Notizen zu Linearer Algebra

Constantin Kogler
Herbstsemester 2018

Inhaltsverzeichnis

Vorwort

Dieses Dokument umfasst Notizen zu meiner Übungsstunde in Linearer Algebra, welche begleitend zu der Vorlesung von Prof. Richard Pink im Herbstsemester 2018 and der ETH Zürich erstellt wurden. Die dreizehn Kapitel entsprechen den dreizehn Übungsstunden. Es wird auf Definitionen, Beispiele und praktische Anwendungen Wert gelegt anstatt die Theorie vollständig und rigoros darzustellen. Jedoch sind auch manche mathematische Sätze, Propositionen und Lemmatas aufgeführt, welche ich persönlich für die Entwicklung der Theorie als besonders grundlegend erachte. Zu diesen Aussagen gibt es meist knappe Beweise, welche ich aus Gründen der Vollständigkeit aufgeschrieben habe.

Die Schnittmenge zwischen diesen Notizen und dem Material, welches ich in der Übungsstunde vorgetragen habe, variierte von Woche zu Woche, wobei insbesondere kaum Beweise in der Übungsstunde behandelt wurden. Hingegen habe ich in der Übungsstunde auch einige Bilder gemalt und Rechenalgorithmen erklärt, welche in diesen Notizen nicht aufgeführt sind.

Bedanken möchte ich mich schliesslich bei all den Teilnehmern meiner Übungsstunde für die aktive Partizipation und die zahlreichen guten Fragen. Auch Nicholas Dykeman gilt es zu danken für das Korrekturlesen dieser Notizen.

Zürich, 2018
Constantin Kogler

1 Mengen, Abbildungen, Relationen und Quantoren

1.1 Mengen

Die Mengenlehre ist an und für sich ein eigenes Gebiet der Mathematik oder Logik, dessen Behandlung nicht das Ziel der Vorlesung Lineare Algebra ist. Wir wählen hier einen naiven Zugang zu Mengen und ignorieren einfach die logischen Schwierigkeiten im Umgang mit Mengen. Für uns ist eine Menge einfach eine Ansammlung an Elementen. Zum Beispiel

$$\{\text{Mathematik}, \text{Faserland}, \text{Donald Trump}\}.$$

Wir werden aber viel eher mathematische Mengen betrachten wie

$$\{3, 1, 4, 1, 5, 9, 2\}.$$

Mengen können auch andere Mengen enthalten, also ist auch folgendes Objekt eine Menge

$$\{\{\}, \{\text{Penne}, \text{Tomaten}, \text{Parmesan}\}\}.$$

Wir geben nun ein paar Regeln oder Konventionen zum Umgang mit Mengen an:

1. Jedes Element einer Menge kommt nur einmal vor. Somit ist

   $$\{3, 1, 4, 1, 5, 9, 2\} = \{3, 1, 4, 5, 9, 2\}.$$

2. Die Reihenfolge der Elemente ist irrelevant. Mengen haben keine intrinsische Ordnung. Also ist

   $$\{3, 1, 4, 5, 9, 2\} = \{1, 2, 3, 4, 5, 9\}.$$

   Mengen, welche aus einzelnen Zahlen bestehen, schreibt man jedoch üblicherweise in aufsteigender Reihenfolge.

3. Mengen dürfen sich nicht selber als Element enthalten. Dies führt nämlich schnell zu Problemen, wie folgende Überlegung zeigt. Nehmen wir für den Moment an, dass Mengen sich selber enthalten können. Betrachten wir dann folgende Menge

   $$\mathcal{F} = \{\text{Mengen die sich nicht selbst enthalten}\}.$$

   Dies ist eine Menge, da es einfach eine Ansammlung von Elementen ist. Wir führen dies nun zu einem Widerspruch. Dazu betrachten wir zwei Fälle und zeigen, dass beide nicht sein können. Nehmen wir zuerst an, dass $\mathcal{F}$ sich nicht selber enthält. Dann ist aber $\mathcal{F}$ nach Definition in $\mathcal{F}$ enthalten. Widerspruch. Nehmen wir nun an, dass $\mathcal{F}$ sich selbst enthält. Also ist $\mathcal{F}$ eine Menge, die sich nicht selbst enthält nach der Definition der Menge $\mathcal{F}$. Widerspruch.

   Als Lösung des Widerspruches folgern wir, dass $\mathcal{F}$ keine Menge ist. Eine natürliche Frage, die sich nun stellt, ist ein widerspruchfreies System von Axiomen anzugeben, welche Mengen axiomatisch charakterisieren. Dies ist jedoch eine schwierige Frage und nicht Ziel der Vorlesung lineare Algebra.

4. Die leere Menge ist die Menge, welche keine Elemente hat. Sie wird üblicherweise bezeichnet mit $\{\}$ oder $\emptyset$. Eine wichtige Eigenschaft der leeren Menge ist, dass jede Aussage über ihre Elemente wahr ist, da sie gar keine Elemente hat über welche eine Aussage getroffen werden kann.

5. Ist x ein Element einer Menge A, so schreiben wir $x \in A$. Ist x kein Element von A, so schreiben wir $x \notin A$.

Als nächstes diskutieren wir ein paar Operationen, welche wir mit Mengen machen können. Im Folgenden bezeichnen wir stets mit A und B zwei Mengen.

1. Vereinigung: Die Vereinigung zweier Mengen ist definiert als

$$A \cup B := \{x \,|\, x \in A \text{ oder } x \in B\}.$$

2. Durchschnitt: Wir definieren

$$A \cap B := \{x \,|\, x \in A \text{ und } x \in B\}.$$

3. Differenz:

$$A \backslash B := \{x \,|\, x \in A \text{ und } x \notin B\}.$$

4. Kartesisches Produkt:

$$A \times B := \{(a, b) \,|\, x \in A \text{ und } x \in B\}.$$

Wir diskutieren nun Abbildungen zwischen Mengen.

Definition. Seien X und Y Mengen. Eine Abbildung von X nach Y, geschrieben $f : X \to Y$, ordnet jedem Element von X ein eindeutiges Element von Y zu.

Ich möchte nun auch noch ein paar Anmerkungen zu der Definition von Abbildungen machen:

1. Vom formalen Standpunkt sind Abbilgungen nur dann gleich, wenn auch die Definitionsmenge X und die Bildmenge Y gleich sind.

2. Betrachten wir

$$f : \mathbb{N} \to \{1\}, \qquad n \mapsto n.$$

Dies ist keine Abbildung, da sie nicht *wohldefiniert* ist. Das Problem ist, dass $n = f(n) \notin \{1\}$. Es ist eine wichtige Bedingung für Abbildungen $f : X \to Y$, dass tatsächlich $f(x) \in Y$ für alle $x \in X$. Falls ihr zeigen müsst, dass eine Abbildung wohldefiniert ist, müsst ihr also zeigen, dass tatsächlich $f(x) \in Y$ für alle $x \in X$.

Wir geben nun noch wichtige Begriffe zu Abbildungen an.

Definition. Seien X und Y Mengen und sei $f : X \to Y$ eine Abbildung. Wir nennen:

1. f injektiv, falls für alle $x, x' \in X$ mit $x \neq x'$ folgt $f(x) \neq f(x')$.

2. surjektiv, falls es es für alle $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$.

3. bijektiv, falls f injektiv und surjektiv ist.
4. invertierbar, falls eine Abbildung $g : Y \to X$ existiert, mit

$$g \circ f = \mathrm{id}_X \qquad \text{und} \qquad f \circ g = \mathrm{id}_Y.$$

Man bezeichnet die Inverse meist mit f^{-1}.

Eine sehr wichtige Eigenschaft für Abbildungen zwischen Mengen ist, dass eine Abbildung bijektiv ist genau dann, wenn sie invertierbar ist. Diese Aussage ist so essentiell, dass ich aus Gründen der Vollständigkeit nicht unterlassen kann einen Beweis dieser Aussage aufzuschreiben.

Satz. *Eine Abbildung zwischen zwei Mengen $f : X \to Y$ ist bijektiv genau dann, wenn sie invertierbar ist.*

Vor dem Beweis möchte ich anmerken, dass immer wenn ein Mathematiker in einer Aussage schreibt *genau dann, wenn* dann meint er, dass Aussage A Aussage B impliziert und auch, dass Aussage B Aussage A impliziert. Obiger Satz heisst also ausgeschrieben Folgendes: *Jede bijektive Abbildung ist invertierbar* und *Jede invertierbare Abbildung ist bijektiv*. Dies müssen wir also im Beweis zeigen.

Beweis. Sei $f : X \to Y$ eine bijektive Abbildung. Wir wollen zeigen, dass f invertierbar ist, also konstruieren wir eine Abbildung $g : Y \to X$ sodass

$$g \circ f = \mathrm{id}_X \qquad \text{und} \qquad f \circ g = \mathrm{id}_Y.$$

Wir definieren g wie folgt:

$$g : Y \to X, \qquad y \mapsto x \text{ sodass } f(x) = y.$$

Also assoziieren wir zu jedem Element $y \in Y$ ein Element $x \in X$ mit $f(x) = y$. Diese Abbildung ist wohldefiniert, weil f surjektiv ist. Denn nach der Definition von Surjektivität existiert für jedes $y \in Y$ ein $x \in X$ mit $f(x) = y$. Dieses Element ist sogar eindeutig, weil f injektiv ist. Falls es nämlich zwei verschiedene Elemente $x, x' \in X$ gäbe mit $y = f(x) = f(x')$, so wäre dies ein Widerspruch zu der Definition von Injektivität. Somit gilt für alle $x \in X$

$$g(f(x)) = x,$$

da x das eindeutige Element x' ist mit $f(x') = f(x)$. Auch gilt für alle $y \in Y$

$$f(g(y)) = y,$$

da nach Definition $g(y)$ ein Element von X ist, sodass f auf dieses Element angewendet gleich y ist. Somit haben wir gezeigt, dass g tatsächlich eine Inverse von f ist.

Nun zeigen wir die andere Richtung. Dazu nehmen wir an, dass f invertierbar ist. Also gibt es eine Abbildung $g : Y \to X$ mit

$$g \circ f = \mathrm{id}_X \qquad \text{und} \qquad f \circ g = \mathrm{id}_Y.$$

Wir wollen zeigen f ist bijektiv, also zeigen wir, dass f injetktiv und surjektiv ist. Wir zeigen zuerst, dass f injektiv ist. Dazu führen wir einen Widerspruchsbeweis durch. Wir nehmen also an, f ist nicht injektiv. Also gibt es zwei verschiedene Elemente $x, y \in X$ mit $f(x) = f(y)$. Damit folgt aber weil $g \circ f = \mathrm{id}_X$, dass

$$x = g(f(x)) = g(f(y)) = y.$$

Wir haben aber angenommen, dass x und y verschieden sind. Das ist somit ein Widerspruch.

Es bleibt zu zeigen, dass f surjektiv ist. Dazu verwenden wir die zweite Eigenschaft: $f \circ g = \mathrm{id}_Y$. Also gilt für alle $y \in Y$, dass

$$f(g(y)) = y$$

und weil $g(y) \in X$ haben wir für jedes Element von Y ein Element von X gefunden, welches auf y abgebildet wird. Somit ist f surjektiv. □

1.2 Quantoren und Aussagenlogik

Quantoren helfen bei der Formalisierung mathematischer Aussagen. Die wichtigsten Quantoren sind die Folgenden:

1. $\wedge$ = und
2. $\vee$ = oder
3. $\neg$ = nicht oder Negation
4. $\exists$ = es existierert
5. $\exists!$ = es existierert genau ein
6. $\forall$ = für alle

Die ersten drei Quantoren beziehen sich auf Aussagen. Nach Definition sind Aussagen stets als wahr oder falsch klassifiziert. Abstrakter können wir jeder Aussage die Zahl 0 oder 1 zuordnen, je nachdem ob sie wahr oder falsch ist. Wir verwenden hierbei die folgende Konvention:

$$0 := \text{falsch}, \qquad 1 := \text{wahr}.$$

Somit können wir nun sogenannte Wahrheitstafeln aufstellen. Als Beispiel stellen wir eine Wahrheitstafel für $A \vee B$ auf, wobei A und B Aussagen sind. Dies sieht dann so aus.

A	B	$A \vee B$
1	1	1
1	0	1
0	1	1
0	0	0

Nun können wir auch definieren, was $A \to B$ bedeutet. Dabei wir $A \to B$ als A *impliziert* B ausgesprochen.

Definition. Die Aussage $A \to B$ ist definiert als $(\neg A) \wedge B$ und wahr genau dann, wenn A falsch oder B wahr ist.

Wir können $A \to B$ auch mit einer Wahrheitstafel ausdrücken.

A	$\neg A$	B	$A \to B = (\neg A) \vee B$
1	0	1	1
1	0	0	0
0	1	1	1
0	1	0	1

Nun können wir auch definieren, was es formal heisst, dass zwei Aussagen äquivalent sind.

Definition. Die Aussage $A \leftrightarrow B$ ist definiert als $(A \to B) \wedge (B \to A)$.

Wir zeigen nun mit Hilfe einer Wahrheitstafel, dass $A \leftrightarrow B$ genau dann wahr ist, wenn A und B den gleichen Wahrheitswert haben.

A	B	$A \leftrightarrow B$
1	1	1
1	0	0
0	1	0
0	0	1

Als weiteres Beispiel behaupten wir nun folgende Aussage.

Lemma. *Die Aussage $A \leftrightarrow (A \wedge \text{wahr})$ ist wahr.*

Beweis. Wir führen den Beweis mit einer Wahrheitstafel.

A	$(A \wedge \text{wahr})$
1	1
0	0

Somit haben die Aussagen A und $(A \wedge \text{wahr})$ stets den gleichen Wahrheitswert. Also gilt $A \leftrightarrow (A \wedge \text{wahr})$. □

1.3 Relationen

Definition. Eine *Relation* $\sim$ auf einer Menge X ist eine Teilmenge des kartesischen Produktes $X \times X$. Ist (x, y) ein Element dieser Teilmenge, so schreiben wir $x \sim y$.

Definition. Eine Relation $\sim$ auf einer Menge X heisst *Äquivalenzrelation*, falls folgende drei Eigenschaften erfüllt sind:

1. Reflexivität: Für alle $x \in X$ gilt $x \sim x$.
2. Symmetrie: Falls $x \sim y$ für $x, y \in X$, so gilt $y \sim x$.
3. Transitivität: Falls $x \sim y$ und $y \sim z$ für $x, y, z \in X$, so folgt $x \sim z$.

Definition. Sei $\sim$ eine Äquivalenzrelation auf einer Menge X. Die Äquivalenzklasse eines Elements $x \in X$ ist die Teilmenge

$$[x] := \{y \in X \,:\, y \sim x\}.$$

Die Menge

$$X/\sim := \{[x] \,:\, x \in X\}$$

bezeichnet die Menge der Äquivalenzklassen.

Wir geben nun ein paar Beispiele an.

Beispiel. Triviale Relation: Auf jeder Menge X können wir die triviale Relation definieren, unter welcher alle Elemente assoziiert sind. In diesem Fall ist $[x] = X$ für alle $x \in X$ und somit besteht $X/\sim$ aus einem Element.

Beispiel. Kreis: Betrachten wir $\mathbb{R}$ und die Äquivalenzrelation

$$\begin{aligned} x \sim y \qquad &\Leftrightarrow \qquad \text{Es gibt ein } n \in \mathbb{Z}, \text{ sodass } x = y + n, \\ &\Leftrightarrow \qquad x \equiv y \mod 1. \end{aligned}$$

Die ist eine Äquivalenzrelation wie man leicht überprüft. Es gilt des Weiteren für $x \in \mathbb{R}$

$$\begin{aligned} [x] &= x + \mathbb{Z} \\ &= \{x + n : n \in \mathbb{Z}\} \\ &= \{\ldots, x-2, x-1, x, x+1, x+2, \ldots\}. \end{aligned}$$

Die Menge der Äquivalenzklassen können wir als Kreis betrachten.

