Lineare Algebra

(Mathematik für Physik I)
Prof. Dr. M.M. Wolf (TUM, WS23/24)

Mengen, Abbildungen, Relationen **3**

Gruppen, Ringe, Körper **12**

Vektorräume **19**
Basis-Ergänzungssatz 26
Dimensionsformel für Untervektorräume 27

Lineare Abbildungen **28**
Dimensionssatz 30
Matrizen 33

Lineare Gleichungssysteme **37**
Gauß-Algorithmus 40
Rangsatz 43

Determinanten **45**
Leibniz-Formel 46
Laplace Entwicklungssatz 51

Eigenwerte **54**
Charakteristisches Polynom 58
Trigonalisierbarkeit 62
Jordan-Normalform 65

Skalarprodukte **67**
Normen 67
Cauchy-Schwarz 69
Orthogonalität 70
Gram-Schmidt Verfahren 72

Matrixgruppen & Orientierung **78**

Matrixzerlegungen **80**
Spektralsatz für normale Abbildungen 83
Spektralsatz für reelle normale Matrizen 86
Hauptachsentransformation 87
Singulärwertzerlegung 93
Polarzerlegung 95

Matrixfunktionen **96**
Cayley-Hamilton 98

Vorstellung & Überblick

Organisation: $\underbrace{\overset{\text{Wolf}}{V} + \overset{\text{Dr. Karrasch}}{Z\bar{U}} + \overset{\text{Karrasch + Tutorinnen}}{\bar{U}}}_{\text{moodle page}}$

Gedanken zum Studium:

- eigenes Skript
- Fragen, Fragen, Fragen
- Lerngruppen
- Literatur parallel zur Vorlesung

Lankham, Nachtergaele, Schilling
"Linear Algebra - As an Introduction to Abstract Mathematics"

Jänich
„Lineare Algebra"
„Mathematik 1&2"

Kemper & Reimers
(weiterführend)

Was ist Algebra? Abstraktion von $+, \cdot$ ohne Grenzprozesse
↳ Analysis

Moderne Algebra geht wesentlich auf Emmy Noether (1882 - 1935) zurück.

Was ist Lineare Algebra? Theorie endl.-dim. Linearer Abbildungen
(Lorentz-Trafos, Quanten-Observable, Layer in neuronalen Netzen, etc.)

I. GRUNDBEGRIFFE – Mengen, Abbildungen, Relationen

I.1. Mengen

Eine Menge M ist eine ungeordnete Zusammenfassung von unterschiedlichen (mathematischen) Objekten, den Elementen der Menge.

Bsp.:

- $M = \{1,2,3\} = \{1,3,2\} = \{3,2,1,3\}$

ähnlich: "set" in Python

- $\emptyset = \{\ \}$ Leere Menge
- $\mathbb{N} := \{1,2,3,4,\dots\}$ natürliche Zahlen

„... ist definiert als ..."

$3 \in \mathbb{N}$, $0 \notin \mathbb{N}$ (ACHTUNG! Def.en in Literatur unterschiedlich)

- $\mathbb{Z} := \{0,-1,1,-2,2,\dots\}$ ganze Zahlen
- $\mathbb{R}$ reelle Zahlen

Mengen werden oft über Eigenschaften angegeben:

$B = \{x \in A \mid \mathcal{E}(x)\}$ bedeutet, dass $x \in B$ g.d.w. $x \in A$ und $\mathcal{E}(x)$ zutrifft.

Bsp.:

- $\mathbb{N} = \{x \in \mathbb{Z} \mid x > 0\}$
- $\mathbb{Q} := \{\frac{z}{n} \mid z \in \mathbb{Z} \wedge n \in \mathbb{N}\}$ rationale Zahlen

logisches 'und' -> siehe Analysis Vlsg. (Junktoren & Quantoren)

$A \subseteq B$ bedeutet: A ist Teilmenge von B, d.h.

$$x \in A \Rightarrow x \in B$$

$A \subsetneq B$ bedeutet: A ist 'echte Teilmenge' von B, d.h.

$$A \subseteq B \wedge A \neq B$$

Bem.:

- $A = B \iff (A \subseteq B) \wedge (B \subseteq A)$
- $\emptyset \subseteq A$ gilt immer

Def.: Seien A, B Mengen. Man definiert:

$A \cap B := \{ x \mid x \in A \wedge x \in B \}$ Durchschnitt

(A, B heißen disjunkt wenn $A \cap B = \emptyset$)

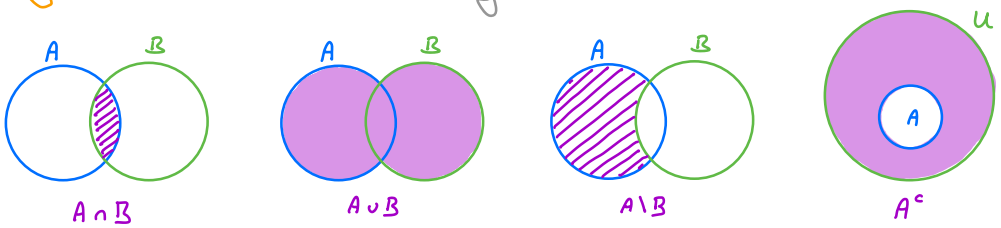
$A \cup B := \{ x \mid x \in A \overset{\text{'oder'}}{\vee} x \in B \}$ Vereinigung

$A \setminus B := \{ x \mid x \in A \wedge x \notin B \}$ Differenz

$\mathcal{P}(A) := \{ M \mid M \subseteq A \}$ Potenzmenge

Sei ferner $A \subseteq U$, dann ist das 'Komplement von A in U' definiert als $A^c := \{ x \in U \mid x \notin A \}$.

Venn Diagramme zur Veranschaulichung:



Das kartesische Produkt von A und B ist

$$A \times B := \{ \underbrace{(a, b)}_{\text{„geordnetes Paar"}} \mid a \in A \wedge b \in B \}$$

Analog definiert man für Mengen $A_1, \dots, A_n$ mit $n \in \mathbb{N}$:

$$A_1 \times A_2 \times \dots \times A_n := \{ \underbrace{(a_1, \dots, a_n)}_{\text{„n-Tupel"}} \mid a_1 \in A_1 \wedge \dots \wedge a_n \in A_n \}$$

ähnlich: "list", "tuple" in Python

Bsp.:

- $\mathbb{N}_0 := \{0\} \cup \mathbb{N} = \{0, 1, 2, \dots\}$
- $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$. Allgemein: $A^n := \overbrace{A \times \dots \times A}^{n \text{ mal}}$
- $M := \{a, c\} \times \{b, a\} = \{(a, b), (a, a), (c, b), (c, a)\}$

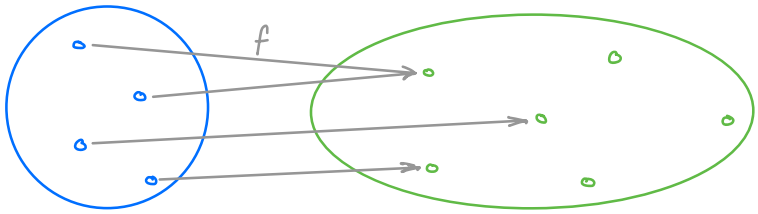
  Da die Paare geordnet sind ist z.B. $(a, c) \notin M$.

I.2. Abbildungen

Eine Abbildung (synonym: Funktion) f von einer Menge A in eine Menge B ordnet jedem $a \in A$ genau ein Element $b \in B$ zu. Man schreibt:

$$f: A \to B \quad \text{oder} \quad A \xrightarrow{f} B \quad \text{bzw.} \quad f: a \mapsto b \quad \text{und} \quad f(a) = b .$$

Dabei heißt A Definitionsbereich und B Wertebereich von f.

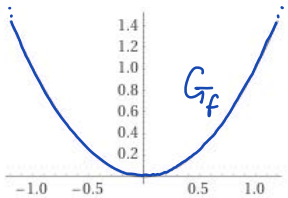


Für die Menge aller Abbildungen von A nach B schreibt man B^A.

Def.: Sei $f: A \to B$ eine Abbildung, $M \subseteq A$, $N \subseteq B$.

- $f(M) := \{ f(x) \mid x \in M \}$ heißt Bild von M unter f.
- $f^{-1}(N) := \{ x \in A \mid f(x) \in N \}$ heißt Urbild von N unter f.
- $f|_M : \begin{cases} M \to B \\ x \mapsto f(x) \end{cases}$ heißt Einschränkung von f auf M.
- Der Graph von f ist definiert als
$$G_f := \{ (x,y) \in A \times B \mid y = f(x) \}$$

Bsp.:
- $id_A : A \to A$, $id_A(x) := x$ Identität
- Für $f: \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$ ist z.B. $f([0,1]) = [0,1]$
$f^{-1}([0,1]) = [-1,1]$ und $f^{-1}(\{1\}) = \{-1,1\}$.
Für $g := f|_{\mathbb{Z}}$ ist $g^{-1}([0,1]) = \{-1,0,1\}$.



Bem.: ∘ Zwei Abbildungen $f: X \to Y$ und $g: X \to Y$ sind gleich

g.d.w. für alle $x \in X$ gilt: $f(x) = g(x)$

- Für $f: X \to Y$ nennt man oft $f(X)$ schlicht das Bild von f.

<u>Def.:</u> Eine Abbildung $f: X \to Y$ heißt

- injektiv, wenn $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$

  (d.h. keine zwei verschiedenen Punkte in X werden auf denselben Punkt in Y abgebildet.)

- surjektiv, wenn $f(X) = Y$

  (d.h. $\overset{\text{'für alle'}}{\forall} y \in Y \ \overset{\text{'existiert'}}{\exists} x \in X: f(x) = y$)

- bijektiv, wenn f surjektiv und injektiv ist.

<u>Bsp.:</u>

- $f: \mathbb{R} \to [0,\infty)$, $f(x) := x^2$ ist surjektiv aber nicht injektiv.

  $f|_{[0,\infty)}$ ist bijektiv. $f|_{\mathbb{N}}: \mathbb{N} \to [0,\infty)$ ist injektiv aber nicht surjektiv.

- Eine Permutation einer endlichen Menge A (z.B. $A = \{1, 2, \dots, n\}$) ist eine bijektive Abb. $\pi: A \to A$.

  Dies lässt sich als 'Umordnung' oder 'Umlabeln' interpretieren.

- Aus jeder injektiven Abbildung $g: X \to Y$ lässt sich über $\tilde{g}: X \to g(X)$, $\tilde{g}(x) := g(x)$ eine bijektive Abbildung (= Bijektion) $\tilde{g}$ konstruieren.

<u>Def.:</u> Für Abbildungen $f: X \to Y$ und $g: Y \to Z$ ist die Komposition (Verknüpfung, Hintereinanderausführung) $g \circ f: X \to Z$ definiert als $g \circ f(x) := g(f(x))$.

Bem.:

- Komposition ist assoziativ, d.h. es gilt $(h\circ g)\circ f = h\circ(g\circ f)$, da $(h\circ g)\circ f(x) = h(g(f(x))) = h\circ(g\circ f)(x) \quad \forall x$.
- Abb.en auf kartesischen Produkten werden oft als Verknüpfungen bezeichnet. Hier: $\circ : Y^X \times Z^Y \to Z^X$, $(f,g) \mapsto g\circ f$

  Aber auch $+ : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $(x,y) \mapsto x+y$

  und $\cdot : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $(x,y) \mapsto x\cdot y$

Satz: Ist $f: X \to Y$ bijektiv, dann existiert eine Abbildung $g: Y \to X$ für die gilt:

$$f\circ g = id_Y \quad \wedge \quad g\circ f = id_X .$$

Diese ist eindeutig, wird als Umkehrabbildung f^{-1} bezeichnet, und es gilt $(f^{-1})^{-1} = f$.

Beweis: Wegen Bijektivität von f existiert zu jedem (surjektiv!) $y \in Y$ genau ein (injektiv!) $x \in X$ mit $f(x) = y$.

Diese eindeutige Zuweisung definiert eine Abbildung $g: Y \to X$, so dass $g(y) = x$ für das x mit $f(x) = y$. Nun gilt

$\forall y \in Y: \quad f\circ g(y) = f(g(y)) = f(x) = y$ also $f\circ g = id_Y$ und

$\forall x \in X: \quad g\circ f(x) = g(f(x)) = g(y) = x$ also $g\circ f = id_X$.

Zur Eindeutigkeit: Ist $\tilde{g}: Y \to X$ eine weitere Abb. mit $f\circ\tilde{g} = id_Y$, dann gilt:

$$\tilde{g} = id_X \circ \tilde{g} = (g\circ f)\circ\tilde{g} \underset{\text{Assoziativität}}{=} g\circ\underbrace{(f\circ\tilde{g})}_{= id_Y} = g$$

Zu $(f^{-1})^{-1} = f$: Per Konstruktion ist f^{-1} bijektiv und erfüllt

$$f^{-1} \circ f = id_X \wedge f \circ f^{-1} = id_Y$$

Wegen Eindeutigkeit der Umkehrabb. ist damit $(f^{-1})^{-1} = f$.

$\square$

„Ende des Beweises"

<u>Bem.:</u>

- Die Umkehrung des Satzes gilt ebenfalls. D.h. die Existenz einer Umkehrabb. impliziert Bijektivität.
  (Hinweis zu einem Beweis: Zeige zunächst, dass $f \circ g = id$ impliziert, dass f surjektiv und g injektiv ist)
- Man zeigt leicht, dass die Verknüpfung zweier Bijektionen wieder bijektiv ist.

<u>Bsp.:</u>

- Für $f: [0,\infty) \to [0,\infty)$, $f(x) := x^2$ ist $f^{-1}(y) = \sqrt{y}$.

<u>Def.:</u> Seien A, B Mengen.

- A und B heißen gleichmächtig und wir schreiben $|A| = |B|$, wenn eine Bijektion $f: A \to B$ existiert.
- A heißt höchstens so mächtig wie B (kurz $|A| \leq |B|$), wenn eine injektive Abbildung $f: A \to B$ existiert.
- B heißt mächtiger als A (kurz $|A| < |B|$), wenn $|A| \leq |B|$ aber nicht $|A| = |B|$ gilt.
- A heißt abzählbar wenn $|A| \leq |\mathbb{N}|$ und überabzählbar sonst.

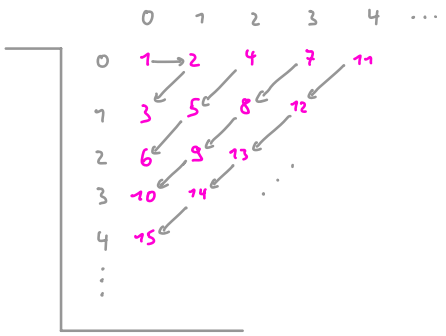
Bem.:

- Hat A $n \in \mathbb{N}$ Elemente, können wir gefahrlos $|A| = n$ schreiben.
- Es gilt: $|A| \leq |B| \wedge |B| \leq |A| \Rightarrow |A| = |B|$.

  Dies ist der Satz von Cantor-Schröder-Bernstein.

Bsp.:

- $|\mathbb{N}_0| = |\mathbb{N}|$, da $f: \mathbb{N}_0 \to \mathbb{N}$, $f(n) := n+1$ bijektiv
- $|\mathbb{N}_0| = |\mathbb{Z}|$, z.B. mit $f: \mathbb{N}_0 \to \mathbb{Z}$, $f(n) := \begin{cases} \frac{n}{2}, & n \text{ gerade} \\ -\frac{n+1}{2}, & n \text{ ungerade} \end{cases}$
- $|\mathbb{N}_0 \times \mathbb{N}_0| = |\mathbb{N}|$:



- $|\mathbb{Q}| = |\mathbb{N}|$. Beweisidee:

  $|\mathbb{Q}| \underset{\text{Def. von } \mathbb{Q}}{\leq} |\mathbb{Z} \times \mathbb{N}| \underset{|\mathbb{Z}| = |\mathbb{N}_0| \wedge |\mathbb{N}| = |\mathbb{N}_0|}{=} |\mathbb{N}_0 \times \mathbb{N}_0| = |\mathbb{N}|$

  Zudem gilt $|\mathbb{N}| \leq |\mathbb{Q}|$ also $|\mathbb{Q}| = |\mathbb{N}|$

- $|\mathbb{N}| < |\mathbb{R}|$. Beweisidee (Cantors Diagonalisierungsargument):

  Sei $z_1, z_2, z_3, \ldots$ eine beliebige 'Aufzählung' reeller Zahlen in $[0,1)$ als Dezimalzahlen. D.h. für $n \in \mathbb{N}$ ist $z_n \in [0,1]$ und $z_{n,k} \in \{0, \ldots, 9\}$ die zugehörige k'te Dezimalstelle. Konstruiere nun eine Dezimalzahl $y \in [0,1)$ mit k'ter Dezimalstelle $y_k := \begin{cases} 4, & \text{wenn } z_{k,k} \neq 4 \\ 5, & \text{wenn } z_{k,k} = 4 \end{cases}$

  Dann gilt $\forall k \in \mathbb{N}: y \neq z_k$. D.h. y kann nicht in der Aufzählung enthalten sein, womit die zugehörige Abb. $\mathbb{N} \to [0,1)$ nicht surjektiv sein kann.

- Ist $|M| = n \in \mathbb{N}_0$, dann gilt $|\mathcal{P}(M)| = 2^n$.

I.3. Relationen

Def.: Eine Relation R zw. zwei Mengen A und B ist eine Teilmenge $R \subseteq A \times B$.

Ist $B = A$ spricht man von einer „Relation auf A".

Eine solche heißt:

(i) reflexiv, wenn $\forall x \in A: \quad (x,x) \in R$

(ii) symmetrisch, wenn $\forall x,y \in A: \quad (x,y) \in R \Rightarrow (y,x) \in R$

(iii) transitiv, wenn $\forall x,y,z \in A:$

$$(x,y) \in R \wedge (y,z) \in R \quad \Rightarrow \quad (x,z) \in R$$

Eine Relation auf A, die (i), (ii) & (iii) erfüllt nennt man Äquivalenzrelation und schreibt dann $x \sim y$ für $(x,y) \in R$.

Bsp.:

- Jede Funktion $f: X \to Y$ kann durch ihren Graphen $G_f \subseteq X \times Y$ als Relation aufgefasst werden.
- Die Zuweisung 'Studierende $\to$ Studienfächer' ist nicht unbedingt eine Funktion (ggfs. nicht eindeutig), kann aber als Relation betrachtet werden.
- „Parallelität" ist eine Äquivalenzrelation auf der Menge der Geraden in der Ebene.
- Auf $\mathbb{Z}$ ist $\quad n \sim m \;:\Leftrightarrow\; 0 = (n-m) \bmod 7$ eine Äquivalenzrelation.
- „Gleichmächtigkeit" ist eine Äquivalenzrelation (auf der Klasse aller Mengen), da die Verknüpfung von Bijektionen wieder bijektiv ist.

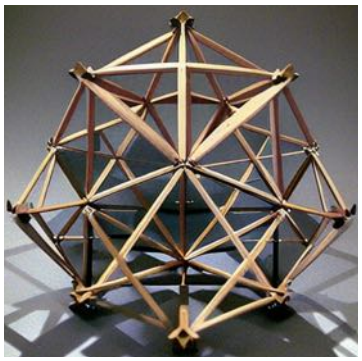
Def.: Sei X eine Menge auf der eine Äquivalenzrelation $\sim$ gegeben ist. Die Äquivalenzklasse $[x]$ eines Elements $x \in X$ bzgl. $\sim$ ist definiert als: $[x] := \{ y \in X \mid y \sim x \}$

Bsp.: Für $X = \mathbb{Z}$ und $x \sim y :\Leftrightarrow 0 = (x-y) \bmod 2$ gibt es zwei disjunkte Äquivalenzklassen: die Mengen der geraden bzw. ungeraden Zahlen:

$$[1] = [3] = [5] \ldots \quad \text{und} \quad [2] = [4] = [6] \ldots .$$

II. ALGEBRAISCHE GRUNDSTRUKTUREN

(Gruppen, Körper, Ringe)



II.1. GRUPPEN

Math. Formalisierung von 'Symmetrien'.

Fundamental in der Physik. z.B.:

- Symmetrien in Eichfeldtheorien ($SU(3) \times SU(2) \times U(1)$ Standardmodell)
- Noether's Theorem : Symmetrien $\rightarrow$ Erhaltungssätze

Def.: ∘ Ein Paar $(G, *)$ einer nichtleeren Menge G und einer Verknüpfung

$* : G \times G \to \underline{\underline{G}}$ heißt Gruppe, wenn:

Sprich "G ist abgeschlossen unter *"

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

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

(iii) $\forall a \in G \ \exists a^{-1} \in G: \quad a^{-1}*a = e$ (Existenz eines inversen Elements)

neutrales Element aus (ii)

∘ Eine Gruppe $(G,*)$ heißt

- abelsch (synonym: kommutativ) wenn für alle $a,b \in G$ gilt:
$$a*b = b*a.$$
- endlich, wenn G endlich ist. $|G|$ heißt dann die Ordnung der Gruppe.

Bsp.:

∘ Zyklische Gruppe $(\mathbb{Z}_n, +)$ für $n \in \mathbb{N}$, wobei $\mathbb{Z}_n := \{0, 1, \dots, n-1\}$ und '+' Addition mod n ist. $(\mathbb{Z}_n, +)$ ist abelsch mit Ordnung n.

∘ $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{C}, +)$ sind abelsche Gruppen.

∘ S_n bezeichnet die Symmetrische Gruppe der Permutationen von $n \in \mathbb{N}$ Elementen. Diese ist von der Ordnung $n!$ und nicht abelsch für alle $n \geq 3$. z.B.:

$$(a,b,c) \begin{cases} \xrightarrow{1 \leftrightarrow 2} (b,a,c) \xrightarrow{2 \leftrightarrow 3} (b,c,a) \\ \xrightarrow{2 \leftrightarrow 3} (a,c,b) \xrightarrow{1 \leftrightarrow 2} (c,a,b) \end{cases} \quad \neq$$

- $(\mathbb{R}\setminus\{0\}, \cdot)$ und $(\mathbb{Q}\setminus\{0\}, \cdot)$ sowie $(\mathbb{C}\setminus\{0\}, \cdot)$ sind abelsche Gruppen.
- Die Gruppe, die durch die elementaren Operationen $\{V, H, L, R, O, U\}$ des Zauberwürfels generiert wird hat Ordnung $\frac{8!\,3^8\,12!\cdot 2^{12}}{2\cdot 2\cdot 3} \approx 4\cdot 10^{19}$ und ist nicht abelsch (z.B. ist $RO \neq OR$).

<u>Korollar:</u> Sei $(G, *)$ eine Gruppe mit neutralem Element e. Dann gilt:

(meint: „einfache Folgerung")

(1) Für jedes $a \in G$ mit inversem Element a^{-1} gilt:

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

(2) $\forall a \in G: \quad a * e = a$

(3) Es existiert kein weiteres neutrales Element, d.h. wenn $\tilde{e} \in G$ so ist, dass $\forall a \in G: \tilde{e} * a = a$, dann ist $\tilde{e} = e$.

(4) Zu jedem $a \in G$ existiert genau ein inverses Element a^{-1}.

<u>Beweis:</u> (1) Zu a^{-1} muss ebenfalls ein inverses Element $(a^{-1})^{-1} \in G$ existieren.

