Mathematik für Informatiker III

Institut für Informatik
Freie Universität Berlin
Dozent: Dr. Klaus Kriegel

Mitschrift: Jan Sebastian Siwy

Wintersemester 2002/03

Inhaltsverzeichnis

Einleitung

Themen der Vorlesung:

- Stochastik (Wahrscheinlichkeitstheorie)
  diskret → kontinuierlich
  Ereignis → messbare Ereignisse
  Erwartungswert: Summe → Integral
  Maße für die Abweichung vom Erwartungswert
- Lineare Algebra
  Vektoren
  Basis
  Matrix
  lineare Gleichungssysteme
- Endliche Körper und Codierungstheorie

Kapitel 1

Stochastik

1.1 Wahrscheinlichkeitsräume

1.1.1 Wiederholung

Erläuterung: Ein endlicher diskreter Wahrscheinlichkeitsraum (Ω, p) besteht aus einer endlichen Menge Ω von elementaren Ereignissen von einer Verteilungsfunktion p:

$$p : \Omega \to [0, 1]$$

für die gilt:

$$\sum_{a \in \Omega} p(a) = 1$$

Jede Teilmenge $A \subseteq \Omega$ ist ein Ereignis. Die Verteilungsfunktion p wird erweitert zu einem Wahrscheinlichkeitsmaß, das ebenfalls mit p bezeichnet wird:

$$p : 2^\Omega \to [0, 1] \quad \text{mit} \quad p(A) = \sum_{a \in A} p(a)$$

Bemerkung: Der Ausdruck 2^Ω bezeichnet die Potenzmenge von Ω:

$$2^\Omega = \mathcal{P}(\Omega)$$

Hinweis: Diese Definitionen sind erweiterbar auf abzählbare Mengen Ω von Elementarereignissen.

Beispiel: Es sei gegeben ein Würfel mit den Ereignissraum $\Omega = \{1, 2, \ldots 6\}$ und der gleichverteilten Verteilungsfunktion p. Das Ereignis A sei der Wurf einer geraden Zahl

$$A = \{2, 4, 6\}$$

Die Wahrscheinlichkeit für das Ereignis A beträgt:

$$p(A) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

Eigenschaften: Für das Verscheinlichkeitsmaß gilt ($\bar{A} = \Omega \backslash A$):

$$\begin{aligned} p(\bar{A}) &= 1 - p(A) \\ p(A \cup B) &= p(A) + p(B) \end{aligned}$$

für alle $A, B \subseteq \Omega$ mit $A \cap B = \emptyset$.

1.1.2 Stetige Wahrscheinlichkeitsräume

Verallgemeinerung: Ergebnisse (eines Experiments) sind reelle Zahlen oder (noch allgemeiner) Elemente einer überabzählbaren Menge (z.B. Punkte in einem Raum, Punkte in einer Kreisscheibe).

Probleme: Aus der bisherigen Definition eines diskreten Wahrscheinlichkeitsraumes ergeben sich folgende Probleme:

- Der Ausdruck
$$\sum_{a \in \Omega} p(a)$$
ist nicht sinnvoll, wenn Ω überabzählbar ist.

- Die Potenzmenge der Ereignisse
$$2^{\Omega} = \mathcal{P}(\Omega)$$
führt als Ereignismenge zu weiteren Schwierigkeiten.

Modell:

- Nicht alle $A \subseteq \Omega$ sind Ereignisse (messbar).

- Die Menge $\mathcal{F}$ der Ereignisse ist eine Teilmenge von 2^{Ω} mit den folgenden Eigenschaften:
$$\begin{aligned} A \in \mathcal{F} &\Rightarrow \bar{A} \in \mathcal{F} \\ A_1, A_2, \ldots \in \mathcal{F} &\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \end{aligned}$$
Anmekung: Eine Mengenfamilie $\mathcal{F}$ mit den beiden Eigenschaften nennt man eine *σ-Algebra*. Ist eine Mengefamilie nur abgeschlossen bezüglich Komplement und *endlichen* Vereinigungen, so nennt man sie eine *Algebra*.

Definition: Ein *Wahrscheinlichkeitsraum* ist ein Trippel $(\Omega, \mathcal{F}, p)$, wobei

- $\mathcal{F} \subseteq 2^\Omega$ ist eine σ-Algebra über Ω
- $p : \mathcal{F} \to [0,1]$ ist ein Wahrscheinlichkeitsmaß mit
  - $p(\bar{A}) = 1 - p(A)$ für alle $A \in \mathcal{F}$
  - für jede Folge $A_1, A_2, \ldots$ von paarweise disjunkten Ereignissen $A_i \in \mathcal{F}$ gilt
$$p\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} p(A_i)$$

Eigenschaften und Folgerungen:

- $\mathcal{F}$ ist abgeschlossen gegen endliche und abzählbare Durchschnitte:
$$\begin{aligned} A, B \in \mathcal{F} &\Rightarrow A \cap B = \overline{\bar{A} \cup \bar{B}} \in \mathcal{F} \\ A_1, A_2, \ldots \in \mathcal{F} &\Rightarrow \bigcap_{i=1}^{\infty} A_i = \overline{\bigcup_{i=1}^{\infty} \bar{A}_i} \in \mathcal{F} \end{aligned}$$
- $\mathcal{F}$ ist abgeschlossen gegen Mengendifferenzen:
$$A \backslash B = A \cap \bar{B}$$
- p ist monoton:
$$A \subseteq B \Rightarrow p(A) \leq p(B)$$
  Denn aus $A \subseteq B$ folgt:
$$B = A \cup (B \backslash A)$$
  Die Vereinigung $A \cup (B \backslash A)$ ist disjunkt. Daraus folgt:
$$p(A) \leq p(A) + p(B \backslash A) = p(B)$$

Satz: Ist $A_1 \subseteq A_2 \subseteq \ldots$ eine aufsteigende Folge von Ereignissen und ist A die Vereinigung dieser Ereignisse
$$A = \bigcup_{i=1}^{\infty} A_i$$
dann gilt die für die Wahrscheinlichkeit von A
$$p(A) = \lim_{n \to \infty} p(A_n)$$

Beweis:

- Die Folge $p(A_1), p(A_2), \ldots$ ist monoton wachsend und beschränkt. Damit ist sie auch konvergent.
- Aus $A_1 \subseteq A_2 \subseteq \ldots$ folgt

$$A = A_1 \cup \underbrace{(A_2 \backslash A_1)}_{B_2} \cup \underbrace{(A_3 \backslash A_2)}_{B_3} \cup \ldots$$

Die Vereinigung ist disjunkt. Damit ergibt sich für die Wahrscheinlichkeit von A:

$$\begin{aligned} p(A) &= p(A_1) + \sum_{i=2}^{\infty} p(B_i) \\ &= p(A_1) + \lim_{n\to\infty} \sum_{i=2}^{n} p(B_i) \\ &= \lim_{n\to\infty} (p(A_1) + (p(A_2) - p(A_1)) + (p(A_3) - p(A_2)) + \ldots \\ &\quad + (\underline{p(A_n)} - p(A_{n-1}))) \\ &= \lim_{n\to\infty} p(A_n) \end{aligned}$$

Hinweis: Analog gilt für $B_1 \supseteq B_2 \supseteq \ldots$ mit B als Schnitt über diese Ereignisse

$$B = \bigcap_{i=1}^{\infty} B_i$$

die Wahrscheinlichkeit von B:

$$p(B) = \lim_{n\to\infty} p(B_n)$$

Beispiel: Betrachtet werden zufällige reelle Zahlen aus dem Intervall $[0, 1]$ mit Gleichverteilung:

- Jedes Intervall $(a, b]$ ist Ereignis. Die Wahrscheinlichkeit für dieses Ereignis beträgt

$$p((a, b]) = b - a$$

- Aus der σ-Algebra-Eigenschaft folgt, dass dann auch offene und abgeschlossenen Intervalle in $\mathcal{F}$ sein müssen.
  - abgeschlossene Intervalle:

$$[a, b] = \bigcap_{i=1}^{\infty} \left(a - \frac{a}{i}, b\right]$$

– offene Intervalle:

$$(a, b) = (a, 1]\backslash[b, 1]$$

– auch jede reelle Zahl aus dem Intervall $[0, 1]$:

$$\forall x \in [0, 1] \quad \{x\} \in \mathcal{F}$$

Die Ereignismenge $\mathcal{F}$ besteht somit aus allen abzählbaren disjunkten Vereinigungen von Intervallen (offen, abgeschlossen, halboffen oder Punkt):

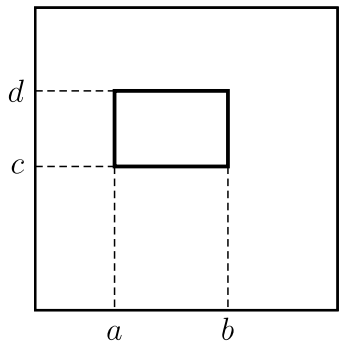
$$\bigcup_{i=1}^{\infty} \langle a_i, b_i \rangle \in \mathcal{F}$$

wobei $b_i \geq a_i$ und $\langle \in \{[, (\}$ und $\rangle \in \{),]\}$ und $b_i \leq a_{i+1}$.

Für die Wahrscheinlichkeit eines solchen Ereignisses gilt somit:

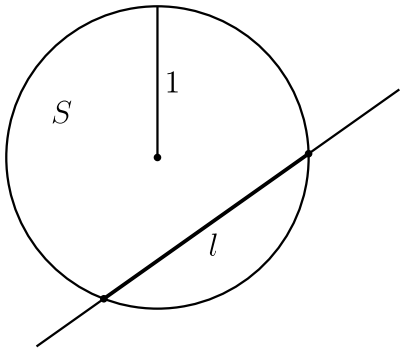
$$p\left(\bigcup_{i=1}^{\infty} \langle a_i, b_i \rangle\right) = \sum_{i=1}^{\infty} (b_i - a_i)$$

Beispiel: Ananlog kann die Gleichverteilung für zufällige Punkte aus dem Einheitsquadrat $[0, 1] \times [0, 1]$ definiert werden.



$$p((a, b] \times (c, d]) = (b - a)(d - c)$$

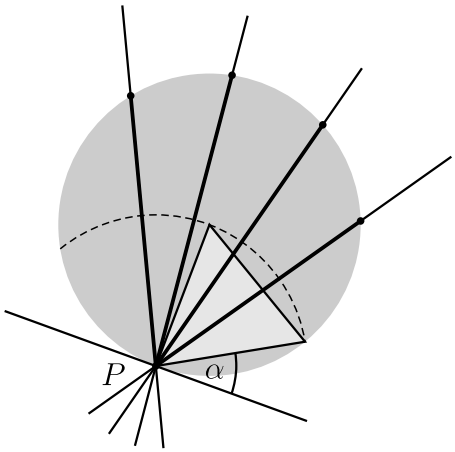
Achtung: Der Begriff Gleichverteilung ist nicht bei jeder Objektklasse so leicht zu beschreiben. Zur Veranschaulichung werden zufällige Geraden betrachtet, die den Einheitskreis S schneiden, und die Länge l der Sehne gemessen.



Das Ereignis A besteht aus allen Geraden, die S schneiden, für die gilt, dass $l \geq 1$.

- **1. Ansatz:**

  Schnittpunkte von Geraden von S sind gleichverteilt auf S. Betrachten nur Geraden durch festen Punkt $P \in S$.

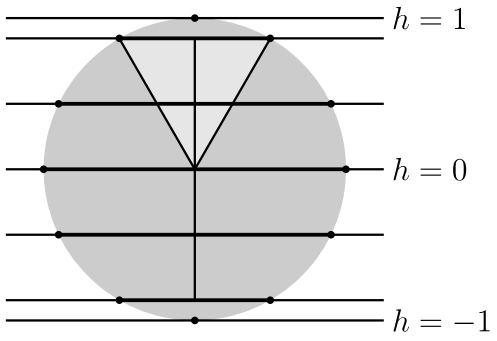


  Geraden durch P sind gleichverteilt bezüglich des Winkels α zur Tangente $(0 \leq \alpha \leq \pi)$.

$$\begin{aligned} l \geq 1 \quad &\Leftrightarrow \quad \frac{\pi}{6} \leq \alpha \leq \frac{5\pi}{6} \\ p(A) \quad &= \quad \frac{\frac{5\pi}{6} - \frac{\pi}{6}}{\pi} = \frac{2}{3} \approx 0{,}67 \end{aligned}$$

- **2. Ansatz:**

  Alle Richtungen sind gleichverteilts, deshalb betrahten wir nur Geraden einer bestimmten Richtung, o.B.d.A. nur horizonrale Geraden.



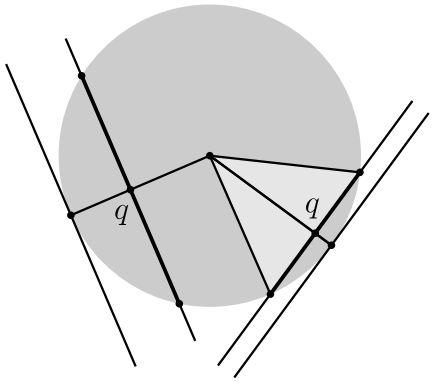
  Geraden sind gleichverteilt bezüglich $-1 \leq h \leq 1$.

$$h^2 + \left(\frac{1}{2}\right)^2 = 1 \quad \Rightarrow \quad h = \frac{\sqrt{3}}{2}$$

$$p(A) \quad = \quad \frac{2 \cdot \frac{\sqrt{3}}{2}}{2} = \frac{\sqrt{3}}{2} \approx 0{,}87$$

- **3. Ansatz:**

  Jede Gerade ist durch den Mittelpunkt q auf ihrer Sehne bestimmt (außer Durchmesser: sind vernachlässigbar).



  Annahme: Punkte q sind gleichverteilt auf Kreisscheibe.

$$l \geq 1 \quad \Leftrightarrow \quad \text{Abstand von } q \text{ zu } (0,0) \leq \frac{\sqrt{3}}{2}$$

$$p(A) \quad = \quad \frac{\text{Fläche von Kreis mit Radius } \frac{\sqrt{3}}{2}}{\text{Fläche von Einheitskreis}} = \frac{\pi \cdot \frac{3}{4}}{\pi \cdot 1} = \frac{3}{4}$$

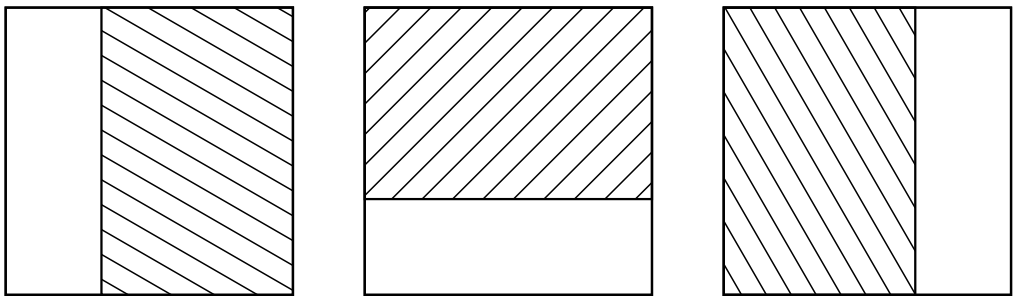
1.1.3 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Definition: Sei $(\Omega, \mathcal{F}, p)$ ein Wahrscheinlichkeitsraum, die Ereignisse $A, B \in \mathcal{F}$ und $p(B) > 0$, so ist die *bedingte Wahrscheinlichkeit* von Ereignis A unter B gegeben durch

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)}$$

Beispiel: Sei $\Omega = [0,1] \times [0,1]$ eine Menge der Ereignisse mit Gleichverteilung:

- $A = \{(x,y) \mid x \geq \frac{1}{3}\}$
- $B = \{(x,y) \mid y \geq \frac{1}{3}\}$
- $C = \{(x,y) \mid x \leq \frac{2}{3}\}$



A B C

$$\begin{aligned} p(A) = p(B) = p(C) &= \frac{2}{3} \\ p(A \cap B) &= \frac{4}{9} \\ p(A \cap C) &= \frac{1}{3} \\ p(A \mid B) &= \frac{\frac{4}{9}}{\frac{2}{3}} = \frac{2}{3} = p(A) \\ p(A \mid C) &= \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2} \neq p(A) \end{aligned}$$

Definition: Zwei Ereignisse A und B sind *unabhängig*, wenn

$$p(A \cap B) = p(A) \cdot p(B)$$

Folgerung: Wenn A und B unabhängig und $p(B) > 0$, dann

$$p(A \mid B) = p(A)$$

Definition: Eine Familie $\{A_i \mid i \in I\}$ von Ereignissen ist unabhängig, wenn für jede endliche Teilmenge $J \subseteq I$

$$p\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} p(A_i)$$

Achtung: Es gibt Familien von paarweise unabhängigen Ereignissen, die nicht unabhängig sind.

Beispiel: Sei $\Omega = \{a, b, c, d\}$ eine Menge der Elementarereignisse mit den Wahrscheinlichkeiten $p(a) = p(b) = p(c) = p(d) = \frac{1}{4}$. Es seien gegeben die folgenden Ereignisse:

- $A = \{a, d\}$
- $B = \{b, d\}$
- $C = \{c, d\}$

$$\begin{aligned} p(A) = p(B) = p(C) &= \frac{1}{2} \\ p(A \cap B) &= \frac{1}{4} = p(A) \cdot p(B) \\ p(A \cap C) &= \frac{1}{4} = p(A) \cdot p(C) \\ p(B \cap C) &= \frac{1}{4} = p(B) \cdot p(C) \end{aligned}$$

Das heißt, dass die Ereignisse paarweise unabhängig sind. Dennoch ist die Familie mit den Ereignissen $\{A, B, C\}$ *nicht* unabhängig:

$$p(A \cap B \cap C) = \frac{1}{4} \quad \neq \quad \frac{1}{8} = p(A) \cdot p(B) \cdot p(C)$$

Satz (Partitionstheorem): Sei $\{B_1, B_2, \ldots\}$ eine Partition von Ω, wobei für alle B_i gilt, dass $p(B_i) > 0$. Dann ist

$$p(A) = \sum_i p(A \mid B_i) \cdot p(B_i) \quad \text{für alle } A \in \mathcal{F}$$

Beweis:

$$\begin{aligned}
p(A) &= p(A \cap \Omega) \\
&= p\left(A \cap \left(\bigcup_i B_i\right)\right) \\
&= p\left(\bigcup_i (A \cap B_i)\right) \quad \text{(disjkunte Vereinigung)} \\
&= \sum_i p(A \cap B_i) \\
&= \sum_i \frac{p(A \cap B_i)}{p(B_i)} \cdot p(B_i) \\
&= \sum_i p(A \mid B_i) \cdot p(B_i)
\end{aligned}$$

Beispiel: Morgen früh regnet es (R) oder schneit (S) oder es gibt keinen Niederschlag (K).

- Bei Regen ist die Wahrscheinlichkeit für eine Busverspätung $\frac{1}{3}$.
- Bei Schnee ist die Wahrscheinlichkeit für eine Busverspätung $\frac{2}{3}$.
- Bei Regen ist die Wahrscheinlichkeit für eine Busverspätung $\frac{1}{6}$.

Die Wettervorhersage:

- $p(R) = \frac{1}{5}$
- $p(S) = p(K) = \frac{2}{5}$

$$p(\text{Busverspätung}) = \frac{1}{3} \cdot \frac{1}{5} + \frac{2}{3} \cdot \frac{2}{5} + \frac{1}{6} \cdot \frac{2}{5} = \frac{2+8+2}{30} = \frac{12}{30} = \frac{2}{5}$$

1.2 Zufallsvariablen

1.2.1 Diskrete Zufallsvariablen

Definition: Sei $(\Omega, \mathcal{F}, p)$ ein Wahrscheinlichkeitsraum. Eine Funktion

$$X : \Omega \to \mathbb{R}$$

ist eine *diskrete Zufallsvariable*, falls

- das Bild der Zufallsvariable

$$\text{Im } X = \{x \in \mathbb{R} \mid \exists\, \omega \in \Omega \quad x = X(\omega)\}$$

  abzählbar ist und

- für alle $x \in \mathbb{R}$:

$$X^{-1}(x) = \{\omega \mid X(\omega) = x\} \in \mathcal{F}$$

Konsequenz: Für alle $T \subseteq \mathbb{R}$ ist

$$X^{-1}(T) \in \mathcal{F}$$

denn

$$X^{-1}(T) = \bigcup_{x \in (T \cap \text{Im } X)} X^{-1}(x)$$

Definition: Sei X eine diskrete Zufallsvariable auf $(\Omega, \mathcal{F}, p)$. Die *Gewichtsfunktion* $p_X : \mathbb{R} \to [0, 1]$ von X sei wie folgt definiert:

$$p_X(x) = p(X^{-1}(x))$$

Oft verwendet man für $p_X(x)$ auch die intuitivere Schreibweise $p(X = x)$.

Konsequenz:

$$\begin{aligned} \sum_{x \in \text{Im } X} p_X(x) &= \sum_{x \in \text{Im } X} p(\{\omega \mid X(\omega) = x\}) \\ &= p\left(\bigcup_{x \in \text{Im } X} \{\omega \mid X(\omega) = x\} \right) \\ &= p(\Omega) = 1 \end{aligned}$$

Die Funktion p_X charakterisiert die Zufallsvariable X sehr genau, in dem Sinne, dass man für jede solche Beschreibung eine Realisierung durch einen Wahrscheinlichkeitsraum und eine Zufallsvariable finden kann.

Sei $S \subseteq \mathbb{R}$ abzählbar und für jedes $s \in S$ sei eine Zahl $\pi_s \in [\,0,1\,]$ gegeben mit

$$\sum_{s \in S} \pi_s = 1$$

Konstruktion:

- $\Omega = S$
- $\mathcal{F} = \mathcal{P}(S)$
- $p(A) = \sum_{s \in A} \pi_s$ für jedes $A \subseteq S$
- $X : \Omega \to \mathbb{R}$
- $X(s) = s$

Beispiele: p ohne Zusatz ist eine Zahl aus $[\,0,1\,]$, $q = 1 - p$

1. *Bernoulli-Verteilung*:
   Eine Zufallsvariable X mit Bernoulli-Verteilung:

   - Im $X = \{0,1\}$
   - $p_X(0) = q$ und $p_X(1) = p$

   Probe: $p_X(0) + p_X(1) = p + q = p + (1-q) = 1$

   z.B. Münzwurf mit „unfairer“ Münze:

   - Wahrscheinlichkeit für Kopf K: p
   - Wahrscheinlichkeit für Zahl Z: $q = 1 - p$

   Die Zufallsvariable wird folgendermaßen definiert:

   $$\begin{aligned} X &: \{\mathrm{K}, \mathrm{Z}\} \to \{0,1\} \\ X(K) &= 1 \\ X(Z) &= 0 \end{aligned}$$

2. *Binominialverteilung*
   Eine Zufallsvariable X mit Binominialverteilung mit den Parametern n und p:

   - Im $X = \{0,1,\ldots,n\}$ und
   - $p_X(k) = \binom{n}{k}\ p^k\ q^{n-k}$ für alle $k \in \{0,1,\ldots,n\}$

Probe: $\sum_{k=0}^{n} p_X(k) = \sum_{k=0}^{n} \binom{n}{k}\ p^k\ q^{n-k} = (p+q)^n = 1^n = 1$

Binominialverteilung tritt auf z.B. bei n-facher Wiederholung eines Bernoulli-Experiments (unabhängig).

- $\Omega = \{\mathrm{K}, \mathrm{Z}\}^n$
- $\omega = (a_1, a_2, \ldots, a_n)$ mit $a_i = \{\mathrm{K}, \mathrm{Z}\}$
- $p(\omega) = p^{k(\omega)} \cdot q^{z(\omega)}$
- $X(\omega) := k(\omega)$

mit $k(\omega)$ = Anzahl der Köpfe in ω und $z(\omega)$ = Anzahl der Zahlen in ω

Wahrscheinlichkeit, dass bei n Münzwürfen genau k-mal Kopf fällt:

$$\begin{aligned} p_X(k) &= p(\{\omega \mid X(\omega) = k\}) \\ &= p(\{\omega \mid k(\omega) = k\}) \\ &= \sum_{\substack{\omega \text{ mit} \\ k(w)=k}} p(\omega) \\ &= \sum_{\substack{\omega \text{ mit} \\ k(w)=k}} p^k\ q^{n-k} \\ &= \binom{n}{k}\ p^k\ q^{n-k} \end{aligned}$$

3. *Geometrische Verteilung*:
   Wiederholung eines Wurfes einer (p, q)-Münze so lange, bis zum erstem mal K auftritt:

   - $\Omega = \{\mathrm{K}, \mathrm{ZK}, \mathrm{ZZK}, \mathrm{ZZZK}, \mathrm{ZZZZK}, \ldots\}$
   - $p((\mathrm{K})) = p$, $p((\mathrm{ZK}) = qp$, $p((\mathrm{Z}^l\mathrm{K})) = q^l\ p$
   - $X(\omega) = |\omega|$ (Anzahl der Würfe)
   - $\mathrm{Im}(X) = \{1, 2, 3, \ldots\}$
   - $p_X(k) = p(\{\omega \mid |\omega| = k\}) = p(\mathrm{Z}^{k-1}K) = q^{k-1} \cdot p$

4. *Poisson-Verteilung*:
   Eine Zufallsvariable X mit Poisson-Verteilung mit Parameter λ:

   - $\mathrm{Im}(X) = \mathbb{N}$ (inkl. der Null)
   - $p_X(k) = \frac{1}{k!}\ \lambda^k\ e^{-\lambda}$

Probe: $\sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} \frac{1}{k!}\, \lambda^k\, e^{-\lambda} = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{1}{k!}\, \lambda^k = e^{-\lambda} \cdot e^{\lambda} = 1$

Die Poisson-Verteilung tritt auf als Grenzwert von Binominialverteilung mit n groß, p klein, $\lambda = n \cdot p$ und $k \ll n$.

$$\begin{aligned} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k (1-p)^{n-k} \\ &\approx \frac{n^k}{k!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &\approx \frac{\lambda^k}{k!} \cdot e^{-\lambda} \end{aligned}$$

1.2.2 Stetige Zufallsvariablen

Definition: Eine *Zufallsvariable* auf $(\Omega, \mathcal{F}, p)$ ist eine Abbildung $X : \Omega \to \mathbb{R}$, so dass für alle $x \in \mathbb{R}$ gilt:

$$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}$$

Definition: Die Funktion

$$F_X(x) = p(\{\omega \in \Omega \mid X(\omega) \leq x\})$$

nennt man die *Verteilungsfunktion* von X.

Lemma: Für jede Verteilungsfunktion $F = F_X$ einer Variablen $X : \Omega \to \mathbb{R}$ gilt:

a) $x \leq y \quad \Rightarrow \quad F(x) \leq F(y)$

b) $\lim_{x \to -\infty} F(x) = 0$ und $\lim_{x \to \infty} F(x) = 1$

c) $\forall\, x \in \mathbb{R} \quad \lim_{h \to 0+} F(x+h) = F(x)$

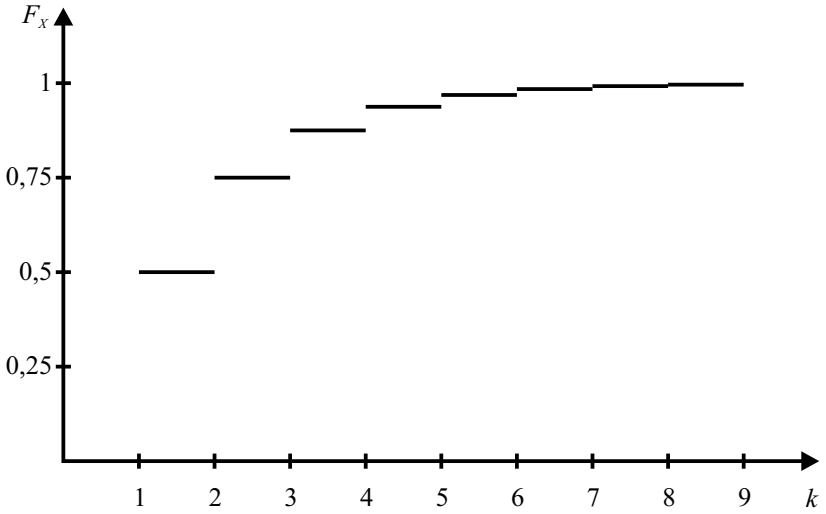
Beweis: Sei $A_x = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}$ und $F(x) = p(A_x)$

a) $x \leq y \quad \Rightarrow \quad A_x \subseteq A_y \quad \Rightarrow \quad \underbrace{p(A_x)}_{F(x)} \leq \underbrace{p(A_y)}_{F(y)}$

b) $\emptyset = \bigcap_{x=1}^{\infty} A_{-x} \quad \Rightarrow \quad 0 = p(\emptyset) = \lim_{x \to \infty} p(A_{-x}) = \lim_{x \to -\infty} F(x)$

$\Omega = \bigcup_{x=1}^{\infty} A_x \quad \Rightarrow \quad 1 = p(\Omega) = \lim_{x \to \infty} p(A_n) = \lim_{x \to \infty} F(x)$

Achtung: Die Funktion F_X ist nicht zwingend stetig, z.B. bei geometrischer Verteilung:



Definition: Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ ist *stetig*, wenn eine Funktion $f : \mathbb{R} \to \mathbb{R}^+$ existiert, so dass

$$F_X(x) = \int_{\infty}^{x} f(n)\, dn$$

f wird *Dichte* der Verteilung genannt.

Beispiel: Gleichverteilung im Intervall $[\,1, 3\,]$.

