Lineare Algebra und Analytische Geometrie (LAAG)

Dr Julia Goedecke
julia.goedecke@tu-dresden.de

SS 24, letzter Update 8. Juli 2024

Inhaltsverzeichnis

Einleitung

Warum sollen Sie Lineare Algebra hören?

- Sie studieren Mathematik, nicht nur Schulmathematik, und Lineare Algebra gehört zur Grundausbildung.
- Sie sollen mathematisches Denken lernen (logisches, abstraktes Denken, Problemlösen, Zusammenhänge finden, usw.). Das lernt man nur an Beispielen, und dies ist eins.
- Teile des Stoffs sind relevant für Schulstoff (je nach Schulart unterschiedlich)

Übersicht

In der linearen Algebra geht es um Vektoren, Matrizen und Skalare (die „Zutaten"), und deren Strukturen und Eigenschaften (Zusammenhänge).
Etwas mehr über die Zutaten: Es gibt

- **Vektoren**, z.B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ oder $\begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}$ oder $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ oder v, ...
- **Skalare**, also reelle Zahlen, z.B. 0.5, $\sqrt{2}$, λ, ...
- **Matrizen**, z.B. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ oder $\begin{pmatrix} 3 & 5 & 2 \\ 6 & 1 & 2 \end{pmatrix}$, ...
- und noch ein paar abstraktere Konzepte.

Einige der wichtigsten Sachen, die wir mit diesen Zutaten machen, sind

- **Linearkombinationen**:
  - **Vektoraddition** z.B. $\begin{pmatrix} 3 \\ 5 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$
  - **skalare Multiplikation** z.B. $3 \cdot \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ 15 \end{pmatrix}$

  In Verbindung geben diese beiden Operationen eine Linearkombination:

$$3 \cdot \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + 5 \cdot \begin{pmatrix} -1 \\ 2 \\ 6 \end{pmatrix} - \sqrt{2} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

  Linearkombinationen sind einer der roten Fäden im Kurs und werden eine wichtige Rolle spielen.
- Matrizen agieren auf Vektoren durch **Matrixmultiplikation**, z.B.

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+2 \\ 0+1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

  Wir werden lernen, wie das geht.

Wofür stehen Vektoren?

- Ein Pfeil/eine Familie von Pfeilen in der Ebene oder im Raum.
- Ein Punkt (in der Ebene oder im Raum, oder in höherdimensionalen „Räumen“).
- Experimentaldaten: n verschiedene Messungen.
- Logistik: z.B. die Position von Lastwagen in verschiedenen Fabrikhallen.
- Bilder: „hugh, saturation and brightness“ und Position des Pixels.
- Viele mehr.

Wofür stehen Matrizen?

- Eine Abbildung, die einen Vektor auf einen anderen Vektor schickt.
- Ein lineares Gleichungssystem.
- Daten (siehe oben)
- ...

Was heißt „linear“?
Intuitiv hat es etwas mit „Linien“ zu tun, also Geraden, im Gegensatz zum Beispiel zu quadratischen Parabeln usw. Genauer heißt linear, dass wir Vektoren oder Variablen addieren können und mit einem Skalar multiplizieren, aber wir können sie nicht quadrieren, die Wurzel ziehen, sin, cos, log, Exponentialfunktion anwenden, ...
Wir werden sehen, dass selbst diese sehr stark scheinende Einschränkung noch ein reiches Teilgebiet der Mathematik liefert.

Motivierende Zusammenhänge

Lineare Gleichungssysteme (LGS) sind auch ein roter Faden in der linearen Algebra. Wir können drei Schreibweisen betrachten:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

oder

$$Ax = b$$

Zeilendarstellung

- übliche Sicht in der Schule
- geometrische Sicht: Schnittpunkte von Geraden, Ebenen, ...

Spaltendarstellung

gibt Kontext für

- Vektorsicht
- Linearkombination und Spann
- lineare Unabhängigkeit
- Basis

Matrixdarstellung

gibt Kontext für

- Sicht der linearen Abbildungen
- Kern und Bild
- Determinante
- Invertierbarkeit

und führt weiter zu

- Eigenvektor, Eigenwert

(die wiederum das Lösen von LGS brauchen.)

Lösungwege für LGS benutzen Zeilenumformungen (und versteckte Matrixmultiplikation). Die selben Algorithmen helfen beim Berechnen von Determinanten und inversen Matrizen.

Die **Struktur der Lösungsmenge**, also die Frage „Wie viele Lösungen gibt es, wie hängen sie zusammen“, führt zu

- Unterräumen,
- linearer Unabhängigkeit und Spann,
- Invertierbarkeit,
- Basis und Dimension.

Anwendungen

Warum ist Lineare Algebra in der Welt nützlich und wichtig? Warum gehört es zur Grundausbildung der Mathematik?
Es wird in vielen vielen Bereichen der Mathematik, Naturwissenschaften, Ingenieurwissenschaften und darüber hinaus benutzt.

- ⋄ Funktionen (mit Hilfe von Tangenten) durch Geraden zu approximieren gibt eine lineare Situation. Dies ist hilfreicher als man vielleicht erst denkt.
- ⋄ Benutzt zum Lösen von Differenzialgleichungen.
- ⋄ Kann in Wahrscheinlichkeitstheorie und Statistik benutzt werden.
- ⋄ Benutzt in Machine Learning/ künstliche Intelligenz.
- ⋄ Benutzt in verschiedenen Methoden in der Numerik.
- ⋄ Benutzt in analytischer Geometrie.
- ⋄ Viele viele mehr...

KAPITEL 1

Vektoren und Matrizen

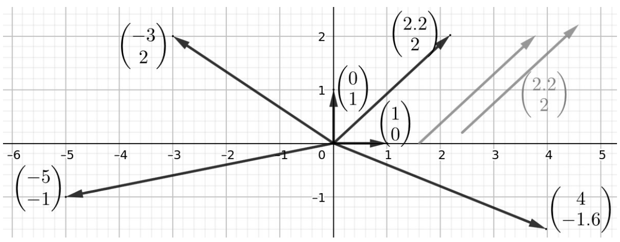
A. Vektoren

Sind Vektoren Pfeile? Familien von Pfeilen? Punkte? Daten?
Um die vielen verschiedenen Interpretationen zu vereinen (und weitere Beispiele darüber hinaus zu ermöglichen), arbeiten wir mit einer algebraischen Darstellung von Vektoren.

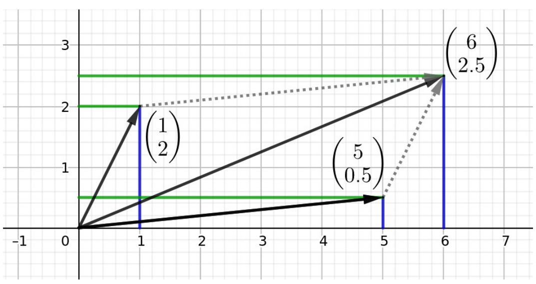
Definition 1.1: Ein **Spaltenvektor** (kurz Vektor) ist eine Spalte aus n Einträgen, die (bei uns) aus reellen Zahlen bestehen.
Die Menge aller Spaltenvektoren mit n reellen Einträgen wird als $\mathbb{R}^n$ bezeichnet.

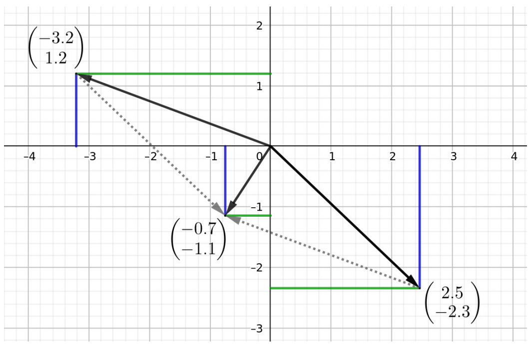
Zum Beispiel $\begin{pmatrix} \sqrt{2} \\ \frac{1}{2} \end{pmatrix}$, $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, $\begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} \in \mathbb{R}^3$, $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Aber auch $5 \in \mathbb{R}^1 = \mathbb{R}$.

Für $n = 2$ oder $n = 3$ können wir uns Vektoren als Punkte in der Ebene oder im Raum vorstellen, oder als Vektor vom Ursprung zu diesem Punkt („Ortsvektor"), oder als Familie von Vektoren, die diese Differenz zwischen Anfangs- und Endpunkt gemeinsam haben.



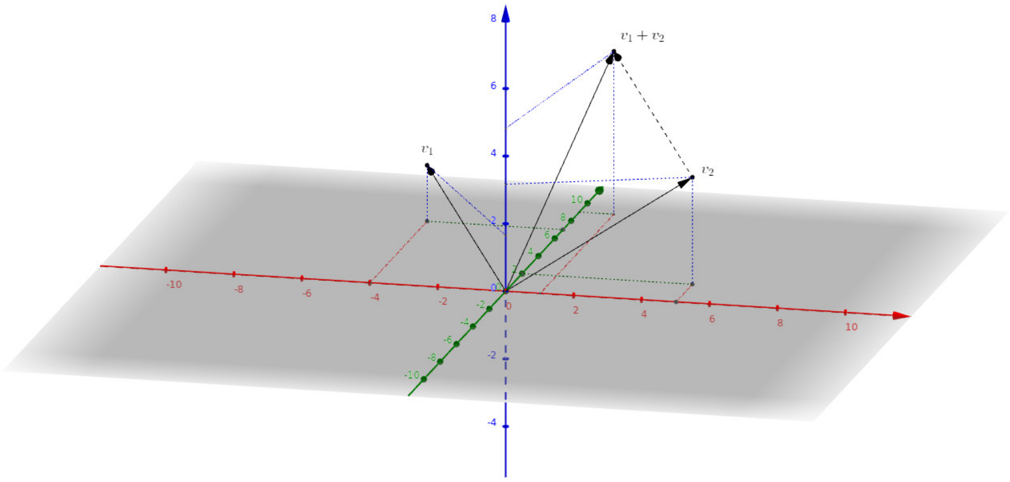
Wir können zwei Vektoren mit derselben Anzahl von Einträgen addieren. Geometrisch addieren wir Vektoren (als Familie von Pfeilen vorgestellt), in dem wir einen Pfeil an das Ende des anderen setzen und den Ursprung mit diesem neuen Endpunkt verbinden.
Algebraisch entspricht dies der Addition in jeder Komponente, wie wir in diesem Beispiel in $\mathbb{R}^2$ sehen.





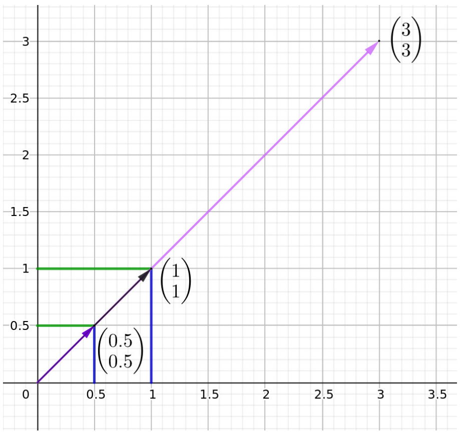
$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 5 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 1+5 \\ 2+0.5 \end{pmatrix} = \begin{pmatrix} 6 \\ 2.5 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -3.2 \\ 1.2 \end{pmatrix} + \begin{pmatrix} 2.5 \\ -2.3 \end{pmatrix} = \begin{pmatrix} -3.2+2.5 \\ 1.2-2.3 \end{pmatrix} = \begin{pmatrix} -0.7 \\ -1.1 \end{pmatrix}$$

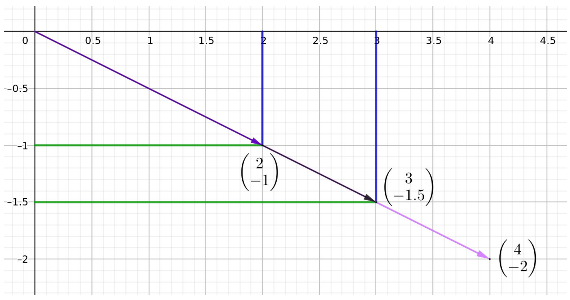
Oder in $\mathbb{R}^3$:



Es ist schwieriger im Raum zu zeichnen. Sie können sich dieses Bild in GeoGebra anschauen und drehen und damit herumspielen, um es besser zu verstehen.
Die algebraische Darstellung der Addition macht auch klar, dass die Reihenfolge keine Rolle spielt, zum Beispiel $\begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}5\\0.5\end{pmatrix} = \begin{pmatrix}5\\0.5\end{pmatrix} + \begin{pmatrix}1\\2\end{pmatrix}$.

Wir können einen Vektor auch mit einer reellen Zahl skalieren: geometrisch behalten wir die Richtung bei und verändern nur die Länge des Vektors.





$$\frac{1}{2} \cdot \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0.5\\0.5\end{pmatrix} \quad \text{und} \quad 3 \cdot \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix}$$

$$\frac{2}{3} \cdot \begin{pmatrix}3\\-1.5\end{pmatrix} = \begin{pmatrix}2\\-1\end{pmatrix} \quad \text{und} \quad \frac{4}{3} \cdot \begin{pmatrix}3\\-1.5\end{pmatrix} = \begin{pmatrix}4\\-2\end{pmatrix}$$

Auch dies funktioniert algebraisch.

Definition 1.2: Die **Vektoraddition** und die **skalare Multiplikation** in $\mathbb{R}^n$ sind komponentenweise definiert:

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_{n-1}\\x_n\end{pmatrix} + \begin{pmatrix}y_1\\y_2\\\vdots\\y_{n-1}\\y_n\end{pmatrix} = \begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_{n-1}+y_{n-1}\\x_n+y_n\end{pmatrix} \qquad \text{und} \qquad \lambda \cdot \begin{pmatrix}x_1\\x_2\\\vdots\\x_{n-1}\\x_n\end{pmatrix} = \begin{pmatrix}\lambda x_1\\\lambda x_2\\\vdots\\\lambda x_{n-1}\\\lambda x_n\end{pmatrix}$$

Wir addieren separat pro Eintrag, und multiplizieren jeden Eintrag des Vektors mit dem Skalar.

Übung 1.3: Berechnen Sie: $\diamond\ 0\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ $\quad\diamond\ \begin{pmatrix}3\\1\\2\\4\\-2\end{pmatrix}+\begin{pmatrix}2\\1\\4\\-2\\3\end{pmatrix}$ $\quad\diamond\ 3\cdot\begin{pmatrix}\sqrt{3}\\ \pi\end{pmatrix}+6\cdot\begin{pmatrix}-\frac{\sqrt{3}}{2}\\ \pi\end{pmatrix}$

Wir haben also für Vektoren mit zwei oder drei Einträgen sowohl eine geometrische als auch eine algebraische Methode für die Vektoraddition und skalare Multiplikation. Für Vektoren mit mehr Einträgen kann man sich die geometrische Version nur noch sehr schwer vorstellen (oder gar nicht). Deshalb heißt es auch lineare *Algebra*, da oft die algebraische Version besser zu handhaben ist. Dennoch hilft uns die geometrische Visualisierung in der Ebene und im Raum auch im übertragenen Sinne für die Situationen mit mehr Einträgen.

Notation 1.4: Bisher haben wir Vektoren immer ausgeschrieben, zum Beispiel als $\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$. Oft ist es nützlich, einen Vektor nur mit einer Variable zu beschreiben, zum Beispiel $u, v \in \mathbb{R}^n$. Wir können dann bei Bedarf immer noch die einzelnen Komponenten betrachten.
Den Vektor, bei dem alle Einträge Null sind, nennen wir auch **Nullvektor** und schreiben ihn als $0 \in \mathbb{R}^n$. Hier muss aus dem Kontext klar sein, wie viele Einträge der Vektor hat.
Es gibt bestimmte Konventionen über den Gebrauch von Buchstaben in der Mathematik. Wir benutzen zum Beispiel oft u, v, w als Vektoren, aber auch manchmal x, y. Diese könnten dann die Komponenten $x_1, x_2, \ldots$ haben. Wir benutzen oft λ, μ, ν als Skalare, aber auch r, s, t sind hierfür üblich.
Oft hängt es auch am Kontext, welcher Buchstabe für ein bestimmtes Objekt am geeignetsten erscheint. Das wichtigste ist, alle Variablen richtig zu definieren, damit wir nicht raten müssen, welches mathematisches Objekt sie repräsentieren.

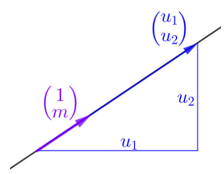
B. Geraden und Ebenen

Aus der Schule kennen Sie (wahrscheinlich) zwei Schreibweisen für Geraden:

$$y = mx + b$$

Hier ist m die Steigung der Gerade, und b der Schnittpunkt mit der y-Achse.
Oder mit Vektoren:

$$\begin{pmatrix}x\\y\end{pmatrix} = t\cdot\begin{pmatrix}u_1\\u_2\end{pmatrix} + \begin{pmatrix}b_1\\b_2\end{pmatrix}$$



Hier ist $\frac{u_2}{u_1} = m$ die Steigung, und die Gerade geht durch den Punkt $\begin{pmatrix}b_1\\b_2\end{pmatrix}$.
Wir nennen t einen **Parameter**, der alle reellen Zahlen durchläuft. Man kann sich das so vorstellen, als ob sich mit t die Zeit ändert, und der Punkt sozusagen auf der Gerade läuft.

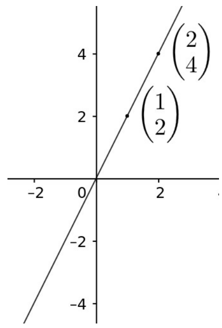
Übung 1.5: Wie würden Sie die Gerade $\left\{ t\begin{pmatrix}0\\1\end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}$ in der Form $y = mx + b$ schreiben?

Von diesen beiden Formen eignet sich die zweite gut zur Verallgemeinerung in höherdimensionale Räume.

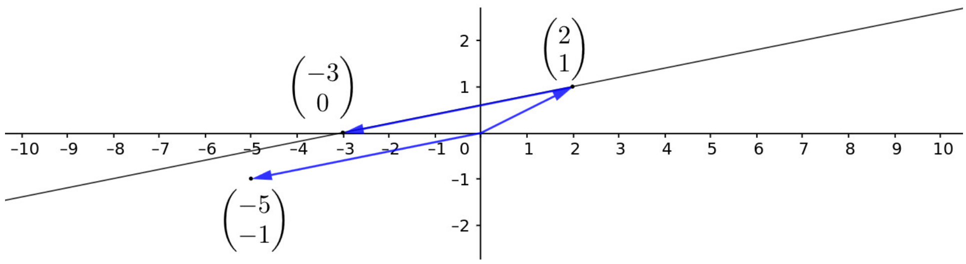
Definition 1.6: Eine **Gerade** in $\mathbb{R}^n$ ist eine Menge der Form $\{b + t\cdot u \mid t \in \mathbb{R}\}$ mit $b, u \in \mathbb{R}^n$ und $u \neq 0$.

Hier ist $b \in \mathbb{R}^n$ ein Punkt auf der Geraden, und u gibt uns die Richtung der Geraden. Wir können b den **Stützvektor** nennen, und u den **Richtungsvektor** oder **Spannvektor**. Die Bedingung $u \neq 0$ garantiert uns, dass die Menge wirklich eine Gerade ist und nicht nur ein Punkt.
Diese Beschreibung einer Gerade nennt man auch **Parameterform**.
Am besten vorstellen können wir uns Geraden in $\mathbb{R}^2$ und $\mathbb{R}^3$, also in der Ebene und im Raum.

Beispiele 1.7:



- $\left\{ \begin{pmatrix}0\\0\end{pmatrix} + t\begin{pmatrix}1\\2\end{pmatrix} \middle| t \in \mathbb{R} \right\}$ ist die Gerade in $\mathbb{R}^2$ durch den Ursprung und den Punkt $\begin{pmatrix}1\\2\end{pmatrix}$. Wir schreiben dann oft einfach $\left\{ t\begin{pmatrix}1\\2\end{pmatrix} \middle| t \in \mathbb{R} \right\}$.
- Jeder Vektor ergibt eine Gerade durch den Ursprung. Doch sind manche dieser Geraden gleich: $\left\{ t\begin{pmatrix}1\\2\end{pmatrix} \middle| t \in \mathbb{R} \right\} = \left\{ t\begin{pmatrix}2\\4\end{pmatrix} \middle| t \in \mathbb{R} \right\}$ oder
$$\left\{ t\begin{pmatrix}1\\2\\3\end{pmatrix} \middle| t \in \mathbb{R} \right\} = \left\{ t\begin{pmatrix}-3\\-6\\-9\end{pmatrix} \middle| t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix}2\\4\\6\end{pmatrix} + t\begin{pmatrix}-1\\-2\\-3\end{pmatrix} \middle| t \in \mathbb{R} \right\}$$
- Es gibt auch Geraden, die nicht durch den Ursprung gehen, zum Beispiel $\left\{ \begin{pmatrix}2\\1\end{pmatrix} + t\begin{pmatrix}-5\\-1\end{pmatrix} \middle| t \in \mathbb{R} \right\}$.Wir sehen im Bild, dass die Gerade auch nicht durch $\begin{pmatrix}-5\\-1\end{pmatrix}$ geht: dies ist der Richtungsvektor, und bei einer Gerade, die nicht durch den Ursprung geht, kein Punkt auf der Geraden. Wenn zum Beispiel $t = 1$ ist, bekommen wir den Punkt $\begin{pmatrix}2\\1\end{pmatrix} + \begin{pmatrix}-5\\-1\end{pmatrix} = \begin{pmatrix}-3\\0\end{pmatrix}$.



In solch einer Geraden kommen also Vektoraddition und skalare Multiplikation zusammen.
Es kann einem zuerst seltsam vorkommen, dass wir uns b als Punkt und u als Pfeil (oder Familie von Pfeilen) vorstellen, obwohl beides Vektoren sind. Diese Flexibilität der Vorstellung ist oft hilfreich, und man kann sich recht gut daran gewöhnen, welche Vorstellung in einer bestimmten Situation am hilfreichsten ist, und wie man zwischen ihnen wechselt. Später in der Didaktik werden Sie noch Probleme dieser Vorstellung „Stützvektor plus Richtungsvektor" kennenlernen, aber für diesen Moment ist es vielleicht hilfreich. Im Zweifelsfall können wir immer noch einfach die Algebra benutzen.

Wir betrachten Geraden als **ein-dimensionale** Objekte. Wir haben diesen Begriff noch nicht formal definiert (das kommt erst in Kapitel 5), aber haben wahrscheinlich schon etwas Intuition: Wir können uns auf einer Geraden nur in eine Richtung bewegen (vor und zurück). Es kann hilfreich sein, sich dies als „einen Freiheitsgrad" vorzustellen: ich kann genau eine reelle Zahl wählen, nämlich den Parameter t, um einen Punkt auf der Geraden genau zu bestimmen.

Übung 1.8:

a) Sei $\{t \cdot u \mid t \in \mathbb{R}\}$ eine Gerade durch den Ursprung in $\mathbb{R}^n$, und seien $v_1 = t_1 u$ und $v_2 = t_2 u$ zwei Punkte auf dieser Geraden. Zeigen Sie, dass die Summe $v_1 + v_2$ auch auf der Geraden liegt, und ein skalares Vielfaches λv_1 (für $\lambda \in \mathbb{R}$) ebenso.

b) Sei nun $\{b + t \cdot u \mid t \in \mathbb{R}\}$ eine Gerade in $\mathbb{R}^n$, die nicht durch den Ursprung geht. Zeigen Sie, dass die Summe von zwei Punkten auf dieser Gerade *nicht* wieder auf der Geraden liegt.

Die obigen Beispiel werfen folgende Fragen auf:

- Wie können wir erkennen, ob eine Gerade durch den Ursprung geht oder nicht?
- Wie können wir erkennen, ob zwei Geraden gleich sind?
- Wie können wir erkennen, ob zwei Geraden parallel sind?

Übung 1.9: Benutzen Sie Beispiele, ausgehend von den oben gegebenen, um selbst Antworten auf diese Fragen zu finden. Finden Sie Formeln, die Ihnen helfen? Können Sie Ihre Aussagen beweisen?

Betrachten wir nun Ebenen. Diese haben zwei Freiheitsgrade, also sollten wir zwei Parameter erwarten.

Definition 1.10: Eine **Ebene** in $\mathbb{R}^n$ ist eine Menge der Form

$$\{b + su + tw \mid s, t \in \mathbb{R}\}$$

wobei $b, u, w \in \mathbb{R}^n$ so sind, dass u und w ungleich Null sind und nicht in dieselbe Richtung gehen.

Koordindatenform und Normalenform folgen später noch.

Bei Geraden hatten wir nur die Bedingung $u \neq 0$ für den Richtungsvektor, die garantiert, dass die Menge nicht nur ein Punkt ist. Nun haben wir eine zusätzliche Bedingung. Wir nehmen an, dass sie dazu da ist, zu verhindern, dass die Menge nur eine Gerade ist.

Was bedeutet „nicht in dieselbe Richtung gehen“ genau? Wir betrachten einige Beispiele.

Beispiele 1.11:

- Ebenso wie für Geraden kann eine Ebene durch den Ursprung gehen. Zum Beispiel ist dies die x,y-Ebene in $\mathbb{R}^3$:
$$\left\{ s\begin{pmatrix}1\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\1\\0\end{pmatrix} \middle| s, t \in \mathbb{R} \right\}.$$
Jeder Vektor in dieser Ebene hat die Form $\begin{pmatrix}s\\t\\0\end{pmatrix}$.
- Ebenso gibt es die x,z-Ebene und die y,z-Ebene:
$$\left\{ s\begin{pmatrix}1\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\0\\1\end{pmatrix} \middle| s, t \in \mathbb{R} \right\} \qquad \left\{ s\begin{pmatrix}0\\1\\0\end{pmatrix} + t\begin{pmatrix}0\\0\\1\end{pmatrix} \middle| s, t \in \mathbb{R} \right\}$$
- Eine weitere Ebene durch den Ursprung ist
$$\left\{ s\begin{pmatrix}1\\1\\0\end{pmatrix} + t\begin{pmatrix}0\\0\\1\end{pmatrix} \middle| s, t \in \mathbb{R} \right\}.$$
Sie können diese Ebene in GeoGebra rotieren um verschiedene Blinkwinkel zu erlangen.

Nicht-Beispiel 1.12: Ist dies auch eine Ebene?

$$\left\{ s \begin{pmatrix}1\\1\\0\end{pmatrix} + t \begin{pmatrix}3\\3\\0\end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$$

Algebraisch können wir dies zu $(s+3t)\cdot\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right)$ vereinfachen. Die zwei Vektoren $\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right)$ und $\left(\begin{smallmatrix}3\\3\\0\end{smallmatrix}\right)$ geben keine verschiedenen Richtungen an, sondern nur eine. Dies ist also keine Ebene, sondern eine Gerade.

Wir können unseren Fragen über Geraden also noch mehr Fragen hinzufügen:

- Wann ist eine Menge $\{b + su + tw \mid s,t \in \mathbb{R}\}$ wirklich eine Ebene?
- Wann geht eine Ebene durch den Ursprung?
- Wann sind zwei Ebenen gleich?

Ein Konzept, was uns bei der ersten Frage hilft, ist die Kollinearität.

Definition 1.13: Zwei Vektoren $u, v \in \mathbb{R}^n$ heißen **kollinear** wenn einer der beiden ein skalares Vielfaches des anderen ist: $u = \lambda v$ für ein $\lambda \in \mathbb{R}$ (oder anders herum).

Der Zusatz „oder anders herum“ deckt den Fall $v = 0$ ab. Dieser Nullvektor ist kollinear mit jedem anderen Vektor, weil $0 = 0 \cdot w$ für jeden Vektor $w \in \mathbb{R}^n$. Man könnte statt dessen auch sagen: $\mu u = \lambda v$ für $\lambda, \mu \in \mathbb{R}$, dann hat man beide Möglichkeiten abgedeckt.

Zwei Vektoren, die nicht kollinear sind, heißen **unabhängig**. Für eine Ebene brauchen wir zwei unabhängige Richtungsvektoren. Dies ist ein Beispiel des Konzepts der **linearen Unabhängigkeit**, das wir später noch genauer untersuchen werden.

Übung 1.14: Bestimmen Sie, welche dieser Mengen Ebenen sind. Für jede Ebene, bestimmen Sie, ob sie durch den Ursprung geht.

a) $\left\{ s \begin{pmatrix}1\\0\\4\end{pmatrix} + t \begin{pmatrix}8\\2\\0\end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$

b) $\left\{ \begin{pmatrix}1\\1\\1\end{pmatrix} + s \begin{pmatrix}1\\2\\0\end{pmatrix} + t \begin{pmatrix}0\\4\\2\end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$

c) $\left\{ \begin{pmatrix}3\\1\\1\end{pmatrix} + s \begin{pmatrix}1\\0\\3\end{pmatrix} + t \begin{pmatrix}-2\\0\\-6\end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$

d) $\left\{ \begin{pmatrix}1\\2\\1\end{pmatrix} + s \begin{pmatrix}3\\6\\2\end{pmatrix} + t \begin{pmatrix}1\\2\\1\end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$

e) $\left\{ \begin{pmatrix}-1\\3\\\frac{1}{2}\end{pmatrix} + s \begin{pmatrix}-2\\6\\2\end{pmatrix} + t \begin{pmatrix}1\\-3\\-\frac{3}{2}\end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$

Man kann auch von Ebenen oder von Geraden und Ebenen usw. Schnittpunkte bestimmen. Dies machen wir etwas später, wenn wir mehr Werkzeuge haben, die das noch viel leichter machen.

C. Vektorraum $\mathbb{R}^n$

Wir betrachten jetzt die Struktur der Vektoren in $\mathbb{R}^n$ etwas genauer.
Die Vektoraddition und skalare Multiplikation können wir vereinen:

Definition 1.15: Eine **Linearkombination** von Vektoren u, v ist ein Ausdruck der Form $\lambda u + \mu v$, mit Skalaren $\lambda, \mu \in \mathbb{R}$.

Dieses Konzept wird eine sehr wichtige Rolle im Kurs spielen.

Beispiele 1.16: ⋄ Mit $\lambda = \mu = 1$ sehen wir, dass die Summe zweier Vektoren ein Beispiel einer Linearkombination ist.

⋄ Mit $\mu = 0$ ist auch ein Vielfaches λu eine mögliche Linearkombination.

⋄ Wir können auch Linearkombinationen von mehr als zwei Vektoren haben:

$$\lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

ist eine Linearkombination von drei Vektoren. (ν wird „nü" ausgesprochen.)

Übung 1.17: ⋄ Wenn $v \in \mathbb{R}^n$ beliebig ist , was ist dann $0 \cdot v$?

⋄ Kann man jeden Vektor $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ als Linearkombination von $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ schreiben?

⋄ Kann man jeden Vektor $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ als Linearkombination von $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ schreiben? Geht dies auch als Linearkombination aus nur den ersten drei Vektoren?

Vektoraddition und skalare Multiplikation erfüllen bestimmte Eigenschaften. Manche dieser Eigenschaften kennen wir schon aus dem Kurs Grundlagen (im weiteren als MAGL abgekürzt): $\mathbb{R}^n$ mit Vektoraddition bildet eine kommutative Gruppe. Die Eigenschaften der skalaren Multiplikation schreiben wir hier auch auf.

⚠ Vorsicht: wir dürfen skalare Multiplikation nicht mit der Multiplikation in einem Ring verwechseln: in einem Ring oder Körper (bekannt aus MAGL) multiplizieren wir zwei Elemente der selben Menge. Bei der Skalarmultiplikation haben wir zwei verschiedene Objekte: einen Skalar und einen Vektor, und das Ergebnis ist wieder ein Vektor.

Wir kommen später noch mal auf diese Vektorraumeigenschaften zurück.

Proposition 1.18: (Eigenschaften von Vektoraddition und skalarer Multiplikation)
Für beliebige Vektoren $u, v, w \in \mathbb{R}^n$ und Skalare $\lambda, \mu \in \mathbb{R}$ gilt:

VA0	$u + v \in \mathbb{R}^n$	*(Abgeschlossenheit unter Vektoraddition)*
VA1	$v + 0 = v = 0 + v.$	*(Nullvektor)*
VA2	*Es gibt* $-v$ *mit* $v + (-v) = 0 = (-v) + v.$	*(Negative)*
VA3	$(u + v) + w = u + (v + w)$	*(Assoziativität der Vektoraddition)*
VA4	$u + v = v + u$	*(Kommutativität der Vektoraddition)*
SM0	$\lambda v \in \mathbb{R}^n$	*(Abgeschlossenheit unter Skalarmultiplikation)*
SM1	$1 \cdot v = v$	*(Einheitsskalar)*
SM2	$\lambda \cdot (\mu v) = (\lambda \cdot \mu) v$	*(Assoziativität der Skalarmultiplikation)*
SM3	$(\lambda + \mu)\, v = \lambda v + \mu v$	*(Distributivität von Skalarmult über reelle Addition)*
SM4	$\lambda(u + v) = \lambda u + \lambda v$	*(Distributivität von Skalarmult über Vektoraddition)*

Wir nennen jede Struktur, die alle obigen Eigenschaften erfüllt, einen ***Vektorraum****.*

Die negativen Vektoren sind also die additiven Inversen in der additiven Gruppe $(\mathbb{R}^n, +)$.

BEWEIS. Da Vektoraddition und Skalarmultiplikation komponentenweise definiert sind, gehen alle diese Eigenschaften auf die Körpereigenschaften von $\mathbb{R}$ zurück. Diese werden dann jeweils in jeder Komponente angewendet. Man kann es auch leicht für $\mathbb{R}^2$ oder $\mathbb{R}^3$ ausprobieren, und dann sieht man, dass sich für Vektoren aus $\mathbb{R}^n$ auch nichts ändert. □

Wir werden später sehen, dass nicht nur $\mathbb{R}^n$ diese Eigenschaften erfüllt, sondern auch andere Strukturen. Solche Vektorräume sind der Kontext oder das Fundament der linearen Algebra.

VA steht für Vektoraddition, und SM steht für Skalarmultiplikation. Warum fangen wir mit 0 an? Weil die Definition von Vektoraddition und Skalarmultiplikation eigentlich schon diese Abgeschlossenheit voraussetzt. Aber es ist sehr nützlich, diese beiden Eigenschaften trotzdem in unsere Liste mit aufzunehmen, damit wir sie später nicht vergessen, wenn wir etwas prüfen wollen.
Ein kleiner Kommentar zur Wichtigkeit von **Assoziativität**: Man könnte meinen, dass diese Eigenschaft nicht sehr wichtig ist, sie ist doch immer wahr, oder? Nein! Die Operation "hoch" erfüllt sie zum Beispiel nicht: Ist $2^{(3^2)} = (2^3)^2$? Nein! $2^{(3^2)} = 2^9$ und $(2^3)^2 = 2^6$.

D. Matrizen

Unsere zweite wichtige Zutat für Lineare Algebra sind Matrizen.

Definition 1.19: Eine **Matrix** ist eine rechteckige Anordnung von Einträgen (aus $\mathbb{R}$):

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Eine Matrix hat **Zeilen** und **Spalten**. Wir bezeichnen die **Größe der Matrix** mit $m \times n$, wobei m die Anzahl der Zeilen und n die Anzahl der Spalten ist. Wir schreiben $\mathcal{M}_{m,n}$ für die Menge aller $m \times n$-Matrizen. Wenn $m = n$, wenn also eine Matrix die gleiche Anzahl Zeilen wie Spalten hat, nennen wir sie eine **quadratische** Matrix.

Wofür steht eine Matrix, oder was bedeutet sie? Die Antwort: es gibt viele Möglichkeiten, abhängig vom Kontext! Eine Matrix kann folgendes repräsentieren:

- eine Ansammlung von Daten,
- eine Abbidlung, die Vektoren einer bestimmten Länge auf andere Vektoren schickt,
- ein lineares Gleichungssystem,
- oder noch andere Möglichkeiten.

In diesem Kurs konzentrieren wir uns zumeist auf Matrizen als Abbildungen, und als lineare Gleichungssysteme. Für beides müssen wir wissen, wie man eine Matrix mit einem Vektor multipliziert. Aber wir machen erst einmal ein paar einfachere Dinge mit Matrizen.

Notation 1.20: Wir schreiben Matrizen üblicherweise mit Großbuchstaben A, B, C usw., oder vielleicht M. Wir schreiben a_{ij} oder A_{ij} für die Einträge von A: der erste Index i gibt die Zeile und der zweite Index j die Spalte.

Beispiele 1.21: ◇ $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ ist eine 1×2-Matrix mit Einträgen $a_{11} = 1$ und $a_{12} = 0$.

◇ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist eine 2×2-Matrix. Dies ist eine quadratische Matrix.

◇ $B = \begin{pmatrix} \frac{3}{4} & -\frac{2}{5} & \frac{1}{2} \\ \frac{9}{102} & \pi & \sqrt{3} \end{pmatrix}$ ist eine 2×3-Matrix. Einige der Einträge sind $b_{12} = -\frac{2}{5}$, $b_{23} = \sqrt{3}$.

◇ Ein Vektor in $\mathbb{R}^n$ kann auch als $n \times 1$-Matrix betrachtet werden.

◇ Eine 1×1-Matrix ist einfach eine Zahl: x. Wir lassen dann für gewöhnlich die Klammern weg.

◇ $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$ ist eine allgemeine 3×2-Matrix.

Viele Beispiele in diesem Skript und auch in Lehrbüchern benutzen unverhältnismäßig oft ganzzahlige Einträge. Das ist nur, weil wir Menschen es leichter finden, mit ganzen Zahlen zu rechnen. Wenn wir also ein neues Konzept erklären wollen, machen wir den Kontext so angenehm wie möglich. Im „echten Leben“ gibt es nicht so viele ganze Zahlen. Aber wir können ja auch Computer zum Rechnen benutzen, wir müssen nur die Konzepte verstehen. Und das können wir genauso gut mit ganzen Zahlen lernen.

Definition 1.22: Zwei Matrizen A und B sind **gleich** wenn sie die selbe Größe haben und alle Einträge übereinstimmen.

Beispiele 1.23: ◇ $\begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix}$ und $\begin{pmatrix} 3 & 2 & 0 \\ 1 & 3 & 0 \end{pmatrix}$ sind nicht gleich, weil sie nicht die selbe Größe haben.

◇ $A = \begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix}$ und $B = \begin{pmatrix} x & 2 \\ 1 & y \end{pmatrix}$ sind nicht gleich, weil nicht alle Einträge übereinstimmen. Können wir Werte für x und y wählen, für die $A = B$ ist?

◇ Seien A und B Matrizen mit Einträgen $A_{ij} = i^j$ und $B_{ij} = i^j$. Muss dann $A = B$ sein? NEIN: Vielleicht ist A eine 3×3-Matrix aber B ist eine 2×3-Matrix. Diese sind dann nicht gleich, aber wir hätten es nicht gemerkt, wenn wir nur die Formel für die Einträge überprüft hätten.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \\ 3 & 9 & 27 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \end{pmatrix}$$

Übung 1.24: Erstellen Sie eine Matrix aus Formeln für die Einträge online in Numbas.

Definition 1.25: In einer quadratischen Matrix A bezeichnen wir die Einträge A_{ii} als die **Diagonaleinträge**. Die Summe der Diagonaleinträge heißt **Spur** der Matrix, geschrieben $\mathrm{Spur}(A)$ (oder auch $\mathrm{tr}A$ für „trace“).

Beispiel 1.26:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \text{ hat } \mathrm{Spur}\, A = a_{11} + a_{22} + a_{33}.$$

Definition 1.27: Eine quadratische Matrix, in der alle Diagonaleinträge 1 und alle anderen Einträge 0 sind, heißt **Einheitsmatrix** (auch **Identität**) und wird als $\mathbb{1}$, $\mathbb{1}_n$ oder als I, I_n geschrieben. (In manchen Quellen auch E_n.) Hierbei sind $\mathbb{1}_n$ und I_n von der Größe $n \times n$.

Beispiele 1.28:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Definition 1.29: Für zwei Matrizen A und B der selben Größe ist deren **Matrixsumme** $A+B$ eintragsweise definiert: $(A+B)_{ij} = A_{ij} + B_{ij}$.
Wir können auch eine Matrix mit einem Skalar multiplizieren, wieder eintragsweise: $(\lambda A)_{ij} = \lambda A_{ij}$.

