

Beweis. (Skizze) Sei A' reduzierte Zeilen-Stufen-Form von A. A' hat n führende Einsen, in jeder Zeile und jeder Spalte genau eine. Da A' in Zeilen-Stufen-Form ist, müssen alle führenden Einsen auf der Hauptdiagonalen liegen. Da A' darüber hinaus auch in reduzierter Zeilen-Stufen-Form ist, gibt es sonst keine von Null verschiedenen Einträge, also $A' = E_n$. □

Definition 17 (Inverse einer Matrix)**.** Sei $A \in \mathbb{R}^{n \times n}$. Gibt es ein $B \in \mathbb{R}^{n \times n}$ mit $BA = AB = E$, so heißt A invertierbar und B heißt Inverse von A.

Satz 15.

- *Gibt es zu $A \in \mathbb{R}^{n \times n}$ ein B mit $BA = E$, so gilt auch $AB = E$, und umgekehrt.*
- *Es gibt höchstens ein $B \in \mathbb{R}^{n \times n}$ mit $AB = E = BA$.*

Beweis. später, folgt aus Gruppeneigenschaft. □

Schreibweise: Ist $A \in \mathbb{R}^{n \times n}$ invertierbar, so schreibt man A^{-1} für die Inverse von A.

Bemerkung. Ist $A \in \mathbb{R}^{n \times n}$ invertierbar, so besitzt das lineare Gleichungssystem

$$Ax = b$$

genau eine Lösung, nämlich $A^{-1}b$, denn
$A(A^{-1}b) = (AA^{-1})b = Eb = b$, d.h., $A^{-1}b$ ist eine Lösung (Existenz), und aus
$Ax = b$ folgt $A^{-1}(Ax) = A^{-1}b$, daraus $(A^{-1}A)x = A^{-1}b$, bzw. $x = A^{-1}b$ (Eindeutigkeit).

Beispiel. Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad - bc \neq 0$, dann ist

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

Satz 16. *Sind $A, B \in \mathbb{R}^{n \times n}$ invertierbar, so gilt:*

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

(ii) falls $AB = BA$, dann ist $A^{-1} + B^{-1} = (AB)^{-1}(A + B)$.

Beweis.

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

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

□

1 Algebraische Grundbegriffe

1.1 Logik

Eine **Aussage** ist ein Ausdruck der wahr oder falsch ist.

Beispiel. „3 ist eine Primzahl"ist wahr, „4 ist Teiler von 7"ist falsch, „2 ist Lösung von $x^2 - 3x + 2 = 0$", „$\{1, 2\}$ ist die Lösungsmenge von $x^2 - 3x + 2 = 0$"sind beide wahre Aussagen.

Zwei Aussagen A und B kann man zu neuen Aussagen verknüpfen:

$A \vee B$	ist die Aussage	„A oder B",
$A \wedge B$	ist die Aussage	„A und B",
$\neg A$	ist die Aussage	„nicht A", bzw. „A ist falsch"
$A \Rightarrow B$	ist die Aussage	„Aus A folgt B", bzw., „A ist hinreichend für B"
$B \Leftarrow A$	ist die Aussage	„B ermöglicht A", bzw., „B ist notwendig für A"
$A \Leftrightarrow B$	ist die Aussage	„A ist äquivalent zu B"

Bemerkung. $A \Rightarrow B$ ist gleichbedeutend mit $B \Leftarrow A$ ist gleichbedeutend mit $\neg A \vee B$. $A \Leftrightarrow B$ ist gleichbedeutend mit $(A \Rightarrow B) \wedge (B \Rightarrow A)$.

Ein **Prädikat** (Aussageform) ist ein Ausdruck, der Variablen enthält und für jede Festlegung der Variablen den Aussagewert wahr oder falsch besitzt.
Ein Prädikat wird zu einer Aussage, wenn

(i) allen Variablen ein Wert zugewiesen wird,
(Bsp.: Ist $A_x :=$ „$x + 2 = 4$", dann ist $A_3 =$ „$3 + 2 = 4$" falsch, aber A_2 ist wahr.
Ist $B_n :=$ „n ist Primzahl", dann ist $B_1 =$ „1 ist Primzahl" falsch, aber B_3 ist wahr.

(ii) alle Variablen durch Quantoren gebunden werden:

$\forall x : A_x$	bedeutet	„für alle x ist A_x wahr" (Allquantor)
$\exists x : A_x$	bedeutet	„es gibt ein x für das A_x wahr ist" (Existenzquantor)

Beispiel. Fast immer schränkt man die Quantoren auf eine „vernünftige" Menge ein: $\boxed{\text{10.11.09}}$

- $A_x =$ „$x + 2 = 4$", dann ist $\forall x \in \mathbb{R} : A_x$ falsch, und $\exists x \in \mathbb{R} : A_x$ wahr.

- $A_{x,y} =$ „$x + y = 4$“, dann ist

$$\begin{array}{rl} \forall x{\in}\mathbb{R}\forall y{\in}\mathbb{R} : A_{x,y} & \text{falsch,} \\ \exists x{\in}\mathbb{R}\exists y{\in}\mathbb{R} : A_{x,y} & \text{wahr,} \\ \forall x{\in}\mathbb{R}\exists y{\in}\mathbb{R} : A_{x,y} & \text{wahr,} \\ \exists x{\in}\mathbb{R}\forall y{\in}\mathbb{R} : A_{x,y} & \text{falsch.} \end{array}$$

1.2 Mengen und Abbildungen

Eine endliche Menge kann durch Angabe ihrer Elemente bestimmt werden, z.B.

$$\{1, 2, 4\}, \quad \{2, 3, 5, 2, 2, 3\}.$$

$$\begin{array}{rcl} x \in M & \text{bedeutet} & \text{„}x\text{ ist ist in der Menge }M\text{ enthalten“, „}x\text{ ist Element von }M\text{“} \\ x \notin M & \text{bedeutet} & \neg(x \in M) \\ M' \subset M & \text{bedeutet} & \forall x : x \in M' \Rightarrow x \in M \quad (M'\text{ ist Teilmenge von }M) \\ M' = M & \text{bedeutet} & (M' \subset M) \wedge (M \subset M') \\ \emptyset := \{\} & \text{ist die} & \text{leere Menge, es gilt } \forall x : x \notin \emptyset. \end{array}$$

Bemerkung. Somit ist z.B. $\{1, 1, 2, 1, 2, 2\} = \{1, 2\} = \{2, 1\}$.

Folgende Zahlenmengen werden wir häufig benutzen:

$$\begin{array}{rcl} \mathbb{N} & = & \{1, 2, 3, \ldots\} \quad \text{ist die Menge der natürlichen Zahlen, } 0 \notin \mathbb{N}, \\ \mathbb{N}_0 & = & \{0, 1, 2, 3, \ldots\} \\ \mathbb{Z} & = & \{\ldots, -2, -1, 0, 1, 2, \ldots\} \quad \text{ist die Menge der ganzen Zahlen,} \\ \mathbb{Q} & = & \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\}, \quad \text{wobei } \frac{p}{q} = \frac{r}{s}\text{, falls } ps = rq \text{ (z.B. } \frac{2}{3} = \frac{4}{6} = \frac{6}{9}), \\ \mathbb{R}^+ & = & \text{Menge der positiven reellen Zahlen,} \\ \mathbb{R} & = & \text{Menge der reellen Zahlen,} \\ \mathbb{C} & = & \text{Menge der komplexen Zahlen.} \end{array}$$

Es gilt

$$\emptyset \subset \mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}^+ \subset \mathbb{R} \subset \mathbb{C}$$

Durch ein Prädikat A_x kann man neue Mengen definieren:

$$\{x|\, A_x\} \quad \text{enthält als Elemente alle } x\text{, die die Eigenschaft } A_x \text{ haben.}$$

Beispiel. $\{n \in \mathbb{N} \,|\, n \text{ ist gerade}\} := \{n \,|\, n \in \mathbb{N} \wedge n \text{ ist gerade}\} \subset \mathbb{N}$,
$\mathcal{P}(M) := \{N \,|\, N \subset M\}$ ist die Menge aller Teilmengen von M (Potenzmenge).

Verknüpfung von Mengen: Seien M, N beliebige Mengen. Dann ist

$$\begin{aligned} M \cup N &:= \{x \,|\, (x \in M) \vee (x \in N)\} && \text{Vereinigungsmenge} \\ M \cap N &:= \{x \,|\, (x \in M) \wedge (x \in N)\} && \text{Schnittmenge, Durchschnitt} \\ M \setminus N &:= \{x \,|\, (x \in M) \wedge (x \notin N)\} && \text{Differenzmenge} \\ M \times N &:= \{p \,|\, p = (x, y) \wedge x \in M \wedge y \in N\} && \text{kartesisches Produkt} \end{aligned}$$

Ist I eine beliebige Menge und für jedes $i \in I$ ist M_i eine Menge. Man nennt dann $(M_i)_{i\in I}$ eine (indizierte) **Familie** von Mengen und I die zugehörige Indexmenge.

$$\begin{aligned} \bigcup_{i\in I} M_i &:= \{x \,|\, \exists i{\in}I : x \in M_i\} \\ \bigcap_{i\in I} M_i &:= \{x \,|\, \forall i{\in}I : x \in M_i\} \end{aligned}$$

Ist $I = \{1, \ldots, n\}$, so schreibt man auch $\bigcup\limits_{i=1}^{n} M_i$, bzw., $\bigcap\limits_{i=1}^{n} M_i$

Abbildungen

Das Konzept der Abbildung einer Menge auf eine andere ist das wohl wichtigste der gesamten Mathematik. Mengen und kompliziertere Strukturen sind tote Objekte. Erst die Möglichkeit diese Objekte abzubilden haucht ihnen Leben ein.
Seien M, N beliebige Mengen. Eine **Abbildung** (oder Funktion) f von M nach N ist eine Vorschrift, die jedem $x \in M$ genau ein $f(x) \in N$ zuordnet. Dafür schreibt man

$$f : M \to N, \quad x \mapsto f(x)$$

Zwei Abbildungen $f : M \to N$, $g : M \to N$, heißen gleich, $f = g$, wenn gilt

$$\forall x{\in}M : f(x) = g(x)$$

Beispiel. • Funktionen mit einer Variablen

$$\begin{aligned} \mathrm{sqr} &: \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2, \\ \sqrt{\ } &: \mathbb{R}_0^+ \to \mathbb{R}, \quad x \mapsto \text{diejenige nichtnegative Zahl, deren Quadrat gleich } x \text{ ist,} \\ \exp &: \mathbb{R} \to \mathbb{R}, \\ \log &: \mathbb{R}^+ \to \mathbb{R}, \\ \sin &: \mathbb{R} \to \mathbb{R}, \\ \cos &: \mathbb{R} \to \mathbb{R}. \end{aligned}$$

• Funktionen mit zwei Variablen

$$\begin{aligned} + &: \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto +(x, y) = (x + y), \text{ die Summe von } x \text{ und } y. \\ \cdot &: \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto \cdot(x, y) = (xy), \text{ das Produkt von } x \text{ und } y \\ \mathrm{pow} &: \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto \mathrm{pow}(x, y) = (x^y), \text{ die } y\text{-te Potenz von } x. \end{aligned}$$

- Eine Familie von reellen Zahlen $(r_i)_{i\in I}$ ist eine Funktion $r : I \to \mathbb{R}$ mit $r_i := r(i)$.
- Ein n-Tupel reeller Zahlen, $x \in \mathbb{R}^n$ ist eine Familie mit Indexmenge $\{1,\dots,n\}$, somit ist $\mathbb{R}^n$ die Menge aller Funktionen von $\{1,\dots,n\}$ nach $\mathbb{R}$.
- Eine Familie mit Indexmenge $\mathbb{N}$ nennt man Folge.
- Eine reelle $m \times n$-Matrix A ist eine Abbildung $A : \{1,\dots,m\} \times \{1,\dots,n\} \to \mathbb{R}$, mit der Schreibweise $A_{ij} = A(i,j)$.

Für die Menge aller Abbildungen von M nach N schreibt man

$$N^M := \{f \mid f : M \to N\}$$

Bemerkung. Somit ist $\mathbb{R}^n$ eine Kurzschreibweise für $\mathbb{R}^{\{1,\dots,n\}}$,
$\mathbb{R}^{m\times n}$ ist Abkürzung für $\mathbb{R}^{\{1,\dots,m\}\times\{1,\dots,n\}}$,

Sei $f : M \to N$, $X \subset M$, $Y \subset N$. Dann heißt

$$f(X) := \{f(x) : x \in X\} := \{y \in N \mid \exists x \in X : y = f(x)\} \subset N$$

das **Bild** von X (unter der Abbildung f) und definiert eine Abbildung von $\mathcal{P}(M)$ nach $\mathcal{P}(N)$.

$$f^{-1}(Y) := \{x \in M \mid f(x) \in Y\} \subset M$$

heißt das **Urbild** von Y (unter der Abbildung f). f^{-1} ist keine Abbildung von N nach M, sondern von $\mathcal{P}(N)$ nach $\mathcal{P}(M)$.

$$f|_X : X \to N, \qquad x \mapsto f(x)$$

heißt die **Einschränkung** von f auf X. Ist $f(M) \subset Y$, so ist die Abbildung

$$\tilde{f} : M \to Y, \qquad x \mapsto f(x)$$

verschieden von f, wird aber, wenn keine Missverständnisse zu befürchten sind, auch mit f bezeichnet.
Die Menge $G_f := \{(x, f(x)) : x \in M\} \subset M \times N$ heißt der **Graph** von f.

Beispiel. Definiert man für $A \in \mathbb{R}^{m\times n}$ die Abbildung $f : \mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^m$, so kann die Lösungsmenge von $Ax = b$, $b \in \mathbb{R}^m$, geschrieben werden als das Urbild von b, $L_{(A|b)} = f^{-1}(\{b\})$. Ist z.B. $L_{(A|b)} = \{x^{(0)}+s_1x^{(1)}+s_2x^{(2)} : s \in \mathbb{R}^2\}$ mit $x^{(0)}, x^{(1)}, x^{(2)} \in \mathbb{R}^n$, so ist $L_{(A|b)} = g(\mathbb{R}^2)$ das Bild des R^2 unter der Abbildung

$$g : \mathbb{R}^2 \to \mathbb{R}^n, \qquad s \mapsto x^{(0)} + s_1x^{(1)} + s_2x^{(2)}.$$

12.11.09

Definition 18. Die Abbildung $f : M \to N$ heißt

- **injektiv**, wenn $\forall x, x' \in M : f(x) = f(x') \Rightarrow x = x'$,
- **surjektiv**, wenn $\forall y \in N \exists x \in M : y = f(x)$,
- **bijektiv**, wenn sie zugleich injektiv und surjektiv ist.

Ist $f : M \to N$ bijektiv, so definiert man die **Umkehrabbildung**

$$f^{-1} : N \to M, \qquad y \mapsto \text{dasjenige } x \in M, \text{ für das } f(x) = y \text{ ist.}$$

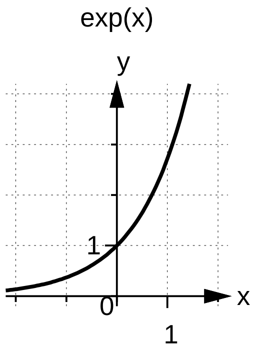
Ein solches x gibt es für jedes $y \in N$, da f surjektiv ist (Existenz). Die Vorschrift ist eindeutig, denn für ein weiteres $x' \in M$ mit $f(x') = y$, folgt wegen der Injektivität aus $f(x) = f(x')$ sofort $x = x'$ (Eindeutigkeit).

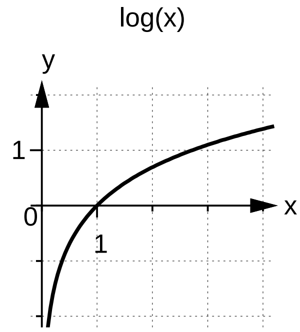
Bemerkung. Ist $f : M \to N$ nicht bijektiv, so ist $f^{-1} : N \to M$ nicht definiert, die Urbildfunktion $f^{-1} : \mathcal{P}(N) \to \mathcal{P}(M)$ hingegen schon. Ist f bijektiv, so gilt

$$\forall y \in N : f^{-1}(\{y\}) = \{f^{-1}(y)\}.$$

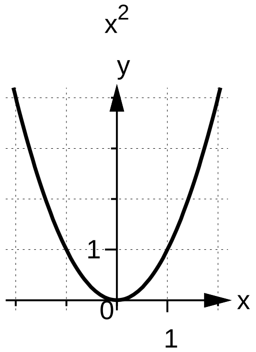
Beispiele.

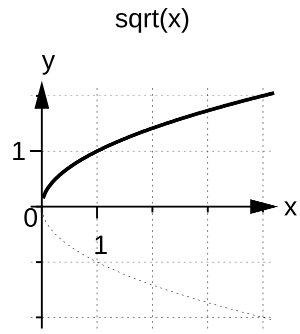
- $\exp : \mathbb{R} \to \mathbb{R}$ ist injektiv, aber nicht surjektiv. $\exp : \mathbb{R} \to \mathbb{R}^+$ ist bijektiv mit Umkehrabbildung $\log : \mathbb{R}^+ \to \mathbb{R}$.



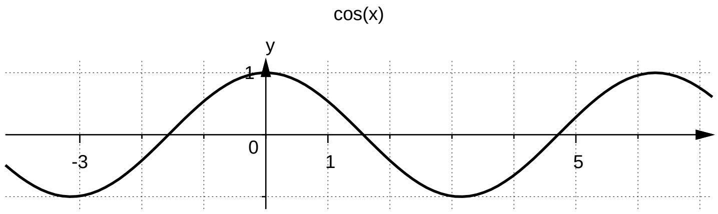


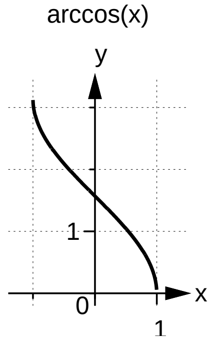
- $\mathrm{sqr} : \mathbb{R} \to \mathbb{R}$ ist weder injektiv noch surjektiv, $\mathrm{sqr}|_{\mathbb{R}_0^+} \to \mathbb{R}_0^+$ ist bijektiv mit Umkehrfunktion $\sqrt{\ }$.





- $\cos : \mathbb{R} \to [-1,1]$ ist nicht injektiv aber surjektiv. $\cos|_{[0,\pi]}$ ist bijektiv, mit Umkehrabbildung $\arccos : [-1,1] \to [0,\pi]$





Definition 19 (Mächtigkeit von Mengen)**.** Zwei Mengen M, N heißen **gleich mächtig**, $|M| = |N|$, wenn es eine Bijektion von M nach N gibt.
Eine Menge M heißt **endlich** ($|M| < \infty$), wenn es ein $n \in \mathbb{N}$ und eine Bijektion von $\{1, \ldots, n\}$ nach M gibt. In diesem Fall sagt man M enthält n Elemente, $|M| = n$.
Andernfalls heißt M unendlich, $|M| = \infty$.
In diesem Fall heißt M **abzählbar unendlich**, wenn es eine Bijektion von $\mathbb{N}$ nach M gibt.
Ist M weder endlich noch abzählbar unendlich, so heißt M **überabzählbar**.

Man beachte, dass aus $|M| = \infty$, $|N| = \infty$ **nicht** $|M| = |N|$ folgt.
Z.B. ist $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$, aber $|\mathbb{Q}| \neq |\mathbb{R}|$, $|\mathbb{N}| \neq |\{0,1\}^{\mathbb{N}}|$). [HA]

Satz 17. *Sei $f : M \to N$. Dann gilt*

(i) f injektiv $\Leftrightarrow \forall y \in Y : |f^{-1}(\{y\})| \leq 1$,

(ii) f surjektiv $\Leftrightarrow \forall y \in Y : |f^{-1}(\{y\})| \geq 1$,

(iii) f bijektiv $\Leftrightarrow \forall y \in Y : |f^{-1}(\{y\})| = 1$,

Beweis.

(i) “$\Leftarrow$” Seien $x, x' \in M$ mit $f(x) = f(x') \Rightarrow x, x' \in f^{-1}(\{f(x)\}) \Rightarrow x = x'$, da $|f^{-1}(\{f(x)\})| \leq 1$. Also ist f injektiv.
“$\Rightarrow$” Widerspruchsbeweis: Sei f injektiv.
Annahme: Es gibt $y \in M$ mit $|f^{-1}(\{y\})| \geq 2 \Rightarrow$ Es gibt $x, x' \in f^{-1}(\{y\})$ mit $x \neq x'$, aber $f(x) = y = f(x')$, im Widerspruch zu f injektiv. Es gibt also kein y, für das $|f^{-1}(\{y\})| \geq 2$, damit folgt die Behauptung.

(ii) Sei $y \in N$ beliebig. $\exists x \in M : f(x) = y \Leftrightarrow f^{-1}(\{y\}) \neq \emptyset \Leftrightarrow |f^{-1}(\{y\})| \geq 1$.

(iii) klar.

$\square$

Definition 20 (Komposition von Abbildungen)**.** Sind $f : M \to N$, $g : N \to P$ Abbildungen, so heißt die Abbildung

$$g \circ f : M \to P, \quad x \mapsto (g \circ f)(x) := g(f(x))$$

Komposition von g mit f.

Ist $g : N' \to P$, so ist $g \circ f$ schon definiert, wenn $f(M) \subset N'$.
Falls definiert, ist $\circ$ assoziativ, $h \circ (g \circ f) = (h \circ g) \circ f$, aber im allgemeinen nicht kommutativ, $g \circ f \neq f \circ g$, z.B. für $f = \mathrm{sqr}$, $g = \cos$.
Zur Menge M definiert man die **identische Abbildung**

$$\mathrm{id}_M : M \to M, \quad x \mapsto x.$$

Für $f : M \to N$ gilt

$$\underbrace{\underbrace{\mathrm{id}_N}_{N \leftarrow N} \circ \underbrace{f}_{N \leftarrow M}}_{N \leftarrow M} = \underbrace{f}_{N \leftarrow M} = \underbrace{\underbrace{f}_{N \leftarrow M} \circ \underbrace{\mathrm{id}_M}_{M \leftarrow M}}_{N \leftarrow M}.$$

Äquivalenzrelationen

Eine Relation $\sim$ setzt jeweils zwei Elemente x, y einer Menge X in Beziehung, $x \sim y$, oder nicht, $\neg(x \sim y)$.

Beispiel.

- $x \leq y$ ist eine Relation auf $\mathbb{R}$.
- Für $f : X \to X$ sei yfx wahr, genau dann, wenn $y = f(x)$ (Relation auf X).
- $m|n$ (m teilt n ohne Rest) ist eine Relation auf $\mathbb{N}$.
- Sei $n \in \mathbb{N}$ fest gewählt. Für $k, l \in \mathbb{Z}$ sei $k \equiv l \pmod n$ genau dann, wenn $n|(k - l)$ (Relation auf $\mathbb{Z}$).

Eine Relation $\sim$ auf X ist eindeutig bestimmt durch ihren Graphen

$$R = \{(x, y) \in X \times X \,|\, x \sim y\} \subset X \times Y.$$

Umgekehrt definiert jede Teilmenge R von $X \times X$ eine Relation auf X durch die Setzung: $x \sim y$ gilt genau dann, wenn $(x, y) \in R$.

Definition 21 (Äquivalenzrelation)**.** Eine Relation $\sim$ auf X heißt **Äquivalenzrelation**, wenn für alle $x, y, z \in X$ gilt

(Ä1) $x \sim x$, (Reflexivität)

(Ä2) $x \sim y \Rightarrow y \sim x$, (Symmetrie)

(Ä3) $x \sim y \wedge y \sim z \Rightarrow x \sim z$. (Transitivität)

In diesem Fall heißt für $x \in X$

$$\overline{x} := [x]_\sim := \{y \in X \mid y \sim x\} \subset X$$

die **Äquivalenzklasse** von x (bzgl. $\sim$), und

$$X/_\sim := \{[x]_\sim : x \in X\} \subset \mathcal{P}(X)$$

heißt **Quotient** von X (bzgl. $\sim$).
Für eine Äquivalenzklasse $A \subset X/_\sim$ heißt jedes Element $x \in A$ auch **Repräsentant** von A.

Definition 22. Eine Familie $(X_i)_{i \in I}$ von nichtleeren Teilmengen von X heißt **Zerlegung** von X, wenn $\bigcup\limits_{i \in I} X_i = X \quad$ und $\quad X_i \cap X_j = \emptyset \;$ für $\; i \neq j$.

Bemerkung. Jede Zerlegung $(X_i)_{i \in I}$ von X definiert eine Äquivalenzrelation $\sim$ auf X vermöge

$$x \sim y \quad :\Leftrightarrow \quad \exists i \in I : x \in X_i \wedge y \in X_i$$

mit dem Quotienten $X/_\sim = \{X_i : i \in I\}$.