Ist $1 \leq y \leq x \leq 3$, so ist

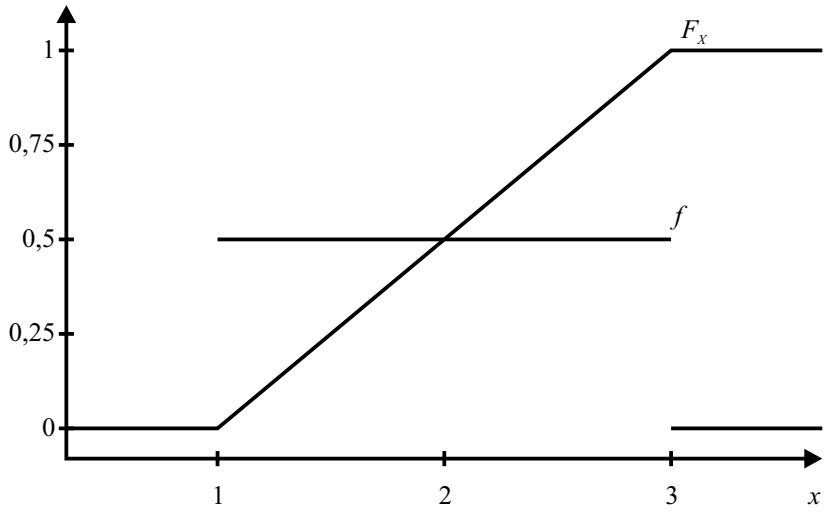
$$p(\{\omega \mid \omega \in [\,y, x\,]\}) = \frac{x - y}{3 - 1}$$

X hat die Verteilung

$$F_X(x) = p(\{\omega \mid X(\omega) \leq x\}) = \frac{x - 1}{3 - 1} = \frac{x - 1}{2}$$

Suche f, so dass

$$F_X(x) = \int_{-\infty}^{\infty} f(n)\, dn$$



Gesuchte Funktion lautet:

$$f(x) = \begin{cases} \frac{1}{2} & \text{falls} \quad x \in [\,1, 3\,] \\ 0 & \text{sonst} \end{cases}$$

Vergleich:

- diskret
  - Gewichtsfunktion p_X
  - Addition der Einzelwahrscheinlichkeiten
- stetig
  - Dichtefunktion f
  - Integrieren über der Dichtefunktion

1.3 Erwartungswert

1.3.1 Bestimmung des Erwartungswertes

Definition: Ist $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable, so ist der der *Erwartungswert* von X definiert durch

$$E(X) = \sum_{x \in \mathrm{Im}X} x \cdot p_X(x) = \sum_{x \in \mathrm{Im}X} x \cdot p(\{\omega \mid X(\omega) = x\})$$

falls diese Reihe absolut konvergiert.

Definition: Ist $X : \Omega \to \mathbb{R}$ eine stetige Zufallsvariable mit der Dichtefunktion f, so ist

$$E(X) = \int_{\infty}^{\infty} x \cdot f(x)dx$$

falls beide uneigentlichen Integrale existieren.

Lemma (Linearität der Erwartungswerte): Sind X und Y Zufallsvariablen über $(\Omega, \mathcal{F}, p)$ mit den Erwartungswerten $E(X)$ und $E(Y)$, dann gilt:

- $E(X + Y) = E(X) + E(Y)$
  mit $(X + Y)(\omega) = X(\omega) + Y(\omega)$
- $E(\alpha \cdot X) = \alpha \cdot E(X)$
  mit $(\alpha \cdot X)(\omega) = \alpha \cdot X(\omega)$ und $\alpha \in \mathbb{R}$

Beispiele:

1. Zufallsvariable X mit Bernoulli-Verteilung mit Parameter p:
   - $\mathrm{Im}X = \{0, 1\}$
   - $p_X(1) = p$ und $p_X(0) = 1 - p$
   - $E(X) = 1 \cdot p + 0 \cdot (1 - p) = p$
2. Zufallsvariable X mit Binominialverteilung mit den Parametern n und p:
   - $X = X_1 + X_2 + X_3 + \ldots X_n \quad \text{mit } X_i = \begin{cases} 1 & \text{falls } i\text{-ter Wurf } K \text{ ist} \\ 0 & \text{falls } i\text{-ter Wurf } Z \text{ ist} \end{cases}$
   - $E(X_i) = p$
   - $E(X) = E(X_1) + E(X_2) + E(X_3) + \ldots + E(X_n) = n \cdot p$

3. Zufallsvariable X mit geometrischer Verteilung mit Parameter p:

   - $\mathrm{Im}X = \mathbb{N}^+$
   - $p_X(k) = (1-p)^{k-1} \cdot p = q^{k-1} \cdot p$

   Bestimmung des Erwartungswertes:

$$\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p \\
&= p \cdot \left(\sum_{k=1}^{\infty} q^{k-1} + \sum_{k=2}^{\infty} q^{k-1} + \sum_{k=3}^{\infty} q^{k-1} + \ldots \right) \\
&= p \cdot \left(\sum_{k=0}^{\infty} q^{k} + q \cdot \sum_{k=0}^{\infty} q^{k} + q^2 \cdot \sum_{k=0}^{\infty} q^{k} + \ldots \right) \\
&= p \cdot \left(\frac{1}{1-q} + \frac{q}{1-q} + \frac{q^2}{1-q} + \ldots \right) \\
&= \frac{p}{1-q} \cdot \left(1 + q + q^2 + \ldots \right) \\
&= 1 \cdot \frac{1}{1-q} \\
&= \frac{1}{p}
\end{aligned}$$

4. Zufallsvariable X mit Poisson-Verteilung mit Paramter λ:

   - $\mathrm{Im}X = \mathbb{N}$
   - $p_X(k) = \frac{1}{k!} \cdot \lambda^k \cdot e^{-\lambda}$

   Bestimmung des Erwartungswertes:

$$\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} k \cdot \frac{1}{k!} \cdot \lambda^k \cdot e^{-\lambda} \\
&= \sum_{k=1}^{\infty} k \cdot \frac{1}{(k-1)!} \cdot \lambda^k \cdot e^{-\lambda} \\
&= \lambda \cdot \sum_{k=1}^{\infty} k \cdot \frac{1}{(k-1)!} \cdot \lambda^{k-1} \cdot e^{-\lambda} \\
&= \lambda
\end{aligned}$$

5. stetige Zufallsvariable X mit Gleichverteilung über einem Intervall $[\,a, b\,]$:

$$\begin{aligned}
E(X) &= \int_{\infty}^{\infty} x \cdot f(x)dx \\
&= \int_{a}^{b} x \cdot \frac{1}{b-a}\, dx \\
&= \frac{1}{b-a} \cdot \int_{a}^{b} x\, dx \\
&= \frac{1}{b-a} \cdot \left[\frac{1}{2}\, x^2\right]_{a}^{b} \\
&= \frac{1}{b-a} \cdot \left(\frac{1}{2}\, b^2 - \frac{1}{2}\, a^2\right) \\
&= \frac{b^2 - a^2}{2(b-a)} \\
&= \frac{(b+a)(b-a)}{2(b-a)} \\
&= \frac{a+b}{2}
\end{aligned}$$

1.3.2 Abweichungen vom Erwartungswert

Satz (Markow-Ungleichung): Sei $X : \Omega \to \mathbb{R}^{\geq 0}$ eine Zufallsvariable mit dem Erwartungswert $E(X)$ unt $t > 0$, dann gilt:

$$p(X \geq t) \leq \frac{E(X)}{t}$$

Beweis für diskrete Variablen:

$$\begin{aligned}
E(X) &= \sum_{x \in \mathrm{Im}X} x \cdot p(X = x) \\
&= \sum_{\substack{x \in \mathrm{Im}X \\ x < t}} x \cdot p(X = x) + \sum_{\substack{x \in \mathrm{Im}X \\ x \geq t}} x \cdot p(X = x) \\
&\geq \sum_{\substack{x \in \mathrm{Im}X \\ x \geq t}} x \cdot p(X = x) \\
&\geq t \cdot \sum_{\substack{x \in \mathrm{Im}X \\ x \geq t}} p(X = x) \\
&= t \cdot p(X \geq t)
\end{aligned}$$

$$\frac{E(X)}{t} \geq p(X \geq t)$$

Satz: Sei $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable und $g : \mathbb{R} \to \mathbb{R}$ eine beliebige Funktion, dann ist $Y = gX : \Omega \to \Omega$ eine Zufallsvariable mit

$$Y(\omega) = g(X(\omega))$$

und

$$E(Y) = \sum_{x \in \mathrm{Im}X} g(x) \cdot p_X(x)$$

falls diese Reihe absolut konvergiert.

Beispiel: Sei X eine Zufallsvariable mit geometrischer Verteilung mit dem Parametern p und $g(x) = x^2$ eine Funktion ($q = 1 - p$):

$$\begin{aligned}
E(gX) &= E(X^2) \\
&= \sum_{k=1}^{\infty} k^2 \cdot q^{k-1} \cdot p \\
&= \sum_{k=1}^{\infty} 1^2 \cdot q^{k-1} \cdot p + \sum_{k=2}^{\infty} \underbrace{(2^2 - 1^2)}_{(2+1)(2-1)} \cdot q^{k-1} \cdot p + \sum_{k=1}^{\infty} \underbrace{(3^2 - 2^2)}_{(3+2)(3-2)} \cdot q^{k-1} \cdot p + \ldots \\
&= 1 + (1+2) \cdot q \cdot \underbrace{\sum_{k=1}^{\infty} p \cdot q^{k-1}}_{1} + (1+4) \cdot q^2 \cdot \underbrace{\sum_{k=1}^{\infty} p \cdot q^{k-1}}_{1} + (1+6) \cdot q^3 \cdot \ldots \\
&= 1 + q + q^2 + q^3 + \ldots + 2 \cdot q + 4 \cdot q^2 + 6 \cdot q^3 + \ldots \\
&= \frac{1}{1-q} + \frac{2q}{p^2} \cdot \sum_{k=0}^{\infty} k \cdot q^k \cdot p \\
&= \frac{1}{p} + \frac{2q}{p^2} \\
&= \frac{p + 2 - 2p}{p^2} \\
&= \frac{2-p}{p^2}
\end{aligned}$$

Definition: Die Erwartungswerte $E(X^i)$ werden i-tes *Moment* von X genannt.

Definition: Die *Varianz* einer Zufallsvariable X mit $E(X) = \mu$ ist

$$\begin{aligned} Var(X) &= E((X-\mu)^2) \\ &= E((X-E(X))^2) \\ &= E\left(X^2 - 2 \cdot X \cdot E(X) + (E(X))^2\right) \\ &= E(X^2) - 2 \cdot E(X) \cdot E(X) + (E(X))^2 \\ &= E(X^2) - 2 \cdot (E(X))^2 + (E(X))^2 \\ &= \underbrace{E(X^2)}_{\text{2. Moment}} - \underbrace{(E(X))^2}_{\text{(1. Moment)}^2} \end{aligned}$$

Die Größe $\sigma = \sqrt{Var(X)}$ wird *Standardabweichung* von X genannt.

Beispiele:

1. Zufallsvariable X mit Bernoulli-Verteilung mit Parameter p:

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= (1^2 \cdot p + 0^2 \cdot (1-p)) - (1 \cdot p + 0 \cdot (1-p))^2 \\ &= p - p^2 \end{aligned}$$

2. Zufallsvariable X mit Binominialverteilung mit den Parametern n und p:

   An der Stelle kann genutzt werden, dass für *unabhängige* Zufallsvariablen gilt:

   - $Var(X+Y) = Var(X) + Var(Y)$
   - $p(X = x \wedge Y = y) = p(X = x) \cdot p(Y = y)$

   Eine Zufallsvariable X mit Binominialverteilung ist eine Summe aus unabhängigen Bernoulli-Variablen:

$$X = X_1 + X_2 + \ldots + X_n$$

   Damit ergibt sich für die Varianz:

$$Var(X) = n \cdot (p - p^2)$$

3. Zufallsvariable X mit geometrischer Verteilung mit Parameter p:

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 \\ &= \frac{1-p}{p^2} \\ &= \frac{1}{p^2} - \frac{1}{p} \end{aligned}$$

Satz (Tschebyscheff-Ungleichung): Sei X eine Zufallsvariable mit dem Erwartungswert $E(X) = \mu$ und der Varianz $Var(X) = \sigma^2$, dann gilt für alle $c > 0$:

$$p(|X - \mu| \geq c) \quad \leq \quad \frac{\sigma^2}{c^2}$$

Spezialfall für $E(X) = \mu = 0$:

$$p(|X| \geq c) \quad \leq \quad \frac{E(X^2)}{c^2}$$

Beispiel: Zufallsvariable X mit Binominialverteilung mit den Parametern n und $p = \frac{1}{2}$:

$$\begin{aligned}
E(X) &= n \cdot \frac{1}{2} = \frac{n}{2} \\
Var(X) &= n \cdot \left(\frac{1}{2} - \left(\frac{1}{2} \right)^2 \right) = \frac{n}{4} = \sigma^2 \\
\text{Wähle: } c &= \frac{n}{4} \\
p\left(\left| X - \frac{n}{2} \right| \geq \frac{n}{4} \right) &\leq \frac{\frac{n}{4}}{\left(\frac{n}{4} \right)^2} = \frac{4}{n} \\
\lim_{n \to \infty} \frac{4}{n} &= 0
\end{aligned}$$

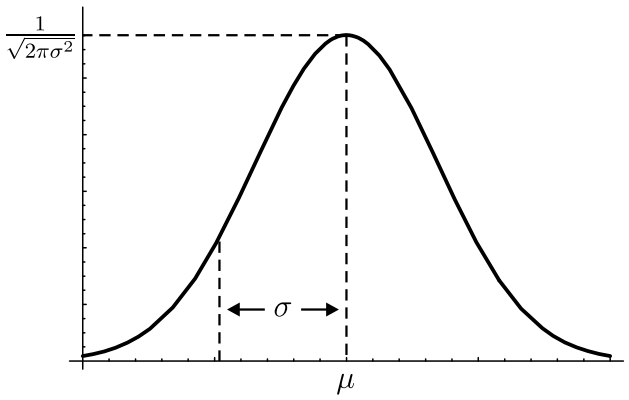
Das heißt, dass die Wahrscheinlichkeit dafür, dass bei n Würfen weniger als $\frac{1}{4}$ oder mehr als $\frac{3}{4}$ der Ergebnisse Köpfe sind, geht für große n gegen 0.

Dagegen die Abschätzung mit der Markow-Ungleichung:

$$p\left(X \geq \frac{3}{4} \cdot n \right) \quad \leq \quad \frac{\frac{n}{2}}{\frac{3}{4} \cdot n} = \frac{2}{3}$$

Gaus'sche Normalverteilung: $N(\mu, \sigma^2)$ mit Dichtefunktion f

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}\cdot(x-\mu)^2}$$



Hinweis: Der Abstand von μ zum Wendepunkt von $f(x)$ beträgt σ.

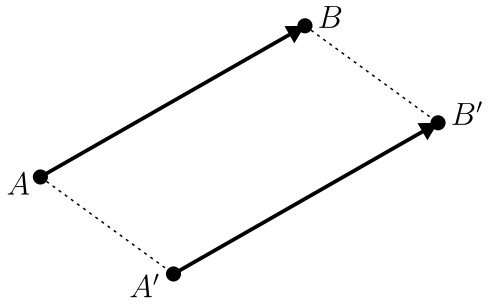
Kapitel 2

Lineare Algebra

2.1 Vektoren – der intuitive Ansatz

2.1.1 Koordinatenfreie Einführung

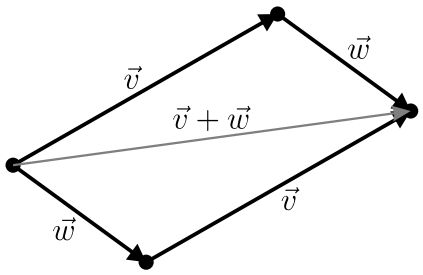
Definition: Ein *Vektor* wird durch geordnetes Punktepaar $\overrightarrow{AB}$ repräsentiert (veranschaulicht durch die gerichtete Strecke von A nach B), wobei zwei Vektoren $\overrightarrow{AB}$ und $\overrightarrow{A'B'}$ gleich sind, wenn es eine Translation (Parallelverschiebung) gibt, die A in A' und B in B' überführt.



Zur Betonung des Aspekts, dass der Anfangspunkt beliebig sein kann, wird der Begriff des *freien Vektors* verwendet.

Legt man einen Bezugspunkt O im Raum fest und betrachtet für einen beliebigen Punkt P den Vektor $\overrightarrow{OP}$, so wird dieser der *Ortsvektor* von P genannt.

Addition von Vektoren: Vertoren werden durch Aneinanderkettung addiert.



Kommutativität kann man durch Parallelogrammeigenschaften sehen.

Nullvektor: Der Vektor $\overrightarrow{OO} = \overrightarrow{AA}$ wird Nullvektor genannt und kurz mit $\vec{0}$ bezeichnet. Der Nullvektor ist das neutrale Element der Addition:

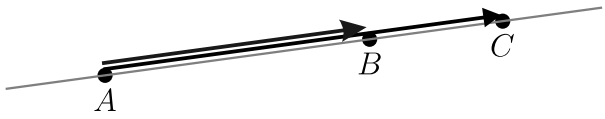
$$\vec{v} + \vec{0} = \vec{v}$$

Inverser Vektor: Der Vektor $\overrightarrow{BA}$ wird als zu $\overrightarrow{AB}$ invers bezeichnet:

$$\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \vec{0}$$

Betrag: Der Betrag (Länge, Norm) des Vektors $\overrightarrow{AB}$ ist der Abstand zwischen A und B und wird mit $\left\|\overrightarrow{AB}\right\|$ bezeichnet.

Multiplikation mit Skalaren:



Liegen drei Punkte A, B und C in dieser Reihefolge auf einer Geraden und ist

$$\left\|\overrightarrow{AB}\right\| = \lambda \cdot \left\|\overrightarrow{AC}\right\|$$

so sagt man

$$\overrightarrow{AB} = \lambda \cdot \overrightarrow{AC}$$

Auf diese Weise wird Multiplikation von reellen Zahlen (Skalaren) mit Vektoren eingeführt.

2.1.2 Koordinatensystem

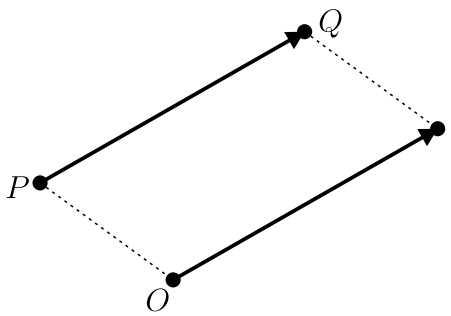
Betrachtet man zusätzlich ein Koordinatensystem im Raum mit dem Ursprung $O = (0, 0, 0)$, so kann jedem Punkt $P = (p_1, p_2, p_3)$ der Ortsvektor

$$\overrightarrow{OP} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

zugeordnet werden.

Kosequenz: Aus $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$ folgt:

$$\overrightarrow{PQ} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix}$$



Addition:

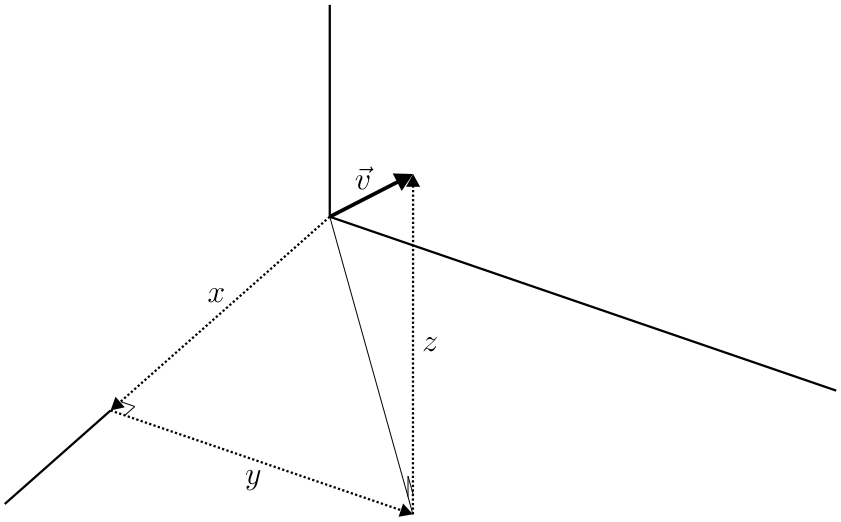
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$$

Multiplikation mit Skalaren:

$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \lambda \cdot x_3 \end{pmatrix}$$

Betrag:

$$\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \Rightarrow \quad \|\vec{v}\| = \sqrt{x^2 + y^2 + z^2}$$



Fazit: Man kann Vektoren im Raum genauso darstellen wie Punkte, aber im Gegensatz zu Punkten können Vektoren addiert und mit Skalaren multipliziert werden.

2.1.3 Zusammenhang zwischen Vektoren und linearen Gleichungssystemen (LGS)

Beispiel: Die folgenden Probleme sind äquivalent:

- Hat dieses lineare Gleichungssystem eine Lösung?

$$\begin{array}{rcrcr} 5\alpha & + & 3\beta & = & -1 \\ 4\alpha & + & \beta & = & 2 \\ 3\alpha & - & \beta & = & 5 \end{array}$$

- Gibt es entsprechende α und β?

$$\alpha \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix} + \beta \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}$$

- Liegt der Punkt $(-1, 2, 5)$ in der Ebene, die von den Punkten $(0, 0, 0)$, $(5, 4, 3)$ und $(3, 1, -1)$ aufgespannt wird?

2.2 Vektorräume

2.2.1 Vektorräume

Definition: Eine Menge K mit zwei Operationen $\oplus$ und $\odot$ und zwei Elementen 0 und 1 (wobei $0 \neq 1$) ist ein *Körper*, falls

- $(K, \oplus)$ ist kommutative Gruppe mit neutralem Element 0
  ($\ominus k$ ist inverses Element zu k bezüglich $\oplus$)
- $(K \backslash \{0\}, \odot)$ ist kommutative Gruppe mit neutralem Element 1
  ($k^{-1} = \frac{1}{k}$ ist inverses Element zu k bezüglich $\odot$)
- $\forall\, k, l, m \in K \quad k \odot (l \oplus m) = (k \odot l) \oplus (k \odot m)$
  (Distributivität)

Beispiele: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ (nicht $\mathbb{Z}$)

Definition: Eine Menge V mit den Operationen

- $\oplus : V \times V \to V$
- $\odot : K \times V \to V$

und dem Element $\vec{0}$ wird *Vektorraum* (VR) über dem Körper K genannt, falls

- $(V, \oplus)$ ist kommutative Gruppe mit neutralem Element $\vec{0}$
  ($\ominus \vec{v}$ ist inverses Element zu $\vec{v}$ bezüglich $\oplus$)
- $\forall\, \lambda, \mu \in K \quad \forall\, \vec{v} \in V \quad \lambda \odot (\mu \odot \vec{v}) = (\lambda \cdot \mu) \odot \vec{v}$
  (Assoziativität der Multiplikationen)
- $\forall\, \vec{v} \in V \quad 1 \odot \vec{v} = \vec{v}$
  (neutrales Element bezüglich der Multiplikation)
- $\forall\, \lambda, \mu \in K \quad \forall\, \vec{v} \in V \quad (\lambda + \mu) \odot \vec{v} = (\lambda \odot \vec{v}) \oplus (\mu \odot \vec{v})$
  (Distributivität 1)
- $\forall\, \lambda \in K \quad \forall\, \vec{v}, \vec{w} \in V \quad \lambda \odot (\vec{v} \oplus \vec{w}) = (\lambda \odot \vec{v}) \oplus (\lambda \odot \vec{w})$
  (Distributivität 2)

Beispiele:

1. Der reelle Vektorraum $\mathbb{R}^n$ über dem Körper $\mathbb{R}$:

$$V = \mathbb{R}^n = \{(x_1, \dots x_n) \mid x_i \in \mathbb{R}\}$$

Addition:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Multiplikation:

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

Nullvektor:

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

Inverser Vektor:

$$- \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}$$

2. Der Vektorraum der stetigen Funktionen über dem Körper $\mathbb{R}$:

$$V = \{f \mid f : [\,0, 1\,] \to \mathbb{R}, \text{ stetig}\}$$

Addition:

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

Multiplikation:

$$(\lambda f)(x) = \lambda(f(x))$$

3. Der Vektorraum $\mathbb{R}$ über dem Körper $\mathbb{Q}$

$$V = \mathbb{R} \quad \text{und} \quad K = \mathbb{Q}$$

Addition:

$$\vec{r}, \vec{s} \in \mathbb{R} \quad \vec{r} + \vec{s} = \overrightarrow{(r + s)}$$

Multiplikation:

$$\lambda \in \mathbb{Q} \quad \vec{r} \in \mathbb{R} \quad \lambda \cdot \vec{r} = \overrightarrow{(\lambda \cdot r)}$$

2.2.2 Unterräume

Definition: Eine Teilmenge $U \neq \emptyset$ eines Vektorraums V über K ist *Unterraum* (Untervektorraum, UR) von V, falls

- $\forall\, \vec{v}, \vec{w} \in U \quad \vec{v} + \vec{w} \in U$
- $\forall\, \vec{v} \in U \quad \forall\, \lambda \in K \quad \lambda\vec{v} \in U$

Beispiele:

1. $V = \mathbb{R}^n = \{(x_1, \ldots x_n) \mid x_i \in \mathbb{R}\}$

   $\rightarrow\ U = \{(x_1, x_2, 0, \ldots 0) \mid x_1, x_2 \in \mathbb{R}\}$

   $\rightarrow$ speziell für R^3: jede Ebene und jede Gerade durch $(0, 0, 0)$

2. $V = \{f \mid f : [0, 1] \rightarrow \mathbb{R},\ \text{stetig}\}$

   $\rightarrow\ U = \{f \mid f : [0, 1] \rightarrow \mathbb{R},\ f \text{ ist konstant}\}$

   $\rightarrow\ U' = \{f \mid f : [0, 1] \rightarrow \mathbb{R},\ f \text{ ist linear, d.h. } f(x) = ax + b\}$

3. $V = \mathbb{R}$ über dem Körper $\mathbb{Q}$

   $\rightarrow\ U = \{q_1 + q_2\sqrt{2} \mid q_1, q_2 \in \mathbb{Q}\}$

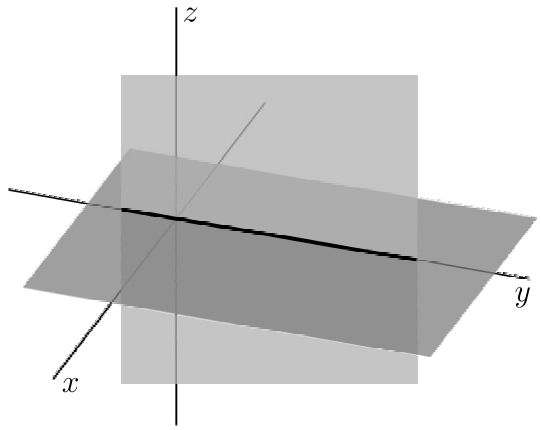
Satz: Sei V ein Vektorraum über einem Körper K und $\{U_i \mid i \in I\}$ eine Familie von Unterräumen, dann ist $\bigcap\limits_{i \in I} U_i$ auch ein Unterraum von V.

Beweis: Sei $\vec{u}$, $\vec{v} \in \bigcap\limits_{i \in I} U_i$ und $\lambda \in K$, dann gilt

- $\vec{u}$ und $\vec{v}$ sind Elemente von allen U_i
- $\vec{u} + \vec{v}$ und $\lambda\vec{u}$ sind Elemente von allen U_i

Daraus folgt, dass $\vec{u} + \vec{v} \in \bigcap\limits_{i \in I} U_i$ und $\lambda\vec{u} \in \bigcap\limits_{i \in I} U_i$.

Beispiel: Durchschnitt von xy-Ebene und der yz-Ebene in $\mathbb{R}^3$ ist die y-Achse.



Folgerung:

- Der Nullvektor $\vec{0}$ gehört zu jeden Unterraum:

$$\forall\, U \;\; \text{UR } V \text{ gilt } \vec{0} \in U$$

- Zu jeden Vektor $\vec{v}$ aus dem Unterraum gehört auch der inverse Vektor $-\vec{v}$ zum Unterraum:

$$\forall\, U \text{ UR } V \;\; \forall\, \vec{v} \in U \text{ gilt } \; -\vec{v} \in U$$

2.2.3 Linearkombinationen und lineare Hülle

Definition: Sind $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k \in V$ und $\lambda_1, \lambda_2, \ldots \lambda_k \in K$, so nennt man den Vektor

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k$$

eine *Linearkombination* (LK) aus $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k$.

Lemma: Die Menge aller Linearkombinationen

$$\{\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k \mid \lambda_i \in K\}$$

der Vektoren $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k$ bildet einen Unterraum.

Beweis: Seien $\vec{v}, \vec{w} \in U$ und $\alpha \in K$ mit

- $\vec{v} = \lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \ldots + \lambda_k\vec{u}_k$
- $\vec{w} = \mu_1\vec{u}_1 + \mu_2\vec{u}_2 + \ldots + \mu_k\vec{u}_k$

dann gilt:

$$\begin{aligned} \vec{v} + \vec{w} &= (\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \ldots + \lambda_k\vec{u}_k) + (\mu_1\vec{u}_1 + \mu_2\vec{u}_2 + \ldots + \mu_k\vec{u}_k) \\ &= (\lambda_1 + \mu_1)\vec{u_1} + (\lambda_2 + \mu_2)\vec{u_2} + \ldots + (\lambda_k + \mu_k)\vec{u}_k \in U \end{aligned}$$

$$\begin{aligned} \alpha\vec{v} &= \alpha(\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \ldots + \lambda_k\vec{u}_k) \\ &= (\alpha\lambda_1)\vec{u}_1 + (\alpha\lambda_2)\vec{u}_2 + \ldots + (\alpha\lambda_k)\vec{u}_k \in U \end{aligned}$$

Definition: Sei $M \subseteq V$ eine Menge von Vektoren, dann ist die *lineare Hülle* (Lin) von M der kleinste (bezüglich Inklusion) Unterraum von V, der M enthält, d.h.

$$\mathrm{Lin}(M) = \bigcap_{\substack{U \text{ UR } V \\ M \subseteq U}} U$$

Satz: Die lineare Hülle einer Menge $M \subseteq V$ ist die Menge aller Linearkombinationen der Vektoren $\vec{v}_i \in M$:

$$\mathrm{Lin}(M) = \{\lambda_1\vec{v}_1 + \ldots \lambda_k\vec{v}_k \mid \lambda_i \in K, \vec{v}_i \in M\}$$

Beweis: Zum Einen bildet die Menge aller Linearkombinationen der Vektoren $\vec{v}_i \in M$ (rechte Seite) einen Unterraum (siehe Lemma). Zum Anderen enthält jeder Unterraum U, der M enthält, auch die Menge aller Linearkombinationen der Vektoren $\vec{v}_i \in M$ (Abgeschlossenheit von Unterräumen bezüglich der Addition und der Multiplikation mit Skalaren). Daraus folgt, dass die Menge aller Linearkombinationen der Vektoren $\vec{v}_i \in M$ der kleinste Unterraum ist, der M enthält.