Beispiel. Konstruktion der rationalen Zahlen aus den ganzen Zahlen: Betrachten wir die ganzen Zahlen

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

sowie die positiven ganzen Zahlen

$$\mathbb{Z}_{>0} = \{1, 2, 3, 4, 5, \ldots\}$$

Wir betrachten nun folgende Äquivalenzrelation auf $\mathbb{Z} \times \mathbb{Z}_{>0}$:

$$(a, b) \sim (c, d) \qquad \Leftrightarrow \qquad ad = bc.$$

Überprüfen wir, dass dies eine Äquivalenzrelation ist.

1. Reflexiv: $(a, b) \sim (a, b)$ da $ab = ab$.
2. Symmetrie: Falls $(a, b) \sim (c, d)$ dann ist $ad = bc$ also auch $bc = ad$ und somit $(c, d) = (a, b)$.
3. Transitivität: Falls $(a, b) \sim (c, d)$ also $ad = bc$ und $(c, d) \sim (e, f)$ also $cf = de$. Also folgt $ade = afc = ebc$ und somit $af = eb$, also $(a, b) \sim (e, f)$.

Wir können nun die rationalen Zahlen als die Menge der Äquivalenzklassen dieser Relation betrachten.

2 Gruppen

Betrachten wir zunächst die rationalen Zahlen

$$\mathbb{Q} := \left\{ \frac{a}{b} \,:\, a \in \mathbb{Z}, b \in \mathbb{N}_{\neq 0} \right\}.$$

Erinnern wir uns, aus dem bekannten Umgang mit rationalen Zahlen, dass folgende Eigenschaften von den rationalen Zahlen und der Addition erfüllt werden:

1. Assoziativität: $x + (y + z) = (x + y) + z$ für alle $x, y, z \in \mathbb{Q}$.
2. Einheitselement: $x + 0 = 0 + x = x$ für alle $x \in \mathbb{Q}$.
3. Additive Inverse: $x + (-x) = (-x) + x = 0$ für alle $x \in \mathbb{Q}$.
4. Kommutativität: $x + y = y + x$ für alle $x, y \in \mathbb{Q}$.

Der Begriff einer Gruppe verallgemeinert diese Eigenschaften. Geben wir nun eine genaue Definition an.

Definition. Eine *Gruppe* ist ein Tripel $(G, \circ, e)$ bestehend aus einer Menge G, einer Abbildung

$$\circ : G \times G \to G$$

und einem Element $e \in G$, genannt Einheitselement, sodass folgende Eigenschaften erfüllt sind:

1. Assoziativität: $x \circ (y \circ z) = (x \circ y) \circ z$ für alle $x, y, z \in G$.
2. Einheitselement: Es gibt ein $e \in G$, sodass $x \circ e = e \circ x = x$ für alle $x \in G$.
3. Inverse: Für jedes $x \in G$ existiert ein Element $x' \in G$ sodass $x \circ x' = x' \circ x = e$. Man bezeichnet mit x^{-1} meist das Inverse.

Des Weiteren nennt man eine Gruppe *kommutativ*, falls für alle $x, y \in G$ gilt, dass

$$x \circ y = y \circ x.$$

Ich mache nun ein paar Anmerkungen zu der Definition von Gruppen.

1. Wir schreiben die Abbildung $\circ$ ein bisschen eigenartig. Lasst euch nicht davon verwirren. Die Alternative wäre, dass wir $\circ(x, y)$ schreiben würden. Dann würde aber beispielsweise das Assoziativgesetz eigenartig aussehen.
2. Mann kann Gruppen auch *multiplikativ* schreiben: Dann schreibt man anstatt $\circ$ ein $\cdot$ und man bezeichnet mit 1 das Einheitselement und mit a^{-1} das Inverse.
3. Kommutative Gruppen werden manchmal *additiv* geschreiben: Dann schreibt man $+$ anstatt $\circ$ und man bezeichnet mit 0 das Einheitselement und mit $-a$ das Inverse.

Wir beweisen nun zwei kleine Lemmata.

Lemma. *Falls für $x, y, z \in G$ gilt, dass*

$$x \circ z = y \circ z,$$

so folgt $x = y$.

Beweis. Falls $x \circ z = y \circ z$ so können wir mit z^{-1} multiplizieren und erhalten dann

$$x = x \circ e = x \cdot (z \circ z^{-1}) = y \circ (z \circ z^{-1}) = y \circ e = y.$$

□

Lemma. *Das Einheitselement in einer Gruppe ist eindeutig.*

Beweis. Falls wir zwei Einheitselemente $e, e' \in G$ haben, so gilt für alle $a \in G$

$$e \circ a = a = e' \circ a.$$

Mit der Kürzungsregel gilt $e = e'$. □

Betrachten wir nun ein paar Beispiele.

Beispiel. $(\mathbb{Q}, +, 0)$ bildet eine kommutative Gruppe. Ebenso $(\mathbb{Z}, +, 0)$ und $(\mathbb{R}, +, 0)$

Beispiel. $(\mathbb{Q}\backslash\{0\}, \cdot, 1)$ bildet eine kommutative Gruppe. Ebenso $(\mathbb{R}\backslash\{0\}, \cdot, 1)$. Hier ist es wichtig anzumerken, dass wir das Nullelement weglassen müssen. Das Problem mit dem Nullelement ist nämlich, dass

$$0 \cdot x = x \cdot 0 = 0$$

für alle $x \in \mathbb{Q}$. Somit ist stets $0 \cdot x \neq 1$. Also kann 0 kein multiplikatives Inveres haben.

Beispiel. Die natürlichen Zahlen bilden mit der Addition **keine** Gruppe, da es kein additives Inverses zu einer positiven Zahl gibt.

Auch bildet $(\mathbb{Z}\backslash\{0\}, \cdot, 1)$ keine Gruppe, da beispielswiese 2 kein Inverses bezüglich der Multiplikation hat.

Beispiel. Betrachten wir nun ein abstrakteres Beispiel. Dazu betrachten wir die Menge $G = \{\$, £\}$ und definieren die Operation $\circ : G \times G \to G$ wie folgt:

$$\$ \circ \$ = \$, \qquad \$ \circ £ = £, \qquad £ \circ \$ = £, \qquad £ \circ £ = \$.$$

Wir können dies auch eleganter in einer sogennanten *Gruppentafel* darstellen:

$\circ$	$\$$	$£$
$\$$	$\$$	$£$
$£$	$£$	$\$$

Man kann nun kombinatorisch überprüfen, dass es sich hierbei um eine kommutative Gruppe handelt.

Wir können diese Gruppe auch einfacher betrachten, sodass sie sich auf mehr Elemente erweitern lässt. Schreiben wir nämlich $\$ = 0$ und $£ = 1$ und betrachten die Abbildung $\circ$ also Addition modulo 2 so erhalten wir die gleiche Gruppe. Diese Gruppe wird auch als $\mathbb{Z}/2\mathbb{Z}$ geschrieben. Wir verallgemeinern dies im nächsten Beispiel.

Beispiel. Fixieren wir eine natürliche Zahl $n \geq 2$. Wir betrachten die Menge

$$\mathbb{Z}/n\mathbb{Z} := \{\overline{0}, \overline{1}, \ldots, \overline{n-2}, \overline{n-1}\}.$$

Wir können auf $\mathbb{Z}/n\mathbb{Z}$ eine Gruppenstruktur definieren, welche die additive Struktur von $\mathbb{Z}$ reflektiert. Nämlich definieren wir für $\overline{x}, \overline{y} \in \mathbb{Z}/n\mathbb{Z}$

$$\overline{x} + \overline{y} = x + y \mod n = \begin{cases} x + y & \text{falls } x + y < n, \\ x + y - n & \text{falls } x + y \geq n. \end{cases}$$

So ist beispielsweise in $\mathbb{Z}/5\mathbb{Z}$

$$\overline{2} + \overline{3} = \overline{0} \qquad \text{und} \qquad \overline{4} + \overline{4} = \overline{3}.$$

In dieser Gruppe schreiben wir $-x$ für das Inverse bezüglicher dieser Gruppenoperation. Also haben wir in $\mathbb{Z}/5\mathbb{Z}$: $(-1) = 4$.

Wir können auch eine Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$ definieren. Dies geschieht analog zu der Addition. Wir definieren also für $\overline{x}, \overline{y} \in \mathbb{Z}$

$$\overline{x} \cdot \overline{y} = x \cdot y \mod n.$$

Beispiel. Für jede Menge X ist die Menge der bijektiven Selbstabbildungen

$$\begin{aligned} \mathrm{Bij}(X) &:= \{f : X \to X \,:\, f \text{ ist bijektiv}\} \\ &= \{f : X \to X \,:\, f \text{ ist invertierbar}\} \\ &= \{f : X \to X \,:\, \exists g : X \to X \text{ mit } f \circ g = g \circ f = \mathrm{id}_X\}. \end{aligned}$$

Die Menge der bijektiven Abbildungen bildet zusammen mit der Komposition von Funktionen eine Gruppe. Überprüfen wir dazu die Eigenschaften. Betrachten wir zunächst die Abbildung

$$\circ : \mathrm{Bij}(X) \times \mathrm{Bij}(X) \to \mathrm{Bij}(X), \qquad (f, g) \mapsto f \circ g.$$

Diese Abbildung ist wohldefiniert, da die Komposition bijektiver Abbildungen bijektiv ist. Die Gruppeneigenschaften folgen:

1. Assoziativität: Es gilt ganz allgemein, dass die Komposition von Funktionen assoziativ ist.

2. Einheitselement: Für jede Menge ist die Identitätsabbildung id_X bijektiv. Auch gilt
$$f \circ \mathrm{id}_X = \mathrm{id}_X \circ f = f.$$
Somit definiert id_X ein Einheitselement.

3. Inverse: Nach Definition gibt es für jede Abbildung $f \in \mathrm{Bij}(X)$ eine Abbildung $g \in \mathrm{Bij}(X)$ mit
$$f \circ g = g \circ f = \mathrm{id}_X.$$
Somit haben wir zu jedem Element aus $\mathrm{Bij}(X)$ ein inverses Element gefunden.

Betrachten wir nun den Spezialfall $X = \{1, 2, 3\}$. Wir behaupten, dass $\mathrm{Bij}(X) = \mathrm{Bij}(\{1, 2, 3\})$ **keine** kommutative Gruppe ist. Um dies zu sehen betrachten wir die beiden Funktionen $f_{213}, f_{132} \in \mathrm{Bij}(\{1, 2, 3\})$, welche wir definieren durch

$$f_{213}(1) = 2, \qquad f_{213}(2) = 1, \qquad f_{213}(3) = 3,$$

und

$$f_{132}(1) = 1, \qquad f_{132}(2) = 3, \qquad f_{132}(3) = 2.$$

Dann haben wir

$$f_{213} \circ f_{132}(1) = 2 \neq 3 = f_{132} \circ f_{213}(1)$$

und somit ist $f_{213} \circ f_{132} \neq f_{132} \circ f_{213}$. Also ist $\mathrm{Bij}(\{1, 2, 3\})$ nicht abelsch. Man kann dies verallgemeinern um zu sehen: Falls die Menge X mehr als zwei Elemente hat, dann ist $\mathrm{Bij}(X)$ nicht kommutativ.

3 Körper, Induktion

3.1 Körper

Wir haben gesehen, dass die rationalen Zahlen zwei Gruppenstrukturen enthalten, nämlich die Addition und die Multiplikation. Des Weiteren gilt

$$1 \neq 0.$$

Und es gilt die Distributivität: Für alle $x, y, z \in \mathbb{Q}$

$$x \cdot (y + z) = x \cdot y + y \cdot z.$$

Ein Objekt, welches all diese Eigenschaften erfüllt, nennen wir einen Körper. Fassen wir dies in der folgenden Definition zusammen.

Definition. Ein Tupel $(K, +, \cdot, 0, 1)$ bestehend aus einer Menge K und zwei Abbildungen

$$\begin{aligned} + &: K \times K \to K,\ (x, y) \mapsto x + y \\ \cdot &: K \times K \to K,\ (x, y) \mapsto x \cdot y \end{aligned}$$

heisst *Körper*, falls Folgendes gilt:

1. $(K, +, 0)$ ist eine kommutative Gruppe.
2. $(K \backslash \{0\}, \cdot, 1)$ ist eine kommutative Gruppe.
3. $1 \neq 0$.
4. Distributivität: Für alle $x, y, z \in K$ gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Folgendes möchte ich anmerken.

1. Eine kommutative Gruppe hat vier Axiome. Sollte man nun alle Axiome eines Körper aufschreiben kommt man auf $2 \cdot 4 + 2 = 10$. Um sich die Körperaxiome zu merken, muss man nur die Gruppenaxiome wissen und dann noch die beiden Axiome hinzufügen, dass $1 \neq 0$ ist und dass Distributivität gilt.
2. In einem Körper schreiben wir $-x$ für das additive Inverse und x^{-1} für das multiplikative Inverse.
3. Teilen heisst Multiplikation mit dem multiplikativen Inversen. Dies bedeutet, dass Brüche in Körper wie folgt definiert sind:

$$\frac{a}{b} := a \cdot b^{-1}.$$

Beispiel. $\mathbb{Q}$ und $\mathbb{R}$ sind Körper.

Beispiel. $\mathbb{Z}/3\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}\}$ ist ein Körper. Wir überprüfen dies nun. Es folgt wie oben, dass $(\mathbb{Z}/3\mathbb{Z}, 0, +)$ eine Gruppe ist. Wir betrachten nun die Multiplikation. Die Assoziativität folgt direkt, ebenso ist klar, dass $\overline{1}$ ein Einheitselement ist. Wir haben auch $\overline{1} \neq \overline{0}$. Die Distributivität folgt aus der Distributivität der Addition und der Multiplikation in $\mathbb{Z}$. Es bleibt also zu zeigen, dass jedes Element ausser $\overline{0}$ ein multiplikatives Inverses hat. Dies ist klar für $\overline{0}$ und für $\overline{1}$. Für $\overline{2}$ bemerken wir, dass $\overline{2} \cdot \overline{2} = \overline{1}$. Also ist $\overline{2}$ sein eigenes multiplikativ Inverses.

Nun zeigen wir noch, wie Brüche in $\mathbb{Z}/3\mathbb{Z}$ zu verstehen sind. Es gilt nämlich:

$$\frac{\overline{1}}{\overline{2}} = \overline{1} \cdot (\overline{2})^{-1} = \overline{1} \cdot \overline{2} = \overline{2}.$$

Beispiel. $\mathbb{Z}/4\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}\}$ ist **kein** Körper. Das Problem ist, dass $\overline{2}$ kein multiplikatives Inverses hat. Es gilt nämlich

$$\overline{0} \cdot \overline{2} = \overline{0}, \qquad \overline{1} \cdot \overline{2} = \overline{2}, \qquad \overline{2} \cdot \overline{2} = \overline{0}, \qquad \overline{3} \cdot \overline{2} = \overline{2}.$$

Allgemeiner kann man folgenden Satz beweisen.

Satz. *$\mathbb{Z}/n\mathbb{Z}$ is ein Körper genau dann, wenn n eine Primzahl ist.*

3.2 Eigenschaften von Körpern

Ihr könnt euch allgemein als Faustregel merken, dass ihr in Körper so rechnen könnt wie ihr von der Schule her gewöhnt seid. Genauer gesagt, kann man jede Rechenregel in Körpern beweisen. Ich werde als Beispiel folgenden Satz beweisen.

Satz. *Sei K ein Körper. Für alle $x \in K$ gilt*

$$0 \cdot x = 0.$$

Beweis. Per Definition ist 0, das neutrale Gruppenelement der additiv geschriebenen Gruppe $(K, +, 0)$. Also gilt nach den Gruppenaxiomen für alle $x \in X$, dass

$$0 + x = x.$$

Setzen wir $x = 0$, so erhalten wir

$$0 + 0 = 0.$$

Somit gilt zusammen mit Distributivität

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x.$$

Subtrahieren wir $0 \cdot x$ auf beiden Seiten, so erhalten wir

$$0 = 0 \cdot x$$

oder äquivalenterweise

$$0 \cdot x = 0.$$

□

Aus diesem Satz folgt auch, dass 0 kein multiplikatives Inverses haben kann. Also kann $(K, \cdot, 1)$ kein Gruppe sein. Dies ist auch der Grund, weshalb man nicht durch 0 teilen darf, denn teilen bedeutet ja nichts anderes als multiplizieren mit dem multiplikativen Inversen.

3.3 Induktion

Das Prinzip *Vollständige Induktion* ist eine Beweisart, mit welcher Aussagen beweisen werden, welche von allen natürlichen Zahlen abhängen. Ich möchte zunächst als Beispiel folgende Aussage beweisen.

Proposition. *Sei K ein Körper und $x \in K\backslash\{0\}$. Dann ist $x^n = 0$ für alle natürlichen Zahlen $n \geq 1$.*

Beweis. Wir beginnen mit der Induktionsverankerung. Dies ist also der Fall $n = 1$, welcher direkt aus der Defintion folgt.

Wir nehmen nun an, dass die Aussage für $i \leq n-1$ gilt. Also ist $x^{n-1} \neq 0$. Für einen Widerspruch nehmen wir nun an, dass

$$x^n = x^{n-1} \cdot x = 0.$$

Da $x \neq 0$ können wir diese Gleichung mit dem multiplikativen Inversen von x multiplizieren und erhalten somit

$$x^{n-1} = 0 \cdot x^{-1} = 0$$

nach dem obigen Satz. Aber dies widerspricht der Induktionsannahme. □

Zusammenfassend besteht ein Induktionsbeweis immer aus drei Schritten. Wir möchten eine Aussage $A(n)$ für alle n beweisen.

1. Induktionsverankerung: Die Aussage wird im Fall $n = 1$ bewiesen.
2. Induktionsannahme: Wir nehmen die Aussage für $A(n-1)$ an.
3. Induktionsschritt: Wir beweisen für $n \geq 2$, dass aus $A(n-1)$ die Aussage $A(n)$ folgt.

Es gibt noch die folgenden Varianten.

1. Vielleicht gilt die Aussage erst ab $n = n_0$. Dann wird in der Induktionsverankerung die Aussage für ein $A(n_0)$ bewiesen.
2. Du kannst in der Induktionsannahme auch annehmen, dass die Aussage für $A(1), \ldots, A(n-1)$ gilt.
3. Manche Mathematiker präferieren die Aussage $A(n)$ anzunehmen.

4 Matrizen und lineare Gleichungsysteme

Betrachten wir ein lineares Gleichungssystem über $\mathbb{R}$:

$$\begin{aligned} x + y &= 4 \\ 2x + y &= 10 \end{aligned}$$

Am einfachsten kann man solch ein Gleichungssystem lösen, indem man zum Beispiel die erste Zeile zweimal von der zweiten abzieht. Dann erhält man das Gleichungssystem mit der gleichen Lösungsmenge:

$$\begin{aligned} x + y &= 4 \\ -y &= 2 \end{aligned}$$

Also sieht man, dass $y = -2$ und da $x + y = 4$, folgt $x = 6$. Somit ist $(x, y) = (6, -2)$ die einzige Lösung dieses Gleichungssystems. Diese Lösung ist eindeutig, da jeder Schritt, den wir gemacht haben, die Lösungsmenge nicht verändert hat. Diesen Prozess nennt man das *Gausseleminationsverfahren*.

Nicht alle linearen Gleichungssysteme haben eine Lösung. Betrachten wir beispielsweise das System:

$$\begin{aligned} x + y &= 4 \\ x + y &= 10 \end{aligned}$$

Dieses System kann keine Lösung haben, weil direkt folgen würde, dass $4 = 10$. Auch wenn man die erste Zeile von der zweiten abzieht erhält man das System:

$$\begin{aligned} x + y &= 4 \\ 0 &= 6 \end{aligned}$$

Dieses System hat ebenso keine Lösung.

Auch ist die Lösung nicht immer eindeutig. Das System

$$\begin{aligned} x + y &= 4 \\ 2x + 2y &= 8 \end{aligned}$$

hat beispielsweise die Lösungsmenge

$$\{(x, 4 - x) : x \in \mathbb{R}\}.$$

Nun ändern wir die Perspektive. Wir können das erste Gleichungssystem auch als Matrix wie folgt schreiben:

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x + y \\ 2x + y \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \end{pmatrix}$$

Wir betrachten nun Matrizen allgemein. Eine Matrix über einem Körper K (meistens $\mathbb{R}$ oder $\mathbb{Q}$) ist ein $n \times m$-Gitter von Zahlen

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

Wir können $n \times m$-Matrizen mit Vektoren der Länge m multiplizieren und erhalten dann n-Vektoren. Als Beispiel, betrachten wir

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{n1} \end{pmatrix}$$

Ebenso können wir $n \times m$-Matrizen mit $m \times k$-Matrizen miteinander multiplizieren und erhalten dann eine $n \times k$-Matrix. Als Übung berechnen wir folgendes Produkt gemeinsam.

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -7 & -8 & -2 & -3 & 9 \\ 0 & 3 & -3 & -1 & -1 \\ 8 & 9 & 7 & 9 & 3 \\ 2 & 3 & 8 & 4 & 6 \\ 0 & 3 & -4 & -1 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 1 & 5 & 9 \\ 2 & 6 & 5 & 3 & 5 \\ 8 & 9 & 7 & 9 & 3 \\ 2 & 3 & 8 & 4 & 6 \\ 2 & 6 & 4 & 3 & 3 \end{pmatrix}$$

Nun möchte ich die Perspektive nochmals verändern: Matrizen können auch als Abbildungen betrachtet werden. Wir haben ja bereits gesehen, dass für einen Spaltenvektor $x \in \mathbb{R}^m$ das Produkt $A \cdot x$ definiert ist und in $\mathbb{R}^n$ enthalten ist. So ist eine Matrix eine Abbildung von $\mathbb{R}^m$ nach $\mathbb{R}^n$. Matrizen erfüllen im Allgemeinen die folgenden Rechenregeln.

Proposition. *Sei A eine $n \times m$-Matrix über $\mathbb{R}$. Dann gilt:*

1. *Für alle $x, y \in \mathbb{R}^m$, dass*
$$A \cdot (x + y) = A \cdot x + A \cdot y.$$
2. *Für alle $\lambda \in \mathbb{R}$, dass*
$$A \cdot (\lambda x) = \lambda (A \cdot x).$$

Beweis. Wir haben gesehen, dass falls $A \in \mathbb{R}^{n \times m}$ und $x \in \mathbb{R}^m$, dann ist für $1 \leq i \leq m$

$$(A \cdot x)_i = \sum_{j=1}^{m} a_{ij} x_j.$$

Somit gilt

$$(A \cdot (x + y))_i = \sum_{j=1}^{m} a_{ij}(x_j + y_j) = \sum_{j=1}^{m} a_{ij} x_j + \sum_{j=1}^{m} a_{ij} y_j 0 (A \cdot x)_i + (A \cdot y)_i$$

und

$$(A \cdot (\lambda x))_i = \sum_{j=1}^{m} a_{ij} \lambda x_j = \lambda \sum_{j=1}^{m} a_{ij} x_j = \lambda (A \cdot x)_i$$

□

Abbildungen, welche diese Eigenschaften erfüllen, nennt man *linear*. Aus der obigen Proposition folgt, dass Matrizen festgelegt sind durch das Bild der Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}.$$

Als nächstes möchte ich ein paar weitere Anmerkungen machen.

1. Matrizenmultiplikation ist nicht kommutativ. Betrachten wir zum Beispiel

   $$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{und} \qquad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

   Dann ist

   $$A \cdot B = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = B \cdot A.$$

2. Im Allgemeinen gilt auch $(A \cdot B)^2 \neq A^2 \cdot B^2$. Nehmen wir für den Moment an, dass A und B invertierbar sind. Dann folgt aus

   $$(A \cdot B)^2 = A \cdot B \cdot A \cdot B = A^2 \cdot B^2$$

   durch Linksmultiplkation von A^{-1}

   $$B \cdot A \cdot B = A \cdot B^2$$

   und durch Rechtsmultiplikation von B^{-1}

   $$B \cdot A = A \cdot B.$$

   Wir haben aber bei 1. zwei invertiere Matrizen gefunden, sodass $A \cdot B \neq B \cdot A$. Also sind diese auch ein Beispiel für $(A \cdot B)^2 \neq A^2 \cdot B^2$.

3. Die Identitätsmatrix hat einige verschieden Schreibweisen

   $$I_n = 1_n = \mathrm{Id}_n = \mathrm{id}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

   Die Identitätsmatrix ist die einzige Matrix, welche erfüllt, dass

   $$A \cdot I_n = A$$

   für alle $n \times n$-Matrizen A.

4. Eine Matrix der Form

   $$\begin{pmatrix} 82 & 10 & 182 \\ 0 & 2 & 10 \\ 0 & 0 & 19 \end{pmatrix}$$

   heisst obere Dreiecksmatrix, während eine Matrix der From

   $$\begin{pmatrix} 82 & 0 & 0 \\ 127 & 2 & 0 \\ 1 & 8923 & 332 \end{pmatrix}$$

   untere Dreiecksmatrix heisst.

5 Invertierbare Matrizen

Definition. Sei A eine $n \times n$ Matrix. Die Matrix A heisst invertierbar, falls es eine weitere $n \times n$-Matrix B gibt, sodass

$$A \cdot B = B \cdot A = I_n.$$

Das Inverse wird meist mit A^{-1} bezeichnet.

Betrachten wir nun eine invertierbare $n \times n$-Matrix und einen Spaltenvektor $y \in \mathbb{R}^n$. Dann hat das Gleichungssystem

$$A \cdot x = y$$

genau die Lösung

$$x = A^{-1} \cdot y.$$

Als nächstes möchte ich an folgenden Satz aus der Vorlesung erinnern.

Satz. *Die Menge* $\mathrm{GL}_n(K)$ *der invertierbaren* $n \times n$*-Matrizen über* K *bilden gemeinsam mit der Matrixmultiplikation und der Identitätsmatrix* I_n *eine Gruppe.*

Beweis. Zeigen wir zunächst, dass die Matrixmultiplikation in $\mathrm{GL}_n(K)$ wohldefiniert ist. Dazu wählen wir zwei invertierbare Matrizen A und B und möchten zeigen, dass $A \cdot B$ auch invertierbar ist. Dazu stellen wir fest, dass $B^{-1} \cdot A^{-1}$ ein Inverses von $A \cdot B$ ist, da mit Assoziativität der Matrixmultiplikation folgt, dass

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot I_n \cdot A^{-1} = A \cdot A^{-1} = I_n.$$

Die Assoziativität der Multiplikation in der Menge der invertierbaren Matrizen folgt aus der allgemeinen Assoziativität der Matrixmultiplikation. Ebenso folgt direkt, dass $I_n \in \mathrm{GL}_n(K)$ ein Einheitselement ist. Für A eine invertierbare Matrix gibt es nach Definition eine Inverse A^{-1}, welche selbst auch invertierbar ist. Somit bildet die Menge der invertierbaren Matrizen eine Gruppe. □

Aus diesem Satz folgen direkt einige Eigenschaften von invertierbaren Matrizen. Dies zeigt, wie nützlich die abstrakten Gruppentheorie am Anfang der Vorlesung war. Es seien im Folgenden A und B invertiebare Matrizen.

1. Die Inverse A^{-1} ist eindeutig.
2. Es gilt $(A^{-1})^{-1} = A$.
3. $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.
4. Falls C eine Linksinverse von A ist, so folgt, dass C auch eine Rechtsinverse ist.

Wir geben nun ein paar Beispiele für invertierbare und nicht invertierbare Matrizen an.

1. Die Null-Matrix ist nie invertierbar.

2. Eine Diagonalmatrix
$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_n \end{pmatrix}$$
ist invertierbar genau dann, wenn alle Diagonaleinträge ungleich 0 sind.

3. Eine obere Dreiecksmatrix
$$\begin{pmatrix} a_1 & * & \dots & * \\ 0 & a_2 & \dots & * \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_n \end{pmatrix}$$
ist invertierbar genau dann, wenn alle Diagonaleinträge ungleich 0 sind. Das Gleiche gilt für untere Dreiecksmatrizen.

4. Für eine 2×2-Matrix
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
ist das Inverse
$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
Dabei ist anzumerken, dass das Inverse nur existiert, falls $ad - bc \neq 0$.

Das Inverse kann mit dem Gaussalgorithmus berechnet werden. Ich erkläre in der Übungsstunde wie und weshalb dies funktioniert.

Wir wechseln jetzt wieder den Standpunkt und betrachten eine $n \times m$-Matrix als eine Abbildung $A : \mathbb{R}^m \to \mathbb{R}^n$. Wir haben dann folgenden Satz.

Satz. *Eine $n \times n$-Matrix A ist invertierbar genau dann, wenn $A : \mathbb{R}^n \to \mathbb{R}^n$ eine bijektive Abbildung ist.*

Beweis. Angenommen A ist invertierbar, dann hat die Abbildung A ein beidseitiges Inverses. Somit ist A eine bijektive Abbildung.

Ist A eine bijektive Abbildung, so hat sie eine inverse Abbildung $A^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. Diese Abbildung kann auch als Matrix dargestellt werden, wie wir später sehen werden. Somit ist A^{-1} eine Inverse. □

6 Beispiele von Vektorräumen

Beginnen wir mit dem wichtigsten Beispiel für einen Vektorraum. Wir betrachten die Menge

$$\mathbb{R}^n = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \,:\, x_1, \dots, x_n \in \mathbb{R} \right\}.$$

Wir können zwei Vektoren addieren

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \dots \\ x_n + y_n \end{pmatrix}$$

und wir können Skalare $\lambda \in \mathbb{R}$ multiplizieren

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

Diese Eigenschaften werden in der Definition eines Vektorraumes verallgemeinert.

Definition. Es sei K ein Körper. Ein *Vektorraum über dem Körper K* ist ein Tupel $(V, +, \cdot, 0)$ bestehen aus einer Menge V und zwei Abbildungen

$$+ : V \times V \to V, \qquad \cdot : K \times V \to V$$

und einem ausgezeichneten Element $0_V \in V$ sodass folgende vier Axiome gelten:

1. $(V, +, 0)$ ist eine kommutative Gruppe.
2. (Distributivität) Für alle $\lambda, \lambda' \in K$ und $v, v' \in V$ gilt
$$\lambda \cdot (v + v') = \lambda \cdot v + \lambda \cdot v' \qquad \text{und} \qquad (\lambda + \lambda') \cdot v = \lambda \cdot v + \lambda' \cdot v.$$
3. (Assoziativität) Für alle $\lambda, \mu \in K$ und $v \in V$ gilt
$$\lambda \cdot (\mu \cdot V) = (\lambda \cdot \mu) \cdot v$$
4. (Einselement) Für alle $v \in V$ gilt
$$1_K \cdot v = v.$$

Ich gebe zunächst ein paar Anmerkungen zu der Definition:

1. Manchmal schreibt man λv anstatt $\lambda \cdot v$.
2. Man nennt die Multiplikation in Vektorräumen manchmal auch Skalarmultiplikation und ein Element aus K nennt man einen Skalar.
3. Bei der Assoziativität in der Definition eines Vektorraumes wird auf der einen Seite die Skalarmultiplikation verwendet und auf der anderen Seite die Körper-Multiplikation.

Wir betrachten nun ein paar Beispiele:

1. $\mathbb{R}^n$ is ein $\mathbb{Q}$-Vektorraum, aber kein $\mathbb{C}$-Vektorraum.
2. $\mathbb{C}^n$ ist ein Vektorraum über $\mathbb{C}$, $\mathbb{R}$ und $\mathbb{Q}$.
3. Betrachten wir $\mathbb{F}_2 = \{\overline{0}, \overline{1}\}$. Wir definieren auf $\mathbb{R}$ folgende Skalare-Multiplikation: Für alle $x \in \mathbb{R}$ definieren wir

   $$\overline{0} \cdot x = 0, \qquad \text{und} \qquad \overline{1} \cdot x = x.$$

   Dann ist $\mathbb{R}$ kein Vektorraum über $\mathbb{F}_2$, da Distributivität nicht gilt. Denn aus Distributivität würde für alle $x \in \mathbb{R}$ folgen

   $$2x = x + x = \overline{1} \cdot x + \overline{1} \cdot x = (\overline{1} + \overline{1}) \cdot x = \overline{0} \cdot x = 0,$$

   was ein Widerspruch ist.
4. Die Menge $M_{n,m}(\mathbb{R})$ der $n \times m$-Matrizen ist ein Vektorraum.
5. Die Menge $\{0\}$ ist ein Vektorraum über jedem Körper K.
6. Jeder Körper K ist selbst ein Vektorraum.
7. Wir betrachten nun folgendes wichtige Beispiel. Betrachten wir die Menge

   $$F(\mathbb{R}) := \{\text{Abbildungen } f : \mathbb{R} \to \mathbb{R}\}.$$

   Wir definieren auf dieser Menge die Addition zweier Funktionen als die punktweise Addition: Falls $f, g \in F(\mathbb{R})$, so definieren wir $f + g$, als die Funktion

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

   für $x \in \mathbb{R}$. Die Multiplikation ist auch punktweise definiert: Für $\lambda \in \mathbb{R}$ und $f \in F(\mathbb{R})$ dann definieren wir λf als die Funktion

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

   für $x \in \mathbb{R}$. Dass dies ein Vektorraum ist folgt aus den entsprechenden Eigenschaften von $\mathbb{R}$. Dieses Beispiel lässt sich auf jede Menge M ausweiten. Genauer gesagt bildet die Menge

   $$F(M) := \{\text{Abbildungen } f : M \to \mathbb{R}\}$$

   auf analoge Weise wie obiges Beispiel einen Vektorraum.
8. Die Menge der Folgen in $\mathbb{R}$ ist dasselbe wie $F(\mathbb{N})$. Somit ist die Menge der Folgen auch ein Vektorraum.

7 Mehr zu Vektorräume

7.1 Unterräume

Ein Unterraum ist eine Teilmenge eines Vektorraumes, welche selber mit der induzierten Struktur einen Vektorraum bildet. Wir geben nun eine genaue Definition an.

Definition. Es sei V ein Vektorraum über einem Körper K. Ein *Unterraum* U des Vektorraumes V ist eine Teilmenge $U \subset V$ mit den Eigenschaften:

1. $U \neq \emptyset$.
2. Falls $u, v \in U$ so ist auch $u + v \in U$.
3. Falls $\lambda \in K$ und $u \in U$ so ist auch $\lambda \cdot u \in U$.

Zunächst möchte ich ein kleines Lemma beweisen.

Lemma. *Es sei V ein Vektorraum über K und U ein Untervektorraum von V. Dann ist $0 \in U$.*

Beweis. Da $U \neq \emptyset$ haben wir ein Element $u \in U$. Somit gilt

$$0_V = 0_K \cdot u \in U.$$

□

Wir geben nun einige Beispiele und Gegenbeispiele an.

1. Jeder Vektorraum hat die Unterräume $\{0\}$ und V selbst. Diese Unterräume werden auch als die trivialen Unterräume bezeichnet.
2. Geraden in $\mathbb{R}^2$, welche durch den Mittelpunkt gehen, sind Untervektorräume. Aber Geraden, welche nicht durch den Mittelpunkt gehen sind keine Unterräume.
3. Der Einheitsball $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \subset \mathbb{R}^2$ ist kein Unterraum von $\mathbb{R}^2$. Allgemeiner ist jede beschränkte Menge, welche nicht gleich $\{0_{\mathbb{R}^2}\}$ ist, kein Unterraum.
4. Die Menge der invertierbaren Matrizen $\mathrm{GL}_n(\mathbb{R})$ ist **kein** Untervektorraum von $\mathrm{M}_{n,n}(\mathbb{R})$ da die Nullmatrix nicht invertierbar ist.
5. Erinnern wir uns an den Raum der Funktionen $F(\mathbb{R})$. Die Menge

   $$F_0(\mathbb{R}) = \{f \in F(\mathbb{R}) : f(0) = 0\}$$

   ist ein Unterraum aber

   $$F_1(\mathbb{R}) = \{f \in F(\mathbb{R}) : f(0) = 1\}$$

   ist **kein** Unterraum.

6. Betrachten wir nochmals den Vektorraum der Folgen

$$F(\mathbb{N}) = \{x = (x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{R} \text{ für alle } n \in \mathbb{N}\}.$$

In Analogie zu dem letzten Beispiel stellen wir fest, dass

$$F^0(\mathbb{N}) = \{x = (x_n)_{n\in\mathbb{N}} \in F(\mathbb{N}) : \lim_{n\to\infty} x_n = 0\}$$

ein Unterraum ist aber

$$F^0(\mathbb{N}) = \{x = (x_n)_{n\in\mathbb{N}} \in F(\mathbb{N}) : \lim_{n\to\infty} x_n = 1\}$$

ist **kein** Unterraum.

Als sehr wichtiges Beipsiel für diese Vorlesung möchte ich nun Folgendes beweisen.

Proposition. *Sei A ein $m \times n$-Matrix über dem Körper K. Dann ist die Menge*

$$U = \{x \in K^n : A \cdot x = 0_n\}$$

ein Untervektorraum von K^n.

Diese Menge ist ein wichtiges Objekt im Studium der linearen Algebra und wird *Kern* genannt.

Beweis. Wir stellen fest, dass $0_n \in K^n$ in dieser Menge enthalten ist. Falls $u, v \in U$, so gilt

$$A \cdot (u + v) = A \cdot u + A \cdot v = 0_n + 0_n = 0_n.$$

Somit ist $u + v \in U$. Auch folgt für alle $\lambda \in K$ und $u \in U$, dass $A \cdot (\lambda u) = \lambda \cdot (A \cdot u) = \lambda \cdot 0_n = 0_n$. Also ist U ein Untervektorraum. □

Erinnern wir uns nun des Weiteren, dass jeder Körper K mit der körpereigenen Addition und Multiplikation einen Vektorraum bildet. Dann haben wir folgende Aussage.

Proposition. *Die einzigen K-Unterräume von K sind die trivialen Unterräume $\{0_K\}$ und K.*

Beweis. Es ist klar, dass $\{0_K\} \subset K$ ein Unterraum ist. Betrachten wir nun einen K-Unterraum U von K und nehmen an, dass $U \neq \{0_K\}$. Dann gibt es ein Element $u \in U\backslash\{0\}$. Für jedes $\lambda \in K$ gilt somit

$$k = \frac{k}{u} \cdot u \in U$$

nach dem 3. Unterraum Axiom. □

Wir geben nun ein Beispiel an, dass diese Proposition untermalt.

1. Wir betrachten $\mathbb{R}$ als $\mathbb{R}$-Vektorraum. Dann sind nach der Proposition $\{0\}$ und $\mathbb{R}$ die einzigen $\mathbb{R}$-Unterräume.

2. Wir betrachten $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum. Dann gibt es sehr viele $\mathbb{Q}$-Unterräume von $\mathbb{R}$. Hier einige Beispiele von $\mathbb{Q}$-Unterräumen von $\mathbb{R}$: $\{0\}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$, Für p eine Primzahl ist $\mathbb{Q}[\sqrt{p}]$ ein Unterraum. Aber auch zum Beispiel interessante Mengen wie

$$\mathbb{Q}[\pi] = \left\{ \sum_{i=0}^{n} q_i \pi^i \,:\, q_i \in \mathbb{Q} \right\}$$

für die Kreiszahl $\pi = 3.141592\ldots$.

7.2 Erzeugnis

Es sei V ein Vektorraum über K.

Definition. Betrachten wir eine Teilmenge $S \subset V$. Dann ist das Erzeugnis von S wie folgt definiert:

$$\begin{aligned}\langle S \rangle &= \left\{ \sum_{s \in S} k_s \cdot s \,:\, k_s \in K \text{ und fast alle } a_s = 0 \right\} \\ &= \{k_1 v_1 + \ldots + k_n v_n \,:\, v_1, \ldots, v_n \in S \text{ und } k_1, \ldots, k_n \in K\}.\end{aligned}$$

Betrachten wir nur eine endliche Teilmenge $S = \{v_1, \ldots, v_n\}$, so vereinfacht sich die Defintion:

$$\langle S \rangle = \langle v_1, \ldots, v_n \rangle = \{k_1 v_1 + \ldots + k_n v_n \,:\, k_1, \ldots, k_n \in K\}.$$

Wir betrachten nun ein paar Beispiele:

1. $\langle 0 \rangle = 0$ und $\langle V \rangle = V$.
2. Es gilt in $\mathbb{R}^2$, dass
$$\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^2.$$
3. Betrachten wir $\mathbb{R}$ als $\mathbb{R}$-Vektorraum: Dann ist
$$\langle 1, \sqrt{2} \rangle = \mathbb{R} = \langle 1 \rangle = \langle \sqrt{2} \rangle.$$
4. Betrachten wir $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum: Dann ist
$$\begin{aligned}\langle 1 \rangle &= \mathbb{Q}, \\ \langle \sqrt{2} \rangle &= \sqrt{2} \cdot \mathbb{Q} = \{\sqrt{2} \cdot q \,:\, q \in \mathbb{Q}\}, \\ \langle 1, \sqrt{2} \rangle &= \mathbb{Q}[\sqrt{2}].\end{aligned}$$

7.3 Lineare Unabhägigkeit

Es sei V ein Vektorraum über K.

Definition. Eine Teilmenge $S \subset V$ heisst *linear abhängig*, falls Elemente $v_1, \ldots, v_n \in S$ existieren und Skalare $k_1, \ldots, k_n$ sodass nicht alle $k_1, \ldots, k_n$ gleich 0 sind mit $k_1 v_1 + \ldots + k_n v_n = 0$.

Falls keine solchen Skalare existieren, so heisst S *linear unabhängig*.

Wir betrachten nun ein paar Beispiele:

1. Sei $V = \mathbb{R}^2$. Die Vektoren
$$\begin{pmatrix}1\\0\end{pmatrix} \quad \text{und} \quad \begin{pmatrix}2\\0\end{pmatrix}$$
sind linear abhängig.
2. Betrachten wir $\mathbb{R}$ als $\mathbb{R}$-Vektorraum. Dann sind $\{1, \sqrt{2}\}$ linear abhängig.
3. Betrachten wir $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum. Dann sind $\{1, \sqrt{2}\}$ linear unabhängig.
4. Sei $V = \mathbb{R}^2$. Die Vektoren
$$\begin{pmatrix}1\\0\end{pmatrix} \quad \text{und} \quad \begin{pmatrix}0\\1\end{pmatrix}$$
sind linear unabhängig.
5. Betrachten wir den Vektorraum
$$F(\mathbb{R}) = \{\text{Abbildungen } f : \mathbb{R} \to \mathbb{R}\}.$$
Dazu betrachten wir für jedes $x \in \mathbb{R}$ die Funktionen δ_x welche wie folgt definiert ist
$$\delta_x(y) := \begin{cases} 1 & x = y, \\ 0 & x \neq y. \end{cases}$$
Wir behaupten nun das Folgende.

Proposition. *Die Funktionen $(\delta_x)_{x\in\mathbb{R}}$ sind linear unabhängig.*

Beweis. Betrachten $x_1, \ldots, x_n \in \mathbb{R}$ und Koeffizienten $r_1, \ldots r_n \in \mathbb{R}$ und nehmen an, dass
$$\sum_{j=1}^{n} r_j \delta_{x_j} = 0.$$
Evaluieren obiger Summe an x_i, so erhalten wir
$$\sum_{j=1}^{n} r_j \delta_{x_j}(x_i) = r_i = 0.$$
Somit sind alle $r_i = 0$, also sind die Funktionen $(\delta_x)_{x\in\mathbb{R}}$ linear unabhängig. □

Proposition. *Betrachten wir für $n \in \mathbb{N}$ die Funktionen*
$$f_n(x) = x^n.$$
Die Menge $(f_n)_{n\in\mathbb{N}}$ ist linear unabhängig.

Beweis. Betrachten wir Zahlen $n_1, \ldots, n_i \in \mathbb{N}$ und Koeffizienten $r_1, \ldots r_i \in \mathbb{R}$ und nehmen an, dass
$$\sum_{j=1}^{i} r_j x^{n_j} = 0$$
für alle $x \in \mathbb{R}$. Da nur das Nullpolynom überall 0 ist, folgt $r_1 = \ldots = r_i = 0$. □

7.4 Basis

Es sei V ein Vektorraum über K.

Definition. Eine Basis eines Vektorraumes ist ein linear unabhängiges Erzeugendensystem.

Folgende Eigenschaften wissen wir aus der Vorlesung:

1. Jeder Vektorraum hat eine Basis.
2. Jede Basis eines Vektorraumes besitzt die gleiche Anzahl Elemente.

Ich möchte nun ausführlich zeigen, dass

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}$$

eine Basis von $\mathbb{R}^n$ ist.

Proposition. *Die Vektoren $e_1 \dots, e_n$ bilden eine Basis von $\mathbb{R}^n$.*

Beweis. Wir zeigen zuerst, dass $e_1 \dots, e_n$ ein Erzeugendensystem ist. Dazu betrachten wir

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Dann ist

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix} + \dots, + x_n \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}$$

und somit erzeugen $e_1, \dots, e_n$ den ganzen Raum $\mathbb{R}^n$. Für die lineare Unabhängigkeit betrachten wir die Gleichung

$$\begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix} + \dots, + x_n \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}.$$

Es folgt somit $x_1 = \dots = x_n = 0$. Somit ist $e_1, \dots, e_n$ eine Basis. □

8 Basen, Dimensionen und direkte Summen

8.1 Drei wichtige Vektorräume

Zuerst betrachten wir den Polynomring in der Variable X über dem Körper K

$$K[X] = \left\{ \sum_{i=0}^{n} a_i X^i \,:\, a_i \in K \right\}.$$

Dieser bildet mit der Addition von Polynomen und der Skalarenmultiplikation einen Vektorraum.

Als nächstes betrachten wir die Menge

$$C([0,1]) := \{f : [0,1] \to \mathbb{R} \text{ stetig}\},$$

welche wieder mit der Addition von Funktionen und der Skalarenmultiplikation einen Vektorraum bildet.

Zuletzt definieren wir den Vektorraum der Folgen

$$F(\mathbb{N}) = \{(x_i)_{i\in\mathbb{N}} \,:\, x_i \in \mathbb{R}\}.$$

8.2 Basis und Dimension

Es sei V ein Vektorraum über K.

Definition. Eine Basis eines Vektorraumes ist ein linear unabhängiges Erzeugendensystem.

Folgende Eigenschaften wissen wir aus der Vorlesung:

1. Jeder Vektorraum hat eine Basis.
2. Jede Basis eines Vektorraumes besitzt die gleiche Anzahl Elemente.

Definition. Es sei V ein Vektorraum über K. Dann ist die Kardinalität jeder Basis die *Dimension* von V über K. Wir schreiben $\dim_K(V)$ für die Dimension von V über K.

Proposition. *Sei V ein Vektorraum der Dimension n über K und $v_1, \ldots, v_n$ eine Basis von V. Dann haben wir eine bijektive Abbildung*

$$\Phi : K^n \to V, \qquad (x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} x_i v_i.$$

Beweis. Φ ist injektiv, da eine Basis aus linear unabhängigen Elementen besteht und sujrektiv, da eine Basis erzeugend ist. □

Proposition. *Seien $v_1, \ldots, v_n \in K^n$. Dann ist $v_1, \ldots, v_n$ eine Basis von K^n genau dann, wenn die Matrix*

$$A = \begin{pmatrix} | & & | \\ v_1 & \ldots & v_n \\ | & & | \end{pmatrix} \in K^{n\times n}$$

invertierbar ist.

Beweis. Ist $v_1, \ldots, v_n$ eine Basis, so ist die Abbildung $L_A : K^n \to K^n$ bijektiv, also ist A invertierbar.

Ist umgekehrt A invertierbar, so ist $L_A : K^n \to K^n$ bijektiv, also ist $v_1, \ldots, v_n$ erzeugend und linear unabhängig. Somit ist $v_1, \ldots, v_n$ eine Basis. □

Betrachten wir ein paar Beispiele für Basen:

1. Betrachten wir $\mathbb{R}^2$. Zwei Vektoren
$$\begin{pmatrix} a \\ b \end{pmatrix} \text{ und } \begin{pmatrix} c \\ d \end{pmatrix}$$
bilden eine Basis genau dann, wenn
$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$
invertierbar ist. Dies ist genau dann der Fall, wenn $ad - bc \neq 0$.
2. Sei K ein Körper. Dann ist jedes Element $k \in K \backslash \{0\}$ ein Basis des Vektorraumes. Das gleiche gilt für jeden Vektorraum der Dimension 1.
3. $1, \sqrt{2}$ ist eine Basis von $\mathbb{Q}[\sqrt{2}]$.
4. $1 + \sqrt{2}$ und $2 + \sqrt{2}$ ist eine Basis von $\mathbb{Q}[\sqrt{2}]$.
5. $1, X, X^2, \ldots$ ist eine Basis des Polynomringes $\mathbb{R}[X]$.
6. $1, X, 2X^2, 3X^3, \ldots$ ist eine Basis des Polynomringes $\mathbb{R}[X]$.
7. $1, X + 1, X^2 + 2, X^3 + 3, \ldots$ ist eine Basis des Polynomringes $\mathbb{R}[X]$.

Nun betrachten wir ein paar Beispiele zu Dimensionen.

1. $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$.
2. $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$ und $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$.
3. Die Dimension von $\mathbb{R}$ über $\mathbb{Q}$ is überabzählbar. Dies folgt, da falls die Dimension von $\mathbb{R}$ über $\mathbb{Q}$ abzählbar wäre, dann würde auch $\mathbb{R}$ abzählbar sein. Dies ist aber nicht der Fall.
4. Die Dimension von $\mathbb{Q}[\sqrt{2}]$ über $\mathbb{Q}$ ist 2.
5. Die Dimension von $\mathbb{R}[X]$ ist unendlich als $\mathbb{R}$-Vektorraum.
6. Die Dimension von $C([0, 1])$ ist unendlich, da die Monome $(x^n)_{n \in \mathbb{N}}$ ein unendliches linear unabhängiges System bilden.
7. Die Dimension von $F(\mathbb{N})$ ist unendlich, da die Folgen $\delta^i \in F(\mathbb{N})$ definiert durch
$$\delta^i_n = \begin{cases} 1 & i = n \\ 0 & i \neq n \end{cases}$$
ein unendliches linear unabhängiges System bilden.

8.3 Direkte Summen und Komplemente

Definition. Seien V_1 und V_2 Unterräume von V mit $V_1+V_2 = V$ und $V_1 \cap V_2 = \emptyset$. Dann ist V die *direkte Summe* von V_1 und V_2 und wir schreiben

$$V = V_1 \oplus V_2.$$

Weiter heisst dann V_2 ein Komplement von V_1 in V.

Es gelten folgende Eigenschaften:

1. Falls $V = V_1 \oplus V_2$, dann ist $\dim(V) = \dim(V_1) + \dim(V_2)$.
2. Allgemeiner gilt:

$$\dim(V_1 + V_2) + \dim(V_1 \cap V_2) = \dim(V_1) + \dim(V_2).$$

Hier ein paar Beispiele:

1. $V = V \oplus \{0\}$.
2. Wir haben
$$\mathbb{R} = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle.$$
3. Es gilt $\mathbb{Q}[\sqrt{2}] = \mathbb{Q} \oplus \sqrt{2}\mathbb{Q}$.

Allgemeiner können wir die n-fache direkte Summe definieren.

Definition. Seien $V_1, \ldots, V_n$ Unterräume von V mit

$$V_1 + \ldots + V_n = V$$

und für alle $i \in \{1, \ldots, n\}$ haben wir

$$V_i \cap \left(\sum_{j \in \{1,\ldots,n\} \setminus \{i\}} V_j \right) = \{0\}.$$

Dann ist V die *direkte Summe* von $V_1, \ldots, V_n$ und wir schreiben

$$V = \bigoplus_{i=1}^{n} V_i.$$

Proposition. *Ist V die direkte Summe der Unterräume $V_1, \ldots, V_n$ so haben wir eine bijektive Abbildung*

$$V_1 \times \ldots \times V_n \to V, \qquad (v_1, \ldots, v_n) \mapsto \sum_{i=1}^{n} v_n.$$

Beweis. Da $V = V_1 + \ldots + V_n$ ist die Abbildung surjektiv. Nehmen wir nun an, wir haben zwei Element $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in V_1 \times \ldots \times V_n$ sodass

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{m} w_i.$$

Dann ist

$$\sum_{i=1}^{n}(v_i - w_i) = 0.$$

Somit ist $v_1 - w_1 = \sum_{i=2}^{n} w_i - v_i$ und somit ist nach Voraussetzung $v_1 - w_1 = 0$ und auch $\sum_{i=2}^{n} w_i - v_i = 0$. Nun sezten wir diesen Prozess fort und folgern dann, dass $v_i = w_i$ für alle $i \in \{1, \ldots, n\}$. Somit ist die Abbildung injektiv. □

Korollar. *Ist $v_1, \ldots, v_n$ eine Basis von K^n, so gilt*

$$K^n = \bigoplus_{i=1}^{n}\langle v_i \rangle.$$

Beweis. Dies folgt direkt aus dem letzten Satz. □

9 Lineare Abbildungen

Definition. Sei K ein Körper und V und W Vektorräume über K. Eine Abbildung $f : V \to W$ heisst linear (oder K-linear), falls

1. Für alle $v_1, v_2 \in V$ gilt $f(v_1 + v_2) = f(v_1) + f(v_2)$.
2. Für alle $v_1 \in V$ und $\alpha \in K$ gilt $f(\alpha v_1) = \alpha f(v_1)$.

Ein paar Anmerkungen zu linearen Abbildungen:

1. Induktiv folgt für jede lineare Abbildung $f : V \to W$ und alle $x_1, \ldots, x_n \in V$ und $\lambda_1, \ldots, \lambda_n \in K$
$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i f(x_i)$$
2. Es ist wichtig, dass beide Vektorräume über dem gleichen Körper sind.
3. Lineare Abbildungen erhalten die Struktur eines Vektorraumes

Das wichtigste Beispiel für eine lineare Abbildung ist eine Matrix. Genauer gesagt, betrachten wir eine $n \times m$-Matrix A mit Koeffizienten in K. Dann definieren wir die Abbildung

$$L_A : K^m \to K^n, \qquad x \mapsto A \cdot x.$$

Gehen wir nun auf die Abbildung L_A für eine Matrix genauer ein. Wir haben schon gesehen, dass für eine Matrix A die Abbildung L_A linear ist. Wir möchten nun erklären, wie Matrixmultiplikation als die Komposition zweier linearer Abbildungen definiert ist.

Dazu ändern wir für den moment den Standpunkt. Gegeben seien zwei Matrizen $A \in K^{m \times n}$ und $B \in K^{n \times l}$. Wir wollen das Produkt $C := B \cdot A$ definieren, sodass

$$L_{B \cdot A} = L_B \circ L_A.$$

Ist dies der Fall, so folgt für $C = (c_{ij})_{1 \le i \le m, 1 \le j \le n}$ durch Einsetzen von e_i, dass

$$\begin{aligned}
\begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{li} \end{pmatrix} &= L_C(e_i) = L_{B \cdot A}(e_i) \\
&= (L_B \circ L_A)(e_i) = L_B(L_A(e_i)) \\
&= L_B\left(\begin{pmatrix} a_{1i} \\ a_{2i} \\ \ldots \\ a_{mi} \end{pmatrix}\right) = B \cdot \begin{pmatrix} a_{1i} \\ a_{2i} \\ \ldots \\ a_{mi} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{j=1}^{n} b_{1j} a_{ji} \\ \sum_{j=1}^{n} b_{2j} a_{ji} \\ \vdots \\ \sum_{j=1}^{n} b_{lj} a_{ji} \end{pmatrix}
\end{aligned}$$

und somit folgt

$$c_{ki} = \sum_{j=1}^{n} b_{kj} a_{ji}.$$

Damit erhalten wir genau die Definition der Matrixmultiplikation. Die ist in folgender Proposition zusammengefasst.

Proposition. *Gegeben seien zwei Matrizen $A \in K^{m\times n}$ und $B \in K^{n\times l}$. Dann gilt*

$$L_{A\cdot B} = L_A \circ L_B.$$

Korollar. *Die Matrixmultiplikation ist assoziativ.*

Beweis. Dies folgt da die Komposition von Abbildungen assoziativ ist ist. □

Ist umgekehrt eine lineare Abbildung

$$f : K^n \to K^m$$

gegeben, dann können wir diese auch auf folgende Weise durch eine Matrix darstellen: Bezeichnen wir mit

$$v_i = f(e_i)$$

für e_i den Standardbasisvektor. Betrachten wir dann die Matrix

$$A = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}.$$

Dann gilt da f linear ist für jedes Element $x = (x_1, \dots, x_n) \in K^n$, dass

$$f(x) = f\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} x_i v_i = Ax.$$

Also ist $f = L_A$. Also wird jede lineare Abbildung zwischen K^n und K^m durch eine Matrix dargestellt. Wir fassen dies in folgender Proposition zusammen, wobei die definieren

$$M_{n,m}(K) := \{n \times m - \text{Matrizen über } K\}$$
$$\mathrm{Hom}_K(K^m, K^n) := \{\text{lineare Abbildungen } f : K^m \to K^n\}$$

Satz. *Die Abbildung*

$$M_{n,m}(K) \to \mathrm{Hom}_K(K^m, K^n), \qquad A \mapsto L_A$$

ist eine lineare Bijektion.

Beweis. Ich überlasse als Übung zu zeigen, dass diese Abbildung injektiv und linear ist. Die Surjektivität folgt aus obiger Konstruktion. □

Als zweites möchte ich folgende Proposition beweisen, für welche wir *Isomorphismen* wiederholen.

Definition. Seien V und W Vektorräume über K. Eine lineare Abbildung $f : V \to W$ heisst ein Isomorphismus, falls f bijektiv ist.

Proposition. *Sei V ein n-dimensionaler Vektorraum mit Basis $v_1, \ldots, v_n$. Dann ist die Abbildung*

$$\varphi : K^n \to V, \qquad (\lambda_1, \ldots, \lambda_n) \mapsto \sum_{i=1}^{n} \lambda_i v_i$$

ein Isomorphismus.

Beweis. Es ist klar, dass diese Abbildung linear ist. Sie ist bijektiv, da $v_1, \ldots, v_n$ eine Basis ist. □

Überlegen wir uns nun ein paar Beispiele:

1. Die Nullabbildung $V \to W, v \mapsto 0_W$ ist linear.
2. Die Identität $V \to V, v \mapsto v$ ist linear.
3. Sei A eine $m \times n$ Matrix. Dann ist die Abbildung
$$L_A : K^n \to K^m, \qquad v \mapsto A \cdot v$$
linear.
4. Die Abbildung
$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto x \cdot y$$
ist **nicht** linear.
5. Die komplexe Konjugation
$$\iota : \mathbb{C} \to \mathbb{C}, \qquad z = a + ib \mapsto \iota(z) = a - ib$$
is $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.
6. Betrachten wir den $\mathbb{R}$-Vektorraum der stetigen Funkionen auf $[0, 1]$
$$C([0, 1]) := \{f : [0, 1] \to \mathbb{R} \text{ stetig}\}.$$
Wir stellen fest, dass jede Funktion in $C([0, 1])$ integriebar ist und ein endliches Integral hat. Die Abbildung
$$I : C([0, 1]) \to \mathbb{R}, \qquad f \mapsto \int_0^1 f(x)\, dx$$
ist linear. Ebenso ist die Abbildung
$$I' : C([0, 1]) \to \mathbb{R}, \qquad f \mapsto \int_0^1 f(x) \cdot x^2\, dx$$
linear. Die Abbildung
$$I'' : C([0, 1]) \to \mathbb{R} \qquad f \mapsto \int_0^1 f(x) + 1\, dx$$
ist jedoch **nicht** linear.

7. Betrachten wir den $\mathbb{R}$-Vektorraum der unendlich oft differenzierbaren Funktionen

$$C^\infty(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \text{ unendlich oft differenzierbar}\}.$$

Die Ableitung

$$D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad f \mapsto f'$$

ist linear. Auch die zweite Ableitung

$$D' : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad f \mapsto f''$$

ist linear sowie beispielsweise

$$D'' : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad f \mapsto f' + f''$$

aber nicht

$$D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad f \mapsto f' + 1.$$

10 Kern, Bild, Direkte Summen und Basiswechselmatrizen

10.1 Kern und Bild

Wir besprechen zunächst ein paar Aspekte von der letzten Serie. Als ersten möchte ich erklären wie man effizient den Kern und das Bild einer Matrix

$$A = \begin{pmatrix} | & | & \dots & | \\ v_1 & v_2 & \dots & v_n \\ | & | & \dots & | \end{pmatrix}$$

bestimmt. Ich schlage folgendes Vorgehen vor:

1. Bringe die Matrix in Zeilenstufenform.
2. Bestimme die Dimension des Bildes: In der Zeilenstufenfrom können wir die Dimension des Bildes leicht ablesen. Dies ist nämlich einfach die Dimesion des Erzeugnisses von den Vektoren in Zeilenstufenform.
3. Finde eine Basis des Bildes: Die Vektoren $v_1, \dots, v_n$ erzeugen das Bild. Also wählen wir einfach entsprechend der Dimension des Bildes viele linear unabhängige Vektoren aus $v_1, \dots, v_n$ aus um eine Basis des Bildes zu finden.
4. Bestimme die Dimension des Kerns: Verwende dazu folgende Formel. Für eine lineare Abbildung $f : V \to W$ gilt
$$\dim(V) = \dim(\text{Kern}(f)) + \dim(\text{Bild}(f)).$$
5. Bestimme eine Basis des Kerns: Wir kennen die Dimension des Kerns. Bemerke, dass die Matrix in Zeilenstufenform den genau gleichen Kern hat wie die ursprüngliche Matrix. Meist ist am effizientesten dann einfach linear unabhängige Elemente des Kerns zu erraten. Somit erhalten wir eine Basis des Kerns.

Betrachten wir nun konkret die Matrix

$$A = \begin{pmatrix} 2 & 7 & 14 \\ 1 & 3 & 6 \\ 2 & 6 & 12 \end{pmatrix}.$$

In unserem Beispiel ist die Zeilenstufenform

$$\begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Also ist die Dimension des Bildes 2 und die des Kerns 1. Da die ersten beiden Spaltenvektoren der Matrix linear unabhängig sind, so folgt, dass

$$\text{Bild}(A) = \left\langle \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 7 \\ 3 \\ 6 \end{pmatrix} \right\rangle.$$

Da die Dimension des Kernes gleich 1 ist, müssen wir nur ein nichttriviales Element des Kerns finden. Durch raten sehen wir

$$\mathrm{Kern}(A) = \left\langle \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \right\rangle.$$

Als zweiten möchte ich Aufgabe 4a) vorlösen. Dazu möchte ich zuerst folgende Aussage beweisen:

Proposition. *Sei $f : A \to B$ eine Abbildung zwischen Mengen. Dann ist f injektiv genau dann, wenn eine Abbildung $g : B \to A$ existiert mit*

$$g \circ f = \mathrm{id}_A$$

Beweis. Gilt diese Eigenschaft, so folgt falls $f(x) = f(y)$ für $x, y \in A$, dass dann auch

$$x = g(f(x)) = g(f(y)) = y.$$

Also ist f injektiv.

Nehmen wir nun an, dass f injektiv ist und betrachten wir

$$\mathrm{Bild}(f) = \{f(x) \,:\, x \in A\}.$$

Für jedes $y \in \mathrm{Bild}(f)$, existiert eindeutiges $x \in A$ sodass $f(x) = y$. Somit können wir folgende Abbildung definieren:

$$g : B \to A, \qquad y \mapsto \begin{cases} x & \text{falls } y \in \mathrm{Bild}(f) \text{ mit } f(x) = y \\ \text{egal} & \text{falls } y \notin \mathrm{Bild}(f). \end{cases}$$

Wir haben also im allgemeinen sehr viele Wahlmöglichkeiten für die Funktion g. Es gilt nach Definition

$$g \circ f = \mathrm{id}_A.$$

Somit ist die Aussage gezeigt. □

Wir können nun die gleiche Aussage für lineare Abbildungen beweisen, aber wir verlangen, dass die Umkehrabbildung auch linear ist.

Proposition. *Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W. Dann ist f injektiv genau dann, wenn eine **lineare** Abbildung $g : W \to V$ existiert mit*

$$g \circ f = \mathrm{id}_V.$$

Beweis. Falls solch eine Abbildung existiert, können wir gleich vorgehen, wie in der letzten Proposition.

Nehmen wir umgekehrt an, dass f injektiv ist. Betrachten wir nun ein Komplement W' von $\mathrm{Bild}(f)$ und betrachten die Abbildung

$$g : W \to V, \qquad v = f(x) + w' \in \mathrm{Bild}(f) \oplus W' \mapsto x,$$

wobei wir anmerken, dass diese Abbildung wohldefiniert ist, da f injektiv ist. Es ist gilt, dass

$$g \circ f = \mathrm{id}_V,$$

jedoch bleibt zu überprüfen, dass g linear ist. Dazu betrachten wir $v_1 = f(x_1)+w_1'$ und $v_2 = f(x_2) + w_2'$. Dann gilt

$$\begin{aligned} g(v_1 + v_2) &= g(f(x_1) + w_1' + f(x_2) + w_2') = g(f(x_1 + x_2) + w_1' + w_2') \\ &= x_1 + x_2 = g(v_1)2g(v_2), \end{aligned}$$

da f linear ist, und für $\lambda \in K$, dass

$$g(\lambda v_1) = g(\lambda(f(x_1) + w_1')) = g(f(\lambda x_1) + \lambda w_1') = \lambda x_1 = \lambda g(v_1).$$

Somit ist g linear und die Aussage somit gezeigt. □

10.2 Eine wichtige Eigenschaft von endlich-dimensionalen Vektorräumen

Satz. *Sei V ein endlich-dimensionaler Vektorraum mit einer geordneten Basis $B = (v_1, \ldots, v_n)$. Dann ist die Abbildung*

$$\varphi_B : K^n \to V, \qquad (x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} x_i v_i$$

ein Isomorphismus.

Beweis. Da eine Basis eine linear unabhängig und erzeugend ist, folgt, dass die Abbildung φ_B surjektiv und injektiv ist. Es folgt direkt, dass diese Abbildung linear ist. □

Korollar. *Jeder endlichdimensionale Vektorraum über K der Dimension n ist isomorph zu K^n.*

Dies bedeutet, dass jeder endlich-dimensinoale Vektorraum die gleiche Vektorraumstruktur hat wie K^n. Anders gesagt gibt es eigentlich nur einen Vektorraum der Dimension n oder die Struktur des Vektorraumes ist komplett durch seine Dimension festgelegt. Also könnte man eigentlich sagen, dass nur *langweilige* endlich-dimensinoale Vektorräume gibt. Diese Eigenschaft ist aber auch sehr nützlich, denn

Es gibt auch nur *langweilige* unendlichdimensionale Vektorräume, wie wir nun mit direkten Summen zeigen werden.

10.3 Direkte Summe

Definition. Betrachten wir eine Indexmenge I und Vektorräume $(V_i)_{i\in I}$. Dann definieren wir die (äussere) *direkte Summe* von $(V_i)_{i\in I}$ als

$$\bigoplus_{i\in I} V_i = \{(v_i)_{i\in I} \,:\, v_i \in V_i \text{ für alle } i \in I \text{ mit fast allen } v_i = 0\}.$$

Zunächst als Beispiel, stellen wir fest, dass

$$K^n = \bigoplus_{i=1}^{n} K.$$

Man könnte dies eigentlich als die Definition von K^n betrachten. Wir verallgemeinern nun den Satz aus Kapitel 10.2.

Satz. *Sei V ein Vektorraum mit einer Basis $B = (v_i)_{i\in I}$. Dann haben wir einen Isomorphismus*

$$\varphi_B : \bigoplus_{i\in I} K \mapsto V, \qquad (x_i)_{i\in I} \mapsto \sum_{i\in I} x_i v_i.$$

Beweis. Diese Abbildung ist wohldefiniert, da wir eine endliche Summe betrachten. Die Abbildung φ_B ist bijektiv, da B eine Basis ist. Es folgt direkt, dass diese Abbildung linear ist. □

Somit sehen wir, dass jeder Vektorraum V isomorph ist zu

$$\bigoplus_{i\in I} K,$$

wobei I eine Indexmenge ist, mit der gleichen Kardinalität wie irgendeine Basis von V. Das Objekt

$$\bigoplus_{i\in I} K$$

hängt nur von der Kardinalität von I ab. Also gilt wieder die Aussage, dass ein Vektorraum komplett durch seine Dimension festgelegt ist.

Ich gebe jetzt noch zwei Beispiele an.

Beispiel. Zunächst der Polynomring $K[X]$ is isomorph

$$K[X] \cong \bigoplus_{i\in\mathbb{N}} K.$$

Beispiel. Zweitens betrachten wir $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum. Es gilt, dass die Dimension von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum gleich der Kardinalität von $\mathbb{R}$ ist. Also haben wir einen Isomorphismus

$$\mathbb{R} \cong \bigoplus_{x\in\mathbb{R}} \mathbb{Q}.$$

10.4 Basiswechselmatrizen

Fangen wir mit einem einfachen Beispiel an. Betrachten wir $\mathbb{R}^2$ die geordneten Basen

$$B = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)$$

und

$$B' = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$$

Angenommen wir haben einen Vektor

$$v = \begin{pmatrix}a\\b\end{pmatrix} = a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} \in \mathbb{R}$$

mit $a, b \in \mathbb{R}^2$ und wir wollen uns überlegen wie wir ihn darstellen können als

$$v = c\begin{pmatrix}1\\0\end{pmatrix} + d\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}c+d\\d\end{pmatrix}$$

In diesem Fall könnten wir jetzt einfach diese beiden Gleichungen gleichsetzen und nach c und d auflösen und erhalten dann $d = b$ und $c = a - b$. Andererseits können wir auch die Abbildungen

$$\varphi_B : \mathbb{R}^2 \to \mathbb{R}^2, \qquad (a, b) \mapsto a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

und

$$\varphi_{B'} : \mathbb{R}^2 \to \mathbb{R}^2, \qquad (c, d) \mapsto c \begin{pmatrix} 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Dann ist also

$$\begin{pmatrix} c \\ d \end{pmatrix} = \varphi_{B'}^{-1} \circ \varphi_B \left(\begin{pmatrix} a \\ b \end{pmatrix} \right)$$

Wie können wir nun (c, d) berechnen? Dazu stellen wir fest, dass die Abbildung $\varphi_{B'}^{-1} \circ \varphi_B$ linear ist und somit durch eine Matrix $A \in M_{2,2}(\mathbb{R})$ gegeben ist und zwar ist

$$A = \begin{pmatrix} \varphi_{B'}^{-1} \circ \varphi_B(e_1) & \varphi_{B'}^{-1} \circ \varphi_B(e_2) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Also sehen wir wieder, dass

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a - b \\ b \end{pmatrix}.$$

Definition. Sei V ein Vektorraum über K der Dimension n und B und B' geordnete Basen. Die Basiswechselmatrix von B nach B' ist die eindeutige Matrix $A \in M_{n,n}(K)$, sodass

$$L_A = \varphi_{B'}^{-1} \circ \varphi_B.$$

Diese Matrix ist die eindeutige Matrix, sodass folgendes Diagram kommutiert:

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle \varphi_B}\uparrow & & \uparrow{\scriptstyle \varphi_{B'}} \\ K^n & \xrightarrow{L_A} & K^n \end{array}$$

Wir geben nun wieder zwei Beispiele an.

Beispiel. Betrachten wir zunächst K^n und sei

$$B = (b_1, \ldots, b_n)$$

eine geordnete Basis mit $b_1, \ldots, b_n \in K^n$ und sei

$$B' = (e_1, \ldots, e_n)$$

die Standardbasis. Dann folgt, dass $\varphi_{B'} = \mathrm{id}$ und somit gilt für die Basiswechselmatrix A, dass

$$L_A = \varphi_B$$

und somit ist

$$A = \begin{pmatrix} | & | & \ldots & | \\ b_1 & b_2 & \ldots & b_n \\ | & | & \ldots & | \end{pmatrix}.$$

Beispiel. Als zweites Beispiel betrachten wir die Polynome in $\mathbb{R}$ mit Grad ≤ 2

$$P_2(\mathbb{R}) = \{f \in \mathbb{R}[X] \,:\, \deg(f) \leq 2\} = \{a + bX + cX^2 \,:\, a, b, c \in \mathbb{R}\}.$$

Wir betrachten die geordnete Standardbasis

$$B = (1, X, X^2)$$

und die Basis

$$B' = (1, X + 1, (X + 1)^2).$$

Wir bemerken

$$L_A\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right) = \varphi_{B'}^{-1}(1) = \begin{pmatrix}1\\0\\0\end{pmatrix}$$
$$L_A\left(\begin{pmatrix}0\\1\\0\end{pmatrix}\right) = \varphi_{B'}^{-1}(X) = \begin{pmatrix}-1\\1\\0\end{pmatrix}$$
$$L_A\left(\begin{pmatrix}0\\1\\0\end{pmatrix}\right) = \varphi_{B'}^{-1}(X^2) = \begin{pmatrix}1\\-2\\1\end{pmatrix}$$

also ist die Basiswechselmatrix

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

11 Dualraum, Quotientenraum

11.1 Dualraum

Definition. Sei V ein Vektorraum über K. Dann definieren wir den *Dualraum* als

$$V^* = \mathrm{Hom}_K(V, K) = \{\text{Lineare Abbildungen } V \to K\}.$$

Elemente des Dualraumes, also lineare Abbildungen von V nach K, nennt man *Linearformen*.

Der Dualraum bildet mit der punktweisen Addition und der punktweisen Skalarmultiplikation einen Vektorraum über K. Im Folgenden diskutiere ich ein paar Beispiele:

1. Betrachten wir $\mathbb{R}^2$ als $\mathbb{R}$-Vektorraum. Ich gebe nun zwei Beispiele für Linearformen an, nämlich

   $$f_1 : \mathbb{R}^2 \to \mathbb{R}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x$$

   und

   $$f_2 : \mathbb{R}^2 \to \mathbb{R}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto y.$$

   Allgemeiner haben wir ja bereits gesehen, dass jede lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}$ durch eine Matrix L_A, in diesem Fall eine 2×1-Matrix, gegeben ist. Also haben wir für jede Linearform $f : \mathbb{R}^2 \to \mathbb{R}$ eine Matrix $A = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$ sodass

   $$f : \mathbb{R}^2 \to \mathbb{R}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = a_1 x_1 + a_2 x_2.$$

2. Dies geht auch ganz allgemein. Falls $V = K^n$ und $f : K^n \to K$ eine lineare Abbildung ist, so gibt es eine Matrix $A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$ sodass $f = L_A$. Dies ist ein Spezialfall des Isomorphismuses

   $$\mathrm{M}_{m,n}(K) \to \mathrm{Hom}_K(K^n, K^m), \qquad A \mapsto L_A.$$

3. Betrachten wir

   $$C([0,1]) = \{f : [0,1] \to \mathbb{R} \text{ stetig}\}.$$

   Dann definiert das Integral

   $$I : C([0,1]) \to \mathbb{F}, \qquad f \mapsto \int f(x)\, dx$$

   eine Linearform auf $C([0,1])$.

Als Beispiel können wir noch eine Aufgabe aus dem Single Choice Test von dieser Woche besprechen.

Proposition. *Sei $V = K^n$. Dann ist die Abbildung*

$$\varphi : K^n \to V^* = \mathrm{Hom}_K(K^n, K), \qquad v \mapsto (w \mapsto f_v(w) = v^T \cdot w)$$

ein Isomorphismus.

Beweis. Um zu zeigen, dass die Abbildung wohldefiniert ist, müssen wir überprüfen, dass für jedes $v \in K^n$ die Abbildung $f_v : K^n \to K$ linear ist. Dazu genügt es festzustellen, dass für jedes $v \in K^n$ die Abbildung $f_v : K^n \to K$ Linksmultiplikation mit der Matrix v^T ist.

Überprüfen wir nun, dass die Abbildung linear ist. Dazu wählen wir $v_1, v_2 \in V$ und $\lambda \in K$. Dann ist

$$f_{v_1+\lambda v_2}(w) = f_{v_1}(w) + \lambda f_{v_2}(w),$$

wie man leicht überprüft und somit ist die Abblildung linear.

Nun zeigen wir, dass die Abbildung bijektiv ist. Da $\dim V = \dim V^* < \infty$ genügt es zu zeigen, dass die Abbildung injektiv ist. Da φ linear ist, müssen wir nur den Kern bestimmen. Nehmen wir also an, dass $v = (v_1, \ldots, v_n)^T \in \mathrm{Kern}(\varphi)$ also gilt für alle $w \in K^n$, dass

$$f_v(w) = 0.$$

Setzen wir als $w = e_i$ den Standardbasisvektor, so erhalten wir

$$0 = f_v(e_i) = v_i$$

und somit ist $v = 0$. □

11.2 Quotientenvektorraum

Gegeben sei ein Vektorraum V über K und ein Unterraum U. Für jedes $v \in V$ betrachten wir die Menge

$$v + U = \{v + u \,:\, u \in U\}.$$

Lemma. *Sei V ein Vektorraum über K und U ein Unterraum und $v_1, v_2 \in V$ zwei Elemente. Dann sind folgende Eigenschaften äquivalent:*

1. $v_1 + U = v_2 + U$
2. $v_1 - v_2 \in U$

Beweis. Angenommen $v_1 + U = v_2 + U$. Dann gibt es $u_1, u_2 \in U$ sodass $v_1 + u_1 = v_2 + u_2$ und somit

$$v_1 - v_2 = u_2 - u_1 \in U.$$

Nehmen wir nun an, dass $v_1 - v_2 \in U$ und betrachten wir $u_1 \in U$. Dann ist

$$v_1 - v_2 + u_1 \in U,$$

da U ein Unterraum ist. Also gibt es $u_2 \in U$, sodass

$$v_1 + u_1 = v_2 + u_2.$$

Also folgt $v_1 + U \subset v_2 + U$. Man zeigt analog, dass $v_2 + U \subset v_1 + U$. □

Definition. Sei V ein Vektorraum über K und U ein Unterraum. Dann definieren wir den *Quotientenvektorraum* als

$$V/U = \{v + U \,:\, v \in V\}.$$

Dieser Menge hat eine natürliche Struktur als Vektorraum, wie in der Vorlesung gezeigt wurde.

Ich möchte zunächst folgende Sachen anmerken:

1. Für die Dimension von V/U gilt

$$\dim(V/U) = \dim(V) - \dim(U).$$

Um dies zu sehen, betrachten wir eine Basis $v_1, \ldots, v_\ell$ von U und erweitern sie zu einer Basis $v_1, \ldots, v_\ell, v_{\ell+1}, \ldots, v_n$ von V. Dann ist

$$v_{\ell+1} + U, \ldots, v_n + U$$

eine Basis von V/U.

2. Falls $f : V \to W$ eine lineare und surjektive Abbildung zwischen Vektorräumen V und W ist, so induziert f einen Isomorphismus

$$\overline{f} : V/\mathrm{Kern}(f) \to W, \qquad v + \mathrm{Kern}(f) \mapsto f(v).$$

Überprüft, dass diese Abbildung wohldefiniert ist. Sie ist auch offensichtlich surjektiv. Nun gilt auch noch, dass

$$\dim V/\mathrm{Kern}(f) = \dim V - \dim \mathrm{Kern}(f) = \dim(\mathrm{Bild}(f)) = \dim(W)$$

und somit ist die Abbildung bijektiv.

Geben wir nun zwei Beispiele an.

1. Betrachten wir $V = \mathbb{R}^2$ als $\mathbb{R}$-Vektorraum und

$$U = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R} \right\}.$$

Bemerke, dass U die Diagonale im Raum $\mathbb{R}^2$ ist. Somit ist $v + U$ eine Verschiebung dieser Diagonale. Damit bestehet der Quotientenraum V/U aus der Diagonalen und allen möglichen Verschiebungen. Somit haben wir diesem interessanten Raum eine Vektorraumstruktur gegeben. Aus der Dimensionsformel, folgt, dass V/U die Dimension 1 hat.

2. Wir geben nun eine explizite Konstruktion von den reellen Zahlen an. Dazu betrachten wir den $\mathbb{Q}$-Vektorraum

$$V = \{(x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{Q} \text{ und die Folge } (x_n)_{n\in\mathbb{N}} \text{ ist eine Cauchyfolge}\},$$

welcher aus allen rationalen Cauchy folgen besteht. Diese Menge bilden mit der komponentenweisen Addition und der komponentenweisen Multiplikation einen Vektorraum über $\mathbb{Q}$. Wir betrachten den Unterraum

$$U = \{(x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{Q} \text{ und } \lim_{n\to\infty} x_n = 0\}$$

der rationalen Nullfolgen. Dann können wir

$$\mathbb{R} = V/U$$

setzten. Man kann sich nun überlegen, wie man alle Eigenschaften von $\mathbb{R}$ anhand dieser Konstruktion überprüft.

12 Determinanten

Wir geben nun eine rekursive Definition der Determinante. Falls $A = (a)$ eine 1×1-Matrix ist, so definieren wir

$$\det(A) = a.$$

Betrachten wir nun eine $n \times n$-Matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ und nehmen an, dass wir die Determinante für quadratische $(n-1)$-Matrizen definiert haben. Bezeichnen wir für jedes (i,j) mit $1 \leq i,j \leq n$ die Matrix A_{ij} die $(n-1) \times (n-1)$-Matrix die erhalten wird indem wir die i-te Zeile und die j-te Spalte streichen. Dann definieren wir

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

Somit ergibt sich

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

und

$$\begin{aligned} \det\left(\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\right) &= a \det\left(\begin{pmatrix} e & f \\ h & i \end{pmatrix}\right) - d \det\left(\begin{pmatrix} b & c \\ h & i \end{pmatrix}\right) + g \det\left(\begin{pmatrix} b & c \\ e & f \end{pmatrix}\right) \\ &= a(ei - hf) - d(bi - ch) + g(bf - ec) \\ &= aei + dhc + gbf - ceg - fha - ibd. \end{aligned}$$

Ich möchte noch auf die folgende geometrische Interpretation der Determinante aufmerksam machen. Es gilt nämlich für eine relle $n \times n$-Matrix A, dass

$$|\det(A)| = \mathrm{vol}(A \cdot [0,1]^n),$$

wobei wir mit $\mathrm{vol}(A \cdot [0,1]^n)$ das euklidische Volumen der Menge $A \cdot [0,1]^n$ meinen. Dies folgt indem wir zunächst diese Aussage für Diagonalmatrizen, Permutationsmatrizen und weitere einfache Fälle beweisen.

Die Determinante hat folgende Eigenschaften:

1. Eine Matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ ist invertierbar genau dann wenn $\det(A) \neq 0$.
2. Für $n \times n$-Matrizen A und B folgt

   $$\det(AB) = \det(A) \det(B).$$

   Aus dieser Formel folgt: Falls A invertierbar ist, so folgt

   $$\det(A) \det\big(A^{-1}\big) = \det\big(AA^{-1}\big) = \det(1_n) = 1,$$

   also folgt

   $$\det\big(A^{-1}\big) = \frac{1}{\det(A)} = \det(A)^{-1}.$$

   Ebenso gilt für $n \times n$-Matrizen A und B mit A invertierbar

   $$\det(B) = \det(A) \det(A)^{-1} \det(B) = \det(A) \det(B) \det\big(A^{-1}\big) = \det\big(ABA^{-1}\big).$$

3. $\det(A) = \det\left(A^T\right)$

4. Für eine obere Dreiecksmatrix

$$A = \begin{pmatrix} \lambda_1 & * & * & \dots & * & * \\ 0 & \lambda_2 & * & \dots & * & * \\ 0 & 0 & \lambda_3 & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_{n-1} & * \\ 0 & 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

gilt

$$\det(A) = \prod_{i=1}^{n} \lambda_i.$$

Nach 3. gilt das Gleiche für untere Dreicksmatrizen.

5. Nach 4. gilt $\det(I_n) = 1$.

6. Sei A eine Matrix. Bezeichnen wir mit A' die Matrix, welche man durch A erhält indem man das Vielfache von einer Zeile von einer anderen Zeile abzieht, so gilt

$$\det(A) = \det(A').$$

7. Sei A eine Matrix. Bezeichnen wir mit A' die Matrix, welche man durch A erhält indem man eine Zeile mit λ multipliziert, so gilt

$$\lambda \det(A) = \det(A').$$

8. Sei A eine Matrix. Bezeichnen wir mit A' die Matrix, welche man durch A erhält indem man zwei Zeilen vertauscht, so gilt

$$\lambda \det(A) = -\det(A').$$

9. Falls wir eine Matrix von folgender Form haben

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

wobei A, B, C selbst Matrizen von passender Grösse sind, so folgt

$$\det\left(\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}\right) = \det(A)\det(D).$$

Als Beispiel möchte ich diskutieren für welche Werte $\alpha \in \mathbb{R}$ die relle Matrix

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 6 & \alpha^2 & 2 \\ 9 & \alpha & 3 \end{pmatrix}$$

invertierbar ist. Dazu berechnen wir die Determinante von A in Abhängigkeit von α. Es ist am einfachsten Gausselemination mit der 3. Spalte durchzuführen. Wir verwenden 8. und erhalten somit

$$\det(A) = -\det\left(\begin{pmatrix} 1 & 1 & 2 \\ 2 & \alpha^2 & 6 \\ 3 & \alpha & 9 \end{pmatrix}\right)$$

Mit Gausselemination und 9. folgt

$$\begin{aligned}\det(A) &= -\det\left(\begin{pmatrix}1 & 1 & 2\\ 0 & \alpha^2-2 & 2\\ 0 & \alpha-3 & 3\end{pmatrix}\right)\\ &= -\det(1)\cdot\det\left(\begin{pmatrix}\alpha^2-2 & 2\\ \alpha-3 & 3\end{pmatrix}\right)\\ &= -(3(\alpha^2-2)-2(\alpha-3))\\ &= -(3\alpha^2-6-2\alpha+6)\\ &= -(3\alpha^2-2\alpha) = \alpha(2-3\alpha).\end{aligned}$$

Somit ist die Matrix A invertierbar, genau dann wenn

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

also falls $\alpha \neq 0, \frac{2}{3}$.

13 Polynome, Eigenwerte und Diagonalisierbarkeit

13.1 Polynome

Definition. Sei K ein Körper. Der *Polynomring* ist definiert als

$$K[X] = \left\{ \sum_{i=0}^{\infty} a_i X^i \,:\, a_i \in K \text{ für alle } i \text{ und fast alle } a_i = 0 \right\}.$$

Ein *Polynom* ist ein Element des Polynomringes.

Ich möchte ein paar Anmerkungen zu der Definition machen.

1. Betrachen wir $K = \mathbb{F}_2 = \{\overline{0}, \overline{1}\}$ und das Polynom

   $$P(X) = X^2 - X.$$

   Dann gilt

   $$P(\overline{0}) = \overline{0} \quad \text{und} \quad P(\overline{1}) = \overline{0}.$$

   Aber das Polynom ist nicht gleich, wie das Nullpolynom.

2. Hat ein Polynom $P(X) \in K[X]$ eine Nullstelle λ, so gibt es ein Polynom $G(X) \in K[X]$ von kleinerem Grad, sodass

   $$P(X) = (X - \lambda)G(X).$$

   Dies impliziert, dass falls ein Polynom $P(X)$ vom Grad n die Nullstellen $\lambda_1, \ldots, \lambda_n$ hat so folgt

   $$P(X) = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n).$$

Eine interessante Klasse von Polynomen sind sogennante *irreduzible Polynome*.

Definition. Sei $F(X) \in K[X]$. Das Polynom $F(X)$ heisst *reduzibel*, falls es sich als Produkt von Polynomen in $K[X]$ vom Grad ≥ 1 schreiben lässt. Ein Polynom $F(X) \in K[X]$ heisst *irreduzibel*, falls es nicht reduzibel ist.

Erinnern wir uns an die Definition einer Primzahl: Eine Zahl heisst Primzahl, falls die einzigen Teiler 1 und die Zahl selbst sind. Äquivalenterweise ist eine Zahl eine Primzahl, wenn sie nicht als Produkt von echt kleineren Zahlen geschrieben werden kann. Mit dieser Umformulierung der Definition, wird klar, dass wir irreduzible Polynome wie Primzahlen im Polynomring $K[X]$ verstehen können.

Die vielleicht wichtigste Eigenschaft von Primzahlen ist, dass wir jede Zahl in ihre Primfaktoren zerlegen können. Obige Analogie zwischen Primzahlen und irreduziblen Polynom legt die Frage nahe, ob dies auch für Polynom gilt. Die affirmative Antwort auf diese Frage ergibt nächste Proposition.

Proposition. *Jedes Polynom in $K[X]$ ist das Produkt von irreduziblen Polynomen.*

Beweis. Sei $F(X) \in K[X]$ ein Polynom. Wir führen einen Induktionsbeweis über den Grad von $F(X)$. Da jedes Polynom vom Grad 0 und 1 irreduzibel ist, gilt die Induktionsannahme. Für den Induktionsschritt, machen wir eine Fallunterscheidung.

Nehmen wir zunächst an, dass $F(X)$ irreduizbel ist. Dann ist $F(X)$ offensichtlicherweise das Produkt von irreduziblen Polynomen.

Nehmen wir nun an, dass $F(X)$ reduizbel ist. Nach Definition gibt es somit Polynome $G(X), H(X) \in K[X]$ von Grad ≥ 1, sodass

$$F(X) = G(X) \cdot H(X).$$

Dann ist der Grad von $G(X)$ und $H(X)$ kleiner gleich dem Grad von F und somit können wir die Induktionsannahme auf $G(X)$ und $H(X)$ anwenden. Also ist $G(X)$ und $H(X)$ das Produkt von irreduziblen Polynomen und somit auch $F(X)$. □

Ich gebe nun ein paar Beispiele für Faktorisierungen von Polynomen in ihre irreduzible Faktoren. Dabei ist es relevant, über welchem Körper man das Polynom betrachtet.

1. Betrachten wir das Polynom

   $$P(X) = X^2 + 1 \in \mathbb{R}[X].$$

   Dieses Polynom ist irreduzibel über $\mathbb{R}$, da es vom Grad 2 ist und keine Nullstellen in $\mathbb{R}$ hat. Aber über den komplexen Zahlen $\mathbb{C}$ haben wir

   $$P(X) = (X - i)(X + i).$$

2. Um die irreduziblen Faktoren eines Polynoms über $\mathbb{Q}$ oder $\mathbb{R}$ zu bestimmen ist es im Allgemeinen eine gute Idee dieses Polynom über $\mathbb{C}$ zu betrachten und dann die komplexen Nullstellen zu berechnen. Als Beispiel betrachten wir nun

   $$P(X) = X^4 - X^2 - 2 \in \mathbb{Q}[X].$$

Wir sehen, dass dieses Polynom über $\mathbb{C}$ die Nullstellen

$$\{\sqrt{2}, -\sqrt{2}, i, -i\}$$

hat. Also ist die Faktorisierung in irreduzible Faktoren des Polynoms über $\mathbb{C}$:

$$P(X) = (X - \sqrt{2})(X + \sqrt{2})(X - i)(X + i) \in \mathbb{C}[X].$$

Über den reellen Zahlen $\mathbb{R}$ ergibt sich somit die Faktorisierung

$$P(X) = (X - \sqrt{2})(X + \sqrt{2})(X^2 + 1) \in \mathbb{R}[X]$$

und schliesslich über den rationalen Zahlen

$$P(X) = (X^2 - 2)(X^2 + 1) \in \mathbb{Q}[X].$$

3. Über $\mathbb{C}$ sind die einzigen irreduziblen Polynome vom Grad 1. Dies folgt, da jedes Polynom über $\mathbb{C}$ eine Nullstelle hat. Also zerfällt jedes Polynom über $\mathbb{C}$ in Linearfaktoren.

13.2 Charakteristisches Polynom, Eigenwerte und Eigenvektoren

Definition. Es sei V ein Vektorraum über K und $f : V \to V$ ein Endomorphismus. Dann heisst $\lambda \in K$ ein *Eigenwert*, falls es $v \in V \backslash \{0\}$ gibt, sodass

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

Der Vektor $v \in V$ heisst dann *Eigenvektor zum Eigenwert* λ.

Sei $\lambda \in K$ ein Eigenwert von f. Dann definieren wir den Eigenraum

$$\mathrm{Eig}_{\lambda,f} = \{v \in V \,:\, f(v) = \lambda v\} = \{v \in V \,:\, (f - \lambda \cdot \mathrm{id}_V)v = 0\} = \mathrm{Kern}(f - \lambda \cdot \mathrm{Id}_V).$$

Diskutieren wir nun ein paar Beispiele:

1. Betrachten wir $\mathrm{id} : V \to V$. Dann ist 1 der einzige Eigenwert und der Eigenraum von 1 ist V.

2. Der Kern einer linearen Abbildung $f : V \to V$ ist der Eigenraum zum Eigenwert $0 \in K$. Des Weiteren ist $0 \in K$ ein Eigenwert von f genau dann, wenn der Kern von f nicht trivial ist.

3. Betrachten wir die lineare Abbildung

$$D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad f \mapsto f'.$$

Wir beschreiben nun die Eigenräume von D. Sei dazu $\lambda \in \mathbb{R}$. Eine Funktion $f \in C^\infty(\mathbb{R})$ ist demnach ein Eigenvektor von D, genau dann, wenn

$$f' = \lambda f.$$

Aus der Analysis ist bekannt, dass die einzige Lösung dieser Gleichung die Möglichkeit

$$f(x) = c \cdot e^{\lambda x}$$

ist für $c \in \mathbb{R}$ eine Konstante. Somit haben wir also

$$\mathrm{Eig}_{\lambda,D} = \{c \cdot e^{\lambda x} \,:\, c \in \mathbb{R}\}.$$

Im folgenden betrachten wir eine $n \times n$-Matrix A über K. Ein Eigenwert ist also ein Skalar $\lambda \in K$ sodass $v \in K^n \backslash \{0\}$ existiert mit

$$Av = \lambda v.$$

Definition. Sei A eine $n \times n$-Matrix über K. Dann definieren wir das charakteristische Polynom von A als

$$\mathrm{char}_A(X) := \det(X \cdot I_n - A) \in K[X].$$

Beispiel. Betrachten wir die 2×2-Matrix

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

mit $a, b, c, d \in K$. Dann ist

$$\begin{aligned} \mathrm{char}_A(X) = \det(X \cdot I_2 - A) &= \det\left(\begin{pmatrix} X-a & -b \\ -c & X-d \end{pmatrix}\right) \\ &= (X-a)(X-d) - cb = X^2 - (a+d)X + ad - bc. \end{aligned}$$

Sei nun $\lambda \in K$ ein Eigenwert. Dann ist also

$$\mathrm{Eig}_{\lambda,A} = \mathrm{Kern}(\lambda \cdot I_n - A)$$

der *Eigenraum* von λ. Die Dimension des Eigenraumes von λ wird die *geometrische Vielfachheit* von λ genannt.

Lemma. *Sei $\lambda \in K$ ein Eigenwert von A. Dann ist der Eigenraum ein Unterraum von K^n.*

Beweis. Es ist klar, dass $0 \in \mathrm{Eig}_{\lambda,A}$. Falls $v_1, v_2 \in \mathrm{Eig}_{\lambda,A}$, dann folgt

$$\begin{aligned} (\lambda \cdot I_n - A)(v_1 + v_2) &= \lambda \cdot v_1 + \lambda \cdot v_2 - Av_1 - Av_2 \\ &= \lambda \cdot v_1 - Av_1 + \lambda \cdot v_2 - Av_2 \\ &= (\lambda \cdot I_n - A)(v_1) + (\lambda \cdot I_n - A)(v_2) = 0 + 0 = 0. \end{aligned}$$

Analog folgt, dass falls $v \in \mathrm{Eig}_{\lambda,A}$ und $\lambda \in K$, dass dann auch $\lambda v \in \mathrm{Eig}_{\lambda,A}$. □

Proposition. *Sei A eine $n \times n$-Matrix über K. Dann sind die Eigenwerte von A genau die Nullstellen des charakteristischen Polynoms.*

Beweis. Nehmen wir an, dass $\lambda \in K$ ein Eigenwert von A ist. Dann gibt es $v \neq 0$, sodass

$$Av = \lambda v$$

oder äquivalenterweise

$$\lambda v - Av = (\lambda \cdot I_n - A)v = 0.$$

Somit ist $v \in \mathrm{Eig}_{\lambda,A}$ und somit

$$\mathrm{Kern}(\lambda \cdot I_n - A) \neq \{0\}.$$

Also ist die Matrix $\lambda \cdot I_n - A$ nicht invertierbar. Dies impliziert

$$\mathrm{char}_A(\lambda) = \det(\lambda \cdot I_n - A) = 0.$$

Die Umkehrung folgt analog. □

Diese Proposition ist nützlich, da sie uns ein Verfahren gibt, die Eigenwerte einer Matrix zu bestimmen. Des Weiteren folgen die nächsten beiden Korollare.

Korollar. *Zwei ähnliche Matrizen haben die gleichen Eigenwerte.*

Beweis. Zeigen wir zunächst, dass zwei ähnliche Matrizen das gleiche charakteristische Polynom. Betrachten wir dazu $n \times n$-Matrizen A, B, C mit C invertierbar und sodass $A = CBC^{-1}$. Dann gilt

$$\begin{aligned}\operatorname{char}_B(X) &= \det(X \cdot I_n - B) \\ &= \det\big(C^{-1}(X \cdot I_n - B)C\big) \\ &= \det\big(X \cdot I_n - C^{-1}BC\big) \\ &= \det(X \cdot I_n - A) = \operatorname{char}_A(X).\end{aligned}$$

Also haben zwei ähnliche Matrizen haben das gleiche charakteristische Polynom und somit nach der letzten Proposition die gleichen Eigenwerte. □

Korollar. *Sei A eine $n \times n$-Matrix über $\mathbb{C}$. Dann hat A einen Eigenwert.*

Beweis. Das charakteristische Polynom von A hat eine Nullstelle in $\mathbb{C}$, da $\mathbb{C}$ algebraisch abgeschlossen ist. Nach der letzten Proposition ist somit diese Nullstelle ein Eigenwert. □

Als konkretes Beispiel betrachten wir die Matrix

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

Um die Eigenwerte von A zu bestimmen, betrachten wir das characteristische Polynom

$$\operatorname{char}_A(X) = \det\left(\begin{pmatrix} X-1 & -1 \\ 0 & X-1 \end{pmatrix}\right) = (X-1)^2.$$

Also ist 1 der einzige Eigenwert.

Definition. Sei $\lambda \in K$ ein Eigenwert der Matrix A. Dann ist nach obigem Satz das charakteristische Polynom von der Form

$$\operatorname{char}_A(X) = (X - \lambda)^{n_\lambda} P(X)$$

mit $n_\lambda \in \mathbb{N}$ und $P(X) \in K[X]$ ein Polynom vom Grad kleiner als $\operatorname{char}_A(X)$. Die natürliche Zahl n_λ nennen wir die *arithmetische Vielfachheit von* λ.

Wir führen nun obiges Beispiel mit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ fort. Wir haben gesehen, dass 1 der einzige Eigenwert ist. Die arithemtische Vielfachheit des Eigenwerts 1 ist 2, da $\operatorname{char}_A(X) = (X-1)^2$. Wir berechnen nun die geometrische Vielfachheit. Dazu berechnen wir den Eigenraum von 1:

$$\begin{aligned}\operatorname{Eig}_{1,A} &= \big\{v \in \mathbb{R}^2 \,:\, Av = v\big\} \\ &= \big\{v \in \mathbb{R}^2 \,:\, (A - I_2)v = 0\big\} \\ &= \left\{v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2 \,:\, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\} \\ &= \left\{\begin{pmatrix} x \\ 0 \end{pmatrix} \,:\, x \in \mathbb{R}\right\} = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle.\end{aligned}$$

Also ist 1 die geometrische Vielfachheit des Eigenwertes 1.

Es ist Anzumerken, dass aus der Definition direkt folgt, dass die geometrische und die arithmetische Vielfachheit beide mindestens 1 sind. Die nächste Proposition bringt die geometrische und die arithmetische Vielfachheit in Verbindung.

Proposition. *Sei $\lambda \in K$ ein Eigenwert der Matrix A. Dann ist die geometrische Vielfachheit von λ kleiner gleich der arithmetischen Vielfachheit.*

Beweis. Sei $\dim(\mathrm{Eig}_{\lambda,A}) = \ell \leq n$. Dann betrachten wir eine Basis $(b_1, \ldots, b_\ell)$ des Eigenraumes $\mathrm{Eig}_{\lambda,A}$. Diese Basis erweitern wir zu einer geordneten Basis $B = (b_1, \ldots, b_n)$ von K^n. Wir stellen fest, dass die Darstellungsmatrix von A bezüglich der Basis B folgende Form hat

$$\begin{pmatrix} \lambda & 0 & 0 & \ldots & 0 & * \\ 0 & \lambda & 0 & \ldots & 0 & * \\ 0 & 0 & \lambda & \ldots & 0 & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda & * \\ 0 & 0 & 0 & \ldots & 0 & C \end{pmatrix},$$

wobei C eine $(n-\ell)\times(n-\ell)$-Matrix ist. Da das charakteristische Polynom von A das gleiche ist, wie das charakteristische Polynom von dieser Darstellungsmatrix, folgt

$$\begin{aligned} \mathrm{char}_{\lambda,A}(X) &= \det\left(\begin{pmatrix} X-\lambda & 0 & 0 & \ldots & 0 & * \\ 0 & X-\lambda & 0 & \ldots & 0 & * \\ 0 & 0 & X-\lambda & \ldots & 0 & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & X-\lambda & * \\ 0 & 0 & 0 & \ldots & 0 & X\cdot I_{n-\ell} - C \end{pmatrix}\right) \\ &= (X-\lambda)^\ell \cdot \mathrm{char}_C(X). \end{aligned}$$

Also ist die algebraische Vielfachheit grösser oder gleich wie die geometrische. □

13.3 Diagonalisierbarkiet

Definition. Eine $n \times n$-Matrix A heisst *diagonalisierbar*, falls es eine Basis von K^n gibt, welche aus Eigenvektoren von A besteht.

Wir zeigen in dem nächsten Lemma, dass eine Matrix genau dann diagonalisierbar ist, wenn sie ähnlich zu einer Diagonalmatrix ist. Dies reflektiert die Tatsache, dass zwei Matrizen genau dann ähnlich sind, wenn sie bis auf einen Basiswechsel die gleichen Matrizen sind.

Lemma. *Eine $n \times n$-Matrix A ist diagonalisierbar, genau dann, wenn es eine Diagonalmatrix D und eine invertierbare Matrix U gibt, sodass*

$$D = U^{-1}AU.$$

Beweis. Nehmen wir an, dass A diagonalisierbar ist, also dass $B = (b_1, \ldots, b_n)$ eine Basis aus Eigenvektoren ist. Es sei λ_i der Eigenwert von b_i für $i \in \{1, \ldots, n\}$

und wir bezeichnen mit D die Diagonalmatrix, welche aus $\lambda_1, \ldots, \lambda_n$ besteht, also $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Betrachten wir die Matrix

$$U = \begin{pmatrix} | & & | \\ b_1 & \ldots & b_n \\ | & & | \end{pmatrix}.$$

Wir behaupten

$$D = U^{-1}AU.$$

Um dies zu beweisen, bezeichnen wir wie üblich mit e_i den Standardbasisvektor. Dann ist

$$U^{-1}AUe_i = U^{-1}(A(Ue_i)) = U^{-1}(Ab_i)) = U^{-1}(\lambda_i b_i) = \lambda_i e_i.$$

Also gilt $D = U^{-1}AU$ und somit ist die erste Richtung gezeigt.

Nehmen nun an, dass solche Matrizen U und $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ existieren. Bezeichen wir mit $(b_1, \ldots, b_n)$ die Spalten von U, also ist

$$U = \begin{pmatrix} | & & | \\ b_1 & \ldots & b_n \\ | & & | \end{pmatrix}.$$

Dann ist $B = (b_1, \ldots, b_n)$ eine Basis von K^n, da U invertierbar ist. Wir behaupten, dass b_i ein Eigenvektor zum Eigenwert λ_i ist. Um dies zu sehen, erinnern wir uns, dass nach Annahme $D = U^{-1}AU$ gilt. Also gilt

$$UDU^{-1} = A.$$

Werten wir dies nun an b_i aus, so erhalten wir

$$Ab_i = U(D(U^{-1}b_i)) = U(De_i) = U(\lambda e_i) = \lambda Ue_i = \lambda b_i.$$

Somit ist die Behauptung gezeigt. □

Der nächste Satz gibt uns eine nützliche Charakterisierung von Diagonalisierbarkeit.

Satz. *Eine $n \times n$-Matrix A ist diagonalisierbar, genau dann wenn das charakteristische Polynom von A über K in Linearfaktoren zerfällt und für jeden Eigenwert $\lambda \in K$ von A die geometrische Vielfachheit von λ gleich wie die arithmetische Vielfachheit von λ ist.*

Beweis. Angenommen A ist diagonalisierbar. Dann gibt es eine geordnete Basis $B = (b_1, \ldots, b_n)$ von K^n, sodass

$$Ab_i = \lambda_i b_i.$$

für alle $i \in \{1, \ldots, n\}$. Nehmen wir ohne Beschränkung der Allgemeinheit an, dass $\lambda_{i+1} \leq \lambda_i$ für alle $i \in \{1, \ldots, n-1\}$. Somit ist die Darstellungsmatrix von A bezüglich dieser Basis die Diagonalmatrix

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

Da das charakteristische Polynom von A das Gleiche ist, wie das charakteristische Polynom von dieser Dastellungsmatrix, folgt, dass das charakteristische Polynom in Linearfaktoren zerfällt. Ebenso folgt, dass für jeden Eigenwert die geometrische Vielfachheit gleich wie die arithmetische ist.

Nehmen wir nun an, dass A diese Bedingungen erfüllt. Seien $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von A. Nach Annahme ist das charakteristische Polynom von der Form

$$\text{char}_A(X) = (X - \lambda_1)^{n_1} \cdot \ldots \cdot (X - \lambda_k)^{n_k}$$

für $n_1, \ldots, n_k \in \mathbb{N}$. Es folgt des Weiteren, dass $n_1 + \ldots + n_k = n$, da n der Grad von $\text{char}_A(X)$ ist. Da die geometrische Vielfachheit gleich ist die wie die arithmetische, folgt, dass

$$\dim(\text{Eig}_{A,\lambda_i}) = n_i$$

für alle $i \in \{1, \ldots, k\}$. Wählen wir nun für jeden Eigenwert λ_i eine Basis $b_1^i, \ldots, b_{n_i}^i$ von Eig_{A,λ_i}. Wir behaupten, dass die Menge

$$B = \{b_1^1, \ldots, b_{n_1}^1, b_1^2, \ldots, b_{n_2}^2, \ldots, b_1^k, \ldots, b_{n_k}^k\}$$

eine Basis von K^n ist. Ist dies der Fall, so bildet B eine Basis von Eigenvektoren von A. Die Menge B ist linear unabhängig, da der Schnitt der Eigenräume der triviale Vektorraum ist. Da die Kardinalität von B gleich n ist, folgt, dass B eine Basis ist. □

Wollen wir nun überprüfen, dass eine gegebene Matrix A diagonalisierbar ist, empfiehlt dieser Satz folgendes Verfahren:

1. Finde das charakteristische Polynom. Damit A diagonalisierbar ist, muss das charakteristische Polynom in Linearfaktoren zerfallen. Ist dies nicht der Fall, so ist A nicht diagonalisierbar.

2. Berechne die Eigenwerte von A (also die Nullstellen des charakteristischen Polynoms).

3. Berechne die arithmetische Vielfachheit der Eigenwerte von A.

4. Falls die arithmetische Vielfachheit aller Eigenwerte 1 ist, so ist A diagonalisierbar, da die geometrische Vielfachheit sets kleiner oder gleich wie die arithmetische Vielfachheit ist, aber zumindest 1 ist.

5. Für alle Eigenwerte mit arithmetischer Vielfachheit grösser als 1, berechne den Eigenraum und überprüfe ob dessen Dimension gleich ist wie die arithmetische Vielfachheit. Ist dies der Fall, so ist A diagonalisierbar. Falls nicht, so ist A nicht diagonalisierbar.

Betrachten wir nun ein paar Beispiele:

1. Die Matrix

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

hat wie in 13.2 berechnet das charakteristische Polynom $(X-1)^2$ aber der Eigenwert 1 hat geometrische Vielfachheit 1. Somit ist A **nicht** diagonalisierbar.

2. Die Matrix
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
hat charakteristisches Polynom
$$\mathrm{char}_A(X) = \det\left(\begin{pmatrix} X & -1 \\ 1 & X \end{pmatrix}\right) = X^2 + 1.$$
Somit ist die Matrix A über $\mathbb{R}$ **nicht** diagonalisierbar. Über den komplexen Zahlen gilt aber
$$X^2 + 1 = (X - i)(X + i),$$
somit ist A über $\mathbb{C}$ diagonalisierbar.