Vorlesungsmitschrift zu Mathematik für Informatiker 3

Naja v. Schmude
www.najas-corner.de/wordpress

16. Februar 2008

Inhaltsverzeichnis

1. Vorlesung, 15.10.2007

1.1. Organisatorisches

Anmeldung zu den Tutorien wird am Mittwoch um ca. 16 Uhr freigeschaltet.
Die Termine - die Tutorien fangen erst nächste Woche an:

Mi	14-16	Alex
Mi	16-18	Alex
Do	10-12	Hanne
Fr	8-10	Leif
Fr	10-12	Leif
Fr	12-14	Hanne

Die Übungszettel sollen immer montags vor der Vorlesung bis 12.15 abgegeben werden.

1.1.1. Aufbau der Vorlesung

Lineare Algebra und Stochastik ist das Thema.

1. Lineare Algebra
2. Codierungstheorie
3. Stochastik

1.2. Grundlegende algebraische Strukturen

Diese kennen wir ja schon aus Mafi 2.

- Halbgruppen, Monoide
- Grupen
- Körper
- Vektorräume

H ist eine Menge von Elementen und $*$ sei eine Operation $H \times H \to H$ mit $(h_1, h_2) \to h_1 * h_2$.

Definition 1 (Halbgruppe). $(H, *)$ *heißt Halbgruppe, falls die Assoziativität gilt:*

$$\forall h_1, h_2, h_3 \in H : (h_1 * h_2) * h_3 = h_1 * (h_2 * h_3)$$

Definition 2 (Monoid). *Eine Halbgruppe* $(H, *)$ *heißt Monoid, falls es zusätzlich ein neutrales Element gibt:*

$$\exists e \in H \forall h \in H : h * e = e * h = h$$

Definition 3 (Gruppe). *Ein Monoid* $(H, *, e)$ *heißt Gruppe, falls es zusätzlich ein inverses Element gibt:*

$$\forall h \in H \exists \overline{h} \in H : h * \overline{h} = \overline{h} * h = e$$

Die Standardnotation für eine Gruppe lautete $(G, *, e)$.

Definition 4 (kommutative Gruppe). $(G, *, e)$ *Gruppe ist kommutativ (abelsch) falls zusätzlich gilt:*

$$\forall g_1, g_2 \in G : g_1 * g_2 = g_2 * g_1$$

Beispiele

1. $(\mathbb{N}, +, 0)$ ist ein Monoid, da wir zur 1 kein inverses Element finden.
   $(\mathbb{N}, \cdot, 1)$ ist auch ein Monoid, weil es keine Brüche gibt, gibt es auch keine inversen Elemente (bis auf die 1).

2. $(\mathbb{Z}, +, 0)$ ist eine kommutative Gruppe.
   $(\mathbb{Z}, \cdot, 1)$ ist immer noch ein Monoid.

3. $(\mathbb{Q}, +, 0)$, $(\mathbb{Q}\backslash\{0\}, \cdot, 1)$ sind Gruppen.

4. Grundmenge ist $\mathbb{B}$, die Wahrheitswerte. Dann ist $(\mathbb{B}, \wedge, 1)$ ein Monoid, aber keien Gruppe, weil es kein inverses Element zur Gruppe gibt.
   $(\mathbb{B}, \vee, 0)$ ist Monoid
   $(\mathbb{B}, \oplus, 0)$ ist eine Gruppe.
   $(\mathbb{B}\backslash 0, \wedge, 1)$ ist Gruppe

5. $f : \mathbb{B}^n \to \mathbb{B}$ ist eine n-stellige Boolesche Funktion mit $(b_1, b_2, \ldots, b_n) \to b$. $\mathbb{B}_n$ ist die Menge aller n-stelliger Boolescher Funktionen.
   Es sei $(\mathbb{B}_n, \vee, \underline{0})$ ein Monoid. Wie definiert man nun $f \vee g$?

   $$(f \vee g)(b_1, \ldots, b_n) = f(b_1, \ldots, b_n) \vee g(b_1, \ldots, b_n)$$

   $\underline{0}$ ist dabei $\underline{0}(b_1, \ldots, b_n) = 0 \in \mathbb{B}$

Definition 5 (Körper). $(K, +, \cdot, 0, 1)$ *heißt Körper, wenn* $(K, +, 0)$ *eine kommutative Gruppe bildet und* $(K\backslash\{0\}, \cdot, 1)$ *auch eine kommutative Gruppe bildet. Zusätzlich fordert man die Distributivität:*

$$\forall a, b, c \in K : a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

Diese drei Forderungen sind die "Körperaxiome"

Beispiel $(\mathbb{Q}, +, \cdot, 0, 1)$, $(\mathbb{R}, +, \cdot, 0, 1$ und $(\mathbb{C}, +, \cdot, 0, 1)$ sind Körper. Auch $(\mathbb{B}, \oplus, \wedge, 0, 1)$ ist ein zweielementiger Körper.

Sei p eine Primzahl. Die Grundmenge $= \{0, 1, 2, \ldots, p-1\}$ sind die Reste der Division durch p. Man definiert eine Addition $+$ als $a+b = (a+b) \mod p$ und die Multiplikation als $a \cdot b = (a \cdot b) \mod p$. 0 ist das neutrale Element der Addition und 1 das der Multiplikation. Insgesamt handelt es sich also um einen Körper.

2. Vorlesung, 17.10.2007

2.0.1. Eigenschaft von Gruppen

Sei $(G, \cdot, 1)$ eine Gruppe.
Für jedes $g \in G$ ist das inverse Element g^{-1} eindeutig.

Beweis 1. *Angenommen $g \in G$ mit inversem Elementen $\overline{g}$ und $\tilde{g}$ und $\overline{g} \neq \tilde{g}$.*

$$\overline{g} = \overline{g} \cdot \cdot 1 = (\overline{g} \cdot g) \cdot \tilde{g} = 1 \cdot \tilde{g} = \tilde{g} \quad \square$$

2.0.2. Rechnen in Körpern

$(K, +, \cdot, 0, 1)$ ist ein Körper mit $(K, +, 0)$ und $(K, \cdot, 1)$ sind abelsche Gruppen. Zudem gilt $\forall a, b, c \in K : a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.
Die Gleichungen der Form $a \cdot x + b = c$ mit $a \neq 0$ hat eine eindeutige Lösung in K.

$$\begin{aligned} a \cdot x + b + (-b) &= c + (-b) \\ a \cdot x &= c + (-b) \\ a \cdot x \cdot a^{-1} &= (c + (-b)) \cdot a^{-1} \\ x &= (c + (-b)) \cdot a^{-1} \end{aligned}$$

Beispiel p sei Primzahl und $\mathbb{Z}_p$ der Körper über $\{0, 1, \ldots, p-1\}$ mit Addition und Multiplikation $\mod p$.
In $\mathbb{Z}_5$ kann z.B. die Formel $2 \cdot x + 1 = 4$ gelöst werden:

$$\begin{aligned} 2x + 1 &= 4 \\ 2x + 1 + 4 &= 3 \\ 2x &= 3 \\ 2 \cdot 3 \cdot x &= 3 \cdot 3 \\ x &= 4 \end{aligned}$$

2.1. Verktorräume

Definition 6 (Vektorräume). *$(K, +, \cdot, 1, 2)$ ist ein Körper und V eine Menge. Wir haben die Operationen $+ : V \times V \to V$ und $\cdot : K \times V \to V$.*

1. $(V, +, \vec{0})$ *ist abelsche Gruppe*
2. $\forall \lambda, \mu \in K \forall v \in V : \lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$
3. $\forall v \in V : 1 \cdot v = v$
4. $\forall \lambda \in K, \forall v, w \in V : \lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w)$
5. $\forall \lambda, \mu \in K \forall v \in V : (\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v)$

V heißt Vektorraum über K, auch K-Vektorraum. Die Elemente in V nennt man Vektoren, die in K heißen Skalare und mit $\vec{0}$ wird der Nullvektor bezeichnet.
Besonders interessant: $K = \mathbb{R}$, der reelle Vektorraum und auch $K = \mathbb{C}$, der komplexe Vektorraum.

Beispiele

1. Für jedes $n \geq 1$: $\mathbb{R}^n$ ist $\mathbb{R}$-Vektorraum und für $\mathbb{C}^n$ ist $\mathbb{C}$-Vektorraum. Dabei ist $\mathbb{R}^n = \{(r_1, \cdot, r_n) | \forall r \in R\}$.

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$
$$(r_1, \ldots r_n) + (r'_1, \ldots, r'_n) = (r_1 + r'_1, r_2 + r'_2, \ldots, r_n + r'_n)$$

2. $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum
   $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ Addition von reelle Zahlen
   $* : \mathbb{Q} \times \mathbb{R} \to \mathbb{R}$ normale Multiplikation.
   Aber $\mathbb{Q}$ ist kein $\mathbb{R}$-Vektorraum bezüglich Multiplikation vom Typ $\mathbb{R} \times \mathbb{Q} \to \mathbb{Q}$. Das kann nämlich aus $\mathbb{Q}$ führen.

3. K^n ist ein K-Vektorraum, wobei $K^n = \{(k_1, \ldots, k_n) | \forall k \in K\}$.
   $+ : K^n \times K^n \to K^n$ als $(k_1, \ldots, k_n) + (k'_1, \ldots, k'_n) = (k_1 + k'_1, \ldots, k_n + k'_n)$ ist die Körperaddition.

4. $\mathbb{Z}_2^n$ ist $\mathbb{Z}_2$-Vektorraum. Dieser Vektorraum ist wichtig für die Kodierungstheorie.

5. Funktionenräume
   $\mathcal{F} : \{f : \mathbb{R} \to \mathbb{R}\}$ ist $\mathbb{R}$-Vektorraum.
   $+ : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$
   $\forall x \in \mathbb{R} : (f + g)(x) = f(x) + g(x)$
   $\cdot : (\lambda) f(x) = \lambda \cdot f(x)$
   $\vec{0} = h$ mit $\forall x \in \mathbb{R} : h(x) = 0$

6. Polynome $K[x]$ über K bilden einen K-Vektorraum. Ein Polynom ist eine Funktion der Form

$$\sum_{i=0}^{n} a_i x^i$$

2.1.1. Unterstruktur

Definition 7 (Unterraum). *V sei ein K-Vektorraum und sei $\emptyset \neq U \subseteq V$. Ist U mit den von V induzierten Operationen ein K-Vektorraum, so heißt U ein Unterraum (Teilraum) von V. äquivalent: $\forall u, v \in U : u + v \in U$ und $\forall u \in U \forall \lambda \in K : \lambda \cdot u \in U$.*

Beispiel

1. V ist ein K-Vektorraum. V und $\{\vec{0}\}$ sind Unterräume.
2. Die Unterräume von $\mathbb{Z}_2^n$ heißen lineare Codes.
3. $\mathbb{R}^2$. Dann beschreib $\{\lambda \cdot v | \lambda \in \mathbb{R}\}$ eine Gerade.
   Allgemein für jeden K-Vektorraum V gilt: $v \in V\{\lambda \cdot v | \lambda \in K\}$.
   Noch allgemeiner: K-Vektorraum über V. $v_1, v_2, \ldots, v_n \in V$, $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. $\sum_{i=1}^n \lambda_i \cdot v_i$ heißt Linearkombination von $v_1, v_2 \ldots v_n$.

Definition 8 (Lineare Hülle). *Die Menge aller Linearkombinationen der Vektoren $v_1, \ldots, v_n$ heißt Lineare Hülle (span) der Vektoren. Also $M = \{v_1, v_2, \ldots v_n\}$ dann wir die Lineare Hülle mit $Lin(M)$ bezeichnet.*

Satz 1. *V ist ein K-Vektorraum. $M = \{v_1, v_2, \ldots, v_n\}$ die Menge von Vektoren. $Lin(M)$ ist Unterraum von V.*

Beweis 2. *$u, w \in Lin(M)$. Dann lässt sich u schreiben als $u = \sum_{i=1}^n \lambda_i v_i$. Genauso kann auch w geschrieben werden: $w = \sum_{i=1}^n \mu_i v_i$*

$$u + w = \sum_{i=1}^{n} (\lambda_i + \mu_i) v_i \in Lin(M)$$

$$\lambda u = \sum_{i=1}^{n} \lambda \lambda_i v_i \in Lin(M)$$

□

3. Vorlesung, 22.10.2007

3.0.2. Lineare (Un)Abhängigkeit

Definition 9 (linear abhängig). $M = \{v_1, \ldots, v_n\} \subseteq V$. $v \in V$ *heißt linear abhängig von* M, *falls*

$$\exists \lambda_i : \sum_{i=1}^{n} \lambda_i * v_i = v$$

Beobachtung Falls v linear abhängig ist von der Menge $M = \{v_1, \ldots, v_n\}$, so ist $Lin(M) = Lin(M \cup \{v\})$

Beweis 3. *Sei* $w \in Lin(M \cup \{v\})$.

$$w = \sum_{i=1}^{n} \lambda_i' v_i + \lambda_v v$$

$$v = \sum_{i=1}^{n} \lambda_i' v_i$$

also

$$w = \sum_{i=1}^{n} \lambda_i' v_i + \sum_{i=1}^{n} \lambda_v \lambda_i v_i = \sum_{i=1}^{n} (\lambda_i' + \lambda_v \lambda_i) v_i$$

also $w \in Lin(M)$. □

Definition 10 (linear unabhängig). $\{v_1, \ldots, v_n\}$ *heißt linear unabhängig, wenn sich kein* v_i *als Linearkombination der anderen darstellen lässt.*

Satz 2 (Kriterium für lineare Unabhängigkeit). V *ist ein* K*-Vektorraum.* $M = \{v_1, \ldots, v_n\}$ *ist linear unabhängig, genau dann wenn für jede Linearkombination mit*

$$\sum_{i=1}^{n} \lambda_i * v_i = \vec{0}$$

gilt: $\forall i : \lambda_i = 0$. *(Jeder Vektor* $v \in Lin(M)$ *lässt sich eindeutig als Linearkombination von* $v_1, \ldots, v_n$ *darstellen.)*

Beweis 4.

=> *Angenommen* $\exists \lambda_1, \ldots, \lambda_n : \sum_{i=1}^{n} \lambda_i v_i = \vec{0}$ *und ein* $\lambda_k \neq 0$

$$\sum_{i=1, i\neq k}^{n} \lambda_i v_i = -\lambda_k v_k \quad | * (-\lambda_k)^{-1}$$

$$\sum_{i=1, i\neq k}^{n} \lambda_i (-\lambda_k)^{-1} v_i = v_k$$

Also ist v_k *linear abhängig von* $\{v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\}$

<= *Sei* $v_1, \ldots, v_n$ *linear abhängig.*

$$\exists k : v_k = \sum_{i=1, i\neq k}^{n} \lambda_i v_i$$

Setze $\lambda_k = -1$.

$$\vec{0} = \sum_{i=1}^{n} \lambda_i v_i \wedge \lambda_k = 0$$

□

Definition 11. *Eine unendliche Menge von Vektoren heißt linear unabhängig, falls jede endliche Teilmenge linear unabhängig ist.*

Beispiel Wir nehmen uns reelle Polynome $\mathbb{R}[x] = \{\sum_{i=1}^{n} r_i x^i | \, r_i \in \mathbb{R}\}$. Dann ist $\{1, x, x^2, x^3, x^4 \ldots\}$ linear unabhängig.

$$\sum_{i=1}^{n} r_i x^i = \vec{0}$$

$\Rightarrow$ alle r_i müssen 0 sein, weil sonst auf linker Seite ein Polynom stehen würde.

3.0.3. Basis eines Vektorraums

Definition 12 (Basis). V *sei ein* K*-Vektorraum. Eine Menge* $B \subset V$ *heißt Basis von* V*, falls*

1. $Lin(B) = V$
2. B *ist linear unabhängig*

Beispiel $\{(2,3),(3,4)\}$ ist eine Basis von $\mathbb{R}^2$. Für jedes $(x,y) \in \mathbb{R}^2$ existiert $\lambda, \mu \in \mathbb{R}$

$$\lambda * (2,3) + \mu * (3,4) = (x,y)$$

also

$$\begin{aligned}
\lambda * 2 + \mu * 3 &= x \\
\lambda * 4 + \mu * 4 &= y \\
9\mu - 8\mu &= 3x - 2y \quad \text{nach Lösen des Lineares Gleichungssystem} \\
\Rightarrow \mu &= 3x - 2y \\
\lambda &= -4x + 3y
\end{aligned}$$

Die Standardbasis von $\mathbb{R}^2$ ist $\{(1,0),(0,1)\}$.

$$(x,y) = x(1,0) + y(0,1)$$

$\mathbb{R}^2$ hat unendlich viele endliche Basen.

Satz 3. *Jeder Vektorraum ($\neq \{\vec{0}\}$) hat eine Basis!*

- Diese Basist ist im Allgemeinen nicht eindeutig.
- die Standardbasis von $\mathbb{R}^n$ lautet

$$\{(1,0,0\ldots,0),(0,1,\ldots,0),\ldots,(0,0,\ldots,0,1)\}$$

- Basen können unendlich groß sein.

Satz 4 (Austauschsatz von Steinitz). *V sein ein K-Vektorraum. Sei $B = \{b_1,\ldots,b_n\}$ eine endliche Basis von V.*
Sei $v \in V$ ein beliebiger Vektor $\neq \vec{0}$.

$$\exists b_k \in B : \{b_1,\ldots,b_{k-1},v,b_{k+1},\ldots,b_n\} \text{ ist Basis von } V$$

Beweis 5. *B ist Basis. Also*

$$v = \sum_{i=1}^{n} \lambda_i b_i$$

Da $v \neq \vec{0} : \exists i : \lambda_i \neq 0$. Ohne Beschränkung der Allgemeinheit können wir sagen $i = 1$.
Zeige $\{v, b_2, \ldots, b_n\}$ ist Basis. D.h.

1. $Lin(\{v,b_2,\ldots,b_n\}) = V$
   B ist Basis: $\forall u \in V : u = \sum_{i=1}^{n} \mu_i b_i$ *(*)*
   insbesondere

$$\begin{aligned}
v &= \sum_{i=1}^{n} \lambda_i b_i \\
&= \lambda_1 b_1 + \lambda_2 b_2 + \ldots + \lambda_n b_n \quad | * (\lambda_1)^{-1} \\
b_1 &= (\lambda_1)^{-1} v - \sum_{i=2}^{n} \lambda_i (\lambda_1)^{-1} b_i
\end{aligned}$$

Einsetzen in (*)

$$u = \mu_1(\lambda_1^{-1}v - \sum_{i=2}^{n} \lambda_i\lambda_1^{-1}b_i) + \mu_2 b_2 + \ldots + \mu_n b_n$$
$$u = \mu_1\lambda_1^{-1}v + (\mu_2\lambda_2\lambda_1^{-1})b_2 + \mu_2 b_2$$
$$=?? + (\mu_2 + (???\lambda_1^{-1}))b_2$$

2. $\{v, b_2, \ldots, b_n\}$ *ist linear unabhängig. Falls* $\beta_1 v + \beta_2 b_2 + \ldots + \beta_n b_n = \vec{0}$*, so ist also* $\beta_1 = \beta_2 = \ldots = \beta_n = 0$*.*

$$\beta_1(\sum_{i=1}^{n} \lambda_i b_i) + \beta_2 b_2 + \ldots + \beta_n b_n = \vec{0}$$
$$\beta_1\lambda_1 b_1 + (\beta_1\lambda_2 + \beta_2)b_2 + \ldots + (\beta_1\lambda_n + \beta_n)b_n = \vec{0}$$

$B = \{b_1, \ldots, b_n\}$ *ist Basis, also ist insbesondere* $\beta_1\lambda_1 = 0$*. Daraus folgt, dass* $\beta_1 = 0$*, da wir gerade* $\lambda_1 \neq 0$ *annahmen.*

4. Vorlesung, 29.10.2007

Korrolar Aus dem Austauschsatz kann man schlussfolgern, dass verschiedene Basen eines Vektorraums gleichmächtig sind.

4.0.4. Dimension eines Vektorraums

Definition 13 (Dimension). *Die Dimension von V ist gleich der Anzahl der Vektoren in einer Basis. Man schreibt $dim_K V$*

Hinweis $dim(\{\vec{0}\}) = 0$

1. $U \subseteq V$ mit U ist Teilraum von V. Dann folgt, dass $dim\ U \leq dim\ V$ (für endliche Vektorräume). Insbesondere gilt für $U \subset V$, dass $dim\ U < dim\ V$.

2. $U, V \subseteq W$ seien Teilräume. Dann sind auch $U \cap V$ und $U + V := \{u+v | u \in U, v \in V\}$ wieder Vektorräume.
   Beispiel $\mathbb{R}^3$. Man nehme sich eine Ebene U und eine Gerade V, die in der Ebene liegt. Dann gilt $dim\ V = 2$,$dim\ U = 1$ und $dim\ U \cap V = 1$ und $dim\ U + V = 2$. Wenn die Gerade aber die Ebene echt schneidet, dann ergibt sich $dim\ U \cup V = 0$ und $dim\ U + V = 3$.

Satz 5 (Basisergänzungssatz). *Eine Menge von linear unabhängigen Vektoren lässt sich zu einer neuen Basis ergänzen.*

Satz 6. $dim(U \cap V) + dim(U + V) = dim\ U + dim\ V$

Beweis 6. *Man benutzt den Basisergänzungssatz. Es gilt $U \cap V \subset U$ und $U \cap V \subset V$. Gleichzeitig gilt $U \subseteq U + V$ und $V \subseteq U + V$.*
Sei $\{v_1, \ldots, v_r\}$ Basis von $U \cap V$. Dann ist nach der Basisergängzung

$$\{v_1, \ldots, v_r\} \subset \{v_1, \ldots, v_r, w_1, w_2, \ldots, w_s\} \text{ Basis von } U$$

$$\{v_1, \ldots, v_r\} \subset \{v_1, \ldots, v_r, z_1, w_2, \ldots, z_t\} \text{ Basis von } V$$

Zeige $\{v_1, \ldots v_r, w_1, \ldots, w_s, z_1, \ldots, z_t\}$ ist Basis von $U + V$.
Dafür müssen folgende Punkte gezeigt werden:

1. $Lin(\{v_1, \ldots v_r, w_1, \ldots, w_s, z_1, \ldots, z_t\}) = U + V$. *Das ist trivial, da hier schon alles drin ist, um Vektoren aus U und V zu bilden.*

2. Die Vektoren $\{v_1, \ldots v_r, w_1, \ldots, w_s, z_1, \ldots, z_t\}$ sind linear unabhängig.

zu zeigen:

$$(*)\lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 w_1 + \ldots + \mu_s w_s + \beta_1 z_1 + \ldots + \beta_t z_t = \vec{0}$$

Daraus folgt, alle Skalare sind 0. Man schreibt die Gleichung um:

$$\beta_1 z_1 + \ldots + \beta_t z_t = -(\lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 w_1 + \ldots + \mu_s w_s) \in U \cap V$$

$\beta_1 z_1 + \ldots + \beta_t z_t = \alpha_1 v_1 + \ldots \alpha_r v_r$ *Basis von $U \cap V$. Also*

$$\beta_1 z_1 + \ldots + \beta_t z_t - (\alpha_1 v_1 + \ldots + \alpha_r v_r) = \vec{0}$$

in (*) einsetzen:

$$\lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 w_1 + \ldots + \mu_s w_s = \vec{0}$$

Dies ist eine Basis und alle Vektoren sind linear unabhängig, also gilt $\lambda_1 = \ldots = \lambda_r = \mu_1 = \ldots = \mu_s = 0$

4.0.5. Lineare Abbildungen (Vektorraumhomomorphismen)

$f : U \to V$ mit U, V sind K-Vektorräume. Diese heißt lineare Abbildung, falls

$$\forall u_1, u_2 \in U \forall \lambda, \mu \in K : f(\lambda u_1 + \mu u_2) = \lambda f(u_1) + \mu f(u_2)$$

Die Lineare Abbildung überführt Linearkombination in Linearkombination.

Beispiel

1. $f : \mathbb{R}^3 \to \mathbb{R}^3$.

   $$f(x_1, x_2, x_3) = (2x_1 + x_3, -x_2)$$

   ist lineare Abbildung.

   $$\begin{aligned} &f(\alpha(x_1, x_2, x_3) + \beta(y_1, y_2, y_3)) = f(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \alpha x_3 + \beta y_3) \\ &= (2\alpha x_1 * 4\beta y_1 + \alpha x_3 + \beta y_3, -\alpha x_2 - \beta y_2) \\ &= (\alpha(2x_1 + x_3) + \beta(2y_1 + y_3), -\alpha x_2 - \beta y_2) \\ &= \alpha * (2x_1 + x_3, -x_2) + \beta * (2y_1 - y_3, -y_2) \\ &= \alpha * (f(x_1, x_2, x_3) + \beta * f(y_1, y_2, y_3) \end{aligned}$$

2. $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x + 1$ ist keine lineare Abbildung, weil $f(\vec{0}) \neq \vec{0}$.
3. $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $(x_1, x_2) \to (x_1, x_1 x_2)$ ist auch keine lineare Abbildung.

Definition 14. *Sei $f : U \to V$ eine lineare Abbildung.*
Den Kern von f ist definiert als

$$Ker(f) = \{u \in U | f(u) = \vec{0}\}$$

und das Bild von f als

$$Im(f) = \{v \in V | \exists u \in U : f(u) = v\}$$

$Ker(f)$ ist Unterraum von U. $Im(f)$ ist Unterraum von V. $Hom(U,V) = \{f : U \to V | f\ lineare\ Abbildung\}$ ist Vektorraum.

Weitere Beispiele

1. $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $(x, y) \to (x, -y)$. Kann man auch als Spiegelung an der x-Achse sehen. Dann ist $Ker(f) = \{\vec{0}\}$ und $Im(f) = \mathbb{R}^2$.
2. $f : \mathbb{R}^3 \to \mathbb{R}^3$ mit $(x, y, z) \to (x, 0, 0)$. Dies ist die Projektion auf die x-Achse. Dann ist $Ker(f) = Lin(\{(0, 1, 0), (0, 0, 1)\})$. $Im(f) = Lin(\{(1, 0, 0)\})$
3. $f : \mathbb{R}^2 \to \mathbb{R}^2$. $(x, y) \to (\frac{\sqrt{2}}{x} - \frac{\sqrt{2}}{y}, \frac{\sqrt{2}}{x} + \frac{\sqrt{2}}{y})$. Dies ist die Drehung von $+45°$ (gegen den Uhrzeigersinn). $Ker(f) = \{\vec{0}\}$ und $Im(f) = \mathbb{R}^2$.

5. Vorlesung, 31.10.2007

Ein Homomorphismus $Hom_K(V, W)$ ist die Menge aller lineare Abbildungen von V nach W. Dies ist wieder ein K-Vektorraum.

Sei $f, g \in Hom_K(V, W)$ was ist dann mit $f + g$?
$f + g : V \to W$ ist definiert durch $(f + g)(v) = f(v) + g(v)$ und $(\lambda f)(v) = \lambda * f(v)$
Gleichzeitig müssen allerdings auch noch die Axiome überprüft werden.

Spezielle Eigenschaften von linearen Abbildungen

$f : V \to W \in Hom_K(V, W)$ heißt Monomorphismus, falls f injektiv ist. Man spricht von einen Epimorphismus, falls f surjektiv und vom Isomorphismus, falls f bijektiv ist. Dann gibt es noch den Endomorphismus, falls $V = W$ und den Automorphismus, falls $V = W$ und dabei das Teil noch isomorph ist.

Folgende einfache Sachverhalte gelten:

1. f ist Monomorphismus, genau dann wenn der Kern $Ker(f) = \{\vec{0}\}$ ist.
2. f ist ein Isomorphismus, daraus folgt f^{-1} ist auch isomorph.
3. Seien $f : V \to W$ und $g : W \to Z$ Isomorphismen. Dann ist $gf : V \to Z$ auch ein Isomorphismus.

Erinnerung Sei eine Basis von V gegeben. Für jedes $v \in V$ ist die Darstellung als Linearkombination der Basisvektoren eindeutig.

5.0.6. Umrechnen der Darstellung eines Vektors bzgl. einer Basis in eine Darstellung bzgl. einer anderen Basis.

Wir haben z.B. $\mathbb{R}^2$ gegeben und die Basen $B = \{b_1, b_2\}$ und $C = \{c_1, c_2\}\cdot$ Zudem ist bekannt, dass $c_1 = 2b_1 + 2b_2 = \binom{2}{2}_B$ und $c_2 = 1b_1 + 2b_2 = \binom{1}{2}_B$.

Nun hat man $v = \binom{\alpha}{\beta}_B$ und es ist nun die Darstellung von v in Basis C gesucht.

$$\begin{aligned}\binom{\alpha}{\beta}_B &= v = \binom{\lambda}{\mu}_C \\ &= \lambda c_1 + \mu c_2 \\ &= \lambda \binom{2}{2}_B + \mu \binom{1}{2}_B \\ &= \binom{2\lambda - \mu}{2\lambda - 2\mu}_B\end{aligned}$$

Darasu ergibt sich ein Gleichungssystem:

$$\begin{aligned}2\lambda - \mu &= \alpha \\ 2\lambda - 2\mu &= \beta \\ \Rightarrow \mu &= -\alpha + \beta \\ \lambda &= \alpha - \frac{1}{2}\beta\end{aligned}$$

Wenn man jetzt das mit konkreten Werten macht, ergibt sich

$$\binom{5}{2}_B = \binom{5 - \frac{1}{2} * 2}{-5 + 2}_B = \binom{4}{-3}_C$$

Satz 7 (Charakterisierung einer linearen Abbildung durch die Wirkung auf die Basis)**.** *Sei $f : U \to V$ und sei $\{b_1, \ldots, b_n\}$ Basis von U. Zudem seinen Vektoren $v_1, \ldots, v_n \in V$ gegeben.*
Dann existiert genau eine lineare Abbildung $f : U \to V$ mit $f(b_i) = v_i \forall i$

Beweis 7. *Für $u \in U$, was ergibt sich dann für $f(u)$? Man schreibe erstmal $u = \sum_{i=1}^n \lambda_i b_i$ als Linearkombination. Also ist $f(u) = \sum_{i=1}^n \lambda_i v_i$. Dies ist jetzt schon unsere lineare Abbildung. Jetzt bleibt noch zu zeigen, dass die Abbildung f eindeutig ist. Angenommen $g : U \to V$ mit $\forall i g(u_i) = v_i$.*

$$\forall u \in U : g(u) = g(\sum_{i=1}^n \lambda_i b_i) = \sum_{i=1}^n \lambda_i g(b_i) = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \lambda_i f(b_i) = f(\sum_{i=1}^n \lambda_i b_i) = f(u)$$

□

Merke Die gesamte Information über eine lineare Abbildung ist in den Bildern der Basisvektoren kodiert.

Satz 8. *$f : V \to W$ ist Isomorphismus, und $\{v_1 \ldots, v_n\}$ Basis von V, genau dann wenn $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ Basis von W ist.*

Beweis 8.

=> *Zeige, das* $\{f(v_1), \ldots, f(v_n)\}$ *linear Unabhängig sind.*
Sei $\lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = \vec{0}$. *Daraus folgt wegen der Linearität*

$$f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \vec{0}$$

Aus dem Monomorphismus von f *folgt widerum*

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = \vec{0}$$

Und daraus ergibt sich, dass alle Skalare

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

sein müssen.
Zeige das $Lin(\{f(v_1), \ldots, f(v_n)\}) = W$. *Sei* $w \in W$. *Dann* $\exists v : f(v) = w$. v *lässt sich mit den Basisvektoren schreiben als* $v = \sum_{i=1}^{n} \lambda_i v_i$. *Dann ist*

$$w = f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i f(v_i)$$

Und somit liegt f *in der linearen Hüllen.*

<= *Zeige* f *injektiv. Sei* $f(v) = \vec{0}$. *Wir schreien* v *als* $v = \lambda_1 v_1 + \ldots \lambda_n v_n$ *und somit*

$$\lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = \vec{0}$$

Hier sind alle Vektoren v_i *linear unabhängig, deshalb müssen alle Skalre* $\lambda_i = 0$ *sein.*
Es ist zu zeigen, dass f *surjektiv ist. Sei* $w \in W$.

$$w = \lambda_i f(v_1) + \ldots + \lambda_n f(v_n) = f(\sum_{i=1}^{n} \lambda_i v_i)$$

□

Korollar Beliebige zwei Vektorräume derselben Dimension über einem Körper K sind isomorph.

$$\forall V, W : dim_K V = dim_K W \Leftrightarrow V \cong W$$

Beispiel $K^n = K \times K \times \ldots \times K$ hat die Dimension n.

Definition 15. *Der Rang von einer linearen Abbildung* f *ist*

$$rg(f) = Dimension\ von\ Im(f)$$

falls die Dimension endlich ist.

Satz 9. *$f : V \to W$ ist lineare Abbildung und $dimV = n$*

$$n = dim(Ker(f)) + rg(f)$$

Beweis 9. *$Ker(f)$ ist Unterraum von V. Sei $\{v_1, \ldots, v_r\}$ die Basis von $Ker(f)$. Wir ergänzen diese Basis zu $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ der Basis von V.*
Zeige $\{f(v_{r+1}), \ldots, f(v_n)\}$ (jetzt auch als $\{w_{r+1}, \ldots, w_n\}$ dargestellt) ist Basis von $Im(f)$.
Sei $v = \sum_{i=1}^{n} \lambda_i v_i$

$$f(v) = \lambda_1 f(v_1) + \ldots + \lambda_r f(v_r) + \lambda_{r+1} f(v_{r+1}) + \ldots + \lambda_n f(v_n) = \lambda_{r+1} f((v_{r+1}) + \ldots + \lambda_n f(v_n)$$

Daraus folgt $Lin(\{f(v_{r+1}), \ldots, f(v_n)\}) = Im(f)$. Jetzt bleibt es zu zeigen, dass die Menge $\{f(v_{r+1}), \ldots, f(v_n)\}$ linear unabhängig ist.

$$\begin{aligned}
\alpha_1 w_{r+1} + \ldots + \alpha_{n-r} w_n &= \vec{0} \\
\Rightarrow f(\alpha_1 v_{r+1} + \ldots + \alpha_{n-r} v_n) &= \vec{0} \\
\Rightarrow \alpha_1 v_{r+1} + \ldots + \alpha_{n-r} v_n &\in Ker(f) \text{ Kann als Linearkomb. geschrieben werden} \\
\alpha_1 v_{r+1} + \ldots + \alpha_{n-r} v_n &= \lambda_1 v_1 + \ldots \lambda_r v_r \\
\Rightarrow \lambda_1 v_1 + \ldots + \lambda_r v_r - \alpha_1 v_{r+1} - \ldots - \alpha_{n-r} v_n &= \vec{0} \\
\Rightarrow \lambda_1 = \ldots = \lambda_r = \alpha_1 = \ldots = \alpha_{n-r} &= 0
\end{aligned}$$

Und somit ist es also linear unabhängig. □

6. Vorlesung, 5.11.2007

6.0.7. Lineare Abbildungen und Matritzen

zunächst Abstrakt: Wir nehmen uns einen Körper K. Dann ist eine $m \times n$ Matrix A über K eine rechteckige Anordnung von $n * m$ Elementen aus K.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} alle\ a_{i,j} \in K$$

i ist dabei der Zeilenindex und j der für die Spalten.

Definition 16. $M(m \times n, K)$ *ist die Menge aller* $m \times n$ *Matritzen über* K.

Für $A, B \in M(m \times n, K)$ gilt

- $A = B \Leftrightarrow \forall i, j a_{i,j} = b_{i,j}$
- $M(m \times n, K)$ ist K-Vektorraum.
  - Addition von Matritzen ergibt

$$A + B = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \dots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \dots & a_{2,n} + b_{2,n} \\ \vdots & & & \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \dots & a_{m,n} + b_{m,n} \end{pmatrix}$$

  - Die Multiplikation mit Skalaren funktioniert so, dass jeder Eintrag mit dem Skalar multipliziert wird:

$$\lambda * A = (\lambda a_{i,j})$$

Matrix-Vektorprodukt

Wir nehmen uns eine Matrix und einen Vektor und wollen folgendes erhalten:

$$M(m \times n, K) \times K^n \to K^m$$

$$(A,(x_1,\dots x_n)) \to A*x = (\sum_{i=1}^{n} a_{1,i}x_i,\dots,\sum_{i=1}^{n} a_{m,i}x_i)$$

Und das jetzt anschaulicher:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} * \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1,i}x_i \\ \sum_{i=1}^{n} a_{2,i}x_i \\ \vdots \\ \sum_{i=1}^{n} a_{m,i}x_i \end{pmatrix}$$

Zusammenhang zwischen lineare Abbildung und Matrix

Satz 10. *Gegeben sei eine Matrix $A \in M(m \times n, K)$. Dann ist die Abbildung $K^n \to K^m$ gegeben durch $x \to Ax$ eine lineare Abbildung.*
Umgekehrt, falls $f : K^n \to K^m$ eine lineare Abbildung ist, dann gibt es dazu genau eine eindeutige Matrix $A \in M(n \times m, K)$ mit $\forall x \in K^n : f(x) = Ax$

Beweis 10.

- $A(x+y) = Ax + Ay$ *und* $A(\lambda x) = \lambda A(x)$ *daraus folgt,* Ax *ist eine lineare Abbildung.*
- *Eindeutigkeit: Sei* $A, B \in M(n \times m, K)$ *mit* $f(x) = Ax + Bx \quad \forall x \in K^n$. *Insbesondere für die Einheitsvektoren* e_i *mit* 1 *an der* i*ten Stelle:*

  $$Ae_i = Be_i \quad i = 1,\dots,n$$

  $$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ \vdots & & & \\ a_{j,1} & a_{j,2} & \dots & a_{j,n} \\ \vdots & & & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} * \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{m,i} \end{pmatrix}$$

  Für B *analog. Demnach gilt* $\forall k, j a_{k,j} = b_{k,j}$
- *Existenz: Sei* $v_i = f(e_i)$.

  $$f(e_i) = \begin{pmatrix} v_{1,i} \\ \vdots \\ v_{m,i} \end{pmatrix} \in K^n$$

  muss i*te Spalte von* A *sein. Damit* $Ae_i = f(e_i)$ *und wegen linearitt von* A *und* f *gilt*

  $$A(\lambda_1 e_1 + \dots + \lambda_n e_n) =?(\lambda_1 e_1 + \dots + \lambda_n e_n)$$

Merke Die Bilder der Einheitvektoren definieren Spalten der Matrix!

$f : V \to W$ und V, W sind K-Vektorräume mit $dimV = n$ und $dimW = m$. Die Basis von V sei $B = \{v_1, \ldots, v_n\}$ und von W $C = \{w_1, \ldots, w_m\}$.

- Die Darstellung von $v \in V$ bezüglich B ist eindeutig.

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

$$v = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}_B$$

- f ist bestimmt durch Wirkung auf Basis.

$$f(\lambda_1 v_1) = f(\lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_B) = \lambda_1 a_{1,1} w_1 + \ldots \lambda_1 a_{m,1} w_n$$

$$f(\lambda_2 v_2) = f(\lambda_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}_B) = \lambda_2 a_{1,2} w_1 + \ldots \lambda_2 a_{m,2} w_n$$

$$f(\lambda_n v_n) = f(\lambda_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}_B) = \lambda_n a_{1,m} w_1 + \ldots \lambda_n a_{m,n} w_n$$

$$\sum = f(v) = f(\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}_B) = \begin{pmatrix} \lambda_1 a_{1,1} + \lambda_2 a_{1,2} + \ldots + \lambda_n a_{1,n} \\ \vdots \\ \lambda_1 a_{m,1} + \lambda_1 a_{m,2} + \ldots + \lambda_n a_{m,n} \end{pmatrix}$$

Die Matrix

$$A_f = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}$$

und das Matrix-Vektorprodukt:

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix} * \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1,i} \lambda_i \\ \vdots \\ \sum_{i=1}^{n} a_{m,i} \lambda_i \end{pmatrix}$$

7. Vorlesung, 7.11.2007

Beispiele

1. $V = W = \mathbb{R}^2$. Wir wollen eine lineare Abbildung haben, die die Drehung um Winkel ϕ realisiert.
   Für Einheitsvektor $(1,0) \to (\cos\phi, \sin\phi)$ und $(0,1) \to (-\sin\phi, \cos\phi)$. Damit sieht die Matrix wie folgt aus:
   $$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

2. Wir wollen um Faktor $\lambda \in \mathbb{R}$ strecken. $(1,0) \to (\lambda, 0)$ und $(0,1) \to (0,\lambda)$. Damit ist die Matrix
   $$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

3. $V = W = \mathbb{R}^3$. Nun wollen wir eine Drehung um ϕ in der x_1, x_3-Ebene. x_2 bleibt dabei fix.
   $$(1,0,0) \to (\cos\phi, 0, \sin\phi)$$
   $$(0,1,0) \to (0,1,0)$$
   $$(0,0,1) \to (-\sin\phi, 0, \cos\phi)$$
   Und damit ist die Matrix
   $$\begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix}$$

4. Spiegelung an der x-Achse in $\mathbb{R}^2$.
   $$(1,0) \to (1,0)$$
   $$(0,1) \to (0,-1)$$
   Dann ist die Matrix
   $$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

5. Wir machen jetzt das gleiche, bloß mit einer anderen Basis $B = \{v_1, v_2\}$. $v_1 = (1,1)$ und $v_2 = (1,-1)$.

   $$(1,0)_B \to (0,1)_B = 0v_1 + 1v_2$$

   $$(0,1)_B \to (1,0)_B$$

   Dann ist die Matrix

   $$A_{BB} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

   $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} * (2,2)_B = (2,2)_B$$

6. Spiegelung an der Geraden durch Ursprung. Die Gerade, durch die wir spiegeln, hat eine Winkel zu x-Achse von $\frac{\phi}{2}$. Dann geht $(1,0) \to (\cos\phi, \sin\phi)$ und $(0,1) \to (\sin\phi, -\cos\phi)$ Die Matrix sieht dann wieder so aus:

   $$\begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

7.0.8. Komposition von linearen Abbildungen

Wir haben zwei Matrizen

$$A_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

$$und \;\; A_\psi = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}$$

dann ist

$$A_{\phi-\psi} = \begin{pmatrix} \cos(\phi+\psi) & -\sin(\phi+\psi) \\ \sin(\phi+\psi) & \cos(\phi+\psi) \end{pmatrix}$$

Situation: $V \to W \to Y$ mit $V \subset K^n$, $W \subset K^m$ und $Y \subset K^r$.
Dann $g \leftrightarrow B \in M(m \times n, K)$ und $f \leftrightarrow A \in M(r \times m, K)$. Dann ist $f \circ g \leftrightarrow C \in M(r \times n, K)$.

Definition 17. *Sei $A \in M(r \times m, K)$ und $B \in M(m \times n, K)$. Dann sei*

$$C = A \cdot B = (c_{i,j}) \in M(r \times n, K) \quad mit\ i = 1, \ldots, r\ ,\ j = 1, \ldots, n$$

mit $c_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}$.

$(0,0,\ldots,1,0,0,0)$ ist e_i mit 1 an Stelle i.

$$e_i \to_B^g (b_{1,i}, b_{2,i}, \ldots (b_{m,i}) \to_A^f \underbrace{\begin{pmatrix} a_{1,1}b_{1,i} + \ldots + a_{1,m}b_{m,i} \\ \vdots \\ a_{r,1}b_{1,i} + \ldots + a_{r,m}b_{m,i} \end{pmatrix}}_{ite\ Spalte\ in\ C}$$

Beispiel Wir wollen uns die Hintereinanderschaltung von Drehungen angucken.

$$\begin{aligned} A_\psi * A_\phi &= \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} * \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \\ &= (...) \\ &= A_\phi * A_\psi \end{aligned}$$

Aber im Allgemeinen ist die Matritzenmultiplikation NICHT kommutativ!.

Eigenschaften der Matritzenmultiplikation

1. Assoziativ. $(A * B) * C = A * (B * C)$
2. Im Allgemeinen nicht kommutativ. Es existieren echte Nullteiler. Das bedeutet $A * B = 0$, wobei 0 die Nullmatrix ist und $A \neq 0, B \neq 0$.
3. Die Einheitsmatrix $E_n \in M(n \times n, K)$ ist eine Matrix, die nur Einsen auf der Diagonalen hat und ansonsten Null ist.

8. Vorlesung, 12.11.2007

Die inverse Matrix zu A wird als A^{-1} bezeichnet und erfüllt die Gleichung

$$A * A^{-1} = E_n$$

Einfache Fakten

1. A invertierbar, dann muss $A \in M(n \times n, K)$
2. $A, B \in M(n \times n, K)$ dann gilt
$$A * B = E \Leftrightarrow B * A = E \Leftrightarrow B = A^{-1}$$
3. A sei invertierbar und A^{-1} auch invertierbar. dann ist $(A^{-1})^{-1} = A$
4. $A, B \in M(n \times n, K)$ invertierbar. Dann ist
$$(A * B)^{-1} = B^{-1} * A^{-1}$$

8.0.9. Invertieren von Matrizen

Wann ist eine Matrix invertierbar? Diese Überlegung führt uns direkt zum Begriff des Ranges. Wir wissen schon, dass für $f : V \to W$ der Rang $rg(f) = dim(Im(f))$ ist.

Rang der Matrix

Für A_f mit $A \in M(m \times n, K)$ wird der Spaltenrang von A definiert als linear unabhängige Spalten in A. Der Zeilenrang von A wäre demnach analog die maximale Anzahl an linear unabhängigen Zeilen in A.

Beispiel

- $$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$
Hier ist der Spaltenrang gleich dem Zeilenrang gleich 3. Demnach ist $rg(A) = 3$.

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

  Hier ist der Spaltenrang gleich Zeilenrang gleich 2. Hier ist $rg(A) = 2$

Satz 11. *Für jede Matrix gilt* $Spaltenrang = Zeilenrang = Rang\ der\ Matrix.$

Beweis 11. $rg(A) = Spaltenrang$ *ist einfach, da die Spalten genau die Bilder der linearen Abbildung sind.*
Für $Zeilenrang = Spaltenrang$ *benutzen wir einen Hilfssatz:*
$A \in M(m \times n, K)$. *Sei* j*te Spalte linear abhängig von anderen Spalten dann gilt*

$$A' \in M(m \times (n-1), K) \ \ und \quad A' = A \backslash jte\ Spalte$$

Dann gilt $Zeilenrang\ A = Zeilenrang\ A'$.
Beweis des Hilfssatzes: Sei $Z = \{z_i | i \in I\}$ *die Menge der linear unabhängigen Zeilen in* A. *Sei* $Z' = \{z_i' | i \in I\}$ *mit* Z' *fehlt* j*te Zeile.*
Sei $\sum_{i \in I} \alpha_i z_i' = \vec{0}$. *Es ist zu zeigen, dass alle* $a_i = 0$ *sind.*

$$\sum_{i \in I} \alpha_i z_{i,j} = \sum_{i \in I} \alpha_i \left(\sum_{k=1 \ldots n, k \neq j} \lambda_k z_{i,k} \right) = \sum_{k \neq j} \lambda_k \underbrace{(\sum_{i \in I} \alpha_i z_{i,k})}_{=\vec{0}} = \sum \lambda_k \vec{0} = \vec{0}$$

Analog dazu verändert das Streichen einer linear abhängigen Zeile den Spaltenrang. Damit ist der Hilfssatz bewiesen.

Wir streichen in A *nun die linear abhängigen Zeilen/ Spalten so lange wie möglich. Daraus bekommen wir eine Matrix* $A' \in (m' \times n', K)$. *Daraus folgt, dass* $Spaltenrang = n' \wedge Zeilenrang = m'$. *Nun eine neue Behauptung* $n' = m'$. *Und indirekt* $n' \leq m'$. *Dies ist aber ein Widerspruch, da man die Spalten der Länge* n *als Vektoren behandeln kann und es kann aber maximal* n *linear unabhängige Vektoren geben. Also muss* $n' = m'$ *gelten.*
$A \in M(m \times n, K)$ *dann ist* $A^t \in M(n \times m, K)$ *die transponierte Matrix zu* A, *wenn Zeilen und Spalten vertauscht werden. Daraus folgt, dass* $rg(A) = rg(A^t)$. □

Bestimmung des Rangs

Wir wollen nun den Rang tatsächlich bestimmen. Dabei beobachtet man, dass man einigen Matrizen den Rang direkt ansieht. Z.B. ist das bei einer Matrix in oberer Dreiecksform so. Da ist der Rang dann $rg(A) = k$, dass ist die Anzahl der Zeilen, die nicht null sind (der Rest hat Einsen auf der Diagonalen, davor nullen und dahinter ist wurscht).
Wenn man nun also eine Matrix in Dreiecksform überführen kann, so dass der Rang sich nicht verändert, dann ist das leicht zu bestimmen.

Überführung in Dreiecksform Elementare Umformungen, die den Rang nicht verändern:

Typ 1 Vertauschung von Zeilen bzw. Vertauschung von Spalten

Typ 2 Multiplikation einer Zeile bzw. einer Spalte mit Skalar $\neq 0$.

Typ 3 Addition des λ-fachen einer Zeile bzw. Spalte zu einer anderen Zeile bzw. Spalte, wobei natürlich immer zu Zeile zu Zeile und Spalte zu Spalte.

Ein Algorithmus für Typ 3 sähe wie folgt aus:

```
Runde 0: A0 = A
Runde k: Falls Matrix die Form
   a11         *
      ...    *
     0   akk  *
   -------------
       0  | B
Falls B=0 -> rg A = k, sonst Existiert brs != 0
  -> bringe brs an die Stelle k+1, k+1,
     mit Addition des l-fachen einer Zeile k+1,k+1 zu 0 machen
      bk+1,k+1 - (bk+i,k+1/bk+1,k+1)*bk+1,k1
```

Umformungen als Matrizenmultiplikation

Typ 1 $T_{i,j}$ ist eine quadratische Matrix mit Größe $m \times m$ bzw. $n \times n$ und sieht so aus, dass man sich die Einheitsmatrix nimmt, und die 1 in Zeile i in die Spalte j verrutscht und die in Zeile j nach Spalte i rutscht. Dann vertauscht $T_{i,j} * A$ die Zeile i mit Zeile j und $A * T_{i,j}$ vertauscht Spalte i mit Spalte j.

Typ 2 $\lambda \neq 0$. Sei $S_{i,\lambda}$ die Matrix, auf der auf der Diagonalen nur Einsen stehen und an Stelle i natürlich dann λ. Dann ist $S_{i,\lambda} * A$ die Skalierung der i-ten Zeile und $A * S_{i,\lambda}$ die Skalierung der i-ten Spalte.

Typ 3 Die Matrix $K_{i,j,\lambda}$ sieht so aus, dass man auf der Diagonalen wieder überall Einsen hat und nur an der Stelle i, j λ steht. Dann ergibt sich durch $K_{i,j,\lambda} * A$, dass die Zeile i mit dem λ-fachen von Zeile j addiert wird. Und dann durch $A * K_{i,j,\lambda}$ dann natürlich die Spalten.

9. Vorlesung, 14.11.2007

Wir rechnen das für den Typ 3 nun einfach mal nach, um zu gucken, ob das tatsächlich stimmt. Sei $a'_{k,j}$ der Eintrag der $k$te Zeile mit $\lambda * j$-te Spalte in $K_{i,j,\lambda}$

$$a'_{k,j} = \sum_{l=1}^{n} a_{k,l} * k_{l,j} = a_{k,i} * \lambda + a_{k,j}$$

Beispiel Wir wollen die Matrix in die Dreiecksform bringen:

$$\begin{pmatrix} 1 & -4 & 2 & 0 \\ 2 & -3 & -1 & -5 \\ 3 & -7 & 1 & -5 \\ 0 & 1 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -4 & 2 & 0 \\ 0 & 5 & -5 & -5 \\ 0 & 5 & -5 & -5 \\ 0 & 1 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -4 & 2 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -4 & 2 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Wie man leicht sieht, ist der Rang also 2.

Jetzt aber mal wieder dahin, was wir eigentlich machen wollten, nämlich die Bestimmung der inversen Matrix.

Bestimmung der inversen Matrix

Wir wollen zu $A \in M(n \times n, K)$ die inverse Matrix bestimmen durch Zeilentranspositionen. Es muss gelten

$$A * A^{-1} = E$$

$$\begin{aligned} K_{i,j,\lambda} * A * A^{-1} &= K_{i,j,\lambda} * E \\ \vdots \qquad & \qquad \vdots \\ E &= A^{-1} \end{aligned}$$

Das heißt, wenn wir links und rechts immer die selben Operationen durchführen, dann erhalten wir irgendwann links die Einheitsmatrix und haben dann rechts automatisch die inverse Matrix zu stehen.

Dazu gucken wir uns folgendes Beispiel an:

$$\overbrace{\begin{pmatrix}1&1&1&1\\1&1&2&1\\0&-1&0&1\\1&0&0&2\end{pmatrix}}^{A}\quad\overbrace{\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}}^{E}$$

$\to 2Z-1Z, 4Z-1Z:$

$$\begin{pmatrix}1&0&1&1\\0&1&1&0\\0&-1&0&1\\0&0&-1&1\end{pmatrix}\quad\begin{pmatrix}1&0&0&0\\-1&1&0&0\\0&0&1&0\\-1&0&0&1\end{pmatrix}$$

$\to 3Z+2Z:$

$$\begin{pmatrix}1&0&1&1\\0&1&1&0\\0&0&1&1\\0&0&-1&1\end{pmatrix}\quad\begin{pmatrix}1&0&0&0\\-1&1&0&0\\-1&1&1&0\\-1&0&0&1\end{pmatrix}$$

$\to 1Z-3Z, 2Z-3Z, 4Z+3Z$

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

$\to 4Z*\frac{1}{2}, 3Z-4Z, 2Z+4Z$

$$\underbrace{\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}}_{E}\quad\underbrace{\begin{pmatrix}2&-1&-1&0\\-1&0,5&-0,5&0,5\\0&0,5&0,5&-0,5\\-1&0,5&0,5&0,5\end{pmatrix}}_{A^{-1}}$$

Und wenn man das jetzt alles mit $A*A^{-1}$ nachrechnet, dann kommt tatsächlich E raus, wow.

9.0.10. Lineare Gleichungssysteme

Definition 18 (LGS). *Ein lineares Gleichungssystem mit Koeffizienten in einem Körper K mit m Gleichungen in insgesamt n Unbekannten wird dargestellt durch eine Matrix $A \in M(m\times n, K)$ und einen Vektor $b \in M(m\times 1, K) = K^m$*

$$\begin{aligned}&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1\\&\vdots\\&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m\end{aligned}$$

Der Vektor $(x_1, \ldots, x_n) \in K^n$ heißt die Lösung des LGS falls $A*(x_1, \ldots, x_n) = (b_1, \ldots, b_m)$. Mit $A|b$ bezeichnet man die $(m \times (n+1))$ Matrix, die als letzte Einträge noch den Vektor b enthält.

Es gilt folgender Satz

Satz 12. *$A * x = b$ ist lösbar, d.h. hat wenigstens eine Lösung, genau dann wenn $rg(A) = rg(A|b)$ gilt.*

Beweis 12. *$rg(A) = rg(A|b)$ gilt, wenn die Spaltenränge gleich sind. Dies gilt jedoch nur, wenn $b \in Lin(a_{11}, a_{21}, \ldots, a_{m1}), \ldots, (a_{1n}, a_{2n}, \ldots, a_{mn}))$. Dann existieren aber λ_i mit*

$$\lambda_1 * (a_{11}, a_{12}, \ldots a_{1n}) + \ldots + \lambda_n * (a_{m1}, a_{m2}, \ldots, a_{mn}) = b$$

Deshalb gilt also $A * (\lambda_1, \ldots, \lambda_n) = b$ □

Definition 19.

- *Ein LGS heißt homogen, falls $b = \vec{0}$.*
- $Loes(A, b) = \{x \in K^n | A * x = b\}$

Satz 13. *Sei $f : K^n \to K^m$ die zu A gehörende lineare Abbildung.*

1. *$Ker(f) = \{x \in K^n | A * x = \vec{0}\}$ ist Unterraum des K^n.*
2. *Seien $x', x'' \in Loes(A, b) \Rightarrow x' - x'' \in Ker(f)$*
3. *Ist $x \in Loes(A, b)$ und $x' \in Ker(f)$ dann $x + x' \in Loes(A, b)$*

Beweis 13.

1. *K-Vektorraum*
2. *$f(x' - x'') = f(x') - f(x'') = b - b = \vec{0}$*
3. *$f(x + x') = f(x) + f(x') = b + \vec{0} = b$*

Satz 14. *$A \in M(m \times n, K)$. Sei $b \in K^m$ und $B \in M(m \times m, K)$ invertierbar.*

$$Loes(A, b) = Loes(B * A, B * b)$$

Beweis 14. *Es wird Hin- und Rückrichtung gezeigt.*

=> *Sei $x \in Loes(A, b)$. Dann ist $A * x = b$ und demnach $B * A = B * b$*

<= *Sei $x \in Loes(B * A, B * b)$. Dann ist $B * A * x = B * b$. Es existiert also B^{-1}, dann ist $B^{-1} * B * A * x = B^{-1} * B * b \Leftrightarrow A * x = b$.*

10. Vorlesung, 19.11.2007

Die in der vorhergehenden Vorlesung bewiesene Tatsache, dass $Loes(A, b) = Loes(B*A, B*b)$ gilt, ist die Grundlage dafür, dass wir die Zeilentransformation anwenden dürfen.

Gaußalgorithmus zur Lösung von LGS

Schritt 1 Bringe Matrix durch elementare Zeilentransformation in Gauß-Jordan–Form. D.h. in eine Zeilenstufenform mit der Eigenschaft, dass alle Leitkoeffizienten sind gleich 1 und oberhalb der Leitkoeffizienten alles 0.
Ein kleines Beispiel

$$A|b = \begin{pmatrix} 1 & 1 & 5 & 0 & 0 & 9 & 6 \\ 0 & 2 & 4 & 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 1 & 0 & 5 & 6 \\ 0 & 3 & 6 & 0 & 1 & 7 & 12 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 3 & 0 & 0 & 8 & 2 \\ 0 & 1 & 2 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 & 0 & 5 & 6 \\ 0 & 3 & 6 & 0 & 1 & 7 & 12 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & b \\ 1 & 0 & 3 & 0 & 0 & 8 & 2 \\ 0 & 1 & 2 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 1 & 4 & 0 \end{pmatrix} = A'|b'$$

Testen der Ränge als Bedingung für Lösbatkeit. In dem Beispiel wäre $rg(A) = rg(A|b) = 4$. Das heißt, es existiert eine Lösung.

Schritt 2 Finde Lösungsmenge $H = Loes(A'|\vec{0})$, d.h. für das homogene LGS. Dies funktioniert so, dass man die Unbekannten, die zu Spalten ohne Leitkoeffizienten gehören, als freie Parameter gewählt werden.
In unserem Beispiel wählen wir $x_3 = \lambda$ und $x_6 = \mu$. Also

$$\begin{aligned} x_1 &= -3\lambda - 8\mu \\ x_2 &= -2\lambda - \mu \\ x_3 &= \lambda \\ x_4 &= -5\mu \\ x_5 &= -4\mu \\ x_6 &= \mu \end{aligned}$$

Das ist die Lösung des homogenen LGS. Man kann es schreiben als

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \lambda * \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -8 \\ -1 \\ 0 \\ -5 \\ -4 \\ 1 \end{pmatrix}$$

H ist also die Lineare Hülle, $H = Lin((-3, -2, 1 \ldots), (-8, -1, 0, \ldots))$.

Schritt 3 Wir suchen nun eine spezielle Lösung für $A'x = b'$. Wir setzten dazu die Unbekannten, die zu Spalten ohne Leitkoeffizienten gehören, gleich 0.
Im Beispiel wird also $x_3 = x_6 = 0$ gesetzt. Daraus folgt gleich

$x_1 = 2$
$x_2 = 4$
$x_4 = 6$
$x_5 = 0$

Schritt 4 Wenn w diese spezielle Lösung von $Ax = b$ ist, so ist $Loes(A, b) = Loes(A', b') = w + H$ die Gesamtlösung des LGS.

Geomentrische Interpretation der Lösung eines LGS

$Ax = b$ mit $A = M(m \times n, K), x \in K^n, b \in K^m$. Mit $A \leftrightarrow f$ folgt $Ker(f) = Loes(A, \vec{0})$ ist Unterraum. Und $dim(f) = dim(Im(f)) = rg(A) = n$. $dim(Loes(A, \vec{0}) = n - rg(A)$. Man bemerke, dass die Lösung kein Vektorraum mehr ist, da der Nullvektor nicht mehr enthalten ist.

10.0.11. Determinanten

Unser Ziel ist, eine Matrix $A \in M(n \times n, K)$ durch ein einziges Element aus K zu repräsentieren. Wir werden hier allerdings nur $K = \mathbb{R}$ angucken. Diese Zahl ist die Determinante und wird mit $det(A)$ bezeichnet.
Die Determinante liefert Aussagen über die Invertierbarkeit von A, die Lösung eine Gleichungssystems und zur Volumenberechnen.

Definition 20. $det : M(n \times n, K) \to K$ *heißt Determinantenfunktion, falls*

1. det *linear in jeder Zeile ist. D.h. Zeile* i *mit* $1 \leq i \leq n$ *und* $\lambda, \mu \in K$

$$det(z_1, z_2, \ldots, \lambda z_i + \mu z_i', \ldots, z_n) = \lambda * det(z_1, z_2, \ldots z_i, \ldots z_n) + \mu * det(z_1, z_2, \ldots, z_i', \ldots, z_n)$$

2. $rg(A) < n \Rightarrow det(A) = 0$.

3. $det(E) = 1$ *für Einheitsmatrix.*

Satz 15. *Für jedes n gibt es genau eine Determinantenfunktion.*

Da der Satz sehr aufwendig ist, werden wir hier keinen Beweis führen.

Satz 16. $n = 2$ *und*

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

Beweis 16. *Zeige, dass diese Funktion die Bedingungen 1,2 und 3 erfüllen.*

1.

$$det\begin{pmatrix} \lambda a_1 + \mu a_2 & \lambda b_1 + \mu b_2 \\ c & d \end{pmatrix} = \lambda * det\begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} + \mu\begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}$$

$$(\lambda a_1 \mu a_2) * d - (\lambda b_1 + \mu b_2) * c = \lambda a_1 d - \lambda b_1 c + \mu a_2 d - \mu b_2 c = \lambda det\begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} + \mu\begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}$$

2. $rg(A) = 0$...
   $rgA = 1$, *geht nur, wenn* $(a, c) = \lambda(b, d)$.

$$det(\begin{pmatrix} \lambda b & b \\ \lambda d & d \end{pmatrix} = \lambda bd - b\lambda d = 0$$

3.

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

□

Satz 17 (Entwicklungssatz von Laplace). $detA = \sum_{i=1}^{n} a_{i,k} * (-1)^{i+k} * det(A_{i,k})$ *ist Entwicklung bezüglich $k$ter Spalte, wobei $A_{i,j} \in M((m-1) \times (n-1), K)$, wobei $i$te Zeile und $j$te Spalte gestrichen werden. Und das ist gleich $\sum_{j=1}^{m} a_{k,j}(-1)^{k+j} det(A_{k,j})$ Entwicklung bezüglich $k$ter Zeile.*

11. Vorlesung, 21.11.2007

Die Determinante für die 3×3 Matrix sieht wie folgt aus nach dem Entwicklungssatz

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

Rechenregeln

1. $$det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{i1} + \lambda g_1 & \dots & a_{in} + \lambda g_n \\ \vdots & & \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{i1} & \dots & a_{in} \\ \vdots & & \\ a_{n1} & \dots & a_{nn} \end{pmatrix} + \underbrace{\lambda det \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}}_{=0}$$

2. Austausch von Zeile i und j. Wir haben A gegeben und wollen zu A' gelangen.

   ```
   1. Schritt: Zeile i= Zeile i+ Zeile j.
   2. Schritt: Zeile j= Zeile j-Zeile i = - Zeile i.
   3. Schritt: Zeile i = Zeile i + Zeile j = Zeile j
   ```

   Bis hierhin haben wir die Determinante nicht geändert.

   ```
   4. Schritt: Zeile j = Zeile j * (-1)
   ```

   Demnach ist $det(A') = -det(A)$!

3. Was ist die Determinante von $\lambda * A$? Hier kann aus jeder Zeile der Faktor λ rausgezogen werden aufgrund der Linearität. Somit ergibt sich

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

4. Wir haben ne Matrix in Dreiecksform

   $$det \begin{pmatrix} a_{11} & & \\ & a_{22} & * \\ \vdots & & \\ 0 & \dots & a_{nn} \end{pmatrix} = a_{11}det(A_{11}) = a_{11}(a_{22}det(?)) = \prod_{i=1}^{n} a_{ii}$$

5. $detA = detA^t$

11.0.12. Determinante und Inverse Matrix

Sei $A \in M(n \times n, K)$ Matrix mit vollem Rang.

Definition 21. *Komplementäre Matrix $\tilde{A}$ mit*

$$\tilde{a}_{i,j} = (-1)^{i+j} det(A_{j,i})$$

Satz 18.

$$A * \tilde{A} = \begin{pmatrix} detA & & & 0 \\ & detA & & \\ & & \ddots & \\ 0 & & & detA \end{pmatrix}$$

Beweis 18.

$$\tilde{A} = \begin{pmatrix} detA_{11} & -detA_{21} & detA_{31} \\ -detA_{12} & detA_{22} & \dots \\ \vdots & & \end{pmatrix}$$

$$\begin{aligned} detA &= \sum_{j=1}^{n} a_{ij} * (-1)^{i+j} detA_{ij} \\ &= \sum_{j=1}^{n} a_{ij} * \tilde{a}_{ij} \\ &= \textit{Eintrag von } A * \tilde{A} \textit{ an Stelle } i,j \end{aligned}$$

□

Korollar Wenn $A \in M(n \times n, K)$ und $detA \neq 0$, dann ist

$$A^{-1} = \frac{1}{detA} * \tilde{A}$$

Mit Hilfe dieser Formel schreiben wir uns jetzt explizit das inverse von einer 2×2 Matrix hin

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

Cramersche Regel zur Lösung von LGS

Funktioniert allerdings nur, wenn die Matrix vollen Rang hat, also $detA \neq 0$.
Die Behauptung ist, dass dann das LGS $Ax = b$ eine eindeutige Lösung hat. Daraus folgt, A

entspricht einem Isomorphismus $K^n \to K^n$ und das heißt wiederum, dass b genau ein Urbild hat. Dieses Urbild (die Lösung) wollen wir jetzt bestimmen.

$$x_1 * \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix} + x_2 * \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{pmatrix} + \ldots + x_n * \begin{pmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nn} \end{pmatrix} = * \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$x_1 * \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix} + \ldots + * \begin{pmatrix} ??a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix} = \vec{0} \Rightarrow \quad det \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} = 0$$

Wegen Linearität von Determinatne bezüglich der Spalten ergibt sich

$$x_i * detA - det \begin{pmatrix} a_{11} & b_1 & a_{1n} \\ \vdots & & \\ a_{n1} & b_n & a_{nn} \end{pmatrix} = 0$$

$$x_i = \frac{det \begin{pmatrix} a_{11} & b_1 & a_{1n} \\ \vdots & & \\ a_{n1} & b_n & a_{nn} \end{pmatrix}}{detA}$$

Determinante für das Matrixprodukt

Das Ziel ist $det(A * B) = det(A) * det(B)$.

Lemma $A \in M(n \times n, K)$ invertierbar, genau dann wenn $detA \neq 0$.

Beweis Wenn invertierbar, dann folgt aus den Axiomen $rg(A) \neq 0$. Wenn der $rg(A) \neq 0$ ist, dann kann man mittels Zeilenumformun zu E gelangen, doch hier ist der Rang nicht 0, sondern 1.

Satz 19. $det(A * B) = det(A) * det(B)$.

Beweis 19. *Wir fixieren das B. Wir definieren uns eine Abbildung $f : M(n \times n, K) \to K$ mit $A \to det(A * B)$. f hat mehrere Eigenschaften*

- *f ist linear in Zeile von A*
- *Abbildung der $i$ten Zeile von $A \to i$te Zeile von $A * B$ ist linear*
- *$rgA < n$, so auch $rg(A * B) < n$*

- $f(E) = det(EB) = detB$
  Falls $detB \neq 0$, *so erfüllt die Abbildung* $A \rightarrow (detB)^{-1} * det(A * B)$ *alle axiomatischen Forderungen an Determinantenfunktion. Es gibt aber nur eine solche Funktion. Daraus folgt* $detA = (detB)^{-1} = det(A * B)$ *also* $detA * detB = det(A * B)$.
  Falls $detB = 0$ *dann folgt dass der* $rg(B) < n$ *ist.* $dim(Ker) > 0$ *und somit auch* $dim(Ker(A * B) > 0$.
  Damit ist also der Rang echt kleiner n *und die Determinante somit* 0.

13. Vorlesung, 28.11.2007

13.0.13. Euklidische Vektorräume

Dies ist ein Vektorraum über $\mathbb{R}$. Die Idee ist, wir wollen Geometrie betreiben. Dazu brauchen wir den Begriff der Länge und den des Winkels. Ebene und Gerade sind ja schon bekannt.

Definition 22. *V ist euklidischer Vektorraum, falls V ein $\mathbb{R}$-Vektorraum ist und ein Skalarprodukt hat.*

$$< , > \quad V \times V \to \mathbb{R}$$

Es hat die folgenden Eigenschaften.

1. *Die Abbildung ist bilinear:*

$$\forall \vec{v} \in V :< -, \vec{v} >: V \to \mathbb{R} \qquad < \vec{v}, - >: V \to \mathbb{R}$$

2. *Symmetrie:* $\forall \vec{u}, \vec{v} \in V :< \vec{u}, \vec{v} >=< \vec{v}, \vec{u} >$
3. *positive Definiertheit* $\forall x \neq \vec{0} :< x, x >\geq 0$ *und* $< \vec{0}, \vec{0} >= 0$.

Beispiele

1. $V = \mathbb{R}^n$

$$< x, y >= \sum_{i=0}^{n} x_i y_i \quad Standardskalarprodukt\ in\ \mathbb{R}^n$$

2. $A \in M(m \times n, K)$ invertierbar. Dann erhält mein ein neues Skalarprodukt durch

$$< x, y >=< Ax, Ay >$$

3. $V = \{stetige\ Funktion\ f[-1, 1] \to \mathbb{R}\}$
   $< f, g >= \int_{-1}^{1} f(x)g(x)dx$ ist Skalarprodukt.

Norm

Definition 23 (Norm). *$(V, <, >)$ sei ein Euklidischer Vektorraum. Dann ist*

$$\forall x \in V : ||x|| = \sqrt{< x, x >}\ heit\ Norm\ von\ x$$

Die Norm gibt übrigens die Länge an, also entspricht ihr.

Beispiel $\mathbb{R}^3$ Standardskalarprodukt. Dann ist also $||x|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$.

Satz 20 (Ungleichung von Cauchy-Schwarz). $\forall x, y : | < x, y > | \leq ||x|| * ||y||$

Beweis 20.

Fall 1 $y = 0$, *dann folgt* $0 = 0$ *ist wahre Aussage.*

Fall 2 $y \neq 0$. *Dann gibt es ein* $\lambda = \frac{<x,y>}{||y||^2} \in \mathbb{R}$.

$$\begin{aligned} 0 \leq & < x - \lambda y, x - \lambda y > = < x, x > -2\lambda < x, y >^2 + \lambda < y, y > \\ & = ||x||^2 - \frac{2 < x, y >^2}{||y||^2} + \frac{< x, y >^2}{||y||^4} * ||y||^2 \\ & = ||x||^2 - \frac{< x, y > 2}{||y||^2} \quad | * ||y||^2 \\ & = \frac{||y||^2}{||x||^2||y||^2 - < x, y >^2} \\ \Leftrightarrow & = | < x, y > \leq ||x|| * ||y|| \end{aligned}$$

Eigenschaften der Norm

1. $\forall x : ||x|| \geq 0$.
2. $||\lambda x|| = |\lambda| * ||x||$ mit $\lambda \in \mathbb{R}$
3. Dreiecksungleichung

$$\forall x, y : ||x + y|| \leq ||x|| + ||y||$$

Jetzt wollen wir den Winkel definieren.

Winkel

$x, y \in V$ und seien beide nicht der Nullvektor. Dann definieren wir

$$\cos \alpha(x, y) = \frac{< x, y >}{||x|| * ||y||}$$

und $0 \leq \alpha(x, y) \leq \pi$. Dies ist wohldefiniert, da

$$-||x|| * ||y|| \leq < x, y > \leq ||x|| * ||y||$$

Hinweis Kosinussatz

$$c^2 = a^2 + b^2 - 2ab\cos\phi$$

$||x-y||^2 =< x-y, x-y >$ Dann folgt aufgrund der Linearität $||x||^2+||y||^2-\frac{2<x,y>||x||*||y||}{||x||*||y||}$.

Definition 24 (Orthogonalität). $u, v \in V$ *heißen orthogonal zueinander, falls* $< u, v >= 0$. *Die Notation ist* $u \perp v$.

Diese Definition verträgt sich mit der des Winkels:

$$\cos\alpha(x,y) = \frac{< x, y >}{||x|| * ||y||} = 0 \quad \Rightarrow \alpha = \frac{\pi}{2}$$

Merke Bei Orthogonaltität können die Vektoren der Nullvektor sein, beim Winkel allerdings nicht!!

Definition 25. $M \subseteq V$ *ist Teilmenge.*

$$M^\perp = \{v \in V | v \perp u \ \ \forall u \in M\}$$

Satz 21. $M^\perp$ *ist Untervektorraum von* V.

Beweis 21. *Dafür müssen alle Kriterien für Untervektorräume überprüft werden:*
Der Nullvektor muss enthalten sein: $\vec{0} \subseteq M^\perp$.
Muss gegenüber der Addition abgeschlossen sein: $x, y \in M^\perp \ u \in M$. *Dann ist* $< x + y, u >=< x, u > + < y, u >= 0 + 0 = 0$ *und schlussendlich gegenüber der Multiplikation mit einem Skalar:* $< \lambda x, u >= \lambda < x, u >= \lambda * 0 = 0$. □

Definition 26 (Orthonormalsystem). *Ein Tupel* $(v_1, \ldots, v_r)$ *von Vektoren aus* V *heißt Orthonormalsystem falls*

1. $\forall i : ||x_i|| = 1$ *und*
2. $\forall i \neq j :< v_i, v_j >= 0$

Man beachte, das so ein Orthonormalsystem keine Basis sein muss, aber trotzdem alle Vektoren linear unabhängig sind, das System insgesamt also auch linear unabhängig ist. Das ist auch ganz einfach zu zeigen:
$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_r v_r = \vec{0}$. Für jedes i wenden wir diese Lineare Abbildung $< -, v_i >$ an. Das einzige, was übrig bleibt ist dann $\lambda_i < v_i, v_i >= 0$, und daraus folgt $\lambda_i = 0$.

Angenommen ein Orthogonalsystem bildet eine Basis. $\forall v_i \in V : v = \sum_{i=1}^{n} < v, v_i > v_i$.

Auch hier ist der Beweis kurz

$$v = \sum_{i=1}^{n} \lambda_i v_i \quad | < v_j, - >$$

$$\begin{aligned} < v_j, v > &= < v_j, \sum_{i=1}^{n} \lambda_i v_i > \\ &= < v_j, \lambda_1 v_1 > + < v_j, \lambda_2 v_2 > + \ldots + \underbrace{< v_j, \lambda_j v_j >}_{=\lambda_j} + \ldots + < v_j, \lambda_n v_n > \\ &= \lambda_j \end{aligned}$$

14. Vorlesung, 3.12.2007

$(v_1, \ldots, v_r)$ Orthonormalsystem in V und $U = Lin\{v_1, \ldots, v_r\}$. Dann ist $\forall v \in V : v = u+w$ mit $u \in U$ und $w \in U^\perp$ auf eindeutige Art gegeben.
Wir beweisen das wieder.

Eindeutigkeit $v = u + w = u' + w'$. Daraus folgt $u - u' + w - w' = \vec{0}$, wobei die us aus U kommen und die ws aus $U^\perp$. Daraus folgt dann $< u - u', w - w' >= 0$. $< u - u', u - u' + w - w' >=< u - u', u - u' >= 0$

Existenz Wir setzen $u = \sum_{i=1}^{r} < v, v_i > *v_i$. w ist entsprechend $w = v - u$. Folgendes muss nachgerechnet werden: $< w, v_j > - < v - \sum_{i=1}^{r} < v, v_i > v_i, v_j >$ ist nach der Linearität gleich $< v, v_j > - \sum_{i=1}^{r} < v, v_i >< v_i, v_j >$ und das ist ungleich null für $i = j$. Es bleibt übrig $< v, v_j > - < v, v_j >= 0$.
Wegen $w,_>$ linear folgt $w \perp U$, also $w \in U^\perp$.

Orthonormalisierung nach Erhard Schmidt Sei V ein euklidischer Vektorraum. $(v_1, \ldots, v_r)$ sei ein linear unabhängiges Tupel von Vektoren. Dann ist $Lin\{v_1, \ldots, v_r\}$ ein Unterraum. Wir suchen das Orthonormalsystem $(\tilde{v}_1, \ldots \tilde{v}_r)$ mit $Lin\{v_1, \ldots, v_n\} = Lin(\tilde{v}_1, \ldots, \tilde{v}_r\}$.
Wir werden das induktiv machen. $\tilde{v}_1 = \frac{v_1}{||v_1||}$. Jetzt hat also der Vektor Länge 1. Wir machen weiter

$$\tilde{v}_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} < v_{k+1}, \tilde{v}_i > *\tilde{v}_i}{||v_{k+1} - \sum_{i=1}^{k} < v_{k+1}, \tilde{v}_i > *\tilde{v}_i||}$$

Daraus folgt folgender Satz

Satz 22. *$(\tilde{v}_1, \ldots, \tilde{v}_r)$ ist Orthonormalsystem und $Lin(v_1, \ldots, v_r) = Lin(\tilde{v}_1, \ldots, \tilde{v}_r)$.*

Wenn $(v_1, \ldots, v_r)$ eine Basis von V ist, dann ist auch $(\tilde{v}_1, \ldots, \tilde{v}_r)$ Ortonormalbasis von V.

Isometrien

Wird auch als Orthogonale Abbildungen bezeichnet.

Definition 27. *$f : V \to W$ mit V, W sind endliche Vektorräume heißt Isometrie, wenn gilt*

$$\forall u, v \in V :< u, v >=< f(u), f(v) >$$

Eine Beobachtung ist, dass solche Abbildungen immer injektiv sind. Und zwar weil $v \in Ker(f) \Rightarrow < v, v >=< f(v), f(v) >=< \vec{0}, \vec{0} >= 0$. Daraus folgt $v = \vec{0}$.
Weiter ist erkennbar, wenn $f : V \to V$ eine Isometrie ist, dann ist bei endlicher Dimension von V sogar eine Isomorphie (oder auch Automorphismus).

Satz 23. *Wenn V, W euklidische Vektorräume sind und $(v_1, \ldots, v_n)$ eine Orthonormalbasis von V ist und $f : V \to W$ eine Isometrie ist, dann ist genau dann $(f(v_1), \ldots, f(v_n))$ Orthogonalbasis von W.*

Beweis 23. *Wir müssen beide Richtungen zeigen*

=> $< f(v_i), f(v_j) > - < v_i, v_j >= \delta_{i,j} = 0$ *für* $i = j$ *und sonst* 1.

<= *Angenommen* $< f(v_i), f(v_j) >= \delta_{i,j}$. *Dann stellen wir uns die Vektoren dar als* $v = \sum_{i=1}^{n} \lambda_i v_i$ *und* $u = \sum_{i=1}^{n} \mu_i v_i$. *Wir müssen nachrechnen, dass* $< u, v >=< f(u), f(v) >$ *gilt:*
$< f(\sum_{i=1}^{n} \mu_i v_i), f(\sum_{i=1}^{n} \lambda_i v_i >=< \sum_{i=1}^{n} \mu_i f(v_i), \sum_{i=1}^{n} \lambda_i f(v_i) >= \sum_{i=1}^{n} \sum_{j=1}^{n} \mu_i \lambda_j < f(v_i), f(v_j) >= \sum_{i=1}^{n} \sum_{j=1}^{n} \mu_i \lambda_j \delta_{i,j} = \sum_{i=1}^{n} \sum_{j=1}^{n} \mu_i \lambda_j < v_i, v_j >=< u, v >$

Korollar Wenn A Matrix einer Isometrie ist, dann ist das genau dann der Fall, wenn die Spalten (die Bilder der Einheitsvektoren) Orthonormalsystem bezüglich des Standardskalarprodukts in $\mathbb{R}^n$ sind.
Für ein Orthonormalsystem kann die inverse Matrix also durch transponieren erhalten werden.

Zusammenfassend gilt A gehört zu einer Isometrie, genau dann wenn die Spalten ein Orthonormalsystem bilden. Und das ist genau dann wenn $A^t * A = E$, d.h. $A^{-1} = A^t$. Und das ist gleichbedeutend damit, wenn die Zeilen ein Orthonormalsystem bilden.
Man muss sich merken, dass Isometrien längen-, winkel, und volumenerhaltend sind, und zwar, weil $|detA| = 1$.

14.0.14. Eigenwerte und Eigenvektoren

Wir sind wieder bei beliebigen Vektorräume. Sei $f : V \to V$, also ein Endomorphismus gegeben und V sei ein K-Vektorraum. Wir interessieren und für Vektoren $v \in V$, bei denen f wie eine Skalierung wirkt. Jetzt formal:

Definition 28. $\lambda \in K$ *und* $v \neq \vec{0} \in V$. *Dann heißt* v *Eigenvektor zum Eigenwert* λ, *falls*

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

Dann ist auch $\alpha * v$ Eigenvektor zum Eigenwert λ, da $f(\alpha v) = \alpha f(v) = \alpha \lambda v$.
Besonders schön ist es, wenn Basisvektoren Eigenvektoren sind. In dem Fall wissen wir, dass die zum Endomorphismus gehörige Matrix eine Diagonalmatrix ist, die nur auf der Diagonalen die λ_i erhält, sonst 0en. Mit der Matrix kann man dann ganz schnell die Determinante berechnen und auch die inverse Matrix.

Definition 29. $f : V \to V$ *heißt diagonalisierbar, falls die Basis* $(v_1, \ldots, v_n)$ *existiert, sodass die zugehörige Matrix eine Diagonalmatrix ist.*

15. Vorlesung, 5.12.2007

Nachtrag zu $n \times n$ Matrizen über Körper K.

$G(n, K)$ Das ist die Gruppe, der invertierbaren $n \times n$ Matrizen über K bezüglich der Matrixmultiplikation. Auch "general linear group" genannt.

$Sl(n, K)$ Das ist die Gruppe der invertierbaren $n \times n$ Matrizen über den Körper, die als $detA = 1$ haben. Diese Gruppe wird auch "special linear group" genannt.

$O(n)$ Nicht zu verwechseln mit der oberen Schranke! Das bezeichnet die orthogonale $n \times n$ Matrizen über $\mathbb{R}$, die alle $|detA| = 1$ haben.

$So(n)$ sind die orthogonalen Matrizen, die genau $detA = 1$ haben.

Einfache Beispiele Sei $V = \mathbb{R}^2$.

1. Wir wollen eine Spiegelung an $Lin(v_1)$ machen. Die Matrix sieht dann so aus:

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

   v_1 ist dann Eigenvektor zu Eigenwert $+1$ und v_2 ist der Eigenvektor zu Eigenwert -1.

2. Bei der Drehung von v_1, v_2 um π gibt es keine Eigenvektoren.

3. Die Scherung horizontal um einen Faktor a.

   $$v = \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x + ay \\ y \end{pmatrix}$$

   Für die Einheitsvektoren sieht die Abbildungsmatrix dann so aus

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

   $\binom{1}{0}$ ist Eigenvektor zum Eigenwert $+1$.

Lemma Wenn $v_1, \ldots, v_r$ Eigenvektoren zu paarweise verschiedenen Eigenwerten sind, so sind sie linear unabhängig.
Das muss jetzt natürlich auch noch bewiesen werden.
Seien λ_i die zugehörigen Eigenwerte. Wir machen einen Beweis per vollständiger Induktion

Induktiosanfang $n = 1$, logisch, ein Vektor kann ja nur unabhängig sein.

Induktionsschritt $n \to n+1$

$\alpha_1 v_1 + \ldots + \alpha_r v_r = \vec{0}$. Es ist zu zeigen, dass alle $\alpha_i = 0$ sind. Man multipliziere mit λ_{r+1}, dann ergibt sich $\alpha_1\lambda_{r+1}v_1 + \ldots + \alpha_r\lambda_{r+1}v_r + \alpha_{r+1}\lambda_{r+1}v_{r+1} = \vec{0}$. Durch Anwenden von f auf diese Gleichung folgt $\alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \ldots + \alpha_r\lambda_r v_r + \alpha_{r+1}\lambda_{r+1} + v_{r+1} = \vec{0}$. $\alpha_1(\lambda_{r+1} - \lambda_1)v_1 + \alpha_2(\lambda_{r+1} - \lambda_2)v_2 + \ldots + \alpha_r(\lambda_{r+1} - \lambda_r)v_r = \vec{0}$. Durch Anwenden der Induktionsvoraussetzung folgt für $v_1, \ldots, v_r$, dass alle $\alpha_i(\lambda_{r+1} - \lambda_i) = 0$. Und nach Voraussetzung sind die λs verschieden. Dann muss aber $\alpha_i = 0$ sein für $i = 1 \ldots r$. $\alpha_{r+1} = 0$ wegen $v_{r+1} \neq \vec{0}$ nach Definition.

□

Fakt Alle Eigenvektoren zu einem Eigenwert λ bilden einen Unterraum von V (also, es müsste noch der Nullvektor mit reingenommen werden).
Um das zu beweisen, muss man zeigen, dass für u, v Eigenvektoren zu λ die Eigenschaften für den Unterraum erfüllt sind. Also dass $f(u+v) = f(u) + f(v) = \lambda u + \lambda v = \lambda(u+v)$ und analog für ein Skalar gilt.

Definition 30. *Dieser Unterraum heißt Eigenraum zum Eigenwert λ.*

Zu diesen Eigenräumen gibt es dann natürlich auch wieder je eine Basis. Und deren Basisvektoren sind dann sogar alle, wie gerade bewiesen, linear unabhängig. f ist genau dann diagonalisierbar, falls $\sum_{i=1}^{r} n_i = n$.

Wie findet man nun also so eine Diagonalmatrix zu f, falls sie existiert?
Natürlich brauchen wir die Eigenwerte. Ein kleiner Trick: Ein Vektor $v \neq \vec{0}$ ist Eigenvektor zu Eigenwert λ, genau dann wenn $v \in Ker(f - \lambda Id_V)$.
Müssen wir natürlich erstmal beweisen. $f(x) = \lambda v \Leftrightarrow f(x) - \lambda v = \vec{0}$. Das ist genau dann wenn, $f(x) - \lambda Id_V(x) = \vec{0}$ Und das entspricht $(f - \lambda Id_V)(x) = \vec{0}$. □
Nach der Dimensionsformel der linearen Abbildung gilt $Ker(f - \lambda Id_V) \neq \vec{0}$ genau dann wenn die zugehörige Determinante verschwindet, d.h. ist null.
Sei A ne Matrix zu f bezüglich einer fixierten Basis von V. Dann ist die Matrix für die Skalierung mit λ die Matrix, die auf der Diagonalen nur λs enthält, also λE. Und $A - \lambda E$ ist dann die Matrix zu $f - Id_V$.

Definition 31. $p_f(\lambda) = det(A - \lambda E)$ *heißt charakteristisches Polynom zu f.*

Beispiel Sei $\mathbb{R}^2$ und die Standardbasis. Und sei f gegeben durch $A = \binom{1\ 1}{1\ 1}$, also $f(\binom{\alpha_1}{\alpha_2} = \binom{\alpha_1+\alpha_2}{\alpha_1+\alpha_2}$. Das charakteristische Polynom ist dann

$$det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 1$$

Die Eigenwerte sind jetzt die reellen Nullstellen des Polynoms. Also $\lambda_1 = 2$ und $\lambda_2 = 0$. Jetzt bestimmen wir die zugehörigen Eigenvektoren als nichttriviale Lösungen der homogenen Gleichungssysteme $(A - \lambda_i E)x_i = \vec{0}$ für $i = 1, 2$. Für $\lambda_1 = 2$ ergibt sich folgendes LGS

$$\begin{aligned} -x_1 + x_2 &= 0 \\ x_1 - x_2 &= 0 \\ x_1 &= x_2 \end{aligned}$$

Also $x = \gamma\binom{1}{1}, \gamma \in \mathbb{R}\backslash\{0\}$ Für $\lambda_2 = 0$ gilt

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 + x_2 &= 0 \\ x_1 &= -x_2 \end{aligned}$$

Und hier also $x = \delta\binom{1}{-1}$ mit $\delta \in \mathbb{R}\backslash\{0\}$. Die Basis aus den Eigenvektoren ist also

$$V = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\}$$

Die Matrix bezüglich dieser Basis ist dann

$$f \leftrightarrow \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Das ist die zugehörige Diagonalmatrix.

16. Vorlesung, 10.12.2007

Hauptachsentransformation

Wir rufen uns nochmals die Spiegelung an einer Achse ins Gedächtnis. Da sah die Transformationsmatrix so aus:

$$A = \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

Fakt Das charakteristische Polynom hängt nicht von der Wahl der Basis ab.

$$\begin{aligned} A' &= T^{-1}AT \quad \text{wobei } T \text{ Transformationsmatrix} \\ p_A(\lambda) &= det(A - \lambda E) = det(T^{-1}) * det(A - \lambda E) * det(T) \\ &= det(T^{-1}(A - \lambda E)) * T) * det(T^{-1}AT) - \lambda T^{-1}ET) \\ &= det(A' - \lambda E) = p_A(\lambda) \end{aligned}$$

Schritte zur Bestimmung einer Basis aus möglichst vielen Eigenvektoren eines Eigenwertes $f : V \to V$ mit $dimV = n$.

1. Bestimme die zu f gehörige Matrix bezüglich einer beliebigen Basis B.
2. $p_f(\lambda) = p_A(\lambda) = det(A - \lambda E)$ ist ein Polynom $n$ten Gerades in Variable λ
3. Bestimme die verschiedenen Nullstellen $\lambda, \ldots, \lambda_k$ mit $k \leq n$
4. Berechne die zugehörigen Eigenräume $E_{\lambda i}$ als Lösungsmenge der Gleichung $f(x) = \lambda_i x$. Dazu berechnen wir die homogene Lösung $(A - \lambda_i E)x = \vec{0}$.
5. Die Basen der Eigenräume vereinigen und gegebenenfalls ergänzen zur Basis von ganz V.
   Falls $\sum_{i=0}^{k} dimE_{\lambda,i} = n$ dann ist das Diagonalform.

Satz 24. *Die algebraischer Vielfachheit von λ_i (= Vielfachheit als Nullstelle von $p_A(\lambda)$) ist größergleich der geometrischen Vielfachheit von λ_i (= Dimension von $E_{\lambda i}$).*

Beweis 24. *λ_i mit geometrischer Vielfachheit $\mu = dimE_{\lambda 0}$. $E_{\lambda 0}$ hat natürlich eine Basis $\{x_1, \ldots, x_\mu\}$ kann ergänzt werden zur Basis $\{x_1, \ldots, x_\mu, x_{\mu+1}, \ldots, x_n\}$ mit $dimV$ bezüglich der Basis von V.*

$$A = \begin{pmatrix} \lambda_1 & 0 & a_{p+1} & \ldots & a_? \\ 0 & \lambda_2 & \vdots & & \end{pmatrix}$$

$p_f(\lambda) = p_A(\lambda) = (\lambda - \lambda_i)^? * \bar{p}(\lambda)$ *und das Restpolynom hat keine reellen Nullstellen mehr.*

Frage Gibt es Bedingungen für f, die sichern, dass die Diagonalisierung garantieren?
Ja, die gibt es! Sei V ein endlicher Vektorraum und wir definieren $f : V \to V$.

Definition 32. *$f : V \to V$ heißt selbstadjugiert, falls*

$$\forall x, y \in V :< f(x), y >=< x, f(y) >$$

Wenn $f : V \to V$ und $dimV < \infty$ dann ist bei f selbstadjugiert die Folge, dass f diagonalisierbar ist.
Jetzt ist die Frage, ob das überhaupt vorkommt.
Noch ein kleiner Fakt: f ist selbstadjugiert und lässt sich diagonolisieren, genau dann wenn die Matrix von f bezüglich einer beliebigen Orthonormalbasis symmetrisch zur Hauptdiagonalen ist.
Das muss jetzt natürlich noch beweisen werden.

$\Rightarrow$ Angenommen, f ist selbstadjugierbar. Wie sieht dann die Matrix aus? Dazu wählen wir eine Orthonormalbasis $(v_1, \ldots, v_n)$, wobei die Spalten die Bilder der Basisvektoren sind. $f(v_j) = \sum_{i=1}^{n} a_{ij} v_i$ Wegen der Orthonormalität gilt $< f(v_j), v_i >= a_{ij}, < f(v_i), v_j >= a_{ji}$ (wird nachgeliefert)

Noch ein Fakt. $f : V \to V$ und selbstadjungiert. $\vec{0} \neq v \in V$ ist Eigenvektor zu λ. Dann ist $f_{|\{v\}^\perp}$ selbstadjungierter Endomoprhismus mit $\{v\}^\perp \to \{v\}^\perp$. Sei $v \in \{v\}^\perp$.

$$< f(v), v >=< u, f(v) >=< u, \lambda v >= \lambda < u, v >= 0$$

Und noch einer: f ist wieder selbstadjungiert und $dimV = n$. Dann hat f einen reellen Eigenwert. Das kann man beweisen: $A : \mathbb{R}^n \to \mathbb{R}^n$ und $A = Tf * T^{-1}$. $p_A(\lambda)$ reelles Polynom $n$ten Grades. Der Fundamentalsatz der Algebra besagt, dass dieses Polynom dann n Nullstellen hat, die komplex sind. Sei $\lambda - \gamma - i \in \mathbb{C}$ eine Nullstelle. Dann ist λ Eigenwert von $A : \mathbb{C}^n \to \mathbb{C}^n$. Dann $\exists z \in \mathbb{C}^n$ mit

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_1 + iy_1 \\ \vdots \\ x_n + iy_n \end{pmatrix}$$

mit $Az = \lambda z$. $A(x+iy) = (\gamma + i\omega)(x - iy)$. Wegen Selbstadjungiertheit gilt $< Ax, y >=< x, Ay >$. Also $T < x - \omega y, y >=< x, \gamma y + \omega x >$. $T < x, y > -\omega < y, y >= T < x, y > +\omega < x, x >$ Es bleibt also $0 = \omega(< y, y > + < x, x >)$ übrig. Also $0 = \omega(||y|| + ||x||)$. Daraus folgt, dass $\omega = 0$.
Also ist der Imaginärteil 0 und somit ist es tatsächlich eine reelle Zahl.

Jetzt kommen wir zu unserem großen Beweis. Wir machen ne Induktion.

Induktionsanfang $n = 1$, das ist die 1×1 Matrix, und die ist nach Definition in Diagonalform

Induktionsschritt Angenommen, das Theorem stimmt für alle $n-1$-dimensionalen Vektorräume und selbstadjugierten Endomorphismen. Aus dem ersten Fakt folgt, dass wenn f selbstadjugiert ist, dann mit symmetrischer $n \times n$ Matrix A bezüglich der Orthonormalbasis. Aus der zweiten Überlegung folgt, dass ein reeller Eigenwert und dazu Eigenvektoren $v \neq \vec{0}$ existieren. Die kann man dann natürlich auch normieren, also $v^* = \frac{v}{||v||}$ hat jetzt Norm 1 und ist Eigenvektor. $f_{|\{v^*\}^\perp}$ ist einstelliger Endomorphismus auf $n-1$-dimensionalem Vektorraum. Wir wenden jetzt die Induktionsvorraussetzung an. $f_{|\{v^*\}^\perp}$ hat Orthonormalbasis und ist diagonalisierbar. Jetzt nimmt man diese Basis plus v^*, was dann unsere gesuchte Basis ist.

17. Vorlesung, 12.12.2007

17.1. Codierungstheorie

Wir wollen Informationen kodieren. Eine Information ist hierbei gegeben als Wörter über einem Alphabet Σ.

Ziele

- Sicher gegen Lauschangriffe soll es sein, damit beschäftigt sich die Kryptographie
- Kompakte Darstellung, da gibt es die Möglichkeiten verlustfrei und Verluste sind möglich
- fehlerkorrigierend

Definition 33. *Σ und Q sind Alphabete. Dann ist $f : \Sigma \to Q^+$ Fortsetzung zu $\phi : \Sigma^+ \to Q^+$ ist dann $\Sigma_1\Sigma_2 \ldots \to \phi(\Sigma_1) \circ \phi(\Sigma_2) \circ \ldots$*

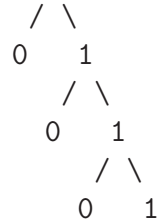
Ein Beispiel wäre $\Sigma = \{a, b\}$ und $Q = \{0, 1\}$ mit $a \to 0$ und $b \to 11$, dann wird aus $aabba \to 0011110$. Und auch wenn es nicht so aussieht, auch $a \to 0$ und $b \to 01$ wäre eindeutig dekodierbar.

17.1.1. Datenkompression

Die Datenkompression verwendet sogenannte Präfixcodes. Nach Herr Hoffmann müsste man allerdings nicht-Präfixcodes sagen.

$f : \Sigma \to q^+$ hat Eigenschaft $\forall \sigma, \sigma' : \phi(\sigma)$ ist kein Präfix von $\phi(\sigma')$

Ein Beispiel für einen Präfixcode ist $C = \{0, 10, 110, 111\}$ Das kann man auch ganz schick in eine Baumdarstellung packen



Wobei jetzt die Wege zu den Blättern die Codierungen sind.
Warum hat das jetzt was mit Kompression zu tun? Das hat was mit variabler Codelänge zu tun. So kann man z.B. häufigen Zeichen mit einem kurzen Codewort verknüpfen.
Die Huffman-Codierung liefert hier bei bekannten Häufigkeiten der Zeichen eine optimale Kompression, das machen wir hier allerdings nicht. Stattdessen machen wir die Blockcodierung.

Blockcodierung

Definition 34. *$C \subset Q^+$ heißt Blockcode, falls alle $c \in C$ die gleiche Länge haben.*

Hier können bei $|Q| = q$ und Blocklänge n genau q^n Zeichen kodiert werden.

Definition 35. *Seien $c, c' \in Q^n$ zwei Codewörter. Dann ist $d(c, c')$ die Anzahl der Stellen, an denen sich c und c' unterscheiden. Das ist der sogenannte Hemming-Abstand zwischen c und c'.*

$$d(c, c') = \{i | q_i \neq q'_i\}$$

Die Kenngröße für einen Blockcode ist $d(C) = \min_{c \neq c'} d(c, c')$ mit $C \leq Q^n$ als Code. $d : Q^n \times Q^n \to \mathbb{N}$ ist eine Metrik auf Q^n, d.h.

1. $\forall q, q' : d(q, q') \geq 0$ und $d(q, q') = 0 \Leftrightarrow q = q'$
2. $\forall q, q' : d(q, q') = d(q', q)$
3. $\forall q, q', q'' : d(q, q') + d(q', q'') \geq d(q, q'')$

Definition 36. *$c \in Q^n$ mit $B_k(c) = \{v \in Q^n | d(c, v) \leq k\}$ ist die Kugel ("Ball") mit Zentrum c und Radius k.*

Wie groß ist denn nun $B_k(c)$? Das ist

$$|B_k(c)| = \sum_{i=0}^{k} \binom{n}{i} (q-1)^i$$

Dekodierung bei Blockcodes $C \subset Q^n$ ist der Code, also die Menge der Codewörter. Wir bekommen ein $v \in Q^n$. Falls $v \in C$ tatsächlich ein Codewort ist, dann wähle σ mit $\phi(\sigma) = v$. Falls $v \notin C$, also ein Fehler bei der Kodierung aufgetreten ist, dann wähle $c \in C$ mit $d(c, v)$ minimal, das muss allerdings nicht eindeutig sein. Diese Funktion nennt man auch "Maximum Likelihood" (ML).

Definition 37. *Ein Code $C \subset Q^n$ heißt k-fehlerkorrigierend, falls ein empfangenes Wort mit $\leq k$ Fehlern bezüglich ML korrekt dekodiert wird.*

Definition 38. *C ist k-fehlerkorrigierend genau dann wenn, $\forall c, c' \in C : B_k(c) \cap B_k(c') = \emptyset$. Und das gilt genau dann, wenn $d(C) \geq 2k + 1$.*

C ist k-fehlerkorrigierend und $|C| = \sum_{i=0}^{k} \binom{n}{i}(q-1)^i \leq q^n$. Wenn tatsächlich die Gleichheit besteht, dann nennt man das einen perfekten Code.

Definition 39. *Der Blockcode C heißt k-fehlererkennend, falls $\forall c \neq c' : c \notin B_k(c')$. Das ist genau dann der Fall, wenn $d(C) \geq k+1$.*

18. Vorlesung, 17.12.2007

Das Verhältnis $\frac{Blocklaenge}{\lceil \log_{|Q|} |\Sigma| \rceil}$ heißt Informationsgehalt. Man verwendet dabei Redundanz um Fehlerkorrektur bzw. -erkennung zu ermöglichen.
Eine weitere Ungleichung, um $d(C)$ in Beziehung mit q zu setzen ist folgendes. Angenommen, C hat den Mindestabstand $d(C)$. Dann ist die Abbildung $Q^m \to Q^{m-d(C)+1}$ mit $(q_1, \ldots, q_m) \to (q_1, \ldots, q_{m-d(C)+1})$ injektiv. D.h., das Bild hat Größe $|C|$. Und somit $|C| \leq q^{m-d(c)+1}$. Durch Logarithmieren erhält man dann $\log_q |C| \leq m - d(C) + 1$. Für die Anzahl der Codewörter gilt dann $d(C) \leq m - \log_q |C| + 1$

18.0.2. Einfache Verfahren zur Fehlererkennnug/ -korrektur

$q^n : \Sigma \to Q^n$ ist der gegebene Blockcode.

1. Wir setzen ein Paritätsbit. $\phi_{par} : Q^m \to Q^{m+1}$ mit $Q = \{0, 1\}$ $(q_1, q_2, \ldots, q_m) \to (q_1, \ldots, q_m, q_1 + \ldots + q_m \mod 2)$ hat eine gerade Anzahl von Einsen. Und der Mindestabstand ist $d(\phi_{par}) = 2$. Es ist also 1-fehlererkennend.

2. Verdoppeln der Nachricht. $\phi_2 : Q^m \to Q^{2m}$ mit $(q_1, \ldots, q_m) \to (q_1, \ldots, q_m, q_1, \ldots, q_m)$. Auch hier ist der Mindestabstand $d(\phi_2) = 2$. Somit ist es auch 1-fehlererkennend.

3. Bei $\phi_3 : Q^m \to Q^{3m}$ ist der Mindestabstand 3. Und das ist dann 1-fehlerkorrigierend.

4. Jetzt verdoppeln wir und setzen das Paritätsbit. Dann haben wir $\phi_{2,par} : Q^m \to Q^{2m+1}$ mit $(q_1, \ldots, q_m) \to (q_1, \ldots, q_m, q_1, \ldots, q_m, q_{m+1})$ mit $q_{m+1} = q_1 + \ldots q_m \mod 2$. Hier ist die Behauptung, dass $d(\phi_{2,par}) = 3$ ist. Das ist auch wieder 1-fehlerkorrigierend.
   Das ist jetzt schonmal natürlich besseres als 3., da wir nicht mehr so viel Redundanz haben. Geht's noch besser? Ja!

5. Und zwar mit dem Kreuzsicherungscode. Es wird vorausgesetzt, dass $m = k^2$. Dann ist $\{0, 1\}^m$ $v = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k^2}$. Man schreibt sich das jetzt so untereinander, dass v_1 und v_k zusammenstehen, dann v_{k+1}, v_{2k} usw. bis $v_?$ nach v_{k^2}. Zudem macht man dann für die Zeile noch ein Paritätsbit und für die Spalten. Dann haben wir $\phi : Q^m \to Q^{m+2k}$. Die Behauptung ist, dass für den Mindestabstand $d(\phi) = 3$ gilt. Das Beweisen wir jetzt noch kurz. $d(u, v) = 3$ mit $u, v \in Q^m$. Dann gilt für $d(\phi(u), \phi(v)) \geq 3$. Was passiert, wenn sie sich vorher an zwei Stellen unterschieden haben, also $d(u, v) = 2$? Dann gilt $d(\phi(u), \phi(v)) \geq 4$. Für $d(u, v) = 1$ gilt dann $d(\phi(u), \phi(v)) = 3$.

6. Jetzt der Hemming-Code. Dafür brauchen wir vier Informationsbits (v_1, v_2, v_3, v_4) mit drei Redundanzbits (v_5, v_6, v_7), wobei $v_5 = v_2 + v_3 + v_4 \mod 2, v_6 = v_1 + v_3 + v_4 \mod 2$ und $v_7 = v_1 + v_2 + v_4 \mod 2$. Das kann als Abbildung $\phi_{hem} = \mathbb{Z}_2^4 \to \mathbb{Z}_2^7$ aufgefasst werden. Jetzt kann man das schick als Matrix schreiben:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix} * \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{pmatrix}$$

Das bezeichnet man auch als Generatormatrix.
Die Dekodierung mittels der Prüfmatrix

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

ist $H * Codewort$ ergibt immer $\vec{0}$. Wenn ein Bit gekippt wird, dann liefert die Prüfmatrix nicht mehr den Nullvektor. Aber sie sagt uns auch, welches Bit gekippt ist, nämlich das ist die Dezimalzahl, die durch das Ergebnis repräsentiert wird.
Was ist $|B_1(Codewort)|$?

$$\sum_{i=0}^{1} \binom{7}{i} (2-1)^i = 1 + 7$$

Das Volumen einer Kugel ist also 8. Wir haben $2^4 = 16$ Codewörter. $2^3 * 2^4 = 2^7$ Der Bildraum ist ja auch genau 2^7, wir schöpfen also den gesamten Raum aus, es ist also ein perfekter Code.

19. Vorlesung, 19.12.2007

19.0.3. Endliche Körper

Ein Körper ist bekanntlich ein Tupel $(K, +, *, 0, 1)$, wobei $(K, +, 0)$ und $(K\backslash\{0\}, *, 1)$ kommutative Gruppen sind. Zudem sind die Operatoren distributiv.
Bei einem endlichen Körper heißt das, dass K endlich ist.
Unser Ziel ist, zu zeigen, dass es für p Prim und $k > 0$ gibt es Körper mit p^k Elementen. $GF(p^k)$ (oder auch $\mathbb{F}_{p^k}$).

$$\forall a, d \in \mathbb{Z}, d \neq 0 : \exists! q, r \in \mathbb{Z} \;\; a = qd + r \mod r \in \{0, \ldots, d-1\} \;\; q = \lfloor \frac{a}{d} \rfloor, r = a \mod d$$

Das braucht man, um den euklidischen Algorithmus zu Berechnen, mit dem der ggT bestimmt wird. Sei $a, b \in \mathbb{Z}^+, a \geq b, r = a \mod b$. Der ggT berechnet sich über

$$r = 0 \Rightarrow ggT(a, b) = b$$
$$sonst\ ggT(a, b) = ggT(b, r)$$

Die Umkehrung liefert $\forall a, b \in \mathbb{Z}^+, \exists s, t \in \mathbb{Z}$ mit $ggT(a, b) = sa - tb$. Das kann man per Induktion beweisen, machen wir hier aber nicht. Steht aber im Skript.

Fakt Falls der $ggT(a, b) = 1$, dann ist die Darstellung $1 = sa - tb$ eindeutig.

Rechnen $\mod n$

Das induziert eine Äquivalentsrelation und das Repräsentantensystem ist $\{0, \ldots, n-1\}$. $a \simeq a' \mod n, b \simeq b' \mod n$ so gilt

$$(a \pm b) \simeq (a' \pm b') \mod n \quad (a * b) \simeq (a' * b') \mod n$$

Fakt p ist Primzahl und $a \in \{1, \ldots, p-1\}$. $\exists! b = \{1, \ldots p-1\}$ mit $ab \simeq 1 \mod p$.
Das beweisen wir jetzt noch hübsch. Sei der $ggT(p, a) = 1$. Dann $\exists s, t : 1 = as - tp$. Wir setzen $b \simeq s \mod p$, so dass $b \in \{0, \ldots, p-1\}$. Zeige $ab \simeq 1 \mod p$, damit $b \neq 0$. Wir wissen $b \simeq a \mod p, a \simeq a \mod p$ und $0 \simeq tp \mod p$. Außerdem ist bekannt $ab \simeq as + dp$ daraus ergibt sich $ab + 0 \simeq (ab - tp \mod p = 1 \mod p$. Und das wollten wir zeigen.
Jetzt fehlt noch die Eindeutigkeit. Angenommen $b' \in \{1, \ldots, p-1\}$: $ab \simeq 1 \mod p$ und $ab' \simeq 1 \mod p$. Daraus folgt $a(b-b') \simeq 0 \mod p$. Und hieraus folgt $p|a(b-b') \Rightarrow p|b-b' \Rightarrow b = b'$.

Beispiel Wir wollen im Körper $\mathbb{F}_{97}$ das multiplikative Inverse zu 25 berechnen.

$$\begin{aligned} 1 &= -31 * 25 - 8 * 17| \mod 97 \\ 26^{-1} &= -31 \mod 97 \\ &= 66 \end{aligned}$$

Theorem Es gibt genau dann einen Körper mit n Elementen, wenn $n = p^k$ mit p Prim und $k \geq 1 \in \mathbb{N}$. Dieser ist bis auf die Isomorphie eindeutig. (p heißt Charakteristik des Körpers mit $1 + 1 + \ldots + 1 = 0$ (wobei p mal 1).

Wir machen folgende Konstruktion. Sei p wieder eine Primzahl. $\mathbb{F}_p[x]$ ist die Menge aller Polynome mit Koeffizient aus $\mathbb{F}_p$.

$$\sum_{i=0}^{d} a_i x^i$$

d gibt dabei den Grad von dem Polynom an. $\mathbb{F}_p[x]$ ist eine kommutative Gruppe bezüglich der Addition und eine kommutative Halbgruppe 1 bezüglich der Multiplikation.
Sei $g(x) \in \mathbb{F}_p[x]$ mit Grad k. Wir betrachten jetzt alle Polynome in $\mathbb{F}_p[x]$ mit Grad $< k$. Auf dieser Grundmenge funktioniert die Addition wie gehabt. Die Multiplikation kann aber nicht mehr so funktionieren, da dort die Grade der zu Multiplizierenden Polynome addiert werden.

Beispiele

1. $g(x) = x^3 - x^2 + 1 \in \mathbb{F}_2[x]$. $(x^2 + 1) * (x^2 + x + 1) = x^4 + x^3 + x^2 + x^2 + x + 1$. In $\mathbb{F}_2[x]$ ist das gleich $x^4 + x^3 + x + 1 = x(x^3 + x^2 + 1) + 1$. $(x^2 + 1)(x^2 + x + 1) \mod g(x) = 1$.

2. $g(x) = x^3 + x^2 + 2$ aber jetzt in $\mathbb{F}_3[x]$. Wir berechnet wieder $(x^2 + 1)(x^2 + x + 1) = x^4 + x^3 + 2x^2 + x + 1. = x(x^3 + x^2 + 2) + 2x^2 + 2x + 1(x^2 + 1)(x^2 + x + 1) = 2x^2 + 2x + 1 \mod g(x)$

Definition 40. *$g(x) \in \mathbb{F}_p[x]$ heißt irreduzibel, genau dann wenn es keine $a(x), b(x) \in \mathbb{F}_p[x] \backslash \mathbb{F}_p$ mit $g(x) = a(x) * b(x)$ gibt.*

$\mathbb{F}_p[x]/g(x)$ ist das Polynom mit Grad $< deg(g(x)$ und Addition und der Multiplikation modulo $g(x)$ bilden Körper.

Additionstafel für 4-Elemente

+	00	01	10	11
00	00	01	10	11
01	01	00	11	10
10	10	11	00	01
11	11	10	01	00

Multiplikationstafel für 4-Elemente

+	00	01	10	11
00	00	00	00	00
01	00	01	10	11
10	00	10	11	01
11	00	11	01	10

20. Vorlesung, 7.1.2008

20.0.4. Lineare Codes

Definition 41. *Ein linearer Code über $\mathbb{F}_q$ ist gegeben durch eine lineare Abbildung (und injektiv sowieso)*

$$g : \mathbb{F}_q^m \to \mathbb{F}_q^n$$

Wegen der Injektivität ist der Kern 0 und somit $rg(Im(g)) = m$ und das Bild wird hierbei als Code C bezeichnet und ist ein Unterraum von $\mathbb{F}_q^n$. Also $\vec{0} \in C$.
g wird berechnet durch Matrix $G \in M(n \times m, \mathbb{F}_q)$ "Generatormatrix", $v \in \mathbb{F}_q^m$ wird notiert als $G * v \in \mathbb{F}_q^n$.

Alternativer Weg um Unterraum C zu beschreiben

$$g : \mathbb{F}_q^m \to \mathbb{F}_q^n$$

wobei $\mathbb{F}_q^n \subset g(\mathbb{F}_q^n)$?. Kern beschreibt überdies als Kern einer Abbildung h

$$h : \mathbb{F}_q^n \to \mathbb{F}_q^{n-m}$$

Matrix $H \in M(n-m \times n, \mathbb{F}_q)$ mit $rg(H) = n - m$. H nennt man auch die Prüfmatrix für den Code ($H * c = \vec{0}$ für $c \in C$).
Weder Generator- noch Prüfmatrix sind eindeutig, man kann durch die eine aber die andere finden.

Beispiele

1. Sei $\mathbb{F}_2$ der Körper. Und $h : \mathbb{F}_2^4 \to \mathbb{F}_2^8$.

   $$(a_1, a_2, a_3, a_4) \to (a_1, a_2, a_3, a_4, a_1 + a_2, a_3 + a_4, a_1 + a_3, a_2 + a_4)$$

   Wir schreiben das als lineares Gleichungssystem

   $$\begin{aligned} a_1 + a_2 + a_5 &= 0 \\ a_3 + a_4 + a_6 &= 0 \\ a_1 + a_3 + a_7 &= 0 \\ a_2 + a_4 + a_8 &= 0 \end{aligned}$$

Der Code ist offensichtlich die Lösung dieses homogenen Gleichungssystems. Wie sieht nun also die zugehörige Matrix aus?

$$H_0 = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Der Code $C_0 = \{x \in \mathbb{F}_2^8 | Hx' = \vec{0}\}$. Es folge $rg(H_0) = 4$ und somit hat C_0 den Rang $8 - 4 = 4$.
Wie werden nun Codewörter erzeugt, also wie sieht die Generatormatrix aus?

$$G_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Die Standardform der Prüf- und Generatormatrix sieht wie folgt aus

$$G = \begin{pmatrix} E_m \\ A \end{pmatrix} \; mit \; A \; (n-m) \times n \; Matrix$$

$$H = \begin{pmatrix} -A & E_{n-m} \end{pmatrix}$$

Beachte, dass $H * G = (0)$.

2. Die Prüfmatrix hat die Form

$$H_1 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \in M(3 \times 5, \mathbb{F}_2)$$

Dann ist

$$G_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \in M(5 \times 2, \mathbb{F}_2)$$

Dann sind die Codewörter

$$G_1 * \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

$$G_1 * \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

außerdem noch der Nullvektor und die Linearkombination

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = 1 * \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + 1 * \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

$d(C_1) = 3$ und somit ist der Code 1-fehlerkorrigierend.

Satz 25. *$G \in M(n \times m, \mathbb{F}_q)$ mit $rg(G) = m$ und $H \in M((n-m) \times m, \mathbb{F}_q)$ mit $rg(H) = n - m$ sind Generator- und Prüfmatrix des ein und desselben linearen Codes genau dann, wenn $H * G = (0)$.*

Beweis 25. *Wir müssen Hin- und Rückrichtung zeigen.*

=> *geht aus Definition hervor*

<= *$H * G = (0)$ gilt genau dann wenn, $Im(G) \leq Ker(H)$. Also $dim(Im(g)) = rg(G) = m \leq dim(Ker(H))$. Also (Dimensionsformel)*

$$\begin{aligned} n &= dim(Ker(H)) + rg(H) \\ &= dim(Ker(H)) + n - m \\ \Rightarrow dim(Ker(H)) &= m \end{aligned}$$

Also folgt tatsächlich $Im(G) = Ker(H)$.

□

Definition 42. *C sei ein lineare Code. Und $v \in C$. Das Gewicht von v ist $w(v)$ = die Anzahl der Komponenten ungleich null. $w(C) = \min\{w(v) | v \in C, v \neq \vec{0}\}$.*

Fakt $w(C) = d(C)$.
Das wird noch kurz bewiesen. Und zwar folgt das aus $d(u, v) = d(u - v, \vec{0})$ (Translationsinvarianz des Hemmingabstands)

Satz 26. *Ist $H \in M((n-m) \times n, \mathbb{F}_q)$ Prüfmatrix von einem linearen Code C, so ist $d(C) =$ minimale Anzahl linear abhängiger Spalten in H.*

Beweis 26. *$H_1, \ldots, H_n$ seien die Spalten von der Prüfmatrix H. Sei $\{H\}_{i \in S} S \subseteq \{1, \ldots, n\}$?? Teilmenge ???*
$\exists \lambda_i \in \mathbb{F}_q : \sum_{i \in S} \lambda_i H = \vec{0}$. Wir definieren uns den Vektor $c = (\lambda_1, \ldots, \lambda_n)$ mit $\lambda_j = 0$ für $j \neq S$.
Dann ist also $H * c = \vec{0}$. $w(c) = |S|$ also $d(C) \leq |S|$.
Sei $c \in C$ mit $c = (\lambda_1, \ldots, \lambda_n)$ mit $w(c) =$ minimal. Sei $S = \{i | \lambda_i \neq 0\}$. Dann ist $\sum_{i \in S} \lambda_i H_i = \vec{0}$ Weil $H * c = \vec{0}$ folgt $\{H\}_{i \in S}$ ist linear abhängig und daraus folgt $d(C) \geq$ minimale Anzahl linear abhäniger Spalten. □

Korollar Ist in einer Prüfmatrix keine Spalte Vielfaches einer anderen Spalte, so ist $d(C) \geq 3$, also 1-fehlerkorrigierend.
Insbesondere gilt für $\mathbb{F}_2$ zwei Spalten sind voneinander linear abhängig, genau dann wenn sie gleich sind.

21. Vorlesung, 9.1.2008

Beispiel

1. Gesucht ist ein 1-fehlerkorregirender, perfekter binärer Code mit 4 Redundanzbits. Wir wissen aus dieser Fragestellung, dass $n - m = 4$. Es müssen nun also n und m bestimmt werden. Wir beginnen mit der Prüfmatrix (in Standardform).

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

   Die Informationsrate ist $\frac{m}{n} = \frac{n-4}{n}$ und soll möglichst nahe an 1 sein, daher muss n so groß wie möglich gewählt werden. ABER es dürfen keine 2 Spalten linear abhängig sein. Wir sind in $\mathbb{F}_2$, daher gilt: 2 Spalten sind linear abhängig, falls sie gleich sind.
   Die Matrix A besteht aus den restlichen $(16 - 1) - 4 = 11$ verschiedenen Spaltenvektoren der Länge 4. (-1 weil der Nullvektor nicht mitkommt und -4, weil wir schon 4 haben). Dabei ist die Reihenfolge egal.
   H ist somit eine $M(4 \times 15, \mathbb{F}_2)$ Matrix. Und somit ist G, die Generatormatrix, eine 15×11 Matrix. Es ist also $|C| = 2^{11}$. $|B_1(c)| = 1 + 15 = 2^4$.

2. Wollen nun einen 1-perfekten Code er $\mathbb{F}_3$ mit 2 Redudanzstellen. Es ist wieder $n - m = 2$. Wir fangen wieder mit der Prüfmatrix an.

$$H = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

   n ist also gleich 4, die Anzahl der Spalten in der Prüfmatrix, somit ist $m = 2$. Wie sieht nun die Generatormatrix aus?

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -A & \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \\ 2 & 2 \end{pmatrix}$$

   $|C| = 3^2 = 9$. Was ist jetzt $|B_1(c)| = 1 + 4 * 2 = 9$. Und $3^4 = 3^2 * 3^2$, daher ist es tatsächlich ein perfekter Code.

21.0.5. Fehlerkorrektur und Prüfmatrix

H sei eine Prüfmatrix $M((n-m) \times n, \mathbb{F}_q)$ mit dem Rang $n-m$. C ist der Code und $d(C) \geq 2k+1$ mit 1-fehlerkorrigierend. $c \in C$ zieht alle $v \in \mathbb{F}_q^k$, falls $d(c,v) \leq k$. Was macht nun H mit diesem Vektor?

$$H * v = Hv - \vec{0} = Hv - H\vec{0} = Hv - Hc = H(v-c)$$

Fakt H angewendet auf $B_k(\vec{0})$ ist injektiv.
Jetzt der Beweis: Angenommen $u, v \in B_k(c)$ und $Hu = Hv$. Daraus folgt $H(u-v) = \vec{0}$, das ist der Kern, also muss $u-v$ ein Codewort sein, da nur diese auf 0 abgebildet werden. Also $w(u-v) = d(u-v,0) \leq 2k$. Das ist ein Widerspruch zu $d(C) \geq 2k+1$.

Folgerung Hv bestimmt $v-c$ eindeutig und somit auch c, weil $c = v-(v-c)$. Bleibt als Aufgabe noch, die Bilder von $B_k(\vec{0})$ unter H zu bestimmen.

Definition 43. *Hv heißt Syndrom von v.*

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

also $n = 7, m = 4, n-m = 3$. Damit ist $d(C) = 3$.

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

$B_1(\vec{0}) = \{(1,\ldots,0),(0,1,\ldots,0),\ldots\}$ insgesamt gibt es also 8 Vektoren.

$$H * \vec{0} = \vec{0}$$

$$Hv_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 1.Spalte\ von\ H$$

$$Hv_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 2.Spalte\ von\ H$$

$$Hv_i = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = i.Spalte\ von\ H$$

Wenn man jetzt das Wort $w = (1, 0, 1, 1)$ Codieren will, rechnet man halt

$$G * w = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Zum Dekodieren muss jetzt natürlich wieder $H * v$ gerechnet werden, wenn v zu dekodieren ist. Das Ergebnis ist dann das Syndrom, um c zu enthalten muss also wieder $v - (v - c)$ berechnet werden.

$\mathbb{F}_2^4 \to \mathbb{F}_2^7$ für die Generatormatrix und $\mathbb{F}_2^7 \to \mathbb{F}_2^3$ als Prüfmatrix ist eine spezielle Hammingcode mit Prüfmatrix. Gibt es das auch noch für andere Parameter? Ja! Und zwar wenn man $m = 2^r - 1$ wählt. Prüfmatrix mit den Parametern $(r \times (2^r - 1))$ in $\mathbb{F}_2$. Die $2^r - 1$ Spalten repräsentieren $1, 2, \ldots, 2^r$ als Binärzahl. Der Code hat $2^{2^r - 1 - r}$ Codewörter. Das Volumen ist $B_k(c) = 1 + 2^r - 1 = 2^r$. $2^{2^r - 1} * 2^r = 2^{2^r - 1}$, der Raum wird also voll ausgenutzt, es handelt sich demnach um einen perfekten Code.

22. Vorlesung, 14.1.2008

22.1. Stochastik

Die Stochastik gliedert sich auf in Wahrscheinlichkeitstheorie und Statistik, hier nur Wahrscheinlichkeitstheorie.

Zentrale Begriffe Das sind her Zufall und Wahrscheinlichkeit.

22.1.1. Axiomatik der Wahrscheinlichkeitstheorie

Das zentrale Hilfsmittel ist die Maßtheorie.

Wahrscheinlichkeitsräume Festlegung der Eigenesraumes. Bei einem Zufallsexperiment ist zu entscheiden, welche Größe zu messen ist. Die Menge ist hier $\Omega \neq \emptyset$ von Elementarergebnissen. Man nennt Ω auch Grundraum, Ereignisraum. Die Teilmenge $A \subseteq \Omega$ heißt Ereignis. Wir wollen im nächsten Schritt die Abgeschlossenheit gegenüber Booleschen Operationen $\neg, \vee$ definieren.

Definition 44. *Ω ist die Menge, das Mengensystem $\mathcal{F} \in \mathcal{P}(\Omega)$ heißt Mengen-Algebra über Ω, falls*

1. $\Omega \in \mathcal{F}$
2. $F \in \mathcal{F} \Rightarrow \overline{F} \in \mathcal{F}$
3. $E, F \in \mathcal{F} \Rightarrow E \cup F \in \mathcal{F}$

Eine Beobachtung ist dabei, dass auch $E, F \in \mathcal{F} \Rightarrow E \cap F \in \mathcal{F}$, wegen de Morgan.
Im Extremfall muss $\mathcal{F} = \{\emptyset, \Omega\}$ bzw. $\mathcal{F} = \mathcal{P}(\Omega)$. Es ist hierbei sinnvoll, sich einzuschränken und nicht immer $\mathcal{P}(\Omega)$ zu wählen.

Definition 45. *Menenalgebra $\mathcal{F}$ über Ω mit Punkt drei erweitert auf $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup A_i \in \mathcal{F}$ heißt σ-Algebra über Ω.*

Wir wollen mit der kleinsten σ-Algebra, die alle gewünschten Teilmengen (Intervalle bei reellen Zahlen) enthält. Sei $(\mathcal{F}_i)$? von σ-Algebra über Ω. Dann ist $\bigcap \mathcal{F}_i$ eine σ-Algebra über Ω.
Der Beweis geht wie folgt: Sei $A_1, A_2, \ldots \in \bigcap \mathcal{F}_i \Rightarrow \forall i \in I A_1, A_2, \ldots \in \mathcal{F}_i$. Und hieraus folgt aufgrund der σ-Algebra dass $\bigcup A_j = \mathcal{F}_i \Rightarrow A_j \in \bigcap_{i \in I} \mathcal{F}_i$. □

Definition 46. $G \subseteq \mathcal{P}(\Omega)$. *Dann heißt*

$$\mathcal{F} + \sigma(G) = \{\mathcal{F} \subseteq \mathcal{P}(\Omega) | \mathcal{F} \text{ } \sigma\text{-Algebra und } G \subseteq \mathcal{F}\}$$

die von G erzeugte σ-Algebra.

"wichtigstes Beispiel" Sei G die Menge der offenen Teilmengen des $\mathbb{R}^n$. Die von G erzeugte σ-Algebra

$$\mathcal{B}(\mathbb{R}^n) = \mathcal{B}^n$$

Borel-Algebra für $\mathbb{R}^n$.

Satz 27. $\mathcal{B}^1 = \mathcal{B}(\mathbb{R}^1) = \sigma\{(-\infty, x] | x \in \mathbb{R}\}$

Beweis 27. *Zeige $\mathcal{B}^n \in \sigma\{...\} = \mathcal{F}$. Sei G ene offene Menge. $G = \cup_{x \in G} U_{\epsilon(x)}(x)$, wobei $U_{\epsilon(x)}(x)$ eine kleine offenen Umgebung von x ist. $G = \bigcup\{(a,b) \ a,b \in G \wedge a,b \in \mathbb{Q}\}$. Daraus folge, $\mathcal{B}^n = \sigma\{(a,b) | a,b \in \mathbb{Q}\}$. Wir zeigen, die offenen Intervalle $(a,b) \in \mathcal{F}$. $(a,b] = (-\infty, b] \backslash (-\infty, a]$. $(a,b) = \bigcup_{k \geq 1} (a, b - \frac{1}{k}]$*
Jetzt noch die andere Richtung $(-\infty, x] = \bigcap_{n \geq 1} (-\infty, x + \frac{1}{n})$. □

Es gilt immer $\forall n \geq 1 : \mathcal{B}^n \subset \mathcal{P}(\mathbb{R}^n)$.

23. Vorlesung, 16.1.2008

Wir ordnen nun jedem $A \in \mathcal{F}$ eine Maßzahl $Pr(A) \in \mathbb{R}$ zu. Dabei nennt man $Pr : \mathcal{F} \to \mathbb{R}$ heißt Maß auf Ω, falls

1. $Pr(A) \geq 0 \forall A \in \mathcal{F}$
2. Für paarweise disjunkte $A1, A_2, \ldots \in \mathcal{F} : Pr(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} Pr(A_i)$

Wenn zusätzlich $Pr(\Omega) = 1$ ist, so heißt Pr Wahrscheinlichkeitsmaß.

Definition 47. $(\Omega, \mathcal{F}, Pr)$ *heißt Wahrscheinlichkeitsraum.*

Einfache Eigenschaften

1. Monotonie: $A, B \in \mathcal{F}, A \subseteq B$ dann folgt $Pr(A) \leq Pr(B)$. Und zwar, gilt das weil $B = A \cup (B \backslash A)$, also $Pr(B) = Pr(A) + Pr(A \backslash B) \Rightarrow Pr(B) \geq Pr(A)$.
2. $Pr(\emptyset) = 0$. Auch das kann gefolgert werden: $Pr(\emptyset) = Pr(\emptyset \cup \emptyset \cup \ldots) = \sum_{i=1}^{\infty} Pr(\emptyset)$ Wenn hier war ungleich 0 stehen würde, dann würden wir ∞ rausbekommen, und das wäre ein Widerspruch, also muss $Pr(\emptyset) = 0$ sein.
3. Identität impliziert für paarweise disjunkte $A_1, \ldots, A_n$ $Pr(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} Pr(A_i)$.

Achtung Nicht jedes Ereignis A mit $Pr(A) = 0$ ist die leere Menge und nicht jedes Ereignis A mit $Pr(A) = 1$ ist Ω!

Beispiele

1. Ω endlich bzw. abzählbar unendlich. $\mathcal{F} = \mathcal{P}(\Omega)$. Jedes $\omega \in \Omega$ habe $Pr(\omega)$ als Wahrscheinlichkeit mit $\sum Pr(\omega) = 1$. Für ein $A \in \mathcal{F}$ gilt $Pr(A) = \sum Pr(\omega)$.
2. $\Omega = \mathbb{R}^n$ und sei $\mathcal{F} = B^n$ die Borelalgebra. Sei zusätzlich $\omega_0 \in \mathbb{R}^n$ ein ausgewähltes Element. $A \in \mathcal{F}$ ist definiert als $Pr(A) = 1$, wenn $\omega_0 \in A$, und 0 sonst.
   Man muss beachten, dass die Menge, die nur aus ω_0 besteht, zu $\mathcal{F}$ gehört. Das heißt, dass das Ergebnis des Experiments ist mit Sicherheit $= \omega_0$.

Satz 28. $(\Omega, \mathcal{F}, Pr)$ *der Warhscheinlichkeitsraum und* $A_1, A_2, \ldots \in \mathcal{F}$.

$$Pr(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} Pr(A_i)$$

Beweis 28. *Schreibe $\bigcup_{i=1}^{\infty} A_i$ als unabhängige, disjunkte Ereigisse B_i und zwar induktiv:*

$$B_1 = A_1$$
$$B_{n+1} = A_{n+1} \backslash \left(\bigcup_{i=1}^{n} B_i \right)$$

Wir müssen noch zeigen, dass $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ gilt. Erstmal wissen wir dass $\bigcup_{i=1}^{\infty} A_i \geq \bigcup_{i=1}^{\infty} B_i$, gilt, ist ja klar. Sei jetzt $\omega \in \bigcup_{i=1}^{\infty} A_i$. Das heißt, es existiert ein minimales i mit $\omega \in A_i$. Daraus folgt, dass $\omega \in B_i$. $Pr(\bigcup_{i=1}^{\infty} A_i) = Pr(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} Pr(B_i) \leq \sum_{i=1}^{\infty} Pr(A_i)$. □

Satz 29 (Stetigkeit). *$(\Omega, \mathcal{F}, Pr)$ ist wieder Wahrscheinlichkeitsraum.*

1. $A_1 \subseteq A_2 \subseteq \ldots$

$$Pr(\bigcup_{i=1}^{\infty} A_i) = \lim_{i \to \infty} Pr(A_i)$$

2. $A_1 \supseteq A_2 \supseteq \ldots$

$$Pr(\bigcap_{i=1}^{\infty} A_i) = \lim_{i \to \infty} Pr(A_i)$$

Beweis 29.

1. *$A = \bigcup_{i=1}^{\infty} A_i$ Wir machen die Ereignisse disjunkt, in dem man sich $B_i = A_i \backslash A_{i-1}$ definiert. $A = A_1 \cup B_2 \cup B_3 \cup \ldots$ ist dann eine disjunkte Vereinigung. Dafür wissen wir $Pr(A) = Pr(A_1) + \sum_{i=1}^{\infty} Pr(B_i) = Pr(A) + \lim_{i \to \infty} \sum_{i=1}^{\infty} Pr(B_i) = \lim_{i \to \infty} Pr(A) - Pr(A_2 \backslash A_1) + \ldots Pr(A_n \backslash A_{n-1}) = \lim_{n \to \infty} Pr(A_n)$.*

2. *Setze $B_i = \Omega \backslash A_i$. Wende nun 1. auf die B_i an, das geht, weil diese nun voneinander Obermengen sind. Dann ist $Pr(\bigcup_{i=1}^{\infty} B_i) = \lim_{i \to \infty} Pr(B_i) = Pr(\Omega \backslash A) = \lim_{i \to \infty} (1 - Pr(A_i))$*

□

24. Vorlesung, 21.1.2008

Beispiele Jetzt mal das Standardbeispiel, mit welcher Wahrscheinlichkeit keine 2 von 25 Leute am selben Tag Geburtstag haben.
Das ist also ein Experiment mit 365 Fächern. $\Omega = \{1, \dots 365\}^n$. Als $\mathcal{F} = \mathcal{P}(\Omega)$. Wir nehmen auch eine Gleichverteilung an. Das Ereignis ist $A = \{\omega \in \Omega | \omega = (\omega_1, \dots, \omega_n) \omega_i \neq \omega_j$ für $i \neq j\}$.

$$\begin{aligned} Pr(A) = \frac{|A|}{\Omega} &= \frac{N * (N-1) * (n-2) * \dots * (N-n+1)}{N^n} \\ &= \frac{N^n}{N^n} * 1 * (1 - \frac{1}{N} \dots (1 - \frac{n-1}{N}) \\ &= e^{\ln Pr(A)} = e^{\ln(\prod_{k=1}^{n-1}(1-\frac{k}{N})} \\ &= e^{\sum_{k=1}^{n-1} \ln(1-\frac{k}{N})} \\ &\leq e^{\frac{1}{N} * \frac{n(n-1)}{2}} \\ &= 0,44 \end{aligned}$$

Jetzt noch ein Beispiel im Euklidischen Raum. $\mathcal{B}^n$ Borelalgebra in $\mathbb{R}^n$. Jetzt ist die Frage, ob man etwas wie eine Gleichverteilung betrachten kann. Die Antwort ist nein. Und zwar weil falls $Pr((a,b)) \geq 0$, dann folgt gleich dass $Pr(\mathbb{R}^n) = \infty$.
Es ist aber durchaus möglich bei Grundmenge mit endlichem Volumen (Lebesgue-Maß: λ^n). Wir wählen $Q \in \mathcal{B}^n$ mit endlichem Volumen. $\mathcal{B}^n$ induziert σ-Algebra auf Ω:

$$\mathcal{F} = \mathcal{B}^n \cap \Omega = \{F \cap \Omega | F \in \mathcal{B}^n\}$$

Definition 48 (Gleichverteilung). $A \in \mathcal{F}$.

$$Pr(A) = \frac{\lambda^{(n)}(A)}{\lambda^{(n)}(\Omega)} = \frac{Volumen\ A}{Volumen\ \Omega}$$

Beispiele

- $\Omega = [0,1] \subset \mathbb{R}$ dann ist $Pr((a,b]) = \frac{b-a}{1} = b - a$
- Das bertrandsches Paradoxon.
  Man hat hier also einen Kreis mit Radius von 1 (o.B.d.A.). Im Kreis schreibt man dann das gleichseitige Dreieck rein, welches durch den Kreis ja eindeutig bestimmt ist. Man

legt jetzt zufällig (gleichverteilt) eine Sehne. Jetzt ist die Aufgabe zu bestimmen, mit welcher Wahrscheinlichkeit die Sehne länger ist als die Seite des Dreiecks.
Die Sehne ist eindeutig bestimmt durch die Normalenrichtung und Abstand zum Mittelpunkt des Kreises. Man wählt also eine beliebige Tangente am Kreis und dadurch ist dann die Normalenrichtung bestimmt (weil rechter Winkel zum Mittelpunkt) und diese Tangente kann man jetzt Richtung Mittelpunkt verschieben und irgendwann stehen bleiben.
Der Winkel der Tangente ϕ liegt zwischen $\phi = [0, 2\pi)$ und $r = [0, 1]$. Dann ist der Wahrscheinlichkeitsraum $\Omega = [0, 1] \times [0, 2\pi) \subset \mathbb{R}^2$. Dann ist $\lambda^{(2)}(\Omega)$ =Fläche $= 2\pi$. $A =$ günstige Fälle $= [0, \frac{1}{2}] \times [0, 2\pi)$. Dann ist das $\lambda^2(A) =$ Volumen von $A = \pi$ und somit die Wahrscheinlichkeit $Pr(A) = \frac{\lambda^2(A)}{\lambda^2(\Omega)} = \frac{1}{2}$.

- Das buffonsche Nadelexperiment.
  Wir betrachten hier Parallelen in $\mathbb{R}^2$ mit Abstand $2a$ voneinander. Wir werfen jetzt eine Nadel der Länge $2l$ in die Ebene und wollen jetzt berechnen, wie hoch die Wahrscheinlichkeit ist, dass die Nadel eine Gerade schneidet.
  Ein Ergebnis ist bestimmt durch den Winkel ϕ zwischen Nadel und Gerade und dem Abstand x zwischen Gerade und Mittelpunkt der Nadel. Demnach ist $\Omega = [0, a] \times [0, \pi)$. Die günstigen Fälle sind $G = \{(x, \phi) | x \leq l \sin \phi\}$. Die Fläche von Ω ist demnach $a * \pi$ und die Fläche von $G = \int_0^\pi l * \sin \phi d\phi = l(\cos 0 - \cos \pi) = 2l$. Dann ist $Pr(G) = \frac{2l}{a\pi}$ für $l = \frac{a}{2}$ ergibt sich also $Pr(G) = \frac{1}{\pi}$.
  Mit dem Versuch kann übrigens π experimentell bestimmt werden.

24.0.2. Bedingte Wahrscheinlichkeit und Unabhängigkeit von Ereignissen

Definition 49. *Sei $(\Omega, \mathcal{F}, Pr)$ der Wahrscheinlichkeitsraum und $A, B \in \mathcal{F}$. Dann ist $Pr(A|B)$ die Wahrscheinlichkeit von A unter der Bedingung A definiert als*

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$$

Elementare Eigenschaften

1. $0 \leq Pr(A|B) \leq 1$
2. $A \cap B = \emptyset \Rightarrow Pr(A|B) = 0$
3. $A = \bigcup A_i \Rightarrow Pr(A|B) = \sum Pr(A_i|B)$
4. $B \subseteq A \Rightarrow Pr(A|B) = 1$.

Man kann beobachten, dass wenn $Pr(-|B)$ ist Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{F})$.

Formel für die totale Wahrscheinlichkeit $\Omega = B_1 \cup B_2 \cup \ldots \cup B_n$ sind paarweise disjunkt.

$$Pr(A) = \sum_{i=1}^{n} Pr(A|B_i) * Pr(B_i)$$

Das wird jetzt noch kurz bewiesen. Man kann zuerst beobachten, dass die $A \cap B_i$ disjunkt sind. Deshalb kann man schreiben

$$\begin{aligned} Pr(A) &= \sum_{i=1}^{n} Pr(A \cap B_i) \\ &= \sum_{i=1}^{n} \underbrace{\frac{Pr(A \cap B_i)}{Pr(B_i)}}_{Pr(A|B)} * Pr(B_i) \end{aligned}$$

Formel von Bayes $A_1, \ldots, A_n$ sei Zerlegung von Ω.

$$Pr(A_i|B) = \frac{Pr(A_i) * Pr(B|A_i)}{\sum_k Pr(A_k) Pr(B|A_k)}$$

Auch wieder ein Beweis:

$$Pr(A_i|B) = \frac{Pr(A_i \cap B)}{Pr(B)} \frac{Pr(A_i)}{Pr(A_i)} = \frac{Pr(B|A_i) * Pr(A_i)}{Pr(B)}$$

Jetzt noch irgendwas mit der Formel für die totale Wahrscheinlichkeit und dann fertig.

Unabhängigkeit Seien $A_1, A_2 \in \mathcal{F}$. Zwei Ereignisse sind unabhängig, wenn das Eintreten von A_1 nicht abhängt von A_2 und andersherum. Also

$$\begin{aligned} Pr(A_1|A_2) &= Pr(A_1) \\ Pr(A_2|A_1) &= Pr(A_2) \end{aligned}$$

oder auch anders gesagt

$$Pr(A_1 \cap A_2) = Pr(A_1) * Pr(A_2)$$

Den Begriff kann man verallgemeiner: Eine Familie heißt so vollständig oder total unabhängig, falls für beliebige $A_{i1} \ldots A_{in}$ $Pr(\bigcap A_{ij}) = \prod Pr(A_{ij})$.

25. Vorlesung, 23.1.2008

Was noch zur Unabhängigkeit zu beachten ist:

- Unabhängigkeit ist eine Eigenschaft von ≥ 2 Ereignissen
- Disjunkte Ereignisse sind niemals unabhängig

25.0.3. Zufallsvariable

Sei $(\Omega, \mathcal{P}(\Omega), Pr)$ der Wahrscheinlichkeitsraum.

Definition 50. *Eine Zufallsvariable ist eine Funktion*

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

Die Zufallsvariable induziert eine Wahrscheinlichkeitsverteilung auf $Im(X) \subseteq \mathbb{R}$. Ein Ereignis ist dann $(X = x) = \{\omega \in \Omega | X(\omega) = x\}$.

Definition 51. *Das Wahrscheinlichkeitsmaß Pr_X ist dann*

$$x \in Im(X) \qquad Pr_X(x) = Pr(X = x) = Pr(\{\omega | X(\omega) = x\}) = \sum Pr(\omega)$$

Wir betrachten dann $(Im(X), \mathcal{P}Im(X)), Pr_X)$ als Wahrscheinlichkeitsraum.
Das Ziel ist, dass Verhalten der Funktion $X : \Omega \to \mathbb{R}$ durch Zahlen zu beschreiben. Die wichtigsten sind hierbei der Erwartungswert $E(X)$ und die Varianz $V(X)$.
Der Erwartungswert ist der durchschnittliche Wert von X und wird berechnet durch

$$E(X) = \sum_{\omega \in \Omega} Pr(\omega) * X(\omega) = \sum_{x \in Im(X)} x * Pr_X(x) = \sum_{x \in Im(X)} x * Pr(x = X)$$

Bei einer unendlichen Menge gilt das nur, falls die Summe absolut konvergent ist.
Wenn das Bild der Zufallsvariable in den natürlichen Zahlen liegt, dann spricht man auch von einer Zählvariable.

$$\begin{aligned} E(X) &= \sum_{i=0}^{\infty} i * Pr(X = i) = \sum_{i=1}^{\infty} iPr(X = i) \\ &= \sum_{i=0}^{\infty} Pr(X > i) \end{aligned}$$

(siehe Mafi 1)

Linearität des Erwartungswertes Wenn $\forall\omega : X(\omega) = X_1(\omega) + X_1(\omega) + \ldots + X_n(\omega)$ bzw. kurz geschrieben $X = X_1 + X_1 + \ldots + X_n$ dann gilt

$$E(X) = \sum_{i=1}^{n} E(X_i) \qquad E(c * X) = c * E(X)$$

Spezielle Verteilungen durch dir Zufallsvariable

Bernoulli-Verteilung Ω ist 2 elementig, nämlich $\Omega = \{0, 1\}$. Dann geht $X : \Omega \to \mathbb{R}$ mit $0 \to 0$ als Misserfolg und $1 \to 1$ als Erfolg. Und $Pr_X(0) = 1 - p \qquad Pr_X(1) = p$. Zusätzlich gilt natürlich $0 \leq p \leq 1$. Der Erwartungswert ist hier dann $E(X) = 0 * (1 - p) + 1 * p = p$.

Binomialverteilung X zählt die Erfolge bei n hintereinander ausgeführten Bernoulli-Experimenten. Dann ist $Im(X) = \{0, 1, \ldots n\}$. Die Wahrscheinlichkeit für genau k mal Erfolg ist

$$Pr_X(k) = \binom{n}{k} p^k * (1 - p)^{n-k}$$

Das ganze kann man über den Binomischen Satz herleiten ...
$X_i : \Omega \to \mathbb{R}$ und $X_i((\omega_1, \ldots, \omega_n)) = 0$ falls $\omega_i = 0$ bzw. $= 1$ falls $\omega_i = 1$. X ist demnach wegen der Linearität $X = X_1 + \ldots + X_n$. Der Erwartungswert ergibt sich also

$$E(X) = \sum_{i=1}^{n} E(X_i) = n * p$$

Geometrische Verteilung Ein Bernoulli-Experiment wird durchgeführt bis ein Erfolg eintritt. Dabei sagt die Zufallsvariable aus, wie oft man das Experiment durchführen muss, bis Erfolg eintritt, also $Im(X) = \{1, 2, \ldots\}$. Die Wahrscheinlichkeit für genau k Versuche $Pr(X = k)$ ergibt sich so, dass man zuerst $k - 1$ mal Misserfolg hat, und dann ein mal einen Erfolg, also

$$Pr(X = k) = (1 - p)^{k-1} * p$$

Der Erwartungswert ist folglich

$$E(X) = \sum_{k=1}^{\infty} k(1 - p)^{k-1} p = \frac{1}{p}$$

Poisson-Verteilung mit $\lambda > 0$. Das Bild der Zufallsvariable ist wieder $Im(X) = \{1, 2, \ldots\}$

$$Pr_X(k) = \frac{1}{k!} \lambda^k e^{-\lambda} \qquad E(X) = \lambda$$

Das ist wichtig für Binomialverteilung bei der n sehr groß ist, p sehr klein und k sehr viel kleiner als n ist.

Allgemeiner Begriff einer Zufallsvariable

$(\Omega, \mathcal{F}, Pr)$ sei jetzt der Wahrscheinlichkeitsraum. Man beachte, dass $\mathcal{F}$ nur noch eine Auswahl an Ereignissen ist.
$X : \Omega \to \mathbb{R}$ heißt Zufallsvariable, falls

$$\forall x \in \mathbb{R} : \{\omega \in \Omega | X(\omega) \leq x\} \in \mathcal{F}$$

Man beachte, dass wir hier nicht mehr die Gleichheit fordern, sonder auf ein Intervall übergehen.
Dann heißt $F_X(x) = Pr(\{\omega \in \Omega | X(\omega) \leq x\}) = Pr(X \leq x)$ die Verteilungsfunktion von X.

Eigenschaften

- Schwach monoton steigend, d.h. aus $x \leq y$ folgt $F_X(x) \leq F_X(y)$
- $\lim_{x \to -\infty} F_X(x) = 0$ und $\lim_{x \to +\infty} F_X(x) = 1$
- Stetigkeit von rechts: $\lim_{\epsilon \to 0^+} F_X(x + \epsilon) = F_X(x)$

26. Vorlesung, 28.1.2008

Definition 52. *Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heißt stetig, falls $\exists f : \mathbb{R} \to \mathbb{R}^{\geq 0}$ mit*

$$F_X(x) = \int_{-\infty}^{x} f(t)dt$$

f nennt man auch die Dichte von X.

Achtung Es gibt Zufallsvariablen, die weder diskret noch stetig sind.

Beispiele

1. Wir haben eine diskrete Zufallsvariable mit $Im(X) = \mathbb{N}$. $Pr(X = k) = p_k$. Dann ist

$$F_X(x) = \begin{cases} 0 & x < 0 \\ p_0 + p_1 + \ldots + p_{\lfloor k \rfloor} & x \geq 0 \end{cases}$$

2. X sei gleichverteilt auf dem Intervall $[a, b] \subseteq \mathbb{R}$ mit $a < b \in \mathbb{R}$.

$$F_X(x) = Pr(\{\omega | X(\omega) \leq x\}) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

Wie sieht das jetzt mit der Dichte aus? Die Dichte ergibt sich als Ableitung von $F_X(x)$, das ist

$$f(x) = \begin{cases} 0 & x < a \\ \frac{1}{b-a} & a \leq x \leq b \\ 0 & x > b \end{cases}$$

Die Dichte kann Werte größer als eins annehmen, z.B. bei $a = 0,5$ und $b = 1$. Dann ist die Dichte 2.

Erwartungswert im diskreten bzw. stetigen Fall

Wir haben wieder $X : \Omega \to \mathbb{R}$.

$$E(X) = \sum_{x \in Im(X)} x * Pr(X \leq x)$$

falls diese Summe absolut konvergiert.

Im stetigen Fall haben wir $X : \Omega \to \mathbb{R}$ sei eine stetige Zufallsvariable mit Dichtefunktion f.

$$E(X) = \int_{-\infty}^{+\infty} x * f(x)dx$$

falls dieses uneigentliche Integral existiert.

Eigenschaften X, Y sind stetige Zufallsvariablen über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, Pr)$. Auch hier gilt dann wieder die Linearität des Erwartungswertes: $E(X+Y) = E(X) + E(Y)$ und $E(c * X) = c * E(X)$ mit $c \in \mathbb{R}$.

Satz 30. *X, Y seien unabhängige, diskrete Zufallsvariablen, d.h. $\forall a, b \in \mathbb{R} : Pr((X \leq a) \cap (Y \leq b)) = Pr(X \leq a) * Pr(Y \leq b)$. Dann ist*

$$E(X * Y) = E(X) * E(Y)$$

Der Satz bleibt ohne Beweis hier.

Satz 31. *X stetige Zufallsvariable mit Dichte f und g stetig. Dann ist $g * X$ eine stetige Zufallsvariable.*

$$E(g * X) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

Beispiel $\Omega = [0,1]^2$, das ist das Einheitsquadrat. Wir wählen zufällig einen Punkt im Einheitsquadrat bezüglich Gleichverteilung. Jetzt bewerten wir dieses Experiment mit zwei Zufallsvariablen. Dabei ist X die Variable, die den Abstand von z zum Rand des Quadrates angibt. Und Y ist die Zufallsvariable, die die Fläche des größten Kreises mit Mittelpunkt z angibt.
Wir interessieren uns jetzt für die Erwartungswerte $E(X), E(Y)$.
Dazu bestimmen wir die Verteilungsfunktion, um daraus die Dichtefunktion zu bekommen.

$$\begin{aligned} F_X(x) &= Pr(\{z \in [0,1]^2 | X(z) \leq x\}) \\ &= \text{Fläche von } \{z \in [0,1]^2 | X(z) \leq x\} \end{aligned}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 4x - 4x^2 & 0 \leq x \leq \frac{1}{2} \\ 1 & x \geq \frac{1}{2} \end{cases}$$

Die Dichte ist wie gesagt die Ableitung der Verteilungsfunktion. Also

$$f(x) = \begin{cases} 0 & x < 0 \\ 4 - 8x & 0 \leq x \leq \frac{1}{2} \\ 0 & x \geq \frac{1}{2} \end{cases}$$

Der Erwartungswert ist also

$$E(X) = \int_{-\infty}^{\infty} x * f(x)dx = \int_0^{\frac{1}{2}} (4x - 8x^2)dx = 2x^2 - \frac{8}{3}x^3|_0^{\frac{1}{2}} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

Der Erwartungswert von Y kann aufgefasst werden als $Y : \omega \to^X \mathbb{R} \to^Y \mathbb{R}$ mit $y(x) = \pi x^2$.

$$E(Y) = \int_{-\infty}^{\infty} \pi x^2 f(x)dx = \pi \int_0^{\frac{1}{2}} 4x^2 - 8x^3 dx = \pi(\frac{4}{3}x^3 - 2x^4)|_0^{\frac{1}{2}} = \pi(\frac{1}{6} - \frac{1}{8}) = \frac{\pi}{24}$$

27. Vorlesung, 30.1.2008

Bedingter Erwartungswert

$$E(X|B) = \sum_{x \in Im(X)} x * Pr(X = x|B)$$

Satz 32. *Wenn $\{B_1, B_2 \ldots\}$ eine abzählbare Menge von Ereignissen mit $Pr(B_i) > 0$ Partition von Ω, so gilt*

$$E(X) = \sum_{i=1}^{\infty} E(X|B_i) Pr(B_i)$$

Beweis 32. $\sum_i \sum_{x \in Im(X)} = \frac{Pr(\{X=x\} \cap B_i) Pr(B_i)}{Pr(B_i)}$. *Die B_i sind disjunkt, daher gilt* $= \sum_? x Pr(\{X = x\} \cap \bigcup B_i) = \sum_? x Pr(\{X = x\}) = E(X)$ □

Beispiel Wir werfen eine faire Münze so lange, bis sie das Bild wechselt. Sei X die Zufallsvariable und $X(\omega)$ die Länge des initialen Rangs.
H sei das Ereignis, dass nach dem ersten Wurf Kopf liegt. Dementsprechend ist H^C das Komplement, dass also Zahl liegt.

$$\begin{aligned}
Pr(X = k|H) &= \frac{Pr(X = k \cap H)}{Pr(H)} = \frac{p^k q}{p} = p^{k-1} q \\
Pr(X = k|H^C) &= \frac{Pr(X = k \cap H^C)}{Pr(H^C)} = \frac{q^k p}{q} = q^{k-1} p \\
E(X) &= E(X|H) * Pr(H) + E(X|H^C) * Pr(H^C) \\
E(X|H) &= \sum_{k=1}^{\infty} k * p^{k-1} q = \frac{q}{(1-p)^2} = \frac{1}{q} \\
E(X|H^C) &= \sum_{k=1}^{\infty} k * q^{k-1} p = \frac{p}{(1-q)^2} = \frac{1}{p} \\
E(X) &= \frac{1}{q} * p + \frac{1}{p} q = \frac{p^+ q^2 + 2pq - 2pq}{pg} \\
&= \frac{1}{pq} - 2
\end{aligned}$$

In unserem Beispiel ist ja $p = q = \frac{1}{2}$ und demnach dann $E(X) = 2$.

Alternativ kann man sich zwei Zufallsvariablen X_H und X_T definieren, die die Anzahl von Zahlen bzw. der Köpfe am Anfang zählen. Dann ist $X = X_H + X_T$, da eins immer null ist. Dann gilt nach der Linearität $E(X) = E(X_H) + E(X_T)$. Dann ist

$$\begin{aligned} E(X_H) &= \frac{1}{q} - 1 \\ E(X_T) &= \frac{1}{p} - 1 \\ E(X) &= \frac{1}{q} + \frac{1}{p} - 2 = \frac{1}{pq} - 2 \end{aligned}$$

Momente

Definition 53. *Sei X eine Zufallsvariable mit dem Erwartungswert $E(X)$. Dann heißt $E(X^k)$ mit $k \in \mathbb{N}^+$ das k-te Moment der Zufallsvariable.*
$E((X - E(X))^k)$ heißt k-tes zentrales Moment der Zufallsvariable.
Die Varianz von X ist $Var(X) = E((X - E(X))^2)$. Das ist die erwartete Größe des quadrierten Abstandes zwischen X und $E(X)$.

Man nimmt übrigens nicht $E(X - E(X))$ das das immer null ist und auch nicht den Betrag davon, da Beträge doof sind. Deshalb quadriert man halt, damit immer schön was positives hat.

$$\begin{aligned} Var(X) &= E((X - E(X))^2) = E(X^2 - 2XE(X) + (E(X))^2) \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2) \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

Beispiele

1. Gleichverteilung über Intervall $[a, b] \in \mathbb{R}$. Die Verteilungsfunktion ist dann

$$F_X(x) = Pr(X \leq x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

Die Dichte $f(x)$ ist dann ja die Ableitung der Verteilungsfunktion demnach

$$f(x) = \begin{cases} 0 & x < a \vee x > b \\ \frac{1}{b-a} & a \leq x \leq b \end{cases}$$

Jetzt wollen wir noch den Erwartungswert.

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{2(b-a)}x^2\Big|_a^b \\ &= \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2} \end{aligned}$$

Die Varianz ist dann

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ E(X^2) &= \int_a^b x^2\frac{1}{b-a}dx = \frac{1}{3(b-a)}x^3\Big|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3} \\ Var(X) &= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+b^2}{4} \\ &= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12} \end{aligned}$$

2. Exponentialverteilung. Sei $\lambda > 0$ ein Parameter. Wir haben als Verteilungsfunktion

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$$

Und jetzt wieder die Dichte als Ableitung

$$f(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x > 0 \end{cases}$$

Dann ist der Erwartungswert

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^{\infty} x\lambda e^{-\lambda x}dx = \frac{1}{\lambda}$$

Die Varianz ist dann

$$Var(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

? X sei exponentiell verteilt mit $\lambda > 0$.

$$\begin{aligned} E(X^k) &= \int_0^{\infty} x^k * f(x)dx = \int_0^{\infty} x^k\lambda e^{-\lambda x}dx \\ &= \int_0^{\infty} x^k\lambda e^{-\lambda x}dx \\ &= \underbrace{(-x^k e^{-\lambda x})|_0^{\infty}}_{\to 0} + \int_0^{\infty} kx^{k-1}e^{-\lambda x} \\ &= \frac{k}{\lambda}(E(X))^{k-1} = \frac{k(k-1}{\lambda^2}E(X)^{k-2} = \frac{k'}{\lambda^k}E(X') = \frac{k'}{\lambda^k} \end{aligned}$$

28. Vorlesung, 4.2.2008

28.0.4. Ungleichungen von Markov und Tschebychev

Markov-Ungleichung

$X : \Omega \to \mathbb{R}^{\geq 0}$ Zufallsvariable. Sei $E(X)$ der Erwartungswert und $t > 0$. Dann ist

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

Beispiele Durchschnittsgröße 1,50m. Was ist jetzt die Wahrscheinlichkeit, dass Person $\geq 2m$ ist $\leq 0,75$. Also hier $E(X) = 1,50$ und $t = 2$. Mit einem kleinen Trick kann man sagen, angenommen dass keiner kleiner als einen Meter ist. Dann ist $Y = X - 1$. Y ist dann auch eine positive Zufallsvariable. $E(Y) = 0,5$. Dann ist $Pr(X \geq 2) = Pr(Y \geq 1)$. Das schätzen wir mit Markov jetzt ab.

$$Pr(Y \geq 1) \leq \frac{0,5}{1} = 0,5$$

Beweis 33. ***Für eine diskrete Zufallsvariable***

$$\begin{aligned}
E(X) &= \sum_{x \in Im(X)} x * Pr(X = x) \\
&= \sum_{x \in Im(X), x < t} xPr(X = x) + \sum_{x \in Im(X), t \geq x} xPr(X = x) \\
&\geq 0 + \sum_{s \in Im(X), t \geq x} x * Pr(X = x) \\
&\geq t \sum_{x \geq t} Pr(X = x) \\
&= t * Pr(X \geq t)
\end{aligned}$$

Für eine stetige Zufallsvariable

$$\begin{aligned} E(X) &= \int_0^\infty x * f_X(x)dx \\ &= \int_0^t x * f_X(x)dx + \int_t^\infty x * f_X(x)dx \\ &\geq 0 + \int_t^\infty x * f_X(x)dx \\ &\geq t * Pr(X \geq t) \end{aligned}$$

□

Tschebytschev Ungleichung

Sei Y eine Zufallsvariable und sei $E(Y) \leq \infty$. Dann gilt

$$Pr(|Y| \geq t) \leq \frac{1}{t^2} E(Y^2)$$

Das ist eine sehr einfache Form. man kann auch speziell sagen $Y = X * E(X)$ und dann ist

$$Pr(|X - E(X)| \geq t) \leq \frac{Var(X)}{t^2}$$

Beweis 34. $A = \{\omega \in \Omega || Y(\omega)| \geq t\}$ *Dann ist* $E(Y^2) = E(Y^2|A) * Pr(A) + E(Y^2|A^C) * Pr(A^C) \geq E(Y^2|A) * Pr(A) \geq t^2 Pr(A)$ □

Achtung Tschebytschev schätz den Betrag ab, also die Abweichung nach oben und nach unten.

Beispiel Es wird eine faire Münze n mal geworfen. Die Zufallsvariable X zählt die Anzahl der Köpfe. Dann ist $E(X) = \frac{n}{2}$. Die Varianz ist hier dann $Var(X) = n(p - p^2) = \frac{n}{4}$.

$$Pr(|X - \frac{n}{2}| \geq \frac{n}{c}) \leq \frac{c^2}{n^2} * \frac{n}{4} = \frac{c^2}{4n}$$

Für $n \to \infty$ geht das ganze gegen 0.

28.0.5. Normalverteilung

Sei X die Zufallsvariable mit Dichtefunktion

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

nennt man auch die Normale Dichte. Die Verteilungsfunktion ist

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{t^2}{2}} dt \quad \text{Standardnormalverteilung}$$

Für dieses Integral gibt es leider keine geschlossene analytische Form. Die Varianz von dem Teil ist 1 und der Erwartungswert 0.
Man schreibt für die Varianz übrigens auch σ^2, σ ist dann die Standardabweichung. Für den Erwartungswert benutzt man auch die Bezeichnung μ. Man kann für alle Standardabweichungen und Erwartungswerte eine Normalverteilung finden, das ist die allgemeine Normalverteilung.

allgemeine Normalverteilung

$N(\mu, \sigma^2)$-Verteilung mit Erwartungswert μ und Varianz σ^2. Die Dichte ist dann

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

Fakt Wenn eine Zufallsvariable X $N(\mu, \sigma^2)$-verteilt ist, so ist $Y - \frac{X-\mu}{\sigma}$ $N(0,1)$-verteilt.

$$Pr(\mu - \sigma \leq X \leq \mu + \sigma) \approx 69\%$$
$$Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 95,5\%$$
$$Pr(\mu - 3\sigma \leq X \leq \mu + 4\sigma) \approx 99,7\%$$

Approximation der Binomialverteilung durch die Normalverteilung

Wir setzen voraus, dass $n * p > 5$ und $n * (1-p) > 5$.

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \underbrace{\Phi\left(\frac{k + \frac{1}{2} - n * p}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - \frac{1}{2} - n * p}{\sqrt{np(1-p)}}\right)}_{\text{Fläche über Intervall } [k-\frac{1}{2}, k+\frac{1}{2}]}$$

Beispiel Flugzeugüberbuchung. Angenommen, wir habne 200 Sitze im Flugzeug. Die Erfahrung sagt, dass 5% der Passagiere nicht erscheinen. Wenn die Gesellschaft bei Überbuchung 1 von 50 Fällen in Verlegenheit kommt, ist das okay. Wir wollen jetzt wissen, wie viele Reservierungen man annehmen darf?
Sei n die Anzahl der Reservierungen und X gibt die Anzahl der erscheinenden Personen an. Wir machen Binomialverteilung, dann ist $p = 0,95$. $\mu = E(X) = n * p = 0,95 * n$. Die Varianz ist $\sigma^2 = Var(X) = n * p(1-p) = 0,0475n$. Dann ist die Standardabweichung $\sigma = 0,2179\sqrt{n}$. Jetzt kann man in so nen Tabellen nachgucken, und erhält $n = 217$.

29. Vorlesung, 6.2.2008

Wir würfeln 6000 mal und unsere Zufallsvariable X zählt die Anzahl der 6en. Dann ist $E(X) = n*p = 6000*\frac{1}{6} = 1000$ und die Varianz ist $Var(X) = n*p(1-p) = 6000*\frac{1}{6}*\frac{5}{6} = 833,\overline{3}$. Somit ist die Standardabweichung $\sigma = \sqrt{Var(X)} = 28,87$
Uns interessiert jetzt die Wahrscheinlichkeit $Pr(X \geq 1100)$. Exakt wäre das

$$Pr(X \geq 1100) = \sum_{k=1100}^{6000} \binom{6000}{k} \left(\frac{1}{6}\right)^k * \left(\frac{5}{6}\right)^{6000-k}$$

Aber das ist ja nen bissel schwierig zu berechnen. Wir schätzen erstmal nach Markov und Tschebytschev ab:

$$Pr(X \geq 1100) \leq \frac{E(X)}{1100} = \frac{1000}{1100} \approx 0,9$$
$$Pr(|X - 1000| \geq 100) \leq \frac{833,\overline{3}}{100^2} = 0,08\overline{3}$$

Da man bei Tschebytschev nach oben und unten abschätzt, kann man das Ergebnis halbieren. Jetzt machen wir noch eine Approximation durch die Normalverteilung. Wir suchen jetzt die Fläche, die die Dichtefunktion mit der x-Achse bildet von $x = 1100$ bis zum Ende.
Wir schreiben das Problem jetzt auf die Standardnormalverteilung um, damit man schön in die Tabellen gucken kann. Wir schreiben $1100 = \underbrace{1000}_{\mu} + 3,46 * \underbrace{28,87}_{\sigma}$

Fakt Wenn eine Zufallsvariable X $N(\mu, \sigma^2)$ verteilt ist, dann ist $Y = \frac{X-\mu}{\sigma}$ $N(0,1)$, also standardverteilt. Demnach gilt $X * \mu = Y * \sigma$.
Wenn man wissen will, was bei der Standardverteilung rechts von einem Wert r liegt, dann entpsricht das bei einer $N(\mu, \sigma^2)$-Verteilung der Frage, was rechts von $\mu + r * \sigma$ liegt.
Für unser Beispiel ergibt sich dann $1 - \Phi(3,46) \approx 0,00028$.

Wahrscheinlichkeit für bestimmte Ereignisse bei Bionimalverteilung X

$X * \mu < r * \sigma$, dann muss man $\Phi(r)$ nach gucken. Für $X + \mu \geq r * \sigma$ muss man dann $1 - \Phi(r)$ rechnen. Wenn wir die Wahrscheinlichkeit zwischen $-r$ und r wissen wollen, also $|X - \mu| \leq r * \sigma$, dann ist das $2\Phi(r) - 1$. Jetzt können wir noch nach dem Komplement dazu fragen, also $|X - \mu| > r\sigma$ und das berechnet sich über $2 - 2\Phi(r)$.

Anwendungen in der Statistik

Wir haben hier nur Stichproben zu Verfügung und schätzen davon ausgehend Parameter ab. Diese sind natürlich fehlerbehaftet. Wir Suchen nun ein Intervall innerhalb dessen das Ergebnis mit gewisser Wahrscheinlichkeit liegt. Das ist das sogenannte Konfidenzintervall.

Beispiel Wir machen ne Wahlumfrage. Wir haben eine Stichprobe von 1000 Befragten, von denen 400 die Partei Z wählen wollen. Wir nahmen an, dass es sich um eine Binomialverteilung handelt, dann ist $p = 0,4$. Das erwartete Ergebnis ist, dass 40% Z wählen. $\mu = 400 * E(X)$, wobei X die Zufallsvariable, die zählt, wie oft Z gewählt wird. Die Varianz ist dann $\sigma^2 = n*p*(1-p) = 240$ und die Standardabweichung dann $\sigma = 15,49$. Wir approximieren wieder durch die Normalverteilung.
Wir wollen jetzt ein Intervall $[\mu - r\sigma, \mu + r\sigma]$, so dass das Integral der Dichtefunktion über dem Intervall gleich $0,95\%$ ist.
Aus der Tabelle für $N(0,1)$ bekommt man $\Phi(r = 1,96) = 0,975$ (Und zwar, weil $0,975 - 0,025 = 0,95$) und man bekommt dann für den Wert von Φ 1,96 heraus. Wegen der Symmetrie ist $\int_{-1,96}^{1,96} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 0,95$. Mit einem Konfidenzniveau von 95% erhält Partei T zwiscen [369,6 ; 430,4] Stimmen.

29.0.6. Grenzwertsätze

Angenommen, es wird ein fairer Würfel 10^6 mal geworfen. Im Durchschnitt wurden $3,500867$ Augen geworfen. Das ist ja auch okay, weil man für einen Wurf 3,5 erwartet. Es sollte nun möglich sein zu beweisen, dass

$$\frac{1}{n}(X_1, X_2 + \ldots + X_n) \ \to \mu \text{ mit } n \to \infty$$

Die erste Schwierigkeit tritt auf, weil wir erstmal betrachten müssen, was Konvergenz überhaupt heißt.

Definition 54. *Eine Sequenz $Z_1, Z_2, \ldots$ von Zufallsvariablen konvergiert bezüglich der Varianz gegen Zufallsvariable Z falls*

$$E((Z_n - Z)^2) \to 0 \text{ für } n \to \infty$$

Für $E(Y^2) = 0$, dann ist $Y = 0$ mit Wahrscheinlichkeit 1.

Satz 33. *Sei $X_1, X_2, \ldots$ Folge von unabhängigen Zufallsvariablen, alle mit Erwartungswert μ und Varianz σ^2. Dann konvergiert*

$$\frac{1}{n}(X_1 + X_2 + \ldots + X_n) \to \mu \text{ Konvergenz bezüglich Varianz}$$

Beweis 35. $S_n = X_1 + \ldots + X_n$ *n-te Partialsumme.* $E(\frac{1}{n}S_n) = \frac{1}{n}E(S_n) = \frac{1}{n} * ? = \mu$.

30. Vorlesung, 11.2.2008

Definition 55. *Die Folge $Z_1, Z_2, \ldots$ konvergiert gegen Zufallsvariable Z mit einer Wahrscheinlichkeit, falls*

$$\forall \epsilon > 0 : Pr(|Z_n - Z| > \epsilon) \to 0 \text{ für } n \to \infty$$

genauer

$$\forall \delta > 0 \forall \epsilon > 0 : |Z_n - Z| \leq \epsilon \text{ mit Wahrscheinlichkeit } \geq 1 - \delta \text{ für } n > n_{\epsilon,\delta}$$

Beobachtung Falls $Z_n \to Z$ bezüglich Varianz geht, dann folgt $Z_n \to Z$ auch bezüglich der Wahrscheinlichkeit. Der Beweis geht hier über die Tschebytschev-Ungleichung.

$$Pr(|Z_n - Z| \geq \epsilon) \leq \frac{1}{\epsilon^2} E((Z_n - Z)^2)$$

30.0.7. Schwaches Gesetz der großen Zahlen

$X_1, X_2, \ldots$ Folge von unabhängigen Zufallsvariablen mit gleichem Erwartungswert μ und alle haben Varianz σ^2. Dann gilt für die Partialsumme

$$\frac{1}{n}(X_1 + \ldots + X_n) \to \mu \text{ mit Wahrscheinlichkeit}$$

Beispiel Die Konvergenz bezüglich der Wahrscheinlichkeit impliziert nicht unbedingt eine Konvergenz bezüglich der Varianz. $Z = 0$ und Z_n habe eine Verteilung

$$Pr(Z_n = 0) = 1 - \frac{1}{n} \qquad Pr(Z_n = n) = \frac{1}{n}$$

Die Konvergenz bezüglich Wahrscheinlichkeit:

$$\forall \epsilon > 0 \forall n > n_\epsilon : Pr(|Z_n| > \epsilon) = Pr(Z_n = n) = \frac{1}{n} \to 0 \text{ für } n \to \infty$$

Jetzt die Konvergenz bezüglich Varianz:

$$E((Z_n - 0)^2) = E(Z_n^2) = 0^2 * (1 - \frac{1}{n}) + n^2 \frac{1}{n} = n \to \infty \text{ für } n \to \infty$$

Zentraler Grenzwertsatz

$X_1, \ldots, X_n$ unabhängige Zufallsvariablen, die alle Erwartungswert μ haben und Varianz $\sigma^2 > 0$.
Bisher hatten wir $S_n = X_1 + \ldots + X_n \to n * \mu$ mit Wahrscheinlichkeit. jetzt wollen wir $S_n = \mu * n$. Die Standartisierte Form von S_n ist

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}}$$

Das Teil hat dann Erwartungswert 0 und die Varianz 1. Und zwar, weil $E(Z_n) = \frac{E(S_n)}{\sqrt{Var(S_n)}} - \frac{E(S_n)}{\sqrt{Var(S_n)}} = 0$.

$$Pr(Z_n \leq x) \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \qquad \text{mit } \; n \to \infty$$

Ein Spezialfall davon ist, dass alle X_i binomialverteilt sind.

Merke Die Normalverteilung ist eine sinnvolle Approximation für Fälle, in denen sich die Zufallsvariable aus vielen gleichartigen Einzelfällen zusammensetzt.

30.0.8. RSA

Chinesischer Restsatz

Da ist eine unbekannte Anzahl x von Dingen: Geteilt durch 3 ist der Rest 2. Geteilt durch 5 ist Rest 3 und geteilt durch 7 bleibt Rest 2. Wie groß ist jetzt x? Die Lösung ist hierfür 23.

Satz 34. *Seien $m_1, \ldots m_n$ sind paarweise koprime positive ganze Zahlen. Koprim bedeutet, dass $ggT(x_i, x_j) = 1$. Dann hat das System*

$$\begin{aligned} x &\equiv a_1 (\mod m_1) \\ x &\equiv a_2 (\mod m_2) \\ &\vdots \\ x &\equiv a_n (\mod m_n) \end{aligned}$$

Genau eine Lösung $0 \leq x < m_1 * m_2 * \ldots * m_n$. Alle anderen Lösungen sind kongruent zu $x \mod m$.

Beweis 36. $M_k = \frac{m}{m_k} = m_1 * \ldots * m_{k-1} * m_{k+1} * \ldots$. *Der größte gemeinsame Teiler ist dann $ggT(m_k, M_k) = 1$. D.h. $\exists y_k : y_k * M_k \equiv 1 (\mod m_k)$.*
Sei $x = a_1 M_1 y_1 + a_2 M_2 y_2 + \ldots + a_n M_n y_n$. Wir zeigen, dass x die Lösung ist.
Wir behaupten, dass $M_j \equiv 0 (\mod m_k)$ für $j = k$. Also $x \equiv a_k M_k y_k (\mod m_k)$ aber daraus folgt gleich $x \equiv a_k (\mod m_k)$.
Jetzt zeigen wir noch, dass das x eindeutig ist. Sei $0 \leq x' < m$ auch Lösung. Dann ist $x - x' \equiv 0 (\mod m_1)$ und so weiter bis $x - x' \equiv 0 (\mod m_n)$. Daraus folgt, dass $x - x' \equiv 0 (\mod m)$ und daraus folgt $x - x' = 0$ und somit $x = x'$.

Der kleine Fermat

p ist Primzahl, $a \in \mathbb{Z}$. Dann ist $a^p \equiv a(\mod p)$, d.h. $a^p - a$ ist Vielfaches von p.

Beweis 37. *Ist p Teiler von a, dann ist $a^p \equiv 0(\mod p)$. Wenn p nicht Teiler von a ist, dann soll gelten $a^{p-1} \equiv 1(\mod p)$. Wir können uns auch auf positive Zahlen beschränken. Das kann man jetzt per Vollständiger Induktion nach a zeigen.*

Induktionsanfang $a = 0$. *Ist einfach. $p|0^p - 0$ gilt natürlich.*

Induktionsschritt *Wir zeigen $a \to a+1$.* $(a+1)^p - (a+1) = a^p + \binom{p}{1}a^{p-1} * 1 + \ldots + \binom{p}{p-1}a * 1^{p-1} + 1^p - (a+1) \equiv a^p + 1 - a - 1 = a^p - a$ *Um zu verstehen, dass die Formel gilt, gucken wir uns mal den Binomialkoeffizient an.*

$$\binom{p}{k} = \frac{p * (p-1) * \ldots * (p-k+1)}{1 * 2 * \ldots * k}$$

Und das ist ein Vielfaches von p. Daher fallen fast alle Summanden raus.

Man kann das aber auch noch eleganter zeigen:
Sei $Z_p^* = \{\overline{1}, \overline{2}, \ldots, \overline{p-1}\}$ sind die Restklassen $\mod p$. Sei $f : Z_p^* \to Z_p^*$ mit $x \to a * x \mod p$. Wir behaupten, dass das eine injektive Abbildung ist.

$$\begin{aligned} ax &\equiv ay \mod p \\ a(x-y) &\equiv 0 \mod p \end{aligned}$$

Also $p|x-y$. Daraus folgt, dass $x = y$.
Damit ist das Ding auch bijektiv, da Definitionsbereich und Wertebereich gleich sind.

31. Vorlesung, 13.2.2008

Das RSA-Kryptosystem

Alice möchte an Bob eine Nachricht verschicken.
Die Nachricht wird dabei als positive Zahl interpretiert, die in gleichgroße Blöcke unterteilt wird. Das Verfahren sieht dann wie folgt aus.

1. Der Empfänger Bob erzeugt zwei große Primzahlen p, q. Er berechnet $n = p * q$ und eine Zahl e mit $ggT(e, (p-1)(q-1)) = 1$. Das Paar (n, e) bildet den so genannten "public key" und wird an Alice weitergegeben.
   Außerdem berechnet Bob eine Zahl d mit $e * d \equiv 1 \mod (p-1)(q-1)$. p, q, d bleiben geheim.

2. Alice zerlegt die Folge in Blöcke der Länge $k = \lfloor \log_2 n \rfloor$. Jeder Block entspricht dabei einer Zahl $m \geq 0$ und $< n$. Alice verschlüsselt einen Block per

   $$E(m) = m^e \mod n$$

   und sendes diese Zahl an Bob.

3. Bob dechiffriert einen Block M als

   $$D(M) = M^d \mod n$$

   Das ergibt eine Zahl, die auf die Blocklänge mit führenden Nullen aufgefüllt wird. Dieser Block ist dann die gesendete Nachricht.

Beispiel Wir rechnen hier mal dezimal.
$p = 43$ und $q = 59$ sind Primzahlen. Demnach ist $n = pq = 2537$. Wir wählen $e = 13$. Die Blocklänge wählen wir als 4 (da der größtmögliche Block 2525, also ZZ ist, und somit unser n noch größer ist). Die Nachricht ist $STOP$, die wir übermitteln wollen. Dabei interpretieren wir $A = 00$, $B = 01$ usw. Also sind die Blöcke 1819 und 1415.

$$E(1819) = 1819^{13} \mod 2537 = 2081$$
$$E(1415) = 1415^{13} \mod 2537 = 2182$$

Jetzt wollen wir auch was entschlüsseln. Die empfangene Nachricht sei 0149 und 0481. Dazu brauchen wir d. Wir machen eine Nebenrechnung:

$$ggT(2437, 13) \quad 5 = 2436 - 167 * 13$$
$$ggT(13, 5) \quad 3 = 13 - 2 * 5$$
$$ggT(3, 2) \quad 2 = 5 - 1 * 3$$
$$ggT(2, 1) \quad 1 = 3 - 1 * 2$$
$$1 = 3 - (5 - 1 * 3) = 2 * 3 - 5$$
$$2 * (13 - 2 * 5) - 5 = 2 * 13 - 5 * 5$$
$$2 * 13 - 5 * (2436 - 187 * 13) = -5 * 2436 + 937 * 13$$

Damit ist $d = 937$. Jetzt kann man schön entschlüsseln

$$D(0149) = 0194^{937} \mod 2537 = 0714$$
$$D(0481) = 0481^{937} \mod 2537 = 1115$$

Die Nachricht war also $HELP$.

Korrektheit

Wir müssen zeigen, dass $D(E(m)) = m$. Es ist ausreichend zu zeigen, dass

$$D(E(m)) \equiv m$$

Der chinesische Restesatz hilft uns jetzt. Falls wir zeigen, dass $D(E(m)) \equiv m(\mod p)$ und $D(E(m)) \equiv m(\mod q)$, dann folgt $D(E(m)) \equiv m \mod n$.

Fall 1 Sei $m \equiv 0(\mod p)$. D.h. $p|m$ und demnach also auch $p|m^e$. Und wegen $n = pq$ folgt, dass $p|m^e \mod n$. Also haben wir $E(m) \equiv 0(\mod p)$. Analog gilt $D(E(m)) \equiv 0 \mod p$.

Fall 2 Sei $n \not\equiv 0 \mod p$. Dann sagt der kleine Fermat, dass $n^{p-1} \equiv 1 \mod p$. Nach Voraussetzung ist $d*e \equiv 1(\mod (p-1)(q-1))$. D.h. $\exists k : d*e = 1+k(p-1)(q-1)$. Jetzt können wir einfach rechnen. $D(E(m)) = (m^e)^d \mod n = m^{e*d} \mod n$. Und das ist $D(E(m)) \equiv m^{de} \mod p$ usw. auf jedenfall kommt man auf $m \mod p$

□

Effizient $\mod$ rechnen

Wir wollen $d = a^b \mod n$ rechnen.

- ganz grausam: Tatsächlich a^b rechnen.
- besser: Verschachteln und immer $((a \mod n)(a \mod n) \mod n)*\ldots$ rechnen. Dabei hat man dann b arithmetische Operationen.
- super: Wir stellen unser b als Binärzahl dar, und dann macht man irgendwas, hab nicht so recht aufgepasst. Zumindest braucht man hier nur noch $O(\log n)$ arithmetische Operationen