2.3 Lineare Unabhängigkeit, Basis und Dimension

2.3.1 Lineare Unabhängigkeit

Definition: Eine Menge $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}$ von k Vektoren heißt *linear abhängig* (l.a.), wenn eine Linearkombination existiert, mit

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k = \vec{0}$$

wobei mindestens ein $\lambda_i \neq 0$ ist.

Definition: Eine Menge $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}$ von k Vektoren heißt *linear unabhängig* (l.u.), wenn sie nicht linear abhängig ist.

Satz: Eine Menge $M \subseteq V$ ist linear unabhängig, wenn jede endliche Teilmenge von M linear unabhängig ist.

Folgerungen:

1. Es kann bei *Aufzählungen* von Vektoren zu Mehrfachnennungen kommen (im Gegensatz zu Mengen). In diesem Fall folgt lineare Abhängigkeit.

   Beispiel: Sei $\vec{v}_1, \vec{v}_2, \ldots$ eine Aufzählung und $\vec{v}_5 = \vec{v}_7$, dann ist die Aufzählung linear abhängig, weil
   $$\vec{0} = 1 \cdot \vec{v}_5 + (-1) \cdot \vec{v}_7$$

2. Aus $\vec{0} \in M$ folgt lineare Abhängigkeit, denn

   $$\vec{0} = \lambda \cdot \vec{0} \quad \text{(auch wenn } \lambda \neq 0\text{)}$$

3. Sei M linear unabhängig und $\lambda_1, \lambda_2, \ldots \lambda_k \in K$, dann folgt aus

   $$\forall\ \vec{v}_1, \vec{v}_2, \ldots \vec{v}_k \in M \qquad (i \neq j \ \rightarrow\ \vec{v}_i \neq \vec{v}_j) \qquad \vec{0} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k$$

   dass für alle λ_i gelten muss

   $$\lambda_i = 0 \qquad (i = 1, 2, \ldots k)$$

   Das heißt: Es existiert *keine* nichttriviale Linearkombination von $\vec{0}$.

4. Wenn M linear abhängig ist, dann existiert eine nichttriviale Linearkombination von $\vec{0}$.

Beispiel: Die Vektoren

$$\begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix} \in \mathbb{R}^3$$

sind linear unabhängig, dann ist der Nullvektor $\vec{0}$ ist eine Linearkombination dieser Vektoren:

$$\vec{0} = \begin{pmatrix}0\\0\\0\end{pmatrix} = \lambda_1 \begin{pmatrix}0\\1\\1\end{pmatrix} + \lambda_2 \begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}\lambda_2\\\lambda_1+\lambda_2\\\lambda_1+\lambda_2\end{pmatrix}$$

$$\begin{aligned} 0 &= \lambda_2 \\ 0 &= \lambda_1 + \lambda_2 \\ 0 &= \lambda_1 + \lambda_2 \end{aligned}$$

Daraus folgt, dass $\lambda_1 = \lambda_2 = 0$.

Satz: Für jede Teilmenge $M \subseteq V$ (über dem Körper K) sind die folgenden Aussagen äquivalent:

- Aussage A:

  Die Menge M ist linear unabhängig.

- Aussage B:

  Kein Vektor $\vec{v} \in M$ kann als Linearkombination aus den übrigen Vektoren aus M dargestellt werden.

- Aussage C:

  Jeder Vektor $\vec{v} \in \mathrm{Lin}(M)$ hat *eindeutige* Darstellung als Linearkombination aus M.

Beweis: Der Satz wird nach folgendem Schema gezeigt:

$$\neg\,\mathrm{A} \underset{\text{1. Schritt}}{\Rightarrow} \neg\,\mathrm{B} \underset{\text{2. Schritt}}{\Rightarrow} \neg\,\mathrm{C} \underset{\text{3. Schritt}}{\Rightarrow} \neg\,\mathrm{A}$$

Die drei Aussagen in ihrer Negation:

- Aussage $\neg$ A:

  Es existiert eine nichttriviale Linearkombination von $\vec{0}$.

- Aussage $\neg$ B:

  Es existiert ein Vektor $\vec{v} \in M$, der eine Linearkombination der übrigen Vektoren ist.

- Aussage $\neg$ C:

  Es existiert ein Vektor $\vec{v} \in \text{Lin}(M)$ mit verschiedenen Linearkombinationen aus M.

Beweisschritte:

- Schritt 1:

  Es existiert folgende nichttriviale Linearkombination von $\vec{0}$:

  $$\vec{0} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k$$

  mit $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k \in M$, $\lambda_1, \lambda_2, \ldots \lambda_k \in K$ und $\exists\, \lambda_i \neq 0$.

  Ohne Beschränkung der Allgemeinheit gelte $\lambda_1 \neq 0$. Damit kann die Gleichung nach $\vec{v}_1$ umgeformt werden:

  $$\vec{v}_1 = \left(-\frac{\lambda_2}{\lambda_1}\right)\vec{v}_2 + \left(-\frac{\lambda_3}{\lambda_1}\right)\vec{v}_3 + \ldots + \left(-\frac{\lambda_k}{\lambda_1}\right)\vec{v}_k$$

  Damit existiert ein Vektor $\vec{v}_1$, der Linearkombination der übrigen Vektoren ist:

  $$\vec{v}_1 \in \text{Lin}(M \setminus \{\vec{v}_1\})$$

- Schritt 2:

  Es existiert ein Vektor $\vec{v}_1$, der Linearkombination der übrigen Vektoren ist:

  $$\vec{v}_1 = \lambda_2\vec{v}_2 + \lambda_3\vec{v}_3 + \ldots + \lambda_k\vec{v}_k$$

  Damit existieren mindestens zwei verschiedene Linearkombinationen von $\vec{v_1}$:

  $$\begin{aligned} \vec{v}_1 &= 1 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2 + 0 \cdot \vec{v}_3 + \ldots + 0 \cdot \vec{v}_k \\ &= 0 \cdot \vec{v}_1 + \lambda_2 \cdot \vec{v}_2 + \lambda_3 \cdot \vec{v}_3 + \ldots + \lambda_k \cdot \vec{v}_k \end{aligned}$$

- Schritt 3:

  Es existiert ein Vektor $\vec{v} \in \text{Lin}(M)$ mit verschiedenen Linearkombinationen aus M:

  $$\begin{aligned} \vec{v} &= \lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \ldots + \lambda_m\vec{u}_m \\ &= \mu_1\vec{w}_1 + \mu_2\vec{w}_2 + \ldots + \mu_n\vec{w}_n \end{aligned}$$

  mit

  $$\{\vec{u}_1, \vec{u}_2, \ldots \vec{u}_m\} \cup \{\vec{w}_1, \vec{w}_2, \ldots \vec{w}_n\} = \{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\} \subseteq M$$

Die beiden Linearkombinationen von $\vec{v}$ ausgedrückt als Linearkombinationen der die Vektoren aus $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}$:

$$\begin{aligned} \vec{v} &= \lambda'_1\vec{v}_1 + \lambda'_2\vec{v}_2 + \ldots + \lambda'_k\vec{v}_k \\ &= \mu'_1\vec{v}_1 + \mu'_2\vec{v}_2 + \ldots + \mu'_k\vec{v}_k \end{aligned}$$

mit

$$\lambda'_i = \begin{cases} \lambda_j & \text{falls } \vec{v}_i = \vec{u}_j \\ 0 & \text{sonst} \end{cases} \quad \text{und} \quad \mu'_i = \begin{cases} \mu_j & \text{falls } \vec{v}_i = \vec{w}_j \\ 0 & \text{sonst} \end{cases}$$

Da es sich um verschiedene Linearkombinationen handelt, existiert ein i_0 mit $\lambda'_{i_0} \neq \mu'_{i_0}$. Daraus folgt:

$$\begin{aligned} \vec{0} &= \vec{v} - \vec{v} \\ &= (\lambda'_1\vec{v}_1 + \lambda'_2\vec{v}_2 + \ldots + \lambda'_k\vec{v}_k) - (\mu'_1\vec{v}_1 + \mu'_2\vec{v}_2 + \ldots + \mu'_k\vec{v}_k) \\ &= (\lambda'_1 - \mu'_1)\vec{v}_1 + \ldots + \underbrace{(\lambda'_{i_0} - \mu'_{i_0})}_{\neq 0}\vec{v}_{i_0} + \ldots + (\lambda'_k - \mu'_k)\vec{v}_k \end{aligned}$$

Damit existiert eine nichttriviale Linearkombination von $\vec{0}$. □

- Beispiel zu Schritt 3:

$$\begin{aligned} \begin{pmatrix} 1 \\ 2 \end{pmatrix} &= 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

Die Menge aller Vektoren, aus beiden Linearkombinationen:

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

Die beiden Linearkombinationen mit Hilfe aller Vektoen aus dieser Menge:

$$\begin{aligned} \begin{pmatrix} 1 \\ 2 \end{pmatrix} &= 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= 0 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

Nichttriviale Linearkombination von $\vec{0}$:

$$\begin{aligned}
\begin{pmatrix}0\\0\end{pmatrix} &= \begin{pmatrix}1\\2\end{pmatrix} - \begin{pmatrix}1\\2\end{pmatrix} \\
&= \left[1 \cdot \begin{pmatrix}1\\0\end{pmatrix} + 2 \cdot \begin{pmatrix}0\\1\end{pmatrix} + 0 \cdot \begin{pmatrix}1\\1\end{pmatrix}\right] - \left[0 \cdot \begin{pmatrix}1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\end{pmatrix} + 1 \cdot \begin{pmatrix}1\\1\end{pmatrix}\right] \\
&= (1-0) \cdot \begin{pmatrix}1\\0\end{pmatrix} + (2-1) \cdot \begin{pmatrix}0\\1\end{pmatrix} + (0-1) \cdot \begin{pmatrix}1\\1\end{pmatrix} \\
&= 1 \cdot \begin{pmatrix}1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\end{pmatrix} - 1 \cdot \begin{pmatrix}1\\1\end{pmatrix}
\end{aligned}$$

2.3.2 Erzeugendensystem und Basis

Definition: Eine Teilmenge $M \subseteq V$ heißt *Erzeugendensystem* von V, wenn die lineare Hülle von M der Vektorraum V ist:

$$\mathrm{Lin}(M) = V$$

Definition: Eine Teilmenge $M \subseteq V$ heißt *Basis*, wenn sie Erzeugendensystem von V und linear unabhängig ist.

Folgerung: Eine Teilmenge $M \subseteq V$ ist genau dann eine Basis von V, wenn jeder Vektor $\vec{v} \in V$ eine *eindeutige* Darstellung als Linearkombination aus M hat.

Beispiele:

- kanonische Basis von $\mathbb{R}^n$:

$$\vec{e}_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \quad \ldots \quad \vec{e}_n = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

  Für die kanonische Basis gilt:

$$\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix} = a_1\vec{e}_1 + a_2\vec{e}_2 + \ldots + a_n\vec{e}_n$$

- weitere Basis von $\mathbb{R}^n$:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad \vec{e}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

- Standardbasis für den Vektorraum $\mathbb{R}[x]$ der Polynome:

$$\vec{e}_1 = 1, \quad \vec{e}_2 = x, \quad \vec{e}_3 = x^2, \quad \ldots$$

Folgerung: Für jede Teilmenge $M \subseteq V$ sind die folgenden Bedingungen äquivalent:

- Aussage A:

  Die Menge M ist Basis von V.

- Aussage B:

  Die Menge M ist minimales Erzeugendensystem von V.

- Aussage C:

  Die Menge M ist eine maximale linear unabhängige Menge.

Bemerkung: Die Begriffe „minimal“ und „maximal“ gelten in Bezug auf Inklusion.

Lemma: Ist die Teilmenge $M \subseteq V$ linear unabhängig und der Vektor $\vec{v}$ Element von V, aber nicht Element aus der linearen Hüllen von M, dann ist die Menge $M \cup \{\vec{v}\}$ ebenfalls linear unabhängig:

$$M \subseteq V \text{ l.u.} \quad \wedge \quad \vec{v} \in V \quad \wedge \quad \vec{v} \notin \text{Lin}(M) \quad \Rightarrow \quad M \cup \{\vec{v}\} \text{ l.u.}$$

Beweis (indirekt): Angenommen die Teilmenge $M \subseteq V$ ist linear unabhängig und der Vektor $\vec{v}$ Element von V, aber nicht Element aus der linearen Hüllen von M, und die Menge $M \cup \{\vec{v}\}$ ist linear abhängig. Damit existiert eine nichttriviale Linearkombination von $\vec{0}$ (mit $\exists\, \lambda_i \neq 0 \ \vee \ \lambda \neq 0$):

$$\vec{0} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \ldots + \lambda_k \vec{v}_k + \lambda \vec{v}$$

Wenn $\lambda \neq 0$, dann lässt sich die Gleichung nach $\vec{v}$ umformen:

$$\vec{v} = \left(-\frac{\lambda_1}{\lambda}\right) \vec{v}_1 + \left(-\frac{\lambda_2}{\lambda}\right) \vec{v}_2 + \ldots + \left(-\frac{\lambda_k}{\lambda}\right) \vec{v}_k$$

Damit ist der Vektor $\vec{v}$ in der linearen Hüllen von M:

$$\vec{v} \in \text{Lin}(M)$$

Dies wäre ein Widerspruch zur Annahme. Damit ist $\lambda = 0$. Daraus folgt:

$$\exists\, \lambda_i \neq 0 \quad \vec{0} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k$$

Dies ist ein Widerspruch, da die Menge M linear unabhängig ist und damit keine nichttriviale Linearkombination von $\vec{0}$ existiert. □

Basisergänzungssatz (Steinitz): Seien

- V ein Vektorraum über dem Körper K
- M eine Teilmenge von V
- N eine Teilmenge von V

Ist die Menge M linear unabhängig und die lineare Hülle von $M \cup N$ der Vektorraum V, dann kann man die Menge M durch eventuelle Hinzunahme von Vektoren aus der Menge N zu einer Basis des Vektorraumes V erweitern.

Beweis: Induktion nach $k = |N|$:

- Induktionsanfang ($k = 0$, das heißt $N = \emptyset$):

  Die Menge M ist Basis, weil M linear unabhängig ist und

  $$\text{Lin}(M) = \text{Lin}(M \cup \emptyset) = \text{Lin}(M \cup N) = V$$

- Induktionsschritt ($k - 1 \to k$):

  - Fall 1: $\text{Lin}(M) = V$

    Daraus folgt, dass die Menge M Basis ist.
  - Fall 2: $\text{Lin}(M) \neq V$

    Sei der Vektor $\vec{v}$ Element von N, aber nicht Element aus der linearen Hüllen von M. Damit ist die Menge $M \cup \{\vec{v}\}$ ebenfalls linear unabhängig.

    Die Menge M wird also um den Vektor $\vec{v}$ erweitert, und die Menge N wird um den Vektor $\vec{v}$ reduziert:

    $$|N \setminus \{\vec{v}\}| = k - 1$$

    Nach Induktionsvoraussetzung existiert eine Erweiterung der Menge M zur Basis.

□

Beispiele:

1. Sei M_1 die folgende linear unabhängige Menge und N_1 die kanonische Basis von $\mathbb{R}^3$:
$$M_1 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$
Man kann jeden der drei Vektoren aus N_1 als Basisergänzung wählen.

2. Sei M_2 die folgende linear unabhängige Menge und N_2 die kanonische Basis von $\mathbb{R}^3$:
$$M_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$
Der Vektor $\vec{e}_1$ ist bereits in M_2 enthalten und damit keine Ergänzung von M_2, der Vektor $\vec{e}_2$ ist auch keine Basisergänzung, weil $M_2 \cup \{\vec{e}_2\}$ linear abhängig wäre, doch der dritte Vektor $\vec{e}_3$ ergänzt die Menge M_2 zu einer Basis.

Austauschlemma: Sind die Mengen $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$ und $\{\vec{w}_1, \vec{w}_2, \ldots \vec{w}_m\}$ Basen von V, dann gibt es für jeden Vektor $\vec{v}_i$ einen Vektor $\vec{w}_j$, so dass die Menge

$$(\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\} \setminus \{\vec{v}_i\}) \cup \{\vec{w}_j\}$$

ebenfalls Basis von V ist.

2.3.3 Dimension

Definition: Besitzt ein Vektorraum V eine endliche Basis $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$, dann ist V *endlich-dimensional* und n heißt die *Dimension* von V:

$$\dim V = n$$

Ein Vektorraum, der keine endliche Basis besitzt, ist *unendlich-dimensional*:

$$\dim V = \infty$$

Folgerung: Ist die Menge $M = \{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}$ Teilmenge eines Vektorraums V und ist $k > \dim V$, so ist M linear abhängig.

Satz: Jeder Vektorraum besitzt eine Basis.

Satz: Ist die Dimension eines Vektorraumes V endlich und U ein Unterraum von V, dann gilt:

$$\dim U \leq \dim V$$

und

$$\dim U < \dim V \quad \Leftrightarrow \quad U \neq V$$

Definition: Sind U_1 und U_2 Unterräume von V, so heißt

$$U_1 + U_2 = \{\vec{x} + \vec{y} \mid \vec{x} \in U_1, \vec{y} \in U_2\}$$

die Summe von U_1 und U_2.

Beispiele:

1. Sei $V = \mathbb{R}^4$:

$$\begin{aligned} U_1 &= \text{Lin}(\{\vec{e}_1, \vec{e}_2, \vec{e}_4\}) \\ U_2 &= \text{Lin}(\{\vec{e}_1, \vec{e}_3, \vec{e}_4\}) \end{aligned}$$

$$U_1 + U_2 = \mathbb{R}^4$$

2. Sei $V = \mathbb{R}^3$:

$$\begin{aligned} U_1 &= \text{Lin}\left(\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right\}\right) \\ U_2 &= \text{Lin}\left(\left\{\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}\right\}\right) \\ U_1 + U_2 &= \text{Lin}\left(\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}\right\}\right) = \text{Lin}\left(\left\{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}\right) \end{aligned}$$

   Bemerkung: $U_1 + U_2$ ist die Ebene senkrecht zu xy-Ebene auf der Geraden $y = -x$ durch den Ursprung $(0, 0, 0)$.

Satz: Die Summe von zwei Unterräumen ist ein Unterraum. Für zwei endlichdimensionale Unterräume U_1 und U_2 gilt:

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Beweisidee:

- Die Basis von $U_1 \cap U_2$ sei:

$$\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_r\}$$

- Ergänzung der Basis von $U_1 \cap U_2$ zur Basis von U_1:

$$\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_r, \vec{u}_1, \vec{u}_2, \dots \vec{u}_s\}$$

- Ergänzung der Basis von $U_1 \cap U_2$ zur Basis von U_2:

$$\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_r, \vec{w}_1, \vec{w}_2, \dots \vec{w}_t\}$$

- Man zeigt: Die Basis von $U_1 + U_2$ ist:

$$\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_r, \vec{u}_1, \vec{u}_2, \dots \vec{u}_s, \vec{w}_1, \vec{w}_2, \dots \vec{w}_t\}$$

- Damit gilt für die Dimenstionen:
  - $\dim(U_1 + U_2) = r + s + t$
  - $\dim U_1 = r + s$
  - $\dim U_2 = r + t$
  - $\dim(U_1 \cap U_2) = r$

  Daraus folgt:

$$\underbrace{\dim(U_1 + U_2)}_{r+s+t} = \underbrace{\dim U_1}_{r+s} + \underbrace{\dim U_2}_{r+t} - \underbrace{\dim(U_1 \cap U_2)}_{r}$$

Beispiele:

1. Sei U_1 eine Ebene durch den Ursprung $(0, 0, 0)$ und U_2 eine Gerade durch den Ursprung $(0, 0, 0)$ mit $U_2 \not\subset U_1$. Daraus heißt:
   - $\dim U_1 = 2$
   - $\dim U_2 = 1$
   - $\dim(U_1 \cap U_2) = 0$ (weil $U_1 \cap U_2 = \vec{0}$)

   Daraus folgt:
   - $\dim(U_1 + U_2) = 3$

   Damit ist $U_1 + U_2 = \mathbb{R}^3$.

2. Seien U_1 und U_2 Ebenen durch den Ursprung $(0,0,0)$ mit $U_1 \neq U_2$. Daraus heißt:

   - $\dim U_1 = 2$
   - $\dim U_2 = 2$
   - $\dim(U_1 \cap U_2) = 1$ (weil $U_1 \cap U_2$ eine Gerade ist)

   Daraus folgt:

   - $\dim(U_1 + U_2) = 3$

   Damit ist $U_1 + U_2 = \mathbb{R}^3$.

2.4 Lineare Abbildungen

2.4.1 Einleitung

Definition: Seien V und W Vektorräume über dem Körper K. Eine Abbildung $f : V \to W$ heißt *linear (Vektorraumhomomorphismus)*, wenn für alle $\vec{v}, \vec{w} \in V$ und für alle $\lambda \in K$ gilt:

$$\begin{aligned} f(\vec{v} + \vec{w}) &= f(\vec{v}) + f(\vec{w}) \\ f(\lambda \cdot \vec{v}) &= \lambda \cdot f(\vec{v}) \end{aligned}$$

$\mathrm{Hom}(V, W)$ bezeichnet die Menge aller linearer Abbildungen $f : V \to W$.

Beobachtungen:

- Sei $f \in \mathrm{Hom}(V, W)$, dann gilt:

$$f(\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k) = \lambda_1 \cdot f(\vec{v}_1) + \lambda_2 \cdot f(\vec{v}_2) + \ldots + \lambda_k \cdot f(\vec{v}_k)$$

- Die Verknüpfung von linearen Abbildungen $f : V \to W$ und $g : W \to Y$ ist eine lineare Abbildung $gf : V \to Y$ mit

$$gf(\vec{v}) = g(f(\vec{v}))$$

- Die Menge aller linearen Abbildungen $\mathrm{Hom}(V, W)$ ist selbst ein Vektorraum mit den Operationen:
  - $(f + g)(\vec{v}) = f(\vec{v}) + g(\vec{v})$
  - $(\lambda \cdot f)(\vec{v}) = \lambda \cdot f(\vec{v})$

Denn für alle $f, g \in \mathrm{Hom}(V, W)$, $\vec{u}, \vec{v} \in V$ und $\lambda \in K$:

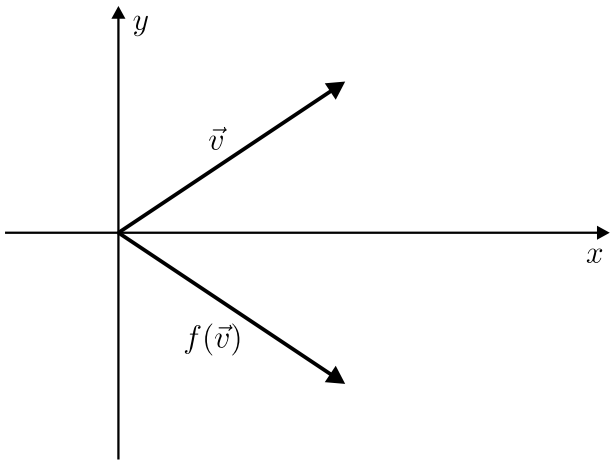
$$\begin{aligned}
(f+g)(\vec{u}) + (f+g)(\vec{v}) &= f(\vec{u}) + g(\vec{u}) + f(\vec{v}) + g(\vec{v}) \\
&= f(\vec{u}) + f(\vec{u}) + g(\vec{v}) + g(\vec{v}) \\
&= f(\vec{u}+\vec{v}) + g(\vec{u}+\vec{v}) \\
&= (f+g)(\vec{u}+\vec{v}) \\
&\\
&\rightarrow \quad f+g \in \mathrm{Hom}(V,W)
\end{aligned}$$

$$\begin{aligned}
\lambda \cdot (f+g)(\vec{u}) &= \lambda \cdot (f(\vec{u}) + g(\vec{u})) \\
&= \lambda \cdot f(\vec{u}) + \lambda \cdot g(\vec{u}) \\
&= f(\lambda\vec{u}) + g(\lambda\vec{u}) \\
&= (f+g)(\lambda\vec{u}) \\
&\\
&\rightarrow \quad \lambda \cdot f \in \mathrm{Hom}(V,W)
\end{aligned}$$

Beispiele: $f, g, h, j : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

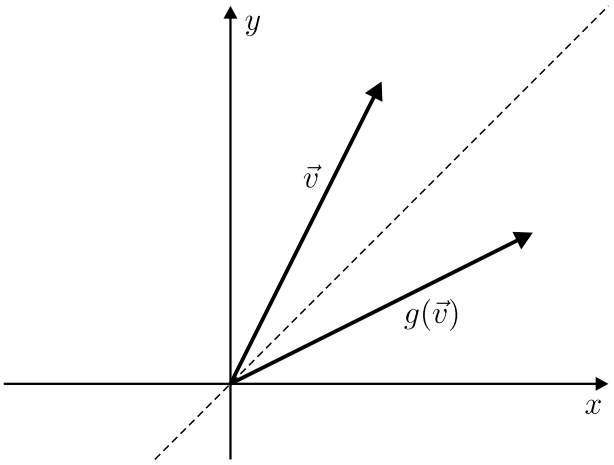
a) Spiegelung an der x-Achse :

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ -y \end{pmatrix}$$



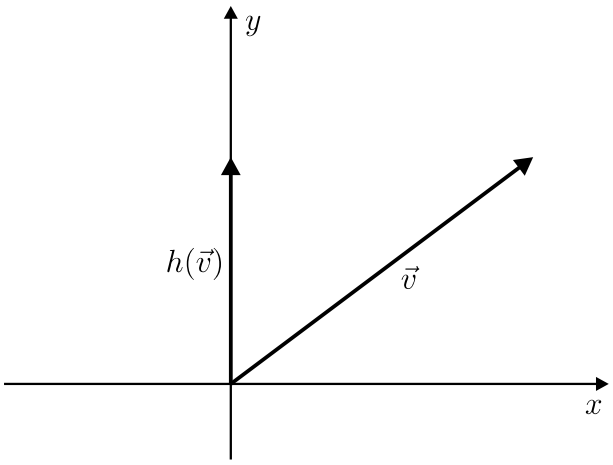
b) Spiegelung an der Geraden $y = x$:

$$g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix}$$



c) Projektion auf die y-Achse:

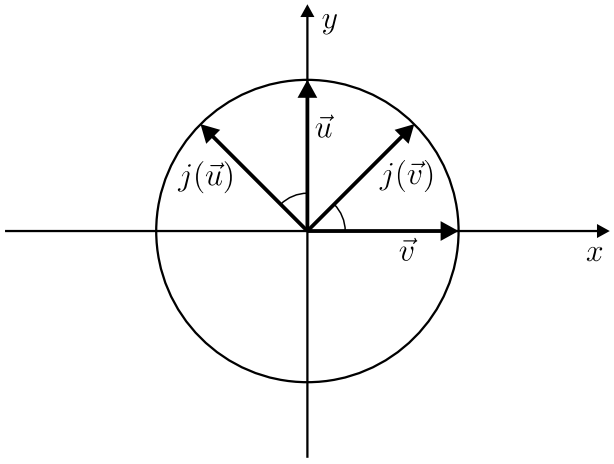
$$h\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 0 \\ y \end{pmatrix}$$



d) Drehung um 45°:

$$
\begin{aligned}
j\left(\begin{pmatrix}1\\0\end{pmatrix}\right) &= \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} \\
j\left(\begin{pmatrix}0\\1\end{pmatrix}\right) &= \begin{pmatrix}-\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
j\left(\begin{pmatrix}x\\y\end{pmatrix}\right) &= j\left(x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix}\right) \\
&= x \cdot j\left(\begin{pmatrix}1\\0\end{pmatrix}\right) + y \cdot j\left(\begin{pmatrix}0\\1\end{pmatrix}\right) \\
&= x \cdot \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} + y \cdot \begin{pmatrix}-\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} \\
&= \begin{pmatrix}\frac{1}{\sqrt{2}}\,x - \frac{1}{\sqrt{2}}\,y\\ \frac{1}{\sqrt{2}}\,x + \frac{1}{\sqrt{2}}\,y\end{pmatrix}
\end{aligned}
$$



2.4.2 Kern und Bild von linearen Abbildungen

Definition: Sei $f \in \mathrm{Hom}(V,W)$ eine lineare Abbildung, so ist ihr *Kern* (Ker f) und ihr *Bild* (Im f) folgendermaßen definiert:

$$\begin{aligned} \text{Ker } f &= \{\vec{v} \in V \mid f(\vec{v}) = \vec{0}\} \\ \text{Im } f &= \{\vec{w} \in W \mid \exists\, \vec{v}\ \ f(\vec{v}) = \vec{w}\} \end{aligned}$$

Beispiele: Kerne und Bilder aus dem obigen Beispiel (siehe 2.4.1)

a) Ker $f = \{\vec{0}\}$:

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad x = 0 \quad \wedge \quad y = 0$$

Im $f = \mathbb{R}^2$:

- f ist eine Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$, und zwar $f^{-1} = f$.

b) Ker $g = \{\vec{0}\}$:

$$g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad x = 0 \quad \wedge \quad y = 0$$

Im $g = \mathbb{R}^2$:

- g ist eine Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$, und zwar $g^{-1} = g$.

c) Ker $h = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$:

$$h\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad y = 0$$

Im $h = \mathrm{Lin}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$:

- die x-Komponente aller Elemente aus dem Bild ist 0.

d) Ker $j = \{\vec{0}\}$:

$$j\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{\sqrt{2}}\, x - \frac{1}{\sqrt{2}}\, y \\ \frac{1}{\sqrt{2}}\, x + \frac{1}{\sqrt{2}}\, y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad x = 0 \quad \wedge \quad y = 0$$

Im $f = \mathbb{R}^2$:

- j ist eine Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$ und j ist bijektiv, so dass folgende Umkehrabbildung existiert:

$$j^{-1}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{\sqrt{2}}\, x + \frac{1}{\sqrt{2}}\, y \\ -\frac{1}{\sqrt{2}}\, x + \frac{1}{\sqrt{2}}\, y \end{pmatrix}$$

Lemma: Der Kern und das Bild einer linearen Abbildung $f \in \mathrm{Hom}(V, W)$ sind Unterräume von V bzw. W:

$$\begin{array}{rll} \mathrm{Ker}\ f & \mathrm{UR} & V \\ \mathrm{Im}\ f & \mathrm{UR} & W \end{array}$$

Beweis (Kern): Seien $\vec{u}, \vec{v} \in \mathrm{Ker}\ f$ und $\lambda \in K$ (Körper zu V).

- Prüfe, ob Ker f mindestens ein Element enthält:

$$\begin{aligned} f(\vec{0}) &= f(\vec{0} - \vec{0}) \\ &= f(\vec{0}) - f(\vec{0}) \\ &= \vec{0} \end{aligned}$$

  Damit ist $\vec{0} \in \mathrm{Ker}\ f$.

- Prüfe Abgeschlossenheit gegenüber der Addition:

$$\begin{aligned} f(\vec{u} + \vec{v}) &= f(\vec{u}) + f(\vec{v}) \\ &= \vec{0} + \vec{0} \quad (\text{da } \vec{u}, \vec{v} \in \mathrm{Ker}\ f) \\ &= \vec{0} \end{aligned}$$

  Damit ist auch $\vec{u} + \vec{v} \in \mathrm{Ker}\ f$.

- Prüfe Abgeschlossenheit gegenüber der Multiplikation mit Skalaren:

$$\begin{aligned} f(\lambda\vec{u}) &= \lambda \cdot f(\vec{u}) \\ &= \lambda \cdot \vec{0} \quad (\text{da } \vec{u} \in \mathrm{Ker}\ f) \\ &= \vec{0} \end{aligned}$$

  Damit ist auch $\lambda\vec{u} \in \mathrm{Ker}\ f$. □

Beweis (Bild): Seien $\vec{u}, \vec{v} \in \mathrm{Im}\ f$ und $\lambda \in K$ (Körper zu W).

- Prüfe, ob Im f mindestens ein Element enthält:

$$f(\vec{0}) = \vec{0} \quad (\text{siehe oben})$$

  Damit ist $\vec{0} \in \mathrm{Im}\ f$.

- Prüfe Abgeschlossenheit gegenüber der Addition:

$$\begin{aligned} \vec{u} + \vec{v} &= f(\vec{p}) + f(\vec{q}) \quad (\text{mit } f(\vec{p}) = \vec{u} \text{ und } f(\vec{q}) = \vec{v}) \\ &= f(\vec{p} + \vec{q}) \\ &= f(\vec{r}) \quad (\text{mit } \vec{r} = \vec{p} + \vec{q} \in V) \end{aligned}$$

  Damit ist auch $\vec{u} + \vec{v} \in \mathrm{Im}\ f$.

- Prüfe Abgeschlossenheit gegenüber der Multiplikation mit Skalaren:

$$\begin{aligned} \lambda\vec{u} &= \lambda \cdot f(\vec{p}) \quad (\text{mit } f(\vec{p}) = \vec{u}) \\ &= f(\lambda\vec{p}) \\ &= f(\vec{r}) \quad (\text{mit } \vec{r} = \lambda\vec{p} \in V) \end{aligned}$$

  Damit ist auch $\lambda\vec{u} \in \text{Im } f$. □

Lemma: Eine lineare Abbildung $f \in \text{Hom}(V, W)$ ist genau dann injektiv, wenn ihr Kern nur aus dem Nullvektor besteht:

$$\text{Ker } f = \{\vec{0}\}$$

Beweis (⇒): Da $f(\vec{0}) = \vec{0}$ und f injektiv ist, bildet kein anderer Vektor auf $\vec{0}$ ab. Damit liegt außer dem Nullvektor kein anderer Vektor im Ker f. □

Beweis durch Widerspruch (⇐): Angenommen Ker $f = \{\vec{0}\}$ und f ist *nicht* injektiv, dann existieren zwei Vektoren $\vec{u}, \vec{v} \in V$, so dass

$$\vec{u} \neq \vec{v} \quad \wedge \quad f(\vec{u}) = f(\vec{v})$$

Daraus folgt:

$$f(\vec{u} - \vec{v}) = f(\vec{u}) - f(\vec{w}) = \vec{0}$$

Damit liegt $\vec{u} - \vec{v} \neq \vec{0}$ in Ker f. Dies ist ein Widerspruch, da Ker $f = \{\vec{0}\}$. □

2.4.3 Spezielle Homomorphismen

Definitionen: Einen Homomorphismus $f \in \text{Hom}(V, W)$ nennt man einen

- *Monomorphismus*, wenn f injektiv ist,
- *Epimorphismus*, wenn f surjektiv ist,
- ***Isomorphismus***, wenn f bijektiv ist,
- *Endomorphismus*, wenn $V = W$,
- ***Automorphismus***, wenn $V = W$ und f bijektiv ist.

Satz: Ist $f \in \text{Hom}(V, W)$ ein Isomorphismus, dann ist auch $f^{-1} \in \text{Hom}(W, V)$ ein Isomorphismus.

Satz: Die Verkettung von Isomorphismen ist auch wieder ein Isomorphismus.

Satz: Seien

- V, W Vektorräume über K,
- die Menge $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\} \subseteq V$ eine Basis von V und
- $\vec{w}_1, \vec{w}_2, \ldots \vec{w}_n \in W$ beliebig,

dann gibt es eine *eindeutige* lineare Abbildung $f \in \mathrm{Hom}(V, W)$ definiert durch

$$f(\vec{v}_i) = \vec{w}_i \quad \text{für} \quad i = 1, 2, \ldots n$$

Beweis: Jeder Vektor $\vec{v} \in V$ hat eine eindeutige Darstellung als Linearkombination aus $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$:

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \ldots + \lambda_n \vec{v}_n$$

- Zu zeigen ist, dass eine entsprechende lineare Abbildung existiert. Dazu wird die Abbildung des Vektors $\vec{v}$ folgendermaßen definiert:

  $$\begin{aligned} f(\vec{v}) &= \lambda_1 \vec{w}_1 + \lambda_2 \vec{w}_2 + \ldots + \lambda_n \vec{w}_n \\ &= \lambda_1 \cdot f(\vec{v}_1) + \lambda_2 \cdot f(\vec{v}_2) + \ldots + \lambda_n \cdot f(\vec{v}_n) \end{aligned}$$

  Außerdem gilt:

  $$f(\vec{v}) = f(\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \ldots + \lambda_n \vec{v}_n)$$

  Daraus folgt:

  $$f(\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \ldots + \lambda_n \vec{v}_n) = \lambda_1 \cdot f(\vec{v}_1) + \lambda_2 \cdot f(\vec{v}_2) + \ldots + \lambda_n \cdot f(\vec{v}_n)$$

  Damit ist f eine lineare Abbildung.

- Außerdem ist zu zeigen, dass f eine eindeutige lineare Abbildung ist:

  Angenommen es existiert eine lineare Abbildung $g \neq f$ mit $g(\vec{v}_i) = \vec{w}_i$. Damit gilt:

  $$\begin{aligned} g(\vec{v}) &= g(\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \ldots + \lambda_n \vec{v}_n) \\ &= \lambda_1 \cdot g(\vec{v}_1) + \lambda_2 \cdot g(\vec{v}_2) + \ldots + \lambda_n \cdot g(\vec{v}_n) \\ &= \lambda_1 \vec{w}_1 + \lambda_2 \vec{w}_2 + \ldots + \lambda_n \vec{w}_n \\ &= f(\vec{v}) \end{aligned}$$

  Das heißt, für alle $\vec{v} \in V$ gilt $g(\vec{v}) = f(\vec{v})$. Damit ist $g = f$. Dies ein Widerspruch zur Annahme. □

Folgerung: Zu zwei n-dimentionalen Vektorräumen existiert mindestens ein Isomorphismus, der den einen Vektorraum in den anderen überführt.

2.4.4 Rang einer linearen Abbildung

Definition: Der *Rang* einer linearen Abbildung $f \in \text{Hom}(V, W)$ ist die Dimension des Bildes von f:

$$\text{rg } f = \dim(\text{Im f})$$

Satz (Dimensionsformel für lineare Abbildungen): Für jede lineare Abbildung $f \in \text{Hom}(V, W)$ gilt:

$$\begin{array}{lcccl} \dim(\text{Ker } f) & + & \dim(\text{Im } f) & = & \dim V \\ \dim(\text{Ker } f) & + & \text{rg } f & = & \dim V \end{array}$$

Beweis: Sei $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}$ eine Basis von Ker $f \subseteq V$:

$$\text{Ker } f = \text{Lin}(\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}) \text{ und } \{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\} \text{ ist l.u.}$$

Diese Basis wird durch die Vektoren $\{\vec{v}_{k+1}, \vec{v}_{k+2}, \ldots \vec{v}_n\}$ zu einer Basis von V erweitert:

$$V = \text{Lin}(\{\vec{v}_1, \ldots \vec{v}_k, \vec{v}_{k+1}, \ldots \vec{v}_n\}) \text{ und } \{\vec{v}_1, \ldots \vec{v}_k, \vec{v}_{k+1}, \ldots \vec{v}_n\} \text{ ist l.u.}$$

Zu zeigen ist, dass $\{f(\vec{v}_{k+1}), f(\vec{v}_{k+2}), \ldots f(\vec{v}_n)\}$ eine Basis von Im f ist.

- Angenommen $\vec{w} \in \text{Im } f$:

  $$\begin{array}{rcl} \vec{w} & = & f(\vec{v}) \\ & = & f(\lambda_1 \cdot \vec{v}_1 + \ldots + \lambda_k \cdot \vec{v}_k + \lambda_{k+1} \cdot \vec{v}_{k+1} + \ldots + \lambda_n \cdot \vec{v}_n) \\ & = & \lambda_1 \cdot f(\vec{v}_1) + \ldots + \lambda_k \cdot f(\vec{v}_k) + \lambda_{k+1} \cdot f(\vec{v}_{k+1}) + \ldots + \lambda_n \cdot f(\vec{v}_n) \end{array}$$

  Da $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k \in \text{Ker } f$, gilt $f(\vec{v}_1) = f(\vec{v}_2) = \ldots = f(\vec{v}_k) = \vec{0}$. Daraus folgt:

  $$\vec{w} = \lambda_{k+1} \cdot f(\vec{v}_{k+1}) + \lambda_{k+2} \cdot f(\vec{v}_{k+2}) + \ldots + \lambda_n \cdot f(\vec{v}_n)$$

  Damit ist $\{f(\vec{v}_{k+1}), f(\vec{v}_{k+2}), \ldots f(\vec{v}_n)\}$ Erzeugendensystem von Im f.

- Der Nullvektor $\vec{0}$ sei eine Linearkombination dieses Erzeugendensystems:

  $$\begin{array}{rcl} \vec{0} & = & \lambda_{k+1} \cdot f(\vec{v}_{k+1}) + \lambda_{k+2} \cdot f(\vec{v}_{k+2}) + \ldots + \lambda_n \cdot f(\vec{v}_n) \\ & = & f(\lambda_{k+1} \cdot \vec{v}_{k+1} + \lambda_{k+2} \cdot \vec{v}_{k+2} + \ldots + \lambda_n \cdot \vec{v}_n) \\ & = & f(\vec{u}) \end{array}$$

  Daraus folgt, dass $\vec{u} \in \text{Ker } f$. Da $\vec{u}$ eine eindeutige Darstellung bezüglich der Basis $\{\vec{v}_1, \ldots \vec{v}_k, \vec{v}_{k+1}, \ldots \vec{v}_n\}$ hat und bereits mit der Basis $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k\}$ darstellbar ist, gilt:

  $$\lambda_{k+1} = \lambda_{k+2} = \ldots = \lambda_n = 0$$

Daraus folgt:

- Die Dimension des Kern von f beträgt k:

$$\dim(\text{Ker } f) = k$$

- Die Dimension des Bildes von f beträgt $n - k$:

$$\dim(\text{Im } f) = n - k$$

- Die Dimension von V beträgt n:

$$\dim V = n$$

Damit gilt:

$$\begin{array}{rccccc} & k & + & n-k & = & n \\ \Leftrightarrow & \dim(\text{Ker } f) & + & \dim(\text{Im } f) & = & \dim V \end{array}$$

2.5 Matrizen

2.5.1 Einleitung

Definition: Eine $m \times n$-*Matrix* über K ist eine Anordnung von $m \times n$ Elementen aus K nach dem folgenden Schema:

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \cdots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m\,1} & a_{m\,2} & \cdots & a_{m\,n} \end{pmatrix}$$

Alternative Schreibweise:

$$A = (a_{i\,j})_{(i,j) \in m \times n}$$

wobei $a_{i\,j}$ die Einträge (Koeffizieten) der Matrix sind.

Definition: Die Menge aller $m \times n$-Matrizen über K wird mit $M(m \times n, K)$ bezeichnet.

Beobachtung: Die Menge $M(m \times n, K)$ ist ein Vektorraum mit den folgenden Operationen ($\lambda \in K$):

$$\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ a_{2\,1} & \cdots & a_{2\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,n} \end{pmatrix} + \begin{pmatrix} b_{1\,1} & \cdots & b_{1\,n} \\ b_{2\,1} & \cdots & b_{2\,n} \\ \vdots & \ddots & \vdots \\ b_{m\,1} & \cdots & b_{m\,n} \end{pmatrix} = \begin{pmatrix} a_{1\,1} + b_{1\,1} & \cdots & a_{1\,n} + b_{1\,n} \\ a_{2\,1} + b_{2\,1} & \cdots & a_{2\,n} + b_{2\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} + b_{m\,1} & \cdots & a_{m\,n} + b_{m\,n} \end{pmatrix}$$

$$\lambda \cdot \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ a_{2\,1} & \cdots & a_{2\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,n} \end{pmatrix} = \begin{pmatrix} \lambda \cdot a_{1\,1} & \cdots & \lambda \cdot a_{1\,n} \\ \lambda \cdot a_{2\,1} & \cdots & \lambda \cdot a_{2\,n} \\ \vdots & \ddots & \vdots \\ \lambda \cdot a_{m\,1} & \cdots & \lambda \cdot a_{m\,n} \end{pmatrix}$$

2.5.2 Multiplikation von Matrizen

Definition: Seien A und B Matrizen folgender Gestalt

- $A = (a_{i\,j})_{(i,j) \in p \times q} \in M(p \times \mathbf{q}, K)$ und
- $B = (b_{i\,j})_{(i,j) \in q \times r} \in M(\mathbf{q} \times r, K)$,

dann ist $C = A \cdot B = (c_{i\,j})_{(i,j) \in p \times r}$ definiert durch:

$$\begin{aligned} c_{i\,j} &= a_{i\,1} \cdot b_{1\,j} + a_{i\,2} \cdot b_{2\,j} + \ldots + a_{i\,q} \cdot b_{q\,j} \\ &= \sum_{k=1}^{q} a_{i\,k} \cdot b_{k\,j} \end{aligned}$$

Regel: „Zeile × Spalte“

Satz: Die Multiplikation von Matrizen ist assoziativ:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

Achtung: Die Multiplikation von Matrizen ist *nicht* kommutativ:

$$A \cdot B \neq B \cdot A$$

2.5.3 Lineare Abbildungen

Definition: Sei

- $f \in \mathrm{Hom}(V, W)$ eine lineare Abbildung,
- die Menge $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$ Basis von V und
- die Menge $\{\vec{w}_1, \vec{w}_2, \ldots \vec{w}_m\}$ Basis von W,

dann wird der Abbildung f eine Matrix $A \in M(m \times n, K)$ zugeordnet durch Darstellung der Bilder der Basisvektoren $f(\vec{v}_i)$ in der Basis $\{\vec{w}_1, \vec{w}_2, \ldots \vec{w}_m\}$ mit

$$\begin{aligned} f(\vec{v}_1) &= a_{1\,1}\ \vec{w}_1 + a_{2\,1}\ \vec{w}_2 + \ldots + a_{m\,1}\ \vec{w}_m \\ f(\vec{v}_2) &= a_{1\,2}\ \vec{w}_1 + a_{2\,2}\ \vec{w}_2 + \ldots + a_{m\,2}\ \vec{w}_m \\ &\vdots \\ f(\vec{v}_n) &= a_{1\,n}\ \vec{w}_1 + a_{2\,n}\ \vec{w}_2 + \ldots + a_{m\,n}\ \vec{w}_m \end{aligned}$$

Umgekehrt bestimmt jede Matrix $A \in M(m \times n, K)$ eine Abbildung f, durch die oberen Formeln.

Regel: Die j-te Spalte der Matrix A stellt $f(\vec{v}_j)$ dar.

Folgerung: Die Vektorräume der linearen Abbildungen $\mathrm{Hom}(V, W)$ und der Matrizen $M(m \times n, K)$ sind isomorph ($n = \mathrm{rg}\ V$ und $m = \mathrm{rg}\ W$).

Festlegung: Für den Vektorraum $V = K^n$ wird die Standardbasis verwendet:

$$\vec{e}_1^{\,(n)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2^{\,(n)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_n^{\,(n)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Entsprechendes gilt für $W = K^m$.

Regel: Die j-te Spalte von der Matrix A entspricht im Folgenden $f\left(\vec{e}_j^{\,(n)}\right)$.

Beobachtung: Sei $A \in M(m \times n, K)$ die zur Abbidlung $f \in \mathrm{Hom}(K^n, K^m)$ zugehörige Matrix wobei für K^n die Standardbasis verwendet wird. Wird zudem ein Vektor mit k Koeffizienten als $k \times 1$-Matrix aufgefasst, dann gilt:

$$A \cdot \vec{v} = f(\vec{v})$$

$$\begin{aligned}
A \cdot \vec{v} &= \begin{pmatrix} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \cdots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m\,1} & a_{m\,2} & \cdots & a_{m\,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} \\
&= \begin{pmatrix} a_{1\,1} \cdot x_1 + a_{1\,2} \cdot x_2 + \ldots + a_{1\,n} \cdot x_n \\ a_{2\,1} \cdot x_1 + a_{2\,2} \cdot x_2 + \ldots + a_{2\,n} \cdot x_n \\ \vdots \\ a_{m\,1} \cdot x_1 + a_{m\,2} \cdot x_2 + \ldots + a_{m\,n} \cdot x_n \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
f(\vec{v}) &= f\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) \\
&= f\left(x_1 \cdot \vec{e}_1^{\,(n)} + x_2 \cdot \vec{e}_2^{\,(n)} + \ldots + x_n \cdot \vec{e}_n^{\,(n)}\right) \\
&= x_1 \cdot f\left(\vec{e}_1^{\,(n)}\right) + x_2 \cdot f\left(\vec{e}_2^{\,(n)}\right) + \ldots + x_n \cdot f\left(\vec{e}_n^{\,(n)}\right) \\
&= x_1 \cdot \begin{pmatrix} a_{1\,1} \\ a_{2\,1} \\ \vdots \\ a_{m\,1} \end{pmatrix} + x_2 \cdot \begin{pmatrix} a_{1\,2} \\ a_{2\,2} \\ \vdots \\ a_{m\,2} \end{pmatrix} + \ldots + x_n \cdot \begin{pmatrix} a_{1\,n} \\ a_{2\,n} \\ \vdots \\ a_{m\,n} \end{pmatrix} \\
&= \begin{pmatrix} a_{1\,1} \cdot x_1 + a_{1\,2} \cdot x_2 + \ldots + a_{1\,n} \cdot x_n \\ a_{2\,1} \cdot x_1 + a_{2\,2} \cdot x_2 + \ldots + a_{2\,n} \cdot x_n \\ \vdots \\ a_{m\,1} \cdot x_1 + a_{m\,2} \cdot x_2 + \ldots + a_{m\,n} \cdot x_n \end{pmatrix}
\end{aligned}$$

$$\Rightarrow \quad A \cdot \vec{v} = f(\vec{v})$$

Satz: Sind $f \in \mathrm{Hom}(K^p, K^q)$ und $g \in \mathrm{Hom}(K^q, K^r)$ lineare Abbildungen und $A \in M(p \times q, K)$ und $B \in M(q \times r, K)$ die zu f und g gehörigen Matrizen bezüglich der Standardbasen von K^p bzw. K^q, dann entspricht das Produkt der Matrizen $A \cdot B$ der Verkettung der Abbildungen fg:

$$C = A \cdot B \in M(p \times r, K) \quad \longleftrightarrow \quad fg \in \mathrm{Hom}(K^p, K^r)$$

Beweis: Für alle Basisvektoren $\vec{e}_k^{\,(p)} \in K^p$ mit $k = 1, 2, \ldots p$ gilt:

$$\begin{aligned}
(gf)\left(\vec{e}_k^{\,(q)}\right) &= f\left(g\left(\vec{e}_k^{\,(q)}\right)\right) \\
&= f\left(\begin{pmatrix} b_{1\,k} \\ b_{2\,k} \\ \vdots \\ b_{r\,k} \end{pmatrix}\right) \\
&= f\left(b_{1\,k} \cdot \vec{e}_1^{\,(r)} + b_{2\,k} \cdot \vec{e}_2^{\,(r)} + \ldots + b_{r\,k} \cdot \vec{e}_r^{\,(r)}\right) \\
&= b_{1\,k} \cdot f\left(\vec{e}_1^{\,(r)}\right) + b_{2\,k} \cdot f\left(\vec{e}_2^{\,(r)}\right) + \ldots + b_{r\,k} \cdot f\left(\vec{e}_r^{\,(r)}\right) \\
&= b_{1\,k} \cdot \begin{pmatrix} a_{1\,1} \\ a_{2\,1} \\ \vdots \\ a_{q\,1} \end{pmatrix} + b_{2\,k} \cdot \begin{pmatrix} a_{1\,2} \\ a_{2\,2} \\ \vdots \\ a_{q\,2} \end{pmatrix} + \ldots + b_{q\,k} \cdot \begin{pmatrix} a_{1\,r} \\ a_{2\,r} \\ \vdots \\ a_{q\,r} \end{pmatrix} \\
&= \begin{pmatrix} a_{1\,1} \cdot b_{1\,k} + a_{1\,2} \cdot b_{2\,k} + \ldots + a_{1\,r} \cdot b_{q\,k} \\ a_{2\,1} \cdot b_{1\,k} + a_{2\,2} \cdot b_{2\,k} + \ldots + a_{2\,r} \cdot b_{q\,k} \\ \vdots \\ a_{q\,1} \cdot b_{1\,k} + a_{q\,2} \cdot b_{2\,k} + \ldots + a_{q\,r} \cdot b_{q\,k} \end{pmatrix}
\end{aligned}$$

$$A \cdot B \cdot \vec{e}_k^{(q)} = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,q} \\ a_{2\,1} & a_{2\,1} & \cdots & a_{2\,q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p\,1} & a_{p\,1} & \cdots & a_{p\,q} \end{pmatrix} \cdot \begin{pmatrix} b_{1\,1} & b_{1\,2} & \cdots & b_{1\,r} \\ b_{2\,1} & b_{2\,1} & \cdots & b_{2\,r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q\,1} & b_{q\,1} & \cdots & b_{q\,r} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k\text{-te Zeile}$$

$$= \begin{pmatrix} \cdots & a_{1\,1} \cdot b_{1\,k} + a_{1\,2} \cdot b_{2\,k} + \ldots + a_{q\,1} \cdot b_{q\,k} & \cdots \\ \cdots & a_{2\,1} \cdot b_{1\,k} + a_{2\,2} \cdot b_{2\,k} + \ldots + a_{2\,r} \cdot b_{q\,k} & \cdots \\ & \vdots & \\ \cdots & a_{q\,1} \cdot b_{1\,k} + a_{q\,2} \cdot b_{2\,k} + \ldots + a_{q\,r} \cdot b_{q\,k} & \cdots \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

$\uparrow$
k-te Spalte

$$= \begin{pmatrix} a_{1\,1} \cdot b_{1\,k} + a_{1\,2} \cdot b_{2\,k} + \ldots + a_{1\,r} \cdot b_{q\,k} \\ a_{2\,1} \cdot b_{1\,k} + a_{2\,2} \cdot b_{2\,k} + \ldots + a_{2\,r} \cdot b_{q\,k} \\ \vdots \\ a_{q\,1} \cdot b_{1\,k} + a_{q\,2} \cdot b_{2\,k} + \ldots + a_{q\,r} \cdot b_{q\,k} \end{pmatrix}$$

Daraus folgt:

$$\forall k \quad (gf)\left(\vec{e}_k^{(q)}\right) = A \cdot B \cdot \vec{e}_k^{(q)}$$

Beispiele:

a) Skalierung des Raumes $\mathbb{R}^n$ um einen Faktor $c \in \mathbb{R}$:

- Definition der Abbildung:

$$f_{\mathbb{R}^n}\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} c \cdot x_1 \\ c \cdot x_2 \\ \vdots \\ c \cdot x_n \end{pmatrix}$$

- Abbildung der Basisvektoren:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ c \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ \vdots \\ c \end{pmatrix}$$

- Matrix:

$$A_{f,\mathbb{R}^n} = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}$$

b) Projektion von $\mathbb{R}^3$ auf die xy-Ebene (nach $\mathbb{R}^3$):

- Definition der Abbildung:

$$g\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

- Abbildung der Basisvektoren:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

- Matrix:

$$B_g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

c) Projektion von $\mathbb{R}^3$ auf die xy-Ebene (nach $\mathbb{R}^2$):

- Definition der Abbildung:

$$g'\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}$$

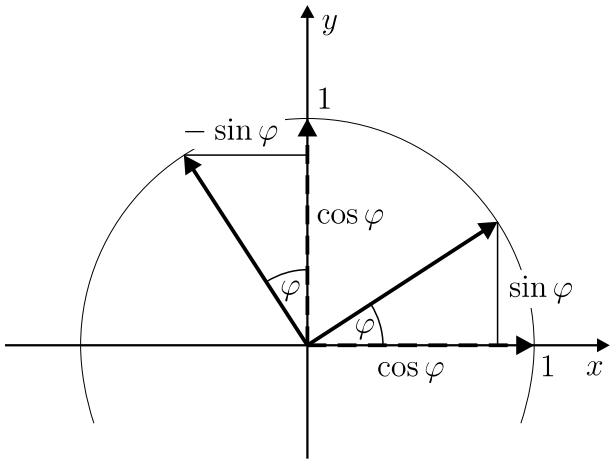
- Abbildung der Basisvektoren:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Matrix:

$$B'_{g'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

d) Drehung ($\circlearrowleft$) von $\mathbb{R}^2$ um einen Winkel φ:



- Definition der Abbildung:

$$h\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \cdot \cos\varphi - y \cdot \sin\varphi \\ x \cdot \sin\varphi + y \cdot \cos\varphi \end{pmatrix}$$

- Abbildung der Basisvektoren:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$$

- Matrix:

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

e) Drehung ($\circlearrowleft$) von $\mathbb{R}^2$ um einen Winkel φ mit anschließender Skalierung um den Faktor $c \in \mathbb{R}$:

- Definition der Abbildung:

$$\begin{aligned} (f_{\mathbb{R}^2} h)\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= f_{\mathbb{R}^2}\left(\begin{pmatrix} x \cdot \cos\varphi - y \cdot \sin\varphi \\ x \cdot \sin\varphi + y \cdot \cos\varphi \end{pmatrix}\right) \\ &= \begin{pmatrix} x \cdot c \cdot \cos\varphi - y \cdot c \cdot \sin\varphi \\ x \cdot c \cdot \sin\varphi + y \cdot c \cdot \cos\varphi \end{pmatrix} \end{aligned}$$

- Matrix:

$$\begin{aligned} A_{f,\mathbb{R}^2} \cdot C_h &= \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \\ &= \begin{pmatrix} c \cdot \cos\varphi & -c \cdot \sin\varphi \\ c \cdot \sin\varphi & c \cdot \cos\varphi \end{pmatrix} \end{aligned}$$

2.6 Rang einer Matrix

2.6.1 Einleitung

Definition: Sei $A \in M(m \times n, K)$ eine Matrix und $f \in \mathrm{Hom}(K^n, K^m)$ die zugehörige lineare Abbildung (bezüglich der Standardbasis), dann ist der *Rang* von A definiert als

$$\mathrm{rg}\ A := \mathrm{rg}\ f = \dim\ (\mathrm{Im}\ f)$$

Der *Zeilenrang* von A ist die maximale Anzahl von linear unabhängigen Zeilenvektoren aus A.

Der *Spaltenrang* von A ist die maximale Anzahl von linear unabhängigen Spaltenvektoren aus A.

Lemma: Ist $\vec{v}_i$ ein Spaltenvektor (Zeilenvektor) von A, der sich als Linearkombination der übrigen Spalten (Zeilen) darstellen lässt und ist A' die Matrix A ohne Spalte (Zeile) $\vec{v}_i$, dann gilt:

$$\text{Spaltenrang } A' = \text{Spaltenrang } A \quad \text{bzw.} \quad \text{Zeilenrang } A' = \text{Zeilenrang } A$$

Satz: Der Rang, der Spaltenrang und der Zeilenrang einer Matrix A sind gleich:

$$\mathrm{rg}\ A = \text{Spaltenrang } A = \text{Zeilenrang } A$$

Beweis:

- Zu zeigen ist, dass $\mathrm{rg}\ A = \text{Spaltenrang } A$:

  Die Spalten von A sind die Bilder der Basisvektoren. Daraus folgt, dass die Spaltenvektoren Erzeugendensystem für Im f sind. Damit ist die maximale linear unabhängige Teilmenge der Spaltenvektoren die Basis von Im f. Also ist der Spaltenrang von A die Dimension von Im f:

  $$\text{Spaltenrang } A = \dim\ (\mathrm{Im}\ f) = \mathrm{rg}\ A$$

- Zu zeigen ist, dass Spaltenrang A = Zeilenrang A:

  Streiche aus A Zeilen und/oder Spalten, die jeweils Linearkombinationen der übrigen Zeilen bzw. Spalten sind, solange das möglich ist.

  $$A \mapsto A' \mapsto A'' \mapsto \ldots \mapsto A^{(\mathrm{end})}$$

  Nach dem Lemma gilt:

  $$\begin{aligned} n &:= \text{Spaltenrang } A = \text{Spaltenrang } A^{(\mathrm{end})} \\ m &:= \text{Zeilenrang } A = \text{Zeilenrang } A^{(\mathrm{end})} \end{aligned}$$

– Angenommen, dass $n < m$:

  Das heißt, dass $A^{(\text{end})}$ m Zeilen hat, aber m Vektoren können in K^n nicht linear unabhängig sein. Damit muss einer der Vektoren eine Linearkombination der übrigen Vektoren sein. Dies ist ein Widerspruch zur Annahme. Also ist $n \geq m$.