Damit gilt
$$\begin{aligned} e &\overset{(iii)}{=} (a^{-1})^{-1} * a^{-1} \\ &\overset{(ii)}{=} (a^{-1})^{-1} * (e * a^{-1}) \\ &\overset{(iii)}{=} (a^{-1})^{-1} * ((a^{-1} * a) * a^{-1}) \\ &\overset{(i)}{=} ((a^{-1})^{-1} * a^{-1}) * (a * a^{-1}) = a * a^{-1} \end{aligned}$$

(2) $a * e \overset{(iii)}{=} a * (a^{-1} * a) \overset{(i)}{=} (a * a^{-1}) * a \overset{(1)}{=} e * a \overset{(ii)}{=} a$

(3) $e \underset{\text{Annahme}}{=} \tilde{e} * e \overset{(2)}{=} \tilde{e}$

(4) Angenommen $\tilde{a} * a = e$. Dann gilt

$$\begin{aligned} \tilde{a} &\overset{(2)}{=} \tilde{a} * e \overset{(1)}{=} \tilde{a} * (a * a^{-1}) \\ &\overset{(i)}{=} (\tilde{a} * a) * a^{-1} = e * a^{-1} \overset{(ii)}{=} a^{-1}. \end{aligned}$$

$\square$

Bem.:

- Manche Autor*innen verlangen schon in der Definition einer Gruppe, dass $a * e = e * a = a$ und $a^{-1} * a = a * a^{-1} = e$.
- Ist die Verknüpfung '+' wird das inverse Element mit '$-a$' (anstatt 'a^{-1}') bezeichnet.
- Ist die Verknüpfung aus dem Kontext klar, wird die Gruppe oft durch alleinige Angabe der Menge G spezifiziert.

Korollar: (Kürzungsregeln) Sei $(G, *)$ ein Gruppe und $a, b, c \in G$.

Dann gilt:

(i) $a * b = c * b \Rightarrow a = c$

(ii) $a * b = a * c \Rightarrow b = c$

Beweis: $\Rightarrow$ Übung.

genauer: $*|_{H \times H}$

Def.: Sei $(G, *)$ eine Gruppe und $H \subseteq G$. $(H, *)$ heißt Untergruppe von $(G, *)$, wenn

(i) $e \in H$,

(ii) $a \in H \Rightarrow a^{-1} \in H$,

(iii) $a, b \in H \Rightarrow a * b \in H$.

Bsp.:

- $H = \{e\}$
- $(G, *) = (\mathbb{R}, +)$, $H = \mathbb{Z}$
- Die Gruppe der Drehungen des $\mathbb{R}^3$ um den Ursprung besitzt die Gruppe der Drehungen in der x-y-Ebene als Untergruppe.

II.2. Körper (engl.: "field")

Def.: Eine Menge K zusammen mit zwei Verknüpfungen $*: K \times K \to K$ und $+: K \times K \to K$ ist ein Körper, wenn gilt:

(i) $(K, +)$ ist eine abelsche Gruppe.

(ii) $(K \setminus \{0\}, *)$ ist eine abelsche Gruppe, wobei '0' das neutrale Element von $(K,+)$ bezeichnet.

(iii) Es gilt das Distributivgesetz. D.h. $\forall\, a,b,c \in K$:

$$a * (b + c) = (a * b) + (a * c)$$

Bem.:
- Eine Konsequenz ist, dass in einem Körper K für $a, b \in K$ sowohl $a * x = b$ (falls $a \neq 0$) als auch $a + x = b$ eindeutig gelöst werden können.
- Für das neutrale Element bzgl. '$*$' schreiben wir natürlich '1'.

Korollar: Sei $(K, *, +)$ ein Körper und '0' das bzgl. '+' neutrale Element. Dann gilt für alle $a, b \in K$:

(i) $a * 0 = 0 * a = 0$,

(ii) $a * b = 0 \Rightarrow a = 0 \vee b = 0$.

Beweis: (i)

$$\underbrace{a * 0}_{\substack{\| \\ a * 0 + \underline{0}}} = a * (0 + 0) \qquad \text{da '0' neutral bzgl. '+'}$$

$$= a * 0 + \underline{a * 0} \qquad \text{Distributivität}$$

Also ist $a * 0 = 0$ Kürzungsregel!

(ii) Angenommen $a \neq 0$. Dann $\exists\, a^{-1}$ s.d. $a^{-1} * a = 1$. Damit:

$$b = 1 * b = a^{-1} * \underbrace{(a * b)}_{=0} = a^{-1} * 0 \overset{(i)}{=} 0.$$

Analog falls $b \neq 0$ angenommen wird. $\square$

Bem.: (ii) nennt man Nullteilerfreiheit des Körpers.

Bsp.:

- Mit der üblichen Multiplikation & Addition sind $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ Körper.
- Für $p \in \mathbb{N}$ Primzahl, ist $\mathbb{Z}_p := \{0, \dots, p-1\}$ mit Mult. & Add. mod p ein Körper, der Restklassenkörper modulo p.
  (Anwendung z.B. in Reed-Solomon-Codes zur Fehlerkorrektur)
  Wenn p keine Primzahl ist, ist dies kein Körper! z.B. ist in $\mathbb{Z}_6$ $2 * 3 = 0$ im Widerspruch zur Nullteilerfreiheit.

II.3. Ringe

Def.: Eine Menge R zusammen mit zwei Verknüpfungen $+ : R \times R \to R$ und $* : R \times R \to R$ heißt Ring, wenn:

(i) $(R, +)$ eine abelsche Gruppe ist,

(ii) $*$ assoziativ ist,

(iii) die Distributivgesetze gelten. D.h. $\forall a, b, c \in R$:

$$(a+b) * c = (a * c) + (b * c)$$

$$c * (a+b) = (c * a) + (c * b)$$

Bem.: Ist '$*$' zusätzlich kommutativ, spricht man von einem kommutativen Ring. Besitzt '$*$' ein neutrales Element, spricht man von einem Ring mit 1.

Bsp.:

- Jeder Körper $(\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p)$ sowie $(\mathbb{Z}, *, +)$ ist ein kommutativer Ring mit 1.
- Die Restklasseuringe $\mathbb{Z}_n$ modulo $n \in \mathbb{N}$ sind kommutative Ringe mit 1.

Def.:

- Sei R ein kommutativer Ring, $n \in \mathbb{N}_0$ und $a_0, a_1, \dots, a_n \in R$ mit $a_n \neq 0$. Dann heißt

$$a_n x^n + \dots + a_1 x + a_0 =: \sum_{k=0}^{n} a_k x^k$$

Polynom vom Grad n über R.

- Die Menge aller Polynome über R (beliebigen Grades) bezeichnet man als $R[x]$.

Bsp.:

- $\mathbb{Q}[x]$, $\mathbb{R}[x]$ sind die Mengen aller Polynome mit rationalen bzw. reellen Koeffizienten.

Bem.:

- Durch folgende Verknüpfungen wird $R[x]$ selbst zu einem kommutativen Ring, dem Polynomring. Für $f(x) = \sum_{k=0}^{n} a_k x^k$, $g(x) = \sum_{l=0}^{m} b_l x^l$ definiert man

$$(f+g)(x) := \sum_k (a_k + b_k) x^k$$

ggfs. durch 0 fortgesetzt

$$(f * g)(x) := \sum_{k=0}^{m+n} c_k x^k \quad \text{wobei} \quad c_k := \sum_{\substack{i,j \in \mathbb{N}_0 \\ \text{mit } i+j=k}} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i}$$

- Genau genommen unterscheidet man zw. dem Polynom $f \in R[x]$, welches man mit der Koeffizientenfolge identifiziert, und der Polynomfunktion $R \ni x \mapsto f(x) \in R$. Für $R \in \{\mathbb{C}, \mathbb{R}\}$ kann man diese Unterscheidung meist ignorieren, da es eine Bijektion dazwischen gibt (Koeffizientenfolge!). I.A. können jedoch unterschiedliche Polynome dieselbe Polynomfkt. darstellen (z.B. x und x^2 als Fkt.en $\mathbb{Z}_2 \to \mathbb{Z}_2$).

<u>Def.:</u> Ein Körper K heißt algebraisch abgeschlossen, wenn jedes nicht-konstante Polynom $f \in K[x]$ in K eine Nullstelle besitzt. D.h. wenn für jedes $f \in K[x]$ vom Grad mindestens 1 ein $z \in K$ existiert, so dass $f(z) = 0$.

<u>Bem.:</u> $\mathbb{Q}$ und $\mathbb{R}$ sind nicht algebraisch abgeschlossen, da z.B. x^2+1 in $\mathbb{Q}$ bzw. $\mathbb{R}$ keine Nullstelle besitzt.

<u>Satz:</u> (Fundamentalsatz der Algebra)
$\mathbb{C}$ ist algebraisch abgeschlossen.

<u>Beweis:</u> $\to$ Analysis III.

<u>Bem.:</u> Tatsächlich hat ein Polynom n-ter Ordnung ($n \in \mathbb{N}$) in $\mathbb{C}$ n Nullstellen $z_1, \dots, z_n \in \mathbb{C}$ und kann stets wie folgt faktorisiert werden:

$$\boxed{f(x) = a_n \prod_{k=1}^{n} (x - z_k)}. \quad (*)$$

(z.B.: $x^2+1 = (x+i)\cdot(x-i)$)

Gilt $(*)$ für $f \in \mathbb{K}[x]$ mit $a_n, z_k \in \mathbb{K}$, sagt man: „Das Polynom zerfällt in Linearfaktoren."

III. VEKTORRÄUME

Def.: Sei $\mathbb{K}$ ein Körper. Ein Vektorraum V über $\mathbb{K}$ (kurz: „$\mathbb{K}$-Vektorraum") ist eine Menge V zusammen mit zwei Abbildungen: einer Vektoraddition $+ : V \times V \to V$, $(v,w) \mapsto v+w$ und einer Skalarmultiplikation $\cdot : \mathbb{K} \times V \to V$, $(\lambda, v) \mapsto \lambda v$, für die gilt:

(i) $(V,+)$ ist eine abelsche Gruppe

(ii) $\forall \lambda, \mu \in \mathbb{K} \quad \forall v, w \in V$:

a) $(\lambda\mu) v = \lambda(\mu v)$

b) $1 v = v$

c) $\lambda(v+w) = \lambda v + \lambda w$ und

$(\lambda+\mu) v = \lambda v + \mu v$

Bem.:

- Achtung: '+' und '·' haben hier jeweils zwei unterschiedliche Bedeutungen (welche?)
- Die neutralen Elemente in $(V,+)$ und $(\mathbb{K},+)$ werden üblicherweise beide mit 0 bezeichnet. Ob es sich dabei um den 'Nullvektor' oder die Null als Zahl handelt, muß aus dem Kontext geschlossen werden.
- Die Elemente aus $\mathbb{K}$ nennt man in diesem Zusammenhang Skalare.
- Vektorräume über $\mathbb{R}$ bzw. $\mathbb{C}$ nennt man reell bzw. komplex.

Bsp.:

- $\mathbb{K}^n$ wird zu einem $\mathbb{K}$-Vektorraum mit

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$
$$\lambda (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n)$$

Hier ist $0 = (0, \dots, 0)$ das neutrale Element und $(-x_1, \dots, -x_n)$ das zu $(x_1, \dots, x_n)$ inverse Element.

Es bietet sich die Schreibweise der n-Tupel als Spaltenvektoren an. D.h.

$$\mathbb{K}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; x_1, \dots, x_n \in \mathbb{K} \right\} \quad \text{und}$$

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

- Die Menge $\mathbb{K}^M := \{ f \mid f : M \to \mathbb{K} \}$ der $\mathbb{K}$-wertigen Funktionen über einer bel. Menge M wird zu einem $\mathbb{K}$-Vektorraum

$$\text{mit} \begin{cases} (f+g)(x) := f(x) + g(x) \\ (\lambda f)(x) := \lambda f(x) \end{cases}$$

(Dies nennt man punktweise Addition bzw. Skalarmultiplik.)

- Die Menge der Polynome $\mathbb{K}[x]$ wird mit koeffizientenweisen Verknüpfungen zu einem $\mathbb{K}$-Vektorraum.

- $\mathbb{R}$ ist ein Vektorraum über $\mathbb{Q}$ (mit $+, \cdot$ wie üblich)

Bem.: $\mathbb{C}^{\mathbb{R}}$, $\mathbb{C}^n$, $\mathbb{C}[x]$ sind Vektorräume über $\mathbb{C}$ aber z.B. auch über $\mathbb{R}$.

Def.: Sei V ein $\mathbb{K}$-Vektorraum. Eine nichtleere Teilmenge $U \subseteq V$ heißt Untervektorraum von V, wenn gilt:

(i) $x, y \in U \Rightarrow x + y \in U$

(ii) $\lambda \in \mathbb{K},\ x \in U \Rightarrow \lambda x \in U$

(D.h. U muss „abgeschlossen" sein bzgl. Addition und Skalarmultiplikation)

Bem.: Oft spricht man schlicht von U als Unterraum.

Bsp.:

- Zu jedem Vektorraum V sind V und $\{0\}$ triviale Untervektorräume.
- $\{(x_1, 0) \mid x_1 \in \mathbb{R}\}$ ist Untervektorraum von $\mathbb{R}^2$.
- Für bel. $n \in \mathbb{N}_0$ bildet die Menge $\mathbb{K}_n[x] := \{\sum_{k=0}^{n} a_k x^k \mid a_0, \dots, a_n \in \mathbb{K}\}$ der Polynome vom Grad höchstens n einen Untervektorraum von $\mathbb{K}[x]$.
- Die Lösungsmenge $\{(x, y) \in \mathbb{R}^2 \mid 2x - 4y = 0\}$ ist ein Untervektorraum von $\mathbb{R}^2$.

Mit Blick auf die Definitionen zeigt man leicht:

Korollar: Sei V ein $\mathbb{K}$-Vektorraum und $U \subseteq V$.
Dann sind äquivalent:

(i) U ist ein Untervektorraum von V.

(ii) U ist, mit der auf V definierten Addition und Skalarmultiplikation, selbst ein $\mathbb{K}$-Vektorraum.

Def.: Seien U_1, U_2 zwei Untervektorräume eines $\mathbb{K}$-Vektorraums V.

- Die Vektorraum-Summe ist definiert als:

$$U_1 + U_2 := \{ x+y \mid x \in U_1 \wedge y \in U_2 \} \subseteq V$$

- Ist $W := U_1 + U_2$ so, dass für jedes $v \in W$ eindeutige Elemente $x \in U_1, y \in U_2$ mit $v = x+y$ existieren, nennen wir W die direkte Summe von U_1 und U_2 und schreiben $W = U_1 \oplus U_2$.

Bsp.:

- $U_1 := \{ (x, y, 0) \in \mathbb{R}^3 \mid x, y \in \mathbb{R} \}$

  $U_2 := \{ (0, 0, z) \in \mathbb{R}^3 \mid z \in \mathbb{R} \}$

  $U_3 := \{ (0, w, z) \in \mathbb{R}^3 \mid w, z \in \mathbb{R} \}$

  Dann gilt: $\mathbb{R}^3 = U_1 + U_2 = U_1 + U_3$ und $\mathbb{R}^3 = U_1 \oplus U_2$, $\mathbb{R}^3$ ist aber keine direkte Summe von U_1 und U_3 da z.B.

$$\begin{aligned}(1,1,1) &= (0,0,1) + (1,1,0) \\ &= \underbrace{(0,1,1)}_{\in U_3} + \underbrace{(1,0,0)}_{\in U_1}\end{aligned}$$

  uneindeutig ist.

Korollar: Seien U_1, U_2 zwei Untervektorräume eines $\mathbb{K}$-Vektorraums V.

(i) $U_1 \cap U_2$ ist Untervektorraum von V.

(ii) $U_1 + U_2$ ist Untervektorraum von V.

(iii) $U_1 \cap U_2 = \{0\}$ g.d.w. $U_1 + U_2 = U_1 \oplus U_2$

Beweis: (i) Wegen $0 \in U_1$, $0 \in U_2$ ist $U_1 \cap U_2 \neq \emptyset$. Weiter gilt:

$x, y \in U_1 \cap U_2 \Rightarrow x+y \in U_1 \wedge x+y \in U_2 \Rightarrow x+y \in U_1 \cap U_2$

$x \in U_1 \cap U_2, \lambda \in \mathbb{K} \Rightarrow \lambda x \in U_1 \wedge \lambda x \in U_2 \Rightarrow \lambda x \in U_1 \cap U_2$.

(ii), (iii) $\rightarrow$ Übung. $\square$

Def.: Sei $n \in \mathbb{N}$, V ein $\mathbb{K}$-Vektorraum und $v_1, \dots, v_n \in V$. Dann ist die lineare Hülle von $v_1, \dots, v_n$ definiert als

$$\operatorname{span}(v_1, \dots, v_n) := \left\{ \sum_{i=1}^{n} \lambda_i v_i \;\middle|\; \lambda_1, \dots, \lambda_n \in \mathbb{K} \right\}.$$

Dabei heißt $\sum_{i=1}^{n} \lambda_i v_i$ eine Linearkombination von $v_1, \dots, v_n$.

Korollar: $U := \operatorname{span}(v_1, \dots, v_n)$ ist Untervektorraum von V.

Beweis: U ist nicht leer da z.B. $v_1 \in U$.

U ist abgeschlossen unter Addition:

$$x = \sum_{i=1}^{n} \lambda_i v_i, \quad y_i = \sum_{i=1}^{n} \mu_i v_i \quad \Rightarrow \quad x + y = \sum_{i=1}^{n} (\lambda_i + \mu_i) v_i$$

U ist abgeschlossen unter Skalarmultiplikation:

$$\alpha x = \sum_{i=1}^{n} (\alpha \lambda_i) v_i. \qquad \square$$

Bem.: Analog definiert man für eine bel. nicht-leere Teilmenge $M \subseteq V$:

$$\operatorname{span}(M) := \left\{ v \in V \;\middle|\; \exists n \in \mathbb{N}\ \exists v_1, \dots, v_n \in M\ \exists \lambda_1, \dots, \lambda_n \in \mathbb{K} : v = \sum_{i=1}^{n} \lambda_i v_i \right\}$$

D.h. $\operatorname{span}(M)$ ist die Menge aller (endlichen!) Linearkombinationen.

Zudem definiert man in einem gegebenen Vektraum $\operatorname{span}(\emptyset) := \{0\}$.

Bsp.:

- Für $v_1, v_2, v_3 \in \mathbb{R}^3$ ist 'typischerweise'

  $\operatorname{span}(v_1)$ eine Gerade

  $\operatorname{span}(v_1, v_2)$ eine Ebene

  $\operatorname{span}(v_1, v_2, v_3) = \mathbb{R}^3$

Das 'typischerweise' nimmt an, das folgendes gilt:

Def.: Sei V ein $\mathbb{K}$-Vektorraum und $n \in \mathbb{N}$.

- Vektoren $v_1, \dots, v_n \in V$ heißen linear unabhängig wenn folgende Implikation für $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ gilt:

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_n = 0$$

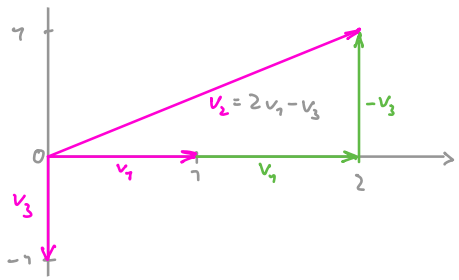
- $M \subseteq V$ heißt linear unabh. wenn jede endliche Teilmenge von M aus linear unabhängigen Vektoren besteht.

Bem.:

- Gilt lin. Unabhängigkeit nicht, spricht man von linearer Abhängigkeit.
- $v_1, \dots, v_n$ sind linear abhängig g.d.w. sich einer der Vektoren als Linearkombination der anderen schreiben lässt.
- Wir betrachten $\emptyset$ als lin. unabh.

Bsp.:

- $v_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 := \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $v_3 := \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ sind linear abhängig, da z.B. $-2v_1 + v_2 + v_3 = 0$.



$\{v_1, v_2\}$ ist linear unabh. da $0 = \lambda_1 v_1 + \lambda_2 v_2 = \begin{pmatrix} \lambda_1 + 2\lambda_2 \\ \lambda_2 \end{pmatrix}$ impliziert, dass $\lambda_2 = 0$ und dann auch $\lambda_1 = 0$ gelten muß.

Def.: Eine Teilmenge B eines $\mathbb{K}$-Vektorraums V heißt **Basis** von V, wenn (i) $\operatorname{span}(B) = V$,

und (ii) B ist lin. unabh.

Bsp.:

- Eine Basis von $\mathbb{K}^n$ ist $\{e_1, \dots, e_n\}$ mit $e_i := (\dots, 0, \underset{\substack{\uparrow \\ \text{i'ter Eintrag}}}{1}, 0, \dots)$. Diese heißt **kanonische Basis** (oder **Standardbasis**).
- Eine Basis von $\mathbb{R}[x]$ ist die Menge aller 'Monome' $B = \{ x^k \mid k \in \mathbb{N}_0 \}$. Beachte: z.B. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \notin \operatorname{span}(B)$

Die Forderung nach lin. Unabh. von B garantiert eine eindeutige Darstellung:

Korollar: Ist $B = \{b_1, \dots, b_n\}$ eine Basis eines $\mathbb{K}$-Vektorraums V, dann lässt sich jedes $v \in V$ auf genau eine Art als $v = \sum_{i=1}^{n} \lambda_i b_i$ mit $\lambda \in \mathbb{K}^n$ darstellen.

Beweis: Angenommen $v = \sum_{i=1}^{n} \lambda_i b_i = \sum_{i=1}^{n} \tilde{\lambda}_i b_i$ mit $\lambda_i, \tilde{\lambda}_i \in \mathbb{K}$.

Dann ist $v - v = \sum_{i=1}^{n} (\lambda_i - \tilde{\lambda}_i) b_i = 0$

und lin. Unabh. impliziert $\lambda = \tilde{\lambda}$. $\square$

Für manche Vektorräume (z.B. $\mathbb{K}^n$, $\mathbb{K}[x]$) ist die Existenz einer Basis offensichtlich. Für andere gilt das nicht (was ist z.B. eine Basis von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum, oder von $\mathbb{R}^{\mathbb{R}}$?) Mit Hilfe des sog. **Zorn'schen Lemmas** lässt sich jedoch zeigen (siehe z.B. [Kemper, Reimers] für einen Beweis):

Satz: (Basis-Ergänzungssatz)

Sei V ein Vektorraum und $M \subseteq V$ linear unabhäng (z.B. $M = \emptyset$).

Dann existiert ein Basis B von V, die M als Teilmenge enthält.

Insbesondere existiert also immer eine Basis.

Satz: Alle Basen eines Vektorraums sind gleichmächtig.

Beweis: siehe z.B. [Kemper, Reimers]. Idee:

Falls eine endliche Basis $B \subseteq V$ existiert, ist die Beweisidee folgende: Ist $\tilde{B}$ eine weitere Basis, lassen sich iterativ einzelne Elemente von $\tilde{B}$ durch einzelne Elemente von B austauschen - unter Beibehaltung der Basiseigenschaft. Die Iteration endet, wenn ganz B in der so neu konstruierten Basis enthalten ist. Dann darf aber kein weiteres Element vorhanden sein, da dies eine Linearkombination der Elemente von B sein würde. Dies zeigt $|\tilde{B}| = |B|$.

Def.: Ein Vektorraum V heißt endlich-dimensional, wenn er eine endliche Basis besitzt. Die Anzahl der Elemente dieser (und damit jeder) Basis nennt man dann die Dimension $\dim(V)$ von V. Existiert keine endliche Basis, nennt man V unendlich-dimensional und schreibt $\dim(V) = \infty$.

Bsp.:

- $\dim(\{0\}) = 0$
- $\dim(\mathbb{R}^n) = n$
- $\dim(\mathbb{R}[x]) = \infty$
- $\dim(\mathbb{R}_n[x]) = n+1$

  wenn, wie üblich, als $\mathbb{R}$-Vektorraum betrachtet
- $\dim(\mathbb{R}) = \begin{cases} 1 & \text{über } \mathbb{R} \\ \infty & \text{über } \mathbb{Q} \end{cases}$

- $\dim(\mathbb{C}) = \begin{cases} 1 & \text{über } \mathbb{C} \\ 2 & \text{über } \mathbb{R} \quad \text{z.B. mit Basis } \{1, i\} \end{cases}$

Satz (Dimensionsformel für Untervektorräume)

Sei V ein Vektorraum mit endlich-dimensionalen Unterräumen U_1 und U_2. Dann ist

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

Bem.: Ganz ähnlich gilt für endliche Mengen A, B:

$$|A \cap B| + |A \cup B| = |A| + |B|$$

Beweis: Wir ergänzen eine Basis $v_1, \dots, v_r$ von $U_1 \cap U_2$ zu einer Basis $v_1, \dots, v_r, w_1, \dots, w_s$ von U_1 und einer Basis $v_1, \dots, v_r, z_1, \dots, z_t$ von U_2. Dann gilt $U_1 + U_2 = \operatorname{span}(B)$ für $B := \{v_1, \dots, v_r, w_1, \dots, w_s, z_1, \dots, z_t\}$. Um lineare Unabhängigkeit zu zeigen, betrachten wir $\underbrace{\sum_{i=1}^{r} \lambda_i v_i + \sum_{j=1}^{s} \mu_j w_j}_{\in U_1} + \underbrace{\sum_{k=1}^{t} \nu_k z_k}_{\in U_2} = 0$.

Also ist $\sum_{k=1}^{t} \nu_k z_k \in U_1 \cap U_2$, so dass $\sum_{k=1}^{t} \nu_k z_k = \sum_{i=1}^{r} \alpha_i v_i$ für geeignete α_i.

Wegen lin. Unabh. der Basiselemente von U_2 folgt $\alpha = 0$ und $\nu = 0$.

— " — — " — von U_1 folgt auch $\lambda = 0$ und $\mu = 0$.

Also ist B eine Basis von $U_1 + U_2$ mit

$r + s + t = (r+s) + (r+t) - r = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$ Elementen. $\square$

Bem.: Damit erhalten wir für endl.-dim Unterräume U_1, U_2:

$$U_1 + U_2 = U_1 \oplus U_2 \iff U_1 \cap U_2 = \{0\}$$

$$\iff \dim(U_1 + U_2) = \dim(U_1) + \dim(U_2)$$

IV. Lineare Abbildungen

Def.: Seien V und W Vektorräume über einem Körper $\mathbb{K}$.

- Eine Abbildung $f: V \to W$ heißt linear oder (Vektorraum-) Homomorphismus, wenn $\forall \lambda \in \mathbb{K}\ \forall x, y \in V$:

$$\boxed{f(x+\lambda y) = f(x) + \lambda f(y)} \iff \begin{cases} f(x+y) = f(x) + f(y) & \forall x,y \\ f(\lambda x) = \lambda f(x) & \forall \lambda\ \forall x \end{cases}$$

- $\mathrm{Hom}(V,W)$ bezeichnet die Menge der Homomorphismen von V nach W
  $\mathrm{End}(V) := \mathrm{Hom}(V,V)$ die der Endomorphismen.
- $f \in \mathrm{Hom}(V,W)$ heißt (Vektorraum-) Isomorphismus wenn f bijektiv ist.
- Existiert ein Isomorphismus zw. V und W, heißen diese isomorph und wir schreiben dafür $V \cong W$.

Bem.:

- Linearität impliziert $f\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i f(x_i)$ $\forall n \in \mathbb{N}, \lambda \in \mathbb{K}^n, x \in V^n$
  Insbesondere: $\underset{\in V}{f(0)} = \underset{\in W}{0}$ und Untervektorräume werden wieder auf solche abgebildet.
- $\mathrm{Hom}(V,W)$ wird selbst zu einem $\mathbb{K}$-Vektorraum mit
  $(f+g)(x) := f(x) + g(x)$ und $(\lambda f)(x) := \lambda f(x)$
- Die Komposition linearer Abbildungen liefert wieder eine lineare Abb.

Bsp.:

- id_V ist linear.
- Für $\varphi \in [0, 2\pi]$ ist $f: \mathbb{R}^2 \to \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos(\varphi) + y\sin(\varphi) \\ -x\sin(\varphi) + y\cos(\varphi) \end{pmatrix}$
  linear und stellt eine Drehung um den Winkel φ dar.
- Die Differentiation $f: \mathbb{R}[x] \to \mathbb{R}[x]$ $f(p) := p'$ ist linear.
- $f: \mathbb{K}[x] \to \mathbb{K}[x]$, $p(x) \mapsto x^3 p(x)$ ist linear.
- $g, f: \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$, $g(x) := 1 + x$ sind nicht linear.
  (z.B. ist $g(x+y) = x+y+1 \neq g(x) + g(y) = x+y+2$)

Satz: Ist $f: V \to W$ ein Isomorphismus, dann ist auch $f^{-1}: W \to V$ ein Isomorphismus.

Beweis: Bijektivität von f^{-1} gilt für jede bijektive Abbildung f. Zur Linearität: Seien $\lambda \in \mathbb{K}$ und $v, w \in W$ mit $f^{-1}(v) =: x$ und $f^{-1}(w) =: y$. Dann gilt:

$$\begin{aligned} f^{-1}(v + \lambda w) &= f^{-1}(f(x) + \lambda f(y)) \\ &\underset{f \text{ linear}}{=} f^{-1}(f(x + \lambda y)) \underset{f^{-1} \circ f = id}{=} x + \lambda y \\ &= f^{-1}(v) + \lambda f^{-1}(w). \end{aligned}$$

$\square$

Satz: Seien V und W $\mathbb{K}$-Vektorräume, $v_1, \dots, v_n \in V$ eine Basis von V und $w_1, \dots, w_n \in W$. Dann existiert genau eine lineare Abbildung $f: V \to W$ für die gilt $\forall i \in \{1, \dots, n\}: f(v_i) = w_i$.

Beweis: Zunächst zur Eindeutigkeit:

Für ein bel. $v \in V$ $\exists! \lambda \in \mathbb{K}^n$ s.d. $v = \sum_{i=1}^{n} \lambda_i v_i$. Wegen der geforderten Linearität ist dann $f(v) = f(\sum_i \lambda_i v_i) = \sum_i \lambda_i f(v_i) = \sum_i \lambda_i w_i$ $(*)$ eindeutig bestimmt.

Existenz von f wird klar, wenn wir $(*)$ als Definition von f verwenden. f ist dann linear und es gilt $f(v_i) = w_i$.

$\square$

Def.: Für $f \in \mathrm{Hom}(V, W)$ definiert man

- den Kern $\ker(f) := \{ x \in V \mid f(x) = 0 \} = f^{-1}(\{0\})$,
- das Bild $f(V) := \{ y \in W \mid \exists x \in V: y = f(x) \}$,
- den Rang $\mathrm{rg}(f) := \dim(f(V))$.

engl.: 'kernel/null space', 'range', 'rank'

Bem.: Wegen Linearität von f, ist $\ker(f)$ Untervektorraum von V und das Bild $f(V)$ Untervektorraum von W.

Satz: Seien V und W $\mathbb{K}$-Vektorräume, $\dim(V) < \infty$ und $f \in \mathrm{Hom}(V, W)$.

Dann gilt:
$$\boxed{\dim(V) = \dim(\ker(f)) + \dim(f(V))}$$

Beweis: Sei $u_1, \dots, u_m$ eine Basis von $\ker(f)$, die wir zu einer Basis $u_1, \dots, u_m, v_1, \dots, v_n$ von V ergänzen. Dann ist $\dim(V) = n + \dim(\ker(f))$.

Wir zeigen nun, dass $f(v_1), \dots, f(v_n)$ Basis von $f(V)$ ist.

Jedes $v \in V$ können wir schreiben als $v = \left(\sum_{i=1}^{m} a_i u_i\right) + \left(\sum_{j=1}^{n} b_j v_j\right)$.

Wegen $u_i \in \ker(f)$ ist also $f(v) = \sum_{j=1}^{n} b_j f(v_j)$, d.h. $f(V) = \mathrm{span}(f(v_1), \dots, f(v_n))$.

Um lineare Unabh. z.z., betrachten wir

$$0 = \sum_{j=1}^{n} \lambda_j f(v_j) \underset{\substack{\uparrow \\ \text{Linearität}}}{=} f\left(\sum_{j=1}^{n} \lambda_j v_j\right)$$

Dies bedeutet, dass $\underbrace{\sum_{j=1}^{n} \lambda_j v_j}_{} \in \ker(f)$

also $= \sum_{i=1}^{m} \mu_i u_i$ für geeignete $\mu_i \in \mathbb{K}$.

Wegen lin. Unabh. von $u_1, \dots, u_m, v_1, \dots, v_n$ muß nun $\lambda_j = 0 = \mu_i \ \forall i, j$. $\square$

Bem.:
- Wenn $\dim(V) = \infty$, gilt dies sinngemäß. D.h.
  $\dim(V) = \infty \Rightarrow \dim(\ker(f)) = \infty \vee \dim(f(V)) = \infty$.
- Für das Bild $f(V)$ schreiben wir zuweilen auch $\mathrm{Bild}(f)$.

Bsp.:
- Für $f: \mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (x, y, 0)$ ist
  $\ker(f) = \{(0, 0, z) \mid z \in \mathbb{R}\}$. Also $\underset{\substack{\shortparallel \\ 3}}{\dim V} = \underset{\substack{\shortparallel \\ 1}}{\dim(\ker(f))} + \underset{\substack{\shortparallel \\ 2}}{\dim(f(V))}$.

Korollar: Sei $f: V \to W$ eine lineare Abb. zw. zwei $\mathbb{R}$-Vektorräumen.

(i) f injektiv $\Leftrightarrow$ $\dim(\ker(f)) = 0$ $\Leftrightarrow$ $\ker(f) = \{0\}$

(ii) — " — $\Rightarrow$ $\dim(V) \leq \dim(W)$

(iii) f surjektiv $\Rightarrow$ $\dim(W) \leq \dim(V)$

Beweis: (i) '$\Rightarrow$': Ist f injektiv und $v \in \ker(f)$, dann gilt

$f(v) = 0 = f(0)$ und wegen Injektivität $v = 0$.

'$\Leftarrow$': Ist $\ker(f) = \{0\}$ und $f(u) = f(v)$, dann gilt

$f(u-v) = f(u) - f(v) = 0$ also $u = v$.

(ii) Ist f injektiv, folgt aus (i), dass $\dim(\ker(f)) = 0$.

Aus der Dimensionsformel des Satzes folgt dann

$$\dim(V) \underset{\ker(f)=\{0\}}{=} \dim(f(V)) \underset{f(V) \text{ Unterraum von } W}{\leq} \dim(W).$$

(iii) Es gilt:

$$\dim(V) = \dim(\underbrace{f(V)}_{=W}) + \dim(\ker(f)) \overset{f \text{ surjektiv}}{=} \dim(W) + \dim(\ker(f)) \geq \dim(W). \qquad \square$$

Korollar: Seien V, W zwei $\mathbb{R}$-Vektorräume. Dann gilt:

(i) $V \cong W \Rightarrow \dim(V) = \dim(W)$

(ii) $\dim(V) = \dim(W) < \infty \Rightarrow V \cong W$

Beweis: (i) $V \cong W$ bedeutet, es existiert eine lineare, bijektive Abb. $V \xrightarrow{f} W$.

Nach dem vorangehenden Korollar ist dann

$$\dim(V) \overset{f \text{ inj.}}{\leq} \dim(W) \overset{f \text{ surj.}}{\leq} \dim V.$$

(ii) Sind $(v_i)_{i=1}^n$, $(w_i)_{i=1}^n$ Basen von V bzw. W, dann existiert eine lin. Abb. $f: V \to W$ mit $f(v_i) = w_i$. Für diese gilt $f(V) = W$ (Surjektivität!) und damit

$$\dim(\ker(f)) = \dim(V) - \dim(f(V)) = \dim(V) - \dim(W) \overset{\text{Annahme}}{=} 0$$

also auch Injektivität. $\square$

<u>Bem.:</u> D.h. es gilt insbesondere: $\dim(V) = n \in \mathbb{N} \Rightarrow V \cong \mathbb{K}^n$

Einen zugehörigen Isomorphismus erhalten wir aus einer bel. Basis:

<u>Def.:</u> Sei $B := (v_1, \dots, v_n)$ eine Basis eines $\mathbb{K}$-Vektorraums V, dann ist der kanonische Basisisomorphismus $\phi_B : \mathbb{K}^n \to V$ definiert als

$$\phi_B(\lambda) := \sum_{i=1}^n \lambda_i v_i .$$

<u>Korollar:</u> Sei V ein endlich-dim. Vektorraum und $f: V \to V$ linear.

1. Folgende Eigenschaften sind äquivalent:
   (i) f ist bijektiv
   (ii) f ist injektiv
   (iii) f ist surjektiv
2. Sind $(b_1, \dots, b_n)$ und $(c_1, \dots, c_n)$ zwei Basen von V und gilt $f(b_i) = c_i \ \forall i$, dann ist f ein Isomorphismus.

<u>Beweis:</u> $\longrightarrow$ Übung.

<u>Bem.:</u> zu 2.: Umgekehrt bildet ein Isomorphismus Basen immer auf Basen ab.

Def.: ∘ Für $m, n \in \mathbb{N}$ ist eine $m \times n$ Matrix A über $\mathbb{K}$ eine 'rechteckige Anordnung'

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix} = (A_{i,j})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = (A_{i,j})_{i,j}$$

Kurzschreibweisen:

mit m Zeilen und n Spalten und Einträgen $\overset{A_{ij}}{\overset{\shortparallel}{A_{ij}}} \in \mathbb{K}$.

Zeilen- Spalten-Index

∘ $\mathbb{K}^{m \times n}$ bezeichnet die Menge aller $m \times n$ Matrizen über $\mathbb{K}$.

∘ Die Transponierte $A^T \in \mathbb{K}^{n \times m}$ einer Matrix $A \in \mathbb{K}^{m \times n}$ ist diejenige Matrix für die $(A^T)_{ij} = A_{ji}$ gilt.

∘ Eine Matrix $A \in \mathbb{K}^{m \times n}$ heißt quadratisch, wenn $m = n$ und symmetrisch, wenn $A^T = A$ gilt.

Bem.: ∘ Ein Spaltenvektor ist damit eine $m \times 1$ Matrix und ein Zeilenvektor eine $1 \times n$ Matrix.

∘ $\mathbb{K}^{m \times n}$ ist $\mathbb{K}$-Vektorraum der Dimension $m \cdot n$ mit $(A+B)_{ij} := A_{ij} + B_{ij}$ und $(\lambda A)_{ij} := \lambda A_{ij}$.

Satz: Seien V, W $\mathbb{K}$-Vektorräume und $b_1, \dots, b_n \in V$ sowie $c_1, \dots, c_m \in W$ Basen von V bzw. W.

Für jede Matrix $A \in \mathbb{K}^{m \times n}$ existiert genau ein $f \in \mathrm{Hom}(V, W)$ mit

$$f : b_j \mapsto \sum_{i=1}^{m} A_{ij} c_i \quad \forall j$$

und umgekehrt existiert zu jedem $f \in \mathrm{Hom}(V,W)$ genau eine Matrix $A \in \mathbb{K}^{m \times n}$ für die dies erfüllt ist.

<u>Beweis:</u> Den ersten Teil hatten wir bereits gezeigt.
Sei nun $f \in \mathrm{Hom}(V,W)$ gegeben, dann lässt sich jedes $f(b_j) \in W$ eindeutig als Linearkombination der Basisvektoren c_i schreiben. Also als $f(b_j) = \sum_{i=1}^{m} A_{ij} c_i$. $\square$

<u>Bem.:</u>
- A heißt Darstellungsmatrix von f bzgl. der Basen $(b_1, \dots, b_n) =: B$ und $(c_1, \dots, c_m) =: C$. Wo es sinnvoll ist, die Abhängigkeiten in die Notation aufzunehmen, schreiben wir für diese Darstellungsmatrix $M_{C,B}(f)$. Die Basen fassen wir dabei stets als Tupel auf, da die Reihenfolge entscheidend ist.

- Die Abbildung $f \mapsto M_{C,B}(f)$ ist ein Isomorphismus zw. $\mathrm{Hom}(V,W)$ und $\mathbb{K}^{m \times n}$ der von den gewählten Basen abhängt.

- Für $v = \sum_{j=1}^{n} \lambda_j b_j$ gilt $f(v) = \sum_{i=1}^{m} \mu_i c_i$ mit

$$\boxed{\mu_i = \sum_{j=1}^{n} A_{ij} \lambda_j \overset{*}{=:} (A\lambda)_i} \qquad \text{kurz: } \mu = A\lambda.$$

D.h. die durch $*$ definierte lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^m$, $\lambda \mapsto A\lambda$ (für die man üblicherweise ebenfalls A schreibt) erfüllt

$$\boxed{\Phi_C \circ A \circ \Phi_B^{-1} = f} \qquad \text{also} \qquad M_{C,B}(f) = \Phi_C^{-1} \circ f \circ \Phi_B$$

wenn Φ_B, Φ_C die kanonischen Basisisomorphismen sind.

Satz: Seien $F \in \mathbb{K}^{m \times n}$ und $G \in \mathbb{K}^{n \times r}$ Darstellungsmatrizen von $f \in \mathrm{Hom}(V, W)$ bzgl. B_V, B_W und $g \in \mathrm{Hom}(U, V)$ bzgl. B_U, B_V. Dann erfüllt die Darstellungsmatrix $H \in \mathbb{K}^{m \times r}$ von $f \circ g$ bzgl. B_U, B_W:

$$\boxed{H_{ik} = \sum_{j=1}^{n} F_{ij} G_{jk}}$$

Dies definiert das Matrixprodukt von F und G und man schreibt $H = FG$.

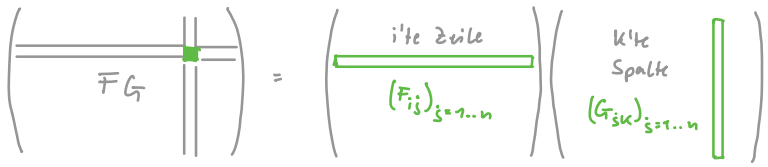
Beweis: Mit $B_U = (u_1, \dots, u_r)$, $B_V = (v_1, \dots, v_n)$, $B_W = (w_1, \dots, w_m)$ gilt $g(u_k) = \sum_{j=1}^{n} G_{jk} v_j$ und $f(v_j) = \sum_{i=1}^{m} F_{ij} w_i$.

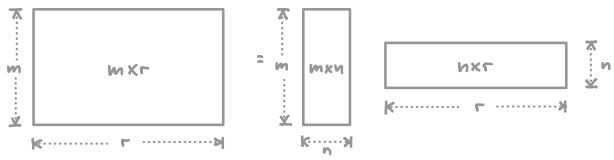
Damit ist
$$\underbrace{f \circ g(u_k)}_{= \sum_{j=1}^{n} G_{jk} f(v_j)} = \sum_{i=1}^{m} \boxed{H_{ik}} w_i \quad \text{per Def. der Darstellungsmatrix von } f \circ g$$
$$= \sum_{i=1}^{m} \boxed{\sum_{j=1}^{n} F_{ij} G_{jk}} w_i$$

Die Eindeutigkeit der Darstellungsmatrix beweist dann die behauptete Gleichheit. $\square$

Bem.:

- Der (i,k)'te Eintrag von FG ergibt sich aus der i'ten Zeile von F und der k'ten Spalte von G:



- 

- Das Matrixprodukt $\mathbb{K}^{m\times n} \times \mathbb{K}^{n\times r} \longrightarrow \mathbb{K}^{m\times r}$, $(F, G) \mapsto FG$ ist assoziativ und distributiv (d.h. z.B. $G(F+H) = GF + GH$) aber i.a. nicht kommutativ. z.B.

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{aber} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

- Mit dem Matrixprodukt wird $\mathbb{K}^{m\times n}$ zu einem Ring.

<u>Bsp.:</u>

- Betrachten wir $A \in \mathbb{K}^{m\times n}$ als Abbildung $\mathbb{K}^n \to \mathbb{K}^m$, $\lambda \mapsto A\lambda$ ist dies schlicht die Matrixmultiplikation einer $m\times n$ Matrix A mit einer $n\times 1$ Matrix (Spaltenvektor!) λ.

  In diesem Fall gilt: Spalten sind die Bilder der Standard-basisvektoren

$$\text{z.B.:} \quad \begin{pmatrix} 3 & 2 & 1 & 8 \\ 4 & 7 & 9 & 2 \\ 8 & 0 & 5 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 7 \\ 0 \end{pmatrix}$$

- A heißt invertierbar, wenn $\lambda \mapsto A\lambda$ ein Isomorphismus ist.

- Die Darstellungsmatrix von $\mathrm{id}_{\mathbb{K}^n}$ bzgl. der Standardbasis ist

$$I \text{ mit } I_{ij} = \underbrace{\delta_{ij}}_{\text{'Kronecker Delta'}} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

- Komplexe Zahlen als reelle 2×2 Matrizen:

  Für $\varphi: \mathbb{C} \to G := \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}$

  $\varphi(a+ib) := \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ gilt:

  (i) φ ist ein Vektorraumisomorphismus (über $\mathbb{R}$)

  (ii) $\varphi(\underbrace{z_1 \cdot z_2}_{\text{Produkt in } \mathbb{C}}) = \varphi(z_1)\underbrace{\varphi(z_2)}_{\text{Matrixprodukt}} \quad \forall z_1, z_2 \in \mathbb{C}$

  Beweis: $\longrightarrow$ Übung.

V. Lineare Gleichungssysteme

Bsp.: Gesucht ist die Schnittmenge dreier Ebenen im $\mathbb{R}^3$, die durch folgende drei Gleichungen gegeben sind:

$$\begin{aligned} x_1 + x_2 + 10x_3 &= -6 \\ 2x_1 - x_2 + x_3 &= -2 \\ 6x_1 - x_2 - x_3 &= 4 \end{aligned} \qquad \text{d.h. } Ax = b \text{ mit } A = \begin{pmatrix} 1 & 1 & 10 \\ 2 & -1 & 1 \\ 6 & -1 & -1 \end{pmatrix},\ b = \begin{pmatrix} -6 \\ -2 \\ 4 \end{pmatrix}.$$

Wäre A in Zeilenstufenform, ließe sich dies einfach lösen. Für $\tilde{A}x = b$ mit $\tilde{A} = \begin{pmatrix} 1 & 1 & 10 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$ erhielten wir z.B. $x_3 = -4$, was wir einsetzen können in $-x_2 + x_3 = -2$ um $x_2 = -2$ zu erhalten, was wir wiederum einsetzen können in $x_1 + x_2 + 10x_3 = -6$ um schließlich $x_1 = 36$ zu erhalten.

Ziel ist, diese 'Zeilenstufenform' zu erreichen, ohne die Lösungsmenge $L := \{ x \in \mathbb{R}^3 \mid Ax = b \}$ zu ändern. Diese wird erhalten bei:

(i) Vertauschen von Zeilen (des Gleichungssystems)

(ii) Multiplizieren einer Zeile mit einem Skalar $\neq 0$.

(iii) Addieren (eines Vielfachen) einer Zeile zu einer anderen.

Anwendung auf die Erweiterte Koeffizientenmatrix $(A \,\vdots\, b) \in \mathbb{R}^{3\times 4}$ ergibt:

$$\left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ 2 & -1 & 1 & -2 \\ 6 & -1 & -1 & 4 \end{array}\right) \begin{array}{l} \\ -2\cdot 1.\text{Zeile} \\ -6\cdot 1.\text{Zeile} \end{array} \overset{(iii)}{\longmapsto} \left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ 0 & -3 & -19 & 10 \\ 0 & -7 & -61 & 40 \end{array}\right) \begin{array}{l} \\ \\ \cdot 3 \end{array} \overset{(ii)}{\longmapsto} \left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ 0 & -3 & -19 & 10 \\ 0 & -21 & -183 & 120 \end{array}\right) \begin{array}{l} \\ \\ -7\cdot 2.\text{Zeile} \end{array} \overset{(iii)}{\longmapsto}$$

Das Gleichungssystem $\hat{A}x = \hat{b}$ $\qquad (\hat{A} \,\vdots\, \hat{b}) := \left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ 0 & -3 & -19 & 10 \\ 0 & 0 & -50 & 50 \end{array}\right)$

hat dieselbe Lösungsmenge wie $Ax = b$.

Durch sukzessives Einsetzen (unten beginnend) lässt sich $\hat{A}x = \hat{b}$ lösen, s.d. $L := \{ x \mid Ax = b \} = \{ x \mid \hat{A}x = \hat{b} \} = \{ (1,3,-1) \}$. D.h. die Ebenen schneiden sich in einem Punkt $(1,3,-1)$. Es wäre auch $L = \emptyset$ (für parallel verschobene Ebenen) oder $|L| = \infty$ (für Ebenen mit gemeinsamer Geraden) möglich gewesen.

Def.: ∘ Ein lineares Gleichungssystem (LGS) von m linearen Gleichungen in n Variablen/Unbekannten $x_1, \dots, x_n \in \mathbb{K}$ ist von der Form $\boxed{Ax = b}$ mit Koeffizientenmatrix $A \in \mathbb{K}^{m \times n}$ und $b \in \mathbb{K}^m$. Ein LGS heißt homogen, wenn $b = 0$ und inhomogen sonst. $\{x \in \mathbb{K}^n \mid Ax = b\}$ heißt die Lösungsmenge des LGS und $(A \vdots b) \in \mathbb{K}^{m \times (n+1)}$ erweiterte Koeffizientenmatrix.

∘ Eine Matrix $A \in \mathbb{K}^{m \times n}$ ist in Zeilenstufenform, wenn folgendes gilt:

(1) Beginnt eine Zeile mit Nullen, stehen unter diesen ausschließlich wieder Nullen.

(2) Unter dem ersten von Null verschiedenen Element jeder Zeile stehen ausschließlich Nullen.

∘ Wir nennen eine Abbildung $\mathbb{K}^{m \times n} \to \mathbb{K}^{m \times n}$, $A \mapsto A'$ elementare Zeilenoperation, wenn A' aus A durch eine der folgenden Operationen hervorgeht:

(i) Vertauschung zweier Zeilen,

(ii) Multiplizieren einer Zeile mit einem Skalar $\lambda \in \mathbb{K} \setminus \{0\}$,

(iii) Addition des λ-fachen ($\lambda \in \mathbb{K}$) einer Zeile zu einer anderen.

Zeilenstufenform:

$$\begin{pmatrix} 0 \dots 0 & A_{1,j_1} & & & \\ 0 \dots 0 & 0 \dots 0 & A_{2,j_2} & & \\ 0 \dots \dots & \dots & 0 \dots 0 & A_{3,j_3} & \\ \vdots & & & 0 & \ddots \\ 0 & & & \vdots & \\ & & & 0 & \end{pmatrix}$$

Bsp.: $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ aber nicht $\begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$

Satz: Sei $A \in \mathbb{K}^{m \times n}$, $b \in \mathbb{K}^m$ und $L := \{x \in \mathbb{K}^n \mid Ax = b\}$ die Lösungsmenge des zugehörigen linearen Gleichungssystems.

(i) $L \neq \emptyset \iff rg(A) = rg(\underbrace{A \vdots b}_{\text{erw. Koeffizientenmatrix}})$

(ii) $x_0 \in L \Rightarrow L = \{x_0 + x \mid x \in ker(A)\}$

D.h. die Lösungsmenge des inhomogenen LGS besteht aus einer inhomogenen Lösung x_0 plus der Lösungsmenge des homogenen LGS.

(iii) Existiert eine Lösung, dann gilt:

$|L| = 1 \iff rg(A) = n \iff ker(A) = \{0\}$

Beweis: (i) Es gilt: $L \neq \emptyset \iff b \in Bild(A)$

Wir verwenden, dass das Bild der linearen Abb. einer Matrix die lineare Hülle der Spaltenvektoren ist. Damit gilt

$Bild(A \vdots b) = Bild(A) + span(b)$

'$\Rightarrow$': Ist $L \neq \emptyset$, also $b \in Bild(A)$, ist damit $rg(A \vdots b) = rg(A)$

'$\Leftarrow$': Ist $rg(A \vdots b) = rg(A)$, dann gilt wegen der Unterraminklusion $Bild(A \vdots b) \supseteq Bild(A)$, dass mit der Dimension auch die Räume gleich sein müssen. Also $b \in Bild(A)$ und damit $L \neq \emptyset$.

(ii) Sicher ist $L \supseteq \{x_0 + x \mid x \in ker(A)\}$, da

$A(x_0 + x) = \underbrace{Ax_0}_{=b} + \underbrace{Ax}_{=0}$.

Umgekehrt, ist $y \in L$, dann gilt $A(y - x_0) = 0$, also $y = x_0 + x$ mit $x \in ker(A)$.

(iii) Existiert eine Lösung, dann folgt aus (ii), dass die Anzahl der freien Parameter in L gleich ist $\dim(\ker(A)) = n - \mathrm{rg}(A)$. Dies ist Null (entsprechend $|L| = 1$), g.d.w. $\mathrm{rg}(A) = n$ ist.

$\square$

<u>Bem.:</u> Die Lösungsmenge eines homogenen LGS ist als stets ein (Unter-) Vektorraum. Die eines inhomogenen LGS ist, wenn nicht leer, ein affiner Unterraum, d.h. ein um einen konstanten Vektor verschobener Vektorraum.

Lösen lässt sich ein LGS $Ax = b$, $A \in \mathbb{K}^{m \times n}$ mit Hilfe des Gaußschen Eliminationsverfahrens (synonym: Gauß-Algorithmus):

<u>kurz:</u> Umformung von $(A \,\vdots\, b)$ mittels elementarer Zeilenoperationen auf Zeilenstufenform und anschließend sukzessives Lösen des resultierenden LGS durch Einsetzen von unten.

<u>Ausführlicher:</u>

0) Beginne bei $k = 1$

1) Suche kleinstes i_k mit $A_{j i_k} \neq 0$ für $j \geq k$. Fertig, wenn keines existiert.

2) Ist $A_{k i_k} = 0$, vertausche Gleichungen (d.h. Zeilen von $(A \,\vdots\, b)$) mit Index $\geq k$, so dass $A_{k i_k} \neq 0$.

3) Ersetze alle Gleichungen mit Index $l > k$ durch

$$\sum_{i=1}^{n} \left(A_{li} - \frac{A_{l i_k}}{A_{k i_k}} A_{ki} \right) x_i = b_l - \frac{A_{l i_k}}{A_{k i_k}} b_k$$

(D.h. das $\frac{A_{l i_k}}{A_{k i_k}}$-fache der k-ten Gleichung wird von der l'ten Gleichung abgezogen.)

4) Fertig, wenn $k = m$. Sonst weiter bei 1) mit $k \leftarrow k+1$.

Bem.: ∘ Struktur der Lösungsmenge:

Ist $(\hat{A} \,\vdots\, \hat{b})$ die aus $(A \,\vdots\, b)$ durch elem. Zeilenoperationen erhaltene Zeilenstufenform und $r \in \{0, \dots, m\}$ die Anzahl der von Null verschiedenen Zeilen in $\hat{A}$, dann sind folgende Fälle möglich:

(i) $r < m$ und $\{\hat{b}_{r+1}, \dots, \hat{b}_m\} \neq \{0\}$: Die Lösungsmenge is leer, da eine Gleichung der Form $0 \cdot x_k \neq 0$ auftaucht.

(ii) $r = m$ oder $\hat{b}_{r+1} = \dots = \hat{b}_m = 0$: Die Lösungsmenge ist nicht leer. (D.h. das LGS ist 'lösbar').

a) $r = n$: es gibt eine eindeutige Lösung

b) $r < n$: es gibt $n-r$ 'freie Parameter'.

∘ Jede elementare Zeilenoperation $A \mapsto \hat{A}$ lässt sich durch $\hat{A} = EA$ realisieren, wobei $E \in \mathbb{K}^{m \times m}$ invertierbar ist. Z.B.:

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & 1 & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}$$

k'te $\quad$ l'te Spalte

vertauscht die Zeilen k & l

$$E = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \lambda & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

multipliziert die k'te Zeile mit λ

$$E = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & 1 & & \\ & & \lambda & \cdots & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

$\leftarrow k$'te Zeile

l'te Spalte

addiert das λ-fache der l'ten Zeile zur k'ten Zeile

Damit führt der Gaußsche Algorithmus auf $\hat{A} = BA$, $\hat{b} = Bb$ wobei $B = E_k \cdots E_2 E_1 \in \mathbb{K}^{m \times m}$ invertierbar ist.

Die durch B dargestellte lineare Abb. beschreibt einen Basiswechsel im Bildraum.

Satz: Für $A \in \mathbb{K}^{m \times n}$ ist $Ax = b$ genau dann für <u>jedes</u> $b \in \mathbb{K}^m$ lösbar, wenn $rg(A) = m$.

Beweis: '$\Rightarrow$': Lösbarkeit für jedes $b \in \mathbb{K}^m$ impliziert, dass $Bild(A) = \mathbb{K}^m$ also $rg(A) = m$.

'$\Leftarrow$': Ist $m = rg(A) = \dim(Bild(A))$, muß in der Inklusion $Bild(A) \subseteq \mathbb{K}^m$ Gleichheit gelten. $\square$

Bem.: Insbesondere ist $Ax = b$ für jedes b lösbar, wenn $A \in \mathbb{K}^{n \times n}$ ein Isomorphismus ist, was $rg(A) = n$ bedeutet.
Die Umkehrabb. lässt sich dann wie folgt bestimmen:

Satz: Ist $A \in \mathbb{K}^{n \times n}$ und lässt sich die Matrix $(A \,\vdots\, I_n) \in \mathbb{K}^{n \times 2n}$ durch eine Abfolge elementarer Zeilenumformungen auf die Form $(I_n \,\vdots\, \tilde{A})$ bringen, dann gilt $\tilde{A} A = I_n$, d.h. $\tilde{A} = A^{-1}$.

Beweis: Es gilt $(I_n \,\vdots\, \tilde{A}) = B(A \,\vdots\, I_n) = (\underbrace{BA}_{I_n} \,\vdots\, \underbrace{BI_n}_{\tilde{A}})$.
Also $B = \tilde{A}$ und damit $\tilde{A} A = I_n$ $\square$

Bem.:
- A^{-1} heißt die zu A Inverse Matrix und es gilt $A^{-1}A = AA^{-1} = I_n$.
- Jede invertierbare Matrix lässt sich auf diese Weise invertieren.

Bsp.:

$$\left(\underbrace{\begin{matrix} 2 & 1 \\ 1 & 1 \end{matrix}}_{A} \,\vdots\, \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right) \xrightarrow[\text{vertauschen}]{} \left(\begin{matrix} 1 & 1 \\ 2 & 1 \end{matrix} \,\vdots\, \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}\right) \xrightarrow[-2 \cdot \text{Zeile 1}]{} \left(\begin{matrix} 1 & 1 \\ 0 & -1 \end{matrix} \,\vdots\, \begin{matrix} 0 & 1 \\ 1 & -2 \end{matrix}\right) \xrightarrow{+\text{Zeile 2}}$$

$$\left(\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix} \,\vdots\, \begin{matrix} 1 & -1 \\ 1 & -2 \end{matrix}\right) \xrightarrow[\cdot(-1)]{} \left(\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \,\vdots\, \underbrace{\begin{matrix} 1 & -1 \\ -1 & 2 \end{matrix}}_{A^{-1}}\right)$$

Kontrolle: $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ✓

Korollar: Sei $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $a,b,c,d \in \mathbb{K}$ und $\Delta := (ad - bc) \neq 0$.

Dann gilt $\boxed{A^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$.

Beweis: Betrachte $a \neq 0$ ($c \neq 0$ geht analog).

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} a & 0 & \frac{ad}{\Delta} & -\frac{ab}{\Delta} \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{array}\right)$$

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{\Delta} & -\frac{b}{\Delta} \\ 0 & 1 & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{array}\right) \longleftarrow$$

$\square$

Lemma: Geht $\tilde{A} \in \mathbb{K}^{m \times n}$ aus $A \in \mathbb{K}^{m \times n}$ durch elementare Zeilenoperationen hervor, dann spannen die Zeilenvektoren von $\tilde{A}$ denselben Raum auf, wie die Zeilenvektoren von A.

Bem.: Dieser Raum wird Zeilenraum genannt.

Beweis: Sind $z_1, \dots, z_m \in \mathbb{K}^n$ die Zeilenvektoren von A, dann ist der zugehörige Zeilenraum $\operatorname{span}(z_1, \dots, z_m) = \{ \sum_{i=1}^{m} c_i z_i \mid c_i \in \mathbb{K} \} \subseteq \mathbb{K}^n$.

Dieser ändert sich offensichtlich nicht, wenn (i) Zeilenvektoren vertauscht werden oder (ii) ein Zeilenvektor z_k durch (λz_k) mit $\lambda \in \mathbb{K} \setminus \{0\}$ ersetzt wird. Wird (iii) z.B. die zweite zur ersten Zeile addiert, dann ist der resultierende Zeilenraum

$$\{ \tilde{c}_1 (z_1 + z_2) + \sum_{i=2}^{m} \tilde{c}_i z_i \mid \tilde{c}_1, \dots, \tilde{c}_m \in \mathbb{K} \} \underset{\substack{\uparrow \\ \tilde{c}_1 + \tilde{c}_2 = c_2 \text{ und } \tilde{c}_i = c_i \text{ für } i \neq 2}}{=} \{ \sum_{i=1}^{m} c_i z_i \mid c_i \in \mathbb{K} \}$$

$\square$

Satz: (Rang = Zeilenrang = Spaltenrang)

Sei $A \in \mathbb{K}^{m \times n}$, z die Anzahl lin. unabh. Zeilen von A und s die Anzahl linear unabh. Spalten von A. Dann gilt:

$\operatorname{rg}(A) = s = z$.

Bem.: Anders ausgedrückt: Zeilenraum, Spaltenraum und Bild einer Matrix haben stets die gleiche Dimension.

Beweis: Wir wissen bereits, dass $rg(A) = s$, da die Spalten von A die Bilder der Einheitsvektoren sind.

Elementare Zeilenoperationen ändern weder z (→ Lemma) noch $\ker(A)$.

Letzteres gilt, da $\ker(BA) = \ker(A)$ für invertierbares B.

Wegen $s = n - \dim(\ker(A))$ bleibt damit auch s unverändert, so dass wir o.B.d.A. Zeilenstufenform annehmen können:

$$\begin{pmatrix} 0 \,..\, 0 & A_{1,j_1} & & \\ 0 \,..\, 0 & 0 \,..\, 0 & A_{2,j_2} & \\ 0 \,..\, . & . & 0 \,..\, 0 & A_{3,j_3} \\ \vdots & & & 0 \quad \ddots \\ 0 & & & \vdots \\ & & & 0 \end{pmatrix}$$

Der von den Spaltenvektoren aufgespannte Raum ist dabei $\mathbb{K}^z \times \{0\}^{m-z}$. Also $z = s$. $\square$

Korollar: Für $A \in \mathbb{K}^{m \times n}$ gilt $rg(A) = rg(A^T)$.

Beweis: Dies folgt unmittelbar aus dem vorangehenden Satz, da $A \mapsto A^T$ lediglich die Rolle von Zeilen & Spalten vertauscht. $\square$

Bsp.: $A = \begin{pmatrix} 1 & 4 & 2 & 3 \\ 2 & 8 & 4 & 6 \\ 7 & 3 & 1 & 6 \\ -1 & -4 & -2 & -3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$ besitzt $rg(A) = 2$, da die Zeilen 1, 2 und 4 Vielfache voneinander sind aber die Zeilen 2 und 3 linear unabh..

Die Anzahl lin. unabh. Spalten ist weniger offensichtlich. Wir wissen aber, dass dies auch 2 sein muss.

VI. Determinanten

Def.:

- Eine bijektive Abbildung $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$ heißt Permutation.
- Eine Permutation heißt Transposition, wenn sie lediglich ein Paar vertauscht, d.h. wenn $\pi(i) = \begin{cases} k & ; i = l \\ l & ; i = k \\ i & ; \text{sonst} \end{cases}$ für ein Paar $k \neq l$.
- S_n bezeichnet die Menge aller Permutationen von $\underbrace{\{1, \dots, n\}}_{=:[n]}$.
- Für $\pi \in S_n$ definieren wir

$$a(\pi) := \left| \left\{ (i,j) \in [n] \times [n] \;\middle|\; i < j \wedge \pi(i) > \pi(j) \right\} \right| \in \left\{ 0, \dots, \frac{n(n-1)}{2} \right\}$$

und das Signum $\operatorname{sgn}(\pi) := (-1)^{a(\pi)}$.

- $\pi \in S_n$ heißt gerade wenn $\operatorname{sgn}(\pi) = 1$ und ungerade bei $\operatorname{sgn}(\pi) = -1$.

Bem.:

- Unter Komposition bildet S_n die Symmetrische Gruppe. Diese besitzt $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$ Elemente.
- Jede Permutation lässt sich als Komposition von Transpositionen realisieren, jede Transposition als Komposition einer ungeraden Anzahl von Nachbarvertauschungen (d.h. Transpositionen mit $k = l+1$).
- Für jede Nachbarvertauschung $\sigma \in S_n$ gilt $\forall \pi \in S_n$:

$$a(\sigma \circ \pi) = a(\pi) \pm 1, \quad \text{d.h.} \quad \operatorname{sgn}(\sigma \circ \pi) = -\operatorname{sgn}(\pi).$$

Damit ist $\pi \in S_n$ (un-)gerade, wenn π sich als Komposition einer (un-)geraden Zahl von Nachbarvertauschungen (bzw. Transpositionen) schreiben lässt. Zudem gilt:

$$\boxed{\forall \pi, \tau \in S_n : \quad \operatorname{sgn}(\pi \circ \tau) = \operatorname{sgn}(\pi) \cdot \operatorname{sgn}(\tau)}$$

Bsp.: Eine Permutation $\pi \in S_n$ ist eindeutig durch das Tupel $(\pi(1), \dots, \pi(n))$ charakterisiert.

Ist $\pi \in S_3$ so gegeben als $(3,2,1)$, dann ist $\operatorname{sgn}(\pi) = -1$, da die Reihenfolge aller drei Paare $(1,2), (1,3), (2,3)$ geändert wurde, d.h. $a(\pi) = 3$.

Def.: Die Determinante einer quadratischen Matrix $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ ist definiert als

$$\boxed{\det(A) := \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)}} \qquad (*)$$

Bem.:

- $\mathbb{K}$ dürfte hier auch ein kommutativer Ring sein.
- Die Formel $(*)$ heißt Leibniz-Formel.
- Jeder der $n!$ Summanden ist ein Produkt, zu dem jede Zeile und jede Spalte von A genau einen Eintrag als Faktor beiträgt.

Bsp.:

- $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{K}^{2 \times 2}$. S_2 hat zwei Elemente: π_1 entsprechend $(1,2)$ mit $\operatorname{sgn}(\pi_1) = 1$ und π_2 entsprechend $(2,1)$ mit $\operatorname{sgn}(\pi_2) = -1$. Also ist
$$\boxed{\det(A) = a_{11}a_{22} - a_{12}a_{21}}.$$
- Für eine Diagonalmatrix $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ mit $a_{ij} = \delta_{ij} a_i$ gilt $\boxed{\det(A) = \prod_{i=1}^{n} a_i}$ da für jedes $\pi \in S_n$ außer dem neutralen Element, für das $\pi(i) = i$ ist, das zugehörige Produkt $\prod_{i=1}^{n} a_{i,\pi(i)}$ eine Null enthält. Insbesondere gilt also $\boxed{\det(I_n) = 1}$.

- Ist $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ eine obere Dreiecksmatrix, d.h. $i > j \Rightarrow a_{ij} = 0$, dann gilt $\boxed{\det(A) = \prod_{i=1}^{n} a_{ii}}$, da alle anderen Produkte in der Leibnizformel wieder einen Faktor Null enthalten. (Analog für 'untere Dreiecksmatrizen')

- Für $A \in \mathbb{K}^{3 \times 3}$ lässt sich die Determinante mit der Sarrus-Regel berechnen:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{cccc} + & + & + & \\ a_{13} & a_{11} & a_{12} & a_{13} \\ a_{23} & a_{21} & a_{22} & a_{23} \\ a_{33} & a_{31} & a_{32} & a_{33} \\ - & - & - & \end{array} \begin{array}{c} a_{11} \\ a_{21} \\ a_{31} \end{array}$$

$$= a_{13}a_{21}a_{32} + a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31}$$
$$- a_{33}a_{21}a_{12} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11}$$

(Für höhere Dimensionen gilt keine analoge Regel. Für die Berechnung der Determinanten gibt es dann bessere Methoden als mittels Leibnizformel.)

Satz: Für $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ gilt:

(i) $\boxed{\det(A) = \det(A^T)}$

(ii) $\det(A)$ ist linear als Funktion einer bel. Spalte von A.

(iii) Geht $\tilde{A}$ aus A durch Vertauschen zweier Spalten hervor, dann gilt $\det(\tilde{A}) = -\det(A)$.

(iv) Hat A zwei identische Spalten, dann gilt $\det(A) = 0$.

(v) Für $B \in \mathbb{K}^{n \times n}$ gilt: $\boxed{\det(AB) = \det(A) \cdot \det(B)}$

Multiplikativität

Bem.: Wegen (i) gelten (ii), (iii) und (iv) auch, wenn man "Spalte(n)" durch "Zeile(n)" ersetzt.

Beweis: (i) Wir können die Faktoren in der Leibnizformel anstatt nach dem ersten Index auch nach dem zweiten ordnen. Genauer gilt:

$$\prod_{i=1}^{n} a_{i,\underbrace{\pi(i)}_{=:j}} = \prod_{j=1}^{n} a_{\pi^{-1}(j),j}$$

Damit bekommen wir:

$$\det(A) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\pi^{-1}(i),i}$$

$$= \sum_{\pi\in S_n} \operatorname{sgn}(\pi^{-1}) \prod_{i=1}^{n} a_{\pi^{-1}(i),i} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\pi(i),i} = \det(A^T)$$

$\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$ da $1 = \operatorname{sgn}(id) = \operatorname{sgn}(\pi\circ\pi^{-1}) = \operatorname{sgn}(\pi)\cdot\operatorname{sgn}(\pi^{-1})$

(ii) Folgt unmittelbar aus der Leibnizformel, denn ist $v \in \mathbb{K}^n$ der k'te Spaltenvektor von A, dann gilt $\det(A) = \sum_{i=1}^{n} c_i v_i$ wobei die $c_i \in \mathbb{K}$ nicht von v abhängen.

(iii) Für $\tilde{A} = (\tilde{a}_{ij})$ mit $\tilde{a}_{ij} = a_{i,\sigma(j)}$, wobei $\sigma \in S_n$ eine Transposition ist, gilt:

$$\det(\tilde{A}) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} \tilde{a}_{i,\pi(i)}$$

$$= \sum_{\pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\sigma\circ\pi(i)}$$

$$\overset{\operatorname{sgn}(\sigma^{-1}) = -1}{=} - \sum_{\pi} \operatorname{sgn}(\sigma\circ\pi) \prod_{i=1}^{n} a_{i,\sigma\circ\pi(i)} = -\det(A)$$

$\pi \mapsto \sigma\circ\pi$ Bijektion auf S_n

(iv) folgt aus (iii), da Vertauschen gleicher Spalten auf $\det(A) = -\det(A)$ führt. (Bem.: wenn "$\operatorname{char}(\mathbb{K}) = 2$", d.h. in $\mathbb{K}$ gilt, dass $1+1=0$ ist, benötigt es noch ein weiteres Argument, welches z.B. in [Reimers, Kemper] zu finden ist.)

(v) Sei $B = (b_{ij})$. Dann gilt:

$$\det(AB) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} \prod_{i=1}^{n} a_{i,k_i} b_{k_i,\pi(i)}$$

$$\prod_{i=1}^{n} \sum_{k=1}^{n} a_{i,k} b_{k,\pi(i)} = \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} \prod_{i=1}^{n} a_{i,k_i} b_{k_i,\pi(i)}$$

$$= \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} \prod_{j=1}^{n} a_{j,k_j} \underbrace{\sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} b_{k_i,\pi(i)}}$$

Für festes $k_1, \dots, k_n$ ist dies die det einer Matrix deren Zeilen die Zeilen $k_1, \dots, k_n$ von B sind. Nach (iv) ist dies Null, wenn die k_i's nicht alle verschieden sind. D.h. o.B.d.A. $(k_1, \dots, k_n) = (\sigma(1), \dots, \sigma(n))$ für ein $\sigma \in S_n$.

$$= \sum_{\sigma \in S_n} \prod_{j=1}^{n} a_{j,\sigma(j)} \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} b_{\sigma(i),\pi(i)}$$

analog zum Beweis von (i)

$$\overset{\downarrow}{=} \underbrace{\sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)}}_{\det(A)} \underbrace{\sum_{\pi} \operatorname{sgn}(\pi \circ \sigma^{-1}) \prod_{i=1}^{n} b_{i,\pi\circ\sigma^{-1}(i)}}_{\det(B)}$$

$$=$$

$\square$

<u>Korollar:</u> Geht $\tilde{A} \in \mathbb{K}^{n\times n}$ aus $A \in \mathbb{K}^{n\times n}$ hervor durch

(i) Vertauschung zweier Zeilen (oder Spalten), dann gilt $\det(\tilde{A}) = -\det(A)$,

(ii) Multiplizieren einer Zeile oder Spalte mit einem Skalar $\lambda \in \mathbb{K}$, dann gilt $\det(\tilde{A}) = \lambda \det(A)$,

(iii) Addieren des Vielfachen einer Zeile zu einer anderen oder einer Spalte zu einer anderen, dann gilt $\det(\tilde{A}) = \det(A)$.

<u>Beweis:</u> (i) & (ii) folgen unmittelbar aus dem Satz.

(iii) folgt (für Zeilen) aus $\det(\tilde{A}) = \det(EA) = \det(E)\det(A)$, da E eine Dreiecksmatrix mit $E_{ii} = 1$ ist, weshalb $\det(E) = 1$.

$\square$

Bem.: Bringen wir eine Matrix aus $\mathbb{K}^{n\times n}$ durch elem. Zeilenop.en auf Zeilenstufenform (in der die Determinante einfach das Produkt der Diagonaleinträge ist), lässt sich über die Änderung der Determinanten also einfach Buch führen.

Korollar: Für $\lambda \in \mathbb{K}$, $A \in \mathbb{K}^{n\times n}$ gilt: $\boxed{\det(\lambda A) = \lambda^n \det(A)}$.

Beweis: Folgt aus (ii) angewendet auf alle n Spalten. $\square$

Satz: Für $A \in \mathbb{K}^{n\times n}$ sind äquivalent:

(i) A ist invertierbar

(ii) $rg(A) = n$

(iii) $\det(A) \neq 0$

Def.: $A \in \mathbb{K}^{n\times n}$ heißt regulär, wenn $\det(A) \neq 0$ und singulär, wenn $\det(A) = 0$.

Beweis: (i) $\Leftrightarrow$ (ii) ist bereits bekannt.

(i) $\Rightarrow$ (iii): Aus $I_n = A \cdot A^{-1}$ folgt

$$1 = \det(I_n) = \det(A \cdot A^{-1}) = \det(A)\det(A^{-1}) \text{ also } \det(A) \neq 0.$$

(iii) $\Rightarrow$ (ii): Ist $\det(A) \neq 0$, dann ändert sich dies durch elementare Zeilenop.en genauso wenig wie $rg(A)$. In Zeilenstufenform bedeutet $\det(A) \neq 0$ jedoch, dass $a_{ii} \neq 0 \ \forall i$. Damit hat A n linear unabh. Spalten, also $rg(A) = n$. $\square$

Aus dem Beweis (i) $\Rightarrow$ (iii) folgt unmittelbar:

Korollar: Ist $A \in \mathbb{K}^{n\times n}$ invertierbar, dann gilt $\boxed{\det(A^{-1}) = \det(A)^{-1}}$.

Bsp.: Liegen die Punkte $(2,-1), (-7,-7), (8,3) \in \mathbb{R}^2$ auf einer Linie?

Eine Linie ist über $\{(x,y) \in \mathbb{R}^2 \mid ax+by+c=0\}$ durch $(a,b,c) \in \mathbb{R}^3$ charakterisiert, wobei $(a,b) \in \mathbb{R}^2 \setminus \{0\}$. Punkte $z_1, z_2, z_3 \in \mathbb{R}^2$ liegen also nur dann auf einer Linie, wenn das homogene LGS

$$\underbrace{\begin{pmatrix} z_{11} & z_{12} & 1 \\ z_{21} & z_{22} & 1 \\ z_{31} & z_{32} & 1 \end{pmatrix}}_{=:Z} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0$$

eine von Null verschiedene Lösung besitzt, also $\det(Z) = 0$ ist.

Im Beispiel ist $\det(Z) = \det \begin{pmatrix} 2 & -1 & 1 \\ -7 & -7 & 1 \\ 8 & 3 & 1 \end{pmatrix} = 0$.

Satz: (Entwicklungssatz von Laplace) Für $A = (a_{ij}) \in \mathbb{K}^{n \times n}$, $n \geq 2$ sei $A_{i,j} \in \mathbb{K}^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Weglassen der i-ten Zeile und j-ten Spalte entsteht. Dann gilt

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

"Entwicklung nach der i'ten Zeile"

(2) $\forall j \in \{1,\dots,n\}$: $$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \cdot \det(A_{i,j})$$

"Entwicklung nach der j'ten Spalte"

Beweis: mittels Leibnizformel $\longrightarrow$ siehe z.B. [Reimers, Kemper].

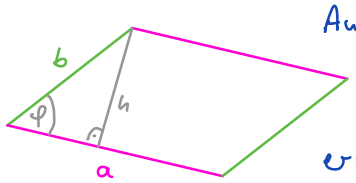
Bsp.:

$$\det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \overset{\text{Entw. nach 1. Zeile}}{=} 0 \cdot \det\begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} - 1 \cdot \det\begin{pmatrix} 3 & 5 \\ 6 & 8 \end{pmatrix} + 2 \cdot \det\begin{pmatrix} 3 & 4 \\ 6 & 7 \end{pmatrix}$$

$$= -(3\cdot 8 - 5\cdot 6) + 2(3\cdot 7 - 6\cdot 4) = 6 - 6 = 0.$$

Geometrische Interpretation der Determinanten

Ist F die Fläche eines durch $a, b \in \mathbb{R}^2$ aufgespannten Parallelogramms, dann gilt $F = \|a\| \cdot h = \|a\| \cdot \|b\| \cdot \sin\varphi$.



Aus $F^2 + \underbrace{(\|a\|\,\|b\| \cos\varphi)}_{= \langle a, b \rangle = a_1 b_1 + a_2 b_2}{}^2 = \|a\|^2 \|b\|^2$

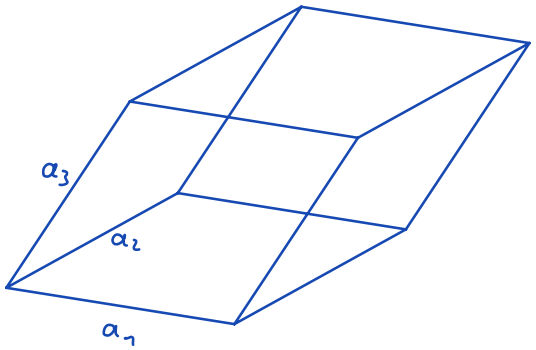
erhält man $F = |a_1 b_2 - a_2 b_1|$

$$= \left| \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \right| .$$

Dies gilt auch allgemeiner: Spannen $a_1, \dots, a_n \in \mathbb{R}^n$ ein Parallelotop auf, dann ist dessen Volumen V gleich

$$\boxed{V = \left| \det \left(a_1 \,\vdots\, \dots \,\vdots\, a_n \right) \right|}$$

(→ Analysis 3)



Das Parallelotop ist dabei das Bild des Einheitswürfels unter der durch $A = \left(a_1 \,\vdots\, \dots \,\vdots\, a_n \right)$ gegebenen linearen Abbildung.

Lemma: (Basiswechsel)

Ist V ein n-dim. $\mathbb{K}$-Vektorraum, $\phi_B : \mathbb{K}^n \to V$ und $\phi_C : \mathbb{K}^n \to V$ zwei kanonische Basisisomorphismen und $F : V \to V$ eine lineare Abbildung mit Darstellungsmatrizen $M_B := \phi_B^{-1} \circ F \circ \phi_B$ und $M_C := \phi_C^{-1} \circ F \circ \phi_C$. Dann gilt:

$$\boxed{M_B = S_{B,C} M_C S_{B,C}^{-1}} \quad \text{mit } S_{B,C} := \phi_B^{-1} \circ \phi_C : \mathbb{K}^n \to \mathbb{K}^n$$

Beweis:

$$M_B = \phi_B^{-1} \circ F \circ \phi_B \underset{F = \phi_C \circ M_C \circ \phi_C^{-1}}{=} (\phi_B^{-1} \circ \phi_C) \circ M_C \circ \underbrace{\phi_C^{-1} \circ \phi_B}_{(\phi_B^{-1} \circ \phi_C)^{-1}} . \qquad \square$$

Bem.: Umgekehrt können wir jede Transformation $M_C \mapsto S M_C S^{-1}$ als Basiswechsel interpretieren/realisieren. Diese Transformationen heißen **Ähnlichkeitstransformationen** und zwei Matrizen —"— **ähnlich**, wenn eine —"— sie ineinander überführt.

Korollar: Sind $M_B, M_C \in \mathbb{K}^{n \times n}$ zwei Darstellungsmatrizen ein- und derselben linearen Abbildung, dann gilt:

$$\det(M_B) = \det(M_C) .$$

Beweis:

$$\det(M_B) \underset{\text{Multiplikativität}}{=} \det(S_{B,C}) \det(M_C) \underbrace{\det(S_{B,C}^{-1})}_{\det(S_{B,C})^{-1}} = \det(M_C) . \qquad \square$$

Als Konsequenz können wir $\det(F)$ für $F : V \to V$ Basisunabhängig definieren, da die Darstellungsmatrizen für F in allen Basen dieselbe Determinante besitzen.

VII. Eigenwerte & Eigenvektoren

Bem.: Eigenwerte spielen eine wichtige Rolle z.B.:

- in der Quantentheorie, wo sie die Menge der (quantisierten) Messwerte darstellen ($\longrightarrow$ z.B. Fraunhoferlinien).
- in der klassischen Mechanik als Hauptträgheitsmomente

Def.: Sei V ein endlich-dim. $\mathbb{K}$-Vektorraum und $F: V \to V$ linear. Gilt die Eigenwertgleichung $\boxed{Fv = \lambda v}$ für ein $\lambda \in \mathbb{K}$ und ein $v \in V \setminus \{0\}$, dann heißt v Eigenvektor zum Eigenwert λ.

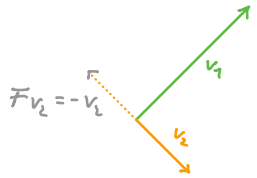
Der Vektorraum $E_\lambda := \ker(F - \lambda \mathrm{id}_V) \subseteq V$ heißt Eigenraum zum Eigenwert λ, seine Dimension $\dim(E_\lambda)$ geometrische Vielfachheit des Eigenwertes λ.

Bem.:

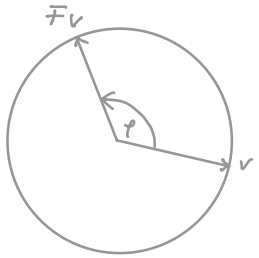
- Es gilt: $v \in E_\lambda \iff Fv = \lambda v$.
  Damit ist E_λ der Unterraum von V, der von allen Eigenvektoren zum Eigenwert λ aufgespannt wird und $\dim(E_\lambda)$ die Anzahl lin. unabh. Eigenvektoren zum Eigenwert λ.
- Ist v ein Eigenvektor, dann auch αv für bel. $\alpha \in \mathbb{K} \setminus \{0\}$.
- Ist ein Eigenwert bekannt, ist die Bestimmung des zugehörigen Eigenraumes, die Lösung eines LGS.
- $\boxed{\text{von nun an betrachten wir ausschließlich endl. dim. Vektorräume.}}$

Bsp.:

- Die Spiegelung im $\mathbb{R}^2$ an einer durch $v_1 \in \mathbb{R}^2 \setminus \{0\}$ gegebenen Achse besitzt zwei Eigenwerte $+1$ und -1 mit Eigenvektoren v_1 und v_2 der senkrecht auf v_1 steht.
  Es existiert also eine <u>Basis</u> aus Eigenvektoren.



- Die Drehung des $\mathbb{R}^2$ um einen Winkel $\varphi \in (0, \pi)$ besitzt offenbar keinen Eigenwert.



- Die durch $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ dargestellte Scherung des $\mathbb{R}^2$ besitzt wegen $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ b \end{pmatrix} \overset{!}{=} \lambda \begin{pmatrix} a \\ b \end{pmatrix}$ nur einen Eigenwert $\lambda = 1$. Der Eigenraum dazu ist ein-dimensional.

Def.: Eine lineare Abbildung $F: V \to V$ heißt diagonalisierbar, wenn eine Basis von Eigenvektoren von F für V existiert.

Bem.:

- Angenommen $v_1, \dots, v_n$ ist so eine Basis, mit $Fv_i = \lambda_i v_i$.
  Der Name „diagonalisierbar" kommt daher, dass F in dieser Basis 'diagonal' ist, nämlich dargestellt durch die Matrix

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

- Für eine Matrix $A \in \mathbb{K}^{n \times n}$ bedeutet diagonalisierbar, dass eine Ähnlichkeitstrafo. existiert, so dass SAS^{-1} diagonal ist. Diese können wir explizit angeben, wenn die Eigenvektoren gegeben sind:

<u>Satz:</u> Sei $v_1, \dots, v_n \in \mathbb{K}^n$ eine Basis von Eigenvektoren von $A \in \mathbb{K}^{n \times n}$ und $X := \left(v_1 \,\vdots\, \cdots \,\vdots\, v_n \right) \in \mathbb{K}^{n \times n}$ die Matrix, deren Spalten durch die v_k's gegeben sind. Dann ist $X^{-1}AX$ diagonal. Genauer:

$$\boxed{X^{-1}AX = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}$$

wobei $A v_k = \lambda_k v_k$.

<u>Beweis:</u> Zu zeigen ist, dass $(X^{-1}AX)e_k = \lambda_k e_k$ gilt, für alle Standardbasisvektoren e_k. Da das Bild von e_k unter X der k'te Spaltenvektor von X und damit v_k ist gilt:

$$(X^{-1}AX)e_k = X^{-1}A\underbrace{(Xe_k)}_{=v_k} \overset{Av_k = \lambda_k v_k}{=} X^{-1}\lambda_k v_k \underset{v_k = Xe_k}{=} \lambda_k X^{-1}Xe_k \underset{X^{-1}X = I_n}{=} \lambda_k e_k.$$

$\square$

Lemma: Sind $v_1, \dots, v_r$ Eigenvektoren zu unterschiedlichen Eigenwerten $\lambda_1, \dots, \lambda_r$ einer linearen Abbildung $F: V \to V$, dann ist $(v_1, \dots, v_r)$ linear unabhängig.

Beweis: durch Induktion:

Beginn: $r=1$. Da $v_1 \neq 0$, ist v_1 in der Tat lin. unabh.

Induktionsannahme: $v_1, \dots, v_k$ sind lin. unabh.

Wir wollen zeigen, dass daraus lin. Unabh. von $v_1, \dots, v_{k+1}$ folgt.

Betrachte daher $\sum_{i=1}^{k+1} \alpha_i v_i = 0$ mit $\alpha_i \in \mathbb{K}$.

Dann gilt $\sum_{i=1}^{k+1} \alpha_i F v_i = 0$ sowie $\lambda_{k+1} \sum_{i=1}^{k+1} \alpha_i v_i = 0$

und für deren Differenz:

$$0 = \sum_{i=1}^{k+1} \alpha_i (\lambda_i - \lambda_{k+1}) v_i = \sum_{i=1}^{k} \alpha_i (\lambda_i - \lambda_{k+1}) v_i$$

Hier muß $\alpha_i (\lambda_i - \lambda_{k+1}) = 0$ sein also $\alpha_i = 0$ da $\lambda_i \neq \lambda_{k+1}$.

$v_1, \dots, v_{k+1}$ sind damit ebenfalls lin. unabh.

$\square$

Satz: Sei V ein n-dim. $\mathbb{K}$-Vektorraum und $F: V \to V$ linear mit r unterschiedlichen Eigenwerten $\lambda_1, \dots, \lambda_r$, deren geometrische Vielfachheiten $n_1, \dots, n_r$ sind. Dann gilt $\boxed{\sum_{i=1}^{r} n_i \leq n}$ und hier Gleichheit genau dann, wenn F diagonalisierbar ist.

Beweis: Sei $v_1^{(i)}, \dots, v_{n_i}^{(i)}$ eine Basis von E_{λ_i}. Wir wollen zunächst

zeigen, dass auch die Vereinigung über i zu einem linear unabhängigen Satz von Vektoren führt.

Sei $\sum_{i=1}^{r} \underbrace{\sum_{k=1}^{n_i} \alpha_k^{(i)} v_k^{(i)}}_{\overset{!!}{=} w_i} = 0$. Dann sind nach dem Lemma alle $= 0$, da $w_i \in E_{\lambda_i}$. Da für festes i aber $v_1^{(i)}, \dots, v_{n_i}^{(i)}$ lin. unabh. sind, müssen alle $\alpha_k^{(i)} = 0$. Also sind $\{v_1^{(i)}, \dots, v_{n_i}^{(i)}\}_{i=1}^{r}$ $\sum_{i=1}^{r} n_i$ linear unabh. Vektoren, so dass $\sum_{i=1}^{r} n_i \leq n$.

Gilt Gleichheit, existiert eine Basis von Eigenvektoren. Existiert, umgekehrt, eine solche -"- -"-, dann muß auch Gleichheit gelten, da $\sum_{i=1}^{r} n_i$ die Anzahl der lin. unabh. Eigenvektoren ist. $\square$

<u>Bem.:</u> Eigenwerte ändern sich nicht unter Ähnlichkeitstransformationen, da $Fv = \lambda v \iff SFS^{-1}\tilde{v} = \lambda \tilde{v}$ mit $\tilde{v} := Sv$.
Wir können eine lineare Abb. $F: V \to V$ also in beliebiger Basis darstellen und die Eigenwerte von F als Eigenwerte der Darstellungsmatrix bekommen. Zu deren Bestimmung definieren wir:

<u>Def.:</u> Zu $A \in \mathbb{K}^{n \times n}$ heißt $\chi_A \in \mathbb{K}[x]$, $\chi_A(x) := \det(x I_n - A)$ das charakteristische Polynom.

<u>Bem.:</u> ∘ In der Literatur ist auch $\det(A - x I_n)$ als Definition für χ_A zu finden.

- χ_A ist ein Polynom vom Grad n. Die beiden führenden Monome ergeben sich gemäß Leibnizformel für $A = (a_{ij})$ aus

$$\prod_{i=1}^{n} (x - a_{ii}) = x^n - x^{n-1} \left(\sum_{i=1}^{n} a_{ii} \right) + \dots$$

Der Term $\sum_{i=1}^{n} a_{ii} =: tr[A]$ heißt Spur (engl. 'trace') von A.

Insgesamt ist damit:

$$\boxed{\chi_A(x) = x^n - tr[A]\, x^{n-1} + \dots + (-1)^n \det(A)}$$

Korollar: 1. Die Eigenwerte von $A \in \mathbb{K}^{n \times n}$ sind genau die Nullstellen des charakteristischen Polynoms χ_A.

2. Ist $\mathbb{K}$ algebraisch abgeschlossen, besitzt $A \in \mathbb{K}^{n \times n}$ mindestens einen Eigenwert.

Beweis: 1.

$$\begin{aligned} 0 = \chi_A(\lambda) &= \det(\lambda I_n - A) \\ \Leftrightarrow \quad & rg(\lambda I_n - A) < n \\ \Leftrightarrow \quad & \dim \ker(\lambda I_n - A) \geq 1 \\ \Leftrightarrow \quad & \exists v \neq 0 : \quad \lambda v = Av \end{aligned}$$

2. χ_A besitzt dann eine Nullstelle. $\square$

Bsp.: $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ besitzt keinen Eigenwert über $\mathbb{R}$, also als $A \in \mathbb{R}^{2 \times 2}$. $A \in \mathbb{C}^{2 \times 2}$ dagegen besitzt Eigenwerte $\pm i$, da $\chi_A(\lambda) = \lambda^2 + 1$ Nullstellen $\pm i$ besitzt.

Bem.: Ähnliche Matrizen besitzen das gleiche charakteristische Polynom, da für $B = SAS^{-1}$ gilt: $\chi_B(x) =$

$$= \det(xI_n - SAS^{-1}) = \det(S(xI_n - A)S^{-1}) = \underbrace{\det(S)\det(S^{-1})}_{=1} \chi_A(x).$$

Def.: Sei $\lambda \in \mathbb{K}$ ein Eigenwert von $A \in \mathbb{K}^{n \times n}$. Die algebraische Vielfachheit von λ ist die Vielfachheit der Nullstelle λ im charakteristischen Polynom von A.

(D.h. das maximale $m \in \mathbb{N}$, s.d. $\chi_A(x) = (x-\lambda)^m f(x)$ mit $f \in \mathbb{K}[x]$)

Bem.: Ist $\mathbb{K}$ algebraisch abgeschlossen, hat das charakteristische Polynom nach dem Fundamentalsatz der Algebra die Form

$$\boxed{\chi_A(x) = \prod_{i=1}^{r} (x-\lambda_i)^{a_i}}$$

wobei $a_i \in \mathbb{N}$ die algebraische Vielfachheit des Eigenwerts λ_i ist und $\sum_{i=1}^{r} a_i = n$ gilt.

Satz: Die algebraische Vielfachheit eines Eigenwerts ist mindestens gleich seiner geometrischen Vielfachheit.

Beweis: Wir verwenden, dass sich die Vielfachheiten unter Basiswechsel nicht ändern. Sei $\lambda \in \mathbb{K}$ der betrachtete Eigenwert mit algebraischer V. $a(\lambda)$ und geom. V. $g(\lambda)$. Dann gibt es $g(\lambda)$ linear unabh. Eigenvektoren $v_1, \dots, v_{g(\lambda)}$. Wir vervollständigen dies zu einer Basis und stellen die Matrix in dieser dar was

auf $\tilde{A} := \left(\begin{array}{c|c} \lambda I_{g(\lambda)} & * \\ \hline 0 & C \end{array}\right)$ führt. Wie $\tilde{A}$ ist $xI_n - \tilde{A}$

eine Blockdreiecksmatrix, deren Determinante gleich

$\chi_{\tilde{A}}(x) = (x-\lambda)^{g(\lambda)} \chi_C(x)$ ist.

(Hier verwenden wir, dass $\det\begin{pmatrix} B & * \\ 0 & C \end{pmatrix} = \det(B)\det(C)$. S.u. ...) $\square$

Def.: Ist $\mathbb{K}$ algebraisch abgeschlossen und $A \in \mathbb{K}^{n \times n}$, dann heißt $\sigma(A) := \{ \lambda \in \mathbb{K} \mid \det(\lambda I_n - A) = 0 \}$ das Spektrum von A.

Korollar: Ist $\mathbb{K}$ algebraisch abgeschlossen, dann ist $A \in \mathbb{K}^{n \times n}$ genau dann diagonalisierbar, wenn für jeden Eigenwert von A gilt, dass algebraische und geometrische Vielfachheit übereinstimmen.

Beweis: Für $\lambda \in \sigma(A)$ seien $a(\lambda)$ und $g(\lambda)$ die alg. bzw. geom. Vielfachheit. Dann gilt:

$$\sum_{\lambda \in \sigma(A)} g(\lambda) \underset{\substack{\uparrow \\ \text{vorigr Satz}}}{\leq} \sum_{\lambda \in \sigma(A)} a(\lambda) \underset{\substack{\uparrow \\ \text{Fundamentalsatz der Algebra}}}{=} n$$

Damit ist die Summe der geom. Vielfachheiten (also die Anzahl der lin. unabh. Eigenvektoren) gleich n, g.d.w. $g(\lambda) = a(\lambda) \quad \forall \lambda \in \sigma(A)$. $\square$

Lemma: Ist $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ eine obere Dreiecksmatrix, dann gilt $\chi_A(x) = \prod_{i=1}^{n} (x - a_{ii})$, d.h. die Eigenwerte von A sind genau die Diagonaleinträge.

Beweis: $xI_n - A$ ist ebenfalls obere Dreiecksmatrix, s.d. die Determinante das Produkt der Diagonaleinträge ist. $\square$

Zwar lassen sich auch für alg. abgeschlossenes $\mathbb{K}$ nicht alle Matrizen diagonalisieren, jedoch ist jede Matrix trigonalisierbar:

<u>Satz:</u> Ist $\mathbb{K}$ algebraisch abgeschl. und $A \in \mathbb{K}^{n \times n}$, dann existiert eine invertierbare Matrix $S \in \mathbb{K}^{n \times n}$, s.d. $S^{-1}AS$ eine obere Dreiecksmatrix ist.

<u>Beweis:</u> Dies lässt sich induktiv beweisen:

Wegen alg. Abgeschlossenheit besitzt A einen Eigenwert, den wir mit λ_1 bezeichnen. Dargestellt in der Basis, die den zug. Eigenvektor als $(1, 0, \dots, 0)$ darstellt, wird A zu

$$A \longmapsto \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & B \end{array}\right) \quad \text{mit } B \in \mathbb{K}^{(n-1)\times(n-1)}$$

Nun verfahren wir mit B genauso. B hat einen Eigenwert λ_2, so dass nach geeigneter Ähnlichkeitstransformation

$$A \longmapsto \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & B \end{array}\right) \longmapsto \left(\begin{array}{c|c|c} \lambda_1 & * & * \\ \hline 0 & \lambda_2 & * \\ \hline 0 & 0 & C \end{array}\right)$$

Iteration führt dann auf eine obere Dreiecksmatrix, mit den Eigenwerten auf der Diagonalen. $\square$

Korollar: $\det\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A)\cdot\det(B)$

Beweis: Sind SAS^{-1} und TBT^{-1} obere Dreiecksmatrizen dann gilt:

$$\det\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det\left[\begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix}^{-1}\right]$$

$$= \det\begin{pmatrix} SAS^{-1} & SCT^{-1} \\ 0 & TBT^{-1} \end{pmatrix} \quad \text{ist ob. D.matrix, s.d.}$$

$$= \det(SAS^{-1})\det(TBT^{-1}) = \det(A)\det(B).$$

$\square$

Korollar: Sei $\mathbb{K}$ alg. abgeschlossen und sind $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ die (nicht notwendigerweise versch.) Eigenwerte von $A \in \mathbb{K}^{n\times n}$.

Dann gilt:

$$\boxed{\begin{aligned} &(i) \quad \operatorname{tr}[A] = \sum_{i=1}^{n} \lambda_i \\ &(ii) \quad \det(A) = \prod_{i=1}^{n} \lambda_i \end{aligned}}$$

d.h. die Eigenwerte werden hier gemäß ihrer alg. Vielfachheit gezählt

Beweis: Wir verwenden, dass sich das charak. Polynom nicht unter Ähnlichkeitstransformationen ändert und dass stets ein S existiert, s.d. $B = SAS^{-1}$ eine obere Dreiecksmatrix ist, für die (i) & (ii) bekannt sind. Für A folgt die Aussage dann aus

$$\chi_A(x) = x^n + \operatorname{tr}[A]\, x^{n-1} + \dots + (-1)^n \det(A)$$

$$\overset{\shortparallel}{\chi_B(x)} = x^n + \operatorname{tr}[B]\, x^{n-1} + \dots + (-1)^n \det(B)$$

durch Koeffizientenvergleich. $\square$

Bem.: Dies zeigt gleichzeitig, dass die Spur 'basisunabhängig' ist, d.h. es gilt $\boxed{\operatorname{tr}[A] = \operatorname{tr}[SAS^{-1}]}$ für bel. invertierbares S.

Def.: Für $\lambda \in \mathbb{K}$, $m \in \mathbb{N}$ heißt die Matrix $J(\lambda) \in \mathbb{K}^{m \times m}$

$$\boxed{J(\lambda) := \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}}$$

Jordanblock (oder 'Jordankästchen') der Größe m zum Eigenwert λ. Für $m=1$ meint dies $J(\lambda) := \lambda$.

Bem.:
- Offensichtlich ist λ der einzige Eigenwert von $J(\lambda)$.
- Manche Texte verwenden die Konvention in der die Einsen unter der Diagonale stehen.

Jordanblöcke der Größe $m \geq 2$ sind paradigmatisch für nicht-diagonalisierbare Matrizen, denn:

Lemma: Der Eigenwert λ eines Jordanblocks $J(\lambda) \in \mathbb{K}^{m \times m}$ besitzt algebraische Vielfachheit m und geom. Vielfachheit 1.

Bem.: D.h. ein Jordanblock besitzt genau einen Eigenvektor (nämlich $(1, 0, \dots, 0)$).

Beweis: Für das charakteristische Polynom von $J(\lambda)$ gilt:

$$\chi(x) = \det(x I_m - J(\lambda)) = (x - \lambda)^m.$$

Damit ist die alg. Vielfachheit von λ gleich m.

Der Eigenraum $E_\lambda = \ker(J(\lambda) - \lambda I_m)$ ist dagegen ein-dimensional, da

$$J(\lambda) - \lambda I_m = \begin{pmatrix} 0 & 1 & & \\ \vdots & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$ Rang $m-1$ besitzt. $\square$

Satz: (Jordan-Normalform)

Sei $A \in \mathbb{K}^{n \times n}$ so, dass das charak. Polynom in Linearfaktoren zerfällt (d.h. $\chi_A(x) = \prod_{i=1}^{n} (x - \lambda_i)$ mit $\lambda_i \in \mathbb{K}$, was für alg. abgeschlossene $\mathbb{K}$ stets zutrifft).

A ist ähnlich zu einer blockdiagonalen Matrix der Form

$$\boxed{\begin{pmatrix} J_1(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_2(\lambda_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J_r(\lambda_r) \end{pmatrix}}$$

wobei jedes $J_k(\lambda_k)$ ein Jordanblock ist. Diese sog. Jordan-Normalform der Matrix A ist eindeutig bis auf die Reihenfolge der Jordanblöcke.

Beweis: z.B. in [Kemper & Reimers].

Bem.:
- Die λ_k's sind dann die Eigenwerte von A, wobei die geometrische Vielfachheit eines Eigenwerts λ gleich der Anzahl der Jordanblöcke ist, die in der Jordan-Normalform zum Eigenwert λ auftauchen.
- A ist genau dann diagonalisierbar, wenn alle Jordanblöcke der Jordan-Normalform Größe 1 besitzen.
- Über $\mathbb{R}$ existiert die Jordan-Normalform nicht notwendigerweise (aber eine modifizierte Version davon ...).

Lemma: Hat $A \in \mathbb{C}^{n \times n}$ ausschließlich reelle Einträge, dann gilt $\lambda \in \operatorname{spec}(A) \Rightarrow \bar{\lambda} \in \operatorname{spec}(A)$

D.h. die Eigenwerte reeller Matrizen sind entweder reell oder sind Teil eines komplex konjugierten Paars.

Beweis: komplex-Konjugation der Eigenwertgleichung $Av = \lambda v$ liefert $\underbrace{A}_{=\bar{A}} \bar{v} = \bar{\lambda} \bar{v}$. $\square$

Bsp.: Für $a, b \in \mathbb{R}$ besitzt $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ die Eigenwerte $a \pm ib$.

Damit kann keine <u>reelle</u> Ähnlichkeitstrafo. existieren, die die Matrix auf Jordan-Normalform bringt, wenn $b \neq 0$ ist.

VIII. Skalarprodukt

Wir definieren zunächst ein Konzept, das die 'Länge' von Vektoren misst:

Def.: Sei V ein reeller oder komplexer Vektorraum. Eine Abbildung $V \ni x \mapsto \|x\| \in \mathbb{R}$ heißt Norm, wenn gilt: $\forall x, y \in V$, $\alpha \in \mathbb{K}$:

(i) $\|x\| = 0 \Rightarrow x = 0$ Definitheit

(ii) $\|\alpha x\| = \underbrace{|\alpha|}_{\text{Betrag von } \alpha} \cdot \|x\|$ Absolute Homogenität

(iii) $\|x+y\| \leq \|x\| + \|y\|$ Dreiecksungleichung

Folgende Konsequenz wird oft gleich mit in die Definition gepackt:

Korollar: Für eine Norm gilt $\forall x \in V: \|x\| \geq 0$ wobei $\|x\| = 0$ genau dann gilt, wenn $x = 0$ ist.

Beweis: Wegen (ii) gilt $\|0\| \overset{(ii)}{=} 0 \cdot \|0\| = 0$. Aus (iii) folgt dann: $0 = \|x - x\| \leq \|x\| + \|-x\| = 2\|x\|$. Zusammen mit (i) folgt die Aussage. $\square$

Bsp.:

- Maximumsnorm für $V \in \{\mathbb{R}^n, \mathbb{C}^n\}$: $\|x\|_\infty := \max_{i \in \{1..n\}} |x_i|$
- Supremumsnorm z.B. für $V = C([0,1])$: $\|f\|_\infty := \sup_{x \in [0,1]} |f(x)|$
- Euklid'sche Norm für $V = \mathbb{R}^n$: $\|x\| := \sqrt{\sum_{i=1}^n x_i^2}$

Def.: • Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Ein Skalarprodukt auf V ist eine Abbildung

$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$, für die gilt: $\forall\, \varphi, \varphi_1, \varphi_2, \psi \in V$:

(i) $\langle \varphi, \varphi \rangle \in [0, \infty)$ wobei $\langle \varphi, \varphi \rangle = 0 \Leftrightarrow \varphi = 0$.

(ii) $\forall \lambda \in \mathbb{K}$: $\langle \psi, \lambda \varphi_1 + \varphi_2 \rangle = \lambda \langle \psi, \varphi_1 \rangle + \langle \psi, \varphi_2 \rangle$

(iii) $\langle \varphi, \psi \rangle = \overline{\langle \psi, \varphi \rangle}$ (bzw. $\langle \varphi, \psi \rangle = \langle \psi, \varphi \rangle$ wenn $\mathbb{K} = \mathbb{R}$)

• V ausgestattet mit einem Skalarprodukt heißt dann Euklid'scher Vektorraum für $\mathbb{K} = \mathbb{R}$ bzw. unitärer Vektorraum für $\mathbb{K} = \mathbb{C}$. Auf diesem ist die (vom Skalarprodukt induzierte) Norm $\|\cdot\| : V \to [0, \infty)$, $\boxed{\|x\| := \sqrt{\langle x, x \rangle}}$ definiert.

Bem.: ∘ Für $\mathbb{K} = \mathbb{R}$ ist das Skalarprodukt also (i) positiv definit, (ii) linear in beiden Argumenten und (iii) symmetrisch.

∘ Für $\mathbb{K} = \mathbb{C}$ ist das Skalarprodukt linear im zweiten Argument und 'konjugiert linear' im ersten Argument, d.h. $\langle \lambda \varphi_1 + \varphi_2, \psi \rangle = \bar{\lambda} \langle \varphi_1, \psi \rangle + \langle \varphi_2, \psi \rangle$.

(Achtung: Manche Mathematiker verwenden die Konvention nach der Linearität im ersten und konj. Lin. im zweiten Argument gilt.)

∘ Unitäre Vektorräume sind für die Quantentheorie zentral.

∘ Für die Norm gilt $\forall x \in V$: $\|x\| = 0 \Leftrightarrow x = 0$ und $\forall \lambda \in \mathbb{K}$: $\|\lambda x\| = |\lambda| \cdot \|x\|$ (absolute Homogenität)

∘ Über allgemeine Körper sind Skalarprodukte nicht definiert. z.B. macht "$\geq$" i. A. keinen Sinn.

Bsp.:

- $V = \mathbb{R}^n$ mit $\langle x, y \rangle := \sum_{i=1}^{n} x_i \cdot y_i$
- $V = \mathbb{C}^n$ mit $\langle x, y \rangle := \sum_{i=1}^{n} \overline{x_i} \cdot y_i$

'Standardskalarprodukt'

d.h. $\langle x, y \rangle = \overline{x}^T \cdot y$ als Matrixprodukt

- $V = \mathbb{R}^{n \times n}$ mit $\langle A, B \rangle := \operatorname{tr}[A^T B]$
- $V = C([0,1])$ mit $\langle f, g \rangle := \int_0^1 \overline{f(x)}\, g(x)\, dx$

(würden wir statt Stetigkeit Integrierbarkeit fordern, wäre dies kein Skalarprodukt, da dann $\langle f, f \rangle = 0$ für $f \neq 0$ sein kann.)

Satz: Ist V ein Vektorraum mit Skalarprodukt, dann gilt $\forall x, y \in V$:

(i)	$\lvert \langle x, y \rangle \rvert \leq \lVert x \rVert \cdot \lVert y \rVert$	Cauchy-Schwarz Ungl.
(ii)	$\lVert x + y \rVert \leq \lVert x \rVert + \lVert y \rVert$	Dreiecksungleichung
(iii)	$\lVert x + y \rVert^2 + \lVert x - y \rVert^2 = 2(\lVert x \rVert^2 + \lVert y \rVert^2)$	Parallelogrammgl.

Beweis: (i) Wir können x durch λx mit $\lvert \lambda \rvert = 1$ ersetzen, so dass o.B.d.A. $\langle x, y \rangle \in \mathbb{R}$. Dann gilt $\forall t \in \mathbb{R}$:

$$0 \leq \lVert tx + y \rVert^2 = t^2 \lVert x \rVert^2 + 2t \langle x, y \rangle + \lVert y \rVert^2$$

Minimieren bzgl. t liefert $t_{min} = -\frac{\langle x, y \rangle}{\lVert x \rVert^2}$ und nach Einsetzen:

$$-\frac{\langle x, y \rangle^2}{\lVert x \rVert^2} + \lVert y \rVert^2 \geq 0 .$$

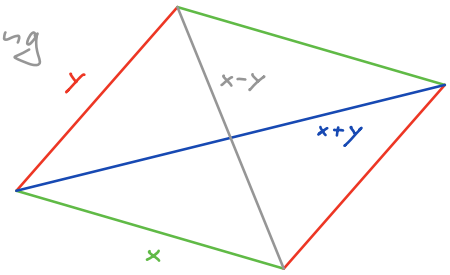
(ii)
$$\begin{aligned} \lVert x + y \rVert^2 &\leq \lVert x \rVert^2 + \lVert y \rVert^2 + 2 \lvert \langle x, y \rangle \rvert \\ &\overset{(i)}{\leq} \lVert x \rVert^2 + \lVert y \rVert^2 + 2 \lVert x \rVert \cdot \lVert y \rVert \\ &= (\lVert x \rVert + \lVert y \rVert)^2 \end{aligned}$$

(iii)
$$\begin{aligned} \lVert x + y \rVert^2 + \lVert x - y \rVert^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= \lVert x \rVert^2 + \lVert y \rVert^2 + \langle x, y \rangle + \langle y, x \rangle + \lVert x \rVert^2 + \lVert y \rVert^2 - \langle x, y \rangle - \langle y, x \rangle \\ &= 2(\lVert x \rVert^2 + \lVert y \rVert^2) \end{aligned}$$

$\square$

Bem.:

- Interpretation / Namensgebung der Parallelogrammgl.:



  Die Gleichung kann als Verallgemeinerung des Satzes von Pythagoras betrachtet werden.

- In einem Euklid'schen Vektorraum erlaubt die C.S. Ungl. die Definition eines Winkels $\theta \in [0, \pi]$, s.d. $\boxed{\cos(\theta) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}}$

- Gleichheit in der Cauchy-Schwarz Ungl. gilt g.d.w. sich x und y nur um einen skalaren Faktor unterscheiden (D.h. $x = \lambda y$ für ein $\lambda \in \mathbb{K}$ oder $y = 0$).

Def.: Sei V ein Vektorraum mit Skalarprodukt, $M \subseteq V$, $x, y \in V$.

- x heißt orthogonal zu y, wenn $\langle x, y \rangle = 0$ ($\Leftrightarrow: x \perp y$)
- $M^\perp := \{ v \in V \mid \forall w \in M : \langle v, w \rangle = 0 \}$ heißt Orthogonalkomplement von M.

Bem.:

- Für $\mathbb{K} = \mathbb{R}$ liefert $\cos(\theta) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$, $\theta \in [0, \pi]$, dass $x \perp y \Leftrightarrow \theta = \frac{\pi}{2}$.
- Es gilt $\{0\}^\perp = V$ und $V^\perp = \{0\}$.

Korollar (Pythagoras): $x \perp y \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$

Beweis: $\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \|y\|^2 + \overbrace{\langle x, y\rangle}^{=0} + \overbrace{\langle y, x\rangle}^{=0}$

$\square$

Korollar: Für $M \subseteq V$ ist $M^\perp$ ein Untervektorraum von V.

Beweis: Es gilt $\varphi \in M^\perp \Rightarrow \lambda\varphi \in M^\perp$ und $\varphi, \psi \in M^\perp \Rightarrow \varphi + \psi \in M^\perp$.

$\square$

Def.: Sei V ein endl.-dim. Vektorraum mit Skalarprodukt. Eine Menge von Vektoren $v_1, \dots, v_n \in V \setminus \{0\}$ heißt

- orthogonal, wenn $v_i \perp v_j$ für alle $i \neq j$,
- orthonormal, wenn $\langle v_i, v_j\rangle = \delta_{ij}$,
- Orthonormalbasis (ONB), wenn sie orthonormal sind und eine Basis von V bilden.

Bsp.:

- Die Standardbasis in $\mathbb{K}^n$ bildet eine ONB bzgl. des Standardskalarprodukts.

Bem.:

- Jede orthogonale Menge von Vektoren ist linear unabh. Denn ist $0 = \sum_{i=1}^n \lambda_i v_i$, dann auch das Skalarprodukt mit v_k, also $0 = \langle v_k, \sum_i \lambda_i v_i\rangle = \sum_i \lambda_i \langle v_k, v_i\rangle = \lambda_k \|v_k\|^2$

- Bzgl. ONBs lassen sich Entwicklungskoeffizienten besonders einfach bestimmen:

Satz: Ist $(v_1, \dots, v_n)$ eine ONB von V, dann gilt für alle $x, y \in V$:

$$\boxed{\begin{aligned} &(i) \quad x = \sum_{i=1}^{n} \langle v_i, x \rangle v_i \\ &(ii) \quad \|x\|^2 = \sum_{i=1}^{n} |\langle v_i, x \rangle|^2 \\ &(iii) \quad \langle x, y \rangle = \underbrace{\sum_{i=1}^{n} \overline{\langle v_i, x \rangle} \langle v_i, y \rangle}_{\text{Standardskalarprodukt der Entwicklungskoeffizienten}} \end{aligned}}$$

Parseval Identität

Beweis: (i) Wegen der Basiseigenschaft existieren $\lambda_i \in \mathbb{K}$, s.d. $x = \sum_{j=1}^{n} \lambda_j v_j$. Skalarprodukt mit v_i ergibt dann: $\langle v_i, x \rangle = \lambda_i$.

(iii) Durch Einsetzen von (i) für x und y:

$$\langle x, y \rangle = \sum_{ij} \overline{\langle v_i, x \rangle} \langle v_j, y \rangle \underbrace{\langle v_i, v_j \rangle}_{= \delta_{ij}} = \sum_{i=1}^{n} \overline{\langle v_i, x \rangle} \langle v_i, y \rangle$$

(ii) Folgt aus (iii) für $x = y$. $\square$

Bem.: Für die Darstellungsmatrix von $A: V \to V$ gilt bzgl. der ONB $(v_i)_{i=1}^{n}$ mit $Av_i =: \sum_j A_{ji} v_j$ dann $\boxed{A_{ij} = \langle v_i, A v_j \rangle}$.

Das Gram-Schmidt Orthonormalisierungsverfahren

Input: Linear unabh. Vektoren $u_1, \dots, u_n$ eines Vekt.raums V mit Skalarprodukt.

Output: Eine ONB $v_1, \dots, v_n$ von $\operatorname{span}(u_1, \dots, u_n)$.

Für $i = 1, \dots, n$ definieren wir:

$$\boxed{v_i := \frac{u_i - \sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j}{\left\| u_i - \sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j \right\|}}$$

Offensichtlich gilt $v_i \in \operatorname{span}\{u_1, \dots, u_i\}$ und insbesondere $\sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j \in \operatorname{span}\{u_1, \dots, u_{i-1}\}$. Wegen lin. Unabh. der u_i's ist die Norm im Nenner also $\neq 0$.

Per Konstruktion gilt:

(i) $\|v_i\| = 1$

(ii) $\langle v_k, v_i \rangle = 0$ für alle $i \neq k$

Es genügt dies für alle $k < i$ zu zeigen.

Induktionsannahme: Orthogonalität gilt für alle $k, i \leq m$

Für $i = m+1$ gilt dann:

$$\langle v_k, v_i \rangle \sim \langle v_k, u_i \rangle - \sum_{j=1}^{i-1} \langle v_j, u_i \rangle \underbrace{\langle v_k, v_j \rangle}_{= \delta_{kj} \text{ da } k, j \leq m} = \langle v_k, u_i \rangle - \langle v_k, u_i \rangle = 0.$$

(iii) $\operatorname{span}\{v_1, \dots, v_i\} = \operatorname{span}\{u_1, \dots, u_i\} \quad \forall i \in \{1, \dots, n\}$

"$\subseteq$" ist klar. Da $v_1, \dots, v_i$ aber orthogonal und damit lin. unabhängig sind, müssen die Dimensionen und damit die Räume gleich sein.

<u>Bem.:</u> ∘ Sind $u_1, \dots, u_k$ bereits orthonormal, dann gilt $v_i = u_i \ \forall i \leq k$.

Korollar: Für jeden endl.-dim. Vektorraum V mit Skalarprodukt gilt: (i) V besitzt eine ONB

(ii) Jede orthonormale Menge von Vektoren $v_1, \dots, v_k \in V$ lässt sich zu einer ONB vervollständigen.

Beweis: (i) Aus jeder Basis von V wird durch Gram-Schmidt eine ONB.

(ii) $v_1, \dots, v_k$ bilden eine ONB von $M := \operatorname{span}\{v_1, \dots, v_k\}$. Diese können wir zu einer Basis $v_1, \dots, v_k, b_{k+1}, \dots, b_n$ von V vervollständigen aus der durch Gram-Schmidt eine ONB $v_1, \dots, v_n$ wird, wobei sich die ersten k Vektoren nicht ändern. $\square$

Satz: Sei U ein Untervektorraum eines endl.-dim. Vektorraums V mit Skalarprodukt. Dann gilt:

(i) $V = U \oplus U^\perp$,

(ii) $(U^\perp)^\perp = U$.

Beweis: (i) Wir müssen zeigen, dass (a) $U + U^\perp = V$ und (b) $U \cap U^\perp = \{0\}$. Sei $v_1, \dots, v_k$ eine ONB von U, die wir zu einer ONB $v_1, \dots, v_n$ von V vervollständigen. Dann gilt für jedes $x \in V$:

$$x = \underbrace{\sum_{i=1}^{k} \langle v_i, x \rangle v_i}_{\in U} + \underbrace{\sum_{j=k+1}^{n} \langle v_j, x \rangle v_j}_{\in U^\perp}.$$

Also $V = U + U^\perp$.

Ist $v \in U \cap U^\perp$, dann muß $v \perp y$ $\forall y \in U$ also auch $\langle v, v \rangle = 0$. Demnach ist $U \cap U^\perp = \{0\}$.

(ii) Ist $u \in U$, gilt für alle $v \in U^\perp$: $\langle u, v \rangle = 0$. Also $u \in (U^\perp)^\perp$ und damit $U \subseteq (U^\perp)^\perp$. Gleichheit gilt, wenn beide dieselbe Dimension besitzen. Dies folgt aber aus (i), da $V = U \oplus U^\perp$ und $V = (U^\perp)^\perp \oplus U^\perp$. $\square$

<u>Def.:</u> Sei U ein Untervektorraum eines endl.-dim. Vektorraums V mit Skalarprodukt. Die Orthogonalprojektion auf U ist definiert als Lineare Abbildung $P_U : V \to V$, so dass $P_U|_U = \mathrm{id}_U$ und $\ker(P_U) = U^\perp$.

<u>Bem.:</u>

- Wegen $V = U \oplus U^\perp$ lässt sich jedes $x \in V$ eindeutig als $x = u + v$ mit $u \in U$, $v \in U^\perp$ schreiben. Dann ist $\boxed{P_U x = u}$ also P_U eindeutig definiert.
- P_U ist eine Projektion, was bedeutet, dass $P_U^2 = P_U$. 'Orthogonalprojektion' bedeutet, dass zusätzlich gilt: $\mathrm{Bild}(P_U) \perp \ker(P_U)$.
- Ist $u_1, \dots, u_d$ ONB von U, dann gilt $\forall x \in V$:
$$P_U x = \sum_{i=1}^{d} \langle u_i, x \rangle u_i .$$
- Die Orthogonalproj. auf $U^\perp$ ist $P_{U^\perp} = \mathrm{id}_V - P_U$.

Satz: Sei U ein Untervektorraum eines endl.-dim. Vektorraums V mit Skalarprodukt und $v \in V$. Dann gilt

$$\boxed{\min_{u \in U} \|v-u\| = \|v - P_U v\|}$$

und $P_U v$ ist das einzige Element von U für das '=' gilt.

Beweis: Sei $P := P_U$ und $u \in U$. Dann gilt:

$$\|v - Pv\|^2 \leq \|\overbrace{v-Pv}^{\in U^\perp}\|^2 + \|\overbrace{Pv-u}^{\in U}\|^2$$

$$\underset{\text{Pythagoras}}{=} \|(v-Pv) + (Pv-u)\|^2 = \|v-u\|^2$$

mit Gleichheit g.d.w. $\|Pv-u\| = 0$ also $u = Pv$. $\square$

Def.: Sind V, W zwei $\mathbb{K}$-Vektorräume mit Skalarprodukt, dann heißt eine lineare Abbildung $F: V \to W$ für die gilt $\forall x, y \in V: \langle F(x), F(y)\rangle = \langle x, y\rangle$ orthogonal (wenn $\mathbb{K} = \mathbb{R}$) bzw. unitär (wenn $\mathbb{K} = \mathbb{C}$).

In beiden Fällen spricht man von einer Isometrie (was bedeutet, dass $\forall x: \|F(x)\| = \|x\|$).

Ist $V = W = \mathbb{K}^n$, versehen mit dem Standardskalarprodukt, werden die Mengen der durch Matrizen gegebenen orthogonalen bzw. unitären Abbildungen $\mathbb{K}^n \to \mathbb{K}^n$ mit $O(n) \subseteq \mathbb{R}^{n \times n}$ bzw. $U(n) \subseteq \mathbb{C}^{n \times n}$ bezeichnet.

Bem.: Die Darstellungsmatrizen von orth./unit. Abbildungen $V \to V$ bzgl. einer ONB sind orth./unit. Matrizen.

Satz: (i) $O(n) = \{ M \in \mathbb{R}^{n\times n} \mid M^T M = I_n \}$

$$= \{ M = (v_1 \,\vdots\, \cdots \,\vdots\, v_n) \in \mathbb{R}^{n\times n} \mid v_1, \dots, v_n \text{ ist ONB} \}$$

(ii) $U(n) = \{ U \in \mathbb{C}^{n\times n} \mid \overline{U}^T U = I_n \}$

$$= \{ U = (v_1 \,\vdots\, \cdots \,\vdots\, v_n) \in \mathbb{C}^{n\times n} \mid v_1, \dots, v_n \text{ ist ONB} \}$$

Beweis: Wir zeigen (ii). (i) ist analog bzgl. des St.skalarprodukts in $\mathbb{R}^n$.

$$U \in U(n) \iff \forall x, y \in \mathbb{C}^n : \langle x, y \rangle = \langle Ux, Uy \rangle$$

$$= \sum_{ijk} \overline{U}_{ik} \overline{x}_k U_{ij} y_j$$

$$= \langle x, \overline{U}^T U y \rangle$$

$$\iff \overline{U}^T U = I_n$$

$$\iff \langle v_i, v_j \rangle = \delta_{ij} \quad \forall i, j \qquad \square$$

Bsp.:
- $O(2) = \left\{ \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \;\middle|\; \varphi \in [0, 2\pi) \right\}$
- $U(1) = \{ e^{i\varphi} \mid \varphi \in \mathbb{R} \}$

Bem.:
- Nach dem Satz sind unitäre/orthogonale Matrizen Isomorphismen mit $U^{-1} = \overline{U}^T$ bzw. $O^{-1} = O^T$.
- Wegen $1 = \det(I_n) = \det(\overline{U}^T) \cdot \det(U) = |\det(U)|^2$ gilt
  $U \in U(n) \Rightarrow |\det(U)| = 1$
  $O \in O(n) \Rightarrow \det(O) \in \{-1, 1\}$

Es gilt sogar folgende, etwas stärkere Aussage:

Satz: Ist $U: V \to V$ unitär, dann gilt $\boxed{\lambda \in \operatorname{spec}(U) \Rightarrow |\lambda| = 1}$.

Beweis: Mit $Ux = \lambda x$ gilt: $\|x\|^2 = \langle x, x\rangle = \langle Ux, Ux\rangle = \langle \lambda x, \lambda x\rangle = |\lambda|^2 \|x\|^2$. $\square$

Bem.:
- Da die definierende Eigenschaft von $O(n)$ bzw. $U(n)$ unter Komposition (Matrixprodukt) erhalten bleibt, handelt es sich hierbei um Gruppen.

Def.: (Matrixgruppen) Bzgl. der Matrixmultiplikation sind folgende Gruppen definiert:

- Allgemeine Lineare Gruppe $GL(n, \mathbb{K}) := \{ M \in \mathbb{K}^{n \times n} \mid \det(M) \neq 0 \}$ (Invertierbare Matrizen)
- Spezielle Lineare Gruppe $SL(n, \mathbb{K}) := \{ M \in \mathbb{K}^{n \times n} \mid \det(M) = 1 \}$
- Orthogonale Gruppe $O(n)$ Drehungen & Spiegelungen
- Spezielle orth. Gruppe $SO(n) := \{ M \in O(n) \mid \det(M) = 1 \}$ (Drehungen)
- Unitäre Gruppe $U(n)$
- Spezielle unitäre Gruppe $SU(n) := \{ M \in U(n) \mid \det(M) = 1 \}$
- Lorentzgruppe $O(1,3) := \left\{ M \in \mathbb{R}^{4 \times 4} \mid M g M^T = g,\ g := \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \right\}$
- ...

Dabei gelten folgende Untergruppenbeziehungen:

$$\begin{array}{ccccc} GL(n, \mathbb{C}) & \supseteq & U(n) & \supseteq & SU(n) \\ \cup\!\!\mid & & \cup\!\!\mid & & \cup\!\!\mid \\ GL(n, \mathbb{R}) & \supseteq & O(n) & \supseteq & SO(n) \end{array}$$

Def.:

- Sind $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ zwei Basen von $\mathbb{R}^n$ und $S_{B,C}$ die zugehörige Basiswechselmatrix (d.h. $S_{B,C} : c_i \mapsto b_i$), dann sagt man, B und C haben dieselbe Orientierung wenn $\det(S_{B,C}) > 0$.
- $M \in GL(n, \mathbb{R})$ heißt
  orientierungserhaltend wenn $\det(M) > 0$,
  orientierungsumkehrend wenn $\det(M) < 0$.

Bem.:

- Damit lassen sich alle Basen des $\mathbb{R}^n$ in zwei Äquivalenzklassen einteilen, die "positiv / negativ orientiert" heißen.
- Alle Elemente von $SO(n)$ sind orientierungserhaltend.

IX. Matrixzerlegungen

Def.: Sei V ein Vektorraum mit Skalarprodukt und $A: W \to V$ eine lineare Abbildung. Eine lineare Abbildung $A^*: V \to W$ heißt zu A adjungiert, wenn für alle $v \in V$, $w \in W$:

$$\boxed{\langle A^* v, w \rangle = \langle v, A w \rangle}$$

Sei nun $W = V$.

- A heißt selbstadjungiert (oder hermitesch), wenn $A^* = A$.
- A heißt normal, wenn $A^* A = A A^*$.

Bem.:
- Sind für eine lineare Abbildung $T: V \to W$ alle Skalarprodukte $\langle Tv, w \rangle$ gegeben, ist T eindeutig bestimmt (da man v, w insbesondere aus einer Basis nehmen kann).
- In der (Quanten-) Physik schreibt man auch $A^\dagger$ statt A^* (dann gesprochen "A dagger").

Korollar: Für eine Matrix $A \in \mathbb{C}^{n \times m}$, aufgefasst als lin. Abb. $\mathbb{C}^m \to \mathbb{C}^n$, gilt bzgl. des Standardskalarprodukts: $\boxed{A^* = \overline{A}^T}$

Beweis: Mit der Standardbasis $e_1, e_2, \ldots$ gilt:

$$(A^*)_{ij} = \langle e_i, A^* e_j \rangle = \overline{\langle A^* e_j, e_i \rangle} = \overline{\langle e_j, A e_i \rangle} = \overline{A_{ji}} . \quad \square$$

Bem.:
- Für jede unitäre Matrix $U \in \mathbb{C}^{n \times n}$ gilt also $U^* = U^{-1}$.
- Unitäre in $U(n)$ sind normal, da wegen $U^* U = I_n$ U^* die (recht & links-) Inverse ist, so dass auch $U U^* = I_n$ gilt und damit $U^* U = U U^*$.

- Für Matrizen kann man $A^* := \overline{A^T}$ als Definition betrachten.
- Für $A \in \mathbb{R}^{n \times n}$ ist dann $A^* = A^T$. In diesem Fall, nennt man selbstadjungierte A symmetrisch.
- Selbstadjungiert impliziert normal, da $A^*A = AA = AA^*$.
- Aus der Definition der adjungierten Abbildung zeigt man:

(i)	$(A+\lambda B)^* = A^* + \overline{\lambda} B^*$		(iv)	$(A^*)^* = A$
(ii)	$id^* = id$		(v)	$(A^*)^{-1} = (A^{-1})^*$
(iii)	$(AB)^* = B^* A^*$			

(vi) Für die Matrixdarstellung $M(A)$ bzgl. einer ONB gilt: $M(A^*) = M(A)^*$.

<u>Satz:</u> Ist $A: V \to V$ selbstadjungiert, dann gilt $\operatorname{spec}(A) \subseteq \mathbb{R}$.

<u>Beweis:</u> Sei $Av = \lambda v$ mit $v \neq 0$. Dann gilt:

$$\lambda \|v\|^2 = \langle v, \lambda v \rangle = \langle v, Av \rangle = \langle A^* v, v \rangle \overset{A = A^*}{=} \langle Av, v \rangle = \overline{\lambda} \|v\|^2. \quad \square$$

<u>Bem.:</u> In der Quantentheorie sind 'Observable', welche Messungen und deren Erwartungswerte beschreiben, selbstadjungiert, da deren Spektrum die möglichen Meßergebnisse beinhaltet und die zug. Größen, wie Ort, Impuls, Drehimpuls, Energie, etc., reell sind.

<u>Bsp.:</u>
- $\begin{pmatrix} 4 & -i & 2 \\ i & 0 & 1+2i \\ 2 & 1-2i & -1 \end{pmatrix}$ ist selbstadjungiert und hat daher ausschl. reelle Eigenwerte.
- Die Pauli-Matrizen $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ sind selbstadjungiert und unitär mit Eigenwerten ± 1. Zusammen mit $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ bilden sie eine Basis des $\mathbb{R}$-Vektorraums der selbstadjungierten 2×2 Matrizen.

Satz: Sei V ein endl.-dim Vektorraum mit Skalarprodukt und $P: V \to V$ linear, so dass $P^2 = P$. Dann gilt mit $U := PV$:
P ist genau dann eine Orthogonalprojektion auf U, wenn $P = P^*$.

Beweis: Da $P|_U = \mathrm{id}_U$ schon aus $P^2 = P$ folgt, ist noch z.z., dass $\ker(P) \perp U \iff P = P^*$.

$$\left(\begin{aligned} x \in U &\Rightarrow \exists y \in V: Py = x \\ &\Rightarrow Px = PPy \underset{P^2=P}{=} Py = x \end{aligned} \right)$$

"$\Rightarrow$": Sei $v_1, \dots, v_n$ ONB von U, s.d. $\forall x \in V: Px = \sum_{i=1}^{n} \langle v_i, x \rangle v_i$. Dann ist

$$\langle y, Px \rangle = \sum_i \langle y, v_i \rangle \langle v_i, x \rangle = \langle Py, x \rangle \quad \forall x, y \in V.$$

"$\Leftarrow$": Sei $x \in \ker(P)$ und $y \in U$, also $Px = 0$ und $y = Py$. Dann gilt:

$$\langle x, y \rangle = \langle x, Py \rangle = \langle P^* x, y \rangle \underset{P = P^*}{=} \langle Px, y \rangle = 0. \qquad \square$$

Lemma: Sei V ein Vektorraum mit Skalarprodukt und $A: V \to V$ linear und normal. Dann gilt:

(i) $\lambda \in \mathrm{spec}(A) \Rightarrow \underbrace{\ker(A - \lambda \mathrm{id}_V)}_{=: E_\lambda(A)} = \underbrace{\ker(A^* - \bar{\lambda} \mathrm{id}_V)}_{E_{\bar{\lambda}}(A^*)}$

(ii) Sind $v, w \in V$ Eigenvektoren zu verschiedenen Eigenwerten λ, μ von A, dann ist $v \perp w$.

(iii) Ist $x \in V$ orthogonal zu jedem Eigenvektor von A, dann besitzt auch Ax diese Eigenschaft.

Beweis: (i) $v \in E_\lambda(A) \iff Av = \lambda v$.

$$\langle v, \bar{\lambda} A v \rangle = \langle v, \bar{\lambda} \lambda v \rangle = \langle \bar{\lambda} v, \bar{\lambda} v \rangle$$

$$\| A^* v - \bar{\lambda} v \|^2 = \underbrace{\langle A^* v, A^* v \rangle}_{\langle v, AA^* v \rangle = \langle v, A^* A v \rangle = \langle v, \lambda A^* v \rangle = \langle \bar{\lambda} v, A^* v \rangle} - \langle A^* v, \bar{\lambda} v \rangle - \langle \bar{\lambda} v, A^* v \rangle + \langle \bar{\lambda} v, \bar{\lambda} v \rangle = 0$$

D.h. $v \in E_{\bar{\lambda}}(A^*)$ und damit $E_\lambda(A) \subseteq E_{\bar{\lambda}}(A^*)$. Durch Anwenden auf $\bar{\lambda}$ und A^* gilt dann
$E_{\bar{\lambda}}(A^*) \subseteq E_\lambda(A^{**}) = E_\lambda(A)$.

(ii) $(\lambda - \mu)\langle v, w\rangle = \langle \bar{\lambda} v, w\rangle - \langle v, \mu w\rangle \overset{(i)}{=} \langle A^* v, w\rangle - \langle v, A w\rangle$
$= \langle v, A w\rangle - \langle v, A w\rangle = 0$

(iii) $Av = \lambda v \neq 0 \Rightarrow \langle v, Ax\rangle = \langle A^* v, x\rangle \overset{(i)}{=} \lambda \langle v, x\rangle = 0.$ $\square$

<u>Satz:</u> (Spektralsatz für normale Abbildungen)
Ist V ein endl.-dim. unitärer Raum und $A: V \to V$ linear und normal. Dann existiert für V eine ONB von Eigenvektoren von A.

<u>Beweis:</u> Sei $\operatorname{spec}(A) = \{\lambda_1, \dots, \lambda_r\}$ und B_i eine ONB für den Eigenraum E_{λ_i}. Nach dem Lemma (ii) ist $\bigcup_{i=1}^{r} B_i =: \{v_1, \dots, v_n\}$ orthonormal, also eine ONB des von den Eigenvektoren v_i aufgespannten Unterraums $L := \operatorname{span}\{v_1, \dots, v_n\}$.
Nach dem Lemma (iii) ist aber $A|_{L^\perp} : L^\perp \to L^\perp$.
Wenn $L^\perp \neq \{0\}$, müsste dies, wie jede lineare Abb. auf einem nicht-trivialen $\mathbb{C}$-Vektorraum, einen Eigenvektor besitzen. Da $L^\perp$ aber orthogonal auf allen Eigenvektoren ist, muß $L^\perp = \{0\}$ also $L = V$ gelten. $\square$

Korollar: (Spektralzerlegung normaler Matrizen)

Für $A \in \mathbb{C}^{n\times n}$ sind äquivalent:

(i) A ist normal.

(ii) Es existiert ein $U \in SU(n)$ und eine Diagonalmatrix D, s.d.

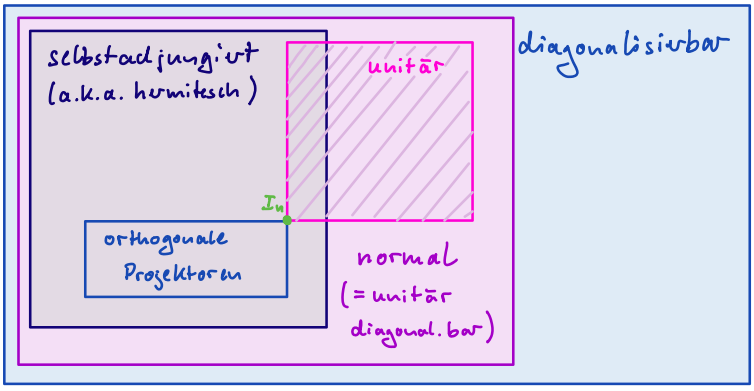
$$\boxed{A = UDU^*}.$$

Beweis: (i) $\Rightarrow$ (ii): Nach dem Spektralsatz ist A diagonalisierbar, also $A = \tilde{U} D \tilde{U}^{-1}$ für ein $\tilde{U} \in GL(n, \mathbb{C})$, und die Eigenvektoren, die wir normiert annehmen können und in den Spalten von $\tilde{U}$ stehen, bilden eine ONB. Damit ist $\tilde{U} \in U(n)$, so dass wir $U := \tilde{U} \cdot \mathrm{diag}(\det(\tilde{U})^{-1}, 1, \dots, 1) \in SU(n)$ wählen können.

(ii) $\Rightarrow$ (i)
$$A^*A = (UDU^*)^* UDU^*$$
$$= U^{**} D^* U^* U D U^* = U D^* D U^*$$
$$AA^* = UDU^*(UDU^*)^*$$
$$= UDU^* U^{**} D^* U^* = U D D^* U^*$$

Diese sind gleich, da für diagonale Matrizen gilt: $D^*D = DD^*$. $\square$

Das Bild in $\mathbb{C}^{n\times n}$ ist also folgendes:



Korollar: Jede normale Matrix $A \in \mathbb{C}^{n \times n}$ lässt sich schreiben als

$$\boxed{A = \sum_{i=1}^{r} \lambda_i P_i},$$

wobei $\operatorname{spec}(A) = \{\lambda_1, \dots, \lambda_r\}$ und P_i die Orthogonalprojektion auf E_{λ_i} darstellt.

Bem.: Die P_i's heißen dann Spektralprojektoren.

Beweis: Es genügt z.z., dass die beiden Seiten dieselbe Wirkung auf eine ONB haben. Wir wählen dazu eine ONB aus Eigenvektoren $v_1, \dots, v_n$. Dann gilt $A v_j = \lambda(v_j) v_j$ und

$$\sum_{i=1}^{r} \lambda_i P_i v_j = \lambda(v_j) v_j .$$

Eigenräume zu untersch. Eigenwerten sind orthogonal und $P_i|_{E_{\lambda_i}} = \mathrm{id}_{E_{\lambda_i}}$ sowie $\ker(P_i) \perp E_{\lambda_i}$. $\square$

Bem.: Sind $v_1, \dots, v_k \in \mathbb{C}^n$ orthonormal und stellt $P \in \mathbb{C}^{n \times n}$ die Orthogonalprojektion auf $\operatorname{span}\{v_1, \dots, v_k\}$ dar, dann ist

$$\boxed{P = \sum_{i=1}^{k} v_i v_i^*},$$

da das 'Matrixprodukt' $v_i^* v_j = \langle v_i, v_j \rangle = \delta_{ij}$ ist.

D.h., wenn $v_1, \dots, v_n$ eine ONB von Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_n$ ist, dann gilt:

$$\boxed{A = \sum_{i=1}^{n} \lambda_i v_i v_i^*}.$$

Zusammengefasst:

$$\mathbb{C}^{n\times n} \ni A \overset{A \text{ normal}}{=} \underset{\in SU(n)}{U}\, \overset{\text{diag.}}{D}\, U^* = \sum_{i=1}^{n} \lambda_i\, v_i v_i^* = \sum_{\lambda \in \operatorname{spec}(A)} \lambda\, P_{E_\lambda}$$

$$D = \sum_{i=1}^{n} \lambda_i e_i e_i^*$$

$$U e_i = v_i$$

„Spektralproj."

Über $\mathbb{R}$ wissen wir bereits, dass sich nicht jede normale Matrix diagonalisieren lässt, da z.B. $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ für $\varphi \in (0,\pi)$ zwar normal ist, aber als Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ keinen Eigenwert besitzt. Wie der folgende Satz zeigt, ist dies aber i.W. schon die einzige Ausnahme:

Satz: (Spektralzerlegung für reelle normale Matrizen)

Ist $A \in \mathbb{R}^{n\times n}$ normal, dann existiert $X \in O(n)$, so dass XAX^T folgende Form besitzt:

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_r & & & & \\ & & & \boxed{\begin{matrix} a_1 & -b_1 \\ b_1 & a_1 \end{matrix}} & & & \\ & & & & \ddots & & \\ & & & & & \boxed{\begin{matrix} a_s & -b_s \\ b_s & a_s \end{matrix}} \end{pmatrix}$$

wobei $\lambda_i, a_s, b_k \in \mathbb{R}$ und $r, s \in \mathbb{N}_0$.

Beweis: $\longrightarrow$ z.B. [Kemper, Reimers].

Bem.:

- Ist $A \in O(n)$, dann gilt offensichtlich, dass $\lambda_i \in \{\pm 1\}$ und $a_i^2 + b_i^2 = 1$.

- Ist A symmetrisch, d.h. $A^T = A \in \mathbb{R}^{n \times n}$, bedeutet dies, dass $s = 0$ sein muss. Demnach ist XAX^T diagonal. Analog zum komplexen Fall, können wir stets $X \in SO(n)$ wählen.

Def.: Ist $A = A^T \in \mathbb{R}^{n \times n}$, dann heißt $\beta: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ $\beta(x,y) := \langle x, Ay \rangle$ symmetrische Bilinearform und $q: \mathbb{R}^n \to \mathbb{R}$, $q(x) := \langle x, Ax \rangle$ quadratische Form. Für $c \in \mathbb{R}$ heißen die Niveauflächen $q^{-1}(\{c\}) = \{ x \in \mathbb{R}^n \mid q(x) = c \}$ Quadriken.

Bem.: I.A. ist eine 'Quadrik' die Lösungsmenge einer quadratischen Gleichung im $\mathbb{R}^n$.
Bsp.: Ellipse, Hyperbel, Parabel, Zylinder

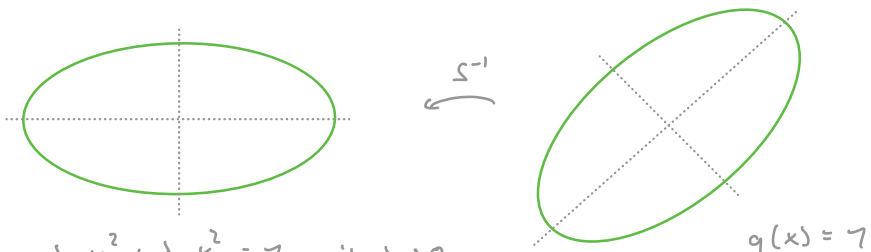
Korollar: (Hauptachsentransformation)
Ist $q: \mathbb{R}^n \to \mathbb{R}$ eine quadratische Form, dann existiert $\lambda \in \mathbb{R}^n$ und $S \in SO(n)$, so dass $\forall x \in \mathbb{R}^n$:

$$\boxed{q(Sx) = \sum_{i=1}^{n} \lambda_i x_i^2}$$

Beweis: Ist $q(x) = \langle x, Ax \rangle$ mit $A = A^T \in \mathbb{R}^{n \times n}$, dann können wir A orthogonal diagonalisieren: $A = SDS^T$ mit $S \in SO(n)$ und $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Dann ist

$$q(Sx) = \langle Sx, ASx \rangle = \langle Sx, SD\underbrace{S^T S}_{= I_n} x \rangle = \langle x, Dx \rangle$$
$$= \sum_i \lambda_i x_i^2 . \qquad \square$$

Bem.: ∘ Die orthogonale Transformation S^{-1} dreht die zugehörigen Quadriken, die je nach Vorzeichen der λ_i's z.B. Hyperbeln oder Ellipsen sind, auf die 'Hauptachsen'.



$\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$ mit $\lambda_i > 0$ $\qquad$ $q(x) = 1$

Anwendungen: ∘ Analysis 2:

Mehrdim. Taylorentwicklung von $f \in C^3(\mathbb{R}^n, \mathbb{R})$:

$$f(x) = f(0) + \langle \nabla f(0), x \rangle + \tfrac{1}{2} \langle x, H x \rangle + O(\|x\|^3)$$

mit Hessematrix $H = H^T$, $H_{ij} := \partial_i \partial_j f(0)$.

Ein lokales Extremum ($\nabla f(0) = 0$) ist durch die Vorzeichen der Eigenwerte von H charakterisiert.

∘ Mechanik: $\underset{\substack{\uparrow \\ \text{Drehimpuls} \\ L \in \mathbb{R}^3}}{L} = \underset{\substack{\uparrow \\ \text{Trägheitstensor}}}{\Theta} \cdot \overset{\substack{\text{Winkelgeschw.} \\ \downarrow}}{\omega}$ $\qquad$ (analog zu $p = m \cdot v$)

$\Theta = \Theta^T \in \mathbb{R}^{3 \times 3}$

Die Hauptachsentrafo. dreht die 'Hauptträgheitsachsen' auf die Standard-ONB und liefert die 'Hauptträgheitsmomente' als Eigenwerte von Θ.

Def.: Eine selbst-adjungierte Matrix $A \in \mathbb{C}^{n\times n}$ heißt

- positiv semi-definit ($A \geq 0$), wenn $\langle x, Ax\rangle \geq 0 \quad \forall x \in \mathbb{C}^n$
- positiv definit ($A > 0$), wenn $\langle x, Ax\rangle > 0 \quad \forall x \in \mathbb{C}^n \setminus \{0\}$

Bem.:
- Mit umgekehrten Vorzeichen definiert man analog 'negativ' (semi-)definit.
- Sind $A, B \in \mathbb{C}^{n\times n}$ selbst-adjungiert, dann lässt sich eine 'partielle Ordnung' $A \geq B \;:\Leftrightarrow\; (A-B) \geq 0$ definieren.
- Aus $\langle x, Ax\rangle \geq 0 \;\; \forall x \in \mathbb{C}^n$ folgt bereits $A = A^*$ (dies gilt jedoch nicht über $\mathbb{R}$, da z.B. $\langle x, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x\rangle = 0 \;\; \forall x \in \mathbb{R}^2$).

Bsp.:
- Für bel. $A \in \mathbb{C}^{n\times m}$ gilt $A^*A \geq 0$, da $\langle x, A^*Ax\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0$.
- Für $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_i \geq 0 \;\forall i$ gilt $D \geq 0$, da $\langle x, Dx\rangle = \sum_{i=1}^{n} \lambda_i |x_i|^2 \geq 0$.

Satz: Für $A = A^* \in \mathbb{C}^{n\times n}$ mit $\operatorname{spec}(A) = \{\lambda_1, \dots, \lambda_r\}$ gilt:

$$A \geq 0 \iff \lambda_i \geq 0 \quad \forall i$$

$$A > 0 \iff \lambda_i > 0 \quad \forall i$$

Beweis: Wir verwenden die Spektralzerlegung $A = U^*DU$ mit $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Dann ist $\langle x, Ax\rangle = \langle x, U^*DUx\rangle = \langle Ux, DUx\rangle > 0 \;\; \forall x \neq 0$

$$\overset{U \text{ bijektiv}}{\iff} \langle y, Dy\rangle > 0 \;\; \forall y \neq 0. \quad \square$$

Bsp.: Stellt $P \in \mathbb{C}^{n\times n}$ eine Orthogonalprojektion dar, d.h. $P = P^2 = P^*$, dann gilt $P \geq 0$, da $\operatorname{spec}(P) \subseteq \{0, 1\}$.

Def.: Sei V ein Vektorraum mit Skalarprodukt. Die Gram-Matrix G eines Tupels von Vektoren $v_1, \dots, v_k \in V$ ist definiert über $\boxed{G_{ij} := \langle v_i, v_j \rangle}$.

Prop.: Eine Gram-Matrix ist stets positiv semidefinit und positiv definit g.d.w. die Vektoren linear unabh. sind.

Beweis:

$$\langle x, Gx \rangle = \sum_{ij} \bar{x}_i \langle v_i, v_j \rangle x_j = \langle \sum_i x_i v_i, \sum_j x_j v_j \rangle$$

$$= \| \sum_i x_i v_i \|^2 .$$

Also ist $G \geq 0$. Lin. Unabhängigkeit bedeutet, dass $\sum_i x_i v_i \neq 0$ für alle $x \neq 0$. Dies ist nun aber gleichbedeutend mit $\langle x, Gx \rangle > 0 \ \forall x \neq 0$. $\square$

Satz: Sei $\langle \cdot, \cdot \rangle_G$ ein beliebiges Skalarprodukt auf $\mathbb{K}^n$. Dann existiert eine positiv definite Matrix $G \in \mathbb{K}^{n \times n}$, s.d. $\forall x, y \in \mathbb{K}^n$:

$$\boxed{\langle x, y \rangle_G = \langle x, Gy \rangle}$$

Beweis: Sei $e_1, \dots, e_n$ eine ONB von $\mathbb{K}^n$ bzgl. $\langle \cdot, \cdot \rangle$ und $G_{ij} := \langle e_i, e_j \rangle_G$. Dann ist $G > 0$ und mit $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$ gilt: $\langle x, y \rangle_G = \sum_{ij} \bar{x}_i y_j \langle e_i, e_j \rangle_G$

$$= \langle x, Gy \rangle . \qquad \square$$

Umgekehrt ist $(x,y) \mapsto \langle x, Gy \rangle$ immer ein Skalarprodukt, wenn $G > 0$. D.h. die positiv definiten Matrizen charakterisieren genau alle Skalarprodukte in $\mathbb{K}^n$.

<u>Korollar:</u> Für jede Basis $x_1, \dots, x_n$ von $\mathbb{K}^n$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ existiert ein Skalarprodukt bzgl. dessen die Basis eine ONB ist.

<u>Beweis:</u> Sei $e_1, \dots, e_n$ die Standardbasis und $A : \mathbb{K}^n \to \mathbb{K}^n$ der Vektorraum-Isomorphismus, der $A : x_i \mapsto e_i \ \forall i$. Dann ist

$$\delta_{ij} = \langle e_i, e_j \rangle = \langle A x_i, A x_j \rangle = \langle x_i, \underbrace{A^* A}_{=: G} x_j \rangle$$

$$=: \langle x_i, x_j \rangle_G$$

wobei $G \geq 0$ und wegen $\det(A) \neq 0 \Rightarrow \det(G) \neq 0$ gilt auch $G > 0$. $\square$

Lemma: Eine selbstadjungierte Matrix $A \in \mathbb{C}^{2\times 2}$ ist genau dann positiv (negativ) semidefinit, wenn $\det(A) \geq 0$ und $\mathrm{tr}[A] \geq 0$ ($\leq$).

Beweis: Als selbstadjungierte Matrix ist A diagonalisierbar mit reellen Eigenwerten λ_1, λ_2. Es gilt $\det(A) = \lambda_1 \cdot \lambda_2$, $\mathrm{tr}[A] = \lambda_1 + \lambda_2$.

Also ist $\det(A) \geq 0 \iff \underbrace{\lambda_1, \lambda_2 \geq 0}_{\text{d.h. } \mathrm{tr}[A] \geq 0} \vee \underbrace{\lambda_1, \lambda_2 \leq 0}_{\text{d.h. } \mathrm{tr}[A] \leq 0}$. $\square$

Der Raum der selbstadjungierten 2×2 Matrizen ist ein 4-dim. $\mathbb{R}$-Vektorraum V. Ein Vektorraumisomorphismus $M: \mathbb{R}^4 \to V$ ist gegeben durch

$$M(x) := \begin{pmatrix} x_0 + x_3 & x_1 - i x_2 \\ x_1 + i x_2 & x_0 - x_3 \end{pmatrix} \quad \text{für} \quad x = (x_0, \dots, x_3) \in \mathbb{R}^4 .$$

Korollar: Für $x, y \in \mathbb{R}^4$ gilt:

$$(x_0 - y_0)^2 \geq \sum_{i=1}^{3} (x_i - y_i)^2 \iff \begin{array}{ll} \text{(i)} & M(x) \geq M(y) \text{ oder} \\ \text{(ii)} & M(x) \leq M(y) \end{array}$$

Ist dies erfüllt, gilt (i) wenn $x_0 \geq y_0$ und (ii) wenn $x_0 \leq y_0$.

Beweis: $M(x) \geq M(y) \iff 0 \leq M(x) - M(y) = M(\underbrace{x-y}_{=:z}) = M(z)$.

$$\det(M(z)) = z_0^2 - z_3^2 - (z_1 + i z_2)(z_1 - i z_2) = (x_0 - y_0)^2 - \sum_{i=1}^{3} (x_i - y_i)^2$$

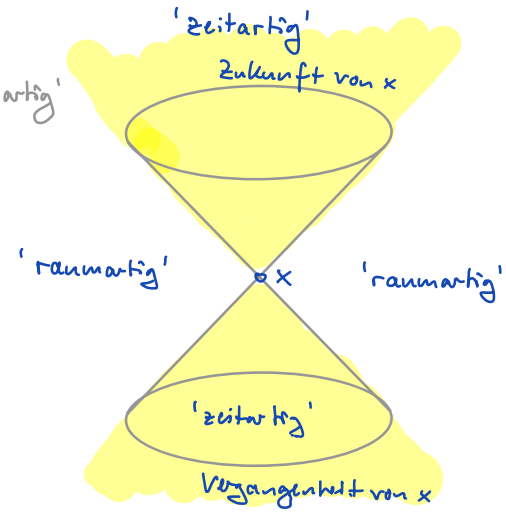
$\mathrm{tr}[M(z)] = 2(x_0 - y_0)$. Die Aussage folgt dann aus dem Lemma. $\square$

Exkursion ...

Interpretation: $\mathbb{R}^4 = \underset{\text{Zeit}}{\mathbb{R}} \times \underset{\text{Raum}}{\mathbb{R}^3}$ betrachtet als 'Minkowskiraum' der spez. Rel.theorie besitzt eine 'partielle Ordnung' für die

$x \geq y \Leftrightarrow M(x) \geq M(y)$

$\det(M(x-y)) \geq 0 \Leftrightarrow$ 'zeitartig'



Lorentz-trafos sind lineare Abb.en $\mathbb{R}^4 \to \mathbb{R}^4$, die die Minkowski-'Metrik' $z_0^2 - z_1^2 - z_2^2 - z_3^2$ invariant lassen. Auf V sind diese einfach $M(z) \mapsto S M(z) S^*$ mit $S \in SL(2, \mathbb{C})$.

Was wenn eine Matrix nicht normal, oder gar nicht quadratisch ist?

Satz: (Singulärwertzerlegung) engl. 'sing. value decomp.' (SVD)

Sei $A \in \mathbb{K}^{n \times m}$ mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $r := \operatorname{rg}(A)$.

Dann existieren $s_1, \dots, s_r > 0$ (Singulärwerte) und orthogonale bzw. unitäre Matrizen $U \in \mathbb{K}^{n \times n}$, $V \in \mathbb{K}^{m \times m}$ s.d.

$$\boxed{A = U D V} \quad \text{mit} \quad D \in \mathbb{K}^{n \times m},\ D_{ij} = \begin{cases} s_i & \text{für } i = j \leq r \\ 0 & \text{sonst} \end{cases}$$

Beweis:

$\mathbb{K}^{m\times m} \ni A^*A \geq 0 \Rightarrow \exists$ ONB $v_1,\dots,v_m \in \mathbb{K}^m$: $A^*A = \sum_{i=1}^{r} \lambda_i v_i v_i^*$ mit $\lambda_i \geq 0$.

Setze $s_i := \sqrt{\lambda_i}$ und $V := \sum_{i=1}^{m} e_i v_i^*$ d.h. $\overline{v_i}$ sind Zeilenvektoren.

Die Vektoren $w_i := \frac{1}{s_i} A v_i$ für $i \in \{1,\dots,r\}$ sind orthonormal,

da $\langle w_i, w_j \rangle = \frac{1}{s_i s_j} \langle v_i, \underbrace{A^*A v_j}_{= \lambda_j v_j} \rangle = \frac{\lambda_j}{s_i s_j} \delta_{ij}$.

Dies lassen sich zu einer ONB $w_1,\dots,w_n$ von $\mathbb{K}^n$ vervollständigen,

s.d. $U := \sum_{l=1}^{n} w_l e_l^* \in \mathbb{K}^{n\times n}$ orthogonal bzw. unitär ist.

Dann ist einerseits $\langle w_i, A v_j \rangle = \frac{1}{s_i} \langle A v_i, A v_j \rangle = \frac{1}{s_i} \langle v_i, A^*A v_j \rangle = s_i \delta_{ij}$

und andererseits $\langle w_i, UDV v_j \rangle \underset{w_i^* U = e_i^*}{=} \langle e_i, DV v_j \rangle \underset{V v_j = e_j}{=} \langle e_i, D e_j \rangle = s_i \delta_{ij}$ $\square$

Bem.:

- Es gilt $A^*A = V^* D^* D V$ und $AA^* = U D D^* U^*$.
  Die Singulärwerte sind die Wurzeln der Eigenwerte von A^*A (& AA^*) und bis auf Umordnung eindeutig.
  V wird aus den Eigenvektoren v_j von A^*A gebildet und U aus den -"- w_j von AA^*.
- Alternative die SVD zu schreiben:

$$\boxed{A = \sum_{i=1}^{r} s_i w_i v_i^*}$$

Ordnet man die Singulärwerte $s_1 \geq s_2 \geq \dots \geq s_r$ erhält man möglich 'Approximationen' $A_k := \sum_{i=1}^{k} s_i w_i v_i^*$ die für $k \ll n$ deutlich weniger Speicherbedarf besitzen

(Stichworte: 'principal component analysis', 'dimensionality reduction')

- Ist $A \geq 0$, dann ist Singulärwertzerlegung = Spektralzerl.
- Ist $A = A^*$ mit $\operatorname{spec}(A) = \{\lambda_1, \dots, \lambda_r\}$, dann gilt $s_i = |\lambda_i|$ ($\to$ Übung)

<u>Korollar:</u> Hat $A \in \mathbb{C}^{n \times n}$ Singulärwerte $s_1, \dots, s_n$, dann ist

$$\boxed{|\det(A)| = \prod_{i=1}^{n} s_i}$$

<u>Beweis:</u>

$$|\det(A)| = |\det(UDV)| = \underbrace{|\det(U)|}_{=1} \cdot \underbrace{\det(D)}_{=\prod_{i=1}^{n} s_i} \cdot \underbrace{|\det(V)|}_{=1}. \qquad \square$$

<u>Konsequenz:</u> Jede lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^n$ ist eine Hintereinanderausführung Drehung - Streckung - Drehung. Volumina werden dabei um einen Faktor $|\det(A)|$ geändert.

<u>Korollar:</u> (Polarzerlegung)

Sei $A \in \mathbb{K}^{n \times n}$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Dann existiert eine positiv semidefinite Matrix $P \in \mathbb{K}^{n \times n}$ und eine unitäre (bzw. orthogonale) Matrix $X \in \mathbb{K}^{n \times n}$, so dass $\boxed{A = XP}$.

<u>Beweis:</u> Die Singulärwertzerlegung liefert $A = UDV = \underbrace{(UV)}_{=:X}\underbrace{(V^*DV)}_{=:P}$.

$\square$

X. Matrix - Funktionen

Def.: Ist $f: \mathbb{C} \to \mathbb{C}$ durch eine auf ganz $\mathbb{C}$ absolut konvergierende Potenzreihe $f(x) = \sum_{k=0}^{\infty} c_k x^k$ gegeben, und $A \in \mathbb{C}^{n \times n}$, dann lässt sich $f(A)$ durch

$$\boxed{f(A) := \sum_{k=0}^{\infty} c_k A^k}$$

definieren, wobei $A^k = \underbrace{A \cdot A \cdot \ldots \cdot A}_{k \text{ mal}}$ und $A^0 = I_n$ ist.

Bem.:

- z.B. kann f Polynom, Exponentialfunktion oder Trigonometrische Fkt. sein.
- Mit den Methoden der Analysis kann man zeigen, dass die Reihe stets konvergiert. D.h. $f(A) \in \mathbb{C}^{n \times n}$ ist wohldefiniert.
- Achtung $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq \begin{pmatrix} f(a) & f(b) \\ f(c) & f(d) \end{pmatrix}$!

  z.B. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ aber $\begin{pmatrix} 0^2 & 1^2 \\ 0^2 & 0^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.
- Im Folgenden schreiben wir $\bigoplus_k B_k$ für eine Blockdiagonale Matrix $\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & B_3 & \\ & & & \ddots \end{pmatrix}$.

Satz: Ist $f: \mathbb{C} \to \mathbb{C}$ durch eine auf ganz $\mathbb{C}$ absolut konvergierende Potenzreihe $f(x) = \sum_{k=0}^{\infty} c_k x^k$ gegeben, und $A \in \mathbb{C}^{n \times n}$ mit Jordan-Normalform $A = X\left(\bigoplus_l J_l(\lambda_l)\right)X^{-1}$, dann ist

$$\boxed{f(A) = X\left[\bigoplus_l f(J_l(\lambda_l))\right]X^{-1}}$$

wobei

$$f: \mathbb{C}^{k \times k} \ni \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \mapsto \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \cdots & \frac{1}{(k-1)!}f^{(k-1)}(\lambda) \\ & f(\lambda) & f'(\lambda) & \cdots & \vdots \\ & & f(\lambda) & & \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}$$

Bem.: D.h. insbesondere wenn A diagonalisierbar ist:

$$\boxed{A = X\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}X^{-1} \Rightarrow f(A) = X\begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix}X^{-1}}$$

Beweisidee: Ist $A = XDX^{-1}$, dann gilt wegen $A^k = XD\underbrace{X^{-1}X}_{I_n}DX^{-1}\cdots$, dass $f(A) = X\left[\sum_{k=0}^{\infty} c_k D^k\right]X^{-1}$.

Zudem ist $D^k = \left(\bigoplus_l J_l\right)^k = \bigoplus_l J_l^k$.

Ist $J = \lambda I_m + N$ ein Jordanblock der Größe m, dann gilt

$$J^k = (\lambda I_m + N)^k = \sum_{i=0}^{k} \binom{k}{i} \lambda^{k-i} N^i$$

wobei alle Terme mit $i \geq m$ wegen $N^m = 0$ Null sind.

Einsetzen liefert nach etwas Umformen (zumindest formal) den gewünschten Ausdruck... $\square$

Bem.: Ein $N \in \mathbb{C}^{n \times n}$ für das ein $k \in \mathbb{N}$ existiert, s.d. $N^k = 0$ ist, heißt *nilpotent*.

Bem.: ∘ Die Form im Satz lässt sich als verallgemeinerte Definition von $f(A)$ verwenden:
Ist $m \in \mathbb{N}$ die Größe des größten Jordanblocks von A, dann genügt es, dass f in einer offenen Umgebung von $\operatorname{spec}(A)$ $m-1$ mal diff.bar ist (bzw. stetig, wenn $m=1$).

Satz: (Cayley-Hamilton)
Für $A \in \mathbb{C}^{n\times n}$ mit charakteristischem Polynom $\chi_A \in \mathbb{C}[x]$ gilt $\boxed{\chi_A(A) = 0}$.

Beweis: Zunächst für diagonalisierbares A:
Für $A = X \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} X^{-1}$ gilt $\chi_A(A) = X \begin{pmatrix} \chi_A(\lambda_1) & & \\ & \ddots & \\ & & \chi_A(\lambda_n) \end{pmatrix} X^{-1} = 0$
da $\chi_A(\lambda_i) = 0 \quad \forall i$. $\square$

Für bel. A können wir eine gegen A konvergierende Folge $A_n \in \mathbb{C}^{n\times n}$ von diagonalisierbaren Matrizen konstruieren (z.B. indem wir die Entartung der Eigenwerte aufheben). Dann gilt für

$$f(A) := \chi_A(A) = f\left(\lim_{n\to\infty} A_n\right) \underset{\substack{\uparrow \\ \text{stetig da Polynom}}}{=} \lim_{n\to\infty} f(A_n) \underset{\substack{\uparrow \\ A_n \text{ diagonalisierbar.}}}{=} 0 .$$

$\square$

Bsp.: Für $A \in \mathbb{C}^{2\times 2}$ ist $\chi_A(x) = x^2 - \operatorname{tr}[A]\, x + \det(A)$.
Nach dem Satz von Cayley-Hamilton gilt demnach:

$$A^2 = \operatorname{tr}[A]\, A - \det(A)\, I_2$$