Beispiele 1.30: Wir können nur Matrizen gleicher Größe addieren!

- $\begin{pmatrix} 3 & 1 \\ 2 & 9 \end{pmatrix} + \begin{pmatrix} 7 & 2 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} 10 & 3 \\ 10 & 12 \end{pmatrix}$
- $3 \cdot \begin{pmatrix} 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 9 \end{pmatrix}$
- $4 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 3 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

Genau wie für Vektoren nennen wir eine Kombination von Matrixaddition und skalarer Multiplikation eine **Linearkombination von Matrizen**.

Übung 1.31: Wenn möglich, addieren Sie folgende Matrizen.

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

Proposition 1.32: (Matrizen bilden einen Vektorraum)
Die Menge $\mathcal{M}_{m,n}$ aller Matrizen der Göße $m \times n$ erfüllt, mit Matrixaddition und Skalarmultiplikation wie oben, die Eigenschaften VA0-4 und SM0-4 und ist somit ein Vektorraum.

BEWEIS. Dies kann man genauso prüfen wir für Vektoren in $\mathbb{R}^n$: alle Operationen passieren nur eintragweise. □

Wir können Matrizen auch „spiegeln“:

Definition 1.33: Für eine $m \times n$-Matrix A ist die **transponierte Matrix** die $n \times m$-Matrix A^T mit Einträgen $(A^T)_{ij} = A_{ji}$.

Wir vertauschen also Zeilen und Spalten.

Beispiele 1.34: Man kann sowohl quadratische als auch nicht-quadratische Matrizen transponieren.

- $\begin{pmatrix} 1 & 3 & 2 \\ 9 & 2 & -1 \end{pmatrix}^T = \begin{pmatrix} 1 & 9 \\ 3 & 2 \\ 2 & -1 \end{pmatrix}$
- $\begin{pmatrix} 2 & 3 \\ 1 & 8 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 \\ 3 & 8 \end{pmatrix}$
- $\begin{pmatrix} 3 & 2 & 1 \end{pmatrix}^T = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$

Man sieht hier, dass die Diagonaleinträge dabei gleich bleiben.

Proposition 1.35: (Eigenschaften der Transponierten)
Für beliebige Matrizen A, B gleicher Größe und $\lambda \in \mathbb{R}$ gilt

(i) $(A^T)^T = A$;
(ii) $(A+B)^T = A^T + B^T$;
(iii) $(\lambda A)^T = \lambda A^T$.

Beweis. Wir prüfen zuerst, dass jeweils die Matrizen links und rechts die gleiche Größe haben. Danach:

(i) $((A^T)^T)_{ij} = (A^T)_{ji} = A_{ij}$
(ii) $((A+B)^T)_{ij} = (A+B)_{ji} = A_{ji} + B_{ji} = (A^T)_{ij} + (B^T)_{ij} = (A^T + B^T)_{ij}$
(iii) $((\lambda A)^T)_{ij} = (\lambda A)_{ji} = \lambda A_{ji} = \lambda (A^T)_{ij} = (\lambda A^T)_{ij}$ □

Beispiel 1.36:

$$\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T\right)^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

E. Matrixmultiplikation

Jetzt lernen wir die wichtige Interaktion zwischen einer Matrix und einem Vektor.

Notation 1.37: Wir bezeichnen die Spalten einer Matrix A mit $A_1, A_2, \dots, A_n$ oder $a_1, a_2, \dots, a_n$:

$$\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix}$$

Definition 1.38: (Matrix mal Vector) Eine $m \times n$-Matrix A bildet einen Vektor $v \in \mathbb{R}^n$ auf einen Vektor $Av \in \mathbb{R}^m$ ab:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{m-11} & a_{m-12} & \cdots & a_{m-1n-1} & a_{m-1n} \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m-11}x_1 + a_{m-12}x_2 + \cdots + a_{m-1n}x_n \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

oder

$$\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} \uparrow \\ a_1 \\ \downarrow \end{pmatrix} + x_2 \begin{pmatrix} \uparrow \\ a_2 \\ \downarrow \end{pmatrix} + \cdots + x_{n-1} \begin{pmatrix} \uparrow \\ a_{n-1} \\ \downarrow \end{pmatrix} + x_n \begin{pmatrix} \uparrow \\ a_n \\ \downarrow \end{pmatrix}$$

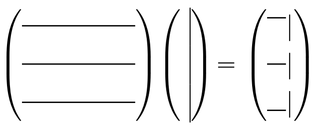
Beispiele 1.39: ◇ Besteht die Matrix nur aus einer Zeile, bekommt man eine einzige Zahl als Antwort. Dies ist vielleicht aus der Schule als „Skalarprodukt von Vektoren“ bekannt:

$$\begin{pmatrix} a & b & c & d & e \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = ax_1 + bx_2 + cx_3 + dx_4 + ex_5$$

◇ 2×2 und 3×3 explizit:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix}$$

◇ Vielleicht hilft dieses Bild zur Visualisierung und/oder Erinnerung der Formel:



Denken Sie dies als „Zeile mal Spalte“: der erste Eintrag einer Zeile mal den ersten Eintrag einer Spalte, plus den zweiten Eintrag der Zeile mal den zweiten Eintrag der Spalten, und so weiter. Wir wenden also einfach das erste Beispiel oben mehrfach an.

◇ Hier ist ein Beispiel mit Zahlen:

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

◇ Wenn wir einen Vektor mit der Einheitsmatrix multiplizieren, ändert er sich nicht:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

◇ Die zweite Sichtweise der Definition kann manchmal sehr hilfreich sein: Eine Matrix mal ein Vektor ist eine Linearkombination der Spalten der Matrix, mit den Einträgen des Vektors als Koeffizienten. Für eine 3×3-Matrix sieht diese zweite Sichtweise so aus:

$$\begin{pmatrix} \uparrow & \uparrow & \uparrow \\ a_1 & a_2 & a_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} \uparrow \\ a_1 \\ \downarrow \end{pmatrix} + x_2 \begin{pmatrix} \uparrow \\ a_2 \\ \downarrow \end{pmatrix} + x_3 \begin{pmatrix} \uparrow \\ a_3 \\ \downarrow \end{pmatrix}$$

Diese Sichtweise ist meist hilfreicher, wenn man bestimmte Eigenschaften und Zusammenhänge verstehen will, als für explizite Berechnungen. Aber wenn Sie diese Sichtweise auch für Berechnungen hilfreich finden, können Sie sie trotzdem benutzen.

Übung 1.40: Berechnen Sie:

◇ $\begin{pmatrix} 3 & 2 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ◇ $\begin{pmatrix} 1 & 3 & 2 \\ 9 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ ◇ $\begin{pmatrix} 1 & 9 & 0 \\ 8 & 0 & 1 \\ 0 & 2 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ ◇ $\begin{pmatrix} 1 & 9 \\ 3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

Oder Matrix mal Vektor in Numbas.

Mit dieser Multiplikation können wir eine Matrix als Abbildung (oder Funktion) betrachten.

Bemerkung 1.41: In diesem Kurs ist eine **Abbildung** oder **Funktion** $f\colon X \longrightarrow Y$ immer linkstotal. Es wird also *jedem* Element $x \in X$ genau ein Funktionswert $f(x) \in Y$ zugeordnet. (Abbildung und Funktion heißen hier genau das gleiche.)

Definition 1.42: Eine $m \times n$-Matrix A bestimmt eine Abbildung $f_A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$: sie bildet einen Vektor $v \in \mathbb{R}^n$ auf $f_A(v) = Av$ ab. Wir nennen diese Abbildung eine **Matrixtransformation** oder **Matrixabbildung**.

Beispiele 1.43:
- $A = \begin{pmatrix} 3 & 2 \\ 5 & 1 \end{pmatrix}$ schickt einen Vektor $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ auf $\begin{pmatrix} 3x+2y \\ 5x+y \end{pmatrix} \in \mathbb{R}^2$. Diese Matrix gibt also eine Matrixabbildung $f_A\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$.
- $B = \begin{pmatrix} 1 & 3 & 2 \\ 9 & 2 & -1 \end{pmatrix}$ gibt eine Matrixabbildlung $f_B\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$.
- $C = \begin{pmatrix} 1 & 9 \\ 3 & 2 \\ 2 & -1 \end{pmatrix}$ gibt eine Matrixabbildung $f_C\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3$, in umgekehrter Richtung zu B.
- Eine Einheitsmatrix $\mathbb{1}_n$ entspricht der **Identitätsabbildung** $\mathrm{id}_n\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$: jeder Vektor wird auf sich selbst geschickt. $\mathbb{1}_n v = v$, also ist für beliebiges $v \in \mathbb{R}^n$ $f_{\mathbb{1}_n}(v) = v$, und somit $f_{\mathbb{1}_n} = \mathrm{id}_n$.

Bemerkung 1.44: Die Identitätsabbildung $\mathrm{id}_X\colon X \longrightarrow X$ mit $\mathrm{id}_X(x) = x$ für jedes $x \in X$ ist eine der wichtigsten Abbildungen überhaupt.
Natürlich ist es nicht nur eine Abbildung, sondern viele: eine für jede Menge X.
Wegen des Beispiels oben kann man die Einheitsmatrix $\mathbb{1}_n$ auch Identität nennen. (Sie ist auch das neutrale Element der Matrixmultiplikation, was auch manchmal Identität heißt.)

Wir können nicht nur eine Matrix mit einem Vektor multiplizieren, sondern auch Matrizen von passenden Größen.

Definition 1.45: (Matrixmultiplikation) Wenn A eine $m \times r$-Matrix ist und B eine $r \times n$-Matrix, dann können wir das **Matrixprodukt** AB mit Einträgen $(AB)_{ij} = \sum_{k=1}^{r} A_{ik}B_{kj}$ bilden. Dies ist eine $m \times n$-Matrix.
Wir können das Matrixprodukt auch mit Hilfe der Spalten der zweiten Matrix aufschreiben: Wenn

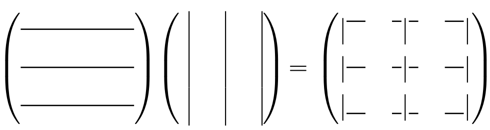
$$B = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ b_1 & \cdots & b_k & \cdots & b_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix},$$

dann ist

$$AB = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ Ab_1 & \cdots & Ab_k & \cdots & Ab_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix},$$

wo wir die schon definierte Operation Matrix mal Vektor benutzen.

Zur Visualisierung:



Der Eintrag in Zeile i und Spalte j des Matrixproduktes AB wird als „Zeile i der Matrix A mal Spalte j der Matrix B“ berechnet.
⚠ Wichtig! Man kann nur Matrizen passender Größen multiplizieren! Zur Berechnung von „Zeile der Matrix A mal Spalte der Matrix B“ brauchen wir genau so viele Einträge in der Zeile von A wie Einträge in der Spalte von B. Also muss A so viele Spalten haben wie B Zeilen hat:

$$\underset{m\times r}{A} \cdot \underset{r\times n}{B} = \underset{m\times n}{AB}$$

Diese mittlere Größe r, die übereinstimmen muss, „hebt sich auf“, und somit bekommt man die Größe $m \times n$ des Matrixprodukts.
Wir sehen ein ähnliches Muster in der Formel zur Berechnung eines Eintrags des Matrixprodukts, mit dem gleichen Index k in der Mitte.

Beispiele 1.46: ◇ Unsere vorherigen Beispiele von „Matrix mal Vektor“ sind hier immer noch Beispiele, weil ein Vektor als Matrix mit nur einer Spalte gesehen werden kann. Insbesondere, wenn zum Beispiel $A = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{pmatrix}$ eine 1×5-Matrix und $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix}$ eine 5×1-Matrix ist, dann ist das Produkt AB eine 1×1-Matrix, also eine Zahl.

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4 + a_5b_5 = \sum_{k=1}^{5} a_k b_k = \sum_k A_{1k} B_{k1}$$

In jedem anderen Beispiel machen wir diese Berechnung mit jeder Zeile der ersten Matrix und jeder Spalte der zweiten Matrix.

◇ Wenn A und B beides 2×2-Matrizen sind, bekommen wir

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Wenn wir genau hinsehen, bemerken wir, dass der (i,j)te Eintrag gleich $a_{i1}b_{1j} + a_{i2}b_{2j}$ ist. Dies ist die Summe aus der Definition, ausgeschrieben für dieses Beispiel.

◇ Wenn A eine 3×3-Matrix ist und B eine 3×2-Matrix, dann ist beispielsweise der Eintrag in der ersten Zeile und zweite Spalte des Produkts gleich

$$(AB)_{12} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} = \sum_k a_{1k} b_{k2}$$

◇ Hier sind ein paar Beispiele mit Zahlen:

$$\underset{3\times 3}{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}} \underset{3\times 2}{\begin{pmatrix} 10 & 11 \\ 12 & 13 \\ 14 & 15 \end{pmatrix}} = \underset{3\times 3}{\begin{pmatrix} 1\cdot 10 + 2\cdot 12 + 3\cdot 14 & 1\cdot 11 + 2\cdot 13 + 3\cdot 15 \\ 4\cdot 10 + 5\cdot 12 + 6\cdot 14 & 4\cdot 11 + 5\cdot 13 + 6\cdot 15 \\ 7\cdot 10 + 8\cdot 12 + 9\cdot 14 & 7\cdot 11 + 8\cdot 13 + 9\cdot 15 \end{pmatrix}}$$

$$\underset{3\times 2}{\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}} \underset{2\times 3}{\begin{pmatrix} 10 & 11 & 12 \\ 13 & 14 & 15 \end{pmatrix}} = \underset{3\times 3}{\begin{pmatrix} 1\cdot 10 + 2\cdot 13 & 1\cdot 11 + 2\cdot 14 & 1\cdot 12 + 2\cdot 15 \\ 3\cdot 10 + 4\cdot 13 & 3\cdot 11 + 4\cdot 14 & 4\cdot 12 + 4\cdot 15 \\ 5\cdot 10 + 6\cdot 13 & 5\cdot 11 + 6\cdot 14 & 5\cdot 12 + 6\cdot 15 \end{pmatrix}}$$

Übung 1.47: Üben Sie Matrixmultiplikation in Numbas.
Vertiefen Sie Ihr Verständnis, indem Sie ein Beispiel suchen: zwei von Null verschiedene Matrizen, deren Produkt die Nullmatrix ist. Überprüfen Sie ihr Beispiel in Numbas.

⚠ VORSICHT! Im Allgemeinen ist $AB \neq BA$. Meist ist die umgekehrte Multiplikation nicht einmal definiert. Aber selbst für quadratische Matrizen sind die beiden Seiten meist nicht gleich.

Beispiel 1.48:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 3 & 8 \end{pmatrix} \quad \text{aber} \quad \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 6 & 8 \end{pmatrix}.$$

Die Einheitsmatrix bekommt wieder eine Sonderstellung:

Proposition 1.49: (Produkt mit Einheitsmatrix)
Für jede $m \times n$-Matrix A gilt $AI_n = A$ and $I_m A = A$: Multiplikation mit der Einheitsmatrix korrekter Größe ändert die Matrix nicht.

Wir können „mal Einheitsmatrix" als Verallgemeinerung von „mal 1" in $\mathbb{R}$ betrachten.

BEWEIS. Nach der Definition der Matrixmulitplikation haben wir $(AI_n)_{ij} = \sum_{k=1}^{n} A_{ik}(I_n)_{kj}$. Aber $(I_n)_{kj} = 0$ für $k \neq j$, und $(I_n)_{jj} = 1$. Also reduziert sich diese Summe zu $(AI_n)_{ij} = \sum_{k=1}^{n} A_{ik}(I_n)_{kj} = A_{ij}(I_n)_{jj} = A_{ij}$. Die Matrizen A und AI_n haben also die gleichen Einträge (und die gleiche Größe), daher ist $A = AI_n$.
Analog für $I_m A$. □

Wenn in einem Beweis „analog" oder „ähnlich" steht, dann ist das eine „versteckte Übung": Sie können für sich überprüfen, ob Sie wirklich alles verstanden haben, indem Sie diesen Teil selbst probieren. Versuchen Sie es erst, ohne den vorherigen Teil des Beweises anzuschauen. Benutzen Sie dann den vorherigen Teil als Hinweis, wenn Sie stecken bleiben.
Bei Matrixmulitplikation müssen wir nicht auf Klammersetzung achten:

Proposition 1.50: (Assoziativität der Matrixmulitplikation)
Matrixmulitplikation ist assoziativ: $(AB)C = A(BC)$ gilt für alle Matrizen, für die diese Multiplikationen definiert sind.

BEWEIS. Zuerst prüfen wir, ob die Größen der Matrizen $(AB)C$ und $A(BC)$ übereinstimmen. (Übung.)
Nun bestimmen wir einen Eintrag der Matrizen auf beiden Seiten.

$$((AB)C)_{ij} = \sum_k (AB)_{ik}C_{kj} = \sum_k \left(\sum_s A_{is}B_{sk} \right) C_{kj} = \sum_k \sum_s A_{is}B_{sk}C_{kj}$$

wobei der letzte Schritt die Klammer ausmultipliziert.

$$(A(BC))_{ij} = \sum_s A_{is}(BC)_{sj} = \sum_s A_{is} \left(\sum_k B_{sk}C_{kj} \right) = \sum_s \sum_k A_{is}B_{sk}C_{kj}$$

Beide Ausdrücke sind gleich, also ist $A(BC) = (AB)C$. Die unterschiedliche Reihenfolge der Summenzeichen ergibt nur unterschiedliche Reihenfolge der Summanden. Aber da Addition in $\mathbb{R}$ kommutativ ist, macht dies nichts aus. □

Bemerkung 1.51: Dies sagt uns, dass die Menge der $n \times n$-Matrizen unter Matrixmultiplikation eine Halbgruppe bilden. (Wir nehmen nur $n \times n$-Matrizen, damit wir alle Matrizen der Menge miteinander multiplizieren können.) Das Produkt mit der Einheitsmatrix (Prop. 1.49) zeigt uns, dass diese Halbgruppe ein neutrales Element hat, nämlich I_n.

Übung 1.52: Seien A eine $m \times r$-Matrix, B eine $r \times l$-Matrix und C eine $l \times n$-Matrix. Schreiben Sie den obigen Beweis aus mit expandierten Summen. Zum Beispiel

$$((AB)C)_{ij} = (AB)_{i1}C_{1j} + (AB)_{i2}C_{2j} + (AB)_{i3}C_{3j} + \cdots + (AB)_{il}C_{lj}$$

und dann das selbe für jedes $(AB)_{ik}$ und so weiter.
Alternativ können Sie dies zuerst für $m = r = l = 2$ aufschreiben.
Sinn dieser Übung: Sie sollen sehen, wie kompliziert Matrixmultiplikation wird; dieses Resultat ist bei weitem nicht so offensichtlich wie $(A + B) + C = A + (B + C)$! Zudem sehen Sie vielleicht hier den Vorteil der Summennotation mit Sigma.

Es gibt einige schöne Formen von Matrizen, die auch unter Multiplikation erhalten werden.

Definition 1.53: Eine **diagonale Matrix** oder **Diagonalmatrix** ist eine quadratische Matrix, in der nur die Diagonaleinträge von Null verschieden sind:

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 & 0 \\ 0 & d_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & d_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & d_n \end{pmatrix}$$

oder A mit $A_{ij} = d_i \delta_{ij}$, oder mit $A_{ij} = 0$ für $i \neq j$.

Übung 1.54: Zeigen Sie, dass das Produkt von zwei Diagonalmatrizen wieder eine Diagonalmatrix ist. Wie hängen die Diagonaleinträge des Produkts mit den Diagonaleinträgen der ursprünglichen Matrizen zusammen?

Definition 1.55: Eine **obere Dreiecksmatrix** ist eine quadratische Matrix, in der alle Einträge unter der Diagonale Null sind:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2(n-1)} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{(n-1)(n-1)} & a_{(n-1)n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

oder A mit $A_{ij} = 0$ für $i > j$.
Eine **untere Dreiecksmatrix** ist eine quadratische Matrix, in der alle Einträge über der Diagonale Null sind: $A_{ij} = 0$ für $i < j$.

Übung 1.56: Zeigen Sie, dass das Produkt von zwei oberen Dreiecksmatrizen wieder eine obere Dreicksmatrix ist.

Eine quadratische Matrix können wir immer wieder mit sich selbst multiplizieren:

Definition 1.57: Für eine quadratische Matrix A ist die k**te Potenz** von A

$$A^k = AA \cdots A,$$

das Produkt von k Kopien von A. Wir definieren $A^0 = I$.

Proposition 1.58: (Transponierte eines Produkts)
$(AB)^T = B^T A^T$ für alle Matrizen A und B, für die das Produkt AB definiert ist.

Beweis. Zuerst prüfen wir, dass $(AB)^T$ und $B^T A^T$ die gleiche Größe haben. (Übung.) Dann schauen wir uns einen Eintrag an.

$$((AB)^T)_{ij} = (AB)_{ji} = \sum_k A_{jk} B_{ki}$$

und

$$(B^T A^T)_{ij} = \sum_k (B^T)_{ik} (A^T)_{kj} = \sum_k B_{ki} A_{jk}.$$

Da beide Audrücke gleich sind, gilt $(AB)^T = B^T A^T$. □

Beispiel 1.59:

$$\left(\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right)^T = \begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}^T = \begin{pmatrix} 1 & 1 \\ 4 & -2 \end{pmatrix}$$

aber

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

Andererseits ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^T \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -2 \end{pmatrix}.$$

Wenn Sie einmal vergessen, wie man $(AB)^T$ berechnet, können Sie es sich erarbeiten, in dem Sie nicht-quadratische Matrizen A und B nehmen und jeweils die Größen betrachten. Dann gibt es nur eine Möglichkeit, wie A^T und B^T überhaupt zusammen passen.

Beispiel 1.60:

$$\underset{2\times 3}{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}} \underset{3\times 4}{\begin{pmatrix} 10 & 11 & 12 & 13 \\ 14 & 15 & 16 & 17 \\ 18 & 19 & 20 & 21 \end{pmatrix}} = AB$$

können wir multiplizieren.

$$\underset{4\times 3}{\begin{pmatrix} 10 & 14 & 18 \\ 11 & 15 & 19 \\ 12 & 16 & 20 \\ 13 & 17 & 21 \end{pmatrix}} \underset{3\times 2}{\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}} = B^T A^T$$

können wir auch multiplizieren, aber

$$\underset{3\times 2}{\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}} \underset{4\times 3}{\begin{pmatrix} 10 & 14 & 18 \\ 11 & 15 & 19 \\ 12 & 16 & 20 \\ 13 & 17 & 21 \end{pmatrix}} = A^T B^T$$

nicht: die Größen passen nicht zusammen.

Wir haben gesehen, dass eine Matrix als Abbildung betrachtet werden kann. Was passiert also, wenn wir zwei Matrizen multiplizieren?

Definition 1.61: Für zwei Abbildungen $f\colon X \longrightarrow Y$ und $g\colon Y \longrightarrow Z$ ist die **Verkettung** oder **Komposition** $g\circ f\colon X \longrightarrow Z$ (lies „g nach f“) von g und f durch $(g\circ f)(x) = g(f(x))$ definiert.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g\circ f"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

Wir führen also erst f und dann g aus. Eselsbrücke: wenn wir $g\circ f$ auf ein Element x anwenden, also $(g\circ f)(x)$, dann steht das f genau neben dem x, wird also zuerst angewendet.

Proposition 1.62: (Komposition von Matrixabbildungen)
Für zwei Matrizen $A \in \mathcal{M}_{m,r}$ und $B \in \mathcal{M}_{r,n}$ ist die Komposition der Matrixabbildungen $f_A\colon \mathbb{R}^r \longrightarrow \mathbb{R}^m$ und $f_B\colon \mathbb{R}^n \longrightarrow \mathbb{R}^r$ genau die Matrixabbildung des Matrixprodukts:

$$f_{AB} = f_A\circ f_B\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "f_B"] \arrow[dr, "f_{AB}"'] & \mathbb{R}^r \arrow[d, "f_A"] \\ & \mathbb{R}^m \end{tikzcd}$$

Beweis. Zuerst prüfen wir, dass f_{AB} und $f_A\circ f_B$ wirklich beide von $\mathbb{R}^n$ nach $\mathbb{R}^m$ gehen (siehe Diagramm).
Für beliebiges $v \in \mathbb{R}^n$ haben wir $f_{AB}(v) = (AB)v$ und $(f_A\circ f_B)(v) = f_A(f_B(v)) = A(Bv)$. Da Matrixmultiplikation assoziativ ist, folgt $f_{AB} = f_A\circ f_B$. □

Nun erinnern wir uns daran, dass wir in der Einleitung gesagt haben, dass Linearkombinationen eins der wichtigsten Konzepte der linearen Algebra sind. Und Matrixmulitplikation erhält Linearkombination!

Proposition 1.63: (Matrixmultiplikation ist linear)
Matrixmultiplikation erhält Linearkombinationen. Wenn A, A' beides $m\times r$-Matrizen und B, B' beides $r \times n$-Matrizen sind, und $\lambda, \mu \in \mathbb{R}$, dann ist

$$A(\lambda B + \mu B') = \lambda AB + \mu AB' \qquad \textit{und} \qquad (\lambda A + \mu A')B = \lambda AB + \mu A'B.$$

Beweis. Wir müssen prüfen, dass die Matrizen auf beiden Seiten einer Gleichung die gleiche Größe und die gleichen Einträge haben.
A ist eine $m \times r$-Matrix, und $\lambda B + \mu B'$ ist eine $r \times n$-Matrix, da sowohl B als auch B' diese Größe haben. Also ist $A(\lambda B+\mu B')$ eine $m\times n$-Matrix. Ebenso sind AB und AB', also auch $\lambda AB+\mu AB'$, $m \times n$-Matrizen. Also stimmen die Größen überein. Nun schauen wir einen Eintrag an.

$$\begin{aligned}
(A(\lambda B + \mu B'))_{ij} &= \sum_{k=1}^{r} A_{ik}(\lambda B + \mu B')_{kj} && \text{Def Matrixmultiplikation} \\
&= \sum_{k=1}^{r} A_{ik}(\lambda B_{kj} + \mu B'_{kj}) && \text{Def Matrixaddition, skalare Mult} \\
&= \sum_{k=1}^{r} (\lambda A_{ik} B_{kj} + \mu A_{ik} B'_{kj}) && \text{Ausmultiplizieren in } \mathbb{R} \\
&= \lambda \sum_{k=1}^{r} A_{ik} B_{kj} + \mu \sum_{k=1}^{r} A_{ik} B'_{kj} && \text{Umstellen, ausklammern in } \mathbb{R} \\
&= \lambda(AB)_{ij} + \mu(AB')_{ij} && \text{Def Matrixmultiplikation} \\
&= (\lambda AB + \mu AB')_{ij} && \text{Def Matrixaddition, skalare Mult}
\end{aligned}$$

Analog für die zweite Gleichung. □

Es ist eine gute Übung für Sie, die zweite Gleichung zu beweisen: Versuchen Sie es zuerst, ohne noch einmal diesen Beweis anzuschauen, und benutzen Sie dies nur als Hinweis wenn nötig.
Sie könnten auch die Summen ohne Summenzeichen ausschreiben. Vielleicht für 2×2-Matrizen.

Korollar 1.64: *Matrixabbildungen sind linear (erhalten Linearkombinationen). Das heißt*

$$f_A(\lambda u + \mu v) = \lambda f_A(u) + \mu f_A(v).$$

Beweis. Dies folgt direkt aus der Definition $f_A(v) = Av$ und dem Resultat, dass Matrixmultiplikation linear ist. $\square$

Wir brauchen das jetzt noch nicht so viel, aber es ist eins der wichtigsten Konzepte der Linearen Algebra. Meine Hoffnung ist, dass Sie sich eher damit anfreunden, wenn Sie es häufiger sehen.

F. Vektoren und Matrizen: Study guide

Am Ende jedes Kapitels sehen Sie einen study guide („Lernleitfaden“) wie diesen, mit Konzeptübersicht und Skills, die zu diesem Kapitel gehören.
Sie sollten alle Konzepte in der Konzeptübersicht verstehen: können Sie die Definition oder Formulierung dieses Konzepts aufschreiben? Können Sie einige Beispiele und Nicht-Beispiele für dieses Konzept angeben?
Die Skills sind Sachen, die sie ausführen können sollten („Fähigkeiten“). Sie müssen sie also üben.
Sie können diese Listen benutzen um zu überprüfen, dass Sie keine wichtigen Sachen aus diesem Kapitel übersehen haben. Entweder beim ersten Lernen oder auch bei der Wiederholung und Klausurvorbereitung.

Konzeptübersicht.

- Vektoren in $\mathbb{R}^2$, $\mathbb{R}^3$ und $\mathbb{R}^n$
- Linearkombination
- Geraden und Ebenen in $\mathbb{R}^2$, $\mathbb{R}^3$ und $\mathbb{R}^n$
- (Informell) „Freiheitsgrad“, zum Beispiel von Gerade, Ebenen
- Kollineare Vektoren
- Vektorraumeigenschaften für $\mathbb{R}^n$
- Matrix, Größe einer Matrix, quadratische Matrix
- Matrixmultiplikation
- Matrix als Funktion: Matrixabbildung
- Diagonale Matrizen, obere und untere Dreiecksmatrizen
- Transponierte Matrix
- Linearität der Matrixmulitplikation.

Skills.

- Vektoren addieren, Vektor mit einem Skalar multiplizieren, Linearkombinationen von Vektoren bilden
- Bestimmen, ob eine gegebene Menge eine Gerade ist oder eine Ebene ist, und ob sie durch 0 geht oder nicht
- Bestimmen, ob zwei Vektoren kollinear sind
- Matrizen addieren, eine Matrix mit einem Skalar multiplizieren, Linearkombinationen von Matrizen bilden
- Matrix transponieren
- Matrix mal Vektor multiplizieren
- Matrizen passender Größen multiplizieren
- Bestimmen, ob zwei gegebene Matrizen multipliziert werden können oder nicht
- Matrixprodukt transponieren

KAPITEL 2

Vektorräume

Der Vektorraum ist das Fundament der linearen Algebra.

A. Vektorraumaxiome und Beispiele

Wir haben in Kapitel 1 gelernt, dass die Menge $\mathbb{R}^n$ aller Spaltenvektoren mit n reellen Einträgen bestimmte Bedingungen erfüllt, die sie zu einem *Vektorraum* machen. Aber die Menge aller $m \times n$-Matrizen erfüllt diese Bedingungen auch! Da steckt also noch mehr dahinter.
Wir schauen uns diese Definition jetzt etwas abstrakter an und lernen noch andere Beispiele kennen.

Definition 2.1: Ein **Vektorraum** über einem Körper K ist eine Menge V mit zwei Operationen $+ \colon V \times V \longrightarrow V$, Addition, und $\cdot \colon K \times V \longrightarrow V$, Skalarmultiplikation, die folgende Bedingungen erfüllen: (es sind jeweils u, v, $w \in V$ und λ, $\mu \in K$)

VA0 $\forall u, v$: $u + v \in V$ (Abgeschlossenheit unter Addition)
VA1 Es existiert ein Nullvektor $0 \in V$, sodass $\forall v$: $v + 0 = v = 0 + v$. (Nullvektor)
VA2 $\forall v$ existiert $-v \in V$ mit $v + (-v) = 0 = (-v) + v$. (Negative)
VA3 $\forall u, v, w$: $(u + v) + w = u + (v + w)$ (Assoziativität der Addition)
VA4 $\forall u, v$: $u + v = v + u$ (Kommutativität der Addition)
SM0 $\forall v, \lambda$: $\lambda v \in V$ (Abgeschlossenheit unter Skalarmultiplikation)
SM1 $\forall v$: $1 \cdot v = v$ (Einheitsskalar)
SM2 $\forall v, \lambda, \mu$: $\lambda \cdot (\mu v) = (\lambda \cdot \mu) v$ (Assoziativität der Skalarmultiplikation)
SM3 $\forall v, \lambda, \mu$: $(\lambda + \mu)\, v = \lambda v + \mu v$ (Distributivität der Skalarmult über Addition im Körper)
SM4 $\forall u, v, \lambda$: $\lambda(u + v) = \lambda u + \lambda v$ (Distributivität der Skalarmult über Addition in V)

Ein **Vektor** ist als Element eines Vektorraums definiert.

Bemerkung 2.2: Die Axiome VA0-4 sagen, dass V mit Addition eine kommutative Gruppe bildet. Die 1 in SM1 ist das neutrale Element der Multiplikation im Körper K. Es kommen in SM2 und SM3 jeweils Operationen im Körper K und im Vektorraum V vor: der Kontext macht klar, welche gemeint sein muss.
Wir sehen, dass wir über jedem Körper einen Vektorraum haben können. In diesem Kurs konzentrieren wir uns auf Vektorräume über $\mathbb{R}$, auch **reelle Vektorräume** genannt.

Beispiele 2.3: ◇ Wir haben gesehen, dass für jede natürliche Zahl n die Menge $\mathbb{R}^n$ ein Vektorraum ist, bei dem Addition und Skalarmultiplikation eintragsweise definiert sind.
◇ Wir haben gesehen, dass die Menge $\mathcal{M}_{m,n}$ der $m \times n$-Matrizen ein Vektorraum ist, wieder mit eintragsweise definierten Operationen.

Nicht-Beispiele 2.4: ◇ Die Menge der positiven reellen Zahlen mit überlicher Addition und mit üblicher reeller Multiplikation als Skalarmultiplikation ist kein Vektorraum: Diese Menge ist zwar abgeschlossen unter Addition, aber nicht unter Skalarmultiplikation. Zum Beispiel ist 1 in der Menge, aber nicht $(-1) \cdot 1$.
◇ Ebenso ist eine Halbebene in $\mathbb{R}^2$ auch kein Vektorraum, weil sie nicht unter Skalarmultiplikation abgeschossen ist.

Beispiele 2.5: (Polynome)

- Sei $P_n = \mathbb{R}[x]_n$ die Menge der Polynome von Grad höchstens n mit reellen Koeffizienten. Also ist $a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n$ ein generisches Element aus $\mathbb{R}[x]_n$, mit Koeffizienten $a_i \in \mathbb{R}$. (Da die Koeffizienten auch 0 sein können, schließt dies Polynome von kleinerem Grad mit ein.)
  Seien $p = a_0 + a_1x + \cdots + a_nx^n$ und $q = b_0 + b_1x + \cdots + b_nx^n$ zwei Elemente von $\mathbb{R}[x]_n$. Dann können wir sie addieren:
  $$p + q = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$
  und erhalten wieder ein Polynom von Grad höchstens n. Also ist VA0 erfüllt: $p+q \in \mathbb{R}[x]_n$.
  Das Nullpolynom $0 = 0 + 0x + 0x^2 + \cdots + 0x^n$ erfüllt $0 + p = 0 = p + 0$ für jedes Polynom $p \in \mathbb{R}[x]_n$. Also ist VA1 erfüllt.
  Wenn p ein Polynom ist, dann ist $-p$, jeweils mit den negativen Koeffizienten, auch ein Polynom vom selben Grad, und $p + (-p) = (-p) + p = 0$. Also ist VA2 erfüllt.
  Man kann leicht zeigen, dass auch VA3 und VA4 erfüllt sind.
  Sei p ein Polynom wie oben, und $\lambda \in \mathbb{R}$. Dann ist
  $$\lambda p = (\lambda a_0) + (\lambda a_1)x + (\lambda a_2)x^2 + \cdots + (\lambda a_n)x^n$$
  wieder ein Polynom von Grad höchstens n, also ist SM0 erfüllt. Nun sind auch SM1-SM4 leicht zu zeigen. Z.B. SM2:
  $$\begin{aligned}\lambda \cdot (\mu p) &= \lambda \cdot ((\mu a_0) + (\mu a_1)x + (\mu a_2)x^2 + \cdots + (\mu a_n)x^n)\\ &= (\lambda(\mu a_0)) + (\lambda(\mu a_1))x + (\lambda(\mu a_2))x^2 + \cdots + (\lambda(\mu a_n))x^n\\ &= ((\lambda\mu)a_0) + ((\lambda\mu)a_1)x + ((\lambda\mu)a_2)x^2 + \cdots + ((\lambda\mu)a_n)x^n\\ &= (\lambda\mu)p\end{aligned}$$
  Also ist die Menge $\mathbb{R}[x]_n$ von Polynomen vom Grad höchstens n ein Vektorraum.
- Die Menge aller (reellen) Polynome $\mathbb{R}[x]$ (unabhängig vom Grad) ist auch ein Vektorraum.

Nicht-Beispiel 2.6: Wenn wir nur die Polynome von Grad genau n nehmen, bekommen wir keinen Vektorraum: Zum Beispiel ist
$$(3 + 2x + 5x^2) + (2 + x - 5x^2) = 5 + 3x$$
eine Summe von zwei Polynomen von Grad 3, die kleineren Grad hat. Also ist VA0 nicht erfüllt.

Beispiele 2.7: (Abbildungen)

- Sei $\mathrm{Abb}([0,1], \mathbb{R})$ die Menge der Abbildungen vom Intervall $[0,1]$ nach $\mathbb{R}$. Wir können solche Funktionen punktweise addieren: $(f+g)(x) = f(x)+g(x)$. Wir können eine Funktion auch mit einem Skalar multiplizieren: $(\lambda f)(x) = \lambda \cdot f(x)$.
  Erinnern Sie sich, dass in diesem Kurs alle Funktionen auf jedem Element der Ausgangsmenge definiert sind!
  Dann gilt:
  - Die Summe von zwei Funktionen $[0,1] \longrightarrow \mathbb{R}$ ist wieder eine Funktion $[0,1] \longrightarrow \mathbb{R}$, also ist VA0 erfüllt.
  - Die Nullfunktion $f(x) = 0$ für $x \in [0,1]$ erfüllt VA1.
  - Da λf wieder eine Funktion $[0,1] \longrightarrow \mathbb{R}$ ist, ist SM0 erfüllt. Die Wahl $\lambda = -1$ gibt jeweils die negative Funktion, die VA2 erfüllt.
  - VA3 und VA4, und SM1-4, können leicht nachgeprüft werden. Zum Beispiel SM3: Um zu prüfen, dass zwei Funktionen $[0,1] \longrightarrow \mathbb{R}$ gleich sind, müssen wir zeigen, dass sie den selben Funktionswert für jedes $x \in [0,1]$ haben:
    $$((\lambda + \mu)f)(x) = (\lambda + \mu) \cdot f(x) = \lambda \cdot f(x) + \mu \cdot f(x)$$

und
$$(\lambda f + \mu f)(x) = (\lambda f)(x) + (\mu f)(x) = \lambda \cdot f(x) + \mu \cdot f(x)$$
Da x beliebig war, gilt also $(\lambda + \mu)f = \lambda f + \mu f$.

- Sei V die Menge aller unendlichen Folgen, mit eintragsweise Addition und Skalarmultiplikation. Das bedeutet $(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} = (a_i + b_i)_{i\in\mathbb{N}}$, und $\lambda(a_i)_{i\in\mathbb{N}} = (\lambda a_i)_{i\in\mathbb{N}}$. Dies ist auch ein Vektorraum. (Man kann eine Folge als Abbildung von $\mathbb{N}$ nach $\mathbb{R}$ definieren.)

Wir sehen an diesen Beispielen, wie man zeigt, dass eine Menge mit gegebenen Operationen ein Vektorraum ist:

- Identifiziere die Menge V, deren Elemente die Vektoren sind.
- Identifiziere the Operationen der Addition und Skalarmultiplikation. (Also wie sind diese auf dem gegebenen Beispiel V definiert?)
- Prüfe, dass diese Operationen VA0 und SM0 erfüllen. Dann ist V **abgeschlossen unter Addition und Skalarmultiplikation**.
- Identifiziere einen Nullvektor, der VA1 erfüllt, und identifizieren, welche Elemente die Rolle der negativen Vektoren übernehmen, um VA2 zu erfüllen.
- Prüfe alle anderen Axiome.

Übung 2.8: Zeigen Sie, dass die beiden Beispiele „unendliche Folgen“ und „Menge aller Polynome“ Vektorräume bilden.