17.11.09

Satz 18. *Ist $\sim$ eine Äquivalenzrelation auf X, so sind alle Äquivalenzklassen nichtleere Mengen und jedes x gehört genau einer Äquivalenzklasse an. Insbesondere gilt für zwei Äquivalenzklassen A, B, dass entweder $A = B$ oder $A \cap B = \emptyset$.*

Beweis. Ist A eine Äquivalenzklasse, so gibt es nach Definition ein $a \in X$ mit $A = [a]_\sim$, also $A \neq \emptyset$. Für jedes $a \in X$ gilt $a \in [a]_\sim = \{x \in X \mid x \sim a\}$, da $a \sim a$. Gehört a einer weiteren Äquivalenzklasse $B = [b]_\sim$, $b \in X$ an, so ist zu zeigen, dass $[a]_\sim = B$.
"$\subset$" Aus $x \in [a]_\sim$ folgt $x \sim a$, aus $a \in B$ folgt $a \sim b$. Wegen der Transitivität ist also $x \sim b$ und somit $x \in B$.
"$\supset$" Aus $x \in B$ folgt $x \sim b$, aus $a \in B$ folgt $a \sim b$, bzw., $b \sim a$. Insgesamt ist also $x \sim a$ und damit $x \in [a]_\sim$
Seien nun $A, B \in X/_\sim$. Dann gibt es $a, b \in X$ mit $A = [a]_\sim$ und $B = [b]_\sim$.
1. Fall: $a \sim b$. Dann ist $a \in B$, und somit $A = [a]_\sim = B$.
2. Fall: $\neg(a \sim b)$. Annahme: Es gibt $x \in A \cap B$. Dann ist $x \sim a$ und $x \sim b$ und somit $a \sim b$, Widerspruch, also ist $A \cap B$ die leere Menge. $\square$

Bemerkung. Ist I eine Menge und $A : I \to X/_\sim$ eine Bijektion, so ist die Familie $(A_i)_{i\in I}$ eine Zerlegung von X. Da man immer $I = X/_\sim$ und $A = \mathrm{id}_{X/_\sim}$ wählen kann, gilt:

Eine Äquivalenzrelation auf X definiert eine Zerlegung von X in Äquivalenzklassen

Anwendung. $A \sim B$ ist eine Äquivalenzrelation auf $\mathbb{R}^{m\times n}$, denn $A \sim A$ (Ä1).
$B \sim A$ heißt, es gibt invertierbare Matrizen $E'_1, \dots, E'_k$ (elementare Zeilenumformungen), so dass $B = E'_k \cdots E'_1 A$. Dann ist $A = (E'_1)^{-1} \cdots (E'_k)^{-1} B$, wobei die $(E'_j)^{-1}$ wieder elementare Zeilenumformungen sind (Übungsblatt 3, Aufgabe 4). Somit ist $A \sim B$ (Ä2). Das aus $C \sim B$ und $B \sim A$ auch $C \sim A$ folgt ist klar (Ä3). Wir werden später sehen, dass sich in jeder der Äquivalenzklasse genau eine Matrix in reduzierter Zeilen-Stufen-Form befindet.

1.3 Gruppen

Eine Verknüpfung $*$ auf einer Menge G ist eine Abbildung

$$*: G \times G \to G, \qquad (a, b) \mapsto *(a, b) =: a * b$$

$*$ heißt **assoziativ**, wenn

$$\forall a, b, c \in G : (a * b) * c = a * (b * c)$$

$*$ heißt **kommutativ**, wenn

$$\forall a, b \in G : a * b = b * a.$$

Beispiele. $+, \cdot : G \times G \to G$, $G = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ sind assoziativ und kommutativ.
$\mathrm{pow} : G \times G \to G$, $(x, y) \mapsto x\hat{\ }y$, $G = \mathbb{N}, \mathbb{R}^+$ ist weder assoziativ noch kommutativ.
$\circ : M^M \times M^M \to M^M$, $(f, g) \mapsto f \circ g$ (Komposition),
$\cdot : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $(A, B) \mapsto AB$ (Matrixmultiplikation)
sind beide assoziativ, aber nicht kommutativ.

Definition 23. Eine Menge G mit Verknüpfung $*$, kurz $(G, *)$ heißt **Gruppe**, wenn gilt

(G1) $*$ ist assoziativ,

(G2) $\exists e \in G \forall a \in G : e * a = a$ (Existenz eines neutralen Elements)

(G3) Ist e wie in (G2), so gilt: $\forall a \in G \exists a' \in G : a' * a = e$. (Existenz des Inversen)

G heißt **abelsche** (oder kommutative) Gruppe, wenn $*$ kommutativ ist.

Satz 19. *Sei $(G, *)$ eine Gruppe.*

(i) G besitzt genau ein neutrales Element, für das gilt:

$$\forall a \in G : a * e = e * a = a \tag{$*$}$$

(ii) Jedes $a \in G$ besitzt genau ein Inverses, bezeichnet mit a^{-1}, für das gilt:

$$a * a^{-1} = a^{-1} * a = e \tag{$**$}$$

(iii) Für $a, b \in G$ hat die Gleichung $a * x = b$ immer genau eine Lösung.

Beweis. Für $a * b$ schreiben wir auch einfach kurz ab.

(a), (b) Zuerst ($**$): Sei e ein neutrales Element von G. Zu $a \in G$ sei a' invers, $a'a = e$ und a'' invers zu a', $a''a' = e$ (G3). Dann ist

$$aa' \overset{(G2)}{=} e(aa') = (a''a')(aa') \overset{(G1)}{=} a''(a'a)a' = a''(ea') \overset{(G2)}{=} a''a' = e.$$

Daraus folgt auch ($*$), nämlich

$$ae = a(a'a) \overset{(G1)}{=} (aa')a \overset{(**)}{=} ea \overset{(G2)}{=} a$$

Eindeutigkeit von e: Sei $\tilde{e}$ auch neutrales Element von G, insbesondere ist $\tilde{e}e = e$. Wegen ($*$) gilt auch $\tilde{e}e = \tilde{e}$, somit ist $\tilde{e} = e$.
Eindeutigkeit de Inversen: Sei $\tilde{a}'$ auch invers zu a, dann ist

$$\tilde{a}' \overset{(*)}{=} \tilde{a}'e \overset{(**)}{=} \tilde{a}'aa' = ea' = a'$$

(c) Existenz: $x = a^{-1}b$ ist Lösung von $ax = b$, denn $a(a^{-1}b) = eb = b$.
Eindeutigkeit: Sei x' eine weitere Lösung, es gilt also $ax' = b$. dann folgt auch $a^{-1}(ax') = a^{-1}b$, bzw., $x' = a^{-1}b$.

$\square$

Bemerkung. Ist G eine abelsche Gruppe, benutzt man auch die additive Schreibweise $a + b$ für die Verknüpfung, 0 (Null) für das neutrale Element und $-a$ (das Negative von a) für das Inverse von a.

Beispiele.

- $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{R}^n, +)$, $(\mathbb{R}^{m\times n}, +)$ sind abelsche Gruppen.
  $(\mathbb{N}, +)$, $(\mathbb{R}_0^+, +)$ sind keine Gruppen, wegen des fehlenden neutralen Elements, bzw., der fehlenden Negativen.

- $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R}^+, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$ sind Gruppen.
  $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$ sind keine Gruppen, da die 0 kein Inverses besitzt.

- $(M^M, \circ)$ ist keine Gruppe, aber $S(M) := \{f : M \to M \,|\, f \text{ ist bijektiv}\} \subset M^M$ ist mit der Komposition eine Gruppe, neutrales Element ist id_M und die Inversen sind die jeweiligen Umkehrabbildungen. $S_n := S(\{1, \ldots, n\}$ mit $\circ$ heißt **Permutationsgruppe**.

- $(\mathbb{R}^{n\times n}, \cdot)$ mit der Matrixmultiplikation ist keine Gruppe, aber

$$\mathrm{GL}_n(\mathbb{R}) := \{A \in \mathbb{R}^{n\times n} \,|\, \mathrm{rg}(A) = n\}$$

  ist mit $\cdot$ eine Gruppe. Neutrales Element ist die Einheitsmatrix E, das Inverse von A ist die durch Gauß-Jordan-Elimination bestimmte Lösung der Matrixgleichung $AX = E$. Nach Satz 19 gibt es genau eine solche Matrix, für die auch $XA = E$ gilt. Sie wird als das Matrixinverse A^{-1} von A bezeichnet. Dies liefert den Beweis von Satz 15. 19.11.09

Definition 24. Sei $(G, \cdot)$ eine Gruppe. $G' \subset G$ heißt **Untergruppe** von G, wenn $G' \neq \emptyset$ und

$$\forall a, b \in G' : a \cdot b \in G' \wedge a^{-1} \in G'.$$

Ist $(H, \odot)$ eine weitere Gruppe, so heißt eine Abbildung $\varphi : G \to H$ mit

$$\forall a, b \in G : \varphi(a \cdot b) = \varphi(a) \odot \phi(b)$$

ein (Gruppen-)**Homomorphismus**. Ein bijektiver Homomorphismus heißt **Isomorphismus**.

Bemerkung.

- Eine Untergruppe G' von $(G, \cdot)$ ist mit der Verknüpfung $\cdot|_{G'\times G'} : G' \times G' \to G'$ wieder eine Gruppe.
  Beweis: Die Einschränkung der Verknüpfung bleibt assoziativ (G1). Da mit $a \in G'$ auch $a^{-1} \in G'$ (G3) folgt auch $a^{-1}a = e \in G'$ (G2). $\square$

- Für einen Gruppenhomomorphismus $\varphi : G \to H$ gilt $\varphi(e) = e$, $\varphi(a^{-1}) = \varphi(a)^{-1}$, wobei die neutralen Elemente von G und H beide mit e bezeichnet werden.
  Beweis: Aus $\varphi(e) = \varphi(e \cdot e) = \varphi(e) \odot \varphi(e)$ folgt durch Multiplikation mit $\varphi(e)^{-1}$, dass $H \ni e = \varphi(e)$. Außerdem ist $\varphi(a^{-1})\varphi(a) = \varphi(a^{-1}a) = \varphi(e) = e$. $\square$

- Ist $\varphi : G \to H$ ein Isomorphismus, so auch $\varphi^{-1} : H \to G$.
  Beweis: φ^{-1} ist bijektiv und für $c, d \in H$ mit $c = \varphi(a)$, $d = \varphi(b)$, $a, b \in G$ ist $\varphi^{-1}(c \odot d') = \varphi^{-1}(\varphi(a) \cdot \varphi(b)) = \varphi^{-1}(\varphi(a \cdot b)) = a \cdot b = \varphi^{-1}(c) \cdot \varphi^{-1}(d)$ $\square$

Beispiele. In den ersten drei Beispielen ist $n \in \mathbb{N}$ fest gewählt.

- Für $n \in \mathbb{N}$ ist $n\mathbb{Z} := \{nk : k \in \mathbb{Z}\}$ eine Untergruppe von $(\mathbb{Z}, +)$.

- Für $n \in \mathbb{N}$, $\mathbb{Z}_n := 0, \ldots, n-1$, ist $(\mathbb{Z}_n, +_n)$ eine abelsche Gruppe, wobei

$$a +_n b := \begin{cases} a + b, & \text{falls } a + b < n, \\ a + b - n & \text{sonst.} \end{cases}$$

Für die Assoziativität erhält man

$$(a +_n b) +_n c = a +_n (b +_n c) = \begin{cases} a+b+c, & \text{falls } a+b+c < n, \\ a+b+c-n, & \text{falls } n \le a+b+c < 2n, \\ a+b+c-2n, & \text{falls } a+b+c \ge 2n. \end{cases}$$

0 ist neutrales Element und $n - a \in \mathbb{Z}_n$ ist invers zu $a \in \mathbb{Z}_n$.

- Auf $\mathbb{Z}$ definiert $k \sim l :\Leftrightarrow k - l \in n\mathbb{Z}$ eine Äquivalenzrelation mit Äquivalenzklassen $\overline{k} = k + n\mathbb{Z} := \{k + nj : j \in \mathbb{Z}\}$. $\overline{k} + \overline{l} := \overline{k+l}$ definiert eine assoziative Verknüpfung auf $\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/_\sim$, mit der $(\mathbb{Z}/n\mathbb{Z}, +)$ eine Gruppe ist.
  Beweis der Wohldefiniertheit von +: Sind $k, k' \in \overline{k}$, $l, l' \in \overline{l}$, dann ist nach Definition einerseits $\overline{k'} + \overline{l'} = \overline{k' + l'}$, andererseits ist $\overline{k'} + \overline{l'} = \overline{k} + \overline{l} := \overline{k+l}$. Damit die Vorschrift wohldefiniert ist, muss $\overline{k' + l'} = \overline{k+l}$ gezeigt werden.
  Dies gilt aber, da aus $k' - k \in n\mathbb{Z}$ und $l' - l \in n\mathbb{Z}$ auch $(k' + l') - (k + l) = (k' - k) + (l' - l) \in n\mathbb{Z}$, da $n\mathbb{Z}$ Untergruppe von Z ist.

- $\mathbb{Z}_n \to \mathbb{Z}/n\mathbb{Z}$, $a \mapsto a + n\mathbb{Z}$ ist ein Gruppenisomorphismus von $(\mathbb{Z}_n, +_m)$ nach $(\mathbb{Z}/n\mathbb{Z}, +)$.

- $\mathbb{Z} \mapsto \mathbb{Z}/n\mathbb{Z}$, $k \mapsto \overline{k}$ ist surjektiver Homomorphismus von $(\mathbb{Z}, +)$ nach $(\mathbb{Z}/n\mathbb{Z}, +)$.

- Die bijektive Abbildung $\exp : \mathbb{R} \to \mathbb{R}^+$ ist ein Isomorphismus von $(\mathbb{R}, +)$ nach $(\mathbb{R}^+, \cdot)$, denn $\exp(x + y) = \exp(x)\exp(y)$.

- Die Abbildung $R : \mathbb{R} \to \mathbb{R}^{2\times 2}$, $R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$
  ist ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ nach $(\mathrm{GL}_n(\mathbb{R}), \cdot)$ [HA]

1.4 Ringe, Körper, Polynome

Definition 25. Eine Menge R mit zwei Verknüpfungen

$$\begin{aligned} + &: R \times R \to R, \quad (a,b) \mapsto a + b, \\ \cdot &: R \times R \to R, \quad (a,b) \mapsto a \cdot b, \end{aligned}$$

kurz $(R, +, \cdot)$, heißt **Ring**, wenn gilt

(R1) $(R, +)$ ist (additiv geschriebene) abelsche Gruppe,

(R2) $\cdot$ ist assoziativ,

(R3) $\forall a, b, c \in R : a \cdot (b + c) = a \cdot b + a \cdot c \wedge (a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivgesetze).

Ein Ring R heißt *kommutativ*, wenn $\forall a, b \in R : a \cdot b = b \cdot a$.
Gibt es ein Element $1 \in R$, für dass $\forall a \in R : 1 \cdot a = a \cdot 1 = a$, so heißt R *Ring mit Eins*.

Das neutrale Element von $(R,+)$ wird als 0 bezeichnet und es gilt $\forall a \in R : 0{\cdot}a = a{\cdot}0 = 0$, denn $0 \cdot a = (0+0)\cdot a = 0\cdot a + 0 \cdot a$.

Beispiele.

- $(\mathbb{Z},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$ sind kommutative Ringe mit 1.
- $(\mathbb{Z}_n,+_n,\cdot_n)$ mit $a \cdot_n b := \operatorname{mod}(ab,n) :=$ (ganzzahliger Rest bei der Division von ab durch n), ist ein kommutativer Ring mit 1.
- $(\mathbb{R}^{n\times n},+,\cdot)$ mit der Matrixmultiplikation ist ein nichtkommutativer Ring mit Eins.

Bemerkung. Ist $(R,+,\cdot)$ ein kommutativer Ring mit Eins und $(a_i)_{i\in\{1,\dots,n\}}$ eine Familie in R, so definiert man rekursiv

$$\begin{aligned}
\sum_{i=1}^{n} a_i &:= \left\{ \begin{matrix} \left(\sum\limits_{i=1}^{n-1} a_i\right) + a_n, & n \geq 1 \\ 0, & n = 0 \end{matrix} \right\} =: a_1 + \cdots + a_n, \\
\prod_{i=1}^{n} a_i &:= \left\{ \begin{matrix} \left(\prod\limits_{i=1}^{n-1} a_i\right) a_n, & n \geq 1 \\ 1, & n = 0 \end{matrix} \right\} =: a_1 \cdots a_n, \\
a^n &:= \prod_{i=1}^{n} a, \quad n \in \mathbb{N}_0.
\end{aligned}$$

24.11.09

In einem Ring mit Eins besitzen die Elemente im Allgemeinen kein multiplikativ Inverses. Die 0 kann nie ein Inverses haben, da ja $a \cdot 0 = 0 \neq 1$.
Gibt es $a,b \in R\setminus\{0\}$ mit $a\cdot b = 0$, so kann weder a noch b ein Inverses besitzen, sonst wäre z.B. $b = a^{-1}\cdot a \cdot b = a^{-1}\cdot 0 = 0$ im Widerspruch zu $b \neq 0$. Analog müsste $a = 0$ sein, wenn b ein Inverses besäße.
Beispiel: In $\mathbb{Z}_6$ ist $2 \cdot_6 3 = 0$. In $(\mathbb{R}^{2\times 2},+,\cdot)$ ist $\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right) = 0$.

Definition 26. Eine Menge $\mathbb{K}$ mit zwei assoziativen Verknüpfungen

$$\begin{aligned}
+ &: \mathbb{K}\times\mathbb{K} \to \mathbb{K}, \quad (a,b) \mapsto a+b, \\
\cdot &: \mathbb{K}\times\mathbb{K} \to \mathbb{K}, \quad (a,b) \mapsto a\cdot b,
\end{aligned}$$

kurz $(\mathbb{K},+,\cdot)$, heißt **Körper**, wenn gilt

(K1) $(\mathbb{K},+)$ ist abelsche Gruppe,

(K2) $(\mathbb{K}\setminus\{0\},\cdot)$ ist abelsche Gruppe (mit 1 als neutralem Element),

(K3) $\forall a,b,c \in \mathbb{K} : a{\cdot}(b{+}c) = a{\cdot}b{+}a{\cdot}c \quad \wedge \quad (a{+}b){\cdot}c = a{\cdot}c{+}b{\cdot}c$ (Distributivgesetze).

Bemerkung. In einem Körper $\mathbb{K}$ gelten die Rechenregeln

(i) $1 \neq 0$,

(ii) $0 \cdot a = a \cdot 0 = 0$,

(iii) $a \cdot b = 0 \Rightarrow (a = 0) \vee (b = 0)$ (Nullteilerfreiheit)

(iv) $-(ab) = a(-b)$, $(-a)(-b) = ab$,

(v) $a \neq 0$ und $x \cdot a = y \cdot a \Rightarrow x = y$ (Kürzungsregel)

Beweis. (i) gilt, da $1 \in \mathbb{K} \setminus \{0\}$, jeder Körper enthält also mindestens zwei Elemente. (ii) gilt, da $\mathbb{K}$ ein Ring ist und (iii) gilt, da für $a \neq 0$ aus $a \cdot b = 0$ folgt dass $a^{-1}ab = a^{-1}0 = 0$ ist, bzw., $b = 0$. (iv) folgt aus
$ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$ und $(-a)(-b) = -((-a)b) = -(-(ab)) = ab$.
(v) folgt aus (K2) durch beidseitige Multiplikation mit a^{-1}. □

Beispiele.

- $(\mathbb{Q}, +, \cdot)$ (Bruchrechnen) und $(\mathbb{R}, +, \cdot)$ (Rechnen mit Grenzwerten) bilden jeweils einen Körper.

- $(\mathbb{Z}_p, +_p, \cdot_p)$ ist ein Körper, wenn p eine Primzahl ist.
  In der diskreten Mathematik lernt man, dass es zu zwei teilerfremden ganzen Zahlen $p, q \in \mathbb{Z}$ immer $r, s \in \mathbb{N}$ gibt, so dass $qr = ps + 1$ (erweiterter Euklidischer Algorithmus). $q \in \mathbb{Z}_n$ ist teilerfremd zur Primzahl p. Es gibt sogar ein $r \in \mathbb{Z}_p$ mit $qr = ps + 1$, d.h. $q \cdot_p r = 1$, m.a.W., $q^{-1} = r$.

- **Die komplexen Zahlen**
  $\mathbb{C} := \{(x + iy) : x, y \in \mathbb{R}\}$ mit der Addition $(x+iy)+(u+iv) := (x+u)+i(y+v)$ und der Multiplikation $(x + iy)(u + iv) := (xu - yv) + i(xv + yu)$ ist ein Körper, der *Körper der komplexen Zahlen.*.
  Beweis: (Skizze) Dass $(\mathbb{C}, +)$ abelsche Gruppe ist, folgt genauso wie für $(\mathbb{R}^2, +)$. $0 + i0$ ist das neutrale Element und $-(x + iy) = (-x) + i(-y) =: -x - iy$ das Negative von $x+iy$. Assoziativität von $\cdot$ und Distributivgesetze durch Nachrechnen, neutrales Element der Multiplikation ist $1 + i \cdot 0$, Inverse:

  $$(x + iy)^{-1} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2},$$

  denn

  $$(x + iy)\left(\tfrac{x}{x^2+y^2} - i\tfrac{y}{x^2+y^2}\right) = \left(x\tfrac{x}{x^2+y^2} - y\tfrac{-y}{x^2+y^2}\right) + i\left(x\tfrac{-y}{x^2+y^2} + y\tfrac{x}{x^2+y^2}\right) = 1 + i \cdot 0.$$

  □

  Für $x + i \cdot 0$ schreibt man einfach x, somit ist $\mathbb{R} \subset \mathbb{C}$, für $0 + iy$ schreibt man iy, für $x + i(-y)$ schreibt man $x - iy$.
  Die Abbildung $\mathbb{R}^2 \ni (x, y) \mapsto x + iy \in \mathbb{C}$ ist ein Gruppenisomorphismus bezüglich der Addition.

Für $z = x + iy \in \mathbb{C}$ heißt $\operatorname{Re}(z) := x$ der *Realteil* von z, $\operatorname{Im}(z) := y$ heißt der *Imaginärteil* von z und

$$\overline{z} := x - iy \in \mathbb{C}$$

heißt *(komplex) konjugierte* von z. Für $w, z \in \mathbb{C}$ gilt offenbar

$$\begin{aligned} \overline{w} + \overline{z} &= \overline{w + z}, \\ \overline{w} \cdot \overline{z} &= \overline{wz}, \\ z = \overline{z} &\Leftrightarrow z \in \mathbb{R} \end{aligned}$$

Für $z = x + iy \in \mathbb{C}$, $x, y \in \mathbb{R}$, gilt

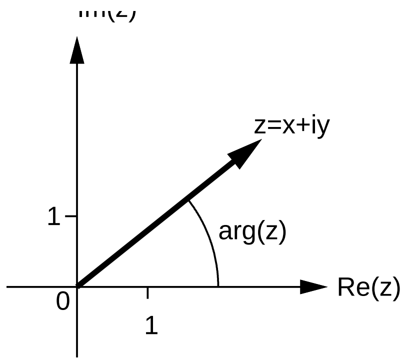
$$z\overline{z} = (x + iy)(x - iy) = x^2 - i^2y^2 = x^2 + y^2 \geq 0.$$

Der Betrag von z ist definiert als

$$|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2} = \left| \begin{pmatrix} x \\ y \end{pmatrix} \right|.$$

Achtung: Auf $\mathbb{C}(\equiv \mathbb{R}^2)$ gibt es keine Ordnung wie auf $\mathbb{R}$!
Es gilt weder $i \leq 1$ noch $i \geq 1$.
Beispiel: $\frac{1}{z} := z^{-1} = \frac{\overline{z}}{z\overline{z}} = \frac{1}{|z|^2}\overline{z}$. Konkret: $\frac{1}{3+4i} = \frac{3-4i}{(3+4i)(3-4i)} = \frac{3-4i}{25}$.

Geometrische Beschreibung (Polarform)



Die Argumentfunktion $\arg : \mathbb{C} \setminus \{0\} \to]-\pi, \pi]$ ist definiert durch

$$\arg z := \begin{cases} \sphericalangle\left(\binom{x}{y}, \binom{1}{0}\right), & \text{falls } y \geq 0, \\ -\sphericalangle\left(\binom{x}{y}, \binom{1}{0}\right), & \text{falls } y < 0 \end{cases}$$

Ist der Betrag $r \geq 0$ und der Winkel $\varphi \in]-\pi, \pi]$ gegeben, dann ist

$$r\underbrace{(\cos\varphi + i\sin\varphi)}_{=:\, e^{i\varphi}} = r\cos\varphi + ir\sin\varphi \in \mathbb{C}$$

Bemerkung. Für $z \neq 0$ gilt

$$z = |z| e^{i \arg(\varphi)}$$

Beweis. (Skizze) Für $z = x + iy$ gilt

$$\cos(\arg(z)) = \begin{cases} \cos\left(\arccos \frac{\binom{x}{y} \cdot \binom{1}{0}}{|\binom{x}{y}|}\right) = \frac{x}{\sqrt{x^2+y^2}} = \frac{\operatorname{Re} z}{|z|}, & \operatorname{Im} z \geq 0 \\ \cdots = \frac{\operatorname{Re} z}{|z|}, & \operatorname{Im} z < 0 \end{cases}$$

Ähnlich erhält man $\sin(\arg z) = \frac{\operatorname{Im} z}{|z|}$. Somit ist

$$|z| e^{i \arg(z)} = |z| \cos(\arg(z)) + i|z| \sin(\arg(z)) = \operatorname{Re} z + i \operatorname{Im} z = z.$$

$\square$

Beispiele: $1 = 1 \cdot e^{i \cdot 0}$, $i = 1 \cdot e^{i\frac{\pi}{2}}$, $1 - i = \sqrt{2} e^{-i\frac{\pi}{4}}$.

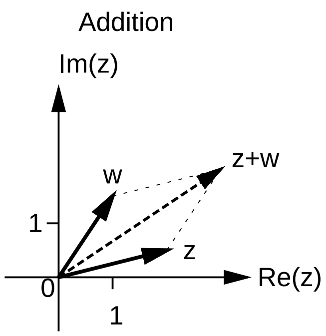
Für $z \in \mathbb{C}$ heißt $z = x + iy$, $x, y \in \mathbb{R}$ die *Normalform* von z und 26.11.09
$z = r\, e^{i\varphi}$, $r \geq 0$, $\varphi \in \,]-\pi, \pi]$ die *Polarform* von z.

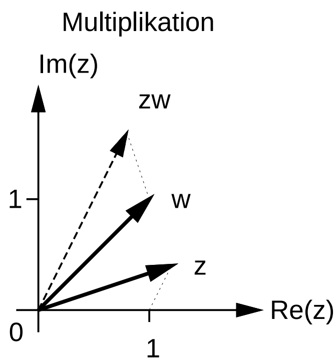
Multiplikation: Seien $z = re^{i\varphi}$, $w = se^{i\psi}$ in Polarform gegeben. Dann ist

$$\begin{aligned} z \cdot w &= r(\cos\varphi + i \sin\varphi) \cdot s(\cos\psi + i \sin\psi) \\ &= rs(\cos\varphi\cos\psi - \sin\varphi\sin\psi) + irs(\sin\varphi\cos\psi + \cos\varphi\sin\psi) \\ &\overset{\text{Add.th.}}{=} rs(\cos(\varphi+\psi) + irs\sin(\varphi+\psi) \\ &= rs\, e^{i(\varphi+\psi)} \end{aligned}$$

Das Produkt zweier komplexen Zahlen ist gegeben durch das Produkt der Beträge, die Winkel zur reellen positiven Achse addieren sich. *Bemerkung:* Daraus ergibt sich direkt die Assoziativität der Multiplikation von komplexen Zahlen.
Insbesondere gilt

$$e^{i(\varphi+\psi)} = e^{i\varphi} e^{i\psi} \quad \text{für alle } \varphi, \psi \in \mathbb{R}.$$





Wozu komplexe Zahlen? Nullstellen von Polynomen.
$z^2 = -1$ hat keine reellen Lösungen. In $\mathbb{C}$ sind $\pm i$ die Lösungen. Anders ausgedrückt: $z^2 + 1 = (z-i)(z+i)$.
$z^2 = re^{i\varphi}$ hat die zwei Lösungen $\pm\sqrt{r}e^{i\frac{\varphi}{2}}$.
$z^2 - 2z + 4 = 0$ hat die Lösungen $1 \pm \sqrt{1-4} = 1 \pm i\sqrt{3}$

Polynome

Sei $\mathbb{K}$ ein Körper, $a_0, a_1, \ldots, a_n \in \mathbb{K}$. Dann heißt

$$p = a_n X^n + \cdots + a_1 X + a_0 = \sum_{k=0}^{n} a_k X^k$$

Polynom (über $\mathbb{K}$) in der Unbestimmten X. $\mathbb{K}[X]$ ist die Menge aller Polynome in X. Der Grad von $p \in \mathbb{K}[X]$ ist

$$\deg p := \begin{cases} -\infty, & \text{falls } p = 0, \\ \max\{k \in \mathbb{N}_0 \mid a_k \neq 0\}, & \text{sonst.} \end{cases}$$

Zwei Polynome $p = a_n X^n + \cdots + a_1 X + a_0$ und $q = b_m X^m + \cdots + b_1 X + b_0$ heißen gleich, $p = q$, wenn $\deg p = \deg q$ und $a_j = b_j$ für $j = 0, \ldots, \deg p$.
Jedes Polynom $p = a_n X^n + \cdots + a_1 X + a_0$ definiert die zugehörige Polynomfunktion

$$\tilde{p} : \mathbb{K} \to \mathbb{K}, \qquad x \mapsto a_n x^n + \cdots + a_1 x + a_0$$

Beispiel: $\mathbb{K} = \mathbb{Z}_2 = \{0, 1\}$. $p = X^2 + X$. $\tilde{p}(0) = 0 + 0 = 0$, $\tilde{p}(1) = 1 \cdot 1 + 1 = 1 + 1 = 0$. D.h., $\tilde{p} = 0$, die Nullfunktion, obwohl $p \neq 0$.

Bemerkung. Für $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ folgt aus $\tilde{p} = \tilde{q}$ schon $p = q$, d.h. die Abbildung $\tilde{\ } : \mathbb{K}[X] \to \mathbb{K}^{\mathbb{K}}$, $p \mapsto \tilde{p}$, ist injektiv.

Addition und Multiplikation von Polynomen

Seien $p, q \in \mathbb{K}[X]$ wie oben, $p = a_n X^n + \cdots + a_1 X + a_0$, $q = b_m X^m + \cdots + b_1 X + b_0$ mit $m \leq n$, dann ist

$$\begin{aligned} p + q &:= (a_n + b_n) X^n + \cdots + (a_1 + b_1) X + (a_0 + b_0), \quad \text{wobei } b_{m+1}, \ldots, b_n := 0, \\ p \cdot q &:= c_{n+m} X^{n+m} + \cdots + c_1 X + c_0, \quad \text{mit } c_k = \sum_{i=0}^{k} a_i b_{k-i}, \\ & \text{wobei } a_j := 0 \text{ für } j > n \text{ und } b_j := 0 \text{ für } j > m. \end{aligned}$$

Es ist $c_0 = a_0 b_0$, $c_1 = a_0 b_1 + a_1 b_0$, $c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$, ..., $c_{n+m} = a_n b_m$.
Beispiel:$(X^2+1) + (2X-2) = X^2 + 2X - 1$, $(X^2+1)(2X-2) = 2X^3 - 2X^2 + 2X - 2$.

Bemerkung. $(\mathbb{K}[X], +, \cdot)$ ist ein kommutativer Ring mit Eins und

$$\deg(p \cdot q) = \deg p + \deg q \quad \text{für } p, q \in \mathbb{K}[X] \setminus \{0\}.$$

Satz 20 (Polynomdivision)**.** *Sind $f, g \in \mathbb{K}[X]$, $g \neq 0$, so gibt es eindeutig bestimmte $q, r \in \mathbb{K}[X]$ mit*

$$f = q \cdot g + r \quad \textit{und} \quad \deg r < \deg g.$$

Man schreibt auch $f \div g$ ist gleich q Rest r.

Beweis.
Eindeutigkeit:
Sei $f = q \cdot g + r = q' \cdot g + r'$, mit $\deg r, \deg r' < \deg g$. Somit ist auch $\deg(r' - r) < \deg g$.
Es folgt $0 = (q - q') \cdot g + (r - r')$, bzw., $(q - q') \cdot g = r' - r$.
Annahme: $q - q' \neq 0$. Dann ist $\deg('-r) = \deg(q - q') + \deg g \geq \deg g$, im Widerspruch zu $\deg(r - r') < \deg g$.
Also folgt $q = q'$ und damit auch $r - r' = (q - q') \cdot g = 0$, d.h., $r = r'$.
Existenz:
Seien $f = a_n X^n + \cdots + a_1 X + a_0$, $g = b_m X^m + \cdots + b_1 X + b_0$ mit $a_n \neq 0$, $b_m \neq 0$.
Ist $n < m$, so ist $q = 0$ und $r = f$, denn $f = 0 \cdot g + r$ und $\deg r = n < m = \deg g$.
Für $n \geq m$ setzt man $p_1 := \frac{a_n}{b_m} X^{n-m}$, $f_1 := f - p_1 \cdot g$. Somit ist $\deg f_1 < \deg f$. Deswegen können wir annehmen, dass der Satz für f_1 schon bewiesen ist.
Sei also $f_1 = q_1 \cdot g + r_1$ mit $\deg r_1 < \deg g$. Dann ist $q := p_1 + q_1$ und $r := r_1$, denn

$$q \cdot g + r = p_1 \cdot g + q_1 \cdot g + r_1 = p_1 \cdot g + f_1 = f \quad \text{und} \quad \deg r < \deg g$$

Für die Polynomdivision von f_1 durch g wendet man das gleiche Verfahren an. Da $\deg f_1 < \deg f$ kommt die Rekursion nach $n - m + 1$ Schritten zum Ende. $\square$

Beispiel. $f = 2X^3 + 3X + 1$, $g = X^2 - 2X$.

$$\begin{aligned}
f \div g \quad &= \quad (2X^3 + 0X^2 + 3X + 1) \div (X^2 - 2X) = \underbrace{\overbrace{2X}^{p_1} + \overbrace{4}^{q_1}}_{q} \quad \text{Rest} \quad \underbrace{11X + 1}_{r} \\
&\quad\; \underline{-(2X^3 - 4X^2)\qquad\qquad} \\
f_1 \quad &= \quad\qquad 4X^2 + 3X + 1 \\
&\quad\qquad \underline{-(4X^2 - 8X)\qquad\qquad} \\
r = r_1 \quad &= \quad\qquad\qquad 11X + 1
\end{aligned}$$

also ist $f = q \cdot g + r$, bzw., $2X^3 + 3X + 1 = (2X + 4)(X^2 - 2X) + (11X + 1)$

Definition 27. Ist $\mathbb{K}$ ein Körper, $f \in \mathbb{K}[X]$. $\lambda \in \mathbb{K}$ heißt Nullstelle von f, wenn $\tilde{f}(\lambda) = 0$.

Beispiele.
- Jedes Polynom vom Grad 1 hat genau eine Nullstelle. *Beweis:* $f = a_1 X + a_0$ mit $a_1 \neq 0$. Dann besitzt die Gleichung $\tilde{f}(\lambda) = 0$, d.i. $a_1 \lambda + a_0 = 0$ die eindeutige Lösung $\lambda = -\frac{a_0}{a_1}$.

- $X^2 - 2$ hat in $\mathbb{Z}_3$, $\mathbb{Q}$ keine, in $\mathbb{Z}_7$, $\mathbb{R}$ jeweils 2 Nullstellen, nämlich $3, 4$, bzw., $\pm\sqrt{2}$. 01.12.09
- $X^2 + 1$ hat in $\mathbb{Z}_3$, $\mathbb{Q}$, $\mathbb{R}$ keine, in $\mathbb{Z}_2$ eine, 1, und in $\mathbb{Z}_5$, $\mathbb{C}$ jeweils zwei Nullstellen, nämlich $2, 3$, bzw., $\pm i$.

Lemma 1. *Ist $\lambda \in \mathbb{K}$ eine Nullstelle von $f \in \mathbb{K}[X]$, dann gibt es genau ein $q \in \mathbb{K}[X]$ mit $f = q(X - \lambda)$ und $\deg q = \deg f - 1$.*

Beweis. Polynomdivision ergibt $f = q \cdot (X - \lambda) + r$ mit eindeutigen q und r. Wegen $\deg(X - \lambda) = 1$ ist $\deg r \leq 0$, d.h., $r = cX^0$ mit $c \in \mathbb{K}$. Nun ist

$$0 = \tilde{f}(\lambda) = \tilde{q}(\lambda)(\lambda - \lambda) + \tilde{r}(\lambda) = \tilde{r}(\lambda).$$

Somit ist $c = 0$ und damit $r = 0$. Außerdem gilt offenbar $\deg f = \deg q + 1$ □

Satz 21. *Sei $\mathbb{K}$ beliebiger Körper. Ein Polynom $f \in \mathbb{K}[X] \setminus \{0\}$ besitzt höchstens $\deg f$ Nullstellen.*

Beweis. Induktion über den Grad von f:
$\deg f = 0$: Dann ist $f = cX^0$ mit $c \in \mathbb{K} \setminus \{0\}$. f hat also keine Nullstellen. $\deg f = k+1$:
1. Fall: f besitzt keine Nullstellen. Dann ist nichts zu zeigen.
2. Fall: f besitzt eine Nullstelle λ. Dann gibt es g mit $g(X - \lambda = f$ und $\deg g = k$. Nach Induktionsvoraussetzung besitzt g höchstens k Nullstellen, somit besitzt f höchstens $k + 1$ Nullstellen. □

Korollar 2. *Ist $|\mathbb{K}| = \infty$, so ist die Abbildung*

$$\tilde{}: \mathbb{K}[X] \to \mathbb{K}^{\mathbb{K}}, \quad f \mapsto \tilde{f}$$

injektiv.

Beweis. Seien $f_1, f_2 \in \mathbb{K}[X]$ mit $\tilde{f}_1 = \tilde{f}_2$, d.h., $\widetilde{f_1 - f_2}(\lambda) = \tilde{f}_1(\lambda) - \tilde{f}_2(\lambda) = 0$ für alle $\lambda \in \mathbb{K}$. $f_1 - f_2$ besitzt also unendlich viele Nullstellen, somit ist $f_1 - f_2 = 0$, bzw. $f_1 = f_2$. □

Da $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ unendlich viele Elemente enthalten, beweist dies die Bemerkung im Anschluss an die Definition der Polynome.

Satz 22 (Fundamentalsatz der Algebra). *Jedes Polynom $f \in \mathbb{C}[X]$ mit $\deg f > 0$ hat mindestens eine Nullstelle.*

ohne Beweis. Der Beweis kann z.B. im Rahmen der Funktionentheorie geführt werden. Er übersteigt den Rahmen dieser Vorlesung.

Satz 23. *Jedes Polynom $f \in \mathbb{C}[X] \setminus \{0\}$ zerfällt in Linearfaktoren. D.h., es gibt $a, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$, wobei $n = \deg f$, so dass*

$$f = a(X - \lambda_1) \cdots (X - \lambda_n)$$

Beweis. Ist f vom Grad 0, so ist $f = a$ mit $a \in \mathbb{C} \setminus \{0\}$.
Sei die Aussage für Polynome vom Grad n bewiesen. Ist f vom Grad $n+1$, so besitzt f eine Nullstelle λ_{n+1}. Es ist also $f = (X - \lambda_{n+1})g$ und der Grad von g ist n. Die Behauptung für f folgt somit aus der Induktionsvoraussetzung. □

Bemerkung.
- Ist $f \in \mathbb{C}[X]$ mit reellen Koeffizienten, so ist mit $\lambda \in \mathbb{C}$ auch $\overline{\lambda}$ eine Nullstelle von f.
- Ein reelles Polynom $f \in \mathbb{R}[X]$ aufgefasst als $f \in \mathbb{C}[X]$ mit reellen Koeffizienten zerfällt in komplexe Linearfaktoren, die soweit nicht schon reell (reelle Nullstellen) paarweise zu reellen quadratischen Faktoren zusammengefasst werden können, denn $(X-\lambda)(X-\overline{\lambda}) = X^2 - (\lambda + \overline{\lambda})X + \lambda\overline{\lambda} = X^2 - 2\mathrm{Re}\,\lambda\, X + |\lambda|^2 \in \mathbb{R}[X]$

1.5 Vektorräume

Definition 28 (Vektorraum)**.** Sei $\mathbb{K}$ ein Körper. Eine Menge V mit zwei Abbildungen

$$\begin{array}{llll} + & : V \times V \to V, & (v,w) \mapsto v + w, & \text{(Vektor-)Addition} \\ \cdot & : \mathbb{K} \times V \to V, & (\lambda, v) \mapsto \lambda \cdot v, & \text{Skalarmultiplikation} \end{array}$$

kurz $(V, +, \cdot)$, heißt **Vektorraum**, wenn gilt

(V1) $(V, +)$ ist abelsche Gruppe,

(V2) Die Skalarmultiplikation hat folgende Verträglichkeitseigenschaften: für alle $\lambda, \mu \in \mathbb{K}$, $v, w \in V$ gilt:

(i) $1 \cdot v = v$, (iii) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,

(ii) $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$, (iv) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

Beispiel. $(\mathbb{R}^2, +, \cdot)$ ist ein $\mathbb{R}$-Vektorraum mit $\binom{x_1}{x_2} + \binom{y_1}{y_2} := \binom{x_1+y_1}{x_2+y_2}$ und $\lambda \cdot \binom{x_1}{x_2} := \binom{\lambda x_1}{\lambda x_2}$.

Schreibweisen

- Das neutrale Element der Gruppe $(V, +)$ heißt Nullvektor. Er wird durch $\mathbf{0}$, $\vec{0}$ oder 0 bezeichnet.
- Das additiv Inverse von $v \in V$ wird mit $-v$ bezeichnet. Man schreibt kurz $v - w$ für $v + (-w)$.
- Für $\lambda \cdot v$ schreibt man auch λv und gelegentlich, wenn keine Missverständnisse zu befürchten sind $v\lambda$.
- Wenn aus dem Kontext klar ist, welche Addition und Skalarmultiplikation gemeint ist, spricht man von dem $\mathbb{K}$-Vektorraum V. Ist auch der Körper ($\mathbb{R}$ oder $\mathbb{C}$) aus dem Kontext ersichtlich, spricht man von dem Vektorraum V.

Bemerkung. In jedem $\mathbb{K}$-Vektorraum V gilt für alle $\lambda \in \mathbb{K}$, $v \in V$

(i) $0 \cdot v = \vec{0}$,

(ii) $\lambda \cdot \vec{0} = \vec{0}$.

(iii) $\lambda \cdot v = 0 \Longrightarrow \lambda = 0$ oder $v = \vec{0}$.

(iv) $(-1) \cdot v = -v$.

Beweis.

(i) $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$. Subtraktion von $0 \cdot v$ auf beiden Seiten ergibt $\vec{0} = 0 \cdot v$.

(ii) , (iii), (iv) [HA]

□

Definition 29. Sei $\mathbb{K}$ ein Körper. $U \subset V$ heißt **Untervektorraum** von V, wenn gilt

(U1) $U \neq \emptyset$,

(U2) $\forall v, w \in U : v + w \in U$,

(U3) $\forall \lambda \in \mathbb{K} \forall v \in U : \lambda v \in U$

Man sagt, ein Untervektorraum ist abgeschlossen bezüglich Addition und Skalarmultiplikation.

Bemerkung.

- Ist U Untervektorraum von V, so ist U eine Untergruppe von $(V, +)$.
- $U \subset V$ ist Untervektorraum von V genau dann, wenn $0 \in U$ und mit $\lambda \in \mathbb{K}$, $v, w \in V$ auch $\lambda v + w \in U$.

Beweis. [HA] □

Beispiele. $U_1 = \{0\}$, $U_2 = \{x \in \mathbb{R}^2 | \, x_1 + x_2 = 0\}$ sind Untervektorräume von $\mathbb{R}^2$.
$U_3 = \{x \in \mathbb{R}^2 | \, x_1 + x_2 = 1\}$ ist kein Untervektorraum, da $0 \notin U_3$.

Satz 24. *Ein Untervektorraum U des Vektorraums V ist mit der auf U eingeschränkten Addition und Skalarmultiplikation ein Vektorraum über dem gleichen Körper.*

Beweis. $+|_{U \times U} : U \times U \to V$ und $\cdot|_{\mathbb{K} \times U} : \mathbb{K} \times U \to V$ haben ihre Bilder wieder in U wegen (U2),(U3). (V2) bleibt für die Einschränkung gültig.
Außerdem ist $(U, +)$ eine Untergruppe von $(V, +)$, also ist auch (V1) erfüllt. □

Beispiele. 03.12.09

- $(\mathbb{R}^n, +, \cdot)$ ist $\mathbb{R}$-Vektorraum. Für $A \in \mathbb{R}^{m \times n}$ ist $L_{(A|0)} = \{x \in \mathbb{R}^n \, | \, Ax = 0\}$ ein Untervektorraum von $\mathbb{R}^n$.
- $(\mathbb{R}^{m \times n}, +, \cdot)$ ist ein $\mathbb{R}$-Vektorraum.
  $\{A \in \mathbb{R}^{n \times n} \, | \, A^T = A\}$ ist ein Unterraum von $\mathbb{R}^{n \times n}$.

- $(\mathbb{R}^I, +, \cdot)$ ist ein $\mathbb{R}$-Vektorraum, wobei für $f, g : I \to \mathbb{R}$, $\lambda \in \mathbb{R}$

$$\begin{aligned}(f+g)(x) &:= f(x)+g(x),\\ (\lambda \cdot f)(x) &:= \lambda f(x).\end{aligned}$$

  Die Menge der Lipschitzstetigen (stetigen, differenzierbaren) reellwertigen Funktionen, definiert auf $\mathbb{R}$, ist ein Unterraum von $\mathbb{R}^{\mathbb{R}}$.

- Die Menge der konvergenten Folgen ist ein Unterraum der Menge aller reellen Zahlenfolgen, $\mathbb{R}^{\mathbb{N}}$.

- $\{0\} \subset \mathbb{R}^3$, die Geraden und Ebenen durch den Nullpunkt und $\mathbb{R}^3$ selbst sind Untervektorräume des $\mathbb{R}^3$.

Definition 30. Sei V ein $\mathbb{K}$-Vektorraum, $v_1, \dots, v_n \in V$, $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. Dann heißt

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = \sum_{k=1}^{n} \lambda_k v_k$$

eine **Linearkombination** der v_k, $k = 1, \dots, n$.
Sei $X \subset V$. Dann ist

$$\operatorname{span}_{\mathbb{K}} X := \left\{ \sum_{k=1}^{n} \lambda_k v_k : n \in \mathbb{N}_0,\ \lambda_1, \dots, \lambda_n \in \mathbb{K},\ v_1, \dots, v_n \in X \right\}$$

der Spann (oder die lineare Hülle) von X.

Bemerkung.

(i) $\operatorname{span}_{\mathbb{K}} X$ ist Unterraum von V.

(ii) Ist U unterraum von V mit $X \subset U$, so ist $\operatorname{span}_{\mathbb{K}} X \subset U$.

Beweis.

(i) $0 \in \operatorname{span}_{\mathbb{K}} X$, Sind v, w jeweils (endliche Linearkombinationen in X, so auch λv, $v + w$.

(ii) Der Unterraum U enthält mit den Elementen von X auch jede Linearkombination von (endlich vielen) Vektoren in X, also $\operatorname{span}_{\mathbb{K}} X \subset U$.

$\square$

Beispiele.

- $V = \mathbb{R}^3$, $v_1, v_2 \in \mathbb{R}^3$.
  Ist $v_1 \neq 0$, so ist $\operatorname{span}_{\mathbb{K}}(v_1)$ die Gerade durch den Ursprung des $\mathbb{R}^3$, in Richtung v_1.
  Ist $v_2 \neq \operatorname{span}_{\mathbb{K}}(v_1)$, so ist $\operatorname{span}_{\mathbb{K}}(v_1, v_2)$ die Ebene durch den Ursprung, die v_1 und v_2 enthält.

- $V = \mathbb{K}[X]$ ist $\mathbb{K}$-Vektorraum mit der Addition von Polynomen und der Skalarmultiplikation $\lambda(a_n X^n + \cdots a_1 X + a_0) = \lambda a_n X^n + \cdots \lambda a_1 X + \lambda a_0$. Setzt man $v_k := X^k$, $k \in \mathbb{N}_0$, so ist $\operatorname{span}(v_1, \ldots, v_n)$ der Raum der Polynome vom Grad $\leq n$ und $\operatorname{span}(v_k)_{k \in \mathbb{N}_0} = \mathbb{K}[X]$.

Man sagt die v_i, $i \in I$, spannen den Unterraum $\operatorname{span}(v_i)_{i \in I}$ auf.
Eine Linearkombination ist immer eine gewichtete Summe **endlich** vieler Vektoren. "Unendliche Summen" werden in der Analysis definiert (Konvergenz von Reihen).

Lineare Unabhängigkeit

Definition 31. Sei V ein $\mathbb{K}$-Vektorraum. Das endliche Tupel $(v_1, \ldots, v_n)$ von Vektoren in V heißt **linear unabhängig**, wenn aus

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

schon

$$\lambda_1 = 0,\ \lambda_2 = 0,\ \ldots,\ \lambda_n = 0$$

folgt.
Eine beliebige Familie von Vektoren $(v_i)_{i \in I}$ heißt **linear unabhängig**, wenn jede endliche Teilfamilie linear unabhängig ist.
Sie heißt **linear abhängig**, wenn sie nicht linear unabhängig ist. In diesem Fall gibt es $v_{i_1}, \ldots, v_{i_n}$, und $\lambda_1, \ldots, \lambda_n \in K$ die nicht alle Null sind mit

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

(mit anderen Worten: Es gibt eine nichttriviale Linearkombination des Nullvektors).

Lemma 2. *Genau dann ist $(v_i)_{i \in I}$ linear unabhängig, wenn sich jedes $v \in \operatorname{span}(v_i)_{i \in I}$ eindeutig als Linearkombination in $(v_i)_{i \in I}$ schreiben lässt.*

Beweis. "$\Leftarrow$": Statt $A \Leftarrow B$ zeigt man $\neg A \Rightarrow \neg B$.
$\neg A$ ist die Aussage, es gibt $v_{i_1}, \ldots, v_{i_n}$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K} \setminus \{0\}$ mit $\lambda_1 v_{i_1} + \cdots + \lambda_n v_{i_n} = 0$. Außerdem ist $0 v_{i_1} + \cdots + 0 v_{i_n} = 0$. Der Nullvektor lässt sich also auf mehr als eine Art als Linearkombination darstellen., das ist aber $\neg B$.
"$\Rightarrow$": Sei $v = \sum\limits_{i \in J_1} \lambda_i v_i = \sum\limits_{i \in J_2} \mu_i v_i \in \operatorname{span}(v_i)_{i \in I}$ mit $J_1, J_2 \subset I$, $|J_1|, |J_2| < \infty$. Dann ist auch $J := J_1 \cup J_2$ endlich und

$$0 = \sum_{i \in J} (\lambda_i - \mu_i) v_i, \qquad \text{wobei } \lambda_i := 0 \text{ für } i \in J \setminus J_2,\ \mu_i := 0 \text{ für } i \in J \setminus J_1.$$

Wegen der linearen Unabhängigkeit folgt $\lambda_i = \mu_i$ für alle $i \in J$, d.h., beide Linearkombinationen sind identisch. $\square$

Beispiele.

- In dem $\mathbb{K}$-Vektorraum $\mathbb{K}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{K} \right\}$ mit komponentenweiser Addition und Skalarmultiplikation,

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix}, \qquad \lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix},$$

definiert man die Einheitsvektoren $e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ mit der 1 in der i-ten Zeile. Dann ist mit $I = \{1, \dots, n\}$

$$\operatorname{span}(e_i)_{i \in I} = \operatorname{span}(e_1, \dots, e_n) = \mathbb{K}^n,$$

denn

$$x \in \mathbb{K}^n \;\Leftrightarrow\; x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\Leftrightarrow\; x = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \operatorname{span}(e_i)_{i \in I}.$$

Außerdem ist $(e_i)_{i \in I}$ linear unabhängig, denn

$$0 = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_n = 0.$$

- Analog gilt für $\mathbb{K}_n = \{\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{K}\}$, dass $K_n = \operatorname{span}(e_i^T)_{i \in I}$ und $(e_i^T)_{i \in I}$ ist linear unabhängig.

- In einer Matrix in Zeilenstufenform sind die von Null verschiedenen Zeilen linear unabhängig. $\boxed{08.12.09}$

Beispiel. Die Matrix $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ & 1 & 2 & 3 & 4 \\ & & & 1 & 2 \\ & & & & 0 \end{pmatrix}$ ist in Zeilen-Stufen-Form.

Aus $\lambda_1 \begin{pmatrix}1 & 2 & 3 & 4 & 5\end{pmatrix} + \lambda_2 \begin{pmatrix}0 & 1 & 2 & 3 & 4\end{pmatrix} + \lambda_3 \begin{pmatrix}0 & 0 & 0 & 1 & 2\end{pmatrix} = 0$ folgt

$$\begin{aligned} \lambda_1 \cdot 1 + \lambda_2 \cdot 0 + \lambda_3 \cdot 0 = 0 &\quad \text{also} \quad \lambda_1 = 0 \\ \lambda_1 \cdot 2 + \lambda_2 \cdot 1 + \lambda_3 \cdot 0 = 0 &\quad \text{also} \quad \lambda_2 = 0 \\ &\vdots \\ \lambda_1 \cdot 4 + \lambda_2 \cdot 2 + \lambda_3 \cdot 1 = 0 &\quad \text{also} \quad \lambda_3 = 0 \end{aligned}$$

- In $\mathbb{K}[X]$ sind die Monome $(X^n)_{n \in \mathbb{N}_0}$ linear unabhängig:

$$a_n X^N + \cdots + a_1 X^1 + a_0 X^0 = 0 \qquad \Rightarrow \qquad a_a = 0, \ldots, a_1 = 0, a_0 = 0.$$

Bemerkung.

(i) Verkleinert man eine linear unabhängige Familie von Vektoren, so bleibt sie linear unabhängig.

(ii) Vergrößert man eine linear abhängige Familie von Vektoren, so bleibt sie linear abhängig.

(iii) (v_1) ist linear abhängig, genau dann, wenn $v_1 = 0$.

(iv) Für $n \geq 2$ sind $v_1, \ldots, v_n$ genau dann linear abhängig, wenn einer dieser Vektoren als Linearkombination der anderen dargestellt werden kann.

Beweis. (i) und (ii) sind klar wegen der Definition.

(iii) "$\Rightarrow$": (v_1) linear abhängig heißt, es gibt ein $\lambda_1 \in \mathbb{K} \setminus \{0\}$ mit $\lambda_1 v_1 = 0$, also ist $v_1 = 0$.
"$\Leftarrow$": $v_1 = 0 \Rightarrow 1 \cdot v_1 = 0$.

(iv) "$\Rightarrow$": $v_1, \ldots, v_n$ linear abhängig heißt, es gibt λ_i mit $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ und für ein $i \in \{1, \ldots, n\}$ ist $\lambda_i \neq 0$. Somit ist

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \cdots - \frac{\lambda_n}{\lambda_i} v_n,$$

das ist eine Linearkombination der anderen Vektoren.
"$\Leftarrow$": Sei $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1} + \mu_{i+1} v_{i+1} + \cdots + \mu_n v_n$. Dann ist

$$\mu_1 v_1 + \cdots + \mu_i v_i + \cdots + \mu_n v_n = 0$$

mit $\mu_i = -1$ eine nichttriviale Linearkombination des Nullvektors.

$\square$

Warnung: (iv) bedeutet, dass $(v_1, \ldots, v_n)$ linear unabhängig sind, wenn keiner der Vektoren als Linearkombination der anderen geschrieben werden kann.
Dies ist eine sehr anschauliche Eigenschaft der linearen Unabhängigkeit. Allerdings ist sie für $n \geq 3$ meist **völlig** unpraktikabel zum Nachweis der linearen Unabhängigkeit.

1.6 Basis und Dimension

Definition 32. Eine Familie $(b_i)_{i\in I}$ von Vektoren eines Vektorraums V heißt **Erzeugendensystem**, wenn $V = \operatorname{span} B$.
Der Vektorraum V heißt **endlich erzeugt**, falls er ein endliches Erzeugendensystem besitzt, d.h., es gibt $B = (b_1, \dots, b_n)$ mit $V = \operatorname{span} B$.
B heißt **Basis** von V, wenn B ein linear unabhängiges Erzeugendensystem ist.
$|B| := |\{b_i : i \in\}|$ heißt die Länge der Basis.

Beispiele.

- $(e_i)_{i\in\{1,\dots,n\}}$ ist Basis des $\mathbb{K}^n$, die **kanonische Basis** oder **Standardbasis** des $\mathbb{K}^n$.
- Die Matrizen $E_{ij} = \begin{pmatrix} & & 0 & & \\ & & \vdots & & \\ 0 & \cdots & 1 & \cdots & 0 \\ & & \vdots & & \\ & & 0 & & \end{pmatrix}$ mit genau einer 1 in der i-ten Zeile und j-ten Spalte bilden eine Basis des $\mathbb{K}^{m\times n}$ mit Länge mn.
- $(X^0, X^1, X^2, \dots)$ ist eine Basis von $K[X]$ mit Länge ∞.
- $(1, i)$ ist Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum.

Satz 25. *Ist $B = (b_1, \dots, b_n)$ eine Basis von V, so ist die Abbildung*

$$\mathbb{K}^n \ni (\lambda_1, \dots, \lambda_n) \mapsto \lambda_1 b_1 + \dots + \lambda_n b_n \in V$$

bijektiv.

Beweis. B ist Erzeugendensystem, also ist die Abbildung surjektiv.
Die Darstellung als Linearkombination ist auch eindeutig, da B linear unabhängig ist (Lemma 2). Daraus folgt die Injektivität. □

Die Umkehrabbildung

$$\cdot_B : V \to \mathbb{K}^n, \qquad v \mapsto {}_B v = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}, \quad \text{so dass } v = \lambda_1 b_1 + \dots + \lambda_n b_n,$$

definiert den **Koordinatenvektor** ${}_B v \in \mathbb{K}^n$ von $v \in V$ bezüglich der Basis B.

Beispiele. Für $E = (e_i)_{i=1,\dots,n}$, die Standardbasis des $\mathbb{K}^n$ gilt ${}_B v = v \in \mathbb{K}^n$.
Für den Unterraum $U = \operatorname{span}(1, X, \dots, X^5)$ von $K[X]$ mit der Basis $(1, X, \dots, X^5)$ ist der Koordinatenvektor, z.B. von $2X^3 - X + 7 \in V$ gleich $(7, -1, 0, 2, 0, 0)$.

Satz 26. *Sei $B = (b_1, \dots, b_n)$ eine Basis von V und $X = (x_1, \dots, x_m)$ in V.*

(i) Ist $m > n$, so ist X linear abhängig.

(ii) Ist $m < n$, so ist X kein Erzeugendensystem.

(iii) Ist X eine Basis von V, so ist $m = n$.

Beweis. (i) Sei $m > n$. Da B Basis ist sind alle x_i Linearkombinationen der b_j, d.h.,

$$\begin{array}{rcl} x_1 & = & a_{11}b_1 + \cdots + a_{1n}b_n \\ \vdots & \vdots & \qquad \vdots \\ x_m & = & a_{m1}b_1 + \cdots + a_{mn}b_n \end{array}$$

Das lineare Gleichungssystem

$$\begin{array}{rcl} a_{11}\lambda_1 + \cdots + a_{m1}\lambda_m & = & 0 \\ \vdots \qquad & \vdots & \vdots \\ a_{1n}\lambda_1 + \cdots + a_{mn}\lambda_m & = & 0 \end{array}$$

besitzt nichttriviale Lösungen $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m$, da es wegen $m > n$ aus weniger Gleichungen als Unbekannten besteht. Somit ist

$$\begin{array}{rcl} \tilde{\lambda}_1 x_1 + \cdots + \tilde{\lambda}_m x_m & = & (\tilde{\lambda}_1 a_{11} + \cdots + \tilde{\lambda}_m a_{m1})b_1 \\ & & + \cdots + \\ & & +(\tilde{\lambda}_1 a_{1n} + \cdots + \tilde{\lambda}_m a_{mn})b_n = 0, \end{array}$$

eine nichttriviale Linearkombination der 0. X ist also linear abhängig.

(ii) Sei $m < n$. *Annahme:* X ist ein Erzeugendensystem, d.h., es gibt Koeffizienten $a_{ij} \in \mathbb{K}$ mit

$$\begin{array}{rcl} b_1 & = & a_{11}x_1 + \cdots + a_{m1}x_n \\ \vdots & \vdots & \qquad \vdots \\ b_n & = & a_{1n}x_1 + \cdots + a_{mn}x_m \end{array}$$

Das lineare Gleichungssystem

$$\begin{array}{rcl} a_{11}\lambda_1 + \cdots + a_{1n}\lambda_n & = & 0 \\ \vdots \qquad & \vdots & \vdots \\ a_{m1}\lambda_1 + \cdots + a_{mn}\lambda_n & = & 0 \end{array}$$

besitzt nichttriviale Lösungen $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$, da es wegen $m < n$ aus weniger Gleichungen als Unbekannten besteht. Somit ist

$$\begin{array}{rcl} \tilde{\lambda}_1 b_1 + \cdots + \tilde{\lambda}_n b_n & = & (a_{11}\tilde{\lambda}_1 + \cdots + a_{1n}\tilde{\lambda}_n)x_1 \\ & & + \cdots + \\ & & +(a_{m1}\tilde{\lambda}_1 + \cdots + a_{mn}\tilde{\lambda}_n)x_m = 0, \end{array}$$

eine nichttriviale Linearkombination der 0. B ist also linear abhängig im Widerspruch dazu, das B eine Basis ist. Somit ist X kein Erzeugendensystem.

(iii) X ist linear unabhängig $\overset{\text{(i)}}{\Rightarrow}$ $m \leq n$. 10.12.09

X ist Erzeugendensystem $\overset{\text{(ii)}}{\Rightarrow}$ $m \geq n$.

□

Bemerkung. Für $A \in \mathbb{K}^{m \times n}$, $m < n$, besitzt $Ax = 0$ immer eine nichttriviale Lösung $x \in \mathbb{K}^n \setminus \{0\}$, denn in der reduzierten Zeilen-Stufen-Form von A gibt es mindestens eine Spalte ohne führende Eins.

Definition 33. Besitzt der $\mathbb{K}$-Vektorraum V eine endliche Basis $(b_1, \ldots, b_n)$, so heißt

$$\dim_{\mathbb{K}} V := n \in \mathbb{N}_0$$

die Dimension von V, andernfalls setzt man $\dim_{\mathbb{K}} V := \infty$.

Beispiele. $\dim_{\mathbb{K}} \mathbb{K}^n = n$, $\dim_{\mathbb{K}} \mathbb{K}^{m \times n} = mn$, $\dim_{\mathbb{K}} \mathbb{K}[X] = \infty$,
$\dim_{\mathbb{R}} \mathbb{C} = 2$, da $(1, i)$ eine Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$ ist. Andererseits gilt natürlich $\dim_{\mathbb{C}} \mathbb{C} = 1$.

Satz 27. *Ist V endlich erzeugt, so kann man*

(i) aus jedem endlichen Erzeugendensystem eine Basis auswählen,

(ii) linear unabhängige Vektoren zu einer Basis ergänzen.

Beweisskizze. (i) Ist das Erzeugendensystem nicht linear unabhängig, so ist mindestens ein Vektor enthalten, der Linearkombination der anderen ist. Durch sein Weglassen bleiben die Restlichen ein Erzeugendensystem. Auf diese Art entfernt man nacheinander einzelne Vektoren bis der Rest nach endlich vielen Schritten linear unabhängig und damit eine Basis ist.

(ii) Bilden die Vektoren kein Erzeugendensystem, dann gibt es einen Vektor, der nicht in ihrem Spann enthalten ist. Zusammen mit diesem bleiben sie linear unabhängig. Auf diese Art fügt man die lineare Unabhängigkeit erhaltend Vektoren hinzu, bis man nach endlich vielen Schritten ein Erzeugendensystem und damit eine Basis erhält.

□

Korollar 3.

- *Jeder endlich erzeugte Vektorraum besitzt eine Basis.*
- *Für jeden Unterraum U von V gilt* $\dim_{\mathbb{K}} U \leq \dim_{\mathbb{K}} V$.

Satz 28. *Sei V ein n-dimensionaler Vektorraum.*

- *Ein Erzeugendensystem mit n Vektoren ist schon eine Basis.*
- *n linear unabhängige Vektoren bilden schon eine Basis.*

Beweis.
- sonst gäbe es eine Basis der Länge $m < n$.
- sonst gäbe es eine Basis der Länge $m > n$.

$\square$

Anwendung auf Matrizen und lineare Gleichungssysteme

Definition 34. Sei $A \in \mathbb{K}^{m \times n}$, und B eine Zeilen-Stufen-Form von A. Weiter sei $A_{*j} := \begin{pmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{pmatrix} \in \mathbb{K}^n$ der j-te Spaltenvektor und $A_{i*} := \begin{pmatrix} A_{i1} & \cdots & A_{in} \end{pmatrix} \in \mathbb{K}_m$ der i-te Zeilenvektor von A. Dann ist

$$\begin{aligned} \mathrm{SR}\,(A) &:= \operatorname{span}(A_{*1}, \ldots, A_{*n}) \subset \mathbb{K}^m \quad \text{Spaltenraum}, \\ \mathrm{ZR}\,(A) &:= \operatorname{span}({A_{1*}}^T, \ldots, {A_{m*}}^T) \subset \mathbb{K}^n \quad \text{Zeilenraum}, \\ \mathrm{NR}\,(A) &:= \{x \in \mathbb{K}^n \mid Ax = 0\} \subset K^n \quad \text{Nullraum} \end{aligned}$$

von A.

Satz 29. *Sei $A \in \mathbb{K}^{m \times n}$, B Zeilen-Stufen-Form von A.*

(i) $\mathrm{NR}\,(A) = \mathrm{NR}\,(B)$, $\mathrm{ZR}\,(A) = \mathrm{ZR}\,(B)$.

(ii) Die Zeilen von B mit führender Eins bilden eine Basis von $\mathrm{ZR}\,(A)$.

(iii) Ist B in reduzierter Zeilen-Stufen-Form, so liest man direkt eine Basis von $\mathrm{NR}\,(A)$ *ab.*

(iv) Diejenigen Spalten von A, in denen B eine führende Eins enthält, bilden eine Basis von $\mathrm{SR}\,(A)$.

Beweisskizze.

(i) $\mathrm{NR}\,(A) = \mathrm{NR}\,(B)$: Elementare Zeilenumformungen erhalten den Lösungsraum von $Ax = 0$ (=Nullraum), Abschnitt 0.5, Satz 1.
$\mathrm{ZR}\,(A) = \mathrm{ZR}\,(B)$: Elementare Zeilenumformungen erhalten den Zeilenraum von A.

(ii) Die Zeilen von B mit führender Eins sind linear unabhängig und spannen $\mathrm{ZR}\,(B) = \mathrm{ZR}\,(A)$ auf, bilden also eine Basis.

(iii) Zu jeder Spalte ohne führende Eins erhält man einen Lösungsvektor von $Bx = 0$. Sie spannen den Nullraum von B und damit auch A auf und sind offenbar linear unabhängig, also eine Basis.

(iv) Spaltenvektoren von A sind genau dann linear unabhängig, wenn die entsprechenden Spaltenvektoren von B linear unabhängig sind. Dies gilt, wenn B elementare Zeilenumformung von A ist, und damit auch, wenn $A \sim B$. Insbesondere ist $\dim \mathrm{SR}\,(A) = \dim \mathrm{SR}\,(B)$.

□

Beispiel. Für die Matrix

$$A = \begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{pmatrix} \sim B = \underbrace{\begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{\text{Zeilen-Stufen-Form}} \sim \underbrace{\begin{pmatrix} 1 & 0 & 0 & -2 & 3 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{\text{reduzierte Zeilen-Stufen-Form}}$$

ist

$$\begin{aligned}
\mathrm{ZR}\,(A) &= \operatorname{span}\left(\begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ -5 \\ -3 \\ -2 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 15 \\ 10 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 18 \\ 8 \\ 6 \end{pmatrix} \right) = \operatorname{span} \underbrace{\left(\begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right)}_{\text{Basis}} \\
&= \operatorname{span} \underbrace{\left(\begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right)}_{\text{Basis}} \\
\mathrm{NR}\,(A) &= \operatorname{span} \underbrace{\left(\begin{pmatrix} 2 \\ 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)}_{\text{Basis}}, \qquad \mathrm{SR}\,(B) = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \end{pmatrix} \right) \\
\mathrm{SR}\,(A) &= \operatorname{span}\left(\begin{pmatrix} 1 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -5 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 15 \\ 18 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 10 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \\ 0 \\ 6 \end{pmatrix} \right) \\
&\qquad = \operatorname{span} \underbrace{\left(\begin{pmatrix} 1 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -5 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 15 \\ 18 \end{pmatrix} \right)}_{\text{Basis}}
\end{aligned}$$

Bemerkung.

(i) $\dim \mathrm{ZR}\,(A) = \dim \mathrm{SR}\,(A) = \operatorname{rg} A$ (Zeilenrang gleich Spaltenrang)

(ii) $\dim \mathrm{SR}\,(A) + \dim \mathrm{NR}\,(A) = n$, die Spaltenzahl von A.

Beweisskizze. (i) folgt daraus, dass die Anzahl der Zeilen mit führender Eins gleich der Anzahl der Spalten mit führender Eins ist.

(ii) $$\underbrace{\dim \mathrm{SR}\,(A)}_{\text{Anzahl der Spalten mit führender Eins}} + \underbrace{\dim \mathrm{NR}\,(A)}_{\text{Anzahl der Spalten ohne führende Eins}} = \underbrace{n}_{\text{Spaltenzahl von } A}$$

□

Zusammenfassung von Kapitel 1

- Mathematische Strukturen (Gruppen, Ringe, Körper) erlauben das Rechnen und das Lösen von Gleichungen mit Objekten, wie wir es mit Zahlen gewohnt sind.
- Polynome über $\mathbb{C}$ besitzen immer Nullstellen und können daher vollständig in Linearfaktoren zerlegt werden.
- In einem Vektorraum ergeben Linearkombinationen von Vektoren wieder Vektoren. Um alle Vektoren eindeutig als Linearkombination einiger weniger darzustellen, braucht man eine Basis.
- In jedem n-dimensionalen Vektorraum gibt es Basen, die alle die Länge n haben.
- Gauß-Elimination bei linearen Gleichungssystemen liefert explizit Basen von Untervektorräumen.

2 Lineare Abbildungen

15.12.09

2.1 Definition und Beispiele

Definition 35. Eine Abbildung $F : V \to W$ zwischen zwei $\mathbb{K}$-Vektorräumen V und W heißt linear (genauer $\mathbb{K}$-linear), wenn gilt:

(L1) $\forall v, w \in V : F(v + w) = F(v) + F(w)$,

(L2) $\forall \lambda \in \mathbb{K} \forall v \in V : F(\lambda v) = \lambda F(v)$.

Sprechweise:

- Zu “F linear” sagt man auch F ist ein Homomorphismus (von $\mathbb{K}$-Vektorräumen)
- Eine lineare Abbildung $F : V \to W$ heißt
  - Isomorphismus, wenn F bijektiv ist,
  - Endomorphismus, wenn $V = W$,
  - Automorphismus, wenn F bijektiv ist und $V = W$.

Bemerkung. Ist $F : V \to W$ linear, dann gilt

(i) $F(0) = 0$,

(ii) $F\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i F(v_i)$

(iii) Sind $V' \subset V$ und $W' \subset W$ Unterräume, so auch $F(V') \subset W$ und $F^{-1}(W') \subset V$,

(iv) Ist F ein Isomorphismus, so auch F^{-1}.

Beweis. (i) $F(0) = F(0 \cdot 0) = 0 \cdot F(0) = 0$,

(ii) Induktion über n,

(iii) Wegen (i) ist $0 \in F(V')$, also nichtleer. Sei $\lambda \in \mathbb{K}$, $w, w' \in F(V')$ mit $w = F(v)$, $w' = F(v')$, $v, v' \in V'$. Dann ist $\lambda w + w' = \lambda F(v) + F(v') = F(\underbrace{\lambda v + v'}_{\in V'}) \in F(V')$.

Somit ist $F(V')$ ein Untervektorraum.
Wegen (i) ist $0 \in F^{-1}(W')$, also nichtleer. Seien $\lambda \in \mathbb{K}$, $v, v' \in F^{-1}(W')$. Dann ist $F(\lambda v + v') = \lambda \underbrace{F(v)}_{\in W'} + \underbrace{F(v')}_{\in W'} \in W'$, also ist $\lambda v + v' \in F^{-1}(W')$.

(iv) Wie bei den Gruppenhomomorphismen.

□

Bemerkung. Sei $F : V \to W$ linear. Dann gilt

(i) $v_1, \dots, v_n \in V$ linear abhängig $\Longrightarrow F(v_1), \dots, F(v_n) \in W$ linear abhängig,

(ii) $F(v_1), \dots, F(v_n) \in W$ linear unabhängig $\Longrightarrow v_1, \dots, v_n \in V$ linear unabhängig.

Beweis. [HA] □

Bemerkung. Sind $F : U \to V$, $G : V \to W$ linear, so auch $G \circ F : U \to W$.

Beweis. Elementar. □

Beispiele.

- Für $A \in \mathbb{K}^{m\times n}$ ist $f : \mathbb{K}^n \times \mathbb{K}^m$, $f(x) = Ax$ linear.
- Die Nullabbildung $0 : V \to W$, $0(v) = 0$, ist linear.
- Die identische Abbildung $\mathrm{id} : V \to V$ ist linear.
- Für $\lambda \in \mathbb{K}$ ist $F : V \to V$, $F(v) = \lambda v$, linear.
- $\frac{d}{dx} : \mathbb{K}[X] \to \mathbb{K}[X]$, $a_nX^n + \dots + a_2X^2 + a_1X + a_0 \mapsto na_nX^{n-1} + \dots + 2a_2X + a_1$ ist linear (Ableitung von Polynomen).
- Die Verschiebung um $v \in \mathbb{R}^2 \setminus \{0\}$, $\mathbb{R}^2 \ni x \mapsto x + v \in \mathbb{R}^2$ ist *nicht* linear.
- Die Rotation um den Winkel φ, $\mathbb{R}^2 \ni x \mapsto \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} x \in \mathbb{R}^2$ ist linear.

Definition 36. Sei $F : V \to W$ linear, dann heißt

$\mathrm{Im}\, F := F(V)$ das Bild von F,

$\mathrm{rg}\, F := \dim \mathrm{Im}\, F$, der Rang von F,

$\mathrm{Ker}\, F := F^{-1}(\{0\})$ der Kern von F.

Bemerkung. (a) $\mathrm{Im}\, F \subset W$ und $\mathrm{Ker}\, F \subset V$ sind Untervektorräume.

(b) F surjektiv $\Leftrightarrow \mathrm{Im}\, F = W$,

(c) F injektiv $\Leftrightarrow \mathrm{Ker}\, F = \{0\}$.

Beweis. Alles klar, bis auf
(c) “$\Leftarrow$”: Sei $\mathrm{Ker}\, F = \{0\}$, $x, y \in V$ mit $F(x) = F(y)$. Dann ist $F(x - y) = 0$, also $x - y \in \mathrm{Ker}\, F$ und damit $x - y = 0$, bzw., $x = y$. □

Definition 37. Seien V, W $\mathbb{K}$-Vektorräume. Dann ist

$$\mathrm{Hom}(V, W) := \{f : V \to W \mid f \text{ ist linear}\}$$

die Menge aller linearen Abbildungen von V nach W.

Bemerkung. $\mathrm{Hom}(V, W)$ ist mit der Addition $(f+g)(v) := f(v) + g(v)$ und der Skalarmultiplikation $(\lambda f)(v) := \lambda f(v)$ ein $\mathbb{K}$-Vektorraum.

2.2 Lineare Abbildungen und Matrizen

Satz 30. *Seien V, W endlichdimensionale Vektorräume, $(v_1, \dots, v_n)$ eine Basis von V, $w_1, \dots, w_n \in W$. Dann gibt es genau eine lineare Abbildung $F : V \to W$ mit $F(v_i) = w_i$, $i = 1, \dots, n$, und es gilt* $\mathrm{Im}\, F = \mathrm{span}(w_1, \dots, w_n)$.
Ist $(w_1, \dots, w_n)$ linear unabhängig, so ist F injektiv.

Beweis. Sei $v \in V$ mit $v = \lambda_1 v_1 + \dots + \lambda_n v_n$. Definiere $F(v) = \lambda_1 w_1 + \dots + \lambda_n w_n$. Für das so definierte F gilt $F(v_i) = w_i$ und F ist linear, denn seien $\lambda \in \mathbb{K}$, $v, v' \in V$ mit $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, $v' = \lambda'_1 v_1 + \dots + \lambda'_n v_n$. Dann ist

$$\begin{aligned} F(\lambda v + v') &= F\big((\lambda\lambda_1 + \lambda'_1)v_1 + \dots + (\lambda\lambda_n + \lambda'_n)v_n\big) \\ &= (\lambda\lambda_1 + \lambda'_1)w_1 + \dots + (\lambda\lambda_n + \lambda'_n)w_n = \lambda F(v) + F(v'). \end{aligned}$$

Zu zeigen: $\mathrm{Im}\, F = \mathrm{span}(w_1, \dots, w_n)$
"$\subset$": Sei $w \in \mathrm{Im}\, F$, d.h. es gibt ein $v \in V$ mit $F(v) = w$. Ist $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, so ist $w = \lambda_1 w_1 + \dots + \lambda_n w_n$, also $w \in \mathrm{span}(w_1, \dots, w_n)$.
"$\supset$": Sei $w = \lambda_1 w_1 + \dots + \lambda_n w_n$, somit ist $w = F(\underbrace{\lambda_1 v_1 + \dots + \lambda_n v_n}_{\in V}) \in \mathrm{Im}\, F$.

Ist $(w_1, \dots, w_n)$ linear unabhängig, so folgt aus $F(v) = 0$ mit $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, dass $\lambda_1 w_1 + \dots + \lambda_n w_n = 0$ ist, also $\lambda_1 = \dots = \lambda_n = 0$. Somit ist $v = 0$. □

Satz 31. *Seien V, W endlichdimensionale $\mathbb{K}$-Vektorräume. Dann ist* $\dim V = \dim W$ *genau dann, wenn es einen Isomorphismus $f : V \to W$ gibt.*

Beweis. "$\Rightarrow$": Sei $(v_1, \dots, v_n)$ Basis von V und $(w_1, \dots, w_n)$ Basis von W. Dann definiert $f(v_i) := w_i$ einen Isomorphismus, denn f ist als lineare Abbildung eindeutig bestimmt, $\mathrm{Im}\, f = \mathrm{span}(w_1, \dots, w_n) = W$ und $\mathrm{Ker}\, F = 0$, da die $w_1, \dots, w_n$ linear unabhängig sind.
"$\Leftarrow$": Sei $f : V \to W$ Isomorphismus und $(v_1, \dots, v_n)$ Basis von V. Dann ist $(w_1, \dots, w_n)$ mit $w_i := f(v_i)$ eine Basis von W. $\mathrm{span}(w_1, \dots, w_n) = W$ gilt, da f surjektiv ist. Andererseits folgt aus $\lambda_1 w_1 + \dots + \lambda_n w_n = 0$, dass $f(\lambda_1 v_1 + \dots + \lambda_n v_n) = 0$, also wegen f injektiv $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ und damit $\lambda_1 = \dots = \lambda_n = 0$. D.h. $(w_1, \dots, w_n)$ ist linear unabhängig und damit Basis, $\dim W = n = \dim V$. □

Bemerkung. 17.12.09

- Man sagt, die Vektorräume V und W sind isomorph, $V \simeq W$, wenn es einen Isomorphismus $f : V \to W$ gibt. Sind V, W endlichdimensional, so ist also $V \simeq W$ genau dann, wenn $\dim V = \dim W$.

- Isomorphie von Vektorräumen ist eine Äquivalenzrelation.
($V \simeq V$, $V \simeq W \;\Rightarrow\; W \simeq V$, $U \simeq V \;\wedge\; V \simeq W \;\Rightarrow\; U \simeq W$)

- Ist V n-dimensional mit Basis $B = (b_1, \ldots, b_n)$, so ist $V \ni v \mapsto {}_B v \in \mathbb{K}^n$ ein Isomorphismus zwischen V und $\mathbb{K}^n$.

- Jeder n-dimensionale $\mathbb{K}$-Vektorraum ist isomorph zum $\mathbb{K}^n$.

Bemerkung. Seien V, W $\mathbb{K}$-Vektorräume mit Basen $b = (b_1, \ldots, b_n)$ von V und $c = (c_1, \ldots, c_m)$ von W.
Für $A \in \mathbb{K}^{m \times n}$ sei die lineare Abbildung $F : V \to W$ definiert durch

$$F(b_j) := A_{1j}c_1 + \cdots + A_{mj}c_j, \qquad j = 1, \ldots, n.$$

Dann ist für beliebiges $v \in V$ mit $v = \lambda_1 b_1 + \cdots \lambda_n b_n$

$$F(v) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} A_{ij}c_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij}\lambda_j \right) c_i.$$

Wegen ${}_b v = \lambda \in \mathbb{K}^n$ gilt also ${}_c(F(v)) = A({}_b v)$.

Satz 32. *Seien V, W $\mathbb{K}$-Vektorräume mit Basen $b = (b_1, \ldots, b_n)$ von V und $c = (c_1, \ldots, c_m)$ von W. Zu jeder linearen Abbildung $F : V \to W$ gibt es genau eine Matrix $A \in \mathbb{K}^{m \times n}$ mit*

$$F(b_j) = A_{1j}c_1 + \cdots + A_{mj}c_m = \sum_{i=1}^{m} c_i A_{ij}, \quad j = 1, \ldots, n.$$

Die Abbildung

$$\mathrm{Hom}(V, W) \ni F \mapsto {}_c F_b := A \in \mathbb{K}^{m \times n}$$

ist ein Vektorraumisomorphismus. ${}_c F_b$ heißt darstellende Matrix von F bezüglich der Basen b und c.

Beweis. Da $(c_1, \ldots, c_m)$ eine Basis von W ist und $F(b_j) \in W$, sind die A_{ij}, $i = 1, \ldots, m$ jeweils für $j = 1, \ldots, n$ eindeutig bestimmt. Die Abbildung $\mathrm{Hom}(V, W) \ni F \mapsto {}_c F_b := A \in \mathbb{K}^{m \times n}$ ist also injektiv. Wegen der vorherigen Bemerkung ist sie auch surjektiv. Seiein nun $F, G \in \mathrm{Hom}(V, W)$. Dann ist ${}_c(F + G)_b = {}_c F_b + {}_c G_b$, da

$$(F + G)(b_j) = F(b_j) + G(b_j) = \sum_{i=1}^{m} c_i A_{ij} + \sum_{i=1}^{m} c_i A'_{ij} = \sum_{i=1}^{m} c_i \underbrace{(A_{ij} + A'_{ij})}_{\text{Koeff. von } {}_c(F+G)_b}$$

Dass ${}_c(\lambda F)_b = \lambda({}_c F_b)$ sieht man analog. $F \mapsto {}_c F_b$ ist also ein injektiver, surjektiver Homomorphismus, d.h., ein Isomorphismus. $\square$

Bemerkung.

(i) $\dim \operatorname{Hom}(V,W) = \dim \mathbb{K}^{m\times n} = m \cdot n$

(ii) Für $\mathrm{id} : V \to V$, $\mathrm{id}(v) = v$ ist ${}_b\mathrm{id}_b = E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$, da

$$\mathrm{id}(b_j) = 0 \cdot b_1 + \cdots + 1 \cdot b_j + \cdots + 0 \cdot b_n, \quad j = 1, \ldots n.$$

(iii) Für $F : V \to W$ linear, $v \in V$ ist $\underbrace{{}_c(F(v))}_{\in \mathbb{K}^m} = \underbrace{{}_cF_b}_{\in \mathbb{K}^{m\times n}} \underbrace{{}_bv}_{\in \mathbb{K}^n}$.

Satz 33. *Ist $F : V \to W$ linear mit $\operatorname{rg} F = r$ $(= \dim \operatorname{Im} F)$, so können die Basen b und c so gewählt werden, dass*

$$ {}_cF_b = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} \quad \textit{mit } r \textit{ Einsen ist.}$$

Beweisskizze. Sei $s = \dim \operatorname{Ker} F$. Man wählt $b_{n+1-s}, \ldots, b_n$ als Basis von $\operatorname{Ker} F$ und ergänzt zu einer Basis $(b_1, \ldots, b_n)$ von V. Nun setzt man $c_i := F(b_i)$ für $i = 1, \ldots, n-s$ und ergänzt zu einer Basis $(C_1, \ldots, c_m)$ von W. Zu zeigen bleibt, dass die $c_1, \ldots, c_{n-s}$ linear unabhängig sind: Sei $0 = \lambda_1 c_1 + \cdots + \lambda_{n-s} c_{n-s} = F(\lambda_1 b_1 + \cdots + \lambda_{n-s} b_{n-s}))$, d.h., $v := \lambda_1 b_1 + \cdots + \lambda_{n-s} b_{n-s} \in \operatorname{Ker} F$, also gibt es $\lambda_{n+1-s}, \ldots, \lambda_n$ mit $-v = \lambda_{n+1-s} b_{n+1-s} + \cdots + \lambda_n b_n$, da diese Vektoren eine Basis des Kerns bilden. Insgesamt erhält man $\lambda_1 b_1 + \cdots + \lambda_n b_n = 0$ und damit $\lambda_1 = \cdots = \lambda_n = 0$. Somit bilden die $c_1, \ldots, c_{n-s}$ eine Basis von $\operatorname{Im} F$. $\square$

Satz 34. *Seien $F : U \to V$, $G : V \to W$ linear mit Basen $b = (b_1, \ldots, b_n)$ von U, $c = (c_1, \ldots, c_m)$ von V, $d = (d_1, \ldots, d_l)$ von W. Dann ist*

$$ {}_d(G \circ F)_b = {}_dG_c \; {}_cF_b, $$

d.h., die Komposition linearer Abbildungen entspricht der Matrixmultiplikation der darstellenden Matrizen.

Beweis. Nach Definition gilt

$$G(c_j) = \sum_{i=1}^{l} d_i ({}_dG_c)_{ij}, \quad j = 1, \ldots, m$$

und

$$F(b_k) = \sum_{j=1}^{m} c_j ({}_cF_b)_{jk}, \quad k = 1, \dots, n.$$

Also ist

$$\begin{aligned} G \circ F(b_k) &= G(F(b_k)) = G\left(\sum_{j=1}^{m} c_j ({}_cG_b)_{jk}\right) = \sum_{j=1}^{m} \underbrace{\left(\sum_{i=1}^{l} d_i ({}_dG_c)_{ij}\right)}_{=G(c_j)} ({}_cF_b)_{jk} \\ &= \sum_{i=1}^{l} d_i \underbrace{\left(\sum_{j=1}^{m} ({}_dG_c)_{ij}, ({}_cF_b)_{jk}\right)}_{=\,({}_d(G\circ F)_b)_{ik}} \end{aligned}$$

und somit $({}_d(G \circ F)_b)_{ik} = ({}_dG_c\ {}_cF_b)_{ik}$. □

Korollar 4. *Ist $F : U \to V$ bijektiv, so ist $\dim U = \dim V$ $(= n)$. ${}_cF_b \in \mathbb{K}^{m\times n}$ ist dann invertierbar mit $({}_cF_b)^{-1} =_b (F^{-1})_c$.*

Beweis.
${}_b(F^{-1})_c\ {}_cF_b = {}_b(F^{-1} \circ F)_b = {}_b\mathrm{id}_b = E_n$ und ${}_cF_b\ {}_b(F^{-1})_c = {}_c(F \circ F^{-1})_c = {}_c\mathrm{id}_c = E_n$ □

Bemerkung. Die Assoziativität der Matrixmultiplikation folgt genauso aus der Assoziativität der Komposition von linearen Abbildungen.

2.3 Koordinatentransformationen

07.01.10

Sei V ein $\mathbb{K}$-Vektorraum mit Basis $b = (b_1, \dots, b_n)$. Der Isomorphismus

$$\Phi_b : \mathbb{K}^n \to V, \quad \Phi_b(\lambda_1, \dots, \lambda_n) = \lambda_1 b_1 + \dots + \lambda_n b_n$$

heißt *Parametrisierung* von V (bezüglich b). Φ_b bildet die kanonische Basis von $\mathbb{K}^n$ auf die Basis b ab,

$$\Phi_b(e_i) = b_i, \quad i = 1, \dots, n.$$

Ihr Inverses liefert den Koordinatenvektor von $v \in V$,

$$\Phi_b^{-1}(v) = {}_bv, \quad \text{bzw.,} \quad \Phi_b({}_bv) = v$$

.

Satz 35. *Ist $b' = (b'_1, \dots, b'_n)$ eine weitere Basis von V, so gilt für alle $\lambda \in \mathbb{K}^n$*

$$(\underbrace{\Phi_b^{-1}}_{\mathbb{K}^n \leftarrow V} \circ \underbrace{\Phi_{b'}}_{V \leftarrow \mathbb{K}^n})(\lambda) = \underbrace{{}_b\mathrm{id}_{b'}}_{\mathbb{K}^{n\times n}} \underbrace{\lambda}_{\mathbb{K}^n}$$

${}_b\mathrm{id}_{b'}$ heißt Transformationsmatrix des Basiswechsels von b' nach b.

Beweis. $\Phi_b^{-1} \circ \Phi_{b'}(e_j) = \Phi_b^{-1}(b'_j) = {}_b(b'_j)$, und $({}_b\mathrm{id}_{b'})e_j = ({}_b\mathrm{id}_{b'})\ {}_{b'}(b'_j) = {}_b(\mathrm{id}(b'_j)) = {}_b(b'_j)$ für $j = 1, \dots, n$. Die beiden linearen Abbildungen stimmen also auf einer Basis des $\mathbb{K}^n$ und damit auf ganz $\mathbb{K}^n$ überein. □

Bemerkung.

(i) Für $v \in V$ ist ${}_bv = ({}_b\mathrm{id}_{b'})({}_{b'}v)$.
Beweis. $({}_b\mathrm{id}_{b'})({}_{b'}v) = {}_b(\mathrm{id}(v)) = {}_bv$.

(ii) Für $f : V \to V$ linear gilt für die darstellenden Matrizen bezüglich verschiedener Basen

$$ {}_bf_b = {}_b\mathrm{id}_{b'}\ {}_{b'}f_{b'}\ {}_{b'}\mathrm{id}_b, \quad \text{wobei } {}_b\mathrm{id}_{b'} = ({}_{b'}\mathrm{id}_b)^{-1}. $$

Beweis. Zunächst ist ${}_b\mathrm{id}_{b'}\ {}_{b'}\mathrm{id}_b = {}_{b'}(\mathrm{id} \circ \mathrm{id})_{b'} = {}_{b'}\mathrm{id}_{b'} = E$.
Außerdem gilt ${}_b\mathrm{id}_{b'}\ ({}_{b'}f_{b'}\ {}_{b'}\mathrm{id}_b) = {}_b\mathrm{id}_{b'}\ {}_{b'}f_b = {}_bf_b$.

(iii) Es gilt ${}_b\mathrm{id}_{b'} = \begin{pmatrix} {}_bb'_1 & \cdots & {}_bb'_n \end{pmatrix} \in \mathbb{K}^{n \times n}$.
Beweis. ${}_bb'_i = {}_b(\mathrm{id}(b'_i)) = {}_b\mathrm{id}_{b'}\ {}_{b'}b'_i = {}_b\mathrm{id}_{b'}\ e_i = ({}_b\mathrm{id}_{b'})_{*i}$, ($i$-te Spalte), $i = 1, \dots, n$.

Anwendung: Im $\mathbb{R}^n$ mit der kanonischen Basis $e = (e_1, \dots, e_n)$ sei $b = (b_1, \dots, b_n)$ eine weitere Basis. Dann ist

$$ {}_e\mathrm{id}_b = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \quad \left(= \begin{pmatrix} {}_eb_1 & \cdots & {}_eb_n \end{pmatrix}\right). $$

Satz 36. *Sei $F : V \to W$ linear, $b = (b_1, \dots, b_n)$ Basis von V, $c = (c_1, \dots, c_n)$ Basis von W. Dann ist*

(i) $\mathrm{Im}\, F = \Phi_c(\mathrm{SR}\,({}_cF_b))$,

(ii) $\mathrm{Ker}\, F = \Phi_b(\mathrm{NR}\,({}_cF_b))$.

Beweis.

(i) $\mathrm{Im}\, F = \mathrm{span}(F(b_1), \dots, F(b_n))$,
$\mathrm{SR}\,({}_cF_b) = \mathrm{span}(({}_cF_b)_{*1}, \dots, ({}_cF_b)_{*1})$, wobei $({}_cF_b)_{*i} = ({}_cF_b)e_i = ({}_cF_b)_bb_i = {}_cF(b_i)$. Also ist

$$ \begin{aligned} \Phi_c(\mathrm{SR}\,({}_cF_b)) &= \Phi_c(\mathrm{span}({}_cF(b_1), \dots, {}_cF(b_n))) \\ &= \mathrm{span}(\Phi_c({}_cF(b_1)), \dots, \Phi_c({}_cF(b_n))) = \mathrm{Im}\, F \end{aligned} $$

(ii) "$\subset$": $v \in \mathrm{Ker}\, F \Rightarrow ({}_cF_b)\ {}_bv = {}_cF(v) = {}_c0 = 0 \in \mathbb{K}^m$, also ist ${}_bv \in \mathrm{NR}\,({}_cF_b)$.
"$\supset$": $v \in \Phi_b(\mathrm{NR}\,({}_cF_b)) \Rightarrow {}_bv \in \mathrm{NR}\,({}_cF_b) \Rightarrow \underbrace{({}_cF_b)\ {}_bv}_{={}_cF(v)} = 0 \Rightarrow F(v) = 0 \Rightarrow v \in \mathrm{Ker}\, F$.

□

Satz 37 (Dimensionsformel für lineare Abbildungen)**.** *Sei $F : V \to W$ linear, V, W endlichdimensional. Dann gilt*

$$ \dim \mathrm{Ker}\, F + \dim \mathrm{Im}\, F = \dim V $$

.

Beweis. Seien $b = (b_1, \dots, b_n)$ Basis von V, $c = (c_1, \dots, c_m)$ Basis von W. $A := {}_cF_b \in \mathbb{K}^{m\times n}$. Laut der letzten Bemerkung im Abschnitt 1.6 gilt

$$\dim \mathrm{NR}\,(A) + \dim \mathrm{SR}\,(A) = n = \dim V.$$

Da Φ_b, Φ_c Isomorphismen sind, liefert der vorhergehende Satz $\dim \mathrm{Ker}\, F = \dim \mathrm{NR}\,(A)$ und $\dim \mathrm{Im}\, F = \dim \mathrm{SR}\,(A)$. □

Definition 38. Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ heißen ähnlich, wenn es ein $S \in \mathrm{GL}_n(\mathbb{K})$ gibt mit

$$B = S^{-1}AS.$$

Bemerkung.

(i) Ähnlichkeit ist eine Äquivalenzrelation.
Beweis. [HA]

(ii) Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ sind genau dann ähnlich, wenn es eine lineare Abbildung $F : V \to V$ und zwei Basen a, b von V gibt, so dass $A = {}_aF_a$ und $B = {}_bF_b$.

Beweis.
"$\Leftarrow$": Ist $A = {}_aF_a$ und $B = {}_bF_b$, so ist $B = {}_b\mathrm{id}_a\ {}_aF_a\ {}_a\mathrm{id}_b$ mit ${}_b\mathrm{id}_a = ({}_a\mathrm{id}_b)^{-1}$.
"$\Rightarrow$": Setze $V = \mathbb{K}^n$, $F(v) := Av$ für $v \in V$, d.h. $A = {}_eF_e$ mit der kanonischen Basis $e = (e_1, \dots, e_n)$. Sei $B = S^{-1}AS$ mit $S \in \mathrm{GL}_n(\mathbb{K})$. Setze $b_i := Se_i$, dann ist $b = (b_1, \dots, b_n)$ Basis des $\mathbb{K}^n$ da S ein Isomorphismus ist. Wegen $b_i = {}_eb_i = {}_e\mathrm{id}_b\ {}_bb_i = {}_e\mathrm{id}_b\ e_i$ für $i = 1, \dots, n$ ist $S = {}_e\mathrm{id}_b$. Somit ist

$$B = S^{-1}AS = ({}_e\mathrm{id}_b)^{-1}\ {}_eF_e\ {}_e\mathrm{id}_b = {}_b\mathrm{id}_e\ {}_eF_e\ {}_e\mathrm{id}_b = {}_bF_b.$$

□ 12.01.10

(iii) Somit sind zwei Matrizen genau dann ähnlich, wenn sie die selbe lineare Abbildung in verschiedenen Basen darstellen.

Schreibweise: Ist $f : \mathbb{K}^n \to \mathbb{K}^n$, mit $f(x) = Ax$, $A \in \mathbb{K}^{n\times n}$ und b eine Basis des $\mathbb{K}^n$, so schreibt man für ${}_bf_b$ auch ${}_bA_b$. Somit ist $A = {}_eA_e$ (vergleiche: für $x \in \mathbb{K}^n$ ist $x = {}_ex$).

Schreibweise: $\mathrm{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in \mathbb{K}^{n\times n}$ heißt Diagonalmatrix.

Definition 39. $A \in \mathbb{K}^{n\times n}$ heißt diagonalisierbar, wenn sie ähnlich zu einer Diagonalmatrix ist.

Bemerkung. A diagonalisierbar heißt, es gibt eine Basis b des $\mathbb{K}^n$, so dass ${}_bA_b$ diagonal ist. D.h., ${}_bA_b = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, d.h., ${}_b(Ab_i) = \lambda_ie_i = \lambda_i\ {}_bb_i$ und somit

$$Ab_i = \lambda_ib_i,$$

man sagt b_i ist ein Eigenvektor von A zum Eigenwert λ_i.

Daumenregel:

„Fast alle Matrizen (über $\mathbb{C}$) sind diagonalisierbar.“
„Alle Matrizen sind fast diagonalisierbar.“

Ausblick: Warum diagonalisieren?
Sei $A \in \mathbb{R}^{n\times n}$. Was ist A^{100}? Ist A diagonal, so ist die Antwort einfach:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^{100} = \begin{pmatrix} \lambda_1^{100} & 0 \\ 0 & \lambda_2^{100} \end{pmatrix}$$

Aus der Analysis: $e^{ta} = \sum\limits_{n=0}^{\infty} \frac{t^n a^n}{n!}$. Es gilt $e^{ta}e^{sa} = e^{(t+s)a}$ für $s, t \in \mathbb{R}$, $a \in \mathbb{R}$.
Analog, das Matrixexponential:
Sei $A \in \mathbb{R}^{n\times n}$: $e^{tA} := \sum\limits_{n=0}^{\infty} \frac{t^n}{n!} A^n$ ist absolut konvergent für $t \in \mathbb{R}$ und $e^{tA}e^{sA} = e^{(t+s)A}$.
Berechnung des Matrixexponentials: Einfach, falls A diagonal ist:

$$\text{Für } A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ ist } e^{tA} = \begin{pmatrix} e^{t\lambda_1} & 0 \\ 0 & e^{t\lambda_2} \end{pmatrix}.$$

3 Determinanten

3.1 Beispiele und Definition

(i) Für $A \in \mathbb{R}^{2\times 2}$ ist $\det A := A_{11}A_{22} - A_{12}A_{21}$.
Ist $\det A \neq 0$, so ist $A^{-1} = \frac{1}{\det A}\begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11}\end{pmatrix}$ Für die Lösung von $Ax = b$ gilt dann

$$x = A^{-1}b = \frac{1}{\det A}\begin{pmatrix} A_{22}b_1 - A_{12}b_2 \\ -A_{21}b_1 + A_{11}b_2\end{pmatrix} = \frac{1}{\det A}\begin{pmatrix} \det\begin{pmatrix} b_1 & A_{12} \\ b_2 & A_{22}\end{pmatrix} \\ \det\begin{pmatrix} A_{11} & b_1 \\ A_{21} & b_2\end{pmatrix}\end{pmatrix}$$

Das ist die „Cramersche Regel“ für $n = 2$.

(ii) Fläche eines Parallelogramms: Sei $v, w \in \mathbb{R}^2$. Dann ist die Fläche des von v und w aufgespannten Parallelogramms

$$F = |v| \cdot |w| \cdot |\sin \sphericalangle(v, w)|.$$

Setzt man $\tilde{v} = \begin{pmatrix} v_1 \\ v_2 \\ 0\end{pmatrix}, \tilde{w} = \begin{pmatrix} w_1 \\ w_2 \\ 0\end{pmatrix}$, so ist $|\tilde{v}| = |v|$, $|\tilde{w}| = |w|$ und $\sphericalangle(\tilde{v}, \tilde{w}) = \sphericalangle(v, w)$.
Dann gilt

$$\begin{aligned} F &= |\tilde{v}| \cdot |\tilde{w}| \cdot |\sin \sphericalangle(\tilde{v}, \tilde{w}) = |v \times w| = \left|\begin{pmatrix} v_1 \\ v_2 \\ 0\end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ 0\end{pmatrix}\right| = \left|\begin{pmatrix} 0 \\ 0 \\ v_1w_2 - v_2w_1\end{pmatrix}\right| \\ &= |v_1w_2 - v_2w_1| = \left|\det\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2\end{pmatrix}\right|. \end{aligned}$$

(iii) Für $A \in \mathbb{R}^{3\times 3}$ ist

$$\begin{aligned} \det A \;:=\; & A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} \\ & -A_{13}A_{22}A_{31} - A_{11}A_{32}A_{23} - A_{12}A_{21}A_{33}. \end{aligned}$$

Für $u, v, w \in \mathbb{R}^3$ ergibt der Betrag des Spatprodukts $u \cdot (v \times w)$ das Volumen des von u, v, w aufgespannten Spates $\{\lambda_1 u + \lambda_2 v + \lambda_3 w \;:\; \lambda_1, \lambda_2, \lambda_3 \in [0, 1]\} \subset \mathbb{R}^3$. Es gilt

$$u \cdot (v \times w) = \det\begin{pmatrix} u & v & w\end{pmatrix} = \det\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3\end{pmatrix}$$

Definition 40. Sei $A \in \mathbb{K}^{n\times n}$. Bezeichne $A^{(ij)} \in \mathbb{K}^{(n-1)\times(n-1)}$ die Matrix, die entsteht, wenn man die i-te Zeile und die j-te Spalte aus A streicht. Dann heißt

$$\det A := \begin{cases} A_{11}, & \text{falls } n = 1, \\ \sum\limits_{k=1}^{n} (-1)^{k+1} A_{1k} \det A^{(1k)}, & \text{falls } n > 1, \end{cases}$$

die Determinante von A (Entwicklung nach der ersten Zeile).

Beispiele.

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

(ii) $\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} = 1\cdot\det\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} - 2\cdot\det\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} + 3\cdot\begin{pmatrix} 2 & 3 \\ 3 & 1 \end{pmatrix} = 1\cdot 5 - 2\cdot 1 + 3\cdot(-7) = -18.$

Satz 38. $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ *ist*

(D1) in jeder Spalte linear, d.h., für $a_1, \dots, a_n \in \mathbb{K}^n$, $j \in \{1, \dots, n\}$, $a_j = a_j' + \lambda a_j''$, $\lambda \in \mathbb{K}$

$$\begin{aligned} &\det\begin{pmatrix} a_1 & \cdots & (a_j' + \lambda a_j'') & \cdots & a_n \end{pmatrix} \\ &= \det\begin{pmatrix} a_1 & \cdots & a_j' & \cdots & a_n \end{pmatrix} + \lambda \det\begin{pmatrix} a_1 & \cdots & a_j'' & \cdots & a_n \end{pmatrix} \end{aligned}$$

(D2) alternierend, d.h. für $1 \le i < j \le n$ *ist*

$$\det\begin{pmatrix} a_1 & \cdots & \underset{i\text{-te}}{a_i} & \cdots & \underset{j\text{-te}}{a_j} & \cdots & a_n \end{pmatrix} = -\det\begin{pmatrix} a_1 & \cdot & \underset{i\text{-te}}{a_j} & \cdots & \underset{j\text{-te}}{a_i} & \cdots & a_n \end{pmatrix}$$

(D3) normiert, d.h., $\det E_n = 1$.

(D1–3) bestimmen die Funktion $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ *eindeutig.*

Beweis von (D1–3). Durch Induktion über n. Induktionsanfang $n = 1$ ist klar. Induktionsschritt exemplarisch von $n-1$ nach n für $n = 3$: Zu (D1):

$$\begin{aligned} \det\begin{pmatrix} \lambda a_1 & b_1 & c_1 \\ \lambda a_2 & b_2 & c_2 \\ \lambda a_3 & b_3 & c_3 \end{pmatrix} &= \lambda a_1 \cdot \det\begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} - b_1 \cdot \det\begin{pmatrix} \lambda a_2 & c_2 \\ \lambda a_3 & c_3 \end{pmatrix} + c_1 \cdot \begin{pmatrix} \lambda a_2 & b_2 \\ \lambda a_3 & b_3 \end{pmatrix} \\ &\overset{\text{I.V.}(n=2)}{=} \lambda a_1 \cdot \det\begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} - \lambda b_1 \cdot \det\begin{pmatrix} a_2 & c_2 \\ a_3 & c_3 \end{pmatrix} + \lambda c_1 \cdot \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \\ &= \lambda \det\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \end{aligned}$$

Additivität analog.
Zu (D2):

$$\begin{aligned}
\det\begin{pmatrix} b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \\ b_3 & a_3 & c_3 \end{pmatrix} &= b_1 \cdot \det\begin{pmatrix} a_2 & c_2 \\ a_3 & c_3 \end{pmatrix} - a_1 \cdot \det\begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} + c_1 \cdot \begin{pmatrix} b_2 & a_2 \\ b_3 & a_3 \end{pmatrix} \\
&\overset{\text{I.V.}(n=2)}{=} -a_1 \cdot \det\begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} + b_1 \cdot \det\begin{pmatrix} a_2 & c_2 \\ a_3 & c_3 \end{pmatrix} - c_1 \cdot \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \\
&= -\det\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}
\end{aligned}$$

Zu (D3):

$$\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \cdot \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 0 \cdot \det\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$$

Für die Eindeutigkeitsaussage werden die folgenden Lemmata benötigt, in denen nur die Eigenschaften (D1–3) der Determinante benutzt werden.

14.01.10

Lemma 3. *Sei $a_1, \dots, a_n \in \mathbb{K}^n$. Dann gilt:*

(i) Sind zwei der Vektoren gleich, so ist $\det\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} = 0$.

(ii) Sind die a_i, $i = 1, \dots, n$, linear abhängig, so ist $\det\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} = 0$.

Beweis. (i) Sei $a_i = a_j$, $i < j$. Dann ist

$$\begin{aligned}
\det\begin{pmatrix} \cdots & a_i & \cdots & a_j & \cdots \end{pmatrix} &\overset{(D2)}{=} -\det\begin{pmatrix} \cdots & a_j & \cdots & a_i & \cdots \end{pmatrix} \\
&\overset{a_i = a_j}{=} -\det\begin{pmatrix} \cdots & a_i & \cdots & a_j & \cdots \end{pmatrix}
\end{aligned}$$

Aus $x = -x$ folgt $x = 0$.

(ii) Sei ohne Einschränkung $a_n = \lambda_1 a_1 + \cdots + \lambda_{n-1} a_{n-1}$. Dann ist

$$\begin{aligned}
&\det\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \\
&\overset{(D1)}{=} \lambda_1 \det\begin{pmatrix} a_1 & \cdots & a_{n-1} & a_1 \end{pmatrix} + \cdots + \lambda_{n-1} \det\begin{pmatrix} a_1 & \cdots & a_{n-1} & a_{n-1} \end{pmatrix} = 0.
\end{aligned}$$

□

Lemma 4. *Sei $A \in \mathbb{K}^{n \times n}$. Dann ist äquivalent:*

(i) A ist invertierbar,

(ii) $\operatorname{rg}(A) = n$,

(iii) A ist ein Produkt von Elementarmatrizen (elementaren Zeilenumformungen).

Beweis. Ringschluss: $\operatorname{rg} A = n$ (ii) $\Rightarrow$ die redzuierte Zeilen-Stufen-Form von A ist E_n $\Rightarrow$ A entsteht aus E_n durch elementare Zeilenumformungen $\Rightarrow$ A ist Produkt von den Elementarmatrizen (iii)

$$P_{ij} = \begin{array}{c} \begin{pmatrix} 1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && 1 &&&&&&& \\ &&& 0 && \cdots && 1 && \\ &&&& 1 &&&&& \\ &&& \vdots && \ddots && \vdots && \\ &&&&&& 1 &&& \\ &&& 1 && \cdots && 0 && \\ &&&&&&&& \ddots & \\ &&&&&&&&& 1 \end{pmatrix} \\ \begin{matrix} \mid & \qquad\qquad & \mid \\ i\text{-te} & & j\text{-te} \end{matrix} \end{array}$$

$$S_{i,\lambda} = \begin{array}{c} \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \\ \mid \\ i\text{-te} \end{array}, \qquad T_{ij,\lambda} = \begin{array}{cl} \begin{pmatrix} 1 &&& \\ & \ddots & \lambda & \\ && \ddots & \\ &&& 1 \end{pmatrix} & -i\text{-te} \\ \begin{matrix} \mid \\ j\text{-te} \end{matrix} & \end{array}$$

Elementarmatrizen sind invertierbar $\Rightarrow$ A ist invertierbar (i) $\Rightarrow$ $Ax = b$ hat für alle $b \in \mathbb{K}^n$ die eindeutige Lösung $x = A^{-1}b \Rightarrow \operatorname{rg} A = n$ (ii). □

Lemma 5. *Sei $A \in \mathbb{K}^{n \times n}$ beliebig. Dann gilt für die Elementarmatrizen*

(a) $\det P_{ij} = -1$, $\det AP_{ij} = \det A \det P_{ij}$,

(b) $\det S_{i,\lambda} = \lambda$, $\det AS_{i,\lambda} = \det A \det S_{i,\lambda}$,

(c) $\det T_{ij,\lambda} = 1$, $\det AT_{ij,\lambda} = \det A \det T_{ij,\lambda}$.

Beweis. Sei $A = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$, $a_1, \ldots, a_n \in \mathbb{K}^n$

(a) $\det AP_{ij} = \det \begin{array}{c} \begin{pmatrix} \cdots & a_j & \cdots & a_i & \cdots \end{pmatrix} \\ \begin{matrix} \mid & \qquad & \mid \\ i\text{-te} & & j\text{-te} \end{matrix} \end{array} = -\det A$. Für $A = E_n$ folgt $\det P_{ij} = -1$,

(b) $\det AS_{i,\lambda} = \det\begin{pmatrix}\cdots & \lambda a_i & \cdots\end{pmatrix} = \lambda \det A$, für $A = E_n$ also $\det S_{i,\lambda} = \lambda$.

(c) $\det AT_{ij,\lambda} = \det\begin{pmatrix}a_1 & \cdots & \underset{\substack{| \\ i\text{-te}}}{(a_i + \lambda a_j)} & \cdots & a_n\end{pmatrix} = \det\begin{pmatrix}a_1 & \cdots & a_n\end{pmatrix} = \det A$, für

$A = E_n$ also $\det T_{ij,\lambda} = 1$

$\square$

Beweis der Eindeutigkeit der Determinantenfunktion.
Sei $\widetilde{\det} : \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Funtion die (D1-3) erfüllt. Zu zeigen ist $\widetilde{\det} = \det$.
Sei $A \in \mathbb{K}^{n\times n}$, $A = \begin{pmatrix}a_1 & \cdots & a_n\end{pmatrix}$.
1. Fall. A ist nicht invertierbar $\Rightarrow \operatorname{rg} A < n \Rightarrow$ die $a_1, \ldots, a_n$ sind linear abhängig $\Rightarrow$ $\widetilde{\det}A = 0$, $\det A = 0$.
2. Fall. A ist invertierbar $\Rightarrow$ es gibt Elemntarmatrizen $E'_1, \ldots, E'_k$ mit $A = E'_1 \cdots E'_k$. Somit ist

$$\begin{aligned}\det A &= \det E'_1 \cdot \det E'_k \\ \widetilde{\det}A &= \det E'_1 \cdot \det E'_k\end{aligned}$$

$\square$

Satz 39 (Rechenregeln)**.** *Seien $A, B \in \mathbb{K}^{n\times n}$, $\lambda \in \mathbb{K}$. Dann gilt*

(a) $\det(\lambda A) = \lambda^n \det A$,

(b) $\det A \neq 0 \Leftrightarrow A$ *ist invertierbar,*

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

(d) $\det A^{-1} = \frac{1}{\det A}$, *falls A invertierbar,*

(e) $\det A^T = \det A$.

Beweis. Sei $A = \begin{pmatrix}a_1 & \cdots & a_n\end{pmatrix}$, $a_1, \ldots, a_n \in \mathbb{K}$.

zu (a) $\det(\lambda A) = \det\begin{pmatrix}\lambda a_1 & \cdots & \lambda a_n\end{pmatrix} = \lambda^n \det\begin{pmatrix}a_1 & \cdots & a_n\end{pmatrix} = \lambda^n \det A$.

zu (b) „$\Rightarrow$“: A nicht invertierbar $\Rightarrow \operatorname{rg} A < n \Rightarrow a_1, \ldots, a_n$ sind linear abhängig $\Rightarrow$ $\det A = 0$.
„$\Leftarrow$“: A invertierbar $\Rightarrow A$ ist Produkt von Elementarmatrizen $E'_1, \ldots, E'_k \Rightarrow$ $\det A = \underbrace{\det E'_1}_{\neq 0} \cdots \underbrace{\det E'_k}_{\neq 0} \neq 0$.

zu (c) $\mathrm{GL}_n(\mathbb{K})$ ist Gruppe. Also gilt

$$AB \text{ invertierbar} \Leftrightarrow A \text{ invertierbar und } B \text{ invertierbar.}$$

1. Fall: AB nicht invertierbar, also $\det AB = 0$. Dann ist A oder B nicht invertierbar, und somit $\det a \cdot \det B = 0$.
2. Fall: A, B invertierbar, sind beide das Produkt von Elementarmatrizen, $A = E'_1 \cdots E'_k$, $B = F_1 \cdots F_l$, und somit

$$\begin{aligned}
\det(AB) &= \det(E'_1 \cdots E'_k F_1 \cdots F_l) = \det(E'_1 \cdots E'_k) \det F_1 \cdots \det F_l \\
&= \det(E'_1 \cdots E'_k) \det(F_1 \cdots F_l) = \det A \cdot \det B
\end{aligned}$$

zu (d) Ist A invertierbar, so ist $1 = \det E_n = \det(AA^{-1}) = \det A \cdot \det A^{-1}$.

zu (e) 1. Fall: $\det A = 0 \Leftrightarrow \operatorname{rg} A < n \Leftrightarrow \operatorname{rg} A^T < n \Leftrightarrow \det A^T = 0$.
2. Fall: $\det A \neq 0$, A ist Produkt von Elementarmatrizen, $A = E'_1 \cdots E'_k$. Dann ist $A^T = {E'_k}^T \cdots {E'_1}^T$. $\det A = \det A^T$ folgt nun aus $\det E' = \det E'^T$, wann immer E' eine Elementarmatrix ist.

□

Bezeichnung: $A \in \mathbb{K}^{n \times n}$ heißt *obere Dreiecksmatrix*, wenn $A_{ij} = 0$ für $i > j$.
$A \in \mathbb{K}^{n \times n}$ heißt *untere Dreiecksmatrix*, wenn $A_{ij} = 0$ für $i < j$.

Satz 40. *Ist* $\in \mathbb{K}^{n \times n}$ *eine Dreiecksmatrix, so ist* $\det A = A_{11} A_{22} \cdots A_{nn}$.

Beweis. Ist A untere Dreiecksmatrix, so erhält man die Aussage durch sukkzesive Entwicklung nach der ersten Zeile.
Ist A obere Dreiecksmatrix, so ist A^T untere Dreiecksmatrix mit den gleichen Einträgen auf der Diagonalen. □

Bemerkung (Berechnung von Determinanten).

- Entwicklung von A nach der 2. Zeile: Sei $A' = P_{12} A$ (1. und 2. Zeile vertauscht), dann ist

$$\begin{aligned}
\det A = -\det A' &= -\sum_{j=1}^{n} (-1)^{j+1} A'_{1j} \det A'^{(1j)} \\
&= \sum_{j=1}^{n} (-1)^{j+2} A_{2j} \det A^{(2j)}
\end{aligned}$$

- Entwicklung von A nach der i-ten Zeile ergibt analog:

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

- Entwicklung nach der j-ten Spalte von A entspricht der Entwicklung nach der j-ten Zeile von A^T.

- Transformation auf unnormierte Zeilen-Stufen-Form ändert nicht die Determinante. Seien $a_1, \ldots, a_n \in \mathbb{K}^n$, $i \neq j$. Dann ist

$$\begin{aligned}
\det \begin{pmatrix} a_1 \\ \vdots \\ a_j + \lambda a_i \\ \vdots \\ a_n \end{pmatrix} &= \det \begin{pmatrix} a_1^T & \vdots & a_j^T + \lambda a_i^T & \vdots & a_n^T \end{pmatrix} \\
&= \det \begin{pmatrix} a_1^T & \cdots & a_j^T & \cdots & a_n^T \end{pmatrix} + \lambda \underbrace{\det \begin{pmatrix} a_1^T & \cdots & a_i^T & \cdots & a_i^T & \cdots & a_n^T \end{pmatrix}}_{=0} \\
&= \det \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}
\end{aligned}$$

Beispiel.

$$\begin{aligned}
\det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \\ 1 & 5 & 6 \end{pmatrix} &\overset{Z2-Z1}{=} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 6 \end{pmatrix} \overset{Z3-Z1}{=} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 4 & 5 \end{pmatrix} \overset{Z3-2\cdot Z2}{=} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix} \\
&= 1 \cdot 2 \cdot (-1) = -2
\end{aligned}$$

- Vertauschen von zwei Zeilen (oder Spalten) ändert das Vorzeichen der Determinante.

3.2 Cramersche Regel

19.01.10

Definition 41. Sei $A \in \mathbb{K}^{n \times n}$. Dann heißt $\widetilde{A} \in \mathbb{K}^{n \times n}$ mit

$$\widetilde{A}_{ij} = (-1)^{i+j} \det A^{(ji)}$$

die Komplementärmatrix von A.

Satz 41. *Sei* $A \in \mathbb{K}^{n \times n}$. *Dann ist*

$$A\widetilde{A} = \det A \cdot E_n.$$

Gilt $\det A \neq 0$, *so ist* $A^{-1} = \frac{1}{\det A}\widetilde{A}$.

Beweis. Entwicklung nach der i-ten Zeile, $i = 1, \ldots, n$, ergibt

$$\begin{aligned}
\det A &= \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det A^{(ij)} \\
&= \sum_{j=1}^{n} A_{ij} \widetilde{A}_{ji} = (A\widetilde{A})_{ii}
\end{aligned}$$

Für $k \neq i$ bildet man die Matrix A' in dem man in A die k-te Zeile durch eine Kopie der i-ten Zeile von A ersetzt. Dann gilt

$$(A\widetilde{A})_{ik} = \sum_{j=1}^{n} A_{ij}\widetilde{A}_{jk} = \sum_{j=1}^{n}(-1)^{i+j}A_{ij}\det A^{(kj)} = \sum_{j=1}^{n}(-1)^{i+j}A'_{kj}\det A'^{(kj)} = \det A' = 0.$$

Insgesamt also $A\widetilde{A} = \det A \cdot E_n$. Ist nun $\det A \neq 0$ so ist $A\underbrace{\left(\tfrac{1}{\det A}\widetilde{A}\right)}_{A^{-1}} = E_n$. □

Satz 42 (Cramersche Regel)**.** *Sei $A \in \mathbb{K}^{n\times n}$ invertierbar, $b \in \mathbb{K}^n$. Dann ist die (eindeutige) Lösung von $Ax = b$ gegeben durch $x \in \mathbb{K}^n$ mit*

$$x_i = \frac{\det\begin{pmatrix}a_1 & \cdots & a_{i-1} & b & a_{i+1} & \cdots & a_n\end{pmatrix}}{\det A}, \quad \text{wobei } A = \begin{pmatrix}a_1 & \cdots & a_n\end{pmatrix}.$$

Beweis. Die Lösung ist gegeben durch $x = A^{-1}b = \frac{1}{\det A}\widetilde{A}b$, d.h.,

$$x_i = \frac{1}{\det A}\sum_{j=1}^{n}\widetilde{A}_{ij}b_j = \frac{1}{\det A}\sum_{j=1}^{n}(-1)^{i+j}b_jA^{(ji)} = \frac{1}{\det A}\sum_{j=1}^{n}(-1)^{i+j}A'_{ji}A'^{(ji)} = \frac{\det A'}{\det A}$$

wobei $A' = \begin{pmatrix}a_1 & \cdots & a_{i-1} & b & a_{i+1} & \cdots & a_n\end{pmatrix}$, für deren Determinante beim letzten Gleichheitszeichen die Entwicklung nach der i-ten Spalte benutzt wurde. □

3.3 Determinante eines Endomorphismus

Satz 43. *Sind $A, B \in \mathbb{K}^{n\times n}$ ähnlich, so gilt* $\det A = \det B$.

Beweis. Ist $B = S^{-1}AS$, so ist $\det B = \det(S^{-1}AS) = \det S^{-1}\cdot\det A\cdot\det S = \det A$ □

Definition 42. Sei $f : V \to V$ linear, $b = (b_1, \ldots, b_n)$ Basis von V. Dann ist

$$\det f := \det {}_bf_b$$

die Determinante von f.

Bemerkung.

- Die Definition ist unabhängig von der Basis von V.
  Denn wenn c eine weitere Basis von V ist, dann ist ${}_cf_c$ ähnlich zu ${}_bf_b$ und somit $\det {}_cf_c = \det {}_bf_b$.
- Für $f, g : V \to V$ linear gilt offenbar $\det(f \circ g) = \det f \cdot \det g$
- f heißt orientierungstreu, wenn $\det f > 0$.

4 Eigenwerte und Eigenvektoren

4.1 Definition und Beispiele

Definition 43. Sei $f : V \to V$ linear. $v \in V \setminus \{0\}$ heißt **Eigenvektor** von f, falls $f(v) \in \operatorname{span}(v)$.
$\lambda \in \mathbb{K}$ heißt **Eigenwert** von f, wenn es einen Eigenvektor v von f gibt mit

$$f(v) = \lambda v.$$

v ist dann ein Eigenvektor von f zum Eigenwert λ.

Bemerkung.

- Ist $v \in \operatorname{Ker} f \setminus \{0\}$, so ist v ein Eigenvektor von f zum Eigenwert 0.
- Ist $A \in \mathbb{K}^{n\times n}$, dann ist $x \in \mathbb{K}^n \setminus \{0\}$ ein Eigenvektor von A mit Eigenwert λ, falls

$$Ax = \lambda x$$

Beispiel. Diagonalmatrizen.
Für $A = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x) = Ax$ gilt

$$\begin{aligned} A\begin{pmatrix}1\\0\end{pmatrix} &= \begin{pmatrix}-1\\0\end{pmatrix} = (-1)\cdot\begin{pmatrix}1\\0\end{pmatrix}, \text{ d.h., } e_1 \text{ ist Eigenvektor von } A \text{ zum Eigenwert } -1, \\ A\begin{pmatrix}0\\1\end{pmatrix} &= \begin{pmatrix}0\\3\end{pmatrix} = 3\cdot\begin{pmatrix}0\\1\end{pmatrix}, \text{ d.h., } e_2 \text{ ist Eigenvektor von } A \text{ zum Eigenwert } 3, \end{aligned}$$

Satz 44. *Sei $f : V \to V$ linear, $b = (b_1, \ldots, b_n)$ eine Basis von V. Dann gilt*

$$b \text{ ist Basis aus Eigenvektoren von } f \Leftrightarrow {}_bf_b \text{ ist Diagonalmatrix.}$$

Beweis. Für $i = 1, \ldots, n$ gilt:

$$b_i \text{ ist Eigenvektor von } f \text{ zum Eigenwert } \lambda_i \Leftrightarrow f(b_i) = \lambda_i b_i \Leftrightarrow {}_bf(b_i) = \begin{pmatrix}0\\ \vdots \\ \lambda_i \\ \vdots \\ 0\end{pmatrix}.$$

Somit ist ${}_bf_b = \big({}_bf(b_1) \quad \cdots \quad {}_bf(b_n)\big) = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$. □

Beispiele.

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

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } 1, \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} &= \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } -1. \end{aligned}$$

Somit ist ${}_bA_b = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(ii) $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. $\operatorname{rg} A = 1$, $b_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathrm{NR}(A)$ ist Eigenvektor zum Eigenwert 0.
$b_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist Eigenvektor zum Eigenwert 2, da $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
Somit ist ${}_bA_b = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$.

(iii) $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ beschreibt eine Rotation um den Ursprung des im $\mathbb{R}^2$ mit Winkel 90° gegen den Uhrzeigersinn. Es ist $Ae_1 = e_2$, $Ae_2 = -e_1$. Betrachtet man $A \in \mathbb{C}^{2\times 2}$, so gilt für

$$\begin{aligned} b_1 := e_1 + ie_2 \quad &: \qquad Ab_1 = e_2 - ie_1 = -i(e_1 + ie_2) = -ib_1, \\ b_2 := e_1 - ie_2 \quad &: \qquad Ab_2 = e_2 + ie_1 = i(e_1 - ie_2) = ib_2. \end{aligned}$$

b_1 ist also (komplexer) Eigenvektor von A zum Eigenwert $-i$,
b_2 ist (komplexer) Eigenvektor von A zum Eigenwert i.
Die Vektoren b_1 und b_2 sind linear unabhängig und bilden daher eine Basis des $\mathbb{C}^2$.
Somit ist ${}_bA_b = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$.

4.2 Das charakteristische Polynom

Satz 45. *Sei $f : V \to V$ linear, $A \in \mathbb{K}^{n\times n}$ eine darstellende Matrix von f. Dann gilt*

(i) v ist Eigenvektor von f zum Eigenwert $\lambda \Leftrightarrow v \in \mathrm{Ker}(f - \lambda \mathrm{id}) \setminus \{0\}$,

(ii) Die folgenden Aussagen sind äquivalent:

(1) λ ist Eigenwert von f, *(3) $\det(A - \lambda E) = 0$,*

(2) $\mathrm{Ker}(f - \lambda \mathrm{id}) \neq \{0\}$, *(4) $\det(f - \lambda \mathrm{id}) = 0$.*

Beweis. 21.01.10

(i) $f(v) = \lambda v \Leftrightarrow f(v) - \lambda v = 0 \Leftrightarrow (f - \lambda \mathrm{id})(v) = 0$

(ii) λ Eigenwert von $f \Leftrightarrow \exists v \neq 0 : f(v) = \lambda v \Leftrightarrow \mathrm{Ker}(f - \lambda \mathrm{id}) \neq \{0\}$
$\Leftrightarrow f - \lambda\mathrm{id}$ ist nicht bijektiv
$\Leftrightarrow A - \lambda E$ ist nicht invertierbar
$\Leftrightarrow \det(A - \lambda E) = 0$
$\Leftrightarrow \det(f - \lambda\mathrm{id}) = 0$,
da $A - \lambda E$ darstellende Matrix von $f - \lambda\mathrm{id}$ ist.

□

Definition 44. Sei $A \in \mathbb{K}^{n\times n}$. Dann heißt

$$\chi_A : \mathbb{K} \to \mathbb{K}, \quad \chi_A(\lambda) = \det(A - \lambda E)$$

das charakteristische Polynom von A. Ist A darstellende Matrix der linearen Abbildung $f : V \to V$, so heißt

$$\chi_f : \mathbb{K} \to \mathbb{K}, \quad \chi_f(\lambda) = \det(f - \lambda\mathrm{id})$$

das charakteristische Polynom von f.

Beispiele.

- $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\chi_A(\lambda) = \det(A - \lambda E) = \det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$, also
$\chi_A(\lambda) = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc = (-\lambda)^2 + \mathrm{tr}A(-\lambda) + \det A$
- $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $\chi_A(\lambda) = \det\begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 1 \\ 1 & 0 & 0-\lambda \end{pmatrix}$, also
$\chi_A(\lambda) = (1-\lambda)\big((1-\lambda)(-\lambda)-0\big) - \big(0\cdot(-\lambda)-1\big) = -(1-\lambda)^2\lambda + 1 = \lambda^3 + 2\lambda^2 - \lambda + 1$

Satz 46. *Für $A \in \mathbb{K}^{n\times n}$ ist $\chi_A(\lambda)$ ein Polynom vom Grad n.*

Beweisskizze. Induktion über n:
$n = 1$: $A = (A_{11})$, $\chi_A(\lambda) = A_{11} - \lambda$ ist ein Polynom vom Grad 1.

$$n-1 \to n: \det\begin{pmatrix} A_{11}-\lambda & A_{12} & * & \cdots & * \\ * & A_{22}-\lambda & & & \\ \vdots & & \ddots & & * \\ \vdots & * & & \ddots & \\ * & & & & A_{nn}-\lambda \end{pmatrix} =$$

$$= (A_{11}-\lambda)\underbrace{\det\begin{pmatrix} A_{22}-\lambda & & * \\ & \ddots & \\ * & & A_{nn}-\lambda \end{pmatrix}}_{\text{Grad } n-1} - A_{12}\underbrace{\det\begin{pmatrix} * & * & \cdots & * \\ * & A_{33}-\lambda & & * \\ \vdots & & \ddots & \\ * & * & & A_{nn}-\lambda \end{pmatrix}}_{\text{Grad } n-2} \pm \cdots$$

□

Bemerkung. Für $A \in \mathbb{K}^{n \times n}$ ist

$$\chi_A(\lambda) = (-\lambda)^n + \operatorname{tr} A(-\lambda)^{n-1} + \cdots + \det A$$

Beweisskizze. Iteriert man die Formel im vorhergehenden Beweis, so erhält man

$$\begin{aligned} \chi_A(\lambda) &= (A_{11} - \lambda) \cdots (A_{nn} - \lambda) + p(\lambda) \\ &= (-\lambda)^n + (A_{11} + \cdots + A_{nn})(-\lambda)^{n-1} + \cdots \end{aligned}$$

wobei $\deg p = n - 2$.
Außerdem ist der Koeffizient von λ^0 gegeben durch $\chi_A(0) = \det(A - 0 \cdot E) = \det A$. □

Bemerkung. $\chi_A(\lambda) = 0 \overset{\text{Def}}{\iff} \det(A - \lambda E) = 0 \overset{\text{Satz}}{\iff} \lambda$ ist Eigenwert von A,
d.h. die Nullstellen von $\chi_A(\lambda)$ sind genau die Eigenwerte von A.

Lemma 6. *Seien $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ paarweise verschiedene Eigenwerte von $f : V \to V$ (d.h., $i \neq j \Rightarrow \lambda_i \neq \lambda_j$) und seien $b_1, \ldots, b_k$ zugehörige Eigenvektoren von f, also $f(b_i) = \lambda_i b_i$, $i = 1, \ldots, k$, dann sind die $b_1, \ldots, b_k$ linear unabhängig.*

Beweis. Induktion nach k: $k = 1$: $b_1 \neq 0$, also ist (b_1) linear unabhängig.
$k - 1 \to k$: Sei $\alpha_1 b_1 + \cdots + \alpha_k b_k = 0$, dann ist auch

$$0 = f(\alpha_1 b_1 + \cdots + \alpha_k b_k) = \alpha_1 f(b_1) + \cdots + \alpha_k f(b_k) = \alpha_1 \lambda_1 b_1 + \cdots + \alpha_k \lambda_k b_k$$

Zieht man hiervon das λ_k-fache der ersten Gleichung ab, erhält man

$$\alpha_1 \underbrace{(\lambda_1 - \lambda_k)}_{\neq 0} b_1 + \cdots + \alpha_{k-1} \underbrace{(\lambda_{k-1} - \lambda_k)}_{\neq 0} b_{k-1} = 0$$

Nach Induktionsvoraussetzung sind die $b_1, \ldots, b_{k-1}$ linear unabhängig, somit ist $\alpha_1 =' \cdots = \alpha_{k-1} = 0$ und damit auch α_k. □

Satz 47. *Für $A \in \mathbb{K}^{n \times n}$ gilt: Besitzt $\chi_A(\lambda)$ n verschiedene Nullstellen, so ist A ähnlich zu einer Diagonalmatrix.*

Beweis. Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ die paarweise verschiedenen Nullstellen von $\chi_A(\lambda)$. Dann gibt es zu jedem Eigenwert λ_i einen Eigenvektor b_i von A, $i = 1, \ldots, n$. Da die λ_i paarweise verschieden sind, sind die b_i linear unabhängig, bilden also eine Basis $b = (b_1, \ldots, b_n)$ des $\mathbb{K}^n$. Somit ist

$${}_bA_b = \begin{pmatrix} {}_bAb_1 & \cdots & {}_bAb_n \end{pmatrix} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

□

Bemerkung. In diesem Fall zerfällt χ_A in Linearfaktoren

$$\chi_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda),$$

jede der Nullstellen ist einfach.

Problem: Hat χ_A mehrfache Nullstellen, so kann es passieren, dass es keine Basis aus Eigenvektoren gibt. A ist dann nicht ähnlich zu einer Diagonalmatrix.

Beispiel. $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\chi_A(\lambda) = (0-\lambda)^2 = \lambda^2$ (doppelte Nullstelle bei 0).

$b_1 = e_1$ ist Eigenvektor zum Eigenwert 0, denn $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
Sei nun b_2 beliebig gewählt, so dass $b = (b_1, b_2)$ eine Basis des $\mathbb{K}^2$ ist. Dann ist $b_2 = \alpha e_1 + \beta e_2$ mit $\beta \neq 0$ und

$$Ab_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \end{pmatrix} = \beta e_1 \neg\in \operatorname{span}(b_2),$$

also kann b_2 kein Eigenvektor sein. A ist **nicht** diagonalisierbar.

Im folgenden sei V immer ein endlichdimensionaler Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Satz 48. *Sei $f : V \to V$ linear.*

χ_f zerfällt in Linearfaktoren $\iff$ f kann durch eine obere Dreiecksmatrix dargestellt werden.

Beweis. „$\Rightarrow$“: Sei die obere Dreiecksmatrix $A \in \mathbb{K}^{n\times n}$ eine darstellende Matrix von f. Dann ist

$$\chi_f(\lambda) = \chi_A(\lambda) = \det\begin{pmatrix} A_{11}-\lambda & & * \\ & \ddots & \\ 0 & & A_{nn}-\lambda \end{pmatrix} = (A_{11}-\lambda)\cdots(A_{nn}-\lambda).$$

„$\Leftarrow$“: Sei $\chi_f(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$, $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. $\lambda = \lambda_1$ ist Nullstelle von χ_f und somit ein Eigenwert von f. Es gibt also einen Eigenvektor b_1 von f zum Eigenwert λ_1. 26.01.10
Induktion über $n = \dim V$:
$n = 1$: $b = (b_1)$ ist Basis von V, ${}_bf_b = (\lambda_1) \in \mathbb{K}^{1\times 1}$ ist obere Dreiecksmatrix.
$n-1 \to n$: Man ergänzt b_1 zu einer Basis $b = (b_1, \dots, b_n)$ von V. Dann ist

$${}_bf_b = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & \boxed{B'} & \\ 0 & & & \end{pmatrix} =: B \qquad \text{mit } B' \in \mathbb{K}^{(n-1)\times(n-1)}.$$

Dann ist $\chi_f(\lambda) = \chi_B(\lambda) = (\lambda_1 - \lambda)\chi_{B'}(\lambda)$ (Entwicklung nach der ersten Spalte). Somit ist $\chi_{B'}(\lambda) = (\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$. Sei nun $b' = (b_2, \dots, b_n)$, $V' = \operatorname{span} b'$ und $\widetilde{f} : V' \to V'$ die lineare Abbildung, für die ${}_{b'}\widetilde{f}_{b'} = B'$ gilt. Da $\chi_{\widetilde{f}}$ in Linearfaktoren zerfällt, gibt es nach Induktionsvoraussetzung eine Basis $c = (c_1, \dots, c_{n-1})$ von V', so dass ${}_c\widetilde{f}_c =: C$ eine obere Dreiecksmatrix ist und es gilt

$$\begin{array}{rcll} {}_c\widetilde{f}_c & = & {}_c\mathrm{id}_{b'}\ {}_{b'}\widetilde{f}_{b'}\ {}_{b'}\mathrm{id}_c, & \text{bzw.} \\ C & = & P^{-1} \quad B' \quad P, & \text{mit } P := {}_{b'}\mathrm{id}_c. \end{array}$$

Nun ist $a := (b_1, c_1, \dots, c_{n-1})'$ eine Basis von V und es gilt

$$\begin{aligned} {}_af_a &= {}_a\mathrm{id}_b\ {}_bf_b\ {}_b\mathrm{id}_a \\ &= \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \boxed{P^{-1}} \\ 0 & \end{pmatrix} \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & \boxed{B'} \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \boxed{P} \\ 0 & \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & \boxed{P^{-1}B'P} \\ 0 & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & \boxed{C} \\ 0 & \end{pmatrix}, \end{aligned}$$

d.h. ${}_af_a$ ist obere Dreiecksmatrix. □

Bemerkung.

(i) Ist $\mathbb{K} = \mathbb{C}$, so zerfällt jedes Polynom in Linearfaktoren (Fundamentalsatz der Algebra), also kann jede lineare Abbildung $f : V \to V$ durch eine obere Dreiecksmatrix dargestellt werden.

(ii) Ist $\chi_f(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$, so gibt es eine Basis b von V mit

$${}_bf_b = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Dies folgt aus dem Beweis des vorherigen Satzes.

4.3 Polynome von Matrizen und linearen Abbildungen

Ist $p \in \mathbb{K}[x]$, also $p = a_nX^n + \cdots + a_1X + a_0$ mit $a_0, \dots, a_n \in \mathbb{K}$, so gilt

$$\begin{array}{rl} \text{für } x \in \mathbb{K}, \text{ dass} & p(x) = a_nx^n + \cdots + a_1x + a_0 \in \mathbb{K} \\ \text{für } A \in \mathbb{K}^{n\times n}, \text{ dass} & p(A) = a_nA^n + \cdots + a_1A + a_0E \in \mathbb{K}^{n\times n} \\ \text{für } f \in \mathrm{Hom}(V,V), \text{ dass} & p(f) = a_n \underbrace{f \circ \cdots \circ f}_{n \text{ mal}} + \cdots + a_1f + a_0\mathrm{id} \in \mathrm{Hom}(V,V) \end{array}$$

Fragestellungen:

(i) Welche Nullstellen hat $p(A)$ für $A \in \mathbb{K}^{n\times n}$?
Ist $p = (X-\lambda_1)(X-\lambda_2)$, so ist z.B. $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$ oder $\begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ jeweils eine Nullstelle von $p(A)$.

(ii) Gibt es zu gegebenem $A \in \mathbb{K}^{n\times n}$ ein $p \in \mathbb{K}[X]$ mit $p(A) = 0 \in \mathbb{K}^{n\times n}$?
Gesucht: $a_0, \dots, a_k \in \mathbb{K}$, $a_k \neq 0$, so dass $a_k A^k + \cdots + a_1 A + a_0 E = 0$.
$\underbrace{E, A, A^2, \dots, A^{n^2-1}, A^{n^2}}_{n^2+1} \in \mathbb{K}^{n\times n}$ sind linear abhängig, da $\dim \mathbb{K}^{n\times n} = n^2$. Es gibt also eine nichttriviale Linearkombination der $0 \in \mathbb{K}^{n\times n}$. Somit gibt es ein Polynom $p \neq 0$ mit $\deg p \leq n^2$.
Gibt es weitere Polynome p, für die $p(A) = 0 \in \mathbb{K}^{n\times n}$ gilt?

Eine Antwort gibt der

Satz 49 (Cayley-Hamilton)**.** *Sei $f : V \to V$ linear. Zerfällt χ_f in Linearfaktoren, so gilt $\chi_f(f) = 0$.*

Beweis. Sei $\chi_f(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$. Somit gibt es eine Basis $b = (b_1, \dots, b_n)$ von V, so dass

$$A := {}_b f_b = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Somit ist

$$(-1)^n \chi_A(A) = (A-\lambda_1 E)(A-\lambda_2 E)(A-\lambda_3 E)\cdots(A-\lambda_n E) =$$

$$= \begin{pmatrix} 0 & * & \dots & * \\ 0 & \lambda_2 & & * \\ \vdots & & \ddots & \\ 0 & 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_2 & * & * & \dots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \lambda_3-\lambda_2 & & * \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & \lambda_n-\lambda_2 \end{pmatrix} (A-\lambda_3 E)\cdots(A-\lambda_n E)$$

$$= \begin{pmatrix} 0 & 0 & * & \dots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & & * \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & * \end{pmatrix} (A-\lambda_3 E)\cdots(A-\lambda_n E) 0 \cdots =$$

$$= \begin{pmatrix} 0 & \cdots & 0 & * \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & * \\ 0 & \cdots & 0 & * \end{pmatrix} \begin{pmatrix} * & \cdots & * & * \\ & \ddots & & \vdots \\ 0 & & * & * \\ 0 & \cdots & 0 & 0 \end{pmatrix} = 0 \in \mathbb{K}^{n\times n}.$$

□

Bemerkung.

- $\mathbb{K} = \mathbb{C}$: Der Satz gilt für alle $A \in \mathbb{C}^{n\times n}$ wegen des *Fundamentalsatzes der Algebra.*
- $\mathbb{K} = \mathbb{R}$: Der Satz gilt für alle $A \in \mathbb{R}^{n\times n}$, da $\mathbb{R}^{n\times n} \subset \mathbb{C}^{n\times n}$ und $p_A \in \mathbb{R}[X] \subset \mathbb{C}[X]$.
- Die $n \times n$-Matrizen $E, A, \dots, A^n$ sind linear abhängig.

4.4 Die Jordan-Normalform einer linearen Abbildung

Definition 45. Sei $f : V \to V$ linear. Ein Unterraum U von V heißt **f-invariant**, wenn $f|_U : U \to U$ (d.h., $\forall v \in U : f(v) \in U$).

Beispiele.

- Ist v Eigenvektor von f zum Eigenwert λ, so ist $\mathrm{span}(v)$ f-invariant.
- Sind $v_1, \dots, v_k$ Eigenvektoren von f zum Eigenwert λ, so ist $\mathrm{span}(v_1, \dots, v_k)$ f-invariant.

Satz 50. *Ist $p \in \mathbb{K}[X]$, so ist* $\mathrm{Ker}(p(f))$ *f-invariant.*

Beweis. Sei $v \in \mathrm{Ker}(p(f))$. Z.z. ist $f(v) \in \mathrm{Ker}(p(f))$.

$$\begin{aligned} p(f)(f(v)) &= f(p(f)(v)) = 0, \quad \text{da} \\ (\underbrace{f \circ \cdots \circ f}_{k\text{-mal}})(f(v)) &= (\underbrace{f \circ \cdots \circ f}_{(k+1)\text{-mal}})(v) = f(\underbrace{f \circ \cdots \circ f}_{k\text{-mal}}(v)) \end{aligned}$$

□

Bemerkung. 28.01.10

- Sind U_1, U_2 Unterräume von V mit $U_1 \cap U_2 = \{0\}$, so heißt der Unterraum $U = \mathrm{span}\,(U_1, U_2) = \{u_1 + u_2 : u_1 \in U_1, u_2 \in U_2\}$ direkte Summe von U_1 und U_2. Man schreibt dann
$$U = U_1 \oplus U_2.$$
- Ist $b = (b_1, \dots, b_k)$ Basis von U_1, $c = (c_1, \dots, c_l)$ Basis von U_2, so ist $a = (b_1, \dots, b_k, c_1, \dots, c_l)$ Basis von $U_1 \oplus U_2$.
Beweis: $\underbrace{\alpha_1 b_1 + \cdots + \alpha_k b_k}_{\in U_1} + \underbrace{\beta_1 c_1 + \cdots + \beta_l c_l}_{\in U_2} = 0 \in U_1 \cap U_2$. Also ist $\alpha_1 b_1 + \cdots + \alpha_k b_k$ auch in U_2 enthalten und damit gleich dem Nullvektor. Da b als Basis von U_1 linear unabhängig ist, folgt $\alpha_1 = \cdots = \alpha_k = 0$. Analog folgt, dass $\beta_1 c_1 + \cdots + \beta_l c_l = 0$, da auch in U_2 enthalten und damit $\beta_1 = \cdots = \beta_l = 0$. □

- Sind U_1, U_2 invariante Unterräume von $f : V \to V$ und $V = U_1 \oplus U_2$, so ist

$$ {}_af_a = \begin{pmatrix} \boxed{*} & 0 \\ 0 & \boxed{*} \end{pmatrix} \begin{matrix} \}k\text{-Zeilen} \\ \}l\text{-Zeilen} \end{matrix} $$

mit zwei von Null verschiedenen Blöcken der Dimension $k \times k$, bzw. $l \times l$.

Bezeichnungen: Sei λ Eigenwert von $f : V \to V$. Dann heißt

(i) $E_\lambda(f) := \operatorname{Ker}(f - \lambda \mathrm{id})$ der **Eigenraum** von λ,

(ii) $\dim E_\lambda(f)$ die **geometrische Vielfachheit** von λ,

(iii) $\nu_\lambda \in \mathbb{N}$, so dass $\chi_f(\mu) = (\mu - \lambda)^{\nu_\lambda} q(\mu)$ mit $q(\lambda) \neq 0$, die **algebraische Vielfachheit** von λ (die Vielfachheit der Nullstelle von χ_f),

(iv) $H_\lambda(f) := \operatorname{Ker}(f - \lambda \mathrm{id})^{\nu_\lambda}$ der **Hauptraum** von λ.

Bemerkung.

(i) $E_\lambda(f)$, $H_\lambda(f)$ sind f-invariante Unterräume von V,

(ii) $E_\lambda(f) \subset H_\lambda(f)$,

(iii) die geometrische Vielfachheit von λ ist immer kleiner gleich der algebraischen Vielfachheit.

Beweis.

(i) klar, da $(f - \lambda \mathrm{id})^j$, $j \in \mathbb{N}$, ein Polynom von f ist.

(ii) $v \in E_\lambda(f) \Rightarrow (f - \lambda \mathrm{id})(v) = 0 \Rightarrow (f - \lambda \mathrm{id})^{\nu_\lambda}(v) = 0 \Rightarrow v \in H_\lambda(f)$.

(iii) folgt aus (ii).

□

Satz 51. *Sei λ ein Eigenwert von $f : V \to V$ mit algebraischer Vielfachheit ν. Dann gilt*

1. $\dim H_\lambda(f) = \nu$,
2. *Ist $\mu \neq \lambda$ ein weiterer Eigenwert von f, so ist*

$$H_\lambda(f) \cap H_\mu(f) = \{0\}$$

.

Der Beweis wird hier nicht geführt. Der Beweis verwendet den Satz von Cayley-Hamilton und Teilerfremdheit von Polynomen, siehe z.B. [Fischer, Lineare Algebra] □

Korollar 5. *Sei $f : V \to V$ linear mit $\chi_f(\lambda) = (\lambda_1 - \lambda)^{\nu_1} \cdots (\lambda_k - \lambda)^{\nu_k}$, $\lambda_1, \ldots, \lambda_k$ paarweise verschieden und $\nu_1 + \cdots + \nu_k = n = \dim V$, so gilt*

(i) $V = H_{\lambda_1}(f) \oplus \cdots \oplus H_{\lambda_k}(f)$.

(ii) Es gibt eine Basis von V, $b = (\underbrace{b_1, \cdots, b_{\nu_1}}_{\text{Basis von } H_{\lambda_1}}, \ldots, \underbrace{b_{n-\nu_k+1}, \cdots, b_n}_{\text{Basis von } H_{\lambda_k}})$, so dass

$$
{}_b f_b = \begin{pmatrix} \boxed{*} & & & 0 \\ & \boxed{*} & & \\ & & \ddots & \\ 0 & & & \boxed{*} \end{pmatrix}
\begin{array}{l} \}\nu_1\text{-Zeilen} \\ \}\nu_2\text{-Zeilen} \\ \quad\vdots \\ \}\nu_k\text{-Zeilen} \end{array}
$$

Beweis. (i) „$\supset$“: klar. „$\subset$“: $\dim(H_{\lambda_1}(f) \oplus \cdots \oplus H_{\lambda_k}(f)) = \nu_1 + \cdots + \nu_k = \dim V$.

(ii) Man wähle zu jedem $H_{\lambda_i}(f)$ eine Basis $(c_1^{(i)}, \ldots, c_{\nu_i}^{(i)})$ und füge sie durch Aneinanderreihung zu einer Basis von V zusammen.

□

Definition 46.

- Die Matrix $J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$ heißt **Jordanblock**.

- Eine Matrix $J \in \mathbb{K}^{n \times n}$ heißt **Jordanmatrix**, wenn es Jordanblöcke $J_{m_1}(\lambda_1), \ldots, J_{m_l}(\lambda_l)$ gibt, so dass

$$
J = \begin{pmatrix} \boxed{J_{m_1}(\lambda_1)} & & & 0 \\ & \boxed{} & & \\ & & \ddots & \\ 0 & & & \boxed{J_{m_l}(\lambda_l)} \end{pmatrix}.
$$

Satz 52. *Jede Matrix $A \in \mathbb{C}^{n\times n}$ ist ähnlich zu einer Jordanmatrix.*

Beweisidee. Für jeden Hauptraum $H_\lambda(A)$ bestimmt man eine Basis wie folgt:
Sei $(b_1, \dots, b_i)$ eine Basis von $E_\lambda(A) = \mathrm{NR}\,(A - \lambda E)$. Man ergänze, falls nötig, zu $(b_1, \dots, b_i, c_1 \dots, c_j)$, einer Basis von $\mathrm{NR}\,(A - \lambda E)^2$.

$\vdots$

Man ergänze, falls nötig, zu $(b_1, \dots, b_i, \cdots, d_1 \dots, d_k)$, einer Basis von $\mathrm{NR}\,(A - \lambda E)^m = H_\lambda(A)$, wobei $m \le \nu_\lambda$. Nun setzt man $d_s^{(l)} := (A - \lambda E)^l d_s$ für $l = 0, \dots, m-1$, $s = 1, \dots, k$. Dann kann man zeigen, dass die $m \cdot k$ Vektoren $d_s^{(0)}, \dots, d_s^{(m-1)}$, $s = 1, \dots, k$, linear unabhängig sind. Bilden diese Vektoren noch keine Basis des Hauptraums $H_\lambda(A)$, so ergänzt man sie durch $(b_1', \dots, b_r')$ zu einer solchen und wiederholt den gesamten Schritt für $H_\lambda'(A) := \mathrm{span}\,(b_1', \dots, b_r')$ an Stelle von $H_\lambda(A)$. Dies wird solange iteriert, bis man eine Basis des Hauptraums $H_\lambda(A)$, bestehend aus Vektorhierarchien $d_s^{(0)}, \dots d_s^{(m_s-1)}$, erhalten hat. Jede dieser Vektorhierarchien bildet nun einen Jordanblock der Dimension $m = m_s$, denn es gilt

$$Ad_s^{(l)} = (A - \lambda E)d_s^{(l)} + \lambda d_s^{(l)} = d_s^{(l+1)} + \lambda d_s^{(l)}.$$

Die darstellende Matrix bezüglich dieser Vektoren lautet also, da $d_s^{(m)} = 0$ ist

$$\begin{array}{c} \begin{pmatrix} \ddots & & & & & \\ & \lambda & 1 & & 0 & \\ & 0 & \lambda & \ddots & & \\ & \vdots & 0 & \ddots & 1 & \\ & \vdots & \vdots & & \lambda & \\ & & & & & \ddots \end{pmatrix} \\ \begin{array}{ccc} \nearrow & \nwarrow & \uparrow \\ d_s^{(m-1)} & d_s^{(m-2)} & d_s^{(0)} \end{array} \end{array}$$

$\square$

5 Euklidische Vektorräume

02.02.10

5.1 Definitionen und Beispiele

Definition 47. Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $s : V \times V \to \mathbb{R}$ heißt **Bilinearform**, wenn für alle $u, v \in V$, $\lambda \in \mathbb{R}$ gilt:

(B1) $s(u + w, v) = s(u, v) + s(w, v)$, $s(\lambda u, v) = \lambda s(u, v)$,

(B2) $s(u, v + w) = s(u, v) + s(v, w)$, $s(u, \lambda v) = \lambda s(u, v)$.

Eine Bilinearform s heißt **symmetrisch**, wenn für alle $u, v \in V$

(S) $s(u, v) = s(v, u)$.

Eine symmetrische Bilinearform s heißt **positiv definit**, wenn

(P) $s(v, v) > 0$ für alle $v \in V \setminus \{0\}$.

Eine positiv definite symmetrische Bilinearform s heißt **Skalarprodukt** auf V, man schreibt $\langle u, v\rangle := s(u, v)$.

Beispiele.

- $V = \mathbb{R}^n$, $s(v, w) = v \cdot w (= v^T w)$ für $v, w \in \mathbb{R}^n$ ist ein Skalarprodukt auf dem $\mathbb{R}^n$.
- $V = \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $s_A(v, w) := v \cdot Aw (= v^T Aw)$ ist eine Bilinearform auf dem $\mathbb{R}^n$.
- Ist $A^T = A$, so ist s_A symmetrisch.
  Beweis. $s_A(v, w) = v^T Aw = v \cdot Aw = Aw \cdot v = (Aw)^T v = w^T A^T v = s_A(w, v)$. □
- Für welche symmetrischen $A \in \mathbb{R}^{n\times n}$ ist s_A positiv definit?

Definition 48. Ein reeller Vektorraum V zusammen mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$ heißt **euklidischer Vektorraum**.

Definition 49. In einem euklidischen Vektorraum V heißt für $v \in V$

$$\|v\| := \sqrt{\langle v, v\rangle}$$

die (durch $\langle\cdot,\cdot\rangle$ induzierte) **Norm** von V.

Satz 53. *In einem euklidischen Vektorraum V gilt für $v, w \in V$, $\lambda \in \mathbb{R}$*

(i) $\|v\| = 0 \iff v = 0$,

(ii) $\|\lambda v\| = |\lambda| \cdot \|v\|$,

(iii) $|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$ *(Cauchy-Schwarz-Ungleichung),*

(iv) $\|v + w\| \leq \|v\| + \|w\|$ *(Dreiecks-Ungleichung)*

Beweis.

(i) $\|v\| \neq 0 \Leftrightarrow \langle v, v\rangle \neq 0 \Leftrightarrow v \neq 0$.

(ii) $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda^2 \langle v, v\rangle} = |\lambda| \cdot \|v\|$.

(iii) Beweis wie in Abschnitt 0.6: Man berechne den positiven Wert des Minimums von $\|v + \alpha w\|^2$ für $\alpha \in \mathbb{R}$.

(iv) $$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w\rangle = \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \\ &\overset{\text{(iii)}}{\leq} \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

□

Bezeichnung.

- $\|v\|$ heißt die **Länge** des Vektors $v \in V$,
- $\sphericalangle(v, w) := \arccos \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \in [0, \pi[$ heißt der Winkel zwischen v und w.

Definition 50. Sei V ein euklidischer Vektorraum.

(a) $v, w \in V$ heißen **orthogonal** ($v \perp w$), wenn $\langle v, w\rangle = 0$,

(b) Zwei Unterräume von V, U_1, U_2, heißen **orthogonal** ($U_1 \perp U_2$), wenn $\forall v \in U_1, w \in U_2 : v \perp w$,

(c) Zum Unterraum U von V definiert man das **orthogonale Komplement**

$$U^\perp := \{v \in V \,|\, \forall u \in U : u \perp v\}$$

(d) $(v_1, \ldots, v_n)$ heißt **Orthogonalsystem**, wenn $v_i \neq 0$, $v_i \perp v_j$, für $i, j = 1, \ldots, n$. Ein Orthogonalsystem mit $\|v_i\| = 1$ heißt **Orthonormalsystem (ONS)**. Bilden die v_i sogar eine Basis von V, so heißt $(v_1, \ldots, v_n)$ **Orthonormalbasis (ONB)** von V.

Bemerkung.

- Für ein Orthonormalsystem $(v_1, \dots, v_k)$ gilt
$$\langle v_i, v_j \rangle = \delta_{ij}$$
mit dem *Kronecker-Symbol* $\delta_{ij} := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$

- Ist $(v_1, \cdots, v_k)$ ein Orthonormalsystem, so sind die v_i linear unabhängig.
Beweis. $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0 \Rightarrow 0 = \langle v_1, 0 \rangle = \alpha_1 \langle v_1, v_1 \rangle + \cdots + \alpha_k \langle v_1, v_k \rangle = \alpha_1$.
$\alpha_i = 0$ folgt analog aus $0 = \langle v_i, 0 \rangle$. $\square$

- Ist $(b_1, \dots, b_n)$ eine Orthonormalbasis von V, so gilt für $v \in V$
$$v = \lambda_1 b_1 + \cdots + \lambda_n b_n \quad \text{mit } \lambda_i = \langle b_i, v \rangle.$$
Mit anderen Worten: Der Koordinatenvektor von v ist ${}_b v = \begin{pmatrix} \langle b_1, v \rangle \\ \vdots \\ \langle b_n, v \rangle \end{pmatrix}$. *Beweis.*
Sei $v = \lambda_1 b_1 + \cdots + \lambda_n v_n$. Dann ist
$\langle b_i, v \rangle = \lambda_1 \langle b_i, b_1 \rangle + \cdots + \lambda_n \langle b_i, b_n \rangle = \lambda_i \langle b_i, b_i \rangle = \lambda_i$. $\square$

Satz 54 (Gram-Schmidt-Orthogonalisierung)**.** *Jeder endlichdimensionale euklidische Vektorraum besitzt eine Orthonormalbasis.*

Beweis. Sei $(v_1, \dots, v_n)$ eine beliebige Basis des euklidischen Vektorraums V. Man setzt

$$\begin{aligned}
\widetilde{b}_1 &:= v_1, \quad b_1 := \frac{\widetilde{b}_1}{\|\widetilde{b}_1\|}, \\
\widetilde{b}_2 &:= v_2 - \langle b_1, v_2 \rangle b_1, \quad b_2 := \frac{\widetilde{b}_2}{\|\widetilde{b}_2\|}, \\
&\vdots \\
\widetilde{b}_k &:= v_k - \sum_{j=1}^{k-1} \langle b_j, v_k \rangle b_j, \quad b_k := \frac{\widetilde{b}_k}{\|\widetilde{b}_k\|}
\end{aligned}$$

für $k = 1, \dots, n$. Zu überprüfen ist, dass $(b_1, \dots, b_n)$ wohldefiniert und eine ONB ist.
Zunächst ist $\widetilde{b}_1 \neq 0$, b_1 ist also wohldefiniert und offenbar ist $\operatorname{span}(b_1) = \operatorname{span}(v_1)$.
Sei nun $b_1, \dots, b_{k-1}$ definiert, mit $\operatorname{span}(b_1, \dots, b_{k-1}) = \operatorname{span}(v_1, \dots, v_{k-1})$.
Dann ist $v_k \notin \operatorname{span}(b_1, \dots, b_{k-1})$ und damit auch $\widetilde{b}_k \notin \operatorname{span}(b_1, \dots, b_{k-1})$, also $\widetilde{b}_k \neq 0$. b_k ist also wohldefiniert. Außerdem gilt für $i = 1, \dots, k-1$

$$\langle b_i, \widetilde{b}_k \rangle = \langle b_i, v_k \rangle - \sum_{j=1}^{k-1} \langle b_j, v_k \rangle \langle b_i, b_j \rangle = \langle b_i, v_k \rangle - \langle b_i, v_k \rangle \langle b_i, b_i \rangle = 0$$

b_k steht also senkrecht auf $b_1, \dots, b_{k-1}$ und wegen $b_k \in \operatorname{span}(b_1, \dots, b_{k-1}, v_k)$ ist $\operatorname{span}(b_1, \dots, b_k) = \operatorname{span}(v_1, \dots, v_k)$. Die so definierten $b_1, \dots, b_n$ bilden also eine Orthonormalbasis von V. □

Beispiele.

- Bestimmung von Orthonormalbasen von Untervektorräumen des $\mathbb{R}^n$ [HA].
- **Basiswechsel.** $(e_1, \dots, e_n)$ ist eine Orthonormalbasis des $\mathbb{R}^n$ mit dem Standardskalarprodukt $\langle x, y\rangle = x \cdot y = x^T y$. 04.02.10
  Ist $b = (b_1, \dots, b_n)$ eine weitere ONB des $\mathbb{R}^n$, so ist
  $$B := {}_e\mathrm{id}_b = \begin{pmatrix} {}_eb_1 & \cdots & {}_eb_n \end{pmatrix} = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \in \mathbb{R}^{n\times n}$$
  Die Inverse $B^{-1} = {}_b\mathrm{id}_e$ ist in diesem Fall ganz einfach zu bestimmen.
  Es ist $B^{-1} = B^T$, denn
  $$B^T B = \begin{pmatrix} {b_1}^T \\ \cdots \\ {b_n}^T \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} {b_1}^T b_1 & \vdots & {b_1}^T b_n \\ \vdots & & \vdots \\ {b_n}^T b_1 & \cdots & {b_n}^T b_n \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = E_n$$
- Allgemeiner gilt: Ist $B = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ wobei $(b_1, \dots, b_n)$ eine Basis paarweise orthogonaler (unnormierter) Vektoren ist, so ist
  $$B^{-1} = \begin{pmatrix} \frac{1}{\|b_1\|^2} {b_1}^T \\ \vdots \\ \frac{1}{\|b_n\|^2} {b_n}^T \end{pmatrix}$$

Bemerkung. Sei V euklidischer Vektorraum mit Orthonormalbasis $b = (b_1, \dots, b_n)$ und $v, w \in V$. Dann ist
$$\langle v, w\rangle = {}_bv \cdot {}_bw = {}_bv^T\, {}_bw$$

Beweis. $\langle v, w\rangle = \sum_{i,j=1}^{n} \langle \langle v, b_i\rangle b_i, \langle w, b_j\rangle b_j = \sum_{i,j=1}^{n} \langle v, b_i\rangle \langle w, b_j\rangle \underbrace{\langle b_i, b_j\rangle}_{=\delta_{ij}}$
$$= \sum_{i=1}^{n} \langle v, b_i\rangle\langle w, b_i\rangle = \begin{pmatrix} \langle v, b_1\rangle \\ \vdots \\ \langle v, b_2\rangle \end{pmatrix} \cdot \begin{pmatrix} \langle w, b_1\rangle \\ \vdots \\ \langle w, b_2\rangle \end{pmatrix} \cdot = {}_bv \cdot {}_bw.$$ □

5.2 Orthogonale Transformationen

Definition 51. Sei V ein euklidischer Vektorraum, $\dim V < \infty$ und $F : V \to V$ linear. Dann heißt F **orthogonal**, wenn
$$\langle F(v), F(w)\rangle = \langle v, w\rangle \quad \text{für alle } v, w \in V.$$

Bemerkung.

(i) $\|F(v)\| = \|v\|$ für alle $v, w \in V$.

(ii) $v \perp w \Rightarrow F(v) \perp F(w)$.

(iii) F ist Isomorphismus und F^{-1} ist orthogonal.

(iv) Ist λ ein Eigenwert von F, so ist $|\lambda| = 1$.

Beweis. (i) und (ii) klar

(iii) F ist injektiv, da aus $F(v) \neq 0$ mit (i) auch $v \neq 0$ folgt. Somit ist F sogar bijektiv und $\|v\| = \|F(F^{-1}(v))\| = \|F^{-1}(v)\|$ für alle $v \in V$.

(iv) Ist v ein Eigenvektor von F zum Eigenwert λ, dann gilt
$\|v\| = \|F(v)\| = \|\lambda v\| = |\lambda|\|v\|$, also $|\lambda| = 1$.

□

Bemerkung. Eine orthogonale Abbildung ist Winkel erhaltend.

Definition 52. Eine Matrix $A \in \mathrm{GL}_n(\mathbb{R})$ heißt **orthogonal**, wenn $A^{-1} = A^T$.

Bemerkung. Ist A eine orthogonale Matrix, so ist $\det A = \pm 1$, denn

$$1 = \det E = \det A^T A = \det A^T \cdot \det A = (\det A)^2.$$

Bemerkung. Die Mengen

$$\begin{aligned} \mathrm{O}(n) &:= \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A^{-1} = A^T\}, && \text{(orthogonale Gruppe)} \\ \mathrm{SO}(n) &:= \{A \in \mathrm{O}(n) \mid \det A = 1\} && \text{(spezielle orthogonale Gruppe)} \end{aligned}$$

sind Untergruppen der $\mathrm{GL}_n(\mathbb{R})$.

Beweis. $A, B \in \mathrm{O}(n) \Rightarrow (AB)^T(AB) = B^T A^T AB = B^T B = E_n \Rightarrow AB \in \mathrm{O}(n)$,
$A^{-1} = A^T \mathrm{O}(n)$, da $(A^T)^T A^T = AA^T = AA^{-1} = E_n$, d.h., $\mathrm{O}(n)$ ist Untergruppe.

$A, B \in \mathrm{SO}(n) \Rightarrow \det(AB) = \det A \cdot \det B = 1$, also auch $AB \in \mathrm{SO}(n)$.
Außerdem ist $\det A^{-1} = \det A^T = \det A = 1$, also $A^T \in \mathrm{SO}(n)$. □

Bemerkung.
$A \in \mathbb{R}^{n \times n}$ orthogonal $\Longleftrightarrow$ die Spalten von A bilden eine Orthonormalbasis des $\mathbb{R}^n$.

Beweis. Sei $A = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$. Dann ist $A^T = \begin{pmatrix} b_1^T \\ \vdots \\ b_n^T \end{pmatrix}$ und somit

$$A^T A = E \quad \Leftrightarrow \quad \begin{pmatrix} b_1^T \\ \vdots \\ b_n^T \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \quad \Leftrightarrow \quad b_i^T b_j = \delta_{ij},$$

Dies bedeutet $\langle b_i, b_j \rangle = \delta_{ij}$, bzw., b ist ONB. □

Satz 55. *Sei V ein euklidischer Vektorraum mit Orthonormalbasis $b = (b_1, \dots, b_n)$ und sei $F : V \to V$ linear. Dann gilt*

$$F \text{ ist orthogonal} \quad \Longleftrightarrow \quad {}_bF_b \text{ ist orthogonal.}$$

Beweis.

$$\begin{aligned}
& \langle F(v), F(w)\rangle = \langle v, w\rangle \\
\Leftrightarrow\ & \forall i, j : \langle F(b_i), F(b_j)\rangle = \langle b_i, b_j\rangle = \delta_{ij} \\
\Leftrightarrow\ & (F(b_1), \dots, F(b_n)) \text{ ist ONB von } V \\
\Leftrightarrow\ & ({}_bF(b_1), \dots, {}_bF(b_n)) \text{ ist ONB von } \mathbb{R}^n \\
\Leftrightarrow\ & {}_bF_b = \begin{pmatrix} {}_bF(b_1) & \cdots & {}_bF(b_n) \end{pmatrix} \in \mathbb{R}^n \text{ ist Orthogonalmatrix.}
\end{aligned}$$

□

Beispiele. Orthogonale Transformationen des $\mathbb{R}^n$

- $n = 1$. $F : \mathbb{R} \to \mathbb{R}$ orthogonal $\Rightarrow |F(x)| = |x| \Rightarrow F(x) = \pm x$,
  d.h. $F = \mathrm{id}_{\mathbb{R}}$ oder $F = -\mathrm{id}_{\mathbb{R}}$ (Spiegelung am Ursprung)

- $n = 2$. Betrachte orthogonale 2×2 Matrizen.

  Lemma 7. *Ist $A \in \mathrm{O}(2)$, so gibt es $\alpha \in \,]-\pi, \pi]$, so dass*

  $$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \textit{oder} \quad \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \quad \textit{ist.}$$

  Beweis. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{O}(2)$. Aus $A^T A = E_2$ folgt

  $$a^2 + c^2 = 1, \quad b^2 + d^2 = 1, \quad ab + cd = 0.$$

  Mit $\alpha := \arg(a + ic) \in]-\pi, \pi]$, $\beta := \arg(d + ib) \in]-\pi, \pi]$ gilt

  $$a = \cos\alpha, \quad c = \sin\alpha, \quad d = \cos\beta, \quad b = \sin\beta$$

  Die letzte Gleichung lautet dann

  $$0 = \cos\alpha \sin\beta + \sin\alpha\cos\beta = \sin(\alpha + \beta).$$

  1. Fall: $\alpha + \beta \in \{0, 2\pi\}$. Dann ist $\sin\beta = -\sin\alpha$, $\cos\beta = \cos\alpha$.
  Somit $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.
  2. Fall: $\alpha + \beta = \pi$. Dann ist $\sin\beta = \sin\alpha$, $\cos\beta = -\cos\alpha$.
  Somit $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$. □

Diagonalisierung von $A \in \mathrm{O}(2)$?
1. Fall: $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. $\chi_A(\lambda) = (\cos\alpha - \lambda)^2 + \sin\alpha^2 = \lambda^2 - 2\lambda\cos\alpha + 1$.
Die Nullstellen $\lambda_{1,2} = \cos\alpha \pm i\sin\alpha = e^{\pm i\alpha}$ sind die komplexen Eigenwerte von A. Keine reellen Eigenvektoren.
A ist *Drehung* um Winkel α, $\det A = 1$, d.h., $A \in \mathrm{SO}(2)$.

2. Fall $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$. $\chi_A(\lambda) = (\cos\alpha - \lambda)(-\cos\alpha - \lambda) - \sin^2\alpha = \lambda^2 - 1 = (\lambda-1)(\lambda+1)$. Es gibt also eine Basis aus Eigenvektoren $b_\pm$ zu den Eigenwerten ± 1, die senkrecht aufeinander stehen ($b_+ \cdot b_- = Ab_+ \cdot Ab_- = -b_+ \cdot b_-$).
A ist eine *Spiegelung* an der von b_+ aufgespannten Achse, $\det A = -1$.

- $n = 3$. $\boxed{09.02.10}$

Beispiel (Rotation im $\mathbb{R}^3$ um eine Achse durch den Ursprung)**.**
Sei $a \in \mathbb{R}^3$, $\|a\| = 1$, $\alpha \in \mathbb{R}$. Zu $x \in \mathbb{R}^3$ bestimme man den Vektor x', der entsteht, wenn man x um die durch a bestimmte Achse mit dem Winkel α dreht.
Man definiert $x_\| := (a \cdot x)a$, $x_\perp := x - x_\|$.
Offenbar ist $x_\| \perp x_\perp$, da $a \cdot x_\perp = a \cdot x - (a \cdot x)(a \cdot a) = 0$. Zusammen mit dem Vektor $x_\perp \times a$, der senkrecht auf $x_\|$ und $x_\perp$ und der die gleiche Länge wie $x_\perp$ hat, ergibt sich anschaulich

$$x \mapsto x' = x_\| + \cos\alpha\, x_\perp + \sin\alpha\,(x_\perp \times a).$$

Die Rotation wird also durch die Abbildung $F : \mathbb{R}^3 \to \mathbb{R}^3$,

$$F(x) = (a \cdot x)a + \cos\alpha(x - (a \cdot x)a) + \sin\alpha(x \times a),$$

modelliert. F ist linear und orthogonal, wie man überprüfen kann. a ist Eigenvektor zum Eigenwert 1, da offenbar $F(a) = a$ ist. Wegen $(a \cdot x)a = aa^T x$ ist

$$F(x) = \underbrace{\Big((1 - \cos\alpha)aa^T + \cos\alpha\, E_3 + \sin\alpha\, A\Big)}_{\in\mathbb{R}^{3\times 3}} x$$

mit $A = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}$, so dass $Ax = x \times a$.

Klassifikation orthogonaler Abbildungen auf dem $\mathbb{R}^3$.
Sei $F : \mathbb{R}^3 \to \mathbb{R}^3$ orthogonal, $A = {}_eF_e$. $\chi_A(\lambda)$ ist ein reelles Polynom vom Grad 3 und hat daher eine reelle Nullstelle (Begründung: z.B. Zwischenwertsatz). Da F orthogonal ist besitzt sie also einen Eigenwert $\lambda_1 = \pm 1$. Sei b_1 ein zugehöriger Eigenvektor mit $\|b_1\| = 1$. Man ergänzt nun zu einer Orthonormalbasis (b_1, b_2, b_3). Dann ist $U = \mathrm{span}\,(b_2, b_3)$ F-invariant, denn

$$v \in U \;\Rightarrow\; \langle v, b_1\rangle = 0 \;\Rightarrow\; \langle F(v), F(b_1)\rangle = 0 \;\Rightarrow\; \pm\langle F(v), b_1\rangle = 0 \;\Rightarrow\; F(v) \in U$$

Insbesondere ist $F(b_2), F(b_3) \in U$. Somit ist

$$B = {}_bF_b = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \multicolumn{2}{c}{\multirow{2}{*}{\boxed{A'}}} \\ 0 & & \end{pmatrix}, \quad \text{mit} \quad A' = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Wegen $B^T B = E_3$ ist $A'^T A' = E_2$; also $A \in \mathrm{O}(2)$. Bei geeigneter Wahl von b_2, b_3 erhält man die vier Fälle

$${}_bF_b = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix} \qquad \text{Spiegelung an der } b_1\text{-}b_2\text{-Ebene}$$

$${}_bF_b = \begin{pmatrix} 1 & & \\ & \cos\alpha & -\sin\alpha \\ & \sin\alpha & \cos\alpha \end{pmatrix} \qquad \text{Drehung um die } b_1\text{-Achse mit Winkel } \alpha$$

$${}_bF_b = \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix} \qquad \begin{array}{l} \text{Spiegelung an der } b_2\text{-Achse, bzw.} \\ \text{Drehung um die } b_2\text{-Achse mit Winkel } \pi \end{array}$$

$${}_bF_b = \begin{pmatrix} -1 & & \\ & \cos\alpha & -\sin\alpha \\ & \sin\alpha & \cos\alpha \end{pmatrix} \qquad \begin{array}{l} \text{Drehspiegelung: Drehung um die } b_1\text{-Achse,} \\ \text{Spiegelung an der } b_1\text{-}b_2\text{-Ebene} \end{array}$$

Für beliebige Dimension gilt der

Satz 56. *Ist $F : V \to V$ orthogonal, dann gibt es eine Orthonormalbasis b von V, so dass*

$${}_bF_b = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{A_1} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \boxed{A_k} \end{pmatrix} \qquad \textit{mit } A_1, \ldots, A_k \in \mathrm{SO}(2).$$

Beweis. siehe z.B. [Fischer, Lineare Algebra]

5.3 Symmetrische lineare Abbildungen

Definition 53. Sei V ein euklidischer Vektorraum. Die lineare Abbildung $F : V \to V$ heißt **symmetrisch**, wenn

$$\langle F(v), w \rangle = \langle v, F(w) \rangle \quad \text{für alle } v, w \in V.$$

Satz 57. *Ist $F : V \to V$ linear und b eine Orthonormalbasis von V. Dann gilt*

$$F \text{ ist symmetrisch} \quad \Longleftrightarrow \quad {}_bF_b = {}_bF_b{}^T.$$

Beweis. Für $v, w \in V$, $x := {}_bv, y := {}_bw \in \mathbb{R}^n$ ist $\langle v, w\rangle = x^Ty$. Sei $A = {}_bF_b$.
$\Leftarrow$: Aus $A^T = A$ folgt $\langle F(v), w\rangle = (Ax)^Ty = x^TA^Ty = x^TAy = \langle v, F(w)\rangle$.
$\Rightarrow$: Für $v = b_i$, $w = b_j$, $i, j = 1, \dots, n$ ist $x = e_i$ und $y = e_j$. Somit folgt

$$(A^T)_{ij} = e_i^TA^Te_j = (Ae_i)^Te_j = \langle F(b_i), b_j\rangle = \langle b_i, F(b_j)\rangle = e_i{}^TAej = A_{ij}.$$

□

Satz 58. *Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, d.h. $A^T = A$, so sind alle Eigenwerte von A reell.*

Beweis. Sei $\lambda \in \mathbb{C}$ Eigenwert von $A \in \mathbb{R}^{n\times n} \subset \mathbb{C}^{n\times n}$ mit Eigenvektor $w \in \mathbb{C}^n$, $w = u+iv$, $u, v \in \mathbb{R}^n$. Dann ist auch $\overline{\lambda}$ ein Eigenwert von A mit Eigenvektor $\overline{w} = u - iv$, denn $A(u - iv) = \overline{A(u + iv)} = \overline{\lambda(u + iv)} = \overline{\lambda}(u - iv)$.
Somit gilt

$$\begin{aligned}
\lambda\underbrace{(u^Tu + v^Tv)}_{>0} &= \lambda(u - iv)^T(u + iv) = (u - iv)^TA(u + iv) = (u - iv)^TA^T(u + iv) \\
&= \big(A(u - iv)\big)^T(u + iv) = \overline{(u - iv)^T(u + iv)\lambda} = \overline{\lambda}(u^Tu + v^Tv).
\end{aligned}$$

Also ist $\lambda = \overline{\lambda}$, d.h., $\lambda \in \mathbb{R}$. □

Satz 59. *Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, dann gibt es eine Orthonormalbasis aus Eigenvektoren.*

Beweis. Sei $F : \mathbb{R}^n \to \mathbb{R}^n$ mit $F(x) = Ax$, d.h., F ist symmetrisch.
Wegen $A \in \mathbb{R}^{n\times n} \subset \mathbb{C}^{n\times n}$ hat das charakteristische Polynom von A eine komplexe Nullstelle. A besitzt also einen Eigenwert $\lambda_1 \in \mathbb{C}$. Da A symmetrisch ist, ist λ_1 reell und damit gibt es zu λ_1 einen Eigenvektor $b_1 \in \mathbb{R}^n$ mit $\|b_1\| = 1$. Man ergänzt zu einer Orthonormalbasis $c = (b_1, c_2, \dots, c_n)$. Dann ist für $j \geq 2$

$$\langle F(c_j), b_1\rangle = \langle c_j, F(b_1)\rangle = \langle c_j, \lambda_1 b_1\rangle = 0,$$

d.h. $F(c_j) \in \operatorname{span}(c_2, \dots, c_n) =: U$ für $j \geq 2$. U ist also F-invariant und damit

$${}_cF_c = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \boxed{\quad A' \quad} & \\ 0 & & & \end{pmatrix} \qquad \text{mit } A' \in \mathbb{R}^{(n-1)\times(n-1)}.$$

Die Einschränkung $F|_U : U \to U$ ist symmetrisch mit darstellender Matrix A' bezüglich der Basis $(c_2, \dots, c_n)$ von U. Somit besitzt auch sie einen reellen Eigenwert $\lambda_2 \in \mathbb{R}$ mit zugehörigem Eigenvektor $b_2 \in U$, $\|b_2\| = 1$, der senkrecht auf b_1 steht...
Fährt man so fort, so erhält man schließlich eine Orthonormalbasis $(b_1, \dots, b_n)$ des R^n, die aus Eigenvektoren von F, bzw., A besteht. □

Bemerkung. Eine symmetrische Matrix ist immer diagonalisierbar, und zwar sogar durch eine orthogonale Transformation des Koordinatensystems (eine Transformation von der kanonischen auf eine Orthonormalbasis aus Eigenvektoren).