– Angenommen, dass $n > m$:

  Das heißt, dass $A^{(\text{end})}$ n Spalten hat, aber n Vektoren können in K^m nicht linear unabhängig sein. Damit muss einer der Vektoren eine Linearkombination der übrigen Vektoren sein. Dies ist ein Widerspruch zur Annahme. Also ist $n = m$. □

Definition: Sei $A = (a_{i\,j}) \in M(m \times n, K)$ eine Matrix, dann ist transponierte Matrix von A definiert durch

$$A^t = (a^t_{i\,j}) \in M(n \times m, K) \quad \text{mit} \quad a^t_{i\,j} = a_{j\,i}$$

Beispiel:

$$\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 4 & 0 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 0 \end{pmatrix}$$

Folgerung: Der Rang einer Matrix A und der transponierten Matrix A^t ist gleich.

$$\text{rg } A = \text{rg } A^t$$

2.6.2 Elementare Umformungen

Feststellung: Der Rang einer Matrix kann mit den folgenden elementaren Umformungen bestimmt werden:

- Typ 1: Vertauschung von zwei Zeilen (Spalen).
- Typ 2: Multiplikation einer Zeile (Spalte) mit einem Skalar $\lambda \neq 0$.
- Typ 3: Addition des λ-fachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte).

Satz: Elementare Umformungen ändern den Rang einer Matrix nicht.

Beweis:

- Typ 1 und 2: trivial
- Typ 3: Sei $\vec{v}_i$ ein Zeilenvektor vor und $\vec{v}_i^*$ nach der Umformung, $\vec{v}_k$ sei ein anderer Zeilenvektor und $\lambda \in K$ ein Skalar:

  $$\vec{v}_i^* = \vec{v}_i + \lambda\vec{v}_k$$

  Die ursprüngliche Matrix sei A und die Matrix nach der Umformung A^*:

  $$A^* = A(\vec{v} \leftrightarrow \vec{v}^*)$$

  Sei $\vec{w}$ darstellbar als Linearkombination aus den Zeilenvektoren von A:

  $$\vec{w} = \mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \ldots + \mu_i\vec{v}_i + \ldots + \mu_k\vec{v}_k + \ldots + \mu_n\vec{v}_n$$

  Damit ist der Vektor $\vec{w}$ auch als Linearkombination aus den Zeilenvektoren von A^* darstellbar:

  $$\vec{w} = \mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \ldots + \mu_i\vec{v}_i^* + \ldots + (\mu_k - \lambda\mu_i)\vec{v}_k + \ldots + \mu_n\vec{v}_n$$

  Daraus folgt:

  $$\text{Lin(Zeilenvektoren von } A) = \text{Lin(Zeilenvektoren von } A^*)$$

  Da die Dimension gleich bleibt, bleibt auch der Rang gleich. $\square$

2.6.3 Obere Dreiecksform

Definition: Die Matrix A ist in oberer Dreiecksform, wenn die Matrix die folgende Form hat (das Symbol $*$ steht für beliebigen Inhalt):

$$A = \left(\begin{array}{ccccc|ccc} a_{1\,1} & * & * & \cdots & * & * & \cdots & * \\ 0 & a_{2\,2} & * & \cdots & * & \vdots & \ddots & \vdots \\ 0 & 0 & a_{3\,3} & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{r\,r} & * & \cdots & * \\ \hline 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{array}\right)$$

Für die Werte $a_{1\,1}, a_{2\,2}, a_{3\,3}, \ldots a_{r\,r}$ muss dabei gelten:

$$a_{1\,1} \cdot a_{2\,2} \cdot a_{3\,3} \cdot \ldots \cdot a_{r\,r} \neq 0$$

Beobachtung: Der Rang einer solchen Matrix ist r.

Verfahren: Überführung einer Matrix $A \in M(m \times n, K)$ in obere Dreiecksform:

- Die Matrix A_0 wird mit der Matrix A initialisiert.

$$A_0 := A$$

- Anschließend A_k mit $k = 0, 1, \ldots \min(m, n)$ das folgende Verfahren angewandt. Dabei muss A_k vor jeder Inkrementierung von k folgende Form haben:

$$A_k = \left(\begin{array}{ccccc|ccc} a_{1\,1} & * & * & \cdots & * & * & \cdots & * \\ 0 & a_{2\,2} & * & \cdots & * & \vdots & \ddots & \vdots \\ 0 & 0 & a_{3\,3} & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{k\,k} & * & \cdots & * \\ \hline 0 & \cdots & \cdots & \cdots & 0 & b_{k+1\,k+1} & \cdots & b_{k+1\,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & b_{m\,k+1} & \cdots & b_{m\,n} \end{array}\right)$$

  Die Koeffzienten $b_{i\,j}$ (mit $i = k+1, k+2, \ldots m$ und $j = k+1, k+2, \ldots n$) sind beliebig, und es gilt $a_{1\,1} \cdot a_{2\,2} \cdot a_{3\,3} \cdot \ldots \cdot a_{k\,k} \neq 0$.

Die Teilmatrix von A_k, die nur aus den Elementen $b_{p\,q}$ besteht, wird im Folgenden mit B bezeichnet:

$$B = \begin{pmatrix} b_{k+1\,k+1} & \cdots & b_{k+1\,n} \\ \vdots & \ddots & \vdots \\ b_{m\,k+1} & \cdots & b_{m\,n} \end{pmatrix}$$

Verfahren für A_k:

- Falls für alle $b_{i\,j}$ aus B gilt
$$b_{i\,j} = 0$$
dann ist das Verfahren abgeschlossen. Die Matrix A_k hat obere Dreiecksform.
- Sonst werden folgende Umformungen durchgeführt:
  a) Vertausche Zeilen und/oder Spalten, die durch B gehen, um einen Koeffizieten $b_{i,j} \neq 0$ an die Stelle $a'_{k+1\,k+1}$ zu bringen:

$$A'_k = \left(\begin{array}{cccc|cccc} a_{1\,1} & * & \cdots & * & * & * & \cdots & * \\ 0 & a_{2\,2} & \ddots & * & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{k\,k} & * & * & \cdots & * \\ \hline 0 & \cdots & \cdots & 0 & a'_{k+1\,k+1} & \cdots & \cdots & a'_{k+1\,n} \\ \vdots & \ddots & \ddots & \vdots & b'_{k+2\,k+1} & \cdots & \cdots & b'_{k+2\,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & b'_{m\,k+1} & \cdots & \cdots & b'_{m\,n} \end{array}\right)$$

b) Für $b'_{k+2\,k+1}, b'_{k+3\,k+1}, \dots b'_{m\,k+1}$ werden durch Typ-3-Umformungen Nullen erzeugt:

$$A'_k = \left(\begin{array}{cccc|cccc} a_{1\,1} & * & \cdots & * & * & * & \cdots & * \\ 0 & a_{2\,2} & \ddots & * & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{k\,k} & * & * & \cdots & * \\ \hline 0 & \cdots & \cdots & 0 & a'_{k+1\,k+1} & \cdots & \cdots & a'_{k+1\,n} \\ \vdots & \ddots & \ddots & \vdots & 0 & b''_{k+2\,k+2} & \cdots & b''_{k+2\,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & b''_{m\,k+2} & \cdots & b''_{m\,n} \end{array}\right)$$

Dies wird durch folgende Operation realisiert ($i = k+2, k+3, \dots m$ und $j = k+1, k+2, \dots n$):

$$b''_{i\,j} := b'_{i\,j} - a'_{i\,k+1} \cdot \frac{b'_{k+1\,j}}{a'_{k+1\,k+1}}$$

Beispiel:

- Folgende Matrix A soll in obere Dreiecksform umgeformt werden:

$$\begin{pmatrix} 0 & -2 & 4 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \\ 2 & 0 & 3 \end{pmatrix}$$

- Vertausche die erste und die dritte Zeile, so dass an der Stelle $a_{1\,1}$ ein Koeffizient $\neq 0$ steht:

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & -2 & 4 \\ 2 & 0 & 3 \end{pmatrix}$$

- Erzeuge an den Stellen $a_{2\,1}$ und $a_{4\,1}$ Nullen durch Typ-3-Umformungen mit der ersten Zeile:

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

- An der Stelle a_{22} befindet sich ein Koeffizient $\neq 0$. Damit muss nur noch an der Stelle a_{23} eine Null erzeugt werden:

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

- An der Stelle a_{33} muss eine Null erzeugt werden:

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

2.6.4 Elementarmatrizen

Beobachtung: Elementare Matrixumformungen können auch durch Multiplikation mit sogenannten Elementarmatrizen realisiert werden.

- Typ 1:

  Für die Vertauschung der i-ten und der j-ten Zeile (Spalte) in einer Matrix A wird folgende Matrix durch Abwandlung der Einheitsmatrix konstruiert:

$$T_{ij} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & 0 & \cdots & \rightarrow & \cdots & 1 & \\ & & & \vdots & 1 & & & \vdots & \\ & & & \downarrow & & \ddots & & \uparrow & \\ & & & \vdots & & & 1 & \vdots & \\ & & & 1 & \cdots & \leftarrow & \cdots & 0 & \\ & & & & & & & & 1 \\ & 0 & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \longleftarrow \ i\text{-te Zeile} \\ \\ \\ \\ \longleftarrow \ j\text{-te Zeile} \\ \\ \\ \end{matrix}$$

  Vertauschung der i-ten und der j-ten Zeile:

$$A' = T_{ij} \cdot A$$

  Vertauschung der i-ten und der j-ten Spalte:

$$A'' = A \cdot T_{ij}$$

- Typ 2:

  Für die Multiplikation einer Zeile (Spalte) mit dem Faktor λ in einer Matrix A wird eine Matrix konstruiert, in welcher in der Diagonalen der Einheitsmatrix in der entsprechenden Zeile (Spalte) eine Eins durch λ ersetzt:

$$S_{i\,\lambda} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \longleftarrow \quad i\text{-te Zeile}$$

  Multiplkation der i-ten Zeile mit λ:

$$A' = S_{i\,\lambda} \cdot A$$

  Multiplkation der j-ten Spalte mit λ:

$$A'' = A \cdot S_{i\,\lambda}$$

- Typ 3:

  Für die Addition des λ-fachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte) in einer Matrix A wird eine Einheitsmatrix um ein λ folgendermaßen erweitert:

$$K_{i\,j\,\lambda} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & \ddots & & & & \\ & & & \ddots & \lambda & & \\ & & & & \ddots & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \longleftarrow \quad i\text{-te Zeile}$$

$$\uparrow$$

$$j\text{-te Spalte}$$

  Addition des λ-fachen der j-ten Zeile zur i-ten Zeile:

$$A' = K_{i\,j\,\lambda} \cdot A$$

  Addition des λ-fachen der i-ten Spalte zur j-ten Spalte:

$$A'' = A \cdot K_{i\,j\,\lambda}$$

2.7 Lineare Gleichungssysteme

2.7.1 Einleitung

Definition: Ein *lineares Gleichungssystem* (LGS) mit Koeffizienten in einem Körper K, mit m Gleichungen und n Unbekannten wird durch eine Matrix

$$A = (a_{i\,j})_{(i,j)\in m\times n} \in M(m \times n, K)$$

und einem Vektor

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in K^m$$

repräsentiert. Das lineare Gleichungssystem wird folgendermaßen interpretiert:

$$\begin{array}{ccccccccc} a_{1\,1}\cdot x_1 & + & a_{1\,2}\cdot x_2 & + & \ldots & + & a_{1\,n}\cdot x_n & = & b_1 \\ a_{2\,1}\cdot x_1 & + & a_{2\,2}\cdot x_2 & + & \ldots & + & a_{2\,n}\cdot x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m\,1}\cdot x_1 & + & a_{m\,2}\cdot x_2 & + & \ldots & + & a_{m\,n}\cdot x_n & = & b_m \end{array}$$

Man bezeichnet mit $(A\,|\,b)$ folgende Matrix:

$$(A\,|\,b) = \left(\begin{array}{cccc|c} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,n} & b_1 \\ a_{2\,1} & a_{2\,2} & \cdots & a_{2\,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \\ a_{m\,1} & a_{m\,2} & \cdots & a_{m\,n} & b_m \end{array}\right)$$

Beobachtung: $x_1, x_2, \ldots x_n$ bilden genau dann eine Lösung des linearen Gleichungssystems, wenn

$$A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \vec{b}$$

Satz: Die Gleichung $A \cdot \vec{x} = \vec{b}$ ist genau dann lösbar, wenn

$$\text{rg } A = \text{rg}(A \,|\, b)$$

Beweis:

$$\begin{aligned}
&\text{rg } A = \text{rg}(A \mid b) \\
\Leftrightarrow\ &\text{Spaltenrang}(A) = \text{Spaltenrang}(A \mid b) \\
\Leftrightarrow\ &\vec{b} \in \text{Lin}\left(\left\{\begin{pmatrix} a_{1\,1} \\ a_{2\,1} \\ \vdots \\ a_{m\,1} \end{pmatrix}, \begin{pmatrix} a_{1\,2} \\ a_{2\,2} \\ \vdots \\ a_{m\,2} \end{pmatrix}, \ldots \begin{pmatrix} a_{1\,n} \\ a_{2\,n} \\ \vdots \\ a_{m\,n} \end{pmatrix}\right\}\right) \\
\Leftrightarrow\ &\exists x_1, x_2, \ldots x_n \quad \vec{b} = x_1 \cdot \begin{pmatrix} a_{1\,1} \\ a_{2\,1} \\ \vdots \\ a_{m\,1} \end{pmatrix} + x_2 \cdot \begin{pmatrix} a_{1\,2} \\ a_{2\,2} \\ \vdots \\ a_{m\,2} \end{pmatrix} + \ldots + x_n \cdot \begin{pmatrix} a_{1\,n} \\ a_{2\,n} \\ \vdots \\ a_{m\,n} \end{pmatrix} \\
\Leftrightarrow\ &\exists x_1, x_2, \ldots x_n \quad A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \vec{b}
\end{aligned}$$

Definition: Ein lineares Gleichungssystem mit $\vec{b} = \vec{0}^{\,(m)}$ wird *homogenes Gleichungssystem* genannt.

Beobachtung: Zu jedem linearen Gleichungssystem kann durch Ersetzung von $\vec{b}$ durch $\vec{0}^{\,(m)}$ ein homogenes Gleichungssystem gefunden werden.

Definition: Die Lösungsmenge eines linaren Gleichungssystems aus A und $\vec{b}$ ist folgermaßen definiert:

$$\text{Lös}(A, \vec{b}) = \{\vec{x} \mid A \cdot \vec{x} = \vec{b}\}$$

Satz: Sei $A \in M(m \times n, K)$ die Matrix einer linearen Abbildung $f : K^n \to K^m$ bezüglich der Standardbasis, dann ist

a) Die Lösungsmenge $\text{Lös}(A, \vec{0}^{\,(m)})$ ist gleich Ker f. Damit ist $\text{Lös}(A, \vec{0}^{\,(m)})$ ein Unterraum von K^n.

b) Sei $\vec{b} \in K^m$ und $\vec{x}, \vec{y} \in \text{Lös}(A, \vec{b})$, dann gilt:

$$\vec{x} - \vec{y} \in \text{Lös}(A, \vec{0}^{(m)})$$

c) Sei $\vec{b} \in K^m$, $\vec{x} \in \text{Lös}(A, \vec{b})$ und $\vec{z} \in \text{Lös}(A, \vec{0}^{(m)})$, dann gilt:

$$\vec{x} + \vec{z} \in \text{Lös}(A, \vec{b})$$

Beweis:

a) Durch Einsetzen der jeweiligen Definitionen folgt:

$$\begin{aligned} \text{Lös}(A, \vec{0}^{(m)}) &= \{\vec{x} \mid A \cdot \vec{x} = \vec{0}^{(m)}\} \\ &= \{\vec{x} \mid f(\vec{x}) = \vec{0}^{(m)}\} \\ &= \text{Ker } f \end{aligned}$$

b) Seien $\vec{x}, \vec{y} \in \text{Lös}(A, \vec{b})$, dann gilt:

$$f(\vec{x}) = \vec{b} \quad \text{und} \quad f(\vec{y}) = \vec{b}$$

Daraus folgt:

$$f(\vec{x} - \vec{y}) = f(\vec{x}) - f(\vec{y}) = \vec{b} - \vec{b} = \vec{0}$$

Das heißt:

$$\vec{x} - \vec{y} \in \text{Lös}(A, \vec{0}^{(m)})$$

c) Seien $\vec{x} \in \text{Lös}(A, \vec{b})$ und $\vec{z} \in \text{Lös}(A, \vec{0}^{(m)})$, dann gilt:

$$f(\vec{x}) = \vec{b} \quad \text{und} \quad f(\vec{z}) = \vec{0}$$

Daraus folgt:

$$f(\vec{x} + \vec{z}) = f(\vec{x}) - f(\vec{z}) = \vec{b} - \vec{0} = \vec{b}$$

Das heißt:

$$\vec{x} + \vec{z} \in \text{Lös}(A, \vec{b})$$

Beobachtung: Sei $A \in M(m \times n, K)$ die Matrix einer linearen Abbildung $f : K^n \to K^m$ bezüglich der Standardbasis, dann ist die Lösungsmenge $\text{Lös}(A, b)$ genau dann *nicht* leer, wenn $\vec{b} \in \text{Im } f$.

2.7.2 Gauß'scher Algorithmus

Verfahren: Ein lineares Gleichungssystem aus A und $\vec{b}$ kann nach folgendem Algorithmus gelöst werden:

a) Überprüfung, ob das lineare Gleichungssystem eine Lösung hat: Man bringt die Matrix A in obere Dreiecksform, aber erweitert alle Schritte auf die Matrix $(A \mid b)$ ohne Elementarumformungen mit letzter Spalte vorzunehmen.

Achtung: Spaltenvertauschungen in A bedeuten Variablenvertauschung im linearen Gleichungssystem.

Ergebnis:

$$A' = \left(\begin{array}{ccccc|ccc|c} a'_{1\,1} & * & * & \cdots & * & * & \cdots & * & b'_1 \\ 0 & a'_{2\,2} & * & \cdots & * & \vdots & \ddots & \vdots & b'_2 \\ 0 & 0 & a'_{3\,3} & \ddots & \vdots & \vdots & \ddots & \vdots & b'_3 \\ \vdots & \vdots & \ddots & \ddots & * & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a'_{r\,r} & * & \cdots & * & b'_r \\ \hline 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & b'_{r+1} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & b'_m \end{array}\right)$$

- Fall 1: Falls mindestens ein $b'_i \neq 0$ mit $i = r+1, r+2, \ldots m$ existiert, dann ist rg $(A) <$ rg$(A \mid b)$, und das System hat *keine* Lösung. Das Verfahren wird abgebrochen.
- Fall 2: Falls $b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0$, dann ist rg $(A) =$ rg$(A \mid b)$, und das System hat eine Lösung.

  Im Folgenden wird das System auf folgende Form reduziert:

$$(T \mid S \mid b') = \left(\begin{array}{ccccc|ccc|c} a'_{1\,1} & * & * & \cdots & * & * & \cdots & * & b'_1 \\ 0 & a'_{2\,2} & * & \cdots & * & \vdots & \ddots & \vdots & b'_2 \\ 0 & 0 & a'_{3\,3} & \ddots & \vdots & \vdots & \ddots & \vdots & b'_3 \\ \vdots & \vdots & \ddots & \ddots & * & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a'_{r\,r} & * & \cdots & * & b'_r \end{array}\right)$$

b) Bestimmung einer speziellen Lösung von $(T \mid S \mid b')$:

Zur Bestimmung *einer* speziellen Lösung von $(T \mid S \mid b')$ wird das lineare Gleichungssystem folgendermaßer geteilt:

$$(T \mid S) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = T \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} + S \cdot \begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \vec{b}'$$

Man wählt $x_{r+1} = x_{r+2} = \ldots = x_n = 0$. Dadurch reduziert sich das lineare Gleichungssystem wiefolgt:

$$T \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \vec{b}' \quad \Rightarrow \quad \begin{pmatrix} a'_{1\,1} & a'_{1\,2} & \cdots & a'_{1\,r} \\ 0 & a'_{2\,2} & \cdots & a'_{2\,r} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a'_{r\,r} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_r \end{pmatrix}$$

Die Werte von $x_1, x_2, \ldots x_r$ können nun direkt bestimmt werden:

$$\begin{array}{rclcrcl} b'_r & = & a'_{r\,r} \cdot x_r & \Rightarrow & x_r & = & \frac{b'_r}{a'_{r\,r}} \\ b'_{r-1} & = & a'_{r-1\,r-1} \cdot x_{r-1} + a'_{r-1\,r} \cdot x_r & \Rightarrow & x_{r-1} & = & \frac{b'_{r-1} - a'_{r-1\,r} \cdot x_r}{a'_{r-1\,r-1}} \\ & & \vdots & & & & \\ b'_1 & = & a'_{1\,1} \cdot x_1 + \ldots + a'_{1\,r} \cdot x_r & \Rightarrow & x_1 & = & \frac{b'_1 - a'_{1\,r} \cdot x_r - \ldots - a'_{1\,2} \cdot x_2}{a'_{1\,1}} \end{array}$$

Der Vektor mit einer speziellen Lösung des linearen Gleichungssystem werde mit $\vec{x}$ bezeichnet:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

c) Bestimmung einer Basis der Lösungsmenge des homogenen Gleichungssystems $(T \mid S \mid \vec{0}^{\,(r)})$:

Zur Bestimmung des j-ten Basisvektors von

$$(T \mid S) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = T \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} + S \cdot \begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \vec{0}^{\,(r)}$$

wählt man

$$x_{r+j} = 1 \quad \text{und} \quad x_{r+1} = \ldots = x_{r+j-1} = x_{r+j+1} = \ldots x_n = 0$$

Im Folgenden gelte:

$$S = (s_{i\,j})_{(i,j) \in r \times (n-r)}$$

Damit erhält man das folgende lineare Gleichungssystem

$$\begin{pmatrix} a'_{1\,1} & a'_{1\,2} & \cdots & a'_{1\,r} \\ 0 & a'_{2\,2} & \cdots & a'_{2\,r} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a'_{r\,r} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} -s_{1\,j} \\ -s_{2\,j} \\ \vdots \\ -s_{r\,j} \end{pmatrix}$$

Bemerkung: Der Vektor

$$\begin{pmatrix} -s_{1\,j} \\ -s_{2\,j} \\ \vdots \\ -s_{r\,j} \end{pmatrix}$$

ist dabei der j-Spaltenvektor aus S, multipliziert mit -1.

Die Werte von $x_1, x_2, \ldots x_r$ können nun wie bei der speziellen Lösung bestimmt werden.

Der j-te Basisvektor des homogenen Gleichungssystem werde mit $\vec{x}_j$ bezeichnet.

Das Verfahren muss für alle $n - r$ Spalten von S durchgeführt werden.

d) Bestimmung der allgemeinen Lösung von $(T \mid S \mid b)$:

Mit Hilfe der speziellen Lösung des linaren Gleichungssystem und der Basisvektoren der Lösungsmenge der homogenen Gleichungssystem lässt sich die Lösungsmenge des linearen Gleichungssystem folgermaßen darstellen:

$$\text{Lös}(A \mid b) = \text{Lös}(T \mid S \mid b') = \left\{ \vec{x} + \sum_{j=1}^{n-r} \lambda_j \cdot \vec{x}_j \mid \lambda_1, \lambda_2, \ldots \lambda_{n-r} \in \mathbb{R} \right\}$$

Beispiel: Gegeben sei das folgende Gleichungssystem:

$$\begin{array}{rcrcrcrcr} & & 2x_2 & + & x_3 & - & x_4 & = & 6 \\ x_1 & - & x_2 & + & 2x_3 & & & = & -1 \\ 2x_1 & & & + & 5x_3 & & & = & 3 \\ - & x_1 & - & x_2 & - & 3x_3 & + & 2x_4 & = & -6 \end{array}$$

a) Die dazugehörige Matrix $(A \mid b)$:

$$\left(\begin{array}{rrrr|r} 0 & 2 & 1 & -1 & 6 \\ 1 & -1 & 2 & 0 & -1 \\ 2 & 0 & 5 & 0 & 3 \\ -1 & -1 & -3 & 2 & -6 \end{array}\right)$$

Diese Matrix muss zuerst in obere Dreiecksform überführt werden.

Erste und zweite Zeile vertauschen:

$$\left(\begin{array}{rrrr|r} 1 & -1 & 2 & 0 & -1 \\ 0 & 2 & 1 & -1 & 6 \\ 2 & 0 & 5 & 0 & 3 \\ -1 & -1 & -3 & 2 & -6 \end{array}\right)$$

In der ersten Spalte unter $a_{1\,1}$ Nullen erzeugen:

$$\left(\begin{array}{rrrr|r} 1 & -1 & 2 & 0 & -1 \\ 0 & 2 & 1 & -1 & 6 \\ 0 & 2 & 1 & 0 & 5 \\ 0 & -2 & -1 & 2 & -7 \end{array}\right)$$

In der zweiten Spalte unter $a_{2\,2}$ Nullen erzeugen:

$$\left(\begin{array}{rrrr|r} 1 & -1 & 2 & 0 & -1 \\ 0 & 2 & 1 & -1 & 6 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

Dritte und vierte Spalte tauschen, um an der Stelle $a_{3\,3}$ einen Wert $\neq 0$ zu erzeugen (Achtung: $x_3 \leftrightarrow x_4$):

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

In der dritten Spalte unter $a_{3\,3}$ Nullen erzeugen:

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

Die Matrix hat nun obere Dreiecksform.

Es existiert eine Lösung, da der untere Teil von $\vec{b}$ aus einer Null besteht. Die Matrix kann nun folgendermaßen reduziert werden:

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

Das lineare Gleichungssystem wird geteilt:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \cdot (x_3) = \begin{pmatrix} -1 \\ 6 \\ -1 \end{pmatrix}$$

b) Bestimmung der speziellen Lösung:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 6 \\ -1 \end{pmatrix}$$

Wähle $x_3 = 0$ und bestimme die übrigen Variablen:

$$\begin{array}{rcrcrcrcl} & & & & 1 \cdot x_4 & = & -1 & \Rightarrow & x_4 = -1 \\ & & 2 \cdot x_2 & + & (-1) \cdot x_4 & = & 6 & \Rightarrow & x_2 = 2{,}5 \\ 1 \cdot x_1 & + & (-1) \cdot x_2 & + & 0 \cdot x_4 & = & -1 & \Rightarrow & x_1 = 1{,}5 \end{array}$$

c) Bestimmung des ersten (und einzigen) Basisvektors von $(A \mid \vec{0}^{(4)})$:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$$

Wähle $x_3 = 1$ und bestimme die übrigen Variablen:

$$\begin{array}{rcrcrcrcr} & & & & 1 \cdot x_4 & = & 0 & \Rightarrow & x_4 = 0 \\ & & 2 \cdot x_2 & + & (-1) \cdot x_4 & = & -1 & \Rightarrow & x_2 = -0{,}5 \\ 1 \cdot x_1 & + & (-1) \cdot x_2 & + & 0 \cdot x_4 & = & -2 & \Rightarrow & x_1 = -2{,}5 \end{array}$$

d) Lösungsmenge:

$$\text{Lös}(A \mid b) = \left\{ \begin{pmatrix} 1{,}5 \\ 2{,}5 \\ 0 \\ -1 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -2{,}5 \\ -0{,}5 \\ 1 \\ 0 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$$

2.7.3 Quotientenraum

Beobachtung: Die Lösungsmenge ist *kein* Unterraum, aber eine „Verschiebung“ eines Unterraums.

Definition: Sei V ein Vektorraum, U ein Unterraum von V und $\vec{v}$ ein Vektor aus V, dann nennt man

$$\vec{v} + U = \{\vec{v} + \vec{u} \mid \vec{u} \in U\}$$

die *Nebenklasse* von $\vec{v}$ bezüglich U.

Satz: Sei V ein Vektorraum, U ein Unterraum von V und $\vec{w}, \vec{w}$ zwei Vektoren aus V, dann gilt:

$$\vec{v} + U = \vec{w} + U \quad \Leftrightarrow \quad \vec{v} - \vec{w} \in U \quad \Leftrightarrow \quad \vec{w} \in \vec{v} + U$$

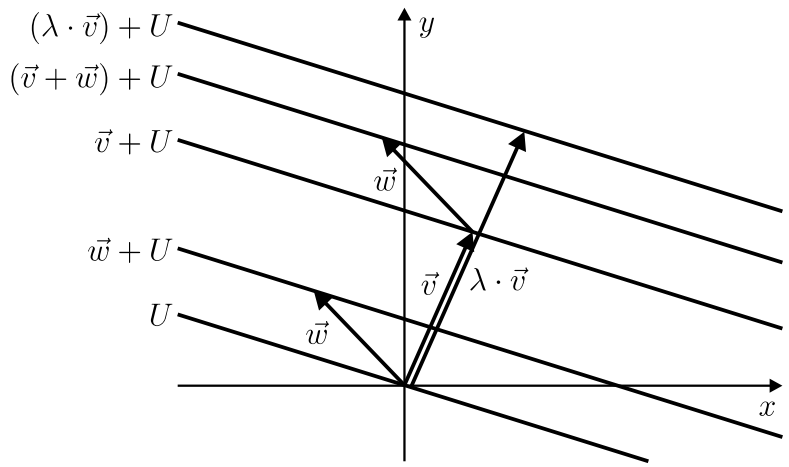
Definition: Sei V ein Vektorraum, U ein Unterraum von V und K der Körper von V, dann bezeichnet man die Menge

$$V_{/U} = \{\vec{v} + U \mid \vec{v} \in V\}$$

als *Quotientenraum* von V nach U.

Beobachtung: Der Quotientenraum $V_{/U}$ ist ein Vektorraum mit folgenden Operationen ($\vec{v}, \vec{w} \in V$, $\lambda \in K$)

$$\begin{aligned}(\vec{v}+U)+(\vec{w}+U) &= (\vec{v}+\vec{w})+U \\ \lambda \cdot (\vec{v}+U) &= (\lambda \cdot \vec{v})+U\end{aligned}$$



und dem neutralen Element $\vec{0} + U = U$.

Beobachtung: Sei $V_{/U}$ ein Quotientenraum, dann ist $\vec{v} + U \in V_{/U}$ eine Äquivalenzklasse von $\vec{v}$ bezüglich der Relation „Differenz ist in U“.

Satz: Sei V ein Vektorraum, U ein Unterraum von V, dann ist

$$\dim V_{/U} = \dim V - \dim U$$

Beweisidee: Man definiert eine Abbildung $\varphi : V \to V_{/U}$ durch

$$\vec{v} \mapsto \vec{v} + U$$

Die Abbidlung φ ist linear und surjektiv (epimorph). Damit ist das Im $\varphi = V_{/U}$. Außerdem ist Ker $\varphi = U$, denn

$$\forall\, \vec{v} \in U \quad \vec{v} + U = U = \vec{0} + U$$

Nach der Dimensionsformel gilt somit:

$$\dim V = \dim(\text{Ker } \varphi) + \dim(\text{Im } \varphi) = \dim U + \dim V_{/U}$$

Beobachtung: Seien $A \in M(m \times n, K)$ eine Matrix, $f \in \mathrm{Hom}(K^n, K^m)$ die dazugehörige Abbildung, $\vec{b}, \vec{c} \in K^n$ sowie $\vec{x}, \vec{y} \in K^m$ Vektoren und $\lambda \in K$ ein Skalar.

- Addition:

$$\mathrm{Lös}(A, \vec{b} + \vec{c}) = (\vec{x} + \vec{y}) + \mathrm{Ker}\ f$$

$$\Updownarrow$$

$$\mathrm{Lös}(A, \vec{b}) = \vec{x} + \mathrm{Ker}\ f \quad \text{und} \quad \mathrm{Lös}(A, \vec{c}) = \vec{y} + \mathrm{Ker}\ f$$

- Multiplikation mit Skalaren:

$$\mathrm{Lös}(A \,|\, \vec{b}) = \vec{x} + \mathrm{Ker}\ f$$

$$\Updownarrow$$

$$\mathrm{Lös}(A \,|\, \lambda\, \vec{b}) = \lambda\, \vec{x} + \mathrm{Ker}\ f$$

Lösungsmengen haben damit die Struktur von Vektorräumen.

Beispiel: Sei $A \in M(2 \times 2, \mathbb{R})$ eine Matrix und $f \in \mathrm{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ die dazugehörige Abbildung (bezüglich der Standardbasis):

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \leftrightarrow \quad f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + 2y \\ 2x + 4y \end{pmatrix}$$

Bestimmung von Ker f (also von $\mathrm{Lös}(A, \vec{0}^{\,(2)})$):

$$x + 2y = 0 \quad \text{und} \quad 2x + 4y = 0$$

Daraus folgt:

$$\mathrm{Lös}(A, \vec{0}^{\,(2)}) = \mathrm{Ker}\ f = \mathrm{Lin}\left(\left\{\begin{pmatrix} 2 \\ -1 \end{pmatrix}\right\}\right)$$

Sei $\vec{b} \in \mathbb{R}^2$ ein Vektor, dann gilt:

$$\mathrm{Lös}(A, \vec{b}) \in \mathbb{R}^2_{/\mathrm{Ker}\ f}$$

2.8 Inverse Matrizen

2.8.1 Einheitsmatrix

Definition: Die Matrix

$$E_n = \begin{pmatrix} \mathbf{1} & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{1} & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{1} \end{pmatrix} \in M(n \times n, K)$$

ist neutrales Element der Matrixmultiplikation in $M(n \times n, K)$ und wird als *Einheitsmatrix* bezeichnet. Das heißt ($A \in M(n \times n, K)$):

$$E_n \cdot A \quad = \quad A \quad = \quad A \cdot E_n$$

2.8.2 Inverse Matrizen

Definition: Sei $A \in M(n \times n, K)$, dann ist A^{-1} die zu A inverse Matrix, wenn

$$A \cdot A^{-1} \quad = \quad E_n \quad = \quad A^{-1} \cdot A$$

Achtung: Die Menge $M(n \times n, K)$ ist *keine* Gruppe. Es gibt z.B. gibt es kein inverses Element für die Nullmatrix:

$$\forall\, A \in M(n \times n, K) \quad \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \cdot A = A \cdot \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \neq E_n$$

Satz: Die Matrix $A \in M(n \times n, K)$ ist genau dann invertierbar, wenn $\operatorname{rg} A = n$.

Beweis: Aufgrund des Zusammenhanges zu linearen Abbildungen lässt sich jeder Matrix eine Funktion $f \in \operatorname{Hom}(K^n, K^n)$ zuordnen, sowie umgekehrt jeder linearen Abbildung eine Matrix zuordnen:

$$A \quad \leftrightarrow \quad f$$

Daraus folgt:

- Surjektivität von f:

$$\operatorname{rg} A = n \quad \Leftrightarrow \quad \dim(\operatorname{Im} f) = n \quad \Leftrightarrow \quad f \text{ ist surjektiv}$$

- Injektivität von f:

$$\dim(\text{Im } f) = n \quad \Leftrightarrow \quad \underbrace{\dim(\text{Ker } f) = 0}_{\dim K^n - \dim(\text{Im } f)} \quad \Leftrightarrow \quad f \text{ ist injektiv}$$

- Umkehrbarkeit von f:

  Da f surjektiv und injektiv ist, ist f bijektiv. Damit existiert eine Umkehrabbildung $f^{-1} \in \text{Hom}(K^n, K^n)$ und damit auch die dazugehörige Matrix $A' \in M(n \times n, K)$:

$$f^{-1} \quad \leftrightarrow \quad A'$$

- Verkettung von Funktionen als Matrixmultiplikation:

$$A \cdot A^{-1} \quad \leftrightarrow \quad f \cdot f^{-1} = \text{Id}_{K^n} \quad \leftrightarrow \quad E^n \quad \Rightarrow \quad A \cdot A^{-1} = E_n$$

□

Beobachtungen:

a) Seien $A, B, C \in M(n \times n, K)$ Matrixen und $A \cdot B = C$.

Überführt man mit den gleichen elementaren *Zeilen*umformungen A in A' und C in C' (ohne B zu verändern), so gilt $A' \cdot B = C'$.

Grund: Zeilenumformungen entsprechen Multiplikation mit Elementarmatrizen von links:

$$A' \cdot B = \underbrace{D_k \cdot \ldots \cdot D_2 \cdot D_1}\cdot A \cdot B = \underbrace{D_k \cdot \ldots \cdot D_2 \cdot D_1}\cdot C = C'$$

b) Sei $A \in M(n \times n, K)$ eine Matrix.

Ist die Matrix A invertierbar, so kann man A mit elementaren *Zeilen*umformungen in E_n überführen.

Grund: Die Matrix A hat vollen Rang haben.

c) Sei $A \in M(n \times n, K)$ eine Matrix.

Überführt man die Matrix A durch Zeilenumformungen in E_n und wendet die gleichen Umformungen auf E_n an, so erhält man A^{-1}.

Grund: Man wendet die Beobachtung a) an:

$$\begin{aligned} A &\rightsquigarrow E_n \\ E_n &\rightsquigarrow X \end{aligned}$$

$$A \cdot A^{-1} = E_n \quad \rightsquigarrow \quad E_n \cdot A^{-1} = \underbrace{D_k \cdot \ldots \cdot D_1}\cdot A \cdot A^{-1} = \underbrace{D_k \cdot \ldots \cdot D_1}\cdot E_n = X$$

$$E_n \cdot A^{-1} = X \quad \Rightarrow \quad X = A^{-1}$$

Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 0 & -1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_n$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 2 \\ 0 & -1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 4 \\ 0 & -1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} -1 & 2 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 4 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} -1 & 2 & 0 \\ -2 & 2 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} -3 & 4 & -2 \\ -2 & 2 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$E_n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} -3 & 4 & -2 \\ 2 & -2 & 1 \\ 0{,}5 & -0{,}5 & 0{,}5 \end{pmatrix} = A^{-1}$$

Probe:

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 0 & -1 & 4 \end{pmatrix} \cdot \begin{pmatrix} -3 & 4 & -2 \\ 2 & -2 & 1 \\ 0{,}5 & -0{,}5 & 0{,}5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2.9 Determinanten

2.9.1 Einleitung

Definition: Die *Determinante* $\det A$ ist eine Kenngröße einer quadratischen Matrix $A \in M(n \times n, K)$, die wie folgt bestimmen kann:

- Fall 1: $n = 1$

$$\det a_{1\,1} = a_{1\,1}$$

- Fall 2: $n > 1$

  Entwicklung nach der ersten Spalte:

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i\,1} \cdot \det A_{i\,1}$$

  Dabei ist $A_{i\,1}$ die Matrix, die man aus A durch Streichen der i-ten Zeile und der ersten Spalte enthält.

Schreibweise: Für die Determinante wird folgende Schreibweise vereinbart:

$$\begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,n} \end{vmatrix} := \det \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,n} \end{pmatrix}$$

Beispiel:

$$\begin{aligned} \begin{vmatrix} \mathbf{0} & 1 & 2 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{4} & \mathbf{6} \\ \mathbf{0} & 1 & 5 & 1 \\ \mathbf{0} & 0 & 2 & 0 \end{vmatrix} &= (-1)^{2+1} \cdot 1 \cdot \begin{vmatrix} 1 & 2 & 0 \\ 1 & 5 & 1 \\ 0 & 2 & 0 \end{vmatrix} \\ &= -\left(1 \cdot \begin{vmatrix} 5 & 1 \\ 2 & 0 \end{vmatrix} - 1 \cdot \begin{vmatrix} 2 & 0 \\ 2 & 0 \end{vmatrix} \right) \\ &= -\left((5 \cdot 0 - 2 \cdot 1) - (2 \cdot 0 - 2 \cdot 0) \right) \\ &= 2 \end{aligned}$$

Beobachtung: Für die Spezialfälle $n = 2$ und $n = 3$ existiert eine einfache Methode zur Bestimmung der Determinenten:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32}) - (a_{13} \cdot a_{22} \cdot a_{31} + a_{11} \cdot a_{23} \cdot a_{32} + a_{12} \cdot a_{21} \cdot a_{33})$$

Achtung: Ab $n = 4$ funktioniert diese Methode nicht mehr!

Definition: Eine Funktion $f : M(m \times n, K) \to L$ heißt *linear in jeder Zeile*, wenn folgende Bedingungen erfüllt sind:

- Seien $A, A' \in M(m \times n, K)$ Matrizen. Wenn sich A und A' nur in der p-ten Zeile unterscheiden, dann gilt:

$$\begin{aligned} f(A) + f(A') &= f\left(\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pn} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\right) + f\left(\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{p1} & \cdots & a'_{pn} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\right) \\ &= f\left(\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} + a'_{p1} & \cdots & a_{pn} + a'_{pn} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\right) \end{aligned}$$

- Sei $A, \in M(m \times n, K)$ eine Matrix und $\lambda \in K$ ein Skalar, dann gilt für alle $p \leq m$:

$$\begin{aligned}
\lambda \cdot f(A) &= \lambda \cdot f\left(\begin{pmatrix}\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{p\,1} & \cdots & a_{p\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,n}\end{pmatrix}\end{pmatrix}\right) \\
&= f\left(\begin{pmatrix}\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ \lambda \cdot a_{p\,1} & \cdots & \lambda \cdot a_{p\,n} \\ \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,n}\end{pmatrix}\end{pmatrix}\right)
\end{aligned}$$

Eine Funktion $f : M(m \times n, K) \to L$ heißt *linear in jeder Spalte*, wenn folgende Bedingungen erfüllt sind:

- Seien $A, A' \in M(m \times n, K)$ Matrizen. Wenn sich A und A' nur in der p-ten Spalte unterscheiden, dann gilt:

$$\begin{aligned}
f(A) + f(A') &= f\left(\begin{pmatrix}\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,p} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,p} & \cdots & a_{m\,n}\end{pmatrix}\end{pmatrix}\right) \\
&\quad + f\left(\begin{pmatrix}\begin{pmatrix} a_{1\,1} & \cdots & a'_{1\,p} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a'_{m\,p} & \cdots & a_{m\,n}\end{pmatrix}\end{pmatrix}\right) \\
&= f\left(\begin{pmatrix}\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,p} + a'_{1\,p} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,p} + a'_{m\,p} & \cdots & a_{m\,n}\end{pmatrix}\end{pmatrix}\right)
\end{aligned}$$

- Sei $A, \in M(m \times n, K)$ eine Matrix und $\lambda \in K$ ein Skalar, dann gilt für alle $p \leq n$:

$$\begin{aligned} \lambda \cdot f(A) &= \lambda \cdot f\left(\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,p} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & a_{m\,p} & \cdots & a_{m\,n} \end{pmatrix}\right) \\ &= f\left(\begin{pmatrix} a_{1\,1} & \cdots & \lambda \cdot a_{1\,p} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m\,1} & \cdots & \lambda \cdot a_{m\,p} & \cdots & a_{m\,n} \end{pmatrix}\right) \end{aligned}$$

Satz: Es gibt genau eine Abbildung $\det : M(n \times n, K) \to K$ mit den folgenden Eigenschaften:

- Die Abbildung det ist linear in jeder Zeile.
- Wenn $\operatorname{rg} A < n$, dann gilt $\det A = 0$.
- Für die Determinante der Einheitsmatrix E_n gilt: $\det E_n = 1$.

Diese Abbildung lässt sich durch die am Anfang angegebene Entwicklungsformel bestimmen.

Beobachtung: Die Abbidlung det ist ebenfalls linear in jeder Spalte.

2.9.2 Eigenschaften von Determinanten

Beobachtung: Die Abbildung det ist invariant gegenüber Typ-3-Zeilen- oder Spaltenumformungen:

- Zeilenumformungen:

  Umformung:

$$A = \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} & \cdots & a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix} \longrightarrow A' = \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} + \lambda \cdot a_{j\,1} & \cdots & a_{i\,n} + \lambda \cdot a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$

Betrachtung der Determinante von A':

$$\begin{aligned}
\det A' &= \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} + \lambda \cdot a_{j\,1} & \cdots & a_{i\,n} + \lambda \cdot a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{vmatrix} \\
&= \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} & \cdots & a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{vmatrix} + \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ \lambda \cdot \mathbf{a_{j\,1}} & \cdots & \lambda \cdot \mathbf{a_{j\,n}} \\ \vdots & \ddots & \vdots \\ \mathbf{a_{j\,1}} & \cdots & \mathbf{a_{j\,n}} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{vmatrix} \\
&= \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} & \cdots & a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{vmatrix} = \det A
\end{aligned}$$

$$\text{weil} \quad \operatorname{rg} \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ \lambda \cdot \mathbf{a_{j\,1}} & \cdots & \lambda \cdot \mathbf{a_{j\,n}} \\ \vdots & \ddots & \vdots \\ \mathbf{a_{j\,1}} & \cdots & \mathbf{a_{j\,n}} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix} < n$$

- Spaltenumformungen: analog

Beobachtung: Vertauschung von zwei Zeilen (Spalten) bewirkt Vorzeichenänderung der Determinanten:

- Vertauschung der i-ten und der j-ten Zeilen in der Matrix $A \in M(n \times n, K)$ mit Hilfe von Typ-2- und Typ-3-Zeilenumformungen:

  1. Matrix A:
  $$A = \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} & \cdots & a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$

  2. Addition der j-ten Zeile zur i-ten Zeile (Typ 3):
  $$\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} + a_{j\,1} & \cdots & a_{i\,n} + a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$

  3. Subtraktion der i-ten Zeile von der j-ten Zeile (Typ 3):
  $$\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} + a_{j\,1} & \cdots & a_{i\,n} + a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} - (a_{i\,1} + a_{j\,1}) & \cdots & a_{j\,n} - (a_{i\,n} + a_{j\,n}) \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$
  $$= \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} + a_{j\,1} & \cdots & a_{i\,n} + a_{j\,n} \\ \vdots & \ddots & \vdots \\ -a_{i\,1} & \cdots & -a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$

4. Addition der j-ten Zeile zur i-ten Zeile (Typ 3):

$$\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} + a_{j\,1} - a_{i\,1} & \cdots & a_{i\,n} + a_{j\,n} - a_{i\,n} \\ \vdots & \ddots & \vdots \\ -a_{i\,1} & \cdots & -a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$

$$= \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ -a_{i\,1} & \cdots & -a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix}$$

5. Multiplikation der j-ten Zeile mit -1 (Typ 2):

$$\begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} & \cdots & a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix} = A'$$

Betrachtung der Determinante von A':

Da die Determinante einer Matrix invariant gegenüber Typ-3-Zeilenumformungen ist, wirkt sich nur die Typ-2-Zeilenumformung, die Multiplikation einer Zeile mit -1, auf die Determinante aus: Aufgrund der Linearität der Determinante in einer Zeile, muss die Determinate durch diese Operation ebenfalls mit -1 multipliziert werden:

$$\begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ -a_{i\,1} & \cdots & -a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{vmatrix} = - \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & \ddots & \vdots \\ a_{j\,1} & \cdots & a_{j\,n} \\ \vdots & \ddots & \vdots \\ a_{i\,1} & \cdots & a_{i\,n} \\ \vdots & \ddots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{vmatrix}$$

- Vertauschung der i-ten und der j-ten Spalten in der Matrix $A \in M(n \times n, K)$ mit Hilfe von Typ-2- und Typ-3-Spaltenumformungen: analog

Folgerung: Vertauschung von zwei Zeilen (Spalten) bewirkt eine Vorzeichenänderung der Determinanten.

Folgerung: Die Determinante kann als Produkt der Diagonalelemente einer oberen Dreiecksmatrix nach elementaren Typ-1- und Typ-3-Zeilenumformungen (Spaltenumformungen) bestimmt werden, wobei für jeden Zeilentausch (Spaltentausch) noch mit -1 multipliziert werden muss.

Beispiel:

- Matrix A:
$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \\ 2 & 3 & 3 \end{pmatrix}$$
- Subtraktion der ersten Zeile von der zweiten Zeile:
$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 2 & 3 & 3 \end{pmatrix}$$
- Subtraktion des zweifachen der ersten Zeile von der dritten Zeile:
$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$
- Vertauschung der zweiten und der dritten Zeile:
$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = A'$$
Anzahl der Vertauschungen: $i = 1$
- Bestimmung der Determinante von A':
$$\det A' = \begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot 1 \cdot 1 = 1$$
- Bestimmung der Determinante von A:
$$\det A = (-1)^i \cdot \det A' = -1 \cdot 1 = -1$$

Beobachtung: Man kann zeigen, dass die Determinante einer Matrix A nach beliebigen Zeilen und beliebigen Spalten entwickelt werden kann:

- Entwicklung nach der k-ten Zeile:

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

- Entwicklung nach der l-ten Spalte:

$$\det A = \sum_{i=1}^{n} (-1)^{i+l} \cdot \det A_{il}$$

Beispiel:

$$\begin{vmatrix} 7 & 3 & \mathbf{0} & -1 \\ 2 & 4 & \mathbf{0} & 5 \\ 8 & -5 & \mathbf{2} & 4 \\ 2 & 1 & \mathbf{0} & 0 \end{vmatrix} = (-1)^{3+3} \cdot 2 \cdot \begin{vmatrix} 7 & 3 & -1 \\ 2 & 4 & 5 \\ \mathbf{2} & \mathbf{1} & \mathbf{0} \end{vmatrix}$$

$$= 2 \cdot \left(2 \cdot \begin{vmatrix} 3 & -1 \\ 4 & 5 \end{vmatrix} - 1 \cdot \begin{vmatrix} 7 & -1 \\ 2 & 5 \end{vmatrix} \right)$$

$$= 2 \cdot (2 \cdot 19 - 1 \cdot 37) = 2$$

Beobachtung: Die Determinanten einer Matrix A und der zu A transponierten Matrix A^t sind gleich:

$$\det A = \det A^t$$

2.9.3 Cramer'sche Regel

Regel: Sei $A \in M(n \times n, K)$ eine Matrix mit rg $A = n$ und $\vec{b} \in K^n$ ein Vektor. Hat das lineare Gleichungssystem aus A und $\vec{b}$ eine *eindeutige* Lösung, dann lässt sich diese Lösung folgermaßen bestimmen:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{mit} \quad x_i = \frac{\det A_i}{\det A}$$

Dabei ist A_i die Matrix, die man erhält, wenn man die i-te Spalte von A durch $\vec{b}$ ersetzt.

Beispiel: Es ist Lösung des folgenden linearen Gleichungssystem zu bestimmen:

$$\begin{array}{rcrcl} 2x_1 & + & 3x_2 & = & 5 \\ x_1 & - & 2x_2 & = & 2 \end{array} \quad \leftrightarrow \quad \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

Anwendung der Cramer'schen Regel:

$$x_1 = \frac{\begin{vmatrix} 5 & 3 \\ 2 & -2 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix}} = \frac{-10-6}{-4-3} = \frac{16}{7}$$

$$x_2 = \frac{\begin{vmatrix} 2 & 5 \\ 1 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix}} = \frac{4-5}{-4-3} = \frac{1}{7}$$

2.9.4 Anwendungen von Determinanten

Festlegung: Im Folgenden werden Punkte wie die Ortsvektoren dieser Punkte behandelt.

Definition: Es wird eine Funktion mydet : $K^2 \times K^2 \times K^2 \to K$ definiert:

$$\text{mydet}(\vec{p}, \vec{q}, \vec{r}) = \begin{vmatrix} p_x & q_x & r_x \\ p_y & q_y & r_y \\ 1 & 1 & 1 \end{vmatrix}$$

Beobachtung: Jeder Ortsvektor $\vec{t} \in \mathbb{R}^2$ zu einem Punkt T lässt sich eindeutig darstellen als Linearkombination der Ortsvektoren $\vec{p}$, $\vec{q}$ und $\vec{r}$ eines Dreiecks PQR:

$$\vec{t} = a \cdot \vec{p} + b \cdot \vec{q} + c \cdot \vec{r} \quad \text{mit} \quad a + b + c = 1$$

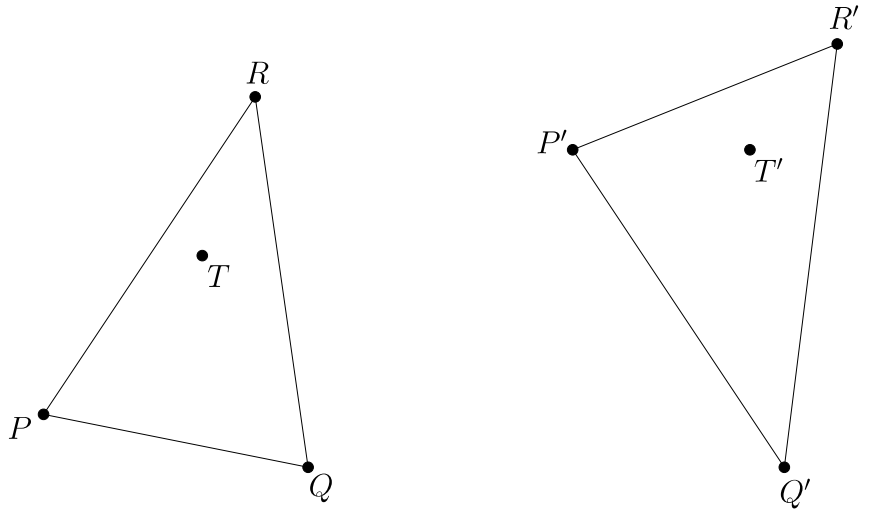
Die Zahlen a, b und c heißen die *baryzentrischen Koordinaten* oder auch *Schwerpunktskoordinaten*. Befinden sich nämlich die Massen a, b und c mit Gesamtmasse 1 an den Punkten P, Q und R, dann ist T der Schwerpunkt. Mit der Cramer'schen Regel werden a, b und c als Lösung eines linearen Gleichungssystem dann wiefolgt bestimmt:

$$a = \frac{\text{mydet}(\vec{t}, \vec{q}, \vec{r})}{\text{mydet}(\vec{p}, \vec{q}, \vec{r})}$$

$$b = \frac{\text{mydet}(\vec{p}, \vec{t}, \vec{r})}{\text{mydet}(\vec{p}, \vec{q}, \vec{r})}$$

$$c = \frac{\text{mydet}(\vec{p}, \vec{q}, \vec{t})}{\text{mydet}(\vec{p}, \vec{q}, \vec{r})}$$

Beobachtung: Mit Hilfe der baryzentrischen Koordinaten lässt sich in der Computergrafik *Warping* realisieren:



Beobachtung: Seien $\vec{p}, \vec{q}, \vec{r} \in \mathbb{R}^2$ Ortsvektoren der Punkte P, Q und R.

- Die Punkte P, Q und R liegen genau dann auf einer Linie, wenn

$$\text{mydet}(\vec{p}, \vec{q}, \vec{r}) = 0$$

- Der Punkt R liegt genau dann links von der gerichteten Geraden $\overrightarrow{PQ}$, wenn

$$\text{mydet}(\vec{p}, \vec{q}, \vec{r}) > 0$$

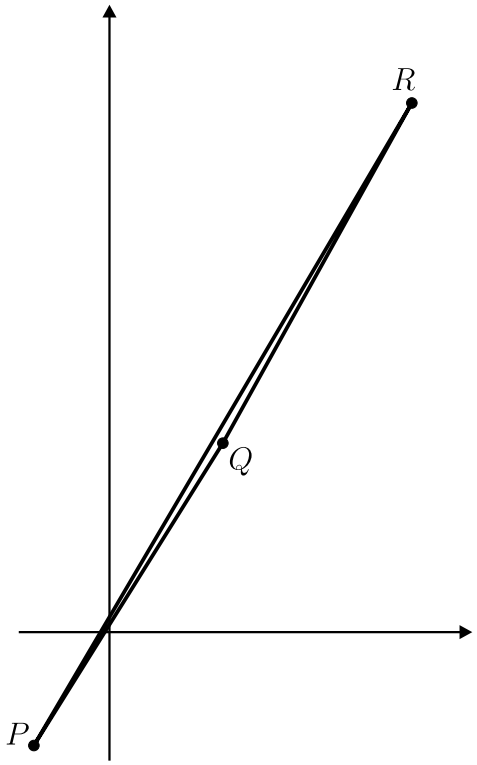
- Der Punkt R liegt genau dann rechts von der gerichteten Geraden $\overrightarrow{PQ}$, wenn

$$\text{mydet}(\vec{p}, \vec{q}, \vec{r}) < 0$$

- Die Fläche es von P, Q und R aufgespannten Dreiecks beträgt:

$$\left| \frac{\text{mydet}(\vec{p}, \vec{q}, \vec{r})}{2} \right|$$

Beispiel: Seien $P = (-2, -3)$, $Q = (3, 5)$ und $R = (8, 14)$ Punkte in $\mathbb{R}^2$:



Bestimmung von mydet:

$$\begin{vmatrix} -2 & 3 & 8 \\ -3 & 5 & 14 \\ 1 & 1 & 1 \end{vmatrix} = -10 + 42 - 24 - 40 + 9 + 28 = 5 > 0$$

Daraus folgt, dass $(8, 14)$ links von der gerichteten Geraden $\overrightarrow{(-2,-3)(3,5)}$ liegt und dass die Fläche des von den drei Punkten aufgespannten Dreiecks 2,5 beträgt.

Beobachtung: Diese Eigenschaften lassen sich auch auf höhere Dimensionen übertragen. Seien P, Q, R und S Punkte in $\mathbb{R}^3$, dann ist

$$\frac{1}{6} \cdot \begin{vmatrix} p_x & q_x & r_x & s_x \\ p_y & q_y & r_y & s_y \\ p_z & q_z & r_z & s_z \\ 1 & 1 & 1 & 1 \end{vmatrix}$$

das Volumen des von den vier Punkten aufgespannten Simplexes. Falls die vier Punkte auf einer Ebene liegen, ist der Wert 0.

Definition: Für $A \in M(n \times n, K)$ wird die *komplementäre Matrix* $\tilde{A} = (\tilde{a}_{ij})_{(i,j) \in n^2}$ definiert durch:

$$\tilde{a}_{ij} = (-1)^{i+j} \cdot \det A_{ji}$$

Dabei ist A_{ji} die Matrix, die man aus A durch Streichen der j-ten Zeile und der i-ten Spalte enthält.

Beobachtung: Man kann leicht nachrechnen, dass auf der Diagonalen von $A \cdot \tilde{A}$ immer $\det A$ steht, denn das is die Zeilenentwicklung von $\det A$. Mit etwas mehr Aufwand kann man zeigen, dass sonst nur Nullen in $A \cdot \tilde{A}$ auftreten:

$$A \cdot \tilde{A} = \begin{pmatrix} \det A & 0 & \cdots & \cdots & 0 \\ 0 & \det A & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \det A & 0 \\ 0 & \cdots & \cdots & 0 & \det A \end{pmatrix}$$

Folgerung: Ist $\det A \neq 0$, dann ist A invertierbar und

$$A^{-1} = \frac{\tilde{A}}{\det A}$$

Spezialfall: Für $A \in M(2 \times 2, K)$ gilt:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{für} \quad ad - bc \neq 0$$

Satz: Für alle $A, B \in M(n \times n, K)$ gilt:

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

Folgerung: Man kann jeden Endomorphismus $f : K^n \to K^n$ eindeutig seine Determinante zuordnen als $\det A$ für diese Matrix A von f bezüglich der Standardbasis.

Satz: Sei $f : K^n \to K^n$ ein Endomorphismus, A die zu f gehörende Matrix bezüglich der Standardbasis und B eine zu f gehörende Matrix bezüglich einer anderen Basis, dann gilt:

$$\det A = \det B$$

... Vorlesung vom 19.12.2002 (fehlt)

2.10 Euklidische Vektorräume

Definition: Sei V ein reeller Vektorraum. Ein Skalarprodukt über V ist eine Abbildung $< , >: V \times V \to \mathbb{R}$ mit folgenden drei Eigenschaften:

1. Biliniarität, d.h. für jedes $\vec{v} \in V$ sind die Abbildungen

   - $< , \vec{v} >: V \to \mathbb{R}$ mit $\vec{w} \mapsto < \vec{w}, \vec{v} >$ oder
   - $< \vec{v}, >: V \to \mathbb{R}$ mit $\vec{w} \mapsto < \vec{v}, \vec{w} >$

   sind linear.

2. Symmetrie, d.h. $< \vec{v}, \vec{w} >=< \vec{w}, \vec{v} >$ für alle $\vec{v}, \vec{w} \in V$.

3. Positive Definitheit, d.h. $< \vec{v}, \vec{v} >> 0$ für alle $\vec{v} \neq \vec{0}$.

Ein Euklidischer Vekttorraum ist ein reeller Vektorraum mit einem Skalarprodukt.

Beispiele:

1. $V = \mathbb{R}^n$, Standardskalarprodukt:

$$\langle (x_1, x_2, \ldots x_n), (y_1, y_2, \ldots y_n) \rangle = x_1 y_1 \cdot x_2 y_2 + \ldots + x_n y_n$$

$$\langle (1, 5, 0, 3), (3, 0, 7, -4) \rangle = 3 + 0 + 0 + (-12) = -9$$

2. $V = \{f : [-1, 1] \to \mathbb{R} \mid \text{stetige Funktion}\}$

$$\langle f, g \rangle := \int_{-1}^{1} f(x) g(x) dx$$

Definition: Die Norm eines Vektors $\vec{v}$ in einem Euklidischen Raum ist definiert durch

$$\|\vec{v}\| = \langle \vec{v}, \vec{v} \rangle$$

Beispiel:

$$\|(1, 2, 0, 2, 4)\| = \sqrt{1 + 4 + 0 + 4 + 16} = \sqrt{25} = 5$$

Das ist auch der Abstand zwischen $(0, 0, \ldots 0)$ und $(1, 2, 0, 2, 4)$ in $\mathbb{R}^5$.

Beispiel: in $\mathbb{R}^3$ ...

Satz (Ungleichung von Cauchy-Schwarz): In jedem Euklidschen Vektorraum gilt für alle $\vec{u}, \vec{v} \in V$

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\|$$

Speziell für $\mathbb{R}^n$ mit $\vec{u} = (a_1, a_2, \ldots a_n)$ und $\vec{v} = (b_1, b_2, \ldots b_n)$:

$$|a_1 b_1 + a_2 b_2 + \ldots + a_n b_b| \leq \sqrt{a_1^2, a_2^2, \ldots a_n^2} \cdot \sqrt{b_1^2, b_2^2, \ldots b_n^2}$$

Für $V = \{f : [-1, 1] \to \mathbb{R} \mid \text{stetige Funktion}\}$:

$$|\int_{-1}^{1} f(x)g(x)dx| \leq \sqrt{\int_{-1}^{1} (f(x))^2 dx} \cdot \sqrt{\int_{-1}^{1} (g(x))^2 dx}$$

Beweis:

1. Fall 1: $\vec{v} = \vec{0}$, dann ergibt die Cauchy-Schwarz-Ungleichung $0 = 0$ (also korrekt)
2. Fall 2: $\vec{v} \neq \vec{0}$

   $$\lambda := \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2}$$

   Nun betrachtet man:

   $$\begin{aligned} 0 &\leq \langle \vec{u} - \lambda \vec{v}, \vec{u} - \lambda \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} - \lambda \vec{v} \rangle - \lambda \langle \vec{v}, \vec{u} - \lambda \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} \rangle - \lambda \langle \vec{u}, \vec{v} \rangle - \lambda \langle \vec{v}, \vec{u} \rangle + \lambda^2 \langle \vec{v}, \vec{v} \rangle \\ &= \|\vec{u}\|^2 - 2 \cdot \frac{\langle \vec{u}, \vec{v} \rangle^2}{\|\vec{v}\|} + \frac{\langle \vec{u}, \vec{v} \rangle^2 \cdot \|\vec{v}\|^2}{\|\vec{v}\|^4} \\ &= \|\vec{u}\|^2 - \frac{\langle \vec{u}, \vec{v} \rangle^2}{\|\vec{v}\|^2} \end{aligned}$$

   Also:

   $$\langle \vec{u}, \vec{v} \rangle^2 \leq \|\vec{u}\|^2 \cdot \|\vec{v}\|^2 \quad \Rightarrow \quad |\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\|$$

Satz: Die Norm in einem Euklidischen Vektorraum hat die folgenden Eigenschaften:

1. $\|\vec{v}\| \geq 0$ für alle $\vec{v} \in V$
2. $\|\vec{v}\| = 0 \quad \Leftrightarrow \quad \vec{v} = \vec{0}$
3. $\|\lambda \vec{v}\| = |\lambda| \|\vec{v}\|$
4. $\|\vec{v} + \vec{u} \leq \|\vec{v}\| + \|\vec{u}\|\|$

Beweis:

1.

$$\begin{aligned}\|\lambda\vec{v}\| &= \sqrt{\langle\lambda\vec{v},\lambda\vec{v}\rangle}\\ &= \sqrt{\lambda\langle\vec{v},\lambda\vec{v}\rangle}\\ &= \sqrt{\lambda^2\langle\vec{v},\vec{v}\rangle}\\ &= \sqrt{\lambda^2}\cdot\sqrt{\langle\vec{v},\vec{v}\rangle}\\ &= |\lambda|\|\vec{v}\|\end{aligned}$$

2.

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

Beispiel: Normale Dreiecksungleichung aus der Geometrie:

GRAFIK

Winkel: Für $\vec{v},\vec{w}\in V$ definieren wir den Öffnungswinkel

$$\alpha(\vec{v},\vec{w}) = \arccos\frac{\langle\vec{v},\vec{w}\rangle}{\|\vec{v}\|\|\vec{w}\|}$$

GRAFIK mit $\vec{v}=(1,0)$ und $\vec{w}=(2,2)$

$$\begin{aligned}\sphericalangle(\vec{v},\vec{w}) &= \arccos\left(\frac{1\cdot 2+0\cdot 2}{\sqrt{1+0}+\sqrt{2^2}2^2}\right)\\ &= \arccos\left(\frac{2}{2\cdot\sqrt{2}}\right)\\ &= \frac{\pi}{4}=45^\circ\end{aligned}$$

In Kosinussatz die Formel einsetzen.

$$\begin{aligned}c^2 &= a^2+b^2-2ab\cos\varphi\\ c^2=\langle\vec{u}-\vec{v},\vec{u}-\vec{v}\rangle=\langle\vec{u},\vec{u}\rangle+\langle\vec{v},\vec{v}\rangle-2\langle\vec{u},\vec{v}\rangle&\\ &= \|\vec{u}\|^2+\|\vec{v}\|^2-2\|\vec{u}\|\|\vec{v}\|\cdot\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}\\ &= a^2+b^2-2ab\cos\varphi\\ \varphi &= \arccos\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}\end{aligned}$$

GRAFIK: Kosinus- $(0\to\pi)$ und Arcuskosinus-Funktion

Definition: Zwei Vektoren $\vec{u}$ und $\vec{v}$ in einem Euklidischen Vektorraum $(V, \langle\ ,\ \rangle)$ heißen orthogonal (senkrecht) zueinander, wenn $\langle \vec{u}, \vec{v} \rangle = 0$ ist.

$$\sphericalangle(\vec{u}, \vec{v}) = \frac{\pi}{2} \quad \Leftrightarrow \quad \langle \vec{u}, \vec{v} \rangle = 0 \quad \Leftrightarrow \quad \vec{u} \text{ steht senkrecht auf } \vec{v}$$

Schreibweise:

$$\vec{u} \perp \vec{v}$$

Für eine Teilmenge $M \subseteq V$ schreibt man $M \perp \vec{u}$, falls $\langle \vec{v}, \vec{u} \rangle = 0$ für alle $\vec{v} \in M$.

Definition: Das orthogonale Komplement $M^\perp$ einer Menge $M \subseteq V$ ist definiert als

$$M^\perp = \{\vec{u} \in V \mid M \perp \vec{u}\}$$

GRAFIK: einige Vektoren aus M (auf einer Linie) und einige aus $M^\perp$ (auch auf einer Linie, aber orthogonal zu den aus M)

Satz: Die Menge $M^\perp$ ist ein Unterraum.

Beweis:

- Der Nullvektoren gehört zu $M^\perp$.
- Addition:

$$\begin{aligned} \vec{u}, \vec{u'} \in M^\perp \quad &\Rightarrow \quad \langle \vec{u}, \vec{v} \rangle = \langle \vec{u'}, \vec{v} \rangle = 0 \quad \text{für alle } \vec{v} \in M \\ &= \quad \langle \vec{u} + \vec{u'}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u'}, \vec{v} \rangle = 0 \\ &\Rightarrow \quad \vec{u} + \vec{u'} \in M^\perp \end{aligned}$$

- Multiplikation: ...

Definition: Eine Menge von Vektoren $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_r$ wird Orthonormalsystem genannt, falls $\|\vec{v}_i\| = 1$ für alle $i = 1, 2, \ldots r$ und $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ für alle $i \neq j$.
kürzer:

$$\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{i\,j} = \begin{cases} 1 & \text{falls} \quad i = j \\ 0 & \text{sonst} \end{cases}$$

Beispiel: Standardbasis für 4D.

Lemma 1: Die Vektoren eines Orthogonalsystem sind linear unabhängig.

Beweis: Angenommen $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_r\vec{v}_r = \vec{0}$.
Zeige: $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$

$$\begin{aligned} 0 &= \langle\vec{0}, \vec{v}_i\rangle \\ &= \langle\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_r\vec{v}_r, \vec{v}_i\rangle \\ &= \underbrace{\lambda_1\langle\vec{v}_1, \vec{v}_i\rangle}_{=0} + \ldots + \underbrace{\lambda_i\langle\vec{v}_i, \vec{v}_i\rangle}_{=1} + \ldots + \underbrace{\lambda_r\langle\vec{v}_r, \vec{v}_i\rangle}_{=0} \\ &= \lambda_i \end{aligned}$$

Lemma 2: Ist $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$ eine orthonormale Basis von V, so gilt für jedes $\vec{v} \in V$ die folgende Entwicklungsformel:

$$\vec{v} = \sum_{i=1}^{n} \langle\vec{v}, \vec{v}_i\rangle\vec{v}_i$$

Beweis: Nachrechnen

Beispiel: $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$

$$\begin{pmatrix}2\\3\\0\end{pmatrix} = (2+0+0)\begin{pmatrix}1\\0\\0\end{pmatrix} + (0+3+0)\begin{pmatrix}0\\1\\0\end{pmatrix} + (0+0+0)\begin{pmatrix}0\\0\\1\end{pmatrix}$$

Lemma 3: Ist $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_r$ ein Orthonormalsystem in V und $U = \mathrm{Lin}(\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_r\})$, so hat jedes $\vec{v} \in V$ eine eindeutige Darstellung

$$\vec{v} = \vec{u} + \vec{w} \text{ mit } \vec{u} \in U \text{ und } \vec{w} \in U^{\perp}$$

Dabei ist

$$\vec{u} = \sum_{i=0}^{r} \langle\vec{v}, \vec{v}_i\rangle\vec{v}_i$$

und

$$\vec{w} = \vec{v} - \vec{u}$$

Beispiel: $V = \mathbb{R}^2$, $r = 1$, $\vec{v}_1 = \begin{pmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{pmatrix}$, $\|\vec{v}_1\| = 1$

$\vec{v} = \begin{pmatrix}3\\2\end{pmatrix} = \vec{u} + \vec{w}$ wobei $\vec{u} \in \mathrm{Lin}(\vec{v}_1)$ und $\vec{w} \perp \vec{v}_1$

GRAFIK: U mit $\vec{v}_1$ und $U^\perp$, außerdem $\vec{v}$, $\vec{v}$ wird auf U und $U^\perp$ projiziert

$$\begin{aligned}
\vec{u} &= \langle \vec{v}, \vec{v}_1 \rangle \vec{v}_1 \\
&= \frac{5\sqrt{2}}{2} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \end{pmatrix} \\
\vec{w} &= \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}
\end{aligned}$$

Satz (Erhard Schmidt'sches Orthonormalisierungsverfahren): Sei $\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_r\}$ linear unabhängig, dann bilden die Vektoren

$$\begin{aligned}
\tilde{v}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} \\
\tilde{v}_{k+1} &= \frac{\vec{v}_{k+1} - \sum_{i=1}^{k} \langle \vec{v}_{k+1}, \tilde{v}_i \rangle \tilde{v}_i}{\left\| \vec{v}_{k+1} - \sum_{i=1}^{k} \langle \vec{v}_{k+1}, \tilde{v}_i \rangle \tilde{v}_i \right\|} \quad \text{für } k = 1, 2, \dots r-1
\end{aligned}$$

ein Orthonormalsystem mit den Eigenschaften, dass

$$\mathrm{Lin}(\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_i\}) = \mathrm{Lin}(\{\tilde{v}_1, \tilde{v}_2, \dots \tilde{v}_i\})$$

für $i = 1, 2, \dots r$.

Beispiel: $\vec{v}_1 = \begin{pmatrix}2\\2\\0\end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix}-3\\-1\\0\end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix}1\\2\\3\end{pmatrix}$

$$\begin{aligned}
\tilde{v}_1 &= \frac{1}{\sqrt{8}}\begin{pmatrix}2\\2\\0\end{pmatrix} = \begin{pmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\\0\end{pmatrix}\\
\tilde{v}_2 - \langle \vec{v}_2, \tilde{v}_1\rangle \tilde{v}_1 &= \begin{pmatrix}-3\\-1\\0\end{pmatrix} - (-2\sqrt{2})\begin{pmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\\0\end{pmatrix}\\
&= \begin{pmatrix}-3\\-1\\0\end{pmatrix} + \begin{pmatrix}2\\2\\0\end{pmatrix}\\
&= \begin{pmatrix}-1\\1\\0\end{pmatrix}\\
\tilde{v}_2 &= \frac{\begin{pmatrix}-1\\1\\0\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\\0\end{pmatrix}\right\|}\\
&= \begin{pmatrix}\frac{-\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\\0\end{pmatrix}
\end{aligned}$$

$$\begin{pmatrix}1\\2\\3\end{pmatrix} - \frac{3\sqrt{2}}{2}\tilde{v}_1 - \frac{\sqrt{2}}{2}\tilde{v}_2$$

$$\begin{aligned}
&= \begin{pmatrix}1\\2\\3\end{pmatrix} - \begin{pmatrix}\frac{3}{2}\\\frac{3}{2}\\0\end{pmatrix} - \begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\\0\end{pmatrix}\\
&= \begin{pmatrix}\cdot\\\cdot\\\cdot\end{pmatrix}\\
\tilde{v}_3 &= \ldots \text{Normieren}
\end{aligned}$$

Definition: Orthogonale Projektion von $\vec{v}$ in den Unterraum $U = \text{Lin}(\vec{v}_1, \vec{v}_2, \ldots \vec{v}_r)$

$$P_U(\vec{v}) = \langle \vec{v}, \vec{v}_1\rangle \cdot \vec{v}_1 + \ldots + \langle \vec{v}, \vec{v}_r\rangle \cdot \vec{v}_r \in U$$

$\vec{w} := \vec{v} - P_U(\vec{v})$, man kann zeigen, dass $\vec{w} \perp U$, d.h. $\vec{w} \in U^\perp$, folglich ist $P_U(\vec{w}) = \vec{0}$

GRAFIK

Die Orthonormalprojektion $P_U : V \to U$ hat die folgenden zwei Eigenschaften:

- P_U beschränkt auf U ist die identische Abbildung

Beispiel: $U = \text{Lin}\left(\begin{pmatrix}2\\1\end{pmatrix}\right)$ in $V = \mathbb{R}^2$

Aufgabe: Berechne die Matrix der Projektion P_U.

GRAFIK: Gerade $y = \frac{1}{2}x$ mit Projektion der Basisvektoren $\begin{pmatrix}1\\0\end{pmatrix}$ und $\begin{pmatrix}0\\1\end{pmatrix}$ auf die Gerade

1. Orthonormalbasis für U

$$\begin{aligned}\vec{v}_1 &= \begin{pmatrix}2\\1\end{pmatrix}\\ \tilde{v}_1 &= \frac{\begin{pmatrix}2\\1\end{pmatrix}}{\left\|\begin{pmatrix}2\\1\end{pmatrix}\right\|} = \frac{\begin{pmatrix}2\\1\end{pmatrix}}{\sqrt{5}} = \begin{pmatrix}\frac{2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{pmatrix}\end{aligned}$$

2. Projektion

$$P_U(\vec{v}) = \langle \vec{v}, \tilde{v}_1 \rangle \cdot \tilde{v}_1$$

Setze für $\vec{v}$ die Basisvektoren $\vec{e}_1$ und $\vec{e}_2$ ein:

$$\begin{aligned}P_U\left(\begin{pmatrix}1\\0\end{pmatrix}\right) &= \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}\frac{2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{pmatrix} \right\rangle \cdot \begin{pmatrix}\frac{2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{pmatrix} = \begin{pmatrix}\frac{4}{5}\\ \frac{2}{5}\end{pmatrix}\\ P_U\left(\begin{pmatrix}0\\1\end{pmatrix}\right) &= \left\langle \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}\frac{2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{pmatrix} \right\rangle \cdot \begin{pmatrix}\frac{2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{pmatrix} = \begin{pmatrix}\frac{2}{5}\\ \frac{1}{5}\end{pmatrix}\end{aligned}$$

3. Matrix:

$$\begin{pmatrix}\frac{4}{5} & \frac{2}{5}\\ \frac{2}{5} & \frac{1}{5}\end{pmatrix} = \begin{pmatrix}0{,}8 & 0{,}4\\ 0{,}4 & 0{,}2\end{pmatrix}$$

2.11 Affiner Raum (intuitiver Zugang)

Motivation: Anwendung des Skalarprodukts in der affinen Geometrie

Mengen:

- V Vektorraum (hier über $\mathbb{R}$, d.h. $V = \mathbb{R}^2, \mathbb{R}^3$)
- A Punktmenge

Operationen:

- Punkt + Vektor $\mapsto$ Punkt
- Punkt $-$ Punkt $\mapsto$ Vektor

GRAFIK: zwei Punkte p und q und ein Verbindungsvektor $\vec{v}$ mit

$$\begin{aligned} \vec{q} &= p + \vec{v} \\ \vec{v} &= q - p \end{aligned}$$

Eigenschaft:

$$p + (\vec{v} + \vec{w}) = (p + \vec{v}) + \vec{w}$$

Standardmodell für A ist V selbst.

affiner Unterraum: $U \subseteq V$ Untervektorraum, $p \in A$

$$p + U = \{p + \vec{u} \mid \vec{u} \in U\}$$

Beispiele:

- affine Unterräume in $A = \mathbb{R}^2$:
  - Punkte (dim $U = 0$)
  - Geraden (dim $U = 1$)
  - ganz $\mathbb{R}^2$ (dim $U = 2$)
- affine Unterräume in $A = \mathbb{R}^3$:
  - Punkte (dim $U = 0$)
  - Geraden (dim $U = 1$)
  - Ebenen (dim $U = 2$)
  - ganz $\mathbb{R}^3$ (dim $U = 3$)

Geradengleichungen: $V = A = \mathbb{R}^2$

1. Gerade durch einen Punkt und parallel zu einem Vektorunterraum
   - $U \subseteq V$ ist ein 1-dimensionaler Vektorunterraum $U = \{\lambda\vec{u} \mid \lambda \in \mathbb{R}\}$
   - $p \in A$

Gerade durch p parallel zu U in Parameterdarstellung:

$$L = \{p + \lambda\vec{u} \mid \lambda \in \mathbb{R}\}$$

2. Gerade durch zweu Punkte:

   - $p = (x, y)$
   - $p' = (x', y')$

   Schreibweisen:

$$\begin{aligned} L &= \{p + \lambda(p' - p) \mid \lambda \in \mathbb{R}\} \\ &= \{q = (x_q, y_q) \mid x_q = x + \lambda(x' - x), y_q = y + \lambda(y' - y), \lambda \in \mathbb{R}\} \\ &= \{q = (x_q, y_q) \mid \underbrace{x_q(y' - y) - x(y' - y)}_{=\lambda(x'-x)(y'-y)} = \underbrace{y_q(x' - x) - y(x' - x)}_{=\lambda(x'-x)(y'-y)}\} \\ &= \{q = (x_q, y_q) \mid ax_q + by_q = c\} \qquad \text{(Koordinantendarstellung von } L\text{)} \end{aligned}$$

   mit $a = y' - y$, $b = x' - x$ und $c = -y(x' - x) + x(y' - y)$

HNS:

GRAFIK: Gerade L und parallele Gerade U_L durch den Ursprung und Normalenvektor $\vec{n}$, außerdem $P_{U_L}(q - 0)$ und $\vec{w} = d \cdot \vec{n}$

- Normalenvektor $\vec{n}$ von L ist senkrecht zu U_L und $\|\vec{n}\| = 1$
- Abstand d von $(0, 0)$ zu L, d.h. $(0, 0) + d \cdot \vec{n} \in L$

Hesse-Normalform von L:

$$L = \{q \mid \langle q - 0, \vec{n}\rangle = d\}$$

Das heißt:

$$\begin{aligned} \langle q - 0, \vec{n}\rangle &= \langle P_{U_L}(q - 0), \vec{n}\rangle + \langle \vec{w}, \vec{n}\rangle \\ &= 0 + \langle d\vec{n}, \vec{n}\rangle \\ &= d\langle \vec{n}, \vec{n}\rangle \\ &= d \end{aligned}$$

Bestimmung der Hesse-Normalform aus der Parameterform:

$$L = \{p + \lambda\vec{v} \mid \lambda \in \mathbb{R}\}$$

Aufgabe: Bestimme $\vec{n}$ und d!

1. $\tilde{v} = \frac{\vec{v}}{\|\vec{v}\|}$ Orthonormalbasis von U_L
2. $\vec{n}' = d\vec{n} = (p - 0) - P_{U_L}(p - 0) = (p - 0) - \langle (p - 0), \tilde{v}\rangle \cdot \tilde{v}$
3. $d = \|\vec{n}'\|$ und $\vec{n} = \frac{\vec{n}'}{d}$

... Vorlesung vom 16.02.2002 (fehlt)

... Vorlesung vom 21.02.2002 (fehlt)

Kapitel 3

Endliche Körper und Codierungstheorie

3.1 Restklassenarithmetik

3.1.1 ...

... Vorlesung vom 21.02.2002 (fehlt)

$$\begin{aligned} 252 &= 1 \cdot 138 + 54 \\ 158 &= 3 \cdot 53 + 36 \\ 54 &= 1 \cdot 36 + 18 \\ 36 &= 2 \cdot 18 + 0 \end{aligned}$$

Die Umkehrung des Euklidischen Algorithmus liefert eine Darstellung des $\mathrm{ggT}(a, b)$ als Linearkombination aus a und b mit ganzzahligen Koeffizienten.

$$\begin{aligned} 18 &= 54 - 1 \cdot 36 \quad (36 = 198 - 3 \cdot 54) \\ &= 54 - (198 - 3 \cdot 54) \\ &= 4 \cdot 54 - 198 \quad (54 = 252 - 198) \\ &= 4 \cdot (252 - 198) - 198 \\ &= 4 \cdot 252 - 5 \cdot 198 \end{aligned}$$

Satz: Sind $a, b \in \mathbb{Z}^+$, dann existieren $r, s \in \mathbb{Z}$, so dass $\mathrm{ggT}(a, b) = r \cdot a + s \cdot b$.

Satz: Seien m und a zwei positive, teilerfremnde Zahlen, dasnn gibt es genau ein $b \in \{1, 2, \ldots m-1\}$, so dass $a \cdot b \equiv 1 (\mathrm{mod} m)$.

Beweis: $\mathrm{ggT}(a, m) = 1 = r \cdot a + s \cdot m$ für geeignete $r, s \in \mathbb{Z}$. Setzen $b := r \mod m \in \{\emptyset, 1, \ldots m-1\}$.

- $b \equiv r (\mathrm{mod} m)$
- $a \equiv a (\mathrm{mod} m)$
- $0 \equiv s \cdot m (\mathrm{mod} m)$
-
- $a \cdot b \equiv r \cdot a (\mathrm{mod} m)$
- $a \cdot b + 0 \equiv r \cdot a + s \cdot m (\mathrm{mod} m) \equiv 1 (\mathrm{mod} m)$

Eindeutigkeit: Angenommen $a \cdot b \equiv a \cdot c \equiv 1 (\mathrm{mod} m)$ und $0 < c \leq b \leq m - 1$

$$a \cdot (b - c) \equiv \underbrace{1 - 1}_{=0} (\mathrm{mod} m)$$

$a \cdot (b - c)$ ist durch m teilbar $\Rightarrow (b - c)$ is durch m teilbar und $0 \leq b - c < m - 1$ $\Rightarrow b - c = 0 \Rightarrow b = c \Rightarrow$ Eindeutigkeit

Folgerung: Ist p eine Primzahl und $a \in \{1, 2, \ldots p-1\}$, dann gibt es ein eindeutiges $b \in \{1, 2, \ldots p-1\}$, so dass $a \cdot b \equiv 1 (\mathrm{mod} p)$. b wird die zu a inverse Zahl bezüglich p genannt.

Folgerung: Die Zahlen $\{0, 1, \ldots p-1\}$, wobei p Primzahl, bilden mit der Addition und Multiplikation modulo p einen Körper. Dieser Körper wird mit $\mathbb{Z}_p$ oder mit GF(p) bezeichnet.

Beispiel: $\mathbb{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$:

$$\begin{aligned} 4+5 &= 2 \\ 4 \cdot 4 &= 2 \end{aligned}$$

$$\begin{aligned} 1 \cdot 1 &= 1 \\ 6 \cdot 6 &= 1 \\ 2 \cdot 8 &= 1 \\ 3 \cdot 5 &= 1 \end{aligned}$$

$$\begin{aligned} 3 \cdot x &= 4 \quad | \cdot 3^{-1} \\ \underbrace{5 \cdot 3}_{=1} \cdot x &= 5 \cdot 4 = 6 \end{aligned}$$

Beispiel: Löse $7 \cdot x = 5$ in $\mathbb{Z}_{17}$.
Inverses zu 7 mod 17:

$$\begin{aligned} 17 &= 2 \cdot 7 + 3 \\ 3 &= 17 - 2 \cdot 7 \\ 7 &= 2 \cdot 3 + 1 \\ 1 &= 7 - 2 \cdot 3 = 7 - 2 \cdot (17 - 2 \cdot 7) = \mathbf{5} \cdot 7 - 2 \cdot 17 \\ 7^{-1} &= (5 \mod 17) = 5 \end{aligned}$$

$$\begin{aligned} 7 \cdot x &= 5 \quad | \cdot 5 \\ 1 \cdot x &= 5 \cdot 5 = 8 \end{aligned}$$

Chinesischer Restklassensatz: Seien $m_1, m_2, \ldots m_n \in \mathbb{Z}^+$ paarweise teilerfremd und $m = m_1 \cdot m_2 \cdot m_n$, dann gibt es für beliebige $a_1, a_2, \ldots a_n \in \mathbb{Z}$ eine Zahl

$x \in \{0, 1, \ldots m-1\}$, so dass die folgenden Komvergenzen erfüllt sind:

$$\begin{aligned} x &\equiv a_1 (\mathrm{mod} m_1) \\ x &\equiv a_2 (\mathrm{mod} m_2) \\ &\vdots \\ x &\equiv a_n (\mathrm{mod} m_n) \end{aligned}$$

Beispiel: $m_1 = 99$, $m_2 = 100$, $m_3 = 101$, $a_1 = 80$, $a_2 = 63$, $a_3 = 27$

Beweis:

1. Finde Zahlen, die $\equiv 1 (\mathrm{mod} m_i)$ und $\equiv 0 (\mathrm{mod} m_j)$ mit $j \neq i$ sind

$$\begin{aligned} M_1 &= \frac{m}{m_1} = m_2 \cdot m_3 \cdot \ldots \cdot m_n \\ M_2 &= \frac{m}{m_2} = \ldots \\ &\vdots \\ M_n &= \frac{m}{m_n} = \ldots \end{aligned}$$

   $\mathrm{ggT}(m_k, M_k) = 1$ d.h. $\exists y_k \equiv 1 (\mathrm{mod} m_k)$

$$M_k \cdot y_k \equiv 1 \pmod{m_k}$$

   Das sind diese Zahlen, denn $M_k \cdot y_k$ is duch jedes m_l $(l \neq k)$ teilbar, d.h. $M_k \cdot y_k \equiv 0 (\mathrm{mod} m_l)$.

   Setze

$$x = (a_1 \cdot y_1 \cdot M_1 + \ldots + a_n \cdot y_n \cdot M_n) \mod m$$

   Prüfe, dass alle Kongruenzen erfüllt sind.

Kleiner Satz von Fermat: Ist p eine Primzahl, dann gilt für jede nicht durch p teilbare Zahl $a \in \mathbb{Z}$:

$$a^{p-1} \equiv 1 \pmod{p}$$

Beispiele:

- Wähle $p = 7$ und $a = 2$:

$$\begin{aligned} a^6 &= 64 \\ 64 &\equiv 1 \pmod{7} \end{aligned}$$

- Wähle $p = 7$ und $a = 3$:

$$\begin{aligned} a^6 &= 729 \\ 729 &\equiv 1 \pmod{7} \end{aligned}$$

3.1.2 RSA-Kryptosysteme

Teilnehmen (party) verschicken Nachrichten $m \in \{0,1\}^n$:

Alice		Bob
m		$D(E(m)) = m$
$\downarrow$		$\downarrow$
$E(m) \in \{0,1\}^{l(n)}$	$\rightsquigarrow$	$E(m)$
	$\downarrow$	
Gegener (adversary) Eve	$E(m)$	
	$\downarrow$	?
	m	

Rivers, Shamir, Adlerman 78

- p, q zwei große Primzahlen
- $n = p \cdot q$
- $e \in \mathbb{Z}$: ggT$(e, (p-1)(q-1)) = 1$
- $d \in \mathbb{Z}$: $d \cdot e \equiv 1 \pmod{(p-1)(q-1)}$

1. Bob gibt n, e bekannt (p, q, d geheim)
2. Alice verschlüsselt Nachricht m ($|m| < \log n$):
$$m' = E(m) := m^e \mod n$$
3. Bob entschlüsselt m':
$$D(m') := m'^d \mod n$$

Behauptung:
$$D(E(m)) = m$$

Beweis:
$$d \cdot e = 1 + k \cdot (p-1)(q-1)$$

Zu zeigen:
$$(m^e)^d = m^{1+k(p-1)(q-1)} \equiv m \pmod{n}$$

genügt zu zeigen (Chinesicher Restsatz):

$$\begin{aligned} (m^e)^d &\equiv m \pmod{p} \\ \text{und } (m^e)^d &\equiv m \pmod{q} \end{aligned}$$

1. Fall 1:

$$m \equiv 0 \pmod{p}$$

   (trivial)

2. Fall 2:

$$\begin{aligned}(m^e)^d &\equiv m \cdot (m^{p-1})^{(q-1)\cdot k} \\ &\equiv m \cdot 1^{(p-1)\cdot k} \pmod{p}\end{aligned}$$

$$(m^d)^d \equiv m \pmod{p}$$

Authetisierung mit RSA:

- 2'): Alice schickt Zufallsstring m
- 3'): Bob schickt $m' = D(m)$ zurück
- 4'): Alice überprüft $E(m') = m$?

Vorteile:

- Parameter leicht zu erzeugen (randomisierter Primzahltest)
- Ver- und Entschlüsselung leicht zu berechnen (Spezialchip)
- sicher in der Praxis (unter Einhaltung bestimmter Regeln)

Nachteile:

- nur sicher, wenn es *keine* effiziente Algorithmen zur Faktorisierung gibt
- möglicherweise auch ohne Faktorisierungsalgorithmus zu brechen
- mit *Quantencomputern* ist Faktorisierung in Polynomialzeit möglich (P. Shor)

Aufgabe: Kryptographie auf Grundlage NP-schwerer Probleme.

3.2 Grundbegriffe der Codierungstheorie

Corierungstheorie – Verschlüsselung von Informationen unter den folgenden Aspekten:

1. Codierung soll helfen, eine Information geheim zu halten.

2. Codierung soll so kurz wie möglich sein

3. Fehler in der Übertragung sollten erkanntund korrigiert werden.

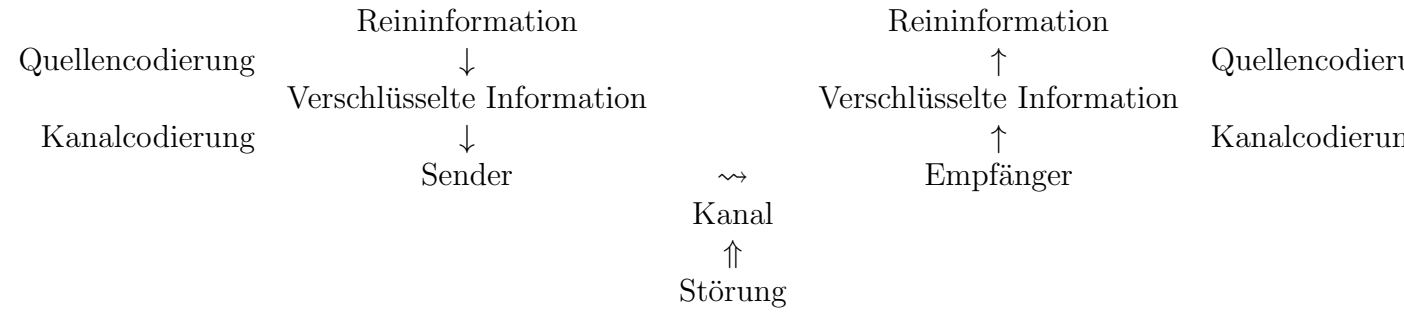
$\rightsquigarrow$ Teilgebiete:

1. Kryptographie (RSA, Einwegfunktionen, Pseudozufallsgeneratoren, ...)

2. Datenkompression (Hufman, ...)

3. Fehlerkorrigierende Codes (Hamming Code, linearer Code, ...)

Modell:



Kanalalphanet: Q, $|Q| = q$, häufig $Q = \{0, 1\}$, $q = 2$

Definition: Eine (Kanal-)Codierung ist eine injektive Funktion $c : I \to Q^n$, wobei I eine Informationsmenge ist (z.B. Alphabet oder bereits Menge der Codewörter aus einer Quellcodierung). Das Bild $C = \text{Im } c$ wird ein Code genannt. Hier besteht der Code nur aus Wörtern gleicher Länge, das nennt man einen Blockcode.

Definition: Seinen $v = (v_1, v_2, \ldots v_n)$ und $w = (w_1, w_2, \ldots w_n) \in Q^n$. Wir definieren den Hamming-Abstand der Worte v und w als Anzahl der Stellen, an denen sie sich unterscheiden:

$$d(v, w) = |\{i \mid 1 \leq i \leq n \text{ und } v_i \neq w_i\}|$$

Beispiel:

$$d((0,1,\mathbf{0},\mathbf{1},\mathbf{0},\mathbf{0},1), (0,1,\mathbf{1},\mathbf{0},\mathbf{1},\mathbf{1},1)) = 4$$

Beobachtung: Der Hamming-Abstand hat alle Eigenschaften einer Abstandsfunktion, d.h. für alle $u, v, w \in Q^n$ gilt:

1. $d(u, v) \geq 0$ und $d(u, v) = 0 \Leftrightarrow u = v$

2. $d(u, v) = d(v, u)$

3. $d(u, v) + d(v, w) \geq d(u, w)$ (Dreiecksungleichung)

Definition: Die Minimalabstand eines Codes $C \subseteq Q^n$ ist

$$d(C) := \min(\{d(c, c') \mid c \neq c', c, c' \in C\})$$

Wir verwenden c, c', c_1, c_2 für Codewörter und allgemein u, w, v für Wörter aus Q^n.

Beispiel:

$$\begin{aligned} c : \{a, b, c, d\} &\to Q^3 \\ a &\mapsto (0,0,0) \\ b &\mapsto (0,1,1) \\ c &\mapsto (1,0,1) \\ d &\mapsto (1,1,0) \end{aligned}$$

$$C = \{(0,0,0), (0,1,1), (1,0,1), (1,1,0)\}$$
$$d(C) = 2$$

Prinzip: Wird ein Wort $w \in Q^n$ empfangen, so sucht man ein (oder besser das) Codewort c mit minimalem Abstand zu w.
Wann ist c richtig?
Empfangen $w \in Q^n$:

1. $w \notin C$, dann ist ein Fehler aufgetreten
2. Wenn wir wissen, dass höchstens 1 Fehler aufgetreten ist, und $c \in C$ ist das einzige Codewort mit $d(w, c) = 1$, dann ist c das ursprüngliche Codewort.

Definition: Ein Code C ist k-fehlererkennend, wenn bei jedem empfangengenen Wort w, das $\leq k$ Fehler, enthält, erkannt wird, *ob* Übertragungsfehler aufgetreten sind.

Definition: Ein Code C ist k-fehlerkorrigierend, wenn bei jedem empfangenen Wort, das $\leq k$ Fehler erhält, die Fehler korrigiert werden könnn, d.h. dass das ursprüngliche Codewort bestimmt werden kann.

Definition: Für $v \in Q^n$ definieren wir die Kugel mit Radius t um v durch

$$B_t(v) = \{w \in Q^n \mid d(v, w) \leq t\}$$

GRAFIK: von Henning übertragen

Satz: C ist k-fehlerkorrigierend genau dann, wenn $\forall c \neq c' \in C\ B_k(c) \cap B_k(c') = \emptyset$ genau dann, wenn $d(C) \geq 2k+1$ (Minimalabstand von C)

Beweis:

- Erste Äquivalenz: bei der Üertragung von c treten $\leq k$ Fehler auf, dann liegt empfangenes Wort $w \in B_k(c)$, d.h. w gehört eindeutig zu c.
- Zweite Äquivalenz:

GRAFIK: von Henning übertragen

Satz: C ist k-fehlererkennend genau dann, wenn $\forall c \in C\ B_k(c) \cap (C \backslash \{c\}) = \emptyset$ genau dann, wenn $d(C) \geq k+1$.

Beweis: wie oben

Beispiele: Einfache Konstruktion mit Paritätsbits und Mehrfachcodierung:

$$I = Q^m \quad \text{mit} \quad Q = \{0,1\}$$

1. Paritätsbit: $c_{\text{par}} : Q^m \to Q^{m+1}$

   $$c_{\text{par}}(v_1, v_2, \ldots v_m) = (\underbrace{v_1, v_2, \ldots v_m, p}_{\text{gerade Anzahl von 1}}) \quad \text{mit} \quad p = v_1 + v_2 + \ldots + v_m \pmod 2$$

   Daraus folgt: $C_{\text{par}} = \text{Im}\ c_{\text{par}}$ hat den Minimalabstand 2, d.h. C_{par} ist 1-fehlererkennend.

2. Doppelcodierung: $c_2 : Q^m \to Q^{2m}$

   $$c_2(v_1, v_2, \ldots v_m) = (v_1, v_2, \ldots v_m, v_1, v_2, \ldots v_m)$$

   Daraus folgt: $C_2 = \text{Im}\ c_2$ hat den Minimalsbstand 2, d.h. C_2 ist 1-fehlererkennend.

3. Dreifachcodierung: $c_3 : Q^m \to Q^{3m}$

   $$c_3(v_1, v_2, \ldots v_m) = (v_1, v_2, \ldots v_m, v_1, v_2, \ldots v_m, v_1, v_2, \ldots v_m)$$

   Daraus folgt: $C_3 = \text{Im}\ c_3$ hat den Minimalsbstand 3, d.h. C_3 ist 1-fehlerkorrigierend.

4. Doppelcodierung mit Paritätsbit: $c_{2+\text{par}} : Q^m \to Q^{2m+1}$

   $$c_{2+\text{par}}(v_1, v_2, \ldots v_m) = (v_1, v_2, \ldots v_m, v_1, v_2, \ldots v_m, p) \quad \text{mit} \quad p = v_1 + v_2 + \ldots v_m \pmod 2$$

   Daraus folgt: $C_{2+\text{par}} = \text{Im}\ c_{2+\text{par}}$ hat den Minimalabstand 3:

- Fall 1: $d(v,w) = 1 \quad (v,w \in Q^m) \Rightarrow d(c_{2+\text{par}}(v), c_{2+\text{par}}(w)) = \underbrace{1+1}_{d(v,w)} + \underbrace{1}_{p}$
- Fall 2: $d(v,w) \geq 2 \quad (v,w \in Q^m) \Rightarrow d(c_{2+\text{par}}(v), c_{2+\text{par}}(w)) \geq 2+2 = 4$

d.h. $C_{2+\text{par}}$ ist 1-fehlerkorrigierend.

5. Kreuzsicherungscode $m = l^2$: $c_{\text{kr}} : Q^m \to Q^{m+2l}$

   Stelle Elemente von Q^m in einer quadratischen Matrix dar und gib für jede Spalte und für jede Zeile das Paritätsbit dazu:

$$\begin{pmatrix} v_1 & v_2 & \cdots & v_l \\ v_{l+1} & v_{l+2} & \cdots & v_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ v_{(l-1)+l+1} & v_{(l-1)+l+2} & \cdots & v_{l^2} \end{pmatrix}$$

   (Bemerkung: Um Paritätsbits in den Zeilen und Spalten erweitern (Zeilen: p_i, Spalten $\bar{p}_j$))

   $C_{\text{kr}} = \text{Im } c_{\text{kr}}$ hat Minimalabstand 3, d.h. ist 1-fehlerkorrigierend.

   $v, w \in Q^m$

   - $d(v,w) \geq 3 \Rightarrow d(c_{\text{cr}}(v), c_{\text{cr}}(w)) \geq 3$
   - $d(v,w) = 2$ dann liegen die zwei Unterschiede in verschiedenen Zeilen i und j oder in verschiedenen Spalten k, k'

$$d(\ldots) \geq 2 + \underbrace{1}_{p_i} + \underbrace{1}_{p_j}$$

   - $d(v,w)1$, sei Unterschied in Zeile i und Spalte j

$$d(\ldots) = \underbrace{1}_{d(u,v)} + \underbrace{1}_{\bar{p}_i} + \underbrace{1}_{p_j} = 3$$

Definition: Die Informationsrate eines Codes $C \subseteq Q^n$ ist der Quotient

$$\frac{\log_q |C|}{n}$$

Das beschreibt das Verhältnis der Längen des Infomationsworts und des Codeworts.

Beispiele:

1. Für $C_{2+\text{par}}$:
$$\frac{m}{2m+1} \lesssim \frac{1}{2}$$
2. Für C_{kr}:
$$\frac{m + 2 \cdot \sqrt{m} - 2\sqrt{m}}{m + 2 \cdot \sqrt{m}} = 1 - \frac{2 \cdot \sqrt{m}}{m + 2 \cdot \sqrt{m}} \approx 1 - \frac{2}{\sqrt{m}}$$

Fragen:

1. Geht es noch besser?
2. Wie korrigiert man 2 und noch mehr Fehler?

Idee für eine Verbesserung (Hamming):

$$v = (v_1, \ldots v_4) \quad Q = \{0, 1\} \quad + 3 \text{ Redundanzbits}$$

$$\begin{aligned} r_1 &= v_2 + v_3 + v_4 \pmod 2 \\ r_2 &= v_1 + v_3 + v_4 \pmod 2 \\ r_3 &= v_1 + v_2 + v_4 \pmod 2 \end{aligned}$$

Minimalabstand 3
Codierung durch Matrixmultiplikation über $\mathbb{Z}_2$:

$$c(v) = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}}_{\text{Generatormatrix}} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ r_1 \\ r_2 \\ r_3 \end{pmatrix}$$

Zur Decodierung mit 1-Fehlerkorrektur verwendet man die folgede Prüfmatrix:

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Berechnung:

$$H \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} v_1 + v_3 + r_1 + r_3 \\ \dots \\ \dots \end{pmatrix}$$

Daraus folgt:

- Fehlererkennung: Es gilt $H \cdot \vec{w} = \vec{0}$ genau dann, wenn keine Fehler (oder ≥ 2 Fehler) aufgetreten sind.
- Fehlerkorrektur: Wenn $H \cdot \vec{w} \neq \vec{0}$, dann gibt $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ die Stelle an, an welcher der Fehler aufgetreten ist:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rightsquigarrow \text{1. Bit falsch } (\vec{v_1})$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rightsquigarrow \text{2. Bit falsch } (\vec{v_2})$$

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rightsquigarrow \text{3. Bit falsch } (\vec{v_3})$$

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rightsquigarrow \text{4. Bit falsch } (\vec{v_4})$$

3.3 Allgemeine Schranken für die Informationsrate

Modell eines binären, symmetrischen Kanals:

- Kanalalphabet $Q = \{0, 1\}$, Bitfolge wird übertragen
- Wahrscheinlichkeit, dass i-tes Bit fehlerhaft übertragen wird, ist gleich $p < \frac{1}{2}$, unabhängig davon, ob dieses Bit 0 oder 1 war.
- Die Ereignisse, dass erstes bzw. zweites, drittes ... Bit falsch übertragen werden, sind unabhängig.

Lemma: Die Wahrscheinlichkeit, dass bei der Übertragung eines Wortes der Länge n *genau* k Fehler auftreten, ist gleich

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Satz von Shannon: Gegeben ein binärer symmetrischer Kanal mit Fehlerwahrscheinlichkeit p und $\varepsilon > 0$:

- Zu jedem
$$R < 1 + p \cdot \log_2(p) + (1-p) \cdot \log_2(1-p)$$
gibt es einen Code C mit Informationsrate $\geq R$, so dass die Wahrscheinlichkeit einer falschen Decodierung (bei „nächster Nachbar"-Suche) höchstens ε ist.

- Zu jedem
$$R < 1 + p \cdot \log_2(p) + (1-p) \cdot \log_2(1-p)$$
gibt es eine Konstante $K_R > 0$, so dass jeder Code mit der Informationsrate $\geq R$ eine Wahrscheinlichkeit $\geq K_R$ für die falsche Decodierung eines Codeworts hat.

Definition: Die *Kapazität* eines binären symmetrischen Kanals mit Fehlerwahrscheinlichkeit p ist:

$$H(p) = 1 + p \cdot \log_2(p) + (1-p) \cdot \log_2(1-p)$$

Nachteile des Shannon-Satzes:

1. Der Satz ist nicht konstruktiv.
2. Wählt man ε klein, dann folgt daraus, dass n sehr groß ist.
3. „Nächster Nachbar"-Suche sehr komplex.

Alternative Ansatz: Wenn $|Q| = q$ und $\vec{r} \in Q^n$, dann gilt

$$B_k(\vec{v}) = \sum_{i=0}^{k} \binom{n}{i} (q-1)^i$$

- $i = 0, 1, \ldots k$: Abstand zu $\vec{v}$
- $\binom{n}{i}$: Stellen, an denen Unterschied zu $\vec{v}$ auftritt
- $(q-1)^i$: Möglichkeiten an diesen Stellen etwas anderes als in $\vec{v}$ zu schreiben

Erinnerung: Ein Code C ist genau dann k-fehlerkorrigierend, wenn $\forall c \neq c' \in C$ gilt

$$B_k(c) \cap B_k(c') = \emptyset$$

genau dann, wenn der Minimalsabstand $d(C) \geq 2k + 1$

Satz: Sei $C \subseteq Q^n$ ein Code mit $d(C) \geq 2k + 1$, dann gilt

$$|C| \cdot \sum_{i=1}^{k} \binom{n}{i} (q-1)^i \leq q^n$$

Beweisidee: Die Kugeln müssen disjunkt sein.

Definition: Ein Code $C \subseteq Q^n$ mit Minimalabstand $d(C) = 2k + 1$ ist *perfekt*, wenn

$$|C| \cdot \sum_{i=1}^{k} \binom{n}{i} (q-1)^i = q^n$$

Beispiel: Der Hamming-Code aus dem letzten Abschnitt ist perfekt. Angaben zu de Code:

- $q = |Q| = 2$
- $|C| = 2^4$
- $n = 7$
- $d(C) = 3$
- $k = 1$

Daraus folgt:

$$|C| \cdot \sum_{i=1}^{k} \binom{n}{i} (q-1)^i = 2^4 \cdot (1+8) = 2^7 = 2^n$$

Folgerung: Aus der Schranke

$$|C| \cdot \sum_{i=1}^{k} \binom{n}{i} (q-1)^i \leq q^n$$

kann man ableiten, dass ein binärer k-fehlerkorrigierender Code der Länge n muss $\approx k \cdot \log_2 n$ Redundanzbits haben.

Satz: Ist $s \leq n$ und g eine Zahl, die

$$g \cdot \sum_{i=0}^{s-1} \binom{n}{i} (q-1)^i \leq q^n$$

erfüllt, dann gibt es in Q^n einen Code c mit Minimalabstand s und $|C| = g$.

Beweisidee: Da $(g-1)$ Kugeln vom Radius $s-1$ Q^n noch nicht überdecken, folgt daraus, dass C erweitert werden kann.

3.4 Linear Codes

Endliche Körper:

- Für jede Primzahl p ist $\mathbb{Z}_p$ ein Körper

$$\mathbb{Z}_p = \{0, 1, \ldots p-1\}$$

- Für jede Primzahlpotenz $q = p^m$ gibt es einen Körper $\mathrm{GF}(q)$, der genau q Elemente hat. Die Körper haben die Charakteristik p, d.h.

$$\underbrace{1+1+\ldots+1}_{p} = 0$$

  $\mathrm{GF}(q)$ ist eine Erweiterung von $\mathbb{Z}_p$

Definition: Ein Code C heißt linear, wenn C ein Untervektorraum eines Hammingraumes $H(n,q)$ ist, wobei $H(n,q)$ die Menge der Wörter der Länge n über $\mathrm{GF}(q)$ ist, d.h.

$$H(n,q) \cong (\mathrm{GF}(q))^n$$

Beispiel: Der Hamming-Code aus 3.2 ist ein linearer Code in $H(7,2)$, denn er ist Bild einer linearen Abbildung

$$(\mathrm{GF}(2))^4 \to (\mathrm{GF}(2))^7$$

Die Dimenstion dieses Codes (Unterraumes) ist 4, wir sprechen von einem $(7,4)$-Code.

Beobachtung: Ein (n,k)-Code in $H(n,q)$ hat q^k Elemente (q^k Linearkombinationen der k Basisvektoren). Dieser Code hat die Informationsrate

$$\frac{1}{n} \cdot \log_q |C| = \frac{1}{n} \cdot \log_q \left(q^k\right) = \frac{k}{n}$$

Definition: Für ein $\vec{v} \in H(n,q)$ ist das *Gewicht* $w(\vec{v})$ die Anzahl der Stellen, an denen $\vec{v}$ ungleich 0 ist.

Definition: Das *Minimalgewicht* von C ist definiert als

$$w(C) = \min_{\vec{v} \neq \vec{0}} \{w(\vec{v}) \mid \vec{v} \in C\}$$

Beispiel: Für das Beispiel aus 3.2 ist $w(C) = 3$.

Satz: Für jeden linearen Code C gilt:

$$w(C) = d(C)$$

Das heißt: Das Minimalgewicht ist gleich dem Minimalabstand.

Beweis:

a) Zu zeigen: $w(C) \geq d(C)$:

$$w(C) = \underbrace{d(\vec{v}, \vec{0})}_{\text{für ein } \vec{v} \in C} \quad \underbrace{\geq d(C)}_{\text{denn } \vec{v}, \vec{0} \in C}$$

b) Zu zeigen: $w(C) \leq d(C)$:
$d(C)$ wird realisiert als $d(\vec{u}, \vec{v})$ für $\vec{u}, \vec{v} \in C$ mit $\vec{u} \neq \vec{v}$:

$$\vec{0} = \vec{v} - \vec{v} \quad \text{und} \quad \vec{u} - \vec{v} \in C \quad \text{(Unterraum)}$$

Dann folgt:

$$w(C) \leq w(\vec{u} - \vec{v}) = d(\vec{u} - \vec{v}, \vec{0}) = d(\vec{u}, \vec{v}) = d(C)$$

Definition: Generatormatrix von C:

$$\forall xin(\mathrm{GF}(q))^k \; G \cdot x \in C \quad \text{und} \quad G \text{ spannt } C \text{ auf}$$

Prüfmatrix/Checkmatix von C

$$\forall v \in C \; H \cdot v = (0) \quad (C = \mathrm{Ker}\ H)$$

Hmmpf: Jeder lineare Code C der Dimension k in $H(n,q)$ kann eindeutig (in Bezug auf den Code) durch eine Generatormatrix $G \in M(n \times k, \mathrm{GF}(q))$ dargestellt werden:

- Wähle Basis von C (als Spaltenvektoren) und stelle aus den k Basisvektoren eine Matrix auf.
- Damit beschreibt G eine Codierung

$$c : (\mathrm{GF}(q))^k \to (\mathrm{GF}(q))^n$$

- Eine Matrix $H \in M((n-k) \times n, \mathrm{GF}(q))$ wird Prüfmatrix (Checkmatrix) von C genannt, wenn C der Kern der von H beschriebenen Abbildung:

$$h : (\mathrm{GF}(q))^n \to (\mathrm{GT}(q))^{n-k}$$

Achtung: Nach der Dimensionsformel ist

$$\begin{aligned} n &= \dim(\mathrm{Ker}\ h) + \dim(\mathrm{Im}\ h) \\ &= \dim C + \mathrm{rg}\ H \\ &= k + \mathrm{rg}\ H \end{aligned}$$

$$\mathrm{rg}\ H = n - k$$

Das heißt: Die Zeilen von H sind linear unabhängig.

Hmmpf: Für $\vec{v} \in H(n,q)$ gilt:

$$\vec{v} \in C \quad \Leftrightarrow \quad H \cdot \vec{v} = \vec{0}$$

Satz: $G \in M(n \times k, \mathrm{GF}(q))$, $H \in M((n-k) \times n, \mathrm{GF}(q))$ mit $\mathrm{rg}\ G = k$ und $\mathrm{rg}\ H = n - k$ bilden genau dann ein Paar Generator/Checkmatrix für einen linearen Code C, wenn

$$H \cdot G = (0)$$

Anwendung: Eine Generatormatrix ist in Standardform, wenn sie die Gestalt

$$G = \begin{pmatrix} E_k \\ A \end{pmatrix}$$

hat. In diesem Fall ist die Matrix

$$H = (-A \quad E_{n-k})$$

eine passende Checkmatrix.

von Henning abscheiben!!!

Satz: C ein (n,k)-Code mit Prüfmatrix H, dann gilt:

$$d(C) \geq d \quad \Leftrightarrow \quad \text{je zwei } d-1 \text{ Spalten von } H \text{ sind linear unabhängig}$$

Beweis ($\Rightarrow$): Angenommen H enthält $d-1$ linear abhängige Spalten (wir nehmen an, die ersten $d-1$), genau dann wenn

$$\exists \alpha_1, \alpha_2, \cdots \alpha_{d-1} (\text{nicht alle } 0)$$

so dass

$$\alpha_1 \cdot H_1 + \alpha_2 \cdot H_2 + \ldots + \alpha_{d-1} \cdot H_{d-1} = 0$$

genau dann, wenn

$$\vec{\alpha} = (\alpha_1, \ldots \alpha_{d-1}, 0, \ldots 0) \neq 0$$

Dann:

$$H \cdot \vec{a} = \sum_{i=1}^{n} \alpha_i \cdot H_i = \sum_{i=1}^{d-1} \alpha_i \cdot H_i + \sum_{i=d}^{n} \alpha_i \cdot H_i = 0$$

d.h. $\vec{a} \in \text{Ker } H = C$

$$d(C) = w(C) \leq w(\vec{\alpha}) \leq d-1$$

Kommentar: $\vec{\alpha} = \bar{a} \leq d-1$

Beweis ($\Leftarrow$): $\vec{v} \in C$, $\vec{v} \neq \vec{0}$, dann $H \cdot \vec{v} = \vec{0}$, wenn $w(\vec{v}) \leq d-1$. Das heißt, wir finden $\leq d-1$ Spalten von H die linear abhängig sind. $\square$

Folgerung: (Fall $d = 3$ des Satzes)
C ist (n,k)-Code mit Prüfmatrix H.

$$d(C) \geq 3 \quad \Leftrightarrow \quad \exists 2 \text{ Spalten von } H \text{ sind linear unabhängig}$$

Wichtig: Sind keine Spalten von H vielfache von einander, dann ist C 1-fehlerkorrigierend.

Beispiel: Code aus 3.2
Prüfmatrix:

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Aus der Prüfmatrix folgt, dass der Hamming-Code ist 1-fehlerkorrigierend.

Weiteres Beispiel: Trippel-Check-Code:
Prüfmatrix:

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

Generatormatrix:

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Das heißt:

$$(a, b, c) \mapsto (a, b, c, a + b, a + c, b + c)$$

Typischerweise wird hier der Körper GF(3) benutzt.

Allgemeine binäre Hamming-Codes: Ziel: 1-fehlerkorrigierend, hohe Informationsrate
n Länge des Codes, H Prüfmatrix

$$n = \underbrace{\dim(\text{Ker } H)}_{\dim C} + \dim(\text{Im } H)$$

Informationsrate:

$$\frac{\dim(\text{Ker } H)}{n} = \frac{n-k}{n}$$

Ist $\dim(\text{Im } H) = k$ fest, dann wollen wir n groß, damit die Informationsrate groß ist.

- Möglichst viele Spalten ($\Rightarrow$ möglichst große Informationsrate)
- Alle Spalten verschieden (und $\neq \vec{0}$)
  damit der Code 1-fehlerkorrigierend ist

Definition: Der binäre Hamming-Code Ham_2k hat als Prüfmatrix die Matrix H_k der n Spalten alle verschiedenen binären Vektoren der Länge k (ohne $\vec{0}$)

Beispiel: Erstes Beispiel:

$$\text{Ham}_2 2 \quad H_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad (a) \mapsto (a, a, a) \quad textdim = 1 \quad \text{Länge} = 3$$

Zweites Beispiel:

$$\mathrm{Ham}_2 3 \quad H_{\S} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Das ist (bis auf Permutation der Spalten) der Hamming-Code von früher

$$\dim = 4 \quad \text{Länge} = 7 \quad d = 3$$

Drittes Beispiel:

$$\mathrm{Ham}_2 4 \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Satz: Ham$k$ ist ein linearer Code mit den Parametern $d(\mathrm{Ham}k) = 3$, Länge$(\mathrm{Ham}k) = 2^k - 1$, $\dim(\mathrm{Ham}(k)) = 2^k - k -^1$
Informationsrate für große Werte:

$$\lim \left(\frac{2^k - k - 1}{2^k - 1} \right) = 1$$

Proposition: Der Hamming-Code Ham$k$ ist 1-perfekt.

Erinnerung: C Code mit $C \subseteq Q^n$ miz $d(C) = 2t - 1$ ist perfekt, wenn

$$|C| \cdot \sum_{i=0}^{t} \binom{n}{k} (q-1)^t = q^n$$

(die t-Kugeln um die Codewörter füllen den Raum Q^n perfekt aus)

Beweis: $t = 1$, $q = 2$, $n = 2^k - 1$

$$|C = d^{\dim C}| = 2^{2^k - k - 1}$$

Größe der Kugeln:

$$\sum_{i=0}^{t} \binom{2^k - 1}{k} 1 = 1 + (2^k - 1) = 2^k$$

$$|C| Erdk = 2^{2^k - k - 1} \cdot 2^k = 2^{2^k - 1} = 2^n$$

- Ham3 kann als Ausgangspunkt für die Konstruktor von 3-perfekten Codes genommen werden $d() = 7$

  Galay Codes

  G_{23} ist binärer $(23, 12)$-Code

- Codes über adenren Körpern inbesondere GF(2^k)

  RCH Codes, RS-Codes