Wir definieren solche abstrakten Strukturen, damit wir die Axiome benutzen können um Resultate zu beweisen, die dann für *alle* Beispiele eines Vektorraums gelten, und nicht jedes mal neu gezeigt werden müssen. Als Beispiel hierfür zeigen wir, dass der Nullvektor und die Negativen sich in *jedem* Vektorraum so verhalten, wie wir erwarten.
Von hier an ist V immer ein Vektorraum, solange nicht anders angegeben.

Proposition 2.9: (Nullvektor und Negative in Vektorräumen)
Sei V ein Vektorraum. Dann gilt:

(i) Der Nullvektor ist eindeutig.
(ii) Für jedes $v \in V$ gilt $0_{\mathbb{R}} \cdot v = 0_V$: Null mal ein beliebiger Vektor ist der Nullvektor.
(iii) Für jedes $\lambda \in \mathbb{R}$ gilt $\lambda 0_V = 0_V$: jedes Vielfache des Nullvektors ist wieder der Nullvektor.
(iv) Wenn $\lambda v = 0_V$, dann ist entweder $\lambda = 0_{\mathbb{R}} \in \mathbb{R}$, oder $v = 0_V \in V$.
(v) Für jedes $v \in V$ ist das Negative $-v$ eindeutig.
(vi) Für jedes $v \in V$ ist $(-1)v = -v$, das Negative von v.

Sie können auch jeweils $\mathbb{R}$ mit einem beliebigen Körper K ersetzen, wenn Sie wollen.
Wir betrachten diese Beweise jetzt als Beispiele, wie man nur mit Hilfe der Axiome einfache Aussagen beweist.

Beweis. Wir wissen über V nichts außer dass es die Axiome VA0-4 und SM0-4 erfüllt. Also dürfen wir in unseren Beweisen nur diese Axiome benutzen. Deswegen schreiben wir sorgfältig auf, welches Axiom wir wo benutzen.

(i) Angenommen wir haben zwei Nullvektoren 0 und $0'$, welche beide das Axiom VA1 erfüllen. Dann gilt

$$\begin{aligned} 0' &= 0' + 0 && \text{wegen VA1 für } 0 \\ &= 0 && \text{wegen VA1 für } 0' \end{aligned}$$

Vergleiche Eindeutigkeit des neutralen Elements in einer Gruppe, Satz 4.2.5 in MAGL.

(ii) Wir haben

$$\begin{aligned}
v &= 1 \cdot v && \text{wegen SM1, Einheitskalar} \\
&= (1 + 0_{\mathbb{R}}) \cdot v && \text{Rechnung in } \mathbb{R} \\
&= 1 \cdot v + 0_{\mathbb{R}} \cdot v && \text{wegen SM3, Distributivität} \\
&= v + 0_{\mathbb{R}} \cdot v && \text{wegen SM1, Einheitskalar} \\
\Rightarrow \quad (-v) + v &= (-v) + (v + 0_{\mathbb{R}} \cdot v) && \text{plus } -v\text{: existiert wegen VA2} \\
\Rightarrow \quad (-v) + v &= ((-v) + v) + 0_{\mathbb{R}} \cdot v && \text{wegen VA3, Assoziativität} \\
\Rightarrow \quad 0_V &= 0_V + 0_{\mathbb{R}} \cdot v && \text{wegen VA2, Negative} \\
\Rightarrow \quad 0_V &= 0_{\mathbb{R}} \cdot v && \text{wegen VA1, Nullvektor}
\end{aligned}$$

(iii) Übung. Es gibt verschiedene Wege, dies zu beweisen: unter Benutzung von $0_V + 0_V = 0_V$ und der Existenz von Negativen, (also via VA1 und VA2, mit ähnlichen Ideen wie (ii)), direkt mit VA2 auf einem anderen Weg, oder unter Benutzung von (vi).

(iv) Wenn $\lambda = 0_{\mathbb{R}}$, dann sind wir fertig. Angenommen $\lambda \neq 0_{\mathbb{R}}$. Also wollen wir zeigen, dass $v = 0_V$. Da $\lambda \neq 0_{\mathbb{R}}$ haben wir $\frac{1}{\lambda} \in \mathbb{R}$. Also ist

$$\begin{aligned}
v &= 1 \cdot v && \text{wegen SM1, Einheitskalar} \\
&= (\frac{1}{\lambda}\lambda) \cdot v && \text{Rechnung in } \mathbb{R} \\
&= \frac{1}{\lambda}(\lambda v) && \text{wegen SM2} \\
&= \frac{1}{\lambda} \cdot 0_V && \text{nach Annahme: } \lambda v = 0_V \\
&= 0_V && \text{wegen (iii)}
\end{aligned}$$

(v) Dies folgt aus der Eindeutigkeit von Inversen in einer Gruppe, Satz 4.2.8 in MAGL.
Ich lasse es hier noch mal stehen, wenn jemand nicht im anderen Skript nachschauen will.
Angenommen $v \in V$ hat zwei Negative: u und w. Also gilt $v + w = 0_V = w + v$ und $v + u = 0_V = u + v$: sowohl u als auch w erfüllen VA2 für v. Dann gilt

$$\begin{aligned}
u &= u + 0_V && \text{wegen VA1, Nullvektor} \\
&= u + (v + w) && \text{wegen VA2 mit } w \\
&= (u + v) + w && \text{wegen VA3, Assoziativität} \\
&= 0_V + w && \text{wegen VA2 mit } u \\
&= w && \text{wegen VA1, Nullvektor.}
\end{aligned}$$

Also gilt $u = w = -v$: das Negative ist eindeutig.
Dies ist natürlich der selbe Beweis wie in Satz 4.2.8 in MAGL.

(vi) Wir haben

$$\begin{aligned}
v + (-1) \cdot v &= 1 \cdot v + (-1) \cdot v && \text{wegen SM1, Einheitskalar} \\
&= (1 + (-1)) \cdot v && \text{wegen SM3, Distributivität} \\
&= 0_{\mathbb{R}} \cdot v && \text{Rechnung in } \mathbb{R} \\
&= 0_V && \text{wegen (ii)}
\end{aligned}$$

Ebenso folgt $(-1) \cdot v + v = 0_V$. Also erfüllt $(-1) \cdot v$ die Bedingung für das Negative von v, und so folgt aus der Eindeutigkeit der Negativen, dass $(-1)v = -v$. □

Dieses Resultat zeigt uns auch, welches der kleinste Vektorraum ist:

Beispiel 2.10: Der kleinstmögliche Vektorraum ist der **Nullvektorraum** (auch Nullraum): $\{0\}$. In diesem Vektorraum gibt es nur ein Element, den Nullvektor. Die Menge ist abgeschlossen unter Addition, weil $0_V + 0_V = 0_V$ nach VA1. Sie ist auch abgeschlossen unter Skalarmultiplikation, weil $\lambda \cdot 0_V = 0_V$, nach Proposition 2.9(iii). Alle anderen Axiome sind leicht zu prüfen, weil wir nur ein Element haben.

B. Lineare Abbildungen

In der Mathematik ist es immer auch sehr wichtig, zu wissen, wie man von einem Objekt zu einem anderen kommt. Wir brauchen dazu *struktur-erhaltenden Abbildungen*. Wir haben ja schon gesehen, dass Linearkombination ein sehr wichtiger roter Faden ist. Unsere struktur-erhaltenden Abbildungen sollen also *Linearkombinationen erhalten*.

Definition 2.11: Eine Abbildung $f\colon V \longrightarrow W$ zwischen zwei Vektorräumen V, W heißt **linear** wenn

$$\text{für alle } u, v \in V, \lambda, \mu \in \mathbb{R}: \qquad f(\lambda v + \mu u) = \lambda f(v) + \mu f(u)$$

Wir nennen den Vektorraum V **Quelle** (oder **Definitionsraum**) von f, und W **Ziel** von f.

Denken Sie daran, dass in diesem Kurs eine Abbildung auf allen Elementen der Quelle definiert sein muss.

Konsequenz 2.12: ◇ Mit $\mu = 0$ bekommen wir $f(\lambda v) = \lambda f(v)$: wir können einen Skalar „herausziehen“.
- ◇ Mit $\lambda = \mu = 1$ bekommen wir $f(v + u) = f(v) + f(u)$: lineare Abbildungen erhalten Addition.
- ◇ Wir könnten eine lineare Abbildung auch definieren, indem wir diese beiden Eigenschaften fordern. Es macht keinen Unterschied, ob wir separat „Erhalt von Skalarmultiplikation“ und „Erhalt von Addition“, oder in einem zusammen „Erhalt von Linearkombination“ überprüfen.

Beispiele 2.13:

a) **(Matrixabbildungen)** Wie wir in Kapitel 1 gesehen haben, gibt jede $m \times n$-Matrix A eine lineare Abbildung $f_A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ via $v \longmapsto Av$.
   - ◇ Die Quelle von f_A ist $\mathbb{R}^n$: damit die Matrixmultiplikation funktioniert, muss der Vektor v so viele Einträge haben wir A Spalten hat.
   - ◇ Das Ziel von f_A ist $\mathbb{R}^m$: der Vektor Av hat so viele Einträge wie A Zeilen hat.

b) **(Streckung)** Für jedes $r \in \mathbb{R}$ ist die Abbildung $f\colon \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}$ linear:

   mit $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $u = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ haben wir

$$\begin{aligned} f(\lambda v + \mu u) &= r(\lambda v + \mu u) = \lambda r v + \mu r u && \text{nach Vektorraumaxiomen} \\ &= \lambda f(v) + \mu f(v). \end{aligned}$$

   Es gibt zwei besonders wichtige Fälle dieses Beispiels:
   - ◇ für $r = 0$ bekommen wir die Nullabbildung $0\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$, die jeden Vektor auf 0 schickt;
   - ◇ für $r = 1$ bekommen wir die Identitätsabbildung $\mathrm{id}\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$, die jeden Vektor auf sich selbst abbildet.

c) **(Nullabbildung)** Für beliebige Vektorräume V, W gibt es immer eine **Nullabbildung** $0\colon V \longrightarrow W$, die jeden Vektor $v \in V$ auf $0 \in W$ schickt, und diese Abbildung ist linear.

d) **(Identität)** Für jeden Vektorraum V ist die **Identität** $\mathrm{id}_V\colon V \longrightarrow V$ linear.

e) **(Projektion)** Die Abbildung $\pi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit Funktionsvorschrift $\binom{x_1}{x_2} \longmapsto x_1$ ist linear:

$$\begin{aligned} \pi\left(\lambda \tbinom{x_1}{x_2} + \mu \tbinom{y_1}{y_2}\right) &= \pi\left(\begin{pmatrix} \lambda x_1 + \mu y_1 \\ \lambda x_2 + \mu y_2 \end{pmatrix}\right) = \lambda x_1 + \mu y_1 \\ &= \lambda \pi\left(\tbinom{x_1}{x_2}\right) + \mu \pi\left(\tbinom{y_1}{y_2}\right) \end{aligned}$$

Nicht-Beispiele 2.14: ⋄ Die Abbildung $\tau\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ mit $\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \longmapsto \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 5 \\ 3 \end{smallmatrix}\right)$ ist *nicht* linear:

$$\tau\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) + \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} x_1+y_1 \\ x_2+y_2 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 5 \\ 3 \end{smallmatrix}\right)$$

$$\text{aber} \qquad \tau\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)\right) + \tau\left(\left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) + \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 10 \\ 6 \end{smallmatrix}\right).$$

„Translationen sind nicht linear."

⋄ Die Abbildung $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit $\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \longmapsto x_1x_2$ ist *nicht* linear:

$$f\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) + \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)\right) = (x_1+y_1)(x_2+y_2) = x_1x_2 + y_1x_2 + x_1y_2 + y_1y_2$$

$$\text{also z.B.} \qquad f\left(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right) = 4$$

$$\text{aber} \qquad f\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)\right) + f\left(\left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)\right) = x_1x_2 + y_1y_2$$

$$\text{also} \qquad f\left(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right) + f\left(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right) = 2.$$

Beispiele 2.15:

⋄ Die Abbildung $f\colon \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ auf dem Vektorraum aller reellen Polynome mit $f(p) = xp$ ist linear: für Polynome $p, q \in \mathbb{R}[x]$ gilt

$$f(\lambda p + \mu q) = x(\lambda p + \mu q) = \lambda xp + \mu xq = \lambda f(p) + \mu f(q).$$

⋄ Sei V der Vektorraum der unendlich oft differenzierbaren Funktionen auf dem Interval $(0,1)$. Dann ist Differenzieren eine lineare Abbildung $D\colon V \longrightarrow V$:

$$D(\lambda f + \mu g) = (\lambda f + \mu g)' = \lambda f' + \mu g' = \lambda D(f) + \mu D(g).$$

Das wissen Sie vielleicht schon aus der Schule, und es kommt nächstes Jahr in Analysis.

In diesem Beispiel haben wir Funktionen (Polynome oder differenzierbare Funktionen) als Elemente, auf die wir eine weitere Funktion anwenden. Das ist kognitiv nicht für alle leicht: etwas, was für viele erst einmal ein Prozess ist, wird auf einmal zu einem Objekt, auf das man einen weiteren Prozess anwenden kann. Diese Stufe von Prozess zu Objekt ist super wichtig. Schauen Sie sich mal in Lara Alcock's Buch über Mathematikstudium das Kapitel 2 „Abstrakte Objekte", und insbesondere Abschnitt 2.2 „Funktionen als abstrakte Objekte" an. Der Abschnitt 2.1 davor ist bestimmt für Sie als angehende Lehrkräfte auch interessant.

Übung 2.16: Bestimmen Sie in Numbas: ist eine gegebene Abbildung linear oder nicht? Wenn Sie dies noch zu schwierig finden, gibt es eine kleine Aufwärmungsaufgabe: Ausdrücke in Funktionen einsetzen.

Übung 2.17: Zeigen Sie: Wenn f linear ist, dann ist

$$f(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n).$$

Folgende Eigenschaften von linearen Abbildungen sind simpel, aber nützlich.

Proposition 2.18: (Lineare Abbildungen erhalten 0 und Differenzen.)
Wenn $f\colon V \longrightarrow W$ eine lineare Abbildung ist, dann gilt

(i) $f(0) = 0$, *und*
(ii) $f(u - v) = f(u) - f(v)$.

BEWEIS. Übung. Gute „Fingerübung" für Sie. □

Proposition 2.19: (Komposition von linearen Abbildungen ist linear.)
Seien $f\colon U \longrightarrow V$ und $g\colon V \longrightarrow W$ lineare Abbidlungen. Die Komposition $g \circ f\colon U \longrightarrow W$ mit $(g \circ f)(u) = g(f(u))$ ist auch linear.

Beweis.

$$\begin{aligned}
(g\circ f)(\lambda_1 u_1 + \lambda_2 u_2) &= g(f(\lambda_1 u_1 + \lambda_2 u_2)) \\
&= g(\lambda_1 f(u_1) + \lambda_2 f(u_2)) && \text{weil } f \text{ linear} \\
&= \lambda_1 g(f(u_1)) + \lambda_2 g(f(u_2)) && \text{weil } g \text{ linear} \\
&= \lambda_1 (g\circ f)(u_1) + \lambda_2 (g\circ f)(u_2)
\end{aligned}$$

□

Beispiel 2.20: Wenn A eine $m \times k$-Matrix und B eine $k \times n$-Matrix ist, dann ist $f_A \circ f_B = f_{AB}$: die Komposition zweiter Matrixabbildungen ist die Matrixabbildung des Matrixproduktes.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "f_B"] \arrow[dr, "f_{AB}"'] & \mathbb{R}^k \arrow[d, "f_A"] \\
& \mathbb{R}^m
\end{tikzcd}$$

Hier nur eine Erinnerung an die wichtige Eigenschaft der Identitätsabbildung:

Proposition 2.21: (Komposition mit Identität)
Für jede lineare Abbildung $f\colon V \longrightarrow W$ gilt

(i) $f\circ \mathrm{id}_V = f\colon V \longrightarrow W$ *und*

(ii) $\mathrm{id}_W \circ f = f\colon V \longrightarrow W$.

$$\begin{tikzcd}
V \arrow[r, "\mathrm{id}_V"] \arrow[dr, "f"'] & V \arrow[d, "f"] \arrow[dr, "f"] & \\
& W \arrow[r, "\mathrm{id}_W"'] & W
\end{tikzcd}$$

Beweis. Übung. □

Komposition von Abbildungen ist im Allgemeinen assoziativ. (Siehe Satz 4.1.1 in MAGL.) Wenn Sie dies in MAGL noch nicht explizit gezeigt haben, dann zeigen Sie dies jetzt:

Übung 2.22: Für $f\colon X \longrightarrow Y$, $g\colon Y \longrightarrow Z$ und $h\colon Z \longrightarrow W$ gilt $(h\circ g)\circ f = h\circ(g\circ f)\colon X \longrightarrow W$. Zeigen Sie dies, indem Sie bestimmen, was jede der beiden Abbildungen links und rechts mit einem beliebigen Element $a \in X$ macht.

Damit wissen wir dann: Die Menge aller linearen Abbildungen $V \longrightarrow V$ von einem Vektorraum auf sich selbst ist mit Komposition eine Halbgruppe, die ein neutrales Element id_V hat.

C. Unterräume: erste Beispiele

Wir haben in Kapitel 1 verschiedene Beispiele von Geraden und Ebenen innerhalb von $\mathbb{R}^3$ kennengelernt. Ein Beispiel war etwa die x, y-Ebene in $\mathbb{R}^3$. Die sieht doch geometrisch so aus wie $\mathbb{R}^2$. Wir könnten uns also fragen, ob es auch Teilmengen von Vektorräumen gibt, die selbst auch wieder ein Vektorraum sind.

Definition 2.23: Sei V ein Vektorraum. Ein **Unterraum** (auch **Untervektorraum**, UVR) von V ist eine Teilmenge $W \subseteq V$, die auch ein Vektorraum ist, mit derselben Addition und Skalarmultiplikation wie V.

Manche der Eigenschaften bleiben automatisch erfüllt, wenn wir eine Teilmenge anschauen: wenn $u + v = v + u$ für *alle* Vektoren $u, v \in V$ gilt, dann ist diese Eigenschaft auch wahr wenn wir u, v nur aus einer bestimmten Teilmenge von V nehmen.

VA3 Assoziativität der Addition
VA4 Kommutativität der Aaddition
SM1 Einheitsskalar
SM2 Assoziativität der Skalarmultiplikation
SM3 Distributivität von Skalarmult über Addition im Körper
SM4 Distributivität von Skalarmult über Addition in V

werden alle automatisch auf einer Teilmenge erfüllt.
Wenn wir also eine Teilmenge von V haben, müssen wir nur prüfen:

VA0 Wenn wir zwei Vektoren der Teilmenge addieren, ist die Summe wieder in der Teilmenge?
VA1 Ist der Nullvektor in der Teilmenge?

VA2 Wenn ein Vektor in der Teilmenge ist, ist sein negativer Vektor auch in der Teilmenge?
SM0 Ist ein skalares Vielfaches eines Vektors aus der Teilmenge immer noch in der Teilmenge?

Wenn diese Eigenschaften alle wahr sind, dann ist die Teilmenge ein Unterraum.

Wir können sogar noch etwas einsparen:

Proposition 2.24: (UVR Bedingungen)
Eine Teilmenge $U \subseteq V$ eines Vektorrams V ist ein Unterraum genau dann wenn
- $0 \in U$ *(Nullvektor ist enthalten)*
- *für alle $u, w \in U$ ist $u + w \in U$* *(abgeschlossen unter Addition)*
- *für alle $u \in U$ und alle $\lambda \in \mathbb{R}$ ist $\lambda u \in U$* *(abgeschlossen unter Skalarmultiplikation)*

BEWEIS. Wenn wir diese drei Bedingungen prüfen, bekommen wir automatisch auch Negative: $-u = (-1) \cdot u$ nach Proposition 2.9. Die Axiome VA3-4 und SM1-4 werden von V geerbt.
Andersherum, wenn U ein Unterraum ist, dann ist es ein Vektorraum mit derselben Addition und Skalarmultiplikation. Also ist U nach Definition abgeschossen unter Addition und Skalarmultiplikation. Es muss auch einen Nullvektor 0_U geben, der für alle $u \in U$ die Bedingung $0_U + u = u$ erfüllt. Aber wie wissen wir, dass dieser Nullvektor von U gleich dem Nullvektor von V sein muss? Wir wissen aus Prop. 2.9, dass in jedem Vektorraum Null mal ein Vektor den Nullvektor ergibt. Also gilt für $u \in U$, dass $0 \cdot u = 0_U$. Aber u ist auch ein Element von V, also gilt auch $0 \cdot u = 0_V$. Daher muss $0_U = 0_V$ sein: der Nullvektor im Unterraum ist der gleiche wie der in V. □

Um zu sehen, dass es wirklich nicht „automatisch“ ist, dass $0_U = 0_V$, also dass das neutrale Element der Teilmenge das selbe sein muss wie das neutrale Element der größeren Menge: Wir nehmen die Menge der 2×2-Matrizen $\mathcal{M}_{2,2}$ mit Matrixmultiplikation. Das ist eine Halbgruppe. Die Teilmenge aller Matrizen der Form $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ mit $a \in \mathbb{R}$ ist abgeschlossen unter Matrixmultiplikation, und hat $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ als neutrales Element der Matrixmultiplikation. Aber $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist nicht neutral in $\mathcal{M}_{2,2}$.

⚠ VORSICHT: Man könnte meinen, dass man auch die Existenz des Nullvektors durch die anderen Bedingungen erhält, weil $0 \cdot u = 0$. Aber dafür brauchen wir erst mal irgendein $u \in U$: die Teilmenge U darf nicht die leere Menge sein. Denn die leere Menge ist trivialerweise unter den Operationen abgeschlossen (weil es keine Fälle zu prüfen gibt). Aber die leere Menge hat auch keinen Nullvektor, und erfüllt somit nicht die Vektorraumeigenschaften.

Beispiele 2.25: (Unterräume, die es immer gibt) Jeder Vektorraum V hat den Nullvektorraum als Unterraum, und sich selbst:

$$\{0\} \text{ UVR von } V \qquad \text{und} \qquad V \text{ UVR von } V$$

Dass $\{0\}$ UVR ist, folgt aus Eigenschaften des Nullvektors (Prop. 2.9): Die Menge ist abgeschlossen unter Addition, weil $0 + 0 = 0$ nach VA1. Sie ist auch abgeschlossen unter Skalarmultiplikation, weil $\lambda \cdot 0 = 0$, nach Proposition 2.9(iii). Und der Nullvektor ist offensichtlich enthalten.

Beispiele 2.26: (Unterräume von $\mathbb{R}^n$)
- Wie wir in Übung 1.8 gesehen haben, ist eine Gerade durch 0 ein UVR. Wir können uns eine solche Gerade als Kopie der Zahlengerade $\mathbb{R}$ innerhalb von $\mathbb{R}^n$ vorstellen: die Wahl des Parameters t, einer einzigen reellen Zahl, bestimmt einen Punkt auf der Gerade.
- Eine Gerade, die nicht durch 0 geht, ist kein Unterraum: sie erfüllt keine der drei Bedingungen.
- Eine Ebene durch 0 ist ein Unterraum. Eine Ebene, die nicht durch 0 geht, ist kein Unterraum.

- Es gibt auch "größere" Unterräume: $\left\{\begin{pmatrix}x_1\\x_2\\x_3\\0\end{pmatrix} \middle| x_1, x_2, x_3 \in \mathbb{R}\right\}$ ist ein UVR von $\mathbb{R}^4$ der so ähnlich aussieht wie $\mathbb{R}^3$. Ebenso

$$\left\{\begin{pmatrix}x_1\\x_2\\0\\x_4\end{pmatrix} \middle| x_1, x_2, x_4 \in \mathbb{R}\right\} \qquad \text{und} \qquad \left\{\begin{pmatrix}0\\x_2\\x_3\\x_4\end{pmatrix} \middle| x_2, x_3, x_4 \in \mathbb{R}\right\} \qquad \text{usw.}$$

- $\left\{\begin{pmatrix}x_1\\x_2\\x_3\\x_1-x_2\\3x_3\end{pmatrix} \middle| x_1, x_2, x_3 \in \mathbb{R}\right\}$ ist ein UVR von $\mathbb{R}^5$.
  - Der Nullvektor ist von dieser Form: benutze $x_1 = x_2 = x_3 = 0$.
  - Die Summe zweier Vektoren dieser Form ist wieder von dieser Form:
$$\begin{pmatrix}x_1\\x_2\\x_3\\x_1-x_2\\3x_3\end{pmatrix} + \begin{pmatrix}y_1\\y_2\\y_3\\y_1-y_2\\3y_3\end{pmatrix} = \begin{pmatrix}x_1+y_1\\x_2+y_2\\x_3+y_3\\(x_1+y_1)-(x_2+y_2)\\3(x_3+y_3)\end{pmatrix}$$
  - Ein skalares Vielfaches erhält die Form:
$$\lambda\begin{pmatrix}x_1\\x_2\\x_3\\x_1-x_2\\3x_3\end{pmatrix} = \begin{pmatrix}\lambda x_1\\\lambda x_2\\\lambda x_3\\(\lambda x_1)-(\lambda x_2)\\3(\lambda x_3)\end{pmatrix}$$
- Die einzigen Unterräume von $\mathbb{R}$ (als Vektorraum $\mathbb{R}^1$) sind $\{0\}$ und $\mathbb{R}$: Wenn ein UVR $W \subseteq \mathbb{R}$ ein Element $w \in W$, $w \neq 0$ enthält, dann bekommen wir (wegen Abgeschlossenheit unter Skalarmultiplikation) auch $\frac{1}{w} \cdot w = 1 \in W$. Und dann ist für jedes $x \in \mathbb{R}$ auch $x \cdot 1 \in W$, also ist $W = \mathbb{R}$.

Übung 2.27: Zeigen Sie, dass die ersten Beispiele wirklich Unterräume sind.

Fakt 2.28: Dies sind alle möglichen Untervektorräume von $\mathbb{R}^2$:

- Nur der Nullvektor $\{\binom{0}{0}\}$.
- Jede Gerade durch den Ursprung.
- Ganz $\mathbb{R}^2$.

Dies sind alle möglichen Untervektorräume von $\mathbb{R}^3$:

- Nur der Nullvektor $\left\{\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$.
- Jede Gerade durch den Ursprung.
- Jede Ebene durch den Ursprung.
- Ganz $\mathbb{R}^3$.

Wir beweisen dies jetzt nicht, aber wir lernen im Kurs Werkzeuge, um dies beweisen zu können.

Beispiele 2.29: (UVR von Polynomräumen)
- ◇ Polynome von Grad $\leqslant 2$ sind ein UVR der Polynome von Grad $\leqslant 3$: $\mathbb{R}[x]_2$ UVR von $\mathbb{R}[x]_3$.
- ◇ Alle Polynome von Grad $\leqslant n$ mit konstantem Term 0 bilden einen UVR von $\mathbb{R}[x]_n$.

Übung 2.30: Zeigen Sie, dass die Beispiele in 2.29 wirklich Unterräume sind.

Beispiele 2.31: (UVR von Matrizenräumen)
- ◇ Die Menge der diagonalen $n \times n$-Matrizen bilden einen UVR von $\mathcal{M}_{n,n}$.
  - Die Nullmatrix ist diagonal.
  - Die Summe von zwei Diagonalmatrizen ist wieder diagonal: Wenn sowohl $A_{ij} = 0$ und $B_{ij} = 0$ für $i \neq j$, dann ist auch $(A + B)_{ij} = A_{ij} + B_{ij} = 0$.
  - Das Vielfache einer Diagonalmatrix ist wieder diagonal: Wenn $A_{ij} = 0$ für $i \neq j$, dann auch $(\lambda A)_{ij} = \lambda \cdot A_{ij} = 0$.
- ◇ Die Menge der oberen Dreiecksmatrizen der Größe $n \times n$ bilden einen UVR von $\mathcal{M}_{n,n}$.
  - Die Nullmatrix ist eine obere Dreiecksmatrix.
  - Die Summe von zwei oberen Dreiecksmatrizen ist wieder eine obere Dreiecksmatrix: Argument wie oben, nur für $i > j$ statt für $i \neq j$.
  - Das Vielfache einer oberen Dreiecksmatrix ist wieder eine obere Dreiecksmatrix.
- ◇ Ebenso bilden die unteren Dreiecksmatrizen der Größe $n \times n$ einen UVR von $\mathcal{M}_{n,n}$.

Hier sind noch ein paar mehr:

Definition 2.32: Eine quadratische Matrix A heißt **symmetrisch** wenn $A = A^T$. Sie heißt **anti-symmetrisch** oder **schief-symmetrisch** wenn $A = -A^T$.

Beispiele 2.33: Folgende Matrizen sind symmetrisch:

◇ $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ ◇ $\begin{pmatrix} -23 & 19 \\ 19 & 37 \end{pmatrix}$ ◇ $\begin{pmatrix} 3 & 4 & 5 \\ 4 & 8 & 2 \\ 5 & 2 & 9 \end{pmatrix}$

Folgende Matrizen sind anti-symmetrisch:

◇ $\begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$ ◇ $\begin{pmatrix} 0 & -19 \\ 19 & 0 \end{pmatrix}$ ◇ $\begin{pmatrix} 0 & 4 & 5 \\ -4 & 0 & -2 \\ -5 & 2 & 0 \end{pmatrix}$

Eine Matrix A ist symmetrisch genau dann wenn $a_{ij} = a_{ji}$ für alle i, j. Und sie ist anti-symmetrisch genau dann wenn $a_{ij} = -a_{ij}$.
Die Diagonaleinträge einer symmetrischen Matrix sind beliebig, aber die Diagonaleintrge einer anti-symmetrischen Matrix müssen Null sein: $a_{ii} = -a_{ii} \Rightarrow 2a_{ii} = 0 \Rightarrow a_{ii} = 0$, weil wir in $\mathbb{R}$ durch 2 teilen können.

Beispiel 2.34: Die Menge Sym_n aller symmetrischen $n \times n$-Matrizen bilden einen UVR von $\mathcal{M}_{n,n}$.
- Die Nullmatrix 0 erfüllt $0^T = 0$, ist also symmetrisch.
- Wenn A, B symmetrisch sind, dann ist $(A + B)^T = A^T + B^T = A + B$, also ist die Summe symmetrisch.
- Wenn A symmetrisch ist und $\lambda \in \mathbb{R}$, dann ist $(\lambda A)^T = \lambda A^T = \lambda A$, also ist dies symmetrisch.

Übung 2.35: Zeigen Sie, dass die Menge ASym_n aller anti-symmetrischen $n \times n$-Matrizen einen UVR von $\mathcal{M}_{n,n}$ bilden.

D. Unterräume: strukturelle Beispiele

Auch lineare Abbildungen geben uns noch mehr Unterräume.

Definition 2.36: Der **Kern** einer linearen Abbildung $f\colon V \longrightarrow W$ ist

$$\operatorname{Ker} f = \{v \in V \mid f(v) = 0\} \qquad \text{alles, was auf 0 geschickt wird}$$

und das **Bild** von f ist

$$\operatorname{Im} f = \{w \in W \mid \exists v \in V \text{ mit } f(v) = w\} \qquad \text{alles in } W \text{ was von } f \text{ getroffen wird}$$

Beispiele 2.37:

a) Eine Matrixabbildung $f_A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ hat $\operatorname{Ker} f_A = \{x \in \mathbb{R}^n \mid Ax = 0\}$. Wir werden in Kapitel 3 lernen, wie man so einen Kern einer Matrixabbildung findet.
Das Bild ist $\operatorname{Im} f_A = \{b \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n : Ax = b\}$. Auch hier werden wir später lernen, wie man dies genauer bestimmt.

b) Wir betrachten die Projektion $\pi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$, die $\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$ auf x_1 schickt (aus Beispiel 2.13).

$$\operatorname{Ker} \pi = \{\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \mid x_1 = 0\} = \left\{\left(\begin{smallmatrix} 0 \\ x_2 \end{smallmatrix}\right) \mid x_2 \in \mathbb{R}\right\}$$
$$\operatorname{Im} \pi = \mathbb{R}$$

weil wir für jedes $x \in \mathbb{R}$ zum Beispiel $\left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right) \in \mathbb{R}^2$ mit $\pi(\left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right)) = x$ haben.

c) Die Abbildung $f\colon \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$, die p auf $f(p) = xp$ schickt, hat

$$\operatorname{Ker} f = \{0\} \qquad \text{— nur 0 landet auf 0}$$
$$\operatorname{Im} f = \{\text{Polynome mit konstantem Term } 0\}.$$

d) Differenzieren $D\colon V \longrightarrow V$, wobei V der Raum der unendlich oft differenzierbaren Funktionen auf $(0,1)$ ist, hat

$$\operatorname{Ker} D = \{\text{konstante Funktionen}\}$$
$$\operatorname{Im} D = V.$$

Übung 2.38: Bestimmen Sie Kern und Bild der Nullabbildung $0\colon V \longrightarrow W$ und der Identität auf V.

Proposition 2.39: (Kern und Bild sind Unterräume.)
Sei $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann ist $\operatorname{Ker} f$ *ein Unterraum der Quelle V und* $\operatorname{Im} f$ *ein Unterraum des Ziels W.*

BEWEIS. Wir prüfen die UVR Bedingungen aus Prop. 2.24.

- ⋄ $0 \in \operatorname{Ker} f$, da $f(0) = 0$ nach Proposition 2.18.
- ⋄ Wenn $u, v \in \operatorname{Ker} f$, dann ist $f(u+v) = f(u) + f(v) = 0 + 0 = 0$. Also auch $u + v \in \operatorname{Ker} f$.
- ⋄ Wenn $v \in \operatorname{Ker} f$ und $\lambda \in \mathbb{R}$, dann ist $f(\lambda v) = \lambda f(v) = \lambda \cdot 0 = 0$. Also auch $\lambda v \in \operatorname{Ker} f$.

Also ist $\operatorname{Ker} f$ ein Unterraum.

Übung: Prüfen Sie die UVR Bedingungen für $\operatorname{Im} f$.

□

Wenn wir zwei Unterräume $U \subseteq V$ und $W \subseteq V$ haben, können wir deren Schnitt und deren Vereinigung (als Mengen) bilden. Wir bekommen aber nicht in allen Fällen wieder einen UVR!

Proposition 2.40: (Schnitt und Vereinigung von Unteräumen)
Sei V ein Vektorraum, und U, W UVR von V. Dann gilt:

(i) Der Schnitt $U \cap W$ ist wieder ein UVR von V.
(ii) Die Vereinigung $U \cup W$ ist ein UVR V genau dann wenn $U \subseteq W$ oder $W \subseteq U$.

Beweis. Übung. □

Dies gilt auch für andere algebraische Strukturen, z.B. Untergruppen und andere, die Sie noch kennenlernen werden.
Die zweite Aussage zeigt uns, dass Vereinigungen fast nie Unterräume sind. Die richtige Weise, zwei Unterräume zu „verbinden", ist nicht mit Vereinigung wie Mengen, sondern mit Summe.

Definition 2.41: Seien U, W zwei Unterräume des Vektorraums V. Die **Summe** $U + W$ der Unterräume ist die Menge der Vektoren

$$U + W = \{u + w \mid u \in U,\ w \in W\}.$$

Proposition 2.42: (Summe von Unterräumen)
Seien U, W Unterräume eines Vektorraums V. Dann ist die Summe $U + W$ ein weiterer UVR von V, der sowohl U als auch W enthält: $U \subseteq U + W$ und $W \subseteq U + W$.

Beweis. Wir prüfen die UVR Bedingungen aus Prop. 2.24.

- ◇ $0 = 0 + 0$, mit $0 \in U$ und $0 \in W$, also ist $0 \in U + W$.
- ◇ Wenn $u_1 + w_1$ und $u_2 + w_2$ zwei Elemente von $U + W$ sind, dann ist deren Summe $(u_1 + u_2) + (w_1 + w_2)$, mit $u_1 + u_2 \in U$ und $w_1 + w_2 \in W$, weil beide UVR unter Addition abgeschlossen sind. Also ist die Summe der Elemente auch in $U + W$.
- ◇ Wenn $u + w \in U + W$ und $\lambda \in \mathbb{R}$, dann ist $\lambda(u + w) = \lambda u + \lambda w$, mit $\lambda u \in U$ und $\lambda w \in W$, weil beide UVR unter Skalarmultiplikation abgeschlossen sind. Also ist $\lambda(u + w) \in U + W$.

Daher ist $U + W$ UVR von V.
Für beliebiges $u \in U$ haben wir $u = u + 0 \in U + W$, weil $0 \in W$. Genauso ist für beliebiges $w \in W$ das Element $w = 0 + w \in U + W$. Also enthält $U + W$ sowohl U als auch W. □

Beispiele 2.43: ◇ Seien $U = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \middle| \ x \in \mathbb{R} \right\}$ und $W = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \middle| \ y \in \mathbb{R} \right\}$ UVR von $\mathbb{R}^2$. Dann ist $U + W = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} \middle| \ x,\ y \in \mathbb{R} \right\} = \mathbb{R}^2$.

Dies zeigt den Unterschied zwischen Vereinigung und Summe: die Vereinigung der x-Achse und der y-Achse ist nur die Menge der beiden Achsen, aber die *Summe* der x-Achse und der y-Achse ergibt die ganze Ebene. Das ist genau das, was wir brauchen, um Unterräume zu kombinieren.

Die Summe von zwei Unterräumen ist der kleinste Unterraum, der beide enthält.

◇ Sei $U = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \middle| \ x,\ y \in \mathbb{R} \right\}$ und $W = \left\{ \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} \middle| \ y, z \in \mathbb{R} \right\}$. Dann ist $U + W = \mathbb{R}^3$, aber es gibt nicht-triviale Überschneidung: $U \cap W = \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \middle| \ y \in \mathbb{R} \right\}$.

Wir sehen: wenn $U \cap W$ nicht der Nullvektorraum ist, dann gibt es etwas „Überflüssigkeit" in der Summe. Manchmal ist es gut, zwischen diesen Situationen unterscheiden zu können.

Definition 2.44: Für Unterräume U und W von V nennen wir die Summe $U + W$ eine **direkte Summe**, und schreiben $U \oplus W$, wenn $U \cap W = 0$.

Beispiele 2.45: In den obigen Beispielen ist das erste Beispiel eine direkte Summe, aber das zweite nicht.

E. Spann und Erzeugendensysteme

Wir können die UVR Bedingungen (Prop. 2.24) auffassen als: ein UVR ist eine nicht-leere Teilmenge, die unter Linearkombinationen abgeschlossen ist. Wir erwarten also, dass die Menge aller Linearkombinationen bestimmter Vektoren einen Unterraum geben müsste.

Definition 2.46: Sei S eine nicht-leere Menge von Vektoren im Vektorraum V. Der **Spann von** S (auch Aufspann oder lineare Hülle), geschrieben

$$\operatorname{Span}(S) = \langle S \rangle$$

ist die Menge aller Linearkombinationen der Vektoren aus S.

$$\langle S \rangle = \{\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_r w_r \mid w_1, \ldots, w_r \in S, \lambda_1, \ldots, \lambda_r \in \mathbb{R}\}$$

Der Spann der leeren Menge ist der Nullvektorraum:

$$\operatorname{Span}(\varnothing) = \{0\}.$$

Hier enthält eine Linearkombination aus unendlich vielen Vektoren nur endlich viele von Null verschiedene Koeffizienten.

Notation 2.47: Wenn $S = \{w_1,\ w_2,\ \ldots,\ w_r\}$ eine endliche Menge ist, schreiben wir auch

$$\operatorname{Span}(S) = \operatorname{Span}(w_1,\ w_2,\ \ldots,\ w_r) = \langle w_1,\ w_2,\ \ldots,\ w_r \rangle .$$

Proposition 2.48: (Spann ist UVR)
Der Spann einer nicht-leeren Teilmenge S von V ist ein Unterraum von V.
Zudem ist dies der kleinste Unterraum von V, der alle Elemente von S enthält.

Der zweite Satz bedeutet: wenn W ein weitere UVR von V mit $S \subset W$ ist, dann ist $\operatorname{Span}(S) \subseteq W$.

Beweis. Es gibt mindestens ein $w \in S$, weil S nicht-leer ist. Seien auch $w_i \in S$, für $i = 1, \ldots, r$.

- $0 \in \operatorname{Span}(S)$ weil $0 = 0 \cdot w$ eine Linearkombination von Vektoren aus S ist.
- Für zwei Linearkombinationen u und v aus Elementen aus S ist deren Summe auch eine Linearkombinaton von Vektoren aus S: in jeder Linearkombination kommen nur endlich viele Elemente aus S vor, also kommen auch nur endlich viele Elemente $w_1, \ldots, w_r \in S$ in beiden Linearkombinationen zusammen vor. Wir können also schreiben: $u = \lambda_1 w_1 + \cdots + \lambda_r w_r$ und $v = \mu_1 w_1 + \cdots + \mu_r w_r$ für $\lambda_i, \mu_i \in \mathbb{R}$ (wobei manche der Koeffizienten Null sein können). Dann ist
$$u + v = (\lambda_1 + \mu_1) w_1 + (\lambda_2 + \mu_2) w_2 + \cdots + (\lambda_r + \mu_r) w_r .$$
Also haben wir $u,\ v \in \operatorname{Span}(S) \Rightarrow u + v \in \operatorname{Span}(S)$.
- Für $u = \lambda_1 w_1 + \cdots + \lambda_r w_r \in \operatorname{Span}(S)$ und einen Skalar $\mu \in \mathbb{R}$ haben wir
$$\mu u = (\mu\lambda_1) w_1 + (\mu\lambda_2) w_2 + \cdots + (\mu\lambda_r) w_r \in \operatorname{Span}(S).$$

Also ist $\operatorname{Span}(S)$ UVR von V.
Für die zweite Aussage erinnnern wir uns daran, dass jeder Unterraum von V abgeschlossen unter Linearkombinationen ist. Wenn also alle Elemente von S in dem Unterraum sind, dann müssen auch alle Linearkombinationen dieser Elemente im Unterraum sein.

Wenn also W ein UVR ist mit $S \subseteq W$, dann ist für eine beliebige Linearkombination $u \in \mathrm{Span}(S)$ auch $u \in W$. Daraus folgt $\mathrm{Span}(S) \subseteq W$, also ist der Spann der kleinste UVR, der alle Elemente von S enthält. □

Übung 2.49: Zeigen Sie, dass die Summe $U + W$ von zwei Unterräumen U und W von V gleich dem Spann der Vereinigung ist: $U + W = \mathrm{Span}(U \cup W)$.
Dies beweist dann unseren Slogan: *Die Summe von zwei Unterräumen ist der kleinste Unterraum, der beide enthält.*

Manchmal ist der Spann einer Teilmenge ganz V.

Definition 2.50: Wir sagen, dass eine Menge S den Vektorraum V **aufspannt** oder **erzeugt**, wenn $\mathrm{Span}(S) = V$. Wenn S den Raum V erzeugt, nennen wir S auch ein **Erzeugendensystem** von V.

Wenn also S den Raum V erzeugt, dann kann jeder Vektor $v \in V$ (in mindestens einer Weise) als Linearkombination von Vektoren aus S geschrieben werden.

Beispiele 2.51: ◇ Die Menge $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ erzeugt $\mathbb{R}^2$: $\begin{pmatrix}x\\y\end{pmatrix} = x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix}$ also ist jeder Vektor in $\mathbb{R}^2$ eine Linearkombination dieser beiden Vektoren.

◇ Die Menge $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$ erzeugt auch $\mathbb{R}^2$: die Linearkombinationen müssen nicht eindeutig sein.

$$\begin{aligned}\begin{pmatrix}x\\y\end{pmatrix} &= x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix} + 0\begin{pmatrix}1\\1\end{pmatrix}\\ &= 0\begin{pmatrix}1\\0\end{pmatrix} + (y-x)\begin{pmatrix}0\\1\end{pmatrix} + x\begin{pmatrix}1\\1\end{pmatrix}\\ &= (x-y)\begin{pmatrix}1\\0\end{pmatrix} + 0\begin{pmatrix}0\\1\end{pmatrix} + y\begin{pmatrix}1\\1\end{pmatrix}\\ &= (-y)\begin{pmatrix}1\\0\end{pmatrix} + (-x)\begin{pmatrix}0\\1\end{pmatrix} + (x+y)\begin{pmatrix}1\\1\end{pmatrix}\\ &= \dots\end{aligned}$$

◇ Die Menge $\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$ erzeugt $\mathbb{R}^3$:

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = \left(\frac{x+y}{2} - z\right)\begin{pmatrix}1\\1\\0\end{pmatrix} + \left(\frac{y-x}{2}\right)\begin{pmatrix}-1\\1\\0\end{pmatrix} + z\begin{pmatrix}1\\1\\1\end{pmatrix}$$

Wir lernen bald, wie man diese Koeffizienten findet.

◇ Die Vektoren $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}$ erzeugen die x, y-Ebene in $\mathbb{R}^3$. Aber die Vektoren $\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix}$ erzeugen die selbe Ebene. Also haben wieder

$$\left\langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \right\rangle = \left\langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix} \right\rangle.$$

◇ In $\mathbb{R}^n$, seien $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$, also ist e_k der Vektor mit Nulleinträgen überall, bis auf eine 1 im Eintrag k. Dann erzeugen $e_1, \dots, e_n$ den Raum $\mathbb{R}^n$.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_{n-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$
$$= x_1 e_1 + x_2 e_2 + \dots + x_{n-1} e_{n-1} + x_n e_n$$

Diese Vektoren $e_1, \dots, e_n$ sind super nützlich und werden uns immer wieder begegnen. Wir nennen Sie die **Standardbasisvektoren**.

◇ Die Monome $\{1, x, x^2, \dots, x^n\}$ erzeugen den Polynomraum $\mathbb{R}[x]_n$: jedes Polynom ist einer Linearkombination dieser Monome.
$$p = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \cdot 1$$

Wir haben in den Beispielen gesehen, dass verschiedene Mengen den selben Unterraum erzeugen können.

Proposition 2.52: (Gleicher Spann)
Für zwei nicht-leere Teilmengen S und S' eines Vektorraums V gilt $\operatorname{Span}(S) = \operatorname{Span}(S')$ genau dann wenn jeder Vektor aus S eine Linearkombination von Vektoren aus S' ist, und jeder Vektor aus S' eine Linearkombination von Vektoren aus S ist.

BEWEIS. Übung. □

Wir könnten uns jetzt fragen: Wie finden wir raus, ob ein bestimmter Vektor im Spann einer bestimmten Teilmenge liegt? Dazu müssen wir ein lineares Gleichungssystem lösen. Und das lernen wir im nächsten Kapitel.

F. Vektorräume: Study guide

Konzeptübersicht.

◇ Vektorraum, Vektorraumaxiome
◇ Vektorraum von Matrizen
◇ Vektorraum von Polynomen
◇ Eigenschaften von Nullvektor und Negativen in einem Vektorraum
◇ Lineare Abbildungen, Quelle, Ziel
◇ Repertoire von und Intuition über lineare Abbildungen
◇ Komposition von linearen Abbildungen
◇ Nullabbildung, Identitätsabbildung
◇ Unterraum, UVR Bedingungen
◇ Unterräume von $\mathbb{R}^n$, besonders für $n = 2, 3$
◇ Symmetrische und anti-symmetrische Matrizen
◇ Verschiedene Matrixunterräume
◇ Schnitt und Vereinigung von Unterräumen
◇ Kern, Bild einer linearen Abbildung
◇ Summe von Unterräumen
◇ Direkte Summe
◇ Spann, Erzeugendensystem
◇ Spann ist UVR

Skills.

- ◇ Bestimmen, ob (und beweisen, dass) eine gegebene Menge mit Operationen ein Vektorraum ist.
- ◇ Einfache Resultate unter Verwendung von Vektorraumaxiomen beweisen.
- ◇ Bestimmen, ob (und beweisen, dass) eine Abbildung linear ist.
- ◇ Bestimmen, ob (und beweisen, dass) eine gegebene Teilmenge in UVR ist oder nicht.
- ◇ Schnitt und Summe von (einfachen) Unterräumen bilden.

(Skills zu den restlichen Konzepten folgen später.)

KAPITEL 3

Lineare Gleichungssysteme

Wir brauchen lineare Gleichungssysteme (LGS) an vielen Stellen in unserem Kurs: zum Beispiel, um den Kern einer Matrixabbildung zu finden, oder um herauszufinden, ob ein bestimmter Vektor im Spann von anderen Vektoren liegt, aber auch für vieles andere. Deswegen ist dieses Kapitel ein sehr wichtiges Werkzeug, auch über diesen Kurs hinaus.
Lineare Gleichungssysteme sind auch eine zweite Interpretation von Matrizen. Wir werden in diesem Kapitel lernen, was ein lineares Gleichungssystem ist, wie man es mit Hilfe von Matrizen lösen kann, und wie viele Lösungen bestimmte Systeme haben können.

A. Lineare Gleichungssysteme

Wir haben gesehen, dass wir Geraden in verschiedenen Schreibweisen aufschreiben können: in Koordinatenform $y = mx + b$, oder in Parameterform $\{b + tu \mid t \in \mathbb{R}\}$.
Die Koordinatenform $y = mx + b$ lässt allerdings keine senkrechten Geraden zu, weshalb dies eine bessere Koordinatenform ist: $a_1x + a_2y = b$. Wenn $a_2 \neq 0$, bekommen wir daraus sehr leicht die aus der Schule übliche Form. Aber mit $a_2 = 0$ bekommen wir auch senkrechte Geraden $x = \frac{b}{a_1}$.
Diese Koordinatenform ist genau eine lineare Gleichung. Und wir werden sehen, dass die Parameterform $\{b + t \cdot u \mid t \in \mathbb{R}\}$ genau einer Lösungsmenge solch einer linearen Gleichung entspricht.

Definition 3.1: Ein **lineares Gleichungssystem (LGS) in n Unbekannten (oder Variablen)** ist ein System von Gleichungen der Form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad & \;\;\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Auf der rechten Seite jeder Gleichung steht eine Linearkombination der Variablen, hier $x_1, \ldots, x_n$, und auf der rechten Seite jeder Gleichung steht eine Konstante.
Ein solches System heißt **homogen** wenn alle b_i auf der rechten Seite gleich Null sind. Ansonsten ist es **inhomogen**.

Eine **Lösung eines LGS** ist ein Vektor $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, dessen Einträge alle Gleichungen gleichzeitig lösen.

Nicht-Beispiele 3.2: Folgende Gleichungen sind *nicht* linear:
- $x^2 + 2y = 0$, $xy = 1$: Produkte der Variablen sind nicht erlaubt.
- $2x - \sqrt{y} = 1$: Wurzeln von Variablen sind nicht erlaubt.
- $\sin(x) - \cos(y) = 1$: Funktionen wie Sinus, Cosinus, Logarithmus usw. sind nicht linear.
- $e^x = 2$: Unbekannten dürfen nicht im Exponenten stehen.

Wir interessieren uns für Lösungen solcher linearen Gleichungssysteme. Fangen wir mit einer einzelnen Gleichung an.

Beispiel 3.3: Die lineare Gleichung $3x - 4y = 0$ hat viele Lösungen: wir können y frei wählen, und dadurch wird x festgelegt. Zum Beispiel $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ ist eine Lösung. Jede andere Lösung ist ein Vielfaches davon: $\begin{pmatrix} 4t \\ 3t \end{pmatrix}$, oder $t\begin{pmatrix} 4 \\ 3 \end{pmatrix}$. Also ist die Lösungsmenge eine Gerade $\left\{ t\begin{pmatrix} 4 \\ 3 \end{pmatrix} \middle| t \in \mathbb{R} \right\}$. Dies ist eine Gerade durch den Ursprung.
Wenn wir dieselbe Gleichung nehmen und inhomogen machen, haben wir etwas ähnliches. $3x - 4y = 2$ hat Lösungsmenge $\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} + t\begin{pmatrix} 4 \\ 3 \end{pmatrix} \middle| t \in \mathbb{R} \right\}$. Diese Gerade geht nicht durch den Ursprung, aber ist parallel zur Geraden, welche die Lösungsmenge der homogenen Gleichung ist.
$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ist eine **spezielle Lösung** dieser inhomogenen Gleichung, aber da $3 \cdot 4 - 4 \cdot 3 = 0$, können wir Vielfache von $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ zu dieser Lösung addieren und erhalten wieder eine Lösung:

$$3(2 + t \cdot 4) - 4(1 + t \cdot 3) = (3 \cdot 2 - 4 \cdot 1) + t(3 \cdot 4 - 4 \cdot 3) = 2 + t \cdot 0.$$

Da wir zwei Variablen und nur eine Gleichung haben, haben wir „einen Freiheitsgrad“: wir können eine der beiden Variablen wählen, und die andere wird dadurch bestimmt. Dies entspricht einer Gerade.

Beispiel 3.4: Die Lösungsmenge einer linearen Gleichung mit drei Variablen ist eine Ebene: wir haben jetzt „zwei Freiheitsgrade“.
$2x_1 - 3x_2 + 6x_3 = 0$ hat Lösungsmenge $\left\{ \begin{pmatrix} \frac{3}{2}s - 3t \\ s \\ t \end{pmatrix} \middle| s, t \in \mathbb{R} \right\} = \left\{ s \cdot \begin{pmatrix} \frac{3}{2} \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}$: wir können zwei der Variablen frei wählen, $x_2 = s$ und $x_3 = t$, und dadurch wird dann x_1 bestimmt.
Wir sehen, dass dies eine Ebene ist: wir können s und t unabhängig voneinander wählen, und dies gibt uns zwei Richtungsvektoren, die nicht kollinear sind.

Bei einem System mit mehreren Gleichungen ist die Lösungmenge dann die Schnittmenge der Lösungen der einzelnen Gleichungen.

Beispiele 3.5: ◇ Wir betrachten das LGS

$$\begin{aligned} 3x - 4y &= 2 \\ 2x - y &= 0 \end{aligned}$$

Die beiden Gleichungen repräsentieren zwei Geraden:

$$\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} + t\begin{pmatrix} 4 \\ 3 \end{pmatrix} \right\} \qquad \text{und} \qquad \left\{ s\begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

Eine Lösung dieses Systems ist also der Schnittpunkt der beiden Geraden. In der Schule haben wir schon Lösungswege gelernt: Die zweite Gleichung gibt $y = 2x$, und dies in die erste Gleichung eingesetzt gibt $3x - 8x = 2$, oder $x = -\frac{2}{5}$. Dann ist $y = -\frac{4}{5}$. Also ist $\begin{pmatrix} -\frac{2}{5} \\ -\frac{4}{5} \end{pmatrix}$ die Lösung des LGS. Die Lösung ist eindeutig.

- Dieses LGS hat keine Lösungen:

$$\begin{aligned} 3x - 4y &= 0 \\ 6x - 8y &= 4 \end{aligned}$$

Die zwei Geraden sind parallel, haben also keinen Schnittpunkt.
- Das LGS

$$\begin{aligned} 3x - 4y &= 2 \\ 6x - 8y &= 4 \end{aligned}$$

hat eine ganze Gerade als Lösungsmenge: die zweite Gleichugn ist ein Vielfaches der ersten Gleichung, also repräsentieren beide Gleichungen die selbe Gerade

$$\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} 4 \\ 3 \end{pmatrix} \right\},$$

und dies ist die Lösungsmenge des LGS.

Wir werden später untersuchen, wann ein LGS eine eindeutige Lösung, keine Lösung oder unendlich viele Lösungen hat.

Definition 3.6: Ein LGS heißt **konsistent** oder **lösbar** wenn es (mindestens) eine Lösung hat. Ein LGS ohne Lösungen heißt **inkonsistent** oder **unlösbar**.

B. Schreibweisen für lineare Gleichungssysteme

Wie wir schon in der Einführung gesehen haben, können wir LGS auf verschiedene Weisen aufschreiben. Wir haben schon in der Definition die Form als Gleichungen gesehen:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n-1}x_{n-1} + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn-1}x_{n-1} + a_{mn}x_n &= b_m \end{aligned}$$

Notation 3.7: Ein LGS kann als Linearkombination der Koeffizienten notiert werden:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Diese Schreibweise brauchen wir, wenn wir das Verhältnis verschiedener Vektoren zueinander untersuchen wollen. Zum Beispiel, ob ein bestimmter Vektor im Spann von anderen Vektoren liegt.
Diese Vektorgleichung enthält genau die selbe Information wie die Gleichungsform des LGS: wenn wir auf der linken Seite die Vektoren zu einem Vektor aufaddieren, haben wir in der Vektorgleichung genau pro Eintrag eine der Gleichungen des Systems.

Notation 3.8: Ein LGS kann in **Matrixform** als

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

notiert werden.

Hier sind die Koeffizienten in einer Matrix zusammengefasst. Wenn A diese Matrix ist, können wir x für den Vektor der Variablen und b für den Vektor der Einträge auf der rechten Seite schreiben. So bekommen wir eine schöne kompakte Schreibweise:

$$Ax = b.$$

Diese Schreibweise hilft uns, wenn wir Eigenschaften von Matrizen und von Abbildungen untersuchen wollen, wie etwa Kern und Bild.
Wenn wir die Matrixmultiplikation auf der linken Seite ausführen, haben wir wieder eine Vektorgleichung, in der jeder Eintrag eine der Gleichungen des LGS darstellt.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n-1}x_{n-1} + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn-1}x_{n-1} + a_{mn}x_n \end{pmatrix}$$

Wir sehen, dass die Matrix so viele Spalten hat wie es Variablen oder Unbekannte gibt, und so viele Zeilen wie es Gleichungen gibt.
Wir wollen in diesem Kapitel einen systematischen Lösungsalgorithmus lernen, um LGS mit Hilfe von Matrizen zu Lösen.
Wir lernen dazu noch eine Kurzschreibweise für ein LGS: Da die Variablennamen selbst nicht wichtig sind (die Lösung ändert sich nicht, wenn die Variablen y_i statt x_i heißen), sind nur die *Anzahl* der Variablen, die Koeffizienten a_{ij} und die Konstanten b_i auf der rechten Seite nötig, um das LGS zu kennen und zu lösen.

Definition 3.9: Die **erweiterte Koeffizientenmatrix** eines LGS ist eine Matrix der Form

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} & b_m \end{array}\right).$$

Dies ist eine andere Schreibweise als die Matrixschreibweise $Ax = b$, denn es ist nicht mehr explizit eine Gleichung. ⚠ Dies ist wirklich „nur“ eine Kurzschreibweise. Der senkrechte Strich erinnert uns an die Position des Gleichheitszeichens.

Konvention 3.10: Bei einem homogenen LGS $Ax = 0$ lassen wir in der erweiterten Koeffizientenmatrix oft die Erweiterung durch Nullen weg:

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} & 0 \end{array}\right) \text{ wird kurz als } \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} \end{pmatrix}$$

geschrieben.

Beispiele 3.11: Hier sind einige LGS und deren erweiterte Koeffizientenmatrizen.

$$\begin{aligned} 3x_1 - x_2 &= 0 \\ 2x_1 - 4x_2 &= 0 \\ -5x_1 + 6x_2 &= 0 \end{aligned} \qquad \begin{pmatrix} 3 & -1 \\ 2 & -4 \\ -5 & 6 \end{pmatrix}$$

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ -x_1 - x_2 + x_3 &= 1 \end{aligned} \qquad \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ -1 & -1 & 1 & 1 \end{array}\right)$$

$$\begin{aligned} 3x_1 - x_2 + 4x_3 \qquad\quad &= 2 \\ -x_1 - 3x_2 \qquad + \sqrt{2}x_4 &= 0 \\ \pi\, x_2 - 3x_3 \qquad\quad &= -5 \end{aligned} \qquad \left(\begin{array}{cccc|c} 3 & -1 & 4 & 0 & 2 \\ -1 & -3 & 0 & \sqrt{2} & 0 \\ 0 & \pi & -3 & 0 & -5 \end{array}\right)$$

Welches LGS ist am leichtesten zu lösen? Eins wo die Gleichungen schon die Lösungen geben:

$$\begin{array}{cccccc} x_1 & & & & = & b_1 \\ & x_2 & & & = & b_2 \\ & & \ddots & & \vdots & \vdots \\ & & & x_n & = & b_n \end{array} \qquad \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_1 \\ 0 & 1 & \cdots & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_n \end{array}\right)$$

In diesem System haben wir genausoviele Gleichungen wie Unbekannte. Also ist die Matrix A (von $Ax = b$) eine quadratische Matrix.
Das nächst-bestlösbare System ist eins wie dieses:

$$\begin{array}{ccccccccc} x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ & & x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & x_n & = & b_n \end{array} \qquad \left(\begin{array}{cccc|c} 1 & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & 1 & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_n \end{array}\right)$$

Wir erhalten x_n mit der letzten Gleichung, substitutieren dies in die vorletzte Gleichung und erhalten x_{n-1}, und so weiter nach oben.
Unser Ziel ist es, die linke Seite der erweiterten Koeffizientenmatrix eines LGS in so eine obere Dreicksform zu überführen, natürlich in einer Weise, die die Lösungsmenge nicht ändert.
Wir können sogar noch eine etwas nützlichere Form finden. Leider können nicht alle Systeme in die schöne Diagonalform überführt werden, aber wir werden sehen, dass wir so nah wie möglich daran kommen werden.

C. Elementare Zeilenumformungen

Es gibt drei recht einfache Typen von Zeilenoperationen, die wir auf ein LGS anwenden können, ohne die Lösungsmenge zu verändern. Es ist egal, ob wir die Operationen auf die Gleichungen oder auf die Zeilen der erweiterten Koeffizientenmatrix ausfüren.

Definition 3.12: Die drei Typen von **elementaren Zeilenumformungen** sind

1. Multipliziere eine Zeile mit einem von Null verschiedenen Skalar.
2. Tausche zwei Zeilen.
3. Addiere ein Vielfaches einer Zeile auf eine andere.

Es ist klar, dass alle diese Umformungen rückgängig gemacht werden können, und dass sie die Lösungsmenge des LGS nicht ändern.

Beispiel 3.13: Wir wenden die elementaren Zeilenumformungen auf das LGS in Gleichungsform und auf die erweiterte Koeffizientenmatrix an.

$$\begin{aligned} 7x_1 + 3x_2 &= 3 \\ x_1 + x_2 &= 1 \end{aligned} \qquad \left(\begin{array}{cc|c} 7 & 3 & 3 \\ 1 & 1 & 1 \end{array}\right)$$

Tausche Zeile 1 und Zeile 2:

$$\begin{aligned} x_1 + x_2 &= 1 \\ 7x_1 + 3x_2 &= 3 \end{aligned} \qquad \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 7 & 3 & 3 \end{array}\right)$$

Addiere -7 mal Zeile 1 zu Zeile 2:

$$\begin{aligned} x_1 + x_2 &= 1 \\ 0 - 4x_2 &= -4 \end{aligned} \qquad \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -4 & -4 \end{array}\right)$$

Multiplizieren Zeile 2 mal $-\frac{1}{4}$ (oder teile sie durch -4):

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_2 &= 1 \end{aligned} \qquad \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 1 \end{array}\right)$$

Addiere -1 mal Zeile 2 zu Zeile 1:

$$\begin{aligned} x_1 &= 0 \\ x_2 &= 1 \end{aligned} \qquad \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right)$$

Dieses LGS hat also eine eindeutige Lösung $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, die wir am Ende direkt ablesen können.

Wir sehen hier, dass die Benutzung der erweiterten Koeffizientenmatrix statt der Gleichungen viel übersichtlicher und effizienter ist: man muss weniger schreiben und hat nur die nötigen Informationen.

In Lineare Algebra können Sie oft Ihre Lösungen selbst überprüfen. Zum Beispiel, setzen Sie die gerade erhaltene Lösung wieder in die ursprünglichen Gleichungen ein:

$$\begin{aligned} 7 \cdot 0 + 3 \cdot 1 &= 3 \\ 0 + 1 &= 1 \end{aligned}$$

stimmt.

Überzeugen Sie sich davon, dass alle diese elemntaren Zeilenumformungen umgekehrt werden können! (Wie?)

⚠ Vorsicht! Stellen Sie sicher, dass Sie nie eine Zeile mit 0 multiplizieren! Dies löscht alle Informationen in dieser Zeile.

Übung 3.14: Zeigen Sie, dass zwei LGS, die durch elementare Zeilenumformungen ineinander überführt werden können, die gleiche Lösungsmenge haben.

D. Gauss Algorithmus

Wir lernen jetzt einen Algorithmus, der uns sagt, welche elementare Zeilenumformung am besten sind, um ein LGS in eine Form zu überführen, in der wir die Lösung leicht ablesen können.

Eine diagonale Form wäre am schönsten, aber das ist nicht immer möglich. Besonders, wenn wir gar keine diagonale Matrix haben.

Beispiel 3.15: Das LGS

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ -x_1 - x_2 + x_3 &= 1 \end{aligned} \qquad \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ -1 & -1 & 1 & 1 \end{array}\right)$$

hat nicht genug Gleichungen, um in eine Diagonalform überführt zu werden.

Die nächstbeste Form ist

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ -1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{\text{II+I}} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{array}\right) \xrightarrow{\text{I}-2\text{II}} \left(\begin{array}{ccc|c} 1 & 0 & -3 & -4 \\ 0 & 1 & 2 & 3 \end{array}\right)$$

korrespondierend zu

$$\begin{aligned} x_1 - 3x_3 &= -4 \\ x_2 + 2x_3 &= 3 \end{aligned}$$

Wir können dann x_3 frei wählen (zum Beispiel als 1, oder als Parameter t), und dann davon abhängend leicht x_1 und x_2 ablesen. Die Lösungsmenge ist dann

$$\left\{ \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}.$$

Wir haben hier notiert, welche elementaren Zeilenoperationen wir durchgeführt haben: die römischen Ziffern stehen für Zeilen.

Definition 3.16: Eine Matrix ist in **Zeilenstufenform (ZSF)** wenn folgende Bedingungen erfüllt sind.

- ◇ Alle Nullzeilen stehen gruppiert am Ende der Matrix.
- ◇ In jeder Zeile, die keine Nullzeile ist, ist der erste von Null verschiedene Eintrag eine 1. Wir nennen dies eine **führende Eins**.
- ◇ In zwei aufeinanderfolgenden, von Null verschiedenen Zeilen steht die führende Eins in der unteren Zeile weiter rechts als die führende Eins in der Zeile darüber.

Wenn zusätzlich über jeder führenden Eins nur Nullen stehen, dann ist die Matrix in **reduzierter Zeilenstufenform**.

Bemerkung 3.17: In einigen Quellen wird bei ZSF nicht verlangt, dass der erste von Null verschiedene Zeileneintrag eine 1 ist. Dieser Eintrag heißt dann Hauptelement oder Pivotelement. Dass der Eintrag 1 ist, wird dann erst bei der reduzierten ZSF verlangt. Es ist aber für unseren Algorithmus hilfreich, wenn wir diese Bedingung schon bei ZSF einführen.

Die dritte Bedingung von ZSF bedeutet, dass unter jeder führenden Eins nur Nullen stehen. Wir könnten also sagen, dass eine Matrix in ZSF eine obere Dreiecksmatrix (oder eine rechteckige Version davon) ist. Aber die führenden Einsen müssen nicht unbedingt auf der Diagonalen stehen (oder, bei rechteckigen Matrizen, in den Einträgen a_{ii}).

Eine Matrix in ZSF sieht so aus:

$$\begin{pmatrix} 1 & * & * & \cdots & * & * & * & \cdots \\ 0 & 1 & * & \cdots & * & * & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \end{pmatrix}$$

Hier steht $*$ für einen beliebigen Eintrag. Eine Marix in *reduzierter* ZSF sieht so aus, mit Nullen über den führenden Einsen:

$$\begin{pmatrix} 1 & \color{red}{0} & * & \cdots & * & \color{red}{0} & * & \cdots \\ 0 & 1 & * & \cdots & * & \color{red}{0} & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \end{pmatrix}$$

Bei einer erweiterten Koeffizientenmatrix betrachten wir die Matrix auf der linken Seite des Strichs, um zu entscheiden, ob sie in (reduzierter) ZSF ist.

Beispiele 3.18: Folgende Matrizen sind in ZSF, aber nicht in reduzierter ZSF.

◇ $\begin{pmatrix} 1 & 8 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$. ◇ $\left(\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$ ◇ $\begin{pmatrix} 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ ◇ $\left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 1 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right)$

Übung 3.19: Benutzen Sie elementare Zeilenumformungen, um die obigen Matrizen in reduzierte ZSF zu überführen.

Konvention 3.20: Wir sagen A **hat reduzierte ZSF** B, wenn A durch elementare Zeilenumformungen in die reduzierte Zeilenstufenform B überführt werden kann.

Beispiele 3.21: Folgende Matrizen sind in reduzierter ZSF.

◇ $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right)$ ◇ $\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right)$ ◇ $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ ◇ $\left(\begin{array}{ccc|c} 1 & -5 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$

Übung 3.22: Bestimmen Sie, ob die oben gegebenen LGS lösbar sind oder nicht. Für jedes lösbare System, geben Sie die Lösungsmenge an.

Gauss-Jordan Algorithmus 3.23: Der Gauss-Jordan Algorithmus überführt eine (erweiterte) Matrix in reduzierte ZSF, mit den folgenden Schritten:

- ◇ Wähle die erste Spalte von links, in der ein von Null verschiedener Eintrag steht.
- ◇ Ist der oberste Eintrag der gewählten Spalte eine Null, so tausche die erste mit einer anderen Zeile, in der in dieser Spalte keine Null steht.
- ◇ Teile die erste Zeile durch den nun obersten Eintrag der gewählten Spalte, um eine führende Eins zu erhalten.
- ◇ Addiere geeignete Vielfache der obersten Zeile zu den darunterliegenden Zeilen, sodass die Einträge unter der führenden Eins zu Null werden.
- ◇ Verdecke nun die oberste Zeile der Matrix und wiederhole die obigen Stufen mit der Restmatrix. Wiederhole so lange, bis die Matrix in Zeilenstufenform ist.
- ◇ Von der ZSF ausgehend, arbeite von hinten, um durch geeignete Additionen von Zeilen auch über den führenden Einsen Nullen zu erzeugen.

Bemerkung 3.24: Der Teil bis ZSF ist der Gauss-Algorithmus, und wenn man dann auch noch bis reduzierte ZSF geht, ist es Gauss-Jordan.
Ich empfehle, *immer* bis zur reduzierten ZSF zu gehen, weil man dort die Lösung sehr viel leichter ablesen kann als bei ZSF.

Informell gesagt macht der Gauss-Jordan Algorithmus folgendes: er vereinfacht die Gleichungen, eliminiert überflüssige Gleichungen (die er in Nullzeilen überführt), und bringt den Rest in eine Form, in der man die Lösungen leicht ablesen kann.
Wir sehen dies am besten an Beispielen:

Beispiele 3.25: Wir benutzen Gauss-Jordan für die folgenden Beispiele.
Sie können auch eine noch ausführlichere Erklärung in meinem Video *Gauss-Jordan-Algorithmus anwenden* anschauen. Dort sehen Sie auch, wie Sie den Gauss-Jordan-Algorithmus Rechner benutzen können, um die verschiedenen möglichen Schritte zu erkunden und den Algorithmus besser zu verstehen, ohne zu viel selbst rechnen zu müssen. (Also mehr verstehen als rechnen).

- Wir wählen die erste Spalte. Dort steht schon eine 1 an oberster Stelle. Wir erzeugen also Nullen unter dieser führenden Eins. Wir machen hier Kästchen um führende Einsen, um sie hervorzuheben.

$$\left(\begin{array}{ccc|c} \boxed{1} & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right) \xrightarrow{\text{III}+4\text{I}} \left(\begin{array}{ccc|c} \boxed{1} & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right)$$

  Nun überdecken wir die erste Zeile und fahren mit der Restmatrix fort. Wir wählen die zweite Spalte. Dort steht eine 2 an der (neuen) obersten Stelle, also teilen wir die zweite Zeile durch 2. Danach erzeugen wir eine Null im Eintrag unter der führenden Eins.

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right) \xrightarrow{\frac{1}{2}\text{II}} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & \boxed{1} & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array}\right) \xrightarrow{\text{III}+3\text{II}} \left(\begin{array}{ccc|c} \boxed{1} & -2 & 1 & 0 \\ 0 & \boxed{1} & -4 & 4 \\ 0 & 0 & \boxed{1} & 3 \end{array}\right)$$

  Weil der Eintrag in der dritten Zeile und dritten Zeile zufällig schon eine 1 ist, ist die Matrix jetzt in ZSF.
  Nun erzeugen wir Nullen *über* den führenden Einsen, angefangen mit der dritten Spalte.

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & \boxed{1} & 3 \end{array}\right) \xrightarrow{\text{II}+4\text{III},\text{I}-\text{III}} \left(\begin{array}{ccc|c} 1 & -2 & 0 & -3 \\ 0 & \boxed{1} & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right) \xrightarrow{\text{I}+2\text{II}} \left(\begin{array}{ccc|c} \boxed{1} & 0 & 0 & 29 \\ 0 & \boxed{1} & 0 & 16 \\ 0 & 0 & \boxed{1} & 3 \end{array}\right)$$

  Wir lesen die eindeutige Lösung des Systems ab: $\begin{pmatrix} 29 \\ 16 \\ 3 \end{pmatrix}$.

- Wir wählen wieder die erste Spalte. Dieses Mal ist der oberste Eintrag 0, also müssen wir Zeilen tauschen. Danach können wir durch Division in der ersten Zeile die führende Eins bekommen. Anschließend erzeugen wir Nullen unter der führenden Eins.

$$\left(\begin{array}{ccc|c} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{array}\right) \xrightarrow{\text{I}\leftrightarrow\text{II}} \left(\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{array}\right) \xrightarrow{\frac{1}{2}\text{I}} \left(\begin{array}{ccc|c} \boxed{1} & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{array}\right) \xrightarrow{\text{III}-5\text{I}} \left(\begin{array}{ccc|c} \boxed{1} & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & 1 & -4 & 8 \\ 0 & -\frac{1}{2} & 2 & -\frac{3}{2} \end{array}\right)$$

  Nun wiederholen wir den Algorithmus für die beiden untersten Zeilen: Wir wählen die zweite Spalte, die in der zweiten Zeile (der neue oberste Eintrag) schon eine führende Eins hat. Wir müssen also nur darunter Nullen erzeugen.

$$\left(\begin{array}{ccc|c} 1 & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \boxed{1} & -4 & 8 \\ 0 & -\frac{1}{2} & 2 & -\frac{3}{2} \end{array}\right) \xrightarrow{\text{III}+\frac{1}{2}\text{II}} \left(\begin{array}{ccc|c} \boxed{1} & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \boxed{1} & -4 & 8 \\ 0 & 0 & 0 & \frac{5}{2} \end{array}\right)$$

  Die dritte Zeile (bis zum Trennstrich) ist eine Nullzeile. Diese steht am Ende der Matrix. Somit ist die Matrix in ZSF.
  Wir sehen hier schon an dieser Stelle, dass das LGS unlösbar ist: in der letzten Zeile steht $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = \frac{5}{2}$, was natürlich unmöglich ist. In diesem Fall kann man sich die letzten Schritte auch sparen.

- Statt dessen ändern wir das Beispiel ein wenig ab: wir fangen mit folgender ZSF an und erzeugen eine Null über der zweiten führenden Eins:

$$\left(\begin{array}{ccc|c} 1 & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \boxed{1} & -4 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{\mathrm{I}+\frac{3}{2}\mathrm{II}} \left(\begin{array}{ccc|c} \boxed{1} & 0 & -5 & \frac{25}{2} \\ 0 & \boxed{1} & -4 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Dies ist nun in reduzierte ZSF, und es gibt eine Spalte ohne führende Einsen, hier in grün. Solch eine Spalte ohne führende Eins bedeutet, dass die zugehörige Variable, hier x_3, einen beliebigen Wert annehmen kann. Dies nennt man eine **freie Variable**. Wir machen sie zu einem Parameter $x_3 = t$.

Wie können wir jetzt die Lösung ablesen? Zur Veranschaulichung übertragen wir das LGS wieder in Gleichungen, damit wir sehen, was hier passiert.

$$\begin{aligned} x_1 - 5x_3 &= \frac{25}{2} \\ x_2 - 4x_3 &= 8 \end{aligned}$$

Wir setzen $x_3 = t$ und bringen es auf die rechte Seite, und so erhalten wir

$$\begin{aligned} x_1 &= \frac{25}{2} + 5t \\ x_2 &= 8 + 4t \end{aligned}$$

Also hat das LGS die Lösungsmenge $\left\{ \begin{pmatrix} \frac{25}{2} \\ 8 \\ 0 \end{pmatrix} + t \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} \middle| \; t \in \mathbb{R} \right\}$.

- Hier ist noch ein Beispiel mit zwei freien Variablen in der Lösung. Dies ist ein homogenes LGS.

$$\begin{pmatrix} \boxed{1} & 2 & 6 & 8 \\ 2 & -1 & 2 & -4 \\ 3 & 1 & 8 & 4 \\ 1 & -3 & -4 & -12 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 2 & 6 & 8 \\ 0 & -5 & -10 & -20 \\ 0 & -5 & -10 & -20 \\ 0 & -5 & -10 & -20 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 6 & 8 \\ 0 & -5 & -10 & -20 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 6 & 8 \\ 0 & \boxed{1} & 2 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 2 & 0 \\ 0 & \boxed{1} & 2 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Wir haben nun zwei Spalten ohne führende Eins, hier eine grün (dritte Spalte) und eine lila (vierte Spalte) markiert. Wir können nun $x_3 = s$ und $x_4 = t$ wählen. Also ist eine allgemeine Lösung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2s \\ -2s - 4t \\ s \\ t \end{pmatrix}$$

und somit ist die Lösungsmenge:

$$\left\{ s \begin{pmatrix} -2 \\ -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ -4 \\ 0 \\ 1 \end{pmatrix} \middle| \; s, t \in \mathbb{R} \right\}.$$

Mehr Zwischenschritte beim Ablesen der Lösung von der reduzierten ZSF sind mit anderen Beispielen im Video *Von reduzierter Zeilenstufenform zu Lösung* auf OPAL (unter Erklärungsvideos) erklärt.

Beispiel 3.26: (Beispielaufgabe mit Lösung: LGS mit Parameter)
Video LGS mit Parameter
Für welche a hat das folgende LGS 0, 1 oder unendliche viele Lösungen? Für jeden lösbaren Fall, finden Sie die Lösungsmenge.

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 7 \\ 2x_2 - 4x_3 &= 8 \\ x_1 + 5x_2 + (a^2-7)x_3 &= a + 14 \end{aligned}$$

Wie üblich schreiben wir das LGS als erweiterte Koeffizientenmatrix und fangen mit dem Gauss-Jordan Algorithmus an.

$$\left(\begin{array}{ccc|c} \boxed{1} & 3 & -2 & 7 \\ 0 & 2 & -4 & 8 \\ 1 & 5 & a^2-7 & a+14 \end{array}\right) \xrightarrow{\text{III}-\text{I}} \left(\begin{array}{ccc|c} \boxed{1} & 3 & -2 & 7 \\ 0 & 2 & -4 & 8 \\ 0 & 2 & a^2-5 & a+7 \end{array}\right) \xrightarrow{\frac{1}{2}\text{II}} \left(\begin{array}{ccc|c} 1 & 3 & -2 & 7 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 2 & a^2-5 & a+7 \end{array}\right)$$

$$\xrightarrow{\text{III}-2\text{II}} \left(\begin{array}{ccc|c} 1 & 3 & -2 & 7 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & a^2-1 & a-1 \end{array}\right) = \left(\begin{array}{ccc|c} \boxed{1} & 3 & -2 & 7 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & (a-1)(a+1) & a-1 \end{array}\right)$$

An dieser Stelle müssen wir verschiedene Fälle betrachten: wann erhalten wir eine führende Eins in der dritten Zeile?

- Fall $a = -1$: In diesem Fall haben wir

$$\left(\begin{array}{ccc|c} 1 & 3 & -2 & 7 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 0 & -2 \end{array}\right)$$

  und sehen, dass dieses LGS unlösbar ist, weil die dritte Zeile eine Nullzeile mit von Null verschiedener Erweiterung ist. Also haben wir hier 0 Lösungen.
- Fall $a = 1$: Nun ist die dritte Zeile inklusive Erweiterung eine Nullzeile. Also können wir Gauss-Jordan zu Ende führen.

$$\left(\begin{array}{ccc|c} 1 & 3 & -2 & 7 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{\text{I}-3\text{II}} \left(\begin{array}{ccc|c} \boxed{1} & 0 & 4 & -5 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

  In diesem Fall bekommen wir die Lösungsmenge

$$\left\{ \begin{pmatrix} -5 \\ 4 \\ 0 \end{pmatrix} + t \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}.$$

  Somit haben wir in diesem Fall unendlich viele Lösungen.
- Fall $a \neq \pm 1$: Nun können wir durch $a^2 - 1$ teilen, um eine führende Eins zu bekommen:

$$\left(\begin{array}{ccc|c} 1 & 3 & -2 & 7 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & a^2-1 & a-1 \end{array}\right) \xrightarrow{\frac{1}{a^2-1}\text{III}} \left(\begin{array}{ccc|c} 1 & 3 & -2 & 7 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & \boxed{1} & \frac{1}{a+1} \end{array}\right)$$

$$\xrightarrow{\text{II}+2\text{III},\text{I}+2\text{III}} \left(\begin{array}{ccc|c} 1 & 3 & 0 & 7+\frac{2}{a+1} \\ 0 & \boxed{1} & 0 & 4+\frac{2}{a+1} \\ 0 & 0 & 1 & \frac{1}{a+1} \end{array}\right) \xrightarrow{\text{I}-3\text{II}} \left(\begin{array}{ccc|c} \boxed{1} & 0 & 0 & \frac{7a+9}{a+1} - 3\frac{4a+6}{a+1} \\ 0 & \boxed{1} & 0 & \frac{4a+6}{a+1} \\ 0 & 0 & \boxed{1} & \frac{1}{a+1} \end{array}\right)$$

  Nun haben wir die eindeutige Lösung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{-5a-9}{a+1} \\ \frac{4a+6}{a+1} \\ \frac{1}{a+1} \end{pmatrix} = \frac{1}{a+1} \begin{pmatrix} -(5a+9) \\ 4a+6 \\ 1 \end{pmatrix}.$$

  ⚠ Vorsicht! Hier ist a ein Parameter des Systems, nicht der Lösung. Für jedes verschiedene a ist es ein anderes LGS, und dieses LGS hat eine eindeutige Lösung.

E. Anzahl von Lösungen eines LGS

Nun können wir mit dem Gauss-Jordan Algorithmus ein LGS lösen. Also denken wir jetzt in der Theorie darüber nach, wie viele Lösungen ein solches System haben kann, und wie die Struktur der Lösungsmenge aussieht.

Proposition 3.27: (Lösungen eines homogenen LGS)
Ein homogenes LGS hat immer mindestens eine Lösung: den Nullvektor.
Wenn ein homogenes LGS mehr als eine Lösung hat, dann hat es unendlich viele Lösungen.

BEWEIS. Angenommen unser LGS hat n Unbekannte $x_1, \ldots, x_n$ und m Gleichungen. Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Dann ist das System $Ax = 0$ mit einer $m \times n$-Matrix A.
Da $A \cdot 0 = 0$, ist also der Nullvektor $x = 0$ immer eine Lösung.
In folgenden Situationen ist der Nullvektor die *einzige* Lösung: Wenn die reduzierte ZSF von A genau n Spalten mit führenden Einsen hat (also genau n von Null verschiedene Zeilen, und danach vielleicht noch Nullzeilen).

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

Jede andere reduzierte ZSF eines LGS mit n Unbekannten kann nur weniger Spalten mit führenden Einsen haben, da es genau n Spalten gibt. Jede Spalte ohne führende Eins führt zu einer freien Variable in der Lösungsmenge, die unendliche viele Werte annehmen kann und somit zu unendlich vielen Lösungen führt.

$$\begin{pmatrix} 1 & 0 & * & \cdots & * & 0 & * & \cdots \\ 0 & 1 & * & \cdots & * & 0 & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \end{pmatrix}$$

□

Die Lösung 0 für ein homogenes LGS nennen wir auch die **triviale Lösung**.

Proposition 3.28: *Die Lösungsmenge eines homogenen LGS ist ein Unterraum von $\mathbb{R}^n$.*

BEWEIS. Sei $Ax = 0$ das LGS, mit $x \in \mathbb{R}^n$ und A von Größe $m \times n$. Dann ist die Lösungsmenge genau der Kern der Matrixabbildung $f_A \colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Da jeder Kern ein UVR der Quelle ist (Prop. 2.39), ist die Lösungsmenge ein Unterraum von $\mathbb{R}^n$. □

Konvention 3.29: Wir bezeichnen den Kern der Matrixabbildung f_A auch manchmal einfach als **Kern der Matrix** A:

$$\operatorname{Ker} A = \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

Proposition 3.30: (Struktur der Lösungsmenge eines inhomogenen LGS)
Wenn man zu einer speziellen Lösung eines inhomogenen LGS eine beliebige Lösung des dazugehörigen homogenen LGS addiert, erhält man einer weitere Lösung des inhomogenen LGS.

Beweis. Das inhomogene LGS sei $Ax = b$, mit $b \neq 0$. Sei a eine spezielle Lösung, also $Aa = b$, und sei $v \in \operatorname{Ker} A$ eine beliebige Lösung des dazugehörigen homogenen LGS $Ax = 0$. Dann ist

$$A(a+v) = Aa + Av = b + 0 = b,$$

also ist $a + v$ eine Lösung des inhomogenen Systems. □

Übung 3.31: Zeigen Sie: sind a und a' zwei Lösungen eines inhomogenen LGS $Ax = b$, so ist $a - a' \in \operatorname{Ker} A$, also eine Lösung des homogenen Systems $Ax = 0$.

Geometrisch gesehen ist die Lösungsmenge eines inhomogenen LGS also eine spezielle Lösung plus den Kern von A (ein Unterraum von $\mathbb{R}^n$). Zum Beispiel eine Gerade, Ebene, ..., die nicht durch den Ursprung geht (sondern parallel verschoben wurde).

Proposition 3.32: (Anzahl von Lösungen eines inhomogenen LGS)
Ein inhomogenes LGS kann null, eine oder unendlich viele Lösungen haben.

Beweis. Wenn das inhomogene LGS unlösbar ist, dann hat es null Lösungen. Dies entspricht dem Fall, dass in der ZSF des LGS eine Nullzeile mit von Null verschiedener Erweiterung vorkommt.

$$\left(\begin{array}{cccc|c} 1 & * & \cdots & * & b_1 \\ 0 & 1 & \cdots & * & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & b_m \neq 0 \end{array}\right)$$

Wenn das inhomogene LGS lösbar ist, dann wissen wir, dass das dazugehörige homogene LGS entweder eine (nur 0) oder unendlich viele Lösungen hat (Prop. 3.27), und dass die Addition einer von Null verschiedener Lösung des homogenen LGS eine neue Lösung für das inhomogene LGS ergibt. (Prop. 3.30). Also hat das inhomogene LGS eintweder eine oder unendlich viele Lösungen. Diese Fälle entsprechen den selben reduzierten ZSF, die wir jeweils für das homogene System angegeben haben, mit dem Zusatz, dass Nullzeilen auch eine Null in der Erweiterung haben.

$$\left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_1 \\ 0 & 1 & \cdots & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_n \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{cccccccc|c} 1 & 0 & * & \cdots & * & 0 & * & \cdots & b_1 \\ 0 & 1 & * & \cdots & * & 0 & * & \cdots & b_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & b_m \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{array}\right)$$

□

Übung 3.33: ◇ Bestimmen Sie die Anzahl der Lösungen eines LGS in red. ZSF.
◇ Bestimmen Sie die Anzahl von freien Variablen in der Lösung eines LGS.

Der Beweis zeigt uns den Zusammenhang zwischen der reduzierten ZSF von A und den verschiedenen Anzahlen von Lösungen. Damit können wir folgende Fragen beantworten:

Übung 3.34: Bestimmen Sie, ob die Aussage wahr oder falsch ist.
- ◇ Ein homogenes LGS mit n Unbekannten, dessen reduzierte ZSF r führende Einsen hat, hat $n-r$ freie Variablen in der Lösungsmenge. wahr ☐ falsch ☐
- ◇ Alle führenden Einsen in einer Matrix in ZSF müssen in verschiedenen Spalten auftreten. wahr ☐ falsch ☐
- ◇ Wenn ein homogenes LGS in n Unbekannten eine reduzierte ZSF mit n führenden Einsen hat, dann hat das LGS nur die triviale Lösung. wahr ☐ falsch ☐
- ◇ Wenn die reduzierte ZSF eines LGS eine Nullzeile hat, dann hat das LGS unendlich viele Lösungen. wahr ☐ falsch ☐
- ◇ Wenn ein LGS mehr Unbekannte als Gleichungen hat, dann muss es unendlich viele Lösungen haben. wahr ☐ falsch ☐

Wir werden diese in der Vorlesung besprechen und Erklärungen oder Beispiele/Gegenbeispiele angeben.

Die letzte Frage ist so wichtig, dass wir auf sie verweisen wollen:

Proposition 3.35: *Wenn ein homogenes LGS mehr Unbekannte als Gleichungen hat, dann hat es unendlich viele Lösungen.*

BEWEIS. Ein homogenes LGS ist immer lösbar, also fällt der Fall von null Lösungen weg: es gibt mindestens eine Lösung.
Das LGS $Ax = 0$ hat eine Matrix der Größe $m \times n$ mit $m < n$. Also kann die reduzierte ZSF höchstens m führende Einsen haben, da jede Zeile nur eine führende Eins haben kann. Da $m < n$, bleiben Spalten ohne führende Eins. Diese geben freie Variablen in der Lösungsmenge. □

F. Lineare Gleichungssysteme: Study guide

Konzeptübersicht.

- ◇ Lineares Gleichungssystem
- ◇ Homogenene und inhomogene LGS
- ◇ Lösung eines LGS
- ◇ Vier Schreibweisen eines LGS: Gleichungen, Linearkombination der Koeffizientenvektoren, Matrixform, erweiterte Koeffizientenmatrix
- ◇ Zeilenstufenform und reduzierte Zeilenstufenform
- ◇ Elementare Zeilenumformungen
- ◇ Gauss-Jordan Algorithmus
- ◇ Anzahl von Lösungen eines LGS
- ◇ Struktur der Lösungsmenge eines inhomogenen LGS
- ◇ Kern einer Matrix

Skills.

- ◇ Bestimmen, ob eine Gleichung linear ist
- ◇ LGS in Matrixform schreiben
- ◇ Erweiterte Koeffizientenmatrix eines LGS aufschreiben
- ◇ Bestimmen ob eine Matrix in ZSF oder in reduzierter ZSF ist
- ◇ Elementare Zeilenumformungen auf einer Matrix ausüben
- ◇ Den Gauss-Jordan Algorithmus anwenden, um eine Matrix in reduzierte ZSF zu bringen
- ◇ Lösungsmenge eines LGS in reduzierter ZSF aufschreiben/ablesen
- ◇ Bestimmen, ob ein inhomogenes LGS lösbar ist oder nicht
- ◇ Bestimmen, wie viele freie Variablen ein LGS hat
- ◇ Mit den obigen Methoden LGS lösen
- ◇ Den Kern einer Matrix(abbildung) finden

KAPITEL 4

Lineare Unabhängigkeit und Basen

Wir haben schon gesehen, dass Linearkombinationen einer der roten Fäden in der Linearen Algebra sind. In diesem Kapitel geht es darum: Wie können wir herausfinden, ob ein Vektor eine Linearkombination von anderen Vektoren ist, oder ob gegebene Vektoren den ganzen Vektorraum aufspannen? Wann gibt es nur eine solche Linearkombination, oder wann gibt es mehrere Möglichkeiten, die denselben Vektor ergeben?

A. Linearkombinationen finden

Wie finden wir Koeffizienten λ_i, um einen Vektor b als Linearkombination $b = \lambda_1 a_1 + \cdots + \lambda_r a_r$ zu schreiben? Dies ist ein LGS, das entweder lösbar ist oder nicht. Und eine Lösung des LGS gibt uns genau die Koeffizienten, die wir suchen.

Linearkombination finden Um zu bestimmen, ob $b \in \mathbb{R}^m$ eine Linearkombination von $a_1, \ldots, a_n \in \mathbb{R}^m$ ist:

- Versuche, das LGS $\lambda_1 a_1 + \cdots + \lambda_n a_n = b$ zu lösen.
- Wenn es lösbar ist, dann gibt die (oder eine) Lösung die Koeffizienten für die Linearkombination.

Die unterschiedlichen Schreibweisen eines LGS als $\lambda_1 a_1 + \cdots + \lambda_r a_r = b$ oder $Ax = b$ gibt uns noch einen interessanten Blickwinkel darauf:

Definition 4.1: Für eine $m \times n$-Matrix A ist der **Spaltenraum** von A der von den Spalten von A erzeugte Unterraum von $\mathbb{R}^m$.

$$\text{Für } A = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} \qquad \text{ist} \qquad \mathrm{SR}(A) = \langle a_1, \ldots, a_n \rangle .$$

Also bekommen wir durch die unterschiedlichen Schreibweisen des selben LGS:

Korollar 4.2: *Für eine $m \times n$-Matrix A und $b \in \mathbb{R}^m$ gilt:*

$$b \in \mathrm{SR}(A) \iff Ax = b \text{ ist lösbar} \iff b \in \operatorname{Im} f_A$$

Insbesondere ist $\operatorname{Im} f_A = \mathrm{SR}(A)$: Das Bild einer Matrixabbildung ist der Spaltenraum der Matrix.

Beispiel 4.3: Ist $b = \left(\begin{smallmatrix} 3 \\ 3 \\ 1 \end{smallmatrix}\right)$ eine Linearkombination von $a_1 = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$ und $a_2 = \left(\begin{smallmatrix} 1 \\ 1 \\ -1 \end{smallmatrix}\right)$?

Wir lösen das LGS $\lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$, oder $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$.

$$\left(\begin{array}{cc|c} \boxed{1} & 1 & 3 \\ 1 & 1 & 3 \\ 1 & -1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} \boxed{1} & 1 & 3 \\ 0 & 0 & 0 \\ 0 & -2 & -2 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} \boxed{1} & 0 & 2 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & 0 \end{array}\right)$$

Es gibt also eine Lösung $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Also ist $b = 2 \cdot a_1 + 1 \cdot a_2$.

Es ist eine gute Idee, die Probe zu machen: $2 \cdot \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right) + 1 \cdot \left(\begin{smallmatrix} 1 \\ 1 \\ -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 \\ 3 \\ 1 \end{smallmatrix}\right)$.

Genauso können wir bestimmen, ob eine gegebene Menge ein Erzeugendensystem von $\mathbb{R}^n$ ist.

Beispiele 4.4: ◇ Erzeugen die Vektoren

$$a_1 = \begin{pmatrix} 1 \\ -4 \\ -3 \end{pmatrix},\ a_2 = \begin{pmatrix} 3 \\ 2 \\ -2 \end{pmatrix},\ a_3 = \begin{pmatrix} 4 \\ -6 \\ -7 \end{pmatrix}$$

ganz $\mathbb{R}^3$? Dies ist der Fall, wenn wir einen beliebigen Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ als Linearkombination von a_1, a_2, a_3 schreiben können.
Wir wollen also folgendes LGS lösen:

$$\lambda_1 \begin{pmatrix} 1 \\ -4 \\ -3 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 2 \\ -2 \end{pmatrix} + \lambda_3 \begin{pmatrix} 4 \\ -6 \\ -7 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\left(\begin{array}{ccc|c} \boxed{1} & 3 & 4 & x \\ -4 & 2 & -6 & y \\ -3 & -2 & -7 & z \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 3 & 4 & x \\ 0 & 14 & 10 & y+4x \\ 0 & 7 & 5 & z+3x \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 3 & 4 & x \\ 0 & 7 & 5 & \frac{y+4x}{2} \\ 0 & 0 & 0 & z+3x-\frac{y+4x}{2} \end{array}\right)$$

Wir sehen hier, dass das LGS im allgemeinen nicht lösbar ist. (Es ist nur für bestimmte Kombinationen von x, y, z lösbar, nicht für alle x, y, z.) Also bilden diese drei Vektoren kein Erzeugendensystem von $\mathbb{R}^3$. Zum Beispiel der Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ liegt nicht in ihrem Spann.

◇ Erzeugen die Vektoren $a_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $a_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $a_3 = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$ ganz $\mathbb{R}^2$?

Wir lösen

$$\lambda_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & x \\ 2 & 1 & 4 & y \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & x \\ 0 & -3 & 6 & y-2x \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & x \\ 0 & \boxed{1} & -2 & \frac{2x-y}{3} \end{array}\right)$$
$$\to \left(\begin{array}{ccc|c} \boxed{1} & 0 & 3 & x - \frac{2(2x-y)}{3} \\ 0 & \boxed{1} & -2 & \frac{2x-y}{3} \end{array}\right)$$

Dieses LGS hat eine Lösung (sogar unendlich viele), also sind die drei Vektoren ein Erzeugendensystem von $\mathbb{R}^2$.
Zum Beispiel können wir $\lambda_3 = 0$ wählen und bekommen dann

$$\begin{pmatrix} x \\ y \end{pmatrix} = \left(-\frac{1}{3}x + \frac{2}{3}y\right) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \left(\frac{2}{3}x - \frac{1}{3}y\right) \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

◇ Erzeugen $a_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$, $a_3 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}$ ganz $\mathbb{R}^3$? Wir lösen $\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$:

$$\left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & x \\ 2 & 1 & 3 & y \\ -3 & 3 & 3 & z \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & x \\ 0 & -1 & -1 & y-2x \\ 0 & 6 & 9 & z+3x \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & x \\ 0 & \boxed{1} & 1 & 2x-y \\ 0 & 0 & 3 & z+3x-(12x-6y) \end{array}\right)$$
$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & x \\ 0 & 1 & 1 & 2x-y \\ 0 & 0 & \boxed{1} & \frac{z+6y-9x}{3} \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 0 & x - (\frac{2}{3}z + 4y - 6x) \\ 0 & \boxed{1} & 0 & 2x - y - (\frac{1}{3}z + 2y - 3x) \\ 0 & 0 & 1 & \frac{1}{3}z + 2y - 3x \end{array}\right)$$
$$\longrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 0 & 0 & 7x - 4y - \frac{2}{3}z - (5x - 3y - \frac{1}{3}z) \\ 0 & \boxed{1} & 0 & 5x - 3y - \frac{1}{3}z \\ 0 & 0 & \boxed{1} & \frac{1}{3}z + 2y - 3x \end{array}\right)$$

Da das LGS lösbar ist, sind die drei Vektoren ein Erzeugendensystem von $\mathbb{R}^3$. Und wir haben

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(2x - y - \frac{1}{3}z\right) \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + \left(5x - 3y - \frac{1}{3}z\right) \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} + \left(\frac{1}{3}z + 2y - 3x\right) \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}.$$

⚠ Wichtig: Wir werden im Kurs immer und immer wieder LGS lösen. Es ist sehr wichtig, dass wir immer die *Bedeutung* des LGS im Kopf behalten: Was bedeutet die Lösung in dem Kontext, der uns zum Lösen dieses bestimmten LGS geführt hat?

Übung 4.5: Bestimmen Sie in Numbas: Ist der Aufspann bestimmter Vektoren ganz V?

B. Linear unabhängige Vektoren

Wir haben gesehen, dass es manchmal nur eine und manchmal mehrere Möglichkeiten gibt, einen Vektor als Linearkombination von gegebenen Vektoren zu schreiben. Zum Beispiel

$$\begin{aligned}\begin{pmatrix} x \\ y \end{pmatrix} &= x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} + (y-x)\begin{pmatrix} 0 \\ 1 \end{pmatrix} + x\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \dots\end{aligned}$$

hat viele Möglichkeiten, aber wenn wir nur die ersten beiden Vektoren nehmen, gibt es nur eine Möglichkeit:

$$\begin{pmatrix} x \\ y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Wir untersuchen nun diesen Unterschied.
Dazu konzentrieren wir uns darauf, auf wie viele Weisen man den Nullvektor als Linearkombination schreiben kann:
Für beliebige Vektoren $v_1, \dots, v_k \in V$ hat die Gleichung

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$$

immer mindestens eine Lösung: $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$, die triviale Lösung. Die Frage ist also, ob es noch andere Lösungen für die λ_i gibt, die auch den Nullvektor ergeben.

Definition 4.6: Vektoren $v_1, \dots, v_k \in V$ eines Vektorraums V heißen **linear unabhängig** wenn $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$ mit $\lambda_i \in \mathbb{R}$ nur die triviale Lösung $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ hat.
Wenn es λ_i, nicht alle Null, gibt, die $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$ erfüllen, dann sind die Vektoren **linear abhängig**.
Wir nennen eine solche nicht-triviale Linearkombination eine **Abhängigkeitsbeziehung**.

Vektoren sind also linear unabhängig genau dann wenn es nur eine Möglichkeit gibt, mit ihnen den Nullvektor zu erzeugen: nämlich nur, indem man alle Koeffizienten Null setzt.

Bemerkung 4.7: In vielen Lehrbüchern lesen Sie als Definition:
$v_1, \dots, v_k \in V$ sind linear unabhängig wenn

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0 \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_k = 0$$

Das ist natürlich dasselbe: wenn es eine Lösung gibt, ist es die triviale (und wir wissen, dass diese immer eine Lösung ist). Ich habe aber die Erfahrung gemacht, dass viele Studierende dann sagen

„$v_1, \dots, v_k \in V$ sind linear unabhängig wenn $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$“

(also nur die Gleichung sehen und nicht die Implikation).
⚠ Das ergibt aber keinen Sinn: Die Gleichung ist nicht die Bedingung, denn die hat ja *immer* eine Lösung. Sondern die Bedingung ist, dass es *nur* die triviale Lösung *und sonst keine* gibt.
Deswegen habe ich mich zu obiger Formulierung entschlossen.

Es gibt es ein paar einfache Fälle:

Proposition 4.8: (Einfach zu erkennende abhängige Vektoren)

(i) Eine Liste von Vektoren, die den Nullvektor enthält, ist linear abhängig.

(ii) Ein einzelner Vektor ist linear unabhängig genau dann wenn er nicht der Nullvektor ist.

(iii) Zwei Vektoren sind linear unabhängig genau dann wenn sie nicht kollinear sind.

BEWEIS. (i) Wenn $v_1, v_2, \ldots, v_k$ den Vektor $v_i = 0$ enthalten, dann haben wir die Abhängigkeitsbeziehung

$$0v_1 + \cdots + 0v_{i-1} + 1 \cdot v_i + 0v_{i+1} + \cdots + 0v_k = 0.$$

(ii) Wenn wir nur einen Vektor v_1 haben, dann ist entweder $v_1 = 0$, in welchem Fall $1 \cdot v_1 = 0$ eine Abhängigkeitsbeziehung ist, oder $v_1 \neq 0$, in welchem Fall wir nur mit $0 \cdot v_1$ den Nullvektor erhalten.

(iii)

$$\begin{aligned} \lambda_1 v_1 + \lambda_2 v_2 = 0 \text{ mit } \lambda_1 \neq 0 \quad &\Rightarrow \quad v_1 = -\frac{\lambda_2}{\lambda_1} v_2 \\ v_1 = \mu v_2 \quad &\Rightarrow \quad 1 \cdot v_1 - \mu v_2 = 0 \end{aligned}$$ □

Der letzte Punkt hilft unserer Intuition in zwei Dimensionen: kollineare Vektoren liegen auf der selben Geraden durch den Ursprung.

Beispiele 4.9: ◇ Die Vektoren $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}$ sind linear abhängig. In einem Koordinatensystem gemalt, liegen beide auf der x-Achse. Wir sehen leicht, dass sie kollinear sind.

Z.B. ist für diese beiden Vektoren $2\begin{pmatrix}1\\0\end{pmatrix} - 1 \cdot \begin{pmatrix}2\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$ eine Abhängigkeitsbeziehung.

◇ $\begin{pmatrix}1\\0\end{pmatrix}$ und $\begin{pmatrix}0\\1\end{pmatrix}$ sind linear unabhängig. Zeichen Sie sie in ein Koordinatensystem.

◇ Sind die Vektoren $\begin{pmatrix}1\\\frac{1}{2}\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}$ linear unabhängig?

Wir wollen wissen, ob es von Null verschiedene Koeffizienten gibt, die

$$\lambda_1 \begin{pmatrix}1\\\frac{1}{2}\end{pmatrix} + \lambda_2 \begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

lösen. Um dies herauszufinden, lösen wir das LGS:

$$\begin{pmatrix}\boxed{1} & 1\\ \frac{1}{2} & 2\end{pmatrix} \longrightarrow \begin{pmatrix}\boxed{1} & 1\\ 0 & \frac{3}{2}\end{pmatrix} \longrightarrow \begin{pmatrix}1 & 1\\ 0 & \boxed{1}\end{pmatrix} \longrightarrow \begin{pmatrix}\boxed{1} & 0\\ 0 & \boxed{1}\end{pmatrix}$$

Also ist die einzige Lösung $\lambda_1 = 0 = \lambda_2$. Daher sind diese Vektoren linear unabhängig.

◇ Die Standardbasisvektoren $e_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, e_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots, e_n = \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$ sind linear unabhängig in $\mathbb{R}^n$:

$$\lambda_1 \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix} + \cdots + \lambda_n \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix} = \begin{pmatrix}0\\\vdots\\0\\0\end{pmatrix}$$

hat eindeutige Lösung $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

- Die Monome $1,\ x,\ \ldots,\ x^n$ sind linear unabhängig in $\mathbb{R}[x]_n$:
$$\lambda_0 \cdot 1 + \lambda_1 \cdot x + \cdots + \lambda_n \cdot x^n = 0$$
hat eindeutige Lösung $\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$.
- Die Polynome $p_1 = 1 + x$, $p_2 = 1 + x^2$ und $p_3 = 2 + x + x^2$ sind linear *abhängig*:
$$1 \cdot p_1 + 1 \cdot p_2 - 1 \cdot p_3 = (1 + x) + (1 + x^2) - (2 + x + x^2) = 0$$
Dies ist eine nicht-triviale Linearkombination, die das Nullpolynom ergibt (also eine Abhängigkeitsbeziehung).
- Die Vektoren $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 4 \\ 5 \\ 4 \end{pmatrix}$ sind nicht linear unabhängig: zum Beispiel ist
$$3 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 4 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
eine Abhängigkeitsbeziehung.
Wie können wir so eine Abhängigkeitsbeziehung finden? Wir lösen
$$\lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 4 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
und erhalten
$$\begin{pmatrix} \boxed{1} & 1 & 4 \\ 1 & 2 & 5 \\ 1 & 1 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 4 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 3 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$
Es gibt also unendlich viele Lösungen, nicht nur $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Das sagt uns schon, dass die Vektoren linear abhängig sind. Wir können dann eine Lösung wählen: zum Beispiel die mit $\lambda_3 = -1$ gibt die obige Abhängigkeitsbeziehung.

Wir sehen in diesen Beispielen, wie wir zeigen können, ob gegebene Vektoren in $\mathbb{R}^n$ linear unabhängig sind.

Lineare Unabhängigkeit von Vektoren $a_1,\ a_2, \ldots, a_n \in \mathbb{R}^m$ bestimmen:
- Löse das homogene LGS $\lambda_1 a_1 + \lambda_2 + \cdots + \lambda_n a_n = 0$.
- Wenn $x = 0$ die einzige Lösung ist, dann sind die Vektoren linear unabhängig.
Wenn es weitere Lösungen gibt (also wenn es eine freie Variable in der Lösung gibt), dann sind die Vektoren linear abhängig.

Übung 4.10: Zeigen Sie, dass $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ in $\mathbb{R}^2$ linear unabhängig sind.

Aber wir bekommen wir einem Vektorraum, der nicht $\mathbb{R}^n$ ist, so ein LGS zum lösen?
Das hängt erst einmal vom Beispiel ab, und wir werden später auch noch eine allgemeine Methode lernen. Im Momemt können wir mit etwas Nachdenken so ein LGS bekommen. Zum Beispiel:
- In einem Matrixraum bekommen wir eine Gleichung pro Eintrag der Matrix.
- In einem Polynomraum können wir **Koeffizientenvergleich** machen und bekommen so eine Gleichung pro Koeffizient. Also zum Beispiel eine Gleichung, indem wir die Koeffizienten von x^n vergleichen, eine weitere von den Koeffizienten von x^{n-1} und so weiter. Diese geben dann unser LGS.

Beispiel 4.11: $\diamond$ Sind die Matrizen $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ linear unabhängig in $\mathcal{M}_{2,2}$? (Dies sind auch symmetrische Matrizen, wir können sie also auch als Element von Sym_2 betrachten.) Wir wollen das folgende LGS lösen:

$$\lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Die vier Matrixeinträge geben uns vier Gleichungen:

$$\begin{aligned} 1 \cdot \lambda_1 + 0 \cdot \lambda_2 + 0 \cdot \lambda_3 &= 0 \\ 0 \cdot \lambda_1 + 1 \cdot \lambda_2 + 0 \cdot \lambda_3 &= 0 \\ 0 \cdot \lambda_1 + 1 \cdot \lambda_2 + 0 \cdot \lambda_3 &= 0 \\ 0 \cdot \lambda_1 + 0 \cdot \lambda_2 + 1 \cdot \lambda_3 &= 0 \end{aligned}$$

oder

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir sehen leicht, dass $\lambda_1 = \lambda_2 = \lambda_3 = 0$ die einzige Lösung ist, also sind diese Matrizen linear unabhängig in $\mathcal{M}_{2,2}$ (und in Sym_2).

$\diamond$ Sind die Polynome $p_1 = x^2 - 3x + 1$, $p_2 = 2x^2 + x - 2$ und $p_3 = x^2 + 4x - 3$ linear unabhängig in $\mathbb{R}[x]_2$?
Wir wollen

$$\lambda_1(x^2 - 3x + 1) + \lambda_2(2x^2 + x - 2) + \lambda_3(x^2 + 4x - 3) = 0$$

lösen. Wir könnten dies auch umschreiben zu

$$(\lambda_1 + 2\lambda_2 + \lambda_3)x^2 + (-3\lambda_2 + \lambda_2 + 4\lambda_3)x + (\lambda_1 - 2\lambda_2 - 3\lambda_3) = 0x^2 + 0x + 0.$$

Wenn wir die Koeffizienten vor x^2 vergleichen, sehen wir, dass wir

$$\lambda_1 + 2\lambda_2 + \lambda_3 = 0$$

brauchen. Die Koeffizienten vor x geben uns

$$-3\lambda_1 + \lambda_2 + 4\lambda_3 = 0.$$

Die konstanten Koeffizienten geben

$$\lambda_1 - 2\lambda_2 - 3\lambda_3 = 0.$$

Also müssen wir folgendes LGS lösen:

$$\begin{pmatrix} 1 & 2 & 1 \\ -3 & 1 & 4 \\ 1 & -2 & -3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Die reduzierte ZSF dieser Matrix ist

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

also gibt es nicht-triviale Lösungen. Daher sind $p_1,\ p_2,\ p_3$ nicht linear unabhängig.
Hier ist zum Beispiel eine Abhängigkeitsbeziehung:

$$p_1 - p_2 + p_3 = (x^2 - 3x + 1) - (2x^2 + x - 2) + (x^2 + 4x - 3) = 0.$$

Neben diesen algebraische Methoden, mit denen wir lineare Unabhängigkeit bestimmen können, sollten wir die Intuition nicht vergessen.

Intuition über lineare Unabhängigkeit

- Für 2 Vektoren: linear abhängig $\Leftrightarrow$ kollinear
- Für 3 Vektoren: linear abhängig heißt sie spannen eine Ebene auf. Linear unabhängig heißt sie spannen den ganzen 3-dimensionalen Raum auf.

Übung 4.12: a) Bestimmen Sie, ob die folgenden Vektoren linear abhängig oder unabhängig sind.

- 0 in $\mathbb{R}^4$.
- $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ in $\mathbb{R}^3$.
- $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ in $\mathbb{R}^2$.
- $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 4 \\ 0 \end{pmatrix}$ in $\mathbb{R}^3$.
- $p_1 = x,\ p_2 = x^2,\ p_3 = 0$ in $\mathbb{R}[x]_2$.

b) Finden Sie ein Beispiel von drei Vektoren, die linear abhängig sind, aber von denen es zwei Vektoren gibt, die linear unabhängig sind. Überprüfen Sie Ihr Beispiel von abhängigen Vektoren in Numbas.

Unsere Intuition über die lineare Unabhängigkeit von drei Vektoren sollte uns auch zeigen:
Drei Vektoren in $\mathbb{R}^2$ müssen auf jeden Fall linear abhängig sein.
Das können wir noch allgemeiner zeigen:

Proposition 4.13: (Zu viele Vektoren)
Mehr als n Vektoren in $\mathbb{R}^n$ sind immer linear abhängig.

BEWEIS. Seien $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ mit $k > n$. Dann hat das LGS $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$ mehr Unbekannte als Gleichungen, und es ist homogen. Also hat die reduzierte ZSF der Koeffizientenmatrix auf jeden Fall Spalten ohne führende Einsen, und somit gibt es unendlich viele Lösungen (Proposition 3.35). Also sind die Vektoren linear abhängig. $\square$

C. Basen

Wir werden jetzt die Konzepte von Erzeugendensystem und linearer Unabhängigkeit verbinden. Zusammen ergeben sie eines der wichtigsten Konzepte der Linearen Algebra.

Definition 4.14: Eine **Basis** in einem Vektorraum V ist ein linear unabhängiges Erzeugendensystem.

Wir schreiben es noch mal so auf, dass man die beiden Bedingungen klar sieht:
Eine Basis ist

- linear unabhängig und
- erzeugt V.

Eine gute Intuition für eine Basis ist, dies als „Koordinatensystem“ für einen Vektorraum zu betrachten. Die erste Bedingung sagt uns, dass es keine Abhängigkeitsbeziehung zwischen den Vektoren gibt, was wir als „es gibt keine überflüssige Information“ interpretieren können. Die zweite Bedingung sagt uns, dass wir jeden Vektor in V mit unserem „Koordinatensystem“ erreichen können.

Beispiele 4.15: (Basen)
Da wir schon viele Beispiele von Erzeugendensystemen und linear unabhängigen Vektoren gesehen haben, können wir sie jetzt verbinden.

a) In $\mathbb{R}^n$ bilden die Standardbasisvektoren $\{e_1, e_2, \dots, e_n\}$ eine Basis: die **Standardbasis von** $\mathbb{R}^n$. Jetzt gibt auch der Name „Standard*basis*vektoren“ einen Sinn! Dies ist meine Lieblingsbasis (und sollte auch Ihre werden).

b) In $\mathbb{R}^2$ bilden $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ eine Basis:

- ◇ Wir haben gesehen, dass sie linear unabhängig sind.
- ◇ Und sie erzeugen ganz $\mathbb{R}^2$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x+y}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{x-y}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

c) Die Monome $1, x, x^2, \dots, x^n$ sind eine Basis von $\mathbb{R}[x]_n$. Wir nennen sie auch die **Standardbasis von** $\mathbb{R}[x]_n$.

Übung 4.16: Zeigen Sie, dass die folgenden Vektoren jeweils eine Basis des angegebenen Vektorraums bilden.

a) $v_1 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ in $\mathbb{R}^2$.
b) $w_1 = \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$, $w_2 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ in $\mathbb{R}^2$.
c) $v_1 = \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix} -1 \\ 1 \\ 0 \end{smallmatrix}\right)$, $v_3 = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$ in $\mathbb{R}^3$.
d) $v_1 = \left(\begin{smallmatrix} 1 \\ 2 \\ 1 \end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix} 2 \\ 9 \\ 0 \end{smallmatrix}\right)$, $v_3 = \left(\begin{smallmatrix} 3 \\ 3 \\ 4 \end{smallmatrix}\right)$ in $\mathbb{R}^3$.
e) Die Matrizen $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ in $\mathcal{M}_{2,2}$.
f) Allgemeiner: Sei E_{ij} die $m \times n$-Matrix mit überall Nullen, bis auf eine 1 im Eintrag i, j. Zeigen Sie, dass $E_{11}, E_{12}, \dots, E_{1n}, E_{21}, E_{22}, \dots, E_{2n}, E_{31}, \dots, E_{m1}, \dots, E_{mn}$ eine Basis von $\mathcal{M}_{m,n}$ bilden.
g) Basis oder nicht? Multiple choice in Numbas.

Wir sehen:

Ein Vektorraum kann verschiedene Basen haben.

Wir werden noch sehr viel mehr über Basen lernen, und auch später noch lernen, wie man besonders schöne Basen für eine gegebene Situation finden kann.
In allen bisherigen Beispielen hatten wir Basen aus endlich vielen Vektoren. Das ist aber nicht immer möglich.

Proposition 4.17: *Der Vektorraum $\mathbb{R}[x]$ aus allen Polynomen hat kein endliches Erzeugendensystem.*

BEWEIS. Nicht in VL: zum freiwilligen eigenen Durchlesen.
Angenommen wir haben ein endliches Erzeugendensystem. Sei k der höchste Grad von allen Polynomen in diesem Erzeugendensystem. Dann können wir das Polynom $x^{k+1} \in \mathbb{R}[x]$ nicht als Linearkombination von Vektoren im Erzeugendensystem schreiben, weil wir durch Linearkombinationen keinen höheren Grad erzeugen können. Also ist es doch kein Erzeugendensystem. □

Manche unserer Resultat funktionieren nur für Vektorräume mit endlicher Basis.

Definition 4.18: Ein Vektorraum V heißt **endlich-dimensional** wenn er ein endliches Erzeugendensystem hat. Wenn V kein endliches Erzeugendensystem hat, sagen wir es ist **unendlich-dimensional**.

Wir benutzen hier das Wort „dimensional“, was wohl etwas mit Dimensionen zu tun hat, die wir noch nicht formal eingeführt haben. Dazu brauchen wir noch ein Resultat über Basen. Aber wir können jetzt schon entscheiden, ob ein Vektorraum endlich-dimensional ist oder nicht.

D. Koordinaten bezüglich einer Basis

Hier ist eine kleine Zusammenfassung:

Sei V ein (reeller) Vektorraum. Die Vektoren $v_1, v_2, \ldots, v_n \in V$
- sind **linear unabhängig** wenn $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$ nur die triviale Lösung $\lambda_i = 0$ für alle i hat;
- **erzeugen** V wenn jeder Vektor aus V als Linearkombination der Vektoren v_i geschrieben werden kann;
- sind eine **Basis** von V wenn sie linear unabhängig sind und V erzeugen.

Wir werden jetzt die Intuition der Basis als Koordinatensystem präzisieren, und wieder auf die Frage eingehen, ob ein Vektor auf mehrere Weisen als Linearkombination dargestellt werden kann.

Theorem 4.19: (Eindeutigkeit der Basisrepräsentation)
Wenn $v_1, \ldots, v_n$ eine Basis für Vektorraum V bilden, dann kann jeder Vektor $v \in V$ auf genau eine Weise als Linearkombination der Vektoren v_i ausgedrückt werden.

BEWEIS. Da $v_1, \ldots, v_n$ eine Basis sind, erzeugen sie V. Also gibt es für jedes $v \in V$ Kooeffizienten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit

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

Angenommen für $\mu_1, \ldots, \mu_n \in \mathbb{R}$ gilt auch

$$v = \mu_1 v_1 + \ldots + \mu_n v_n .$$

also angenommen es gibt eine weitere Linearkombination, die v darstellt.
Wir wollen nun zeigen, dass $\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n$, also dass es doch nur eine Linearkombination gibt.
Wenn wir eine Gleichung von der anderen abziehen, bekommen wir

$$0 = (\lambda_1 - \mu_1)v_1 + \ldots + (\lambda_n - \mu_n)v_n .$$

Da $v_1, \ldots, v_n$ linear unabhängig sind, folgt daraus $\lambda_1 - \mu_1 = \ldots = \lambda_n - \mu_n = 0$, was unsere Behauptung zeigt. □

Wir sehen im Beweis die zwei Teile:
- Erzeugendensystem gibt Existenz einer Linearkombination: jedes $v \in V$ kann auf *mindestens eine Weise* als Linearkombination eines Erzeugendensystems geschrieben werden.
- Lineare Unabhängigkeit gibt Eindeutigkeit der Linearkombination: jedes $v \in V$ kann auf *höchstens eine Weise* als Linearkombination von linear unabhängigen Vektoren geschrieben werden.

In einer Basis kombiert haben wir also beides: jedes $v \in V$ kann auf *genau eine Weise* als Linearkombination einer Basis geschrieben werden. Wir haben eine eindeutige Basisrepräsentation $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$.

Definition 4.20: Wenn $\mathcal{B}\colon v_1, \ldots, v_n$ eine Basis für Vektorraum V ist und

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$

dann nennen wir die Koeffizienten $\lambda_1, \ldots, \lambda_n$ die **Koordinaten** von v bezüglich Basis $\mathcal{B}$. Der Vektor

$${}_{\mathcal{B}}(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

heißt **Koordinatenvektor** von v bezüglich $\mathcal{B}$.

Anders gesagt: Es gibt eine bijektive Abbildung ${}_{\mathcal{B}}(-)\colon V \longrightarrow \mathbb{R}^n$, die jedem Element $v \in V$ einen eindeutigen Vektor ${}_{\mathcal{B}}(v) \in \mathbb{R}^n$ zuordnet. Wir werden später noch sehen, dass diese Bijektion linear ist.

Beispiel 4.21: Sei $V = \mathbb{R}^2$ und sei $e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ die Standardbasis $\mathcal{E}$. Dann ist der Koordinatenvektor von $v = x \cdot e_1 + y \cdot e_2$ gegeben durch

$$_{\mathcal{E}}(v) = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Sei nun $\mathcal{B}$ die Basis aus den Vektoren $v_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ und $v_2 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$. Dann ist

$$\begin{pmatrix} x \\ y \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (x-y) \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

also ist der Koordinatenvektor von $v = \begin{pmatrix} x \\ y \end{pmatrix}$ bezüglich $\mathcal{B}$ der Vektor

$$_{\mathcal{B}}(v) = \begin{pmatrix} y \\ x-y \end{pmatrix}.$$

Wie finden wir diese Koordinatenvektoren? Das haben wir schon in „Linearkombinationen finden“, Abschnitt 44A, gelernt.

Koordinatenvektoren finden
Für eine Basis $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ eines Vektorraums V finden wir den Koordinatenvektor $_{\mathcal{B}}(v)$ von $v \in V$, indem wir das LGS

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = v.$$

lösen.

- Wenn $V = \mathbb{R}^n$, löse $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = v$ für den gesuchten Vektor v, (ein spezifischer Vektor oder allgemein $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$). Wenn $V \neq \mathbb{R}^n$, benutze die vorher erwähnten Methoden, um ein LGS aufzustellen.
- Das LGS hat eine eindeutige Lösung $\lambda_1, \lambda_2, \dots, \lambda_n$ (wenn $\mathcal{B}$ wirklich eine Basis ist).
- Dann ist $_{\mathcal{B}}(v) = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$ der Koordinatenvektor.

Übung 4.22: a) Finden Sie in Numbas Koordinaten bezüglich einer Basis in $\mathbb{R}^3$. und Koordinaten bezüglich einer Basis in $\mathbb{R}^4$.

b) Finden Sie den Koordinatenvektor von $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ bezüglich Basis $\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ von $\mathbb{R}^2$.

c) Finden Sie den Koordinatenvektor des Polynoms

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n$$

bezüglich der Standardbasis von $\mathbb{R}[x]_n$.

Übung 4.23: Seien $v, w \in V$, $\lambda, \mu \in \mathbb{R}$ und $v_1, v_2, \dots, v_n$ eine Basis $\mathcal{B}$ von V. Zeigen Sie, dass

$$_{\mathcal{B}}(\lambda v + \mu w) = \lambda \cdot {}_{\mathcal{B}}(v) + \mu \cdot {}_{\mathcal{B}}(w).$$

In Worten: der Koordinatenvektor einer Linearkombination ist die Linearkombination der Koordinatenvektoren.
Dies zeigt, dass die Bijektion $_{\mathcal{B}}(-): V \longrightarrow \mathbb{R}^n$ eine lineare Abbildung ist (weil sie Linearkombinationen erhält).

E. Lineare Abbildungen und Basen

Die eindeutige Basisrepräsentation hat auch Konsequenzen für lineare Abbildungen.

Proposition 4.24: (Lineare Abbildungen sind eindeutig bestimmt durch Bilder einer Basis.)
Sei $f\colon V \longrightarrow W$ eine lineare Abbildung und $v_1, \cdots, v_n$ eine Basis von V. Dann gilt für beliebiges $v \in V$

$$f(v) = \lambda_1 f(v_1) + \cdots \lambda_n f(v_n),$$

wobei die λ_i durch die eindeutige Basisrepräsentation $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ gegeben sind. Jede Wahl von Bildern $w_1, \dots w_n \in W$ mit $f(v_i) = w_i$ definiert in dieser Weise eine lineare Abbildung.

BEWEIS. Jedes $v \in V$ hat eine eindeutige Basisrepräsentation $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Also folgt aus der Linearität, dass $f(v) = f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n)$.
Für den zweiten Teil, sei $f(v_i) = w_i$ für $i = 1, \dots, n$. Wir definieren $f(v) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n)$. Dies ist wohl-definiert, weil die Repräsentation $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ *eindeutig* ist.
Wir müssen prüfen, dass die so definierte Abbildung f wirklich linear ist. Seien $u, v \in V$. Dann ist $u = \mu_1 v_1 + \ldots + \mu_n v_n$ und $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ für eindeutige $\mu_i, \lambda_i \in \mathbb{R}$. Dann gilt für beliebige Skalare $\nu_1, \nu_2 \in \mathbb{R}$:

$$\begin{aligned}\nu_1 u + \nu_2 v &= \nu_1(\mu_1 v_1 + \ldots + \mu_n v_n) + \nu_2(\lambda_1 v_1 + \ldots + \lambda_n v_n)\\ &= (\nu_1\mu_1 + \nu_2\lambda_1)v_1 + (\nu_1\mu_2 + \nu_2\lambda_2)v_2 + \ldots + (\nu_1\mu_n + \nu_2\lambda_n)v_n.\end{aligned}$$

Also ist

$$\begin{aligned}f(\nu_1 u + \nu_2 v) &= (\nu_1\mu_1 + \nu_2\lambda_1)f(v_1) + (\nu_1\mu_2 + \nu_2\lambda_2)f(v_2) + \ldots + (\nu_1\mu_n + \nu_2\lambda_n)f(v_n)\\ \text{und}\quad \nu_1 f(u) + \nu_2 f(v) &= \nu_1\left[\mu_1 f(v_1) + \ldots + \mu_n f(v_n)\right] + \nu_2\left[\lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)\right]\\ &= (\nu_1\mu_1 + \nu_2\lambda_1)f(v_1) + (\nu_1\mu_2 + \nu_2\lambda_2)f(v_2) + \cdots + (\nu_1\mu_n + \nu_2\lambda_n)f(v_n),\end{aligned}$$

und somit $f(\nu_1 u + \nu_2 v) = \nu_1 f(u) + \nu_2 f(v)$, also ist f linear. □

Intuition in $\mathbb{R}^2$: Video „Linear transformations and matrices"– Chapter 3, Essence of linear algebra, von Three Blue One Brown auf YouTube.

Beispiele 4.25:

a) Die Matrix $A = \left(\begin{smallmatrix}3 & 1\\ 5 & 6\end{smallmatrix}\right)$ schickt die Standardbasisvektoren $e_1 = \left(\begin{smallmatrix}1\\ 0\end{smallmatrix}\right)$ und $e_2 = \left(\begin{smallmatrix}0\\ 1\end{smallmatrix}\right)$ auf $\left(\begin{smallmatrix}3\\ 5\end{smallmatrix}\right)$ bzw. $\left(\begin{smallmatrix}1\\ 6\end{smallmatrix}\right)$. Für jedes $v = \left(\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right) \in \mathbb{R}^2$ haben wir

$$v = x_1 \begin{pmatrix}1\\ 0\end{pmatrix} + x_2 \begin{pmatrix}0\\ 1\end{pmatrix}$$

also ist

$$Av = x_1 \begin{pmatrix}3\\ 5\end{pmatrix} + x_2 \begin{pmatrix}1\\ 6\end{pmatrix} = \begin{pmatrix}3x_1 + x_2\\ 5x_1 + 6x_2\end{pmatrix}.$$

b) Betrachten wir die Basis

$$v_1 = \left(\begin{smallmatrix}1\\ 1\\ 1\end{smallmatrix}\right),\ v_2 = \left(\begin{smallmatrix}1\\ 1\\ 0\end{smallmatrix}\right),\ v_3 = \left(\begin{smallmatrix}1\\ 0\\ 0\end{smallmatrix}\right)$$

von $\mathbb{R}^3$. Wir können eine lineare Abbildung $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definieren, indem wir

$$f(v_1) = \left(\begin{smallmatrix}1\\ 0\end{smallmatrix}\right),\ f(v_2) = \left(\begin{smallmatrix}2\\ -1\end{smallmatrix}\right),\ f(v_3) = \left(\begin{smallmatrix}4\\ 3\end{smallmatrix}\right)$$

setzen. Für jedes $v = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \mathbb{R}^3$ können wir dann $f(v)$ bestimmen, indem wir zuerst die Koordinaten $\lambda_1, \lambda_2, \lambda_3$ finden, sodass $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 + \lambda_3 \\ \lambda_1 + \lambda_2 \\ \lambda_1 \end{pmatrix}$$

Also ist $\lambda_1 = x_3$, $\lambda_2 = x_2 - x_3$ und $\lambda_3 = x_1 - x_2$. Also ist

$$f(v) = x_3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (x_2 - x_3) \begin{pmatrix} 2 \\ -1 \end{pmatrix} + (x_1 - x_2) \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

Dies bestimmt die lineare Abbildung f auf eindeutige Weise.

Bemerkung 4.26: Das Beispiel a) in 4.25 zeigt uns noch eine andere Sichtweise auf $\mathrm{SR}(A) = \mathrm{Im}\, f_A$ (aus Kor. 4.2).

Die Spalten einer Matrix sind genau die Bilder der Standardbasis (unter der Matrixabbildung).

Da eine lineare Abbildung durch die Bilder einer Basis bestimmt ist, sehen wir genau, warum das Bild $\mathrm{Im}\, f_A$ der Matrixabbildung genau die Linearkombinationen der Spalten von A sind.

Formel für lineare Abbildung von Bildern einer Basis finden
Angenommen $v_1, \ldots, v_n$ ist eine Basis für V, und wir wissen, dass $f(v_i) = w_i \in W$, also wir kennen die Bilder dieser Basis unter der linearen Abbildung f. Wie finden wir heraus, wohin f einen beliebigen Vektor $v \in V$ abbildet? Im Fall $V = \mathbb{R}^n$:

- ◇ Finde die Koordinaten $\lambda_1, \ldots, \lambda_n$ für v, also $\left(\begin{smallmatrix} x_1 \\ \vdots \\ x_n \end{smallmatrix}\right) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$.

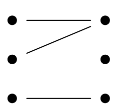
  (Manchmal sieht man sie, ansonsten löse das inhomogene LGS.)
- ◇ Dann ist $f\left(\left(\begin{smallmatrix} x_1 \\ \vdots \\ x_n \end{smallmatrix}\right)\right) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n$.

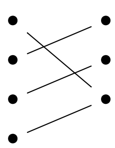
Übung 4.27: Finden Sie in Numbas die Formel einer linearen Abbildung, die durch Bilder einer Basis gegeben ist.

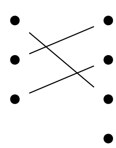
Wir untersuchen noch weitere Eigenschaften von linearen Abbildungen. Zuerst eine Wiederholung von Begriffen aus MAGL, mit einer etwas anderen Formulierung der Definition.

Definition 4.28: Eine Abbildung $f\colon X \longrightarrow Y$ heißt
injektiv wenn für $a, b \in X$ gilt: aus $f(a) = f(b)$ folgt $a = b$;
surjektiv wenn $\mathrm{Im}(f) = Y$.
bijektiv wenn sie injektiv und surjektiv ist.

Slogan für injektiv: *Verschiedene Elemente haben verschiedene Bilder.*
Slogan für surjektiv: *f erreicht alles im Ziel.*


weder injektiv noch surjektiv


surjektiv, nicht injektiv


injektiv, nicht surjektiv

Beispiele 4.29: Von den Beispielen lineare Abbildung in 2.13:

b) (Streckungen) Wenn $r \neq 0$, dann ist $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mit $f(v) = rv$ injektiv:

$$\begin{aligned} & & rv &= rw \\ &\Rightarrow & r(v-w) &= 0 \\ &\Rightarrow & v-w &= 0 \\ &\Rightarrow & v &= w. \end{aligned}$$

f ist auch surjektiv: Für $w \in \mathbb{R}^n$ ist auch $\frac{1}{r}w \in \mathbb{R}^n$, und $r \cdot (\frac{1}{r}w) = w$.

c) Die Nullabbildung $0\colon V \longrightarrow W$ ist
 nicht injektiv solange $V \neq \{0\}$;
 nicht surjektiv solange $W \neq \{0\}$.

d) Die Identität $\mathrm{id}\colon V \longrightarrow V$ ist bijektiv.

e) (Projektion) $\pi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit $\pi\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)\right) = x_1$ ist surjektiv:
 für $x \in \mathbb{R}$ ist $\pi\left(\left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right)\right) = x$.
 Aber π ist nicht injektiv: $\pi\left(\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\right) = \pi\left(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right) = 1$.

Aus Beispiel 2.15:
$f\colon \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ with $f(p) = xp$ ist injektiv:
Angenommen wir haben $p = a_0 + a_1x + \cdots + a_nx^n$ und $q = b_0 + b_1x + \cdots + b_mx^m$ und $xp = xq$.

$$\begin{aligned} &\Rightarrow & a_0x + a_1x^2 + \cdots a_nx^{n+1} &= b_0x + \cdots + b_mx^{m+1} \\ &\Rightarrow & n = m, \quad \text{und} \quad a_k &= b_k \text{ für alle } k, \text{ durch Koeffizientenvergleich.} \\ &\Rightarrow & p &= q. \end{aligned}$$

Also ist f injektiv. Aber f ist nicht surjektiv: Das Polynom $p = 1$ hat kein Urbild (es gibt kein Polynom q mit $xq = 1$).

Proposition 4.30: (Komposition injektiver oder surjektiver Abbildungen)
Für Abbildungen $f\colon X \longrightarrow Y$ und $g\colon Y \longrightarrow Z$ gilt:

(i) Wenn f und g beide injektiv sind, dann ist auch $g \circ f\colon X \longrightarrow Z$ injektiv.
(ii) Wenn f und g beide surjektiv sind, dann ist auch$g \circ f\colon X \longrightarrow Z$ surjektiv.

BEWEIS. Übung. □

Für lineare Abbildungen gibt es eine einfachere Bedingung für Injektivität.

Proposition 4.31: (Injektivität via Kern)
Eine lineare Abbildung $f\colon V \longrightarrow W$ ist injektiv genau dann wenn $\operatorname{Ker} f = \{0\}$.

BEWEIS. Wenn f injektiv ist, dann gilt

$$v \in \operatorname{Ker} f \quad \Rightarrow \quad f(v) = 0 \quad \Rightarrow \quad f(v) = f(0) \quad \Rightarrow \quad v = 0.$$

Also ist $\operatorname{Ker} f = \{0\}$.
Andersherum, wenn $\operatorname{Ker} f = \{0\}$, betrachten wir v, w mit $f(v) = f(w)$. Dann ist

$$f(v) - f(w) = 0 \quad \Rightarrow \quad f(v-w) = 0 \quad \Rightarrow \quad v - w \in \operatorname{Ker} f \quad \Rightarrow \quad v - w = 0 \quad \Rightarrow \quad v = w.$$

Also ist f injektiv. □

Hiermit können wir auch Eigenschaften einer linearen Abbildung an den Bildern einer Basis erkennen.

Proposition 4.32: *Sei $f\colon V \longrightarrow W$ linear und $v_1, \ldots, v_n$ eine Basis von V.*

(i) f ist injektiv genau dann wenn $f(v_1), \ldots, f(v_n)$ linear unabhängig in W.
(ii) f ist surjektiv genau dann wenn $f(v_1), \ldots, f(v_n)$ Erzeugendensystem von W.

BEWEIS. Übung. □

F. Lineare Unabhängigkeit und Basen: Study guide

Konzeptübersicht.

- ◇ Spaltenraum einer Matrix
- ◇ Verbindung von Lösung von LGS und Spann/Spaltenraum
- ◇ Linear unabhängige Vektoren, linear abhängige Vektoren, Abhängigkeitsbeziehung
- ◇ Intuition, wann ein, zwei oder drei Vektoren linear unabhängig sind.
- ◇ Lineare Abhängigkeit von mehr als n Vektoren in $\mathbb{R}^n$.
- ◇ Basis eines Vektorraums
- ◇ Endlich-dimensionaler Vektorraum
- ◇ Eindeutigkeit der Basisrepršentation
- ◇ Koordinatenvektoren
- ◇ Lineare Abbildungen eindeutig bestimmt durch Bilder einer Basis
- ◇ Injektive, surjektive, bijektive Abbildungen und deren Komposition
- ◇ Injektivität via Kern
- ◇ Verbindung zwischen Injektivität bzw. Surjektivität und Eigenschaften der Bilder einer Basis

Skills.

- ◇ Bestimmen, ob (und beweisen, dass) ein Vektor im Spann von anderen Vektoren liegt
- ◇ Bestimmen, ob (und beweisen, dass) gegebene Vektoren ein Erzeugendensystem von $\mathbb{R}^n$ bilden
- ◇ Bestimmen, ob (und beweisen, dass) gegebene Vektoren linear unabhängig sind
- ◇ Bestimmen, ob (und beweisen, dass) gegebene Vektoren eine Basis bilden
- ◇ Obige auch für Polynome und Matrizen als Elemente des Vektorraums
- ◇ Koordinatenvektor bezüglich einer Basis finden
- ◇ Funktionsvorschrift einer linearen Abbildung finden, die durch Bilder einer Basis gegeben ist
- ◇ Bestimmen, ob (und beweisen, dass) eine lineare Abbildung injektiv ist (z.B. via Kern)
- ◇ Bestimmen, ob (und beweisen, dass) eine lineare Abbildung surjektiv ist

KAPITEL 5

Basen und Dimension

Die Resultate dieses Kapitels sind entscheidend für das Verständnis von Linearer Algebra, da viel Theorie und auch Intuition von ihnen abhängt.
Wir interessieren uns für die Beziehungen zwischen linear unabhängigen Mengen, Erzeugendensystemen und Basen. Insbesondere das Verhältnis der möglichen Größen solcher Mengen ist wichtig für das Konzept der *Dimenion*.

A. Dimension

Wir haben schon einen informelle Intuition für Dimension mit dem „Freiheitsgrad" von Geraden und Ebenen. Und wir stellen fest, dass die Standardbasis von $\mathbb{R}$, $\mathbb{R}^2$ und $\mathbb{R}^3$ jeweils so viele Elemente enthält, wie unsere Intuition des Freiheitsgrades angibt. Dies können wir jetzt etwas formalisieren.

Proposition 5.1: (Basen eines Vektorraums sind gleich groß.)
Sei V ein endlich-dimensionaler Vektorraum, und $\mathcal{B} : v_1, \ldots, v_n$ eine beliebige Basis von V.
(i) Eine Liste aus mehr als n Vektoren aus V ist linear abhängig.
(ii) Eine Liste aus weniger als n Vektoren aus V erzeugt nicht ganz V.

Beweis. (i) Seien $w_1, \ldots, w_k$ Vektoren aus V, mit $k > n$. Da $\mathcal{B}$ eine Basis ist, können wir jeden Vektor w_i als Koordinatenvektor ${}_{\mathcal{B}}(w_i) \in \mathbb{R}^n$ schreiben.
Um zu zeigen, dass $w_1, \ldots, w_k$ linear unabhängig sind, betrachten wir

$$\lambda_1 w_1 + \cdots + \lambda_k w_k = 0.$$

Wenn wir Koordinatenvektoren bezüglich $\mathcal{B}$ benutzen, ist dies äquivalent zu

$$\lambda_1 \cdot {}_{\mathcal{B}}(w_1) + \cdots + \lambda_k \cdot {}_{\mathcal{B}}(w_k) = 0$$

(weil ${}_{\mathcal{B}}(-)$ eine lineare Bijektion ist, siehe Übung 4.23). Aber ${}_{\mathcal{B}}(w_1), \ldots, {}_{\mathcal{B}}(w_k)$ sind k Vektoren in $\mathbb{R}^n$ mit $k > n$, also folgt aus Proposition 4.13, dass wir zu viele Vektoren haben, und ${}_{\mathcal{B}}(w_1), \ldots, {}_{\mathcal{B}}(w_k)$ sind linear abhängig. Also gibt es $\lambda_1, \ldots, \lambda_k$ nicht alle null, sodass

$$\lambda_1 w_1 + \cdots + \lambda_k w_k = 0\,.$$

Daher sind $w_1, \ldots, w_k$ linear abhängig.
Zusammenfassung: Wir betrachten die Koordinatenvektoren von w_i, damit wir voheriges Wissen aus $\mathbb{R}^n$ benutzen können.

(ii) Seien $w_1, \ldots, w_m$ Vektoren in V mit $m < n$. Wenn wir annehmen, dass $\mathrm{Span}(w_1, \ldots, w_m) = V$, dann können wir jeden Vektor $v_i \in \mathcal{B}$ als Linearkombination der Vektoren $w_1, \ldots, w_m$ darstellen, also

$$\begin{aligned} v_1 &= a_{11} w_1 + \cdots + a_{m1} w_m \\ \vdots &\quad \vdots \\ v_n &= a_{1n} w_1 + \cdots + a_{mn} w_m \end{aligned}$$

Wir betrachten die Gleichung $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. Wenn wir die obigen Audrücke für v_i in diese Gleichung einsetzen und dann etwas umsortieren, haben wir

$$(\lambda_1 a_{11} + \cdots + \lambda_n a_{1n}) w_1 + \cdots + (\lambda_1 a_{m1} + \cdots + \lambda_n a_{mn}) w_m = 0\,.$$

Wir nennen dies

$$\mu_1 w_1 + \cdots + \mu_m w_m = 0,$$

also wir setzen $\lambda_1 a_{11} + \cdots + \lambda_n a_{1n} = \mu_1$ und so weiter. Nun betrachten wir das homogene LGS

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

mit n Unbekannten und m Gleichungen, wobei $n > m$. Nach Proposition 3.35 gibt es also eine nicht-triviale Lösung $\lambda_1, \dots, \lambda_n$, weil es mehr Unbekannte als Gleichungen gibt. Aber dann sind die Koeffizienten μ_i in der Gleichung

$$\mu_1 w_1 + \cdots + \mu_m w_m = 0$$

alle Null (aber mit von Null verschiedenen λ_i), also können wir dies wieder zu

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

umsortieren und sehen, dass dies eine nicht-triviale Lösung hat. Dies ist aber ein Widerspruch dazu, dass $\mathcal{B}$ eine Basis ist: jede Basis ist linear unabhängig. □

Zusammenfassung: Wir nehmen an, dass die Vektoren ganz V erzeugen, und zeigen dann, dass $\mathcal{B}$ linear abhängig sein muss. Widerspruch. Also können die Vektoren nicht ganz V erzeugen.

Also hat jede Basis von V gleich viele Vektoren.

Definition 5.2: Die **Dimension** eines endlich-dimensionalen Vektorraums V, geschrieben $\dim(V)$ oder $\dim V$, ist definiert als die Anzahl der Vektoren in einer Basis von V. Der Nullvektorraum hat Dimension 0.

Beispiel 5.3: $\dim \mathbb{R}^n = n$, $\dim \mathbb{R}[x]_n = n+1$, $\dim \mathcal{M}_{m,n} = mn$.

Beispiel 5.4: Wenn die Vektoren $v_1, \dots, v_k$ linear unabhängig sind, dann bilden sie eine Basis für $\operatorname{Span}(v_1, \dots, v_k)$. Also ist $\dim \operatorname{Span}(v_1, \dots, v_k) = k$. Die Bedingung „linear unabhängig" ist hier entscheidend.
In Worten: Die Dimension eines Vektorraums, der als Spann von linear unabhängigen Vektoren angegeben ist, ist die Anzahl dieser linear unabhängigen Vektoren.
Zum Beispiel:

- ◇ $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in $\mathbb{R}^3$ ist eine Basis des ein-dimensionalen Unterraums $\operatorname{Span}(v_1) = \langle v_1 \rangle$.
- ◇ $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ in $\mathbb{R}^4$ bilden eine Basis des zwei-dimensionalen Unterraums $W = \langle v_1, v_2 \rangle$, den sie aufspannen.
- ◇ $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$ sind *nicht* linear unabhängig (da $v_3 = v_1 + v_2$). Wir haben $\langle v_1, v_2, v_3 \rangle = \langle v_1, v_2 \rangle = W$, aber v_1, v_2, v_3 bilden *keine* Basis von W. Beliebige zwei dieser Vektoren sind linear unabhängig, also bilden v_1, v_2 eine Basis von W, und v_1, v_3 bilden auch eine Basis von W, und v_2, v_3 ebenso.

Um eine Dimension eines bestimmten Vektorraums zu bestimmen, hilft manchmal die Vorstellung von „Freiheitsgrad" oder „wie viele Zahlen kann ich wählen, bis ein Element des Vektorraums eindeutig bestimmt ist". Zum Beispiel:

Beispiel 5.5: Der Vektorraum Sym_n aller symmetrischen $n \times n$-Matrizen hat Dimension $n + \frac{1}{2}(n^2 - n) = \frac{1}{2}(n^2 + n)$. Wir können dies auf zwei Weisen zählen:

- ⋄ Wir können alle Einträge auf der Diagonale wählen, also n Einträge. Dann können wir von den restlichen $n^2 - n$ Einträgen die Hälfte wählen, weil diese dann die andere Hälfte bestimmt (da die Matrix symmetrisch ist).
- ⋄ Oder: Sagen wir, wir wählen die untere Hälfte der Matrix, inklusive Diagonale. In der ersten Zeile können wir also nur einen Eintrag wählen. In der zweiten Zeile zwei Einträge, in der dritten drei, usw., bis wir in der $n$ten Zeile alle n Einträge wählen dürfen. Wir wählen also insgesamt $1 + 2 + 3 + \cdots + n = \frac{1}{2}n(n+1)$ Einträge.

Dies ist eine hilfreiche Weise, darüber nachzudenken. Um zu beweisen, dass es wirklich korrekt ist, müssen wir eine Basis geben, die ebensoviele Elemente hat. Wir bekommen pro Eintrag, den wir wählen dürfen, ein Element der Basis.
Für symmetrische 2×2-Matrizen gibt dies zum Beispiel die schöne Basis

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, S_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Für symmetrische 4×4-Matrizen haben wir

$$E_1 = \begin{pmatrix} 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \end{pmatrix}, \quad E_4 = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \end{pmatrix},$$

$$S_{12} = \begin{pmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}, \quad S_{13} = \begin{pmatrix} 0&0&1&0 \\ 0&0&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \end{pmatrix}, \quad S_{14} = \begin{pmatrix} 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \\ 1&0&0&0 \end{pmatrix},$$

$$S_{23} = \begin{pmatrix} 0&0&0&0 \\ 0&0&1&0 \\ 0&1&0&0 \\ 0&0&0&0 \end{pmatrix}, \quad S_{24} = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \\ 0&1&0&0 \end{pmatrix}, \quad S_{34} = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix}$$

Übung 5.6: Schreiben Sie eine Basis für die symmetrischen 3×3-Matrizen auf.

Wir können ähnliche Ideen benutzen, um eine Basis eines Kerns eine Matrix (oder der Lösungsmenge eines homogenen LGS) zu finden:

Beispiele 5.7: ⋄ Sei $A = \begin{pmatrix} 1 & 5 & 3 \\ 2 & 7 & 9 \\ 1 & 2 & 6 \end{pmatrix}$. Wir suchen eine Basis von $\operatorname{Ker} A = \operatorname{Ker} f_A$, der Lösungsmenge von $Ax = 0$.
Wir wissen schon, wie wir das System lösen und eine Lösungsmenge aufschreiben.

$$\begin{pmatrix} \boxed{1} & 5 & 3 \\ 2 & 7 & 9 \\ 1 & 2 & 6 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 5 & 3 \\ 0 & -3 & 3 \\ 0 & -3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 5 & 3 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 8 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Also ist $\operatorname{Ker} A = \left\{ t \begin{pmatrix} -8 \\ 1 \\ 1 \end{pmatrix} \right\} = \left\langle \begin{pmatrix} -8 \\ 1 \\ 1 \end{pmatrix} \right\rangle$. Da wir eine freie Variable haben, kann $\operatorname{Ker} A$ als Spann eines einzigen Vektors geschrieben werden. Da der Vektor nicht Null ist, ist dies also der Spann von linear unabhängigen Vektoren. Also ist $\begin{pmatrix} -8 \\ 1 \\ 1 \end{pmatrix}$ eine Basis von $\operatorname{Ker} A$.

Also ist $\dim(\operatorname{Ker} A) = 1$.

◇ Sei $A = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -4 & -3 \\ 2 & 8 & 6 \end{pmatrix}$. Diese Matrix hat reduzierte ZSF $\begin{pmatrix} 1 & 4 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, also ist

$$\operatorname{Ker} A = \left\{ s \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\} = \left\langle \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Unsere Intuition sagt, dass wir zwei Freiheitsgrade haben, also sollten diese beiden Vektoren eine Basis von $\operatorname{Ker} A$ bilden. Wie können wir sicher sein?
Offensichtlich erzeugen sie $\operatorname{Ker} A$, also müssen wir nur lineare Unabhängigkeit zeigen. Das LGS

$$\lambda_1 \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat einzige Lösung $\lambda_1 = \lambda_2 = 0$, was wir von der zweiten und dritten Zeile ablesen können. Also bilden diese beiden Vektoren eine Basis von $\operatorname{Ker} A$, und somit ist $\dim(\operatorname{Ker} A) = 2$.
Dies passiert immer, wenn wir freie Variablen haben: wir können die unterschiedlichen freien Variablen unabhängig voneinander wählen. Deswegen ist es hilfreich, den allgemeinen Lösungsvektor in die Linearkombination $s \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$ zu separieren, und ihn nicht als $\begin{pmatrix} -4s - 3t \\ s \\ t \end{pmatrix}$ stehen zu lassen.

Wir halten fest:

$$\dim(\operatorname{Ker}(A)) = \text{ Anzahl der Spalten ohne führende Eins in der red. ZSF von } A.$$

Übung 5.8: Finden Sie in Numbas eine Basis von $\operatorname{Ker} A$.

Wir untersuchen später im Kapitel noch die Dimensionen von Unterräumen etwas genauer.

B. Plus/Minus Theorem

Als nächstes zeigen wir ein Resultat, welches uns erlaubt, linear unabhängige Mengen zu vergrößern und Erzeugendensysteme zu verkleinern. Die Konsequenzen diese Resultats sind sehr wichtig für fast alles, was wir danach machen.

Theorem 5.9: (Plus/Minus Theorem)
Seien $v_1, \ldots, v_k$ mit $k \geqslant 1$ Vektoren eines Vektorraums V.

(i) Wenn $v_1, \ldots, v_k$ linear unabhängig sind und $v \in V$ nicht im Spann von $v_1, \ldots, v_k$ liegt, dann sind $v_1, \ldots, v_k, v$ immer noch linear unabhängig.

(ii) Wenn $v_1, \ldots, v_k$ ganz V aufspannen und ein v_i als Linearkombination der anderen Vektoren geschrieben werden kann, dann spannen $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$ immer noch ganz V auf.

BEWEIS. (i) Wir wollen zeigen, dass $v_1, \ldots, v_k, v$ linear unabhängig sind. Dazu betrachten wir

$$\lambda_1 v_1 + \cdots + \lambda_k v_k + \lambda v = 0.$$

Wenn $\lambda \neq 0$, dann ist $v = -\frac{1}{\lambda}(\lambda_1 v_1 + \cdots + \lambda_k v_k)$. Aber da v nicht im Spann von $v_1, \ldots, v_k$ liegt, kann dies nicht sein: es muss also $\lambda = 0$ sein. Dann haben wir nur noch die Linearkombination

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0.$$

Aber $v_1, \ldots, v_k$ sind linear unabhängig, also sind alle $\lambda_i = 0$. Also sind $v_1, \ldots, v_k, v$ linear unabhängig.
Zusammenfassung: betrachte v und $v_1, \ldots, v_k$ separat beim Zeigen der linearen Unabhängigkeit.

(ii) Wir wissen, dass $V = \langle v_1, v_2, \ldots, v_k \rangle$, und ein v_i als Linearkombination der anderen geschrieben werden kann. Ohne Beschränkung der Allgemeinheit (oBdA) (was hier bedeutet, dass wir sie wenn nötig umnummerieren können) haben wir $v_k = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1}$. Wir wollen zeigen, dass $v_1, \ldots, v_{k-1}$ immer noch V erzeugen. Für $w \in V$ gibt es $\mu_i \in \mathbb{R}$ mit

$$w = \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_k v_k,$$

weil $v_1, \ldots, v_k$ ein Erzeugendensystem von V ist. Dann können wir v_k ersetzen und bekommen

$$\begin{aligned} w &= \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_k(\lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1}) \\ &= (\mu_1 + \mu_k \lambda_1) v_1 + \cdots + (\mu_{k-1} + \mu_k \lambda_{k-1}) v_{k-1}. \end{aligned}$$

Also kann w als Linearkombination von $v_1, \ldots, v_{k-1}$ geschrieben werden. Da dies für alle $w \in V$ funktioniert, erzeugen $v_1, \ldots, v_{k-1}$ immer noch V. □
Zusammenfassung: ersetze v_k mit seiner Linearkombination.

Wir brauchen dieses Theorem für einige andere wichtige Resultate. Aber hier ist schon mal ein kleines Beispiel, wie man es auch benutzen kann.

Beispiele 5.10: a) Seien $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ und $v_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$. Wir sehen:
- ◇ v_1 und v_2 sind linear unabhängig: es sind nur zwei Vektoren und sie sind nicht kollinear.
- ◇ v_1 und v_2 haben 0 im dritten Eintrag, also liegen sie in der x, y-Ebene.
- ◇ v_3 ist nicht im Spann von v_1, v_2: der dritte Eintrag ist nicht Null.

Also folgt aus dem Plus/Minus Theorem (5.9), dass v_1, v_2, v_3 linear unabhängig sind.

b) Genauso können wir das Theorem auf Polynome $p_1 = 1 - x^2$, $p_2 = 2 - x^2$, $p_3 = x^3$ anwenden. Warum ist p_3 nicht im Spann von p_1, p_2?

c) Wenn v_1, v_2, v_3, v_4 einen Vektorraum V erzeugen und $v_3 = v_1 - 4v_2 + 3v_4$ ist (oder eine andere ähnliche Linearkombination), dann können wir v_3 entfernen, und v_1, v_2, v_4 erzeugen immer noch ganz V.

d) Wenn v_1, v_2, v_3, v_4 den Vektorraum V erzeugen und $v_2 = 2v_4$, dann können wir entweder v_2 oder v_4 entfernen (aber nicht beide!). Also erzeugen v_1, v_3, v_4 ganz V, aber ebenso erzeugen v_1, v_2, v_3 ganz V.

Eine der schönen Konsequenzen des Plus/Minus Theorems ist folgende: wenn wir schon die richtige Anzahl von Vektoren für eine Basis haben, reicht es, eine von den zwei Bedingungen für eine Basis zu überprüfen.

Proposition 5.11: (Zwei für Eins bei Basen)
Sei V ein Vektorraum mit Dimension n. Dann gilt:
(i) n linear unabhängige Vektoren erzeugen automatisch ganz V, und
(ii) ein Erzeugendensystem aus n Vektoren ist automatisch linear unabhängig.

Beweis. (i) Seien $v_1, \ldots, v_n$ linear unabhängig. Angenommen, sie erzeugen nicht V: dann gibt es ein $v \in V$ welches nicht im Spann von $v_1, \ldots, v_n$ liegt. Nach dem Plus/Minus Theorem (Theorem 5.9) sind $v_1, \ldots, v_n, v$ immer noch linear unabhängig. Aber dies sind mehr als n Vektoren, also können sie nicht linear unabhängig sein (nach Theorem 5.1). Also erzeugen $v_1, \ldots, v_n$ ganz V.

(ii) Angenommen $v_1, \ldots, v_n$ erzeugen V aber sind linear abhängig. Dann gibt es λ_i nicht alle Null mit

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0.$$

ObdA haben wir $\lambda_n \neq 0$. Aber dann kann v_n als Linearkombination der anderen Vektoren $v_1, \dots, v_{n-1}$ geschrieben werden, also erzeugen nach dem Plus/Minus Theorem (Theorem 5.9) die Vektoren $v_1, \dots, v_{n-1}$ immer noch ganz V. Aber dies sind weniger als n Vektoren, also können sie nicht V erzeugen (nach Theorem 5.1). Also sind $v_1, \dots, v_n$ linear unabhängig. □

Wenn wir also schon auf andere Weise die Dimension von V kennen (⚠ und nur dann!) und wir eine entsprechende Anzahl von Vektoren haben, die eine Basis sein sollen, brauchen wir nur *eins* von linear unabhängig und Erzeugendensystem zeigen, nicht beide.

Beispiele 5.12: a) Wir sehen ohne Rechnung, dass $v_1 = \binom{-3}{7}$, $v_2 = \binom{5}{5}$ eine Basis von $\mathbb{R}^2$ bilden. Die beiden Vektoren sind offensichtlich nicht kollinear, also sind sie linear unabhängig. Da $\mathbb{R}^2$ Dimension 2 hat, sind diese zwei linear unabhängigen Vektoren also auch eine Basis.

b) Wir sehen ohne Rechnung, dass

$$v_1 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}, \ v_2 = \begin{pmatrix} 4 \\ 0 \\ 7 \end{pmatrix}, \ v_3 = \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$$

eine Basis von $\mathbb{R}^3$ bilden. Die ersten beiden Vektoren liegen in der x, z-Ebene (können Sie erklären, warum?) und sind offenbar nicht kollinear. Der dritte Vektor v_3 liegt nicht in der x, z-Ebene, liegt also auch nicht im Spann von v_1, v_2. Also sind nach dem Plus/Minus Theorem (Theorem 5.9) die Vektoren v_1, v_2, v_3 linear unabhängig. Da $\mathbb{R}^3$ Dimension 3 hat, bilden sie automatisch eine Basis.

Das nächste Resultat werden wir oft benutzen.

Theorem 5.13: (Zu Basis ergänzen oder reduzieren)
Sei V ein endlich-dimensionaler Vektorraum. Dann gilt:

(i) Jede Liste von linear unabhängigen Vektoren in V kann zu einer Basis von V ergänzt werden, und

(ii) jedes Erzeugendensystem von V enthält eine Basis von V.

BEWEIS. (i) Angenommen $v_1, \dots, v_k$ sind linear unabhängig in V, mit $k < n = \dim V$. Wir wollen zeigen, dass wir Vektoren hinzufügen können, um eine Basis von V zu erhalten.
Da weniger als n Vektoren V nicht aufspannen können (Theorem 5.1), gibt es ein $v_{k+1} \in V$, das nicht im Spann von $v_1, \dots, v_k$ liegt. Nach dem Plus/Minus Theorem (Theorem 5.9) sind dann $v_1, \dots, v_k, v_{k+1}$ immer noch linear unabhängig. Wenn $k + 1 = n$, haben wir n linear unabhängige Vektoren, und diese sind dann eine Basis (nach „Zwei für eins“, Prop. 5.11). Wenn $k + 1 < n$, dann können $v_1, \dots, v_{k+1}$ immer noch nicht V erzeugen, also gibt es ein v_{k+2} welches nicht im Spann liegt, also sind $v_1, \dots, v_{k+1}, v_{k+2}$ immer noch linear unabhängig. So fahren wir fort, bis wir n linear unabhängige Vektoren haben, die dann eine Basis von V bilden.

(ii) Sei $v_1, \dots, v_m$ ein Erzeugendensystem, mit $m > n$. Da es mehr als n Vektoren sind, können sie nicht linear unabhängig sein (nach Theorem 5.1). Also gibt es λ_i nicht alle Null, sodass

$$\lambda_1 v_1 + \ldots + \lambda_m v_m = 0;$$

oBdA ist $\lambda_m \neq 0$. Dann kann v_m als Linearkombination der anderen Vektoren geschrieben werden, also folgt aus dem Plus/Minus Theorem (Theorem 5.9), dass $v_1, \dots, v_{m-1}$ immer noch V erzeugen. Wenn $m - 1 = n$, dann bilden die Vektoren automatisch eine Basis (nach „Zwei für eins“, Proposition 5.11). Wenn $m - 1 > n$, dann können wir den Schritt wiederholen und noch einen Vektor entfernen und haben immer noch ein Erzeugendensystem $v_1, \dots, v_{m-2}$. So fahren wir fort, bis wir ein Erzeugendensystem aus n Vektoren haben, was dann automatisch eine Basis ist. □

Beispiele 5.14: Dieses Beispiel nicht Teil des Moduls, nur für Interessierte.

Wir wollen die linear unabhängigen Vektoren

$$v_1 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \; v_2 = \begin{pmatrix} 0 \\ 2 \\ 6 \end{pmatrix}$$

zu einer Basis von $\mathbb{R}^3$ ergänzen. Wir sehen, dass ein Vektor mit von Null verschiedenem ersten Eintrag nicht im Spann von v_1, v_2 liegen kann. Also gibt uns die Ergänzung mit dem Vektor

$$v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

eine Basis v_1, v_2, v_3 von $\mathbb{R}^3$. Dieser weitere Vektor ist nicht eindeutig! Wir können jeden Vektor wählen, der einen von Null verschiedenen ersten Eintrag hat, zum Beispiel $v_3 = \left(\begin{smallmatrix} 4 \\ 3 \\ -5 \end{smallmatrix}\right)$. Aber um uns das Leben etwas leichter zu machen, ist es sinnvoll, nach geeigneten Standardbasisvektoren zu schauen, mit denen wir ergänzen können.

b) Wir suchen Standardbasisvektoren, die folgende zwei Vektoren zu einer Basis von $\mathbb{R}^4$ ergänzen.

$$v_1 = \begin{pmatrix} 1 \\ 3 \\ 6 \\ -3 \end{pmatrix}, \; v_2 = \begin{pmatrix} -1 \\ -2 \\ -4 \\ 2 \end{pmatrix}.$$

Hier sieht man nicht sofort, welche Standardbasisvektoren man wählen könnte oder sollte. Wir können entweder einfach etwas ausprobieren und dann prüfen, ob die resultierenden vier Vektoren linear unabhängig sind oder nicht. Oder wir können etwas genauer hinsehen: das Muster der unteren drei Einträge sieht bei beiden Vektoren gleich aus. Das gibt uns einen Ansatz zum Ausprobieren, und wirklich ist $2v_1 + 3v_2 = -e_1$. Also ist $e_1 \in \operatorname{Span}(v_1, v_2)$ und wir können diesen Vektor nicht hinzufügen. Wir fügen also e_2, e_3 hinzu und überprüfen, ob die resultierenden vier Vektoren wirklich linear unabhängig sind, indem wir das relevante LGS lösen.

$$\begin{pmatrix} \boxed{1} & -1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ 6 & -4 & 0 & 1 \\ -3 & 2 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

Hier sehen wir schon, dass die Vektoren linear unabhängig sind.
Also bilden v_1, v_2, e_2, e_3 eine Basis von $\mathbb{R}^4$.
(Wir hätten auch e_2, e_4 oder e_3, e_4 hinzufügen können. Rechnen Sie dies nach.)

c) Wir betrachten die Polynome

$$p_1 = x - 1, \; p_2 = x + 3$$

in $\mathbb{R}[x]_3$. Wir sehen, dass $p_3 = x^2$ nicht im Spann dieser zwei Polynome liegt. Nach dem the Plus/Minus Theorem sind also p_1, p_2, p_3 immer noch linear unabhängig. Wir können immer noch kein Polynom von Grad 3 erzeugen, also ergänzen wir noch $p_4 = x^3$, und dann bilden p_1, p_2, p_3, p_4 eine Basis von $\mathbb{R}[x]_3$.

Beim Ergänzen zu einer Basis wählt man am besten Standardbasisvektoren und probiert etwas aus. Aber um ein Erzeugendensystem zu einer Basis zu reduzieren gibt es eine schöne Methode. Dies ist ja auch, was man machen muss, um eine Basis eines Spaltenraums einer Matrix zu finden, also auch eine Basis des Bildes einer Matrixabbildung.

Basis von Bild/Spaltenraum finden oder **Erzeugendensystem zu Basis reduzieren**
Wir wollen ein Erzeugendensystem $v_1, \ldots, v_l$ im Vektorraum V mit Dimension n zu einer Basis reduzieren.

(Oder eine Basis des Spaltenraums einer Matrix/ des Bildes einer Matrixabbildung finden. Das fängt im vierten Punkt an.)

- Wenn $V = \mathbb{R}^n$ ist, haben wir schon Spaltenvektoren $v_1, \dots, v_l \in \mathbb{R}^n$.
- Wenn $V \neq \mathbb{R}^n$, dann wähle eine Basis $\mathcal{B}$ von V (möglichst eine schön einfache), und schreibe die v_i als Koordinatenvektoren: dann haben wir also wieder Spaltenvektoren in $\mathbb{R}^n$.
- Bilde eine Matrix A aus den obigen Spaltenvektoren.
- Benutze Gauss-Jordan, um A in (reduzierte) ZSF zu bringen.
- Die Spalten in der ursprünglichen Matrix A, die in der ZSF führende Einsen haben, bilden eine Basis von $\mathrm{SR}(A) = \mathrm{Im}\, f_A$.

Übung 5.15: a) Erklären Sie warum obige Methode eine Basis gibt:
- Warum spannen die gegebenen Spalten von A den ganzen Spaltenraum auf?
- Warum sind die gegebenen Spalten von A linear unabhängig?

Hierfür müssen Sie das LGS $Ax = 0$ und dessen Lösung auf eine bestimmte Weise interpretieren.
Hinweis für den zweiten Punkt: In den Umformungen von A zur red. ZSF, löschen Sie alle Spalten, die am Ende keine führende Eins haben. Was sagt Ihnen diese Rechnung dann? Wäre die Rechnung anders, wenn man die gelöschten Spalten von Anfang an gar nicht dabei hätte?

b) Finden Sie die Dimension eines Aufspanns in Numbas.

c) Finden Sie Basen für Kern und Bild einer Matrixabbildung in Numbas.

d) Finden Sie ein Beispiel einer nicht surjektiven linearen Abbildung, deren Kern Null ist. Überprüfen Sie Ihr Beispiel einer injektiven, nicht surjektiven Abbildung in Numbas.

Da jedes Erzeugendensystem eine Basis enthält, wissen wir nun auch auch:

Korollar 5.16: *Jeder endlich-dimensionale Vektorraum hat eine (endliche) Basis.*

BEWEIS. Sei V endlich-dimensional, also hat V ein endliches Erzeugendensystem. Dies enthält eine Basis (nach Thm 5.13(ii)). Diese Basis ist endlich, als Teilmenge einer endlichen Menge. □

Das Resultat erlaubt uns auch, Basen und Dimensionen von Unterräumen genauer anzuschauen.

C. Basen und Dimensionen von Unterräumen

Proposition 5.17: (Dimensionen von Unterräumen)
Wenn W ein Unterraum eines endlich-dimensionalen Vektorraums V ist, dann gilt:

(i) *W ist endlich-dimensional.*
(ii) $\dim W \leqslant \dim V$.
(iii) *$W = V$ genau dann wenn* $\dim W = \dim V$.

BEWEIS. (i) Da V endlich-dimensional ist, hat es eine endliche Basis. Sei n die Anzahl der Elemente in einer Basis, also $\dim V = n$.
Nun wollen wir zeigen, dass auch W ein endliches Erzeugendensystem hat. Wenn $W = \{0\}$, dann ist die leere Menge ein Erzeugendensystem (oder auch $\{0\}$), welches endlich ist. Wir nehmen also an, dass $W \neq \{0\}$. Sei also $w_1 \neq 0 \in W$. Wenn $\langle w_1 \rangle = W$ (also wenn w_1 ganz W erzeugt), dann haben wir ein endliches Erzeugendensystem und sind fertig.
Wenn $\langle w_1 \rangle \neq W$, dann gibt es $w_2 \in W$, das nicht in diesem Spann liegt: $w_2 \notin \langle w_1 \rangle$. Nach dem Plus/Minus Theorem (Theorem 5.9) ist also w_1, w_2 linear unabhängig. Wenn nun $\langle w_1, w_2 \rangle = W$, dann sind wir fertig. Andernfalls können wir $w_3 \in W$ finden mit $w_3 \notin \langle w_1, w_2 \rangle$, und so weiter.
Da die Vektoren, die wir nach und nach bekommen, immer linear unabhängig bleiben (nach dem Plus/Minus Theorem), können wir nicht mehr als n Vektoren bekommen, weil mehr als

n Vektoren in V auf jeden Fall linear abhängig sind (Prop. 5.1). Also kommt dieser Prozess zu einem Ende, und gibt uns somit ein endliches Erzeugendensystem für W.

(ii) **Übung.** *Hinweis:* Eine Basis von W ist linear unabhängig in V.

(iii) Wenn $W = V$, dann ist offensichtlich $\dim W = \dim V$.
Umgekehrt, angenommen $\dim W = \dim V = n$ und $W \subseteq V$ ist ein UVR. Angenommen es gibt ein $v \in V$, $v \notin W$. Sei $w_1, w_2, \dots, w_n$ eine Basis von W (die existiert, weil W nach (i) endlich-dimensional ist und somit eine endliche Basis hat). Wenn $v \notin W$, dann sind nach dem Plus/Minus Theorem (Theorem 5.9) die Vektoren $w_1, w_2, \dots, w_n, v$ immer noch linear unabhängig in V. Aber dies sind $n+1$ linear unabhängige Vektoren in einem Vektorraum von Dimension n, was nicht möglich ist (Prop. 5.1). Also gibt es kein $v \in V$ welches nicht in W ist, also ist $W = V$. $\square$

⚠ Vorsicht! Wenn W kein Unterraum von V ist, dann impliziert $\dim W = \dim V$ nicht $W = V$. Zum Beispiel der Vektorraum $\mathbb{R}[x]_2$ von Polynomen von Grad höchstens 2 hat Dimension 3, aber ist nicht gleich $\mathbb{R}^3$.

Das Verhältnis $\dim W \leqslant \dim V$ der Dimensionen zwischen Vektorraum und Unterraum ist recht simpel, aber das Verhältnis zwischen *Basen* ist komplizierter.

Beispiele 5.18: (Basen von Unterräumen)
Wenn $v_1, v_2, \dots, v_n$ eine Basis $\mathcal{B}$ von V ist, und $W \subseteq V$ ein UVR, dann gibt es keinen Grund, anzunehmen, dass eine Teilmenge der Basis $\mathcal{B}$ eine Basis von W gibt. Zum Beispiel:

- Sei $\mathcal{E}$ die Standardbasis $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ von $\mathbb{R}^2$, und sei der Unterraum $W = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$ die Gerade, die von $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ aufgespannt wird. Dann gibt es keine Teilmenge von $\mathcal{E}$, die eine Basis von W ist: Eine Basis von W muss ein einziger Vektor der Form $\begin{pmatrix} x \\ x \end{pmatrix}$ sein. (Zum Beispiel $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ oder $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ oder $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$.) e_1 und e_2 sind nicht einmal Elemente von W.
- Sei $W = \mathrm{Sym}_2$ der Vektorraum der symmetrischen 2×2-Matrizen, welches ein Unterraum von $V = \mathcal{M}_{2\times 2}$ ist. Die Standardbasis von $\mathcal{M}_{2\times 2}$ ist
$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$
Wir haben in Beispiel 5.5 gesehen, dass W eine Basis aus 3 Elementen hat, also ist $\dim W = 3$. Aber es gibt nur zwei symmetrische Matrizen in der Standardbasis von $\mathcal{M}_{2\times 2}$, also gibt es keine Teilmenge von 3 Elementen, die eine Basis von W geben könnte.

Wenn wir also zwei Unterräume U, W von V haben, können wir nicht einfach einen Schnitt der Basen nehmen, um eine Basis des Schnitts $U \cap W$ zu bekommen. Wie finden wir also eine Basis des Schnitts? Manchmal haben wir Glück und es geht doch:

Beispiele 5.19:
Bei $V = \mathbb{R}^3$, $U = \left\langle \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right) \right\rangle$, $W = \left\langle \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right) \right\rangle$ ist $U \cap W = \left\langle \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \right\rangle$.

Hier war die Basis des Schnitts leicht zu sehen.
Es gibt noch eine Situation, wo wir die Basis des Schnitts relativ leicht finden können.

Beispiele 5.20: $V = \mathbb{R}^3$, und
$$U = \left\{ \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \,\middle|\, x_1 + x_2 + x_3 = 0 \right\}, \quad W = \left\{ \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \,\middle|\, x_1 + 2x_2 + x_3 = 0 \right\}.$$

Hier sind U und W jeweils via ein LGS gegeben. Um die Elemente in $U \cap W$ zu finden, müssen wir einfach beide LGS zusammenfügen und gleichzeitig lösen: $U \cap W$ ist die Lösungsmenge von

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 0. \end{aligned}$$

Wir lösen:

$$\begin{pmatrix} \boxed{1} & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & \boxed{1} & 1 \\ 0 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{1} & 0 \end{pmatrix}$$

und sehen $U \cap W = \left\{ t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$. Also ist $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ eine Basis von $U \cap W$.

Aber manchmal ist es nicht ganz so leicht.

Beispiel 5.21: Wir suchen eine Basis vom Schnitt von

$$U = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\rangle, \quad W = \left\langle \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \right\rangle.$$

Sei $v \in U \cap W$. Dann ist

$$v = \begin{pmatrix} \lambda_1 \\ \lambda_1 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda_2 \\ 0 \\ 2\lambda_2 \end{pmatrix} = \begin{pmatrix} 2\mu_1 \\ \mu_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2\mu_2 \end{pmatrix}.$$

Aus der zweiten Zeile sehen wir $\lambda_1 = \mu_1$, und aus der dritte Zeile $\lambda_2 = \mu_2$. Also haben wir

$$v = \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 \\ 2\lambda_2 \end{pmatrix} = \begin{pmatrix} 2\lambda_1 \\ \lambda_1 \\ 2\lambda_2 \end{pmatrix}.$$

Nun gibt die erste Zeile $\lambda_1 = \lambda_2$. Also ist

$$v = \begin{pmatrix} 2\lambda_1 \\ \lambda_1 \\ 2\lambda_1 \end{pmatrix} = \lambda_1 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

Mit der Wahl $\lambda_1 = 1$ bekommen wir einen Vektor $\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$, der eine Basis von $U \cap W$ bildet.

Wir halten fest:

Basis eines Schnitts finden

Wenn U und W via LGS gegeben sind:

- Forme ein LGS aus beiden gegebenen LGS zusammen.
- Die Lösungsmenge des größeren LGS ist der Schnitt. Wir finden eine Basis, wie wir es bei dem Kern einer Matrix gewohnt sind.

Wenn U und W jeweils als Spann gegeben sind:

Sei $u_1, u_2, \ldots, u_k$ ein Erzeugendensystem von U und $w_1, w_2, \ldots, w_l$ ein Erzeugendensystem von W. Dann

- Ein $v \in U \cap W$ erfüllt $v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k = \mu_1 w_1 + \cdots + \mu_l w_l$.
- Löse das LGS $\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k - \mu_1 w_1 - \cdots - \mu_l w_l = 0$.
- Benutze die Lösung, um v aufzuschreiben, und finde wie oben eine Basis.

Hier müssen wir die Lösung sorgfältig interpretieren und uns an die Bedeutung aller Koeffizienten im System erinnern.

Übung 5.22: Oben finden wir nur Schnitte von UVR, also zum Beispiel von Geraden und Ebenen durch den Ursprung. Formulieren Sie, wie wir die Schritte leicht abändern können, um Schnitte von Geraden und Ebenen zu finden, die nicht durch den Ursprung gehen.

Hier ist ein Extrabeispiel für Sie, ganz durchgearbeitet, was wir nicht in der Vorlesung machen.

Beispiel 5.23: $V = \mathbb{R}^4$, $U = \left\langle u_1 = \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, u_2 = \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}\right\rangle$, $W = \left\langle w_1 = \begin{pmatrix}3\\1\\1\\0\end{pmatrix}, w_2 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}\right\rangle$.

Sei $v \in U \cap W$, also ist

$$v = \lambda_1 \begin{pmatrix}1\\1\\1\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}1\\-1\\0\\0\end{pmatrix} \quad \text{da } v \in U \quad \text{und} \quad v = \mu_1 \begin{pmatrix}3\\1\\1\\0\end{pmatrix} + \mu_2 \begin{pmatrix}1\\1\\0\\0\end{pmatrix} \quad \text{da } v \in W.$$

Also haben wir

$$v = \begin{pmatrix}\lambda_1+\lambda_2\\ \lambda_1-\lambda_2\\ \lambda_1\\ 0\end{pmatrix} = \begin{pmatrix}3\mu_1+\mu_2\\ \mu_1+\mu_2\\ \mu_1\\ 0\end{pmatrix} \quad \text{oder} \quad \begin{pmatrix}\lambda_1+\lambda_2\\ \lambda_1-\lambda_2\\ \lambda_1\\ 0\end{pmatrix} - \begin{pmatrix}3\mu_1+\mu_2\\ \mu_1+\mu_2\\ \mu_1\\ 0\end{pmatrix} = 0.$$

Dies gibt ein LGS mit vier Unbekannten und vier Gleichungen, welches wir lösen.

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

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

Hier ist μ_2 eine freie Variable. Mit $\mu_2 = t$ haben wir $\mu_1 = -t$, $\lambda_1 = \lambda_2 = -t$. Also ist eine allgemeine Lösung

$$v = \begin{pmatrix}-t-t\\ -t-(-t)\\ -t\\ 0\end{pmatrix} = -t\begin{pmatrix}2\\0\\1\\0\end{pmatrix}.$$

Also ist $v_0 = \left(\begin{smallmatrix}2\\0\\1\\0\end{smallmatrix}\right)$ eine Basis von $U \cap W$.

Wie können wir allgemeiner dieses Problem mit Basen von Unterräumen umgehen?

Wenn wir mit UVR $W \subseteq V$ und Basen arbeiten, fangen wir mit einer Basis $\mathcal{B}_W$ des Unterraums an. Diese ist linear unabhängig in V, also können wir sie zu einer Basis $\mathcal{B}_V$ von V ergänzen (möglich wegen Theorem 5.13). So bekommen wir eine Basis von V, die eine Basis des UVR enthält.
⚠ Damit dies funktioniert, müssen wir immer mit einer Basis des kleinsten UVR anfangen!

Wir werden dieses Prinzip im nächsten Abschnitt ausprobieren.

D. Lineare Abbildungen und Dimension

Die Dimensionen von Kern und Bild einer linearen Abbildung haben auch einen sehr schönen Zusammenhang.

Theorem 5.24: (Dimensionsformel für lineare Abbildungen)
Wenn $f: V \longrightarrow W$ eine lineare Abbildung ist und V endlich-dimensional, dann gilt

$$\dim \operatorname{Ker}(f) + \dim \operatorname{Im}(f) = \dim V.$$

BEWEIS. Sei $\dim V = n$. Sei $v_1, \dots, v_k$ eine Basis von $\operatorname{Ker} f$ (wir fangen mit einer Basis des UVR an), und ergänze sie zu einer Basis $v_1, \dots, v_n$ von V (möglich nach Theorem 5.13). Dann ist $f(v_1) = f(v_2) = \cdots = f(v_k) = 0$, da diese Vektoren im Kern von f liegen. Wir zeigen, dass $f(v_{k+1}), f(v_{k+2}), \dots, f(v_n)$ eine Basis von $\operatorname{Im} f$ bilden.
Die Vektoren erzeugen $\operatorname{Im} f$: für $w \in \operatorname{Im} f$ gibt es $v \in V$ mit $w = f(v)$. Da $v_1, \dots, v_n$ eine Basis von V ist, gibt es $\lambda_i \in \mathbb{R}$ mit $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Also ist

$$\begin{aligned} w = f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_n v_n) \\ &= \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) && \text{da } f \text{ linear} \\ &= 0 + \cdots + 0 + \lambda_{k+1} f(v_{k+1}) + \cdots + \lambda_n f(v_n). \end{aligned}$$

Also erzeugen $f(v_{k+1},) \dots, f(v_n)$ das Bild $\operatorname{Im} f$.
Um lineare Unabhängigkeit zu zeigen, betrachten wir

$$\mu_{k+1} f(v_{k+1}) + \cdots + \mu_n f(v_n) = 0 \quad \text{für } \mu_{k+1}, \dots, \mu_n \in \mathbb{R}.$$

Wegen Linearität ist dann $f(\mu_{k+1} v_{k+1} + \cdots + \mu_n v_n) = 0$, also ist $\mu_{k+1} v_{k+1} + \cdots + \mu_n v_n \in \operatorname{Ker} f$. Aber $v_1, \cdots, v_k$ ist eine Basis von $\operatorname{Ker} f$, also ist

$$\mu_{k+1} v_{k+1} + \cdots + \mu_n v_n = \mu_1 v_1 + \cdots \mu_k v_k.$$

Aber dann ist

$$\mu_1 v_1 + \cdots + \mu_k v_k - \mu_{k+1} v_{k+1} - \cdots - \mu_n v_n = 0.$$

Da $v_1, \dots, v_n$ eine Basis von V bilden, sind sie linear unabhängig. Also sind alle $\mu_j = 0$. Dies zeigt, dass $f(v_{k+1}), \dots, f(v_n)$ linear unabhängig sind. Also ist $\dim(\operatorname{Im} f) = n - k$, und somit

$$n = \dim V = k + (n-k) = \dim \operatorname{Ker}(f) + \dim \operatorname{Im}(f).$$ □

Beispiele 5.25: Schauen Sie die bisherigen Beispiele von linearen Abbildungen an (auch auf Übungsblättern) und überzeugen Sie sich, dass die Dimensionen von Kern und Bild in diesen Beispielen in der Tat diese Dimensionsformel erfüllen. Beginnen Sie mit den Beispielen a) und b) aus 2.37.

Definition 5.26: Eine bijektive lineare Abbildung $f: V \longrightarrow W$ heißt **(linearer) Isomorphismus**. Wenn es einen Isomorphismus zwischen V und W gibt, sagen wir, dass die Räume **isomorph** sind und schreiben $V \cong W$.

Beispiele 5.27:

a) Die Identität $\operatorname{id}: V \longrightarrow V$ auf einem beliebigen Vektorraum V ist ein Isomorphismus.
b) Die Streckung $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mit $f(v) = rv$ ist für $r \neq 0$ ein Isomorphismus.
c) Die Projektion $\pi: \mathbb{R}^2 \longrightarrow \mathbb{R}$ ist kein Isomorphismus: sie ist nicht injektiv.
d) Sei $\mathcal{B}$ eine Basis von V mit $\dim V = n$. Dann ist ${}_{\mathcal{B}}(-): V \longrightarrow \mathbb{R}^n$, die Abbildung, die Koordinatenvektoren gibt, ein Isomorphismus (siehe Übung 4.23).

Übung 5.28: Bestimmen Sie in Numbas, ob eine Abbildung injektiv, surjektiv, ein Isomorphismus ist. (Der letzte Teil der Aufgabe fragt nach der inversen Abbildung: das kommt im nächsten Kapitel. Sie können es jetzt noch auslassen.)

Durch den Dimensionssatz für lineare Abbildungen bekommen wir jetzt noch ein „zwei für eins“ Resultat.

Proposition 5.29: (Zwei für eins bei Isomorphismen)
Für eine lineare Abbildung $f\colon V \longrightarrow V$ von einem endlich-dimensionalen Vektorraum zu sich selbst gilt:

$$f \text{ ist injektiv} \iff f \text{ ist surjektiv} \iff f \text{ ist ein Isomorphismus.}$$

Beweis.

$$\begin{array}{rcl}
 & f \text{ injektiv} & \\
\Leftrightarrow & \operatorname{Ker} f = 0 & \text{nach Injektivität via Kern, Prop. 4.31} \\
\Leftrightarrow & \dim \operatorname{Ker}(f) = 0 & \\
\Leftrightarrow & \dim \operatorname{Im}(f) = \dim V & \text{nach Dimensionssatz, Theorem 5.24} \\
\Leftrightarrow & \operatorname{Im} f = V & \text{weil } \operatorname{Im} f \text{ UVR von } V \text{ (siehe Theorem 5.17)} \\
\Leftrightarrow & f \text{ surjektiv.} &
\end{array}$$

Dies beweist f injektiv genau dann wenn f surjektiv. Da ein Isomorphismus nach Defition eine lineare Abbildung ist, die sowohl injektiv als auch surjektiv ist, folgt auch die letzte Äquivalenz. □

Ein Isomorphismus hat auch eine inverse Abbildung.

Definition 5.30: Die **Inverse** einer linearen Abbildung $f\colon V \longrightarrow W$ ist eine lineare Abbildung $g\colon W \longrightarrow V$ mit

$$g{\circ}f = \mathrm{id}_V : V \longrightarrow V \qquad \text{und}$$
$$f{\circ}g = \mathrm{id}_W : W \longrightarrow W.$$

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[dr, "\mathrm{id}_V"'] & W \arrow[d, "g"] \arrow[dr, "\mathrm{id}_W"] & \\
 & V \arrow[r, "f"'] & W
\end{tikzcd}$$

Sie kennen schon inverse Abbildungen oder Umkehrabbildungen aus MAGL. Beachten Sie, dass in unserem Kurs Abbildungen immer auf jedem Element der Quelle definiert sind!

Proposition 5.31: (Inverse sind eindeutig.)
Wenn die lineare Abbildung $f\colon V \longrightarrow W$ invertierbar ist, dann ist die Inverse eindeutig.

Beweis. Übung. □

Notation 5.32: Da Inverse eindeutig sind, können wir „die Inverse“ statt „eine Inverse“ sagen, und schreiben f^{-1} für die Inverse, wenn sie existiert.

Fakt 5.33: Eine (lineare) Abbildung $f\colon V \longrightarrow W$ ist invertierbar genau dann wenn f bijektiv ist.

In MAGL war es anders in Satz 2.2.32: dort reicht im Prinzip nur Injektivität. Aber weil bei uns die Inverse auch auf ganz W definiert sein muss, brauchen wir auch Surjektivität.

Beweis. Für Interessierte; nicht in der Vorlesung
Wenn f invertierbar ist, dann gilt $f{\circ}f^{-1} = \mathrm{id}_W$ und $f^{-1}{\circ}f = \mathrm{id}_V$.
Wir zeigen, dass f injektiv ist: Wenn $f(u) = f(v)$, dann folgt $f^{-1}(f(u)) = f^{-1}(f(v))$, also $u = v$.
Wir zeigen, dass f surjektiv ist: Für jedes $w \in W$ gibt es $f^{-1}(w) \in V$ mit $f(f^{-1}(w)) = w$.
Also ist f bijektiv.
Umgekehrt, angenommen f ist bijektiv, also sowohl injektiv als auch surjektiv. Wir können die Inverse $f^{-1}\colon W \longrightarrow V$ definieren durch $f^{-1}(w) = v$ für $v \in V$ mit $f(v) = w$. Solch ein v existiert

weil f surjektiv ist, und es ist eindeutig weil f injektiv ist. Durch diese Definition bekommen wir offensichtlich $f \circ f^{-1} = \mathrm{id}_W$ und $f^{-1} \circ f = \mathrm{id}_V$. □

Proposition 5.34: *Die Inverse eines Isomorphismus ist auch ein Isomorphismus.*

Beweis. Es bleibt zu zeigen, dass die Inverse einer invertierbaren linearen Abbildung auch linear ist.
Angenommen $f(v_1) = w_1$ und $f(v_2) = w_2$. Dann ist $f^{-1}(w_1) = v_1$ und $f^{-1}(w_2) = v_2$.
Da f linear ist, gilt $f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 w_1 + \lambda_2 w_2$. Durch die Definition der Inversen gilt also $f^{-1}(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 f^{-1}(w_1) + \lambda_2 f^{-1}(w_2)$.
Also ist f^{-1} auch linear. □

Zusammenfassung: die Eindeutigkeit der Urbilder unter f, die die Definition von f^{-1} zulässt, überträgt auch automatisch die Linearität von f auf Linearität von f^{-1}.

Beispiele 5.35: Aus Beispiele 5.27:

a) Die Identität $\mathrm{id}\colon V \longrightarrow V$ ist selbst-invers.
b) Die Streckung $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mit $f(v) = rv$ für $r \neq 0$ hat Inverse $f^{-1}(v) = \frac{1}{r}v$.
c) Die Abbildung ${}_{\mathcal{B}}(-)\colon V \longrightarrow \mathbb{R}^n$, die Koordinatenvektoren gibt, hat als Inverse die Abbildung $x \longmapsto \sum_i x_i b_i$, wobei x_i die Einträge von $x \in \mathbb{R}^n$ sind, und b_i die Vektoren der Basis $\mathcal{B}$.

Das letzte Beispiel sagt uns auch:

Theorem 5.36: (Dimension n bedeutet isomorph zu $\mathbb{R}^n$.)
Jeder reelle Vektorraum von Dimension n ist isomorph zu $\mathbb{R}^n$.

Beweis. Sei V ein Vektorraum von Dimension n und $\mathcal{B}$ eine Basis von V. Dann ist die Abbildung ${}_{\mathcal{B}}(-)\colon V \longrightarrow \mathbb{R}^n$ eine Isomorphismus (siehe Beispiele 5.27). □

Übung 5.37: Prüfen Sie die Aussage in den Beispielen nach, indem Sie direkt zeigen, dass die Inverse, die in Beispiel 5.35c) gegeben ist, wirklich eine Inverse ist.

Isomorphismen sind etwas besonderes, weil sie Inverse haben, aber auch, weil sie Basen auf Basen schicken.

Proposition 5.38: (Isos erhalten Basen.)
Wenn $f\colon V \longrightarrow W$ ein Isomorphismus zwischen endlich-dimensionalen Vektorräumen ist, und $v_1, \ldots, v_n$ eine Basis von V bilden, dann bilden $f(v_1), \ldots, f(v_n)$ eine Basis von W.

Beweis. Aus Proposition 4.32 folgt: Weil f injektiv ist, sind $f(v_1), \ldots, f(v_n)$ injektiv in W. Weil f surjektiv ist, sind $f(v_1), \ldots, f(v_n)$ ein Erzeugendensystem von W. Daher bilden sie eine Basis von W. □

Insbesondere bedeutet dies: Wenn $f\colon V \longrightarrow W$ ein Isomorphismus zwischen endlich-dimensionalen Vektorräumen ist, dann ist $\dim V = \dim W$.
Wir bekommen sogar die Umkehrung.

Korollar 5.39: (Isomorph ⇔ gleiche Dimension.)
Zwei Vektorräume V und W sind isomorph genau dann wenn sie die gleiche Dimension haben.

Beweis. Eine Richtung haben wir oben schon gezeigt.
Umgekehrt, wenn $\dim V = \dim W = n$, dann sind nach Theorem 5.36 beide Vektorräume zu $\mathbb{R}^n$ isomorph. ALso gibt es Isomorphismen $f\colon V \longrightarrow \mathbb{R}^n$ und $g\colon W \longrightarrow \mathbb{R}^n$. Da die Inverse eines

Isomorphismus wieder ein Isomorphismus ist (Prop. 5.34) und die Komposition von Isomorphismen wieder ein Isomorphismus ist (Prop. 4.30), haben wir den Isomorphismus $g^{-1}\circ f\colon V \longrightarrow W$. $\square$

Übung 5.40: Finden Sie einen alternativen Beweis für $V \cong W \Rightarrow \dim V = \dim W$, der den Dimensionssatz für lineare Abbildungen benutzt.

Korollar 5.41: *Eine Matrix mit invertierbarer Matrixabbildung ist quadratisch.*

BEWEIS. Für eine $m \times n$-Matrix A ist die zugehörige Matrixabbildung $f_A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Aber f_A Isomorphismus impliziert $\dim \mathbb{R}^n = \dim \mathbb{R}^m$, also $m = n$. $\square$

Da die Komposition von injektiven Abbildungen injektiv und die Komposition von surjektiven Abbildungen surjektiv ist (Prop. 4.30), ist sofort klar, dass die Komposition von Isomorphismen wieder ein Isomorphismus ist. Aber es gibt noch eine schöne Art, die Inverse zu finden.

Lemma 5.42: (Socken und Schuhe)
Wenn $f\colon U \longrightarrow V$ und $g\colon V \longrightarrow W$ Isomorphismen sind, dann ist $(g\circ f)^{-1} = f^{-1}\circ g^{-1}$.

Wenn man zuerst die Socken und dann die Schuhe anzieht, muss man beim rückgängig machen erst die Schuhe ausziehen und dann die Socken.

$$\begin{tikzcd}
U \arrow[r, "f"] \arrow[ddrr, bend right, "\mathrm{id}_U"'] & V \arrow[d, "g"] \arrow[dr, "\mathrm{id}_V"] & \\
 & W \arrow[r, "g^{-1}"'] & V \arrow[d, "f^{-1}"] \\
 & & U
\end{tikzcd}$$

BEWEIS. Da Inverse eindeutig sind, reicht es zu zeigen, dass $(g\circ f)\circ(f^{-1}\circ g^{-1}) = \mathrm{id}_W$ und $(f^{-1}\circ g^{-1})\circ(g\circ f) = \mathrm{id}_U$. Wir haben

$$(f^{-1}\circ g^{-1})\circ(g\circ f) = f^{-1}\circ(g^{-1}\circ g)\circ f = f^{-1}\circ \mathrm{id}_V \circ f = f^{-1}\circ f = \mathrm{id}_U$$

und genauso für die andere Reihenfolge. $\square$

E. Basen und Dimension: Study guide

Konzeptübersicht.

- ◇ Dimension eines Vektorraums
- ◇ Mögliche Größen von linear unabhängigen Mengen, Erzeugendensystemen, Basen in gegebenem Vektorraum
- ◇ Beziehungen zwischen den Konzepten lineare Unabhängigkeit, Erzeugendensystem, Basis und Dimension
- ◇ Plus/Minus Theorem
- ◇ Zwei für Eins bei Basen
- ◇ Zu Basis ergänzen oder reduzieren
- ◇ Dimension von Unterräumen
- ◇ Zusammenhang der Dimensionen von Kern und Bild
- ◇ Isomorphismen und Inverse
- ◇ Eindeutigkeit der Inversen
- ◇ Zusammenhang zwischen Injektivität und Surjektivität einer linearen Abbildung von VR zu sich selbst
- ◇ Zusammenhänge von Isomorphismen, Dimensionen und Basen
- ◇ Bedingung unter welcher zwei VR isomorph sind
- ◇ Inverse einer Komposition von invertierbaren Abbildungen

Skills.

- ◇ Dimension eines Vektorraums oder Unterraums finden
- ◇ Dimension benutzen, um zu entscheiden, ob gegebene Vektoren eine Basis eines endlichdimensionalen Vektorraums bilden
- ◇ Erzeugendensystem zu Basis reduzieren
- ◇ Basis für Kern und Bild einer linearen Abbildung finden
- ◇ (Basis für Spaltenraum einer Matrix finden — ist das selbe wie Bild einer linearen Abbildung)
- ◇ Basis und Dimension von Schnitt von UVR finden
- ◇ Dimensionen zur Hilfe nehmen, um finden von (Basen von) Kern und Bild zu erleichtern
- ◇ Bestimmen, ob (und zeigen, dass) eine Abbildung ein Isomorphismus ist
- ◇ Isomorphismus zwischen VR der gleichen Dimension finden

KAPITEL 6

Inverse Matrizen und Determinanten

Lineare Abbildungen sind eine Interpretation von Matrizen. Wie überträgt sich also die Invertierbarkeit von Abbildungen auf die Matrizen? Und was für Konsequenzen hat dies auf die Interpretation von Matrizen als LGS?

A. Inverse Matrizen

Definition 6.1: Eine **Inverse** einer quadratischen Matrix A ist eine Matrix B der gleichen Größe, die $BA = I = AB$ erfüllt. Hierbei ist I die Einheitsmatrix der gleichen Größe.

Vergleichen Sie mit inversen Elementen in Gruppen oder Halbgruppen, aus MAGL.

Bemerkung 6.2: Vergleichen Sie dies mit der Komposition von Matrizabbildungen (Prop. 1.62): $f_A \circ f_B = f_{AB}$. Wenn also $AB = I$, dann ist $f_A \circ f_B = f_I = \mathrm{id}$. Ebenso gibt $BA = I$ auch $f_B \circ f_A = \mathrm{id}$, also sind f_A und f_B zueinander inverse Isomorphismen.

Die Bedingung ist in A und B symmetrisch. Wir sagen auch, dass A und B zueinander invers sind.

Übung 6.3: Zeigen Sie, dass $A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$ und $B = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ zueinander invers sind.

Beispiele 6.4:
- Die Inverse einer Einheitsmatrix ist die Matrix selbst: $I_n I_n = I_n$.
- Die Inverse einer Diagonalmatrix mit von Null verschiedenen Diagonaleinträgen ist die Diagonalmatrix mit den Kehrwerten auf der Diagonale:
$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} \begin{pmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$
und anders herum. (Hier ist $d_i \neq 0$ für $i \in \{1, \dots, n\}$.)

Nicht-Beispiel 6.5: Wie nicht alle Abbildungen invertierbar sind, haben auch nicht alle Matrizen eine Inverse: Zum Beispiel ist eine Diagonalmatrix mit einer Nullzeile nicht invertierbar: Es gibt keine Zahl x mit $0 \cdot x = 1$.

Übung 6.6: Invertieren Sie eine Diagonalmatrix in Numbas.

Formel 6.7: Eine $2{\times}2$-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad{-}bc \neq 0$ hat Inverse $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

BEWEIS.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & -ab+ab \\ cd-cd & -cb+ad \end{pmatrix}$$

also ist

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

und anders herum. □

Offensichtlich brauchen wir hierfür $ad - bc \neq 0$. Wir werden später sehen, dass eine 2×2-Matrix mit $ad - bc = 0$ nicht invertierbar ist.

Übung 6.8: Invertieren Sie eine 2×2-Matrix in Numbas.

Definition 6.9: Eine Matrix heißt **invertierbar** oder **regulär**, wenn sie eine Inverse hat. Eine Matrix, die nicht invertierbar ist, heißt **singulär**.

Wir erinnern uns von Kor. 5.41, dass nur quadratische Matrizen überhaupt die Chance haben, invertierbar zu sein.
Genau wie inverse Abbildungen sind auch inverse Matrizen eindeutig.

Proposition 6.10: (Eindeutigkeit von Inversen)
Wenn eine Matrix eine Inverse hat, dann ist diese Inverse eindeutig.

BEWEIS. Angenommen eine Matrix A hat zwei Inverse, B und C. Dann ist

$$\begin{aligned} B &= BI && \text{Mult. mit Einheitsmatrix ändert die Matrix nicht} \\ &= B(AC) && \text{da } C \text{ eine Inverse ist} \\ &= (BA)C && \text{Matrixmult. ist assoziativ} \\ &= IC && \text{da } B \text{ eine Inverse ist} \\ &= C \end{aligned}$$

□

Wegen dieses Results sagen wir *die* Inverse (statt eine Inverse) von A ist A^{-1}. Also gilt

$$AA^{-1} = I = A^{-1}A.$$

Vergleichen Sie den Beweis oben mit der Eindeutigkeit von Inversen in Gruppen, Satz 4.2.8 aus MAGL: dies ist derselbe Beweis. Hier liegt auch eine Gruppe zugrunde: siehe Übungsblatt.

Proposition 6.11: (Inverse Matrixabbildung)
Wenn A eine invertierbare Matrix ist, dann ist die Matrixabbildung, die zur inversen Matrix A^{-1} gehört, die inverse Abbildung der Matrixabbildung, die zu A gehört: $f_A^{-1} = f_{A^{-1}}$.

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "f_A"] \arrow[dr, "\mathrm{id}"'] & \mathbb{R}^n \arrow[d, "f_{A^{-1}}"'] \arrow[dr, "\mathrm{id}"] & \\ & \mathbb{R}^n \arrow[r, "f_A"'] & \mathbb{R}^n \end{tikzcd}$$

BEWEIS. Sei A eine $n \times n$-Matrix. Dann ist $f_{A^{-1}}(f_A(v)) = A^{-1}Av = Iv = v$ für jedes $v \in \mathbb{R}^n$, und $f_A(f_{A^{-1}}(v)) = AA^{-1}v = Iv = v$ für jedes $v \in \mathbb{R}^n$. □

Die inverse Matrixabbildung ist die Matrixabbildung der Inversen.
Wir werden später sehen, dass dies sogar eine „genau dann wenn“ Aussage ist: A ist invertierbar genau dann wenn f_A ein Isomorphismus ist.

Die Eindeutigkeit der Inversen erlaubt uns auch zu zeigen:

Proposition 6.12: (Eigenschaften von Inversen)

(i) Jede invertierbare Matrix ist die Inverse ihrer Inversen: $(A^{-1})^{-1} = A$

(ii) Wenn A invertierbar ist, dann auch A^T, und $(A^T)^{-1} = (A^{-1})^T$.

*(iii) **("Socken und Schuhe")** Wenn zwei $n \times n$-Matrizen A und B beide invertierbar sind, dann auch ihr Produkt AB, mit Inverser $(AB)^{-1} = B^{-1}A^{-1}$.*

BEWEIS. (i) Folgt aus der Eindeutigkeit von Inversen und $AA^{-1} = I = A^{-1}A$.

(ii) Wir transponieren diese beiden Gleichungen und erhalten

$$(AA^{-1})^T = I^T, \qquad \text{also} \qquad (A^{-1})^T A^T = I,$$

und

$$(A^{-1}A)^T = I^T, \qquad \text{also} \qquad A^T(A^{-1})^T = I.$$

Also erfüllt $(A^{-1})^T$ die Definition einer Inversen von A^T, also ist wegen der Eindeutigkeit von Inversen $(A^T)^{-1} = (A^{-1})^T$.

(iii) Erste Version (direkt): Wir prüfen, ob $B^{-1}A^{-1}$ die Definition der Inversen von AB erfüllt:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

$$\text{und} \qquad (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

Also folgt aus der Eindeutigkeit von Inversen, dass $(AB)^{-1} = B^{-1}A^{-1}$.
Zweite Version (via Abbildungen): Da $f_A \circ f_B = f_{AB}$, haben wir

$$f_{(AB)^{-1}} = f_{AB}^{-1} = (f_A \circ f_B)^{-1} = f_B^{-1} \circ f_A^{-1} = f_{B^{-1}} \circ f_{A^{-1}} = f_{B^{-1}A^{-1}}.$$

Hier benutzen wir, dass die inverse Matrixabbildung die Matrixabbildung der Inversen ist (Prop 6.11). □

Wir notieren dieses letzte Resultat auch als:

Das Produkt invertierbarer Matrizen ist invertierbar.

Beispiel 6.13: ⋄ Inverse einer Inversen: $\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^{-1} = -\frac{1}{3}\begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$, und

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}^{-1} = -\frac{1}{\frac{1}{3}}\begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$$

⋄ Inverse einer Transponierten: $\left(\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^T\right)^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}^T$

⋄ Inverse eines Produkts: $\left(\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}^{-1} = -\frac{1}{6}\begin{pmatrix} -2 & -4 \\ -1 & 1 \end{pmatrix}$

aber

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{-1} = -\frac{1}{3}\begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} -2 & -2 \\ -2 & 1 \end{pmatrix}.$$

Andererseits ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \left(-\frac{1}{3}\right)\begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} -2 & -4 \\ -1 & 1 \end{pmatrix}.$$

Übung 6.14: Benutzen Sie das vorherige Resultat, um in Numbas einen Audruck mit Inversen zu vereinfachen.

Manche Matrizen sind offensichtlich nicht invertierbar:

Proposition 6.15: *Eine Matrix mit einer Nullspalte oder mit einer Nullzeile ist nicht invertierbar.*

BEWEIS. Übung. $\square$

Übung 6.16: Zeigen Sie, dass eine 2×2-Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad - bc = 0$ nicht invertierbar ist.
Hinweis: Sie müssen wahrscheinlich verschiedene Fälle betrachten. Die Position dieser Übung im Skript ist auch ein Hinweis.

Wir können hier auch festhalten:

Bemerkung 6.17: Weil $AA^{-1} = I$ und $A^{-1}A = I$, ist das Multiplizieren einer Matrixgleichung mit einer invertierbaren Matrix eine Äquivalenzumformung:

$$AX = B \Leftrightarrow X = A^{-1}B \qquad \text{für beliebige Matrizen } X, B.$$

Aber dies gilt wirklich nur bei Multiplikation mit einer *invertierbaren* Matrix!

⚠ Wichtig!!! Wenn Sie mit Matrixgleichungen (einschließlich Vektoren) arbeiten, müssen Sie immer genau sagen, ob sie eine Gleichung mit einer Matrix von links oder von rechts multiplizieren! Da Matrixmultiplikation nicht kommutativ ist, macht dies einen Unterschied.
Nun zu linearen Gleichungssystemen.

Proposition 6.18: *Ein LGS $Ax = b$ mit invertierbarer Matrix A hat eine eindeutige Lösung.*

BEWEIS. Da A invertierbar ist, können wir die Matrixgleichung $Ax = b$ von links mit A^{-1} multiplizieren:

$$Ax = b \qquad \Leftrightarrow \qquad A^{-1}Ax = A^{-1}b \qquad \Leftrightarrow \qquad x = A^{-1}b$$

Wegen der Äquivalenzen ist $x = A^{-1}b$ die einzige (also eindeutige) Lösung. $\square$

Dies zeigt uns auch, dass für eine invertierbare Matrix A das LGS $Ax = b$ *für jedes* b lösbar ist!

Bemerkung 6.19: Wenn wir schon wissen, dass A invertierbar ist, und wir haben $AB = I$, dann folgt sofort, dass $B = A^{-1}$, durch Multiplikation mit A^{-1} von links. Vergleichen Sie mit MAGL Satz 4.2.1 (Links-Inverse sind Rechts-Inverse). Aber wenn wir noch *nicht* wissen, ob A überhaupt invertierbar ist, können wir dieses Argument nicht anwenden!
Wir brauchen noch etwas mehr Wissen, bis wir auch zeigen können, dass eine Richtung $AB = I$ genug für die Existenz einer Inversen ist.

Das Resultat über LGS mit invertierbaren Matrizen (Prop. 6.18) sagt uns etwas über die reduzierte ZSF, die eine invertierbare Matrix haben muss: sie muss n führende Einsen haben (wenn A eine $n \times n$-Matrix ist), und muss somit die Einheitsmatrix sein.
Dies gibt uns einen Algorithmus zur Berechnung einer inversen Matrix.

B. Invertierungsalgorithmus

Wir können die Inverse einer Matrix mit dem folgenden Algorithmus berechnen:

Der **Invertierungsalgorithmus** angewendet auf eine quadratische Matrix A hat folgende Schritte:

- Schreibe die Matrix A und die Einheitsmatrix derselben Größe nebeneinander.
- Führe den Gauss-Jordan Algorithmus auf A aus, und führe dieselben Zeilenumformungen auf der Einheitsmatrix daneben aus.
- Wenn die reduzierte ZSF der Matrix A die Einheitsmatrix ist, dann ist die Matrix, die nun rechts daneben steht, die Inverse A^{-1}.

Wir werden beweisen müssen, dass wir wirklich immer die inverse Matrix erhalten. Aber wir schauen uns zuerst ein paar Beispiele an.

Beispiel 6.20: Wir berechnen die Inverse von $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$.

$$\left.\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}\right| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{\mathrm{II}-2\mathrm{I},\,\mathrm{III}-\mathrm{I}} \left.\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -3 \\ 0 & -2 & 5 \end{pmatrix}\right| \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{\mathrm{III}+2\mathrm{II}} \left.\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -3 \\ 0 & 0 & -1 \end{pmatrix}\right| \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -5 & 2 & 1 \end{pmatrix}$$

$$\xrightarrow{(-1)\cdot\mathrm{III}} \left.\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}\right| \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & -1 \end{pmatrix}$$

$$\xrightarrow{\mathrm{II}+3\mathrm{III},\,\mathrm{I}-3\mathrm{III}} \left.\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right| \begin{pmatrix} -14 & 6 & 3 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}$$

$$\xrightarrow{\mathrm{I}-2\mathrm{II}} \left.\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right| \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}$$

Also ist $A^{-1} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}$.

Wir sollten immer die Probe machen: man verrechnet sich sehr leicht, und die Probe kann dies aufzeigen. Das Produkt AA^{-1} sollte uns I geben:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}\begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix} = \begin{pmatrix} -40+26+15 & 16-10-6 & 9-6-3 \\ -80+65+15 & 32-25-6 & 18-15-3 \\ -40+40 & 16-16 & 9-8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Also ist unsere inverse Matrix korrekt.

Übung 6.21: Benutzen Sie den Invertierungsalgorithmus, um $A = \begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & 0 \\ -1 & 2 & 6 \end{pmatrix}$ zu invertieren. Sie können dazu den Invertierungsrechner benutzen.

Wir wissen via Resultate über die Anzahl von Lösungen von LGS, dass bei einer invertierbaren Matrix die Einheitsmatrix als reduzierte ZSF herauskommen muss. Aber warum ist die Matrix auf der rechten Seite dann die Inverse?
Unser Weg, dies zu zeigen, verbindet elementare Zeilenumformungen mit bestimmten Matrizen, die wir später auch noch mal brauchen werden.

Definition 6.22: Es gibt drei Typen von **Elementarmatrizen**, die den elemenentaren Zeilenumformungen entsprechen.

- Eine Diagonalmatrix $M_i(\lambda)$ mit Einsen auf der Diagonale, bis auf ein $\lambda \neq 0$ in der $i$ten Zeile, erzeugt von der Einheitsmatrix durch Multiplikation der $i$ten Zeile mit λ.
- Eine Matrix T_{ij} mit Einsen auf der Diagonalen und sonst Nullen, bis auf eine 1 in der $i$ten Zeile in Spalte j statt auf der Diagonalen, und eine 1 in der $j$ten Zeile in Spalte i statt auf der Diagonalen, erzeugt von der Einheitsmatrix durch den Tausch der Zeilen i und j.
- Eine Matrix $A_{ij}(\lambda)$ mit Einsen auf der Diagonalen und zusätzlich einem λ an Stelle i, j, und sonst Nullen, erzeugt von der Einheitsmatrix durch Addition von λ mal Zeile j zu Zeile i.

$$M_i(\lambda) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \quad T_{ij} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \quad A_{ij}(\lambda) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

Proposition 6.23: *Die Multiplikation einer Matrix von links mit einer Elmentarmatrix hat den Effekt der dazugehörigen elementaren Zeilenumformung.*

BEWEIS. Lassen wir aus. Sie können dies bei 2×2 oder 3×3-Matrizen ausprobieren, um sich zu überzeugen. Interessierte kann ich mit einer Aufgabe versorgen, die dies (mit Hinweisen) formal zeigt. □

Beispiel 6.24: Seien $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$, $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$.
E_1, E_2, E_3 sind Elementarmatrizen, und wir haben

$$E_1 A = \begin{pmatrix} a & b & c \\ d & e & f \\ 2g & 2h & 2i \end{pmatrix}, \quad E_2 A = \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}, \quad E_3 A = \begin{pmatrix} a & b & c \\ d & e & f \\ g+4a & h+4b & i+4c \end{pmatrix}.$$

Dies bedeutet, wenn wir die elementaren Zeilenumformungen im Invertierungsalgorithmus machen, dann multiplizieren wir sowohl A also auch I sukzessive mit Elementarmatrizen von links.

$$\begin{array}{cc} & A \mid I \\ \longrightarrow & E_1 A \mid E_1 I \\ \longrightarrow & E_2 E_1 A \mid E_2 E_1 I \\ \vdots & \vdots \mid \vdots \\ \longrightarrow & E_k \cdots E_2 E_1 A \mid E_k \cdots E_2 E_1 I \end{array}$$

Wenn wir fertig sind, haben wir $(E_k \cdots E_2 E_1) A = I$ links, also ist das Produkt $E_k \cdots E_2 E_1 = B$ der Elementarmatrizen ein guter Kandidat für die Inverse A^{-1}. Und genau diese Matrix steht rechts: $E_k \cdots E_2 E_1 I = B$.

Um zu beweisen, dass dieses B wirklich die Inverse von A ist, müssen wir entweder einen Weg finden, die andere Reihenfolge AB auszurechnen (dies muss auch I ergeben), oder einen anderen Weg finden, dies zu zeigen.

Lemma 6.25: *Alle Elementarmatrizen sind invertierbar.*

BEWEIS. Die Inverse einer Elementarmatrix ist die Elementarmatrix, die zur zugehörigen inversen Zeilenumformung gehört.

$$\begin{aligned} M_i(\lambda)^{-1} &= M_i(\lambda^{-1}) && \text{teile Zeile i durch } \lambda \ (\lambda \neq 0) \\ T_{ij}^{-1} &= T_{ij} && \text{selbst-invers: tausche die Zeilen zurück} \\ A_{ij}(\lambda) &= A_{ij}(-\lambda) && \text{subtrahiere } \lambda \text{ mal Zeile } j \text{ von Zeile } i \end{aligned}$$ □

Nun können wir zeigen

Proposition 6.26: (Invertierungsalgorithmus funktioniert)
Wenn die reduzierte ZSF von A die Einheitsmatrix I ist, dann produziert der Invertierungsalgorithmus wirklich die inverse Matrix von A.

BEWEIS. Der Invertierungsalgorithmus produziert eine Matrix $B = E_k \cdots E_2 E_1$ mit $BA = I$. Da jede Elementarmatrix invertierbar ist (Lemma 6.25) und das Produkt von invertierbaren Matrizen invertierbar ist (Prop. 6.12, „Socken und Schuhe"), ist die Matrix B invertierbar mit $B^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$, und daher haben wir

$$BA = I \quad \Rightarrow \quad B^{-1}BA = B^{-1}I \quad \Rightarrow \quad A = B^{-1} \quad \Rightarrow \quad A^{-1} = B \quad \text{da } (B^{-1})^{-1} = B.$$

Also ist wirklich $A^{-1} = E_k \cdots E_2 E_1$, welches die Matrix auf der rechten Seite am Ende des Algorithmuses ist. □

Nun haben wir einige Bedingungen, die uns sagen, ob A invertierbar ist.

Theorem 6.27: (Invertierbarkeitsbedingungen)
Für eine quadratische Matrix A sind folgende Aussagen äquivalent:

(i) *A ist invertierbar.*
(ii) $\operatorname{Ker} A = \{0\}$.
(iii) *$Ax = 0$ hat nur die triviale Lösung $x = 0$.*
(iv) *Die reduzierte ZSF von A ist die Einheitsmatrix I.*
(v) *A ist das Produkt von Elementarmatrizen.*

BEWEIS. Da (ii) und (iii) die gleiche Aussage, nur anders formuliert sind, beweisen wir (nur) den Zykelschluss (i) ⇒(iii)⇒(iv) ⇒(v)⇒(i). Sei A eine $n \times n$-Matrix.

◇ (i) ⇒(iii): Dies war Prop. 6.18: wenn A invertierbar ist, dann gilt

$$Ax = 0 \quad \Leftrightarrow \quad A^{-1}Ax = A^{-1}0 \quad \Leftrightarrow \quad x = 0,$$

weil jede Matrix mal den Nullvektor gleich dem Nullvektor ist.

◇ (iii) ⇒(iv): Wenn $Ax = 0$ nur die triviale Lösung hat, dann zeigt der Beweis von Prop. 3.27 (Lösungen eines homogenen LGS), dass die reduzierte ZSF von A n führende Einsen hat. Da A auch n Zeilen hat, ist die red. ZSF also I_n.

◇ (iv) ⇒(v): Der Invertierungsalgorithmus (oder Gauss-Jordan) produziert $E_k \cdots E_2 E_1 A$ als die reduzierte ZSF von A. Wenn dies die Einheitsmatrix ist, dann ist $E_k \cdots E_2 E_1 A = I$, und daher $A = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$, da alle Elementarmatrizen invertierbar sind (Lemma 6.25). Die Inverse einer Elementarmatrix ist wieder eine Elementarmatrix, also ist A das Produkt von Elementarmatrizen.

◇ (v)⇒(i): Wenn A das Produkt von Elementarmatrizen ist, dann ist A invertierbar: das Produkt invertierbarer Matrizen ist invertierbar (Prop. 6.12, „Socken und Schuhe"). □

Übung 6.28: Ist $A = \begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{pmatrix}$ invertierbar? Berechnen Sie die reduzierte ZSF um dies zu überprüfen.

Nun können wir recht leicht zeigen:

Proposition 6.29: (Zwei für eins für invertierbare Matrizen)
Wenn A und B zwei $n \times n$-Matrizen mit $BA = I_n$ sind, dann sind A und B invertierbar und zueinander invers.

Wir brauchen also nicht noch zusätzlich $AB = I_n$, wie in der Definition gefordert wird.

BEWEIS. Angenommen $Av = 0$ für ein $v \in \mathbb{R}^n$. Wir wollen zeigen, dass dann $v = 0$, damit A eine der Invertierbarkeitsbedingungen (Theorem 6.27) erfüllt.
Wir haben $B(Av) = B0 = 0$, aber auch $(BA)v = I_n v = v$, also ist $v = 0$. Daher ist $\operatorname{Ker} A = \{0\}$, und A ist invertierbar.
Da A^{-1} existiert, gibt Multiplikation von $BA = I_n$ mit A^{-1} von rechts sofort $B = A^{-1}$. Daraus folgt auch direkt $B^{-1} = A$, da $(A^{-1})^{-1} = A$. □

Übung 6.30: Benutzen Sie „zwei für eins für invertierbare Matrizen", um zu zeigen:
Wenn ein Produkt AB von zwei $n \times n$-Matrizen A und B invertierbar ist, dann sind sowohl A als auch B separat invertierbar.
Das können wir später auch noch mal anders beweisen, aber hier gibt es auch schon eine Möglichkeit.

Wir notieren auch die Auswirkungen, die eine invertierbare Matrix auf ein LGS hat.

Theorem 6.31: (Invertierbares LGS)
Für eine quadratische Matrix A sind folgende Aussagen äquivalent:
- *(i) A ist invertierbar.*
- *(ii) Das homogene LGS $Ax = 0$ hat nur die triviale Lösung $x = 0$.*
- *(iii) Das inhomogene LGS $Ax = b$ ist für jedes $b \in \mathbb{R}^n$ lösbar.*
- *(iv) das inhomogene LGS $Ax = b$ hat für jedes $b \in \mathbb{R}^n$ genau eine Lösung.*

BEWEIS. Wir haben schon (i) ⇔(ii) bewiesen. Wir zeigen nun (i) ⇒(iv) ⇒(iii) ⇒(i).

◇ (i) ⇒(iv): Dies ist Proposition 6.18:
$$Ax = b \quad \Rightarrow \quad A^{-1}Ax = A^{-1}b$$
gibt eindeutige Lösung $x = A^{-1}b$.

◇ (iv) ⇒(iii): Wenn das LGS $Ax = b$ eine eindeutige Lösung hat, dann ist es lösbar.

◇ (iii)⇒(i): Wir bauen eine Inverse für A aus Lösungen von verschiedenen inhomogenen LGS $Ax = b$: Seien $x_1, x_2, \ldots, x_n$ Lösungen zu den LGS $Ax = e_1$, $Ax = e_2$, …, $Ax = e_n$. Sei $B = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ x_1 & \cdots & x_k & \cdots & x_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix}$ die Matrix mit diesen Lösungen als Spalten. Dann ist
$$AB = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ Ax_1 & \cdots & Ax_k & \cdots & Ax_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ e_1 & \cdots & e_k & \cdots & e_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} = I_n.$$
Also folgt aus „zwei für eins für invertierbare Matrizen" (Prop. 6.29), dass A invertierbar ist. □

C. Inverse Matrizen: Study guide

Konzeptübersicht.

- ◇ Inverse Matrix
- ◇ Invertierbare und singuläre Matrizen
- ◇ Inverse einer 2×2-Matrix
- ◇ Eindeutigkeit der Inversen
- ◇ Inverse Matrixabbildung
- ◇ Inverse einer Inversen, Inverse einer Transponierten, Inverse eines Produkts
- ◇ Matrizen mit Nullspalten oder Nullzeilen nicht invertierbar
- ◇ Invertierungsalgorithmus
- ◇ Elementarmatrizen und deren Eigenschaften
- ◇ Invertierbarkeitsbedingungen
- ◇ Zwei für Eins für invertierbare Matrizen
- ◇ Invertierbare LGS

Skills.

- ◇ Inverse einer 2×2-Matrix finden
- ◇ Leichte Eigenschaften von Inversen mit Hilfe der Definition und Eindeutigkeit beweisen
- ◇ Inverse Matrix mit Invertierungsalgorithmus berechnen
- ◇ LGS mit Hilfe einer inversen Matrix lösen

D. Movitation für Determinanten

Wir haben gesehen, dass eine 2×2-Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertierbar ist genau dann wenn $ad - bc \neq 0$. Dies ist eine nützliche Bedingung, die wir auch auf $n \times n$-Matrizen ausweiten wollen.

Definition 6.32: Die **Determinante** einer 2×2-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist $\det A = ad - bc$.

Beispiele 6.33:

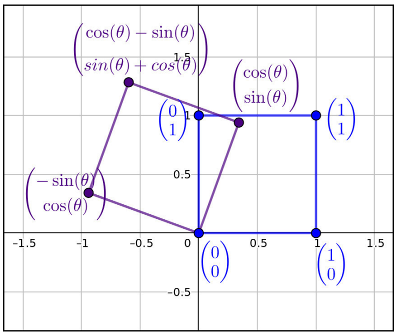
◇ $\det \begin{pmatrix} 1 & 4 \\ 8 & 3 \end{pmatrix} = 3 - 32 = -29$ ◇ $\det \begin{pmatrix} 7 & 0 \\ 0 & 3 \end{pmatrix} = 21$ ◇ $\det \begin{pmatrix} a & b \\ 2a & 2b \end{pmatrix} = 2ab - 2ab = 0$

Die Determinante wird uns über Invertierbarkeit Auskunft geben, aber sie hat auch eine geometrische Bedeutung.

Fakt 6.34: Betrachten Sie das Quadrat mit Ecken $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Dies hat Flächeninhalt 1. Wenn wir das Quadrat mit einer Matrixabbildung f_A transformieren, dann hat das resultierende Viereck Flächeninhalt $\det A$. Eine negative Determinante signalisiert hier, dass die Orientierung der Ecken umgekehrt wurde.

Beispiele 6.35:

$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ hat $\det A = \cos^2(\theta) + \sin^2(\theta) = 1$.



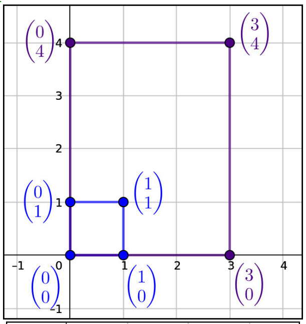
Die Vektoren

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

gehen auf

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta) - \sin(\theta) \\ \sin(\theta) + \cos(\theta) \end{pmatrix}, \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}.$$

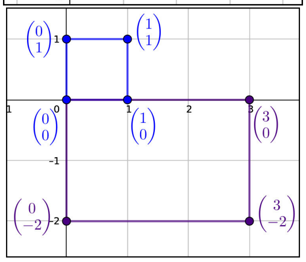
Dies ist eine Drehung um Winkel θ.
Sie können in GeoGebra den Winkel ändern und sehen, was passiert.



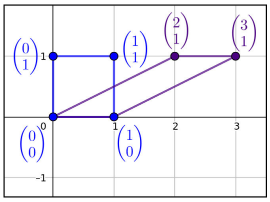
$A = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$ hat $\det A = 12$. Diese Abbildung streckt alles in x-Richtung um Faktor 3 und alles in y-Richtung um Faktor 4, also landet das Einheitsquadrat auf

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix},$$

ein Rechteck mit Flächeninhalt 12.



$A = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$ hat negative Determinante $\det A = -6$. Dies entspricht der Umkehrung der Orientierung der Eckpunkte.



$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ ist eine Scherung: $\det A = 1$: Das Quadrat wird in ein flächengleiches Parallelogramm verwandelt. Sie können sich vorstellen, dass das Quadrat die Vorderseite eines Stapels aus vielen waagerechten Seiten Papier ist, die dann seitlich verrutschen.

Die 2×2-Determinante hat folgende Eigenschaften:

Formel 6.36: (Eigenschaften der 2×2-Determinante)

(1) $\det \begin{pmatrix} ra & b \\ rc & d \end{pmatrix} = rad - rcb = r \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

(2) $\det \begin{pmatrix} a+x & b \\ c+z & d \end{pmatrix} = (a+x)d - b(c+z) = ad - bc + xd - bz = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det \begin{pmatrix} x & b \\ z & d \end{pmatrix}$

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

(4) $\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$.

Wir wollen jetzt die Determinante einer $n \times n$-Matrix definieren, sodass A invertierbar ist genau dann wenn die Determinante ungleich Null ist. Sie soll auch die selben Eigenschaften haben, die wir gerade aufgeschrieben haben.

E. Determinanten via Entwicklung in Zeile oder Spalte

Definition 6.37: (Entwicklung in der ersten Zeile) Sei $A = (a_{ij})$ eine $n \times n$-Matrix, $n \geqslant 2$. Die **Determinante** von A ist definiert als

$$\det A = \sum_{j=1}^{n} a_{1j}(-1)^{1+j} \det(\widehat{A}_{1j}),$$

wobei $\widehat{A}_{ij}$ die Matrix ist, die aus A durch Löschen der i-ten Zeile und der j-ten Spalten entsteht.
Die Determinante der Matrix $\widehat{A}_{ij}$ heißt **Minor** des Eintrags a_{ij}, und $(-1)^{i+j}\det(\widehat{A}_{ij})$ heißt **Kofaktor** des Eintrags a_{ij}.

Mathematiker benutzen so einen Hut oft, wenn etwas ausgelassen wird. In $\widehat{A}_{ij}$ lassen wir Zeile i und Spalte j aus.

Bemerkung 6.38: Wir haben hier die Determinante rekursiv definiert. Um die Determinante einer 3×3-Matrix zu definieren, brauchen wir schon die Determinante einer 2×2-Matrix, die wir oben definiert haben. Und die Determinante einer $n \times n$-Matrix benötigt in dieser Definition wiederum die Determinante einer $n-1 \times n-1$-Matrix.

Notation 6.39: Wenn wir die Einträge einer Matrix explizit ausschreiben, dann benutzen wir oft „gerade Klammern“ zur Kennzeichnung der Determinante:

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \qquad \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

⚠ Wir benutzen diese Notation nicht, wenn die Matrix einen Variablennamen hat! Wir schreiben nur $\det A$, nicht $|A|$.

Beispiel 6.40: In Größe 3×3 haben wir

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \cdot \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \cdot \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \cdot \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Beispiel 6.41: Sei $A = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{pmatrix}$. Dann ist

$$\widehat{A}_{11} = \begin{pmatrix} 4 & -1 \\ -2 & 0 \end{pmatrix}, \quad \widehat{A}_{12} = \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix}, \quad \widehat{A}_{13} = \begin{pmatrix} 2 & 4 \\ 0 & -2 \end{pmatrix},$$

und

$$\begin{aligned}\det A &= \sum_{j=1}^{3} (-1)^{1+j} a_{1j} \det \widehat{A}_{1j} \\ &= \underbrace{(-1)^{1+1}}_{=1} \underbrace{a_{11}}_{=1} \det \widehat{A}_{11} + \underbrace{(-1)^{1+2}}_{=-1} \underbrace{a_{12}}_{=5} \det \widehat{A}_{12} + \underbrace{(-1)^{1+3}}_{=1} \underbrace{a_{13}}_{=0} \det \widehat{A}_{13} \\ &= 1 \cdot 1 \cdot \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} + (-1) \cdot 5 \cdot \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 1 \cdot 0 \cdot \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} \\ &= -2 - 5 \cdot 0 + 0 = -2.\end{aligned}$$

Statt in der ersten Zeile können wir auch in einer beliebigen Zeile oder Spalte entwickeln.

Proposition 6.42: (Entwicklungsformeln)
Sei $A = (a_{ij})$ eine $n \times n$-Matrix, $n \geqslant 2$. Dann ist

$$\begin{aligned}\det A &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\widehat{A}_{ij}) \quad \text{für jedes } i \qquad \text{(in Zeile } i \text{ entwickeln)} \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\widehat{A}_{ij}) \quad \text{für jedes } j \qquad \text{(in Spalte } j \text{ entwickeln)}\end{aligned}$$

BEWEIS. Lassen wir hier aus. Dies kommt von einer anderen Definition der Determinante, die Permutationen und Vorzeichen von Permutationen benutzt. Dazu müssten wir etwas mehr Gruppentheorie machen, als wir Zeit haben. □

Slogan: Die Determinante von A ist eine „alternierende“ Summe von allen möglichen Produkten, in denen je ein Eintrag jeder Zeile und ein Eintrag jeder Spalte multipliziert wird.

Beispiel 6.43: Berechne die Determinante von $A = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{pmatrix}$ durch Entwicklung in der 3ten Spalte:

$$\begin{aligned}\det A &= (-1)^{1+3} \cdot 0 \cdot \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} + (-1)^{2+3} \cdot (-1) \cdot \begin{vmatrix} 1 & 5 \\ 0 & -2 \end{vmatrix} + (-1)^{3+3} \cdot 0 \cdot \begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} \\ &= -2.\end{aligned}$$

Wir sehen also, dass diese Methode sehr hilfreich sein kann, wenn wir eine Zeile oder Spalte mit vielen Nulleinträgen wählen.

Proposition 6.44: *Wenn A eine Nullspalte oder Nullzeile enthält, dann ist* $\det A = 0$.

BEWEIS. Entwickle in der Nullzeile oder -spalte. □

Die Entwicklungsformeln geben uns auch eine schöne Formel zur Berechnung einer 3×3-Determinante.

Formel 6.45: (Gartenzaunregel für 3×3-Matrizen)

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

Übung 6.46: Verifizieren Sie diese Formel, indem Sie in der erste Zeile entwicklen.

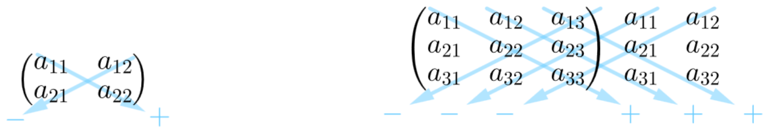


ABBILDUNG 1. „Gartenzaunregel“ für 2×2 und 3×3 Determinanten.

⚠ Vorsicht! Diese einfache Regel funktioniert nur für 2×2 und 3×3-Matrices. Für größere Matrizen gibt es keine solch einfache Regel.

F. Eigenschaften der Determinante

Da wir entlang Zeilen oder Spalten entwickeln können, ändert Transponieren die Determinante der Matrix nicht:

Proposition 6.47: (Determinante der Transponierten)
Sei $A \in \mathcal{M}_{n,n}$. Dann ist
$$\det A^T = \det A$$

BEWEIS. Entwicklung von $\det A^T$ in Zeile 1 gibt das selbe Ergebnis wie Entwicklung von $\det A$ Spalte 1. Für 2×2-Determinanten gilt die Eigenschaft, weil $ad - bc = ad - cb$. □

Nun können wir zeigen, dass die Eigenschaften, die wir für 2×2-Determinanten berechnet haben, auch für die Determinante einer $n \times n$-Matrix gelten:

Proposition 6.48: (Eigenschaften der Determinante)
Seien $a_1, \dots, a_n$ und b_k Spalten von $n \times n$-Matrizen (also Vektoren aus $\mathbb{R}^n$), und $a'_1, \dots, a'_n$ und b'_k Zeilen von $n \times n$-Matrizen. Dann gilt

(i) Man kann einen Skalar aus einer Zeile oder Spalte herausziehen:
$$\det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & \lambda a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} = \lambda \det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} \text{ und}$$
$$\det\begin{pmatrix} \leftarrow & a'_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \lambda a'_k & \rightarrow \\ & \vdots & \\ \leftarrow & a'_n & \rightarrow \end{pmatrix} = \lambda \det\begin{pmatrix} \leftarrow & a'_1 & \rightarrow \\ & \vdots & \\ \leftarrow & a'_k & \rightarrow \\ & \vdots & \\ \leftarrow & a'_n & \rightarrow \end{pmatrix}$$

(ii) Die Determinante ist linear in Zeilen und Spalten:
$$\det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k+b_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} = \det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} + \det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & b_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix} \text{ und}$$
$$\det\begin{pmatrix} \leftarrow & a'_1 & \rightarrow \\ & \vdots & \\ \leftarrow & a'_k+b'_k & \rightarrow \\ & \vdots & \\ \leftarrow & a'_n & \rightarrow \end{pmatrix} = \det\begin{pmatrix} \leftarrow & a'_1 & \rightarrow \\ & \vdots & \\ \leftarrow & a'_k & \rightarrow \\ & \vdots & \\ \leftarrow & a'_n & \rightarrow \end{pmatrix} + \det\begin{pmatrix} \leftarrow & a'_1 & \rightarrow \\ & \vdots & \\ \leftarrow & b'_k & \rightarrow \\ & \vdots & \\ \leftarrow & a'_n & \rightarrow \end{pmatrix}$$

(iii) Der Tausch von zwei Zeilen oder zwei Spalten erzeugt ein Minus:
$$\det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_l & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \end{pmatrix} = -\det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_l & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \end{pmatrix} \text{ und}$$

$$\det\begin{pmatrix} \leftarrow a_1' \rightarrow \\ \vdots \\ \leftarrow a_k' \rightarrow \\ \vdots \\ \leftarrow a_l' \rightarrow \\ \vdots \\ \leftarrow a_n' \rightarrow \end{pmatrix} = -\det\begin{pmatrix} \leftarrow a_1' \rightarrow \\ \vdots \\ \leftarrow a_l' \rightarrow \\ \vdots \\ \leftarrow a_k' \rightarrow \\ \vdots \\ \leftarrow a_n' \rightarrow \end{pmatrix}$$

(iv) Eine Matrix mit zwei gleichen Spalten oder Zeilen hat Determinante 0:

$$\det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \end{pmatrix} = 0 \qquad \textit{und} \qquad \det\begin{pmatrix} \leftarrow a_1' \rightarrow \\ \vdots \\ \leftarrow a_k' \rightarrow \\ \vdots \\ \leftarrow a_k' \rightarrow \\ \vdots \\ \leftarrow a_n' \rightarrow \end{pmatrix} = 0$$

(v) Das Addieren eines Vielfachen einer Spalte oder Zeile zu einer anderen ändert die Determinante nicht:

$$\det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k+\lambda a_l & \cdots & a_l & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \end{pmatrix} = \det\begin{pmatrix} \uparrow & & \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_l & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \end{pmatrix} \ \textit{und}$$

$$\det\begin{pmatrix} \leftarrow & a_1' & \rightarrow \\ & \vdots & \\ \leftarrow & a_k'+\lambda a_l' & \rightarrow \\ & \vdots & \\ \leftarrow & a_l' & \rightarrow \\ & \vdots & \\ \leftarrow & a_n' & \rightarrow \end{pmatrix} = \det\begin{pmatrix} \leftarrow a_1' \rightarrow \\ \vdots \\ \leftarrow a_k' \rightarrow \\ \vdots \\ \leftarrow a_l' \rightarrow \\ \vdots \\ \leftrightarrow a_n' \rightarrow \end{pmatrix}$$

(vi) $\det I_n = 1$.

BEWEIS. (i) Man kann dies durch Induktion zeigen, oder indem man in der richtigen Zeile oder Spalte entwickelt.

Seien $A = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix}$ und $B = \begin{pmatrix} \uparrow & & \uparrow & & \uparrow \\ a_1 & \cdots & \lambda a_k & \cdots & a_n \\ \downarrow & & \downarrow & & \downarrow \end{pmatrix}$, also $b_{ik} = \lambda a_{ik}$ für $i = 1, \ldots, n$. Entwicklung in Spalte k gibt:

$$\begin{aligned} \det B &= \sum_{i=1}^{n} (-1)^{k+i} b_{ik} \det(\widehat{B}_{ik}) \\ &= \sum_{i=1}^{n} (-1)^{k+i} \lambda a_{ik} \det(\widehat{A}_{ik}) \\ &= \lambda \det A. \end{aligned}$$

Für Zeilen folgt dies aus $\det A^T = \det A$, Proposition 6.47.

(ii) Übung. Entwickeln Sie in Spalte/Zeile k.

(iii) Übung . Man kann dies durch Induktion und Entwicklung in einer nicht getauschten Zeile oder Spalte zeigen.

(iv) Folgt aus (iii): Wenn zwei Spalten gleich sind, bekommen wir durch Spaltentausch ein Minus, aber haben gleichzeitig immer noch dieselbe Matrix.

(v) Kombiniere (i), (ii) und (iv).

(vi) Wir berechnen die Determinante der Einheitsmatrix durch Entwicklung in der ersten Zeile:

$$\begin{aligned} \det I_n &= (-1)^{1+1} \underbrace{(I)_{11}}_{=1} \det I_{n-1} + \sum_{j=2}^{n} (-1)^{1+j} \underbrace{(I)_{1j}}_{=0} (\widehat{I}_n)_{1j} \\ &= 1 \cdot \det I_{n-1} = \ldots = \det I_2 = 1. \end{aligned}$$

□

⚠ Vorsicht: Im Allgemeinen ist die Determinante einer Summe ⚠ nicht die Summe der Determinanten:

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

haben Determinanten $\det A = 1, \det B = 5$, sodass $\det A + \det B = 6$, aber

$$A + B = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}$$

hat Determinante $\det(A + B) = 0 \neq 6$.

Proposition 6.49: (Determinante einer oberen Dreiecksmatrix)
Die Determinante einer oberen Dreiecksmatrix ist das Produkt der Diagonaleinträge.

$$\begin{vmatrix} a_{11} & * & \dots & * \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22}\dots a_{nn}$$

Ebenso für untere Dreiecksmatrizen.

BEWEIS. Wir entwickeln immer wieder entlang der ersten Spalte:

$$\det A = a_{11} \det \begin{pmatrix} a_{22} & * & \dots & * \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} = \dots = a_{11}a_{22}\dots a_{nn}.$$ □

Übung 6.50: Berechnen Sie folgende einführenden Determinanten in Numbas.

Entwicklung entlang einer Zeile oder Spalte ist nur sinnvoll, wenn es viele Nullen gibt.

G. Determinante via Zeilenumformungen

Wir lernen nun, wie man Determinanten durch Umformung in eine obere Dreiecksmatrix berechnen kann. Dies ist üblicherweise ein weniger rechen-intensiver Weg als Entwicklen und eignet sich deshalb auch für große Matrizen.

Proposition 6.51: (Determinanten der Elementarmatrizen)
Wir haben $\det(M_i(\lambda)) = \lambda$, $\det(T_{ij}) = -1$, $\det(A_{ij}(\lambda) = 1$.

BEWEIS. Folgt direkt aus Proposition 6.48, Eigenschaften der Determinante. □

Visuell:

$$\det \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = \lambda \qquad \det \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = -1 \qquad \det \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = 1$$

Insbesondere haben alle Elementarmatrizen eine von Null verschiedene Determinante.
Mit Hilfe dieser Elementarmatrizen können wir nun einige Teile von Proposition 6.48 etwas umformulieren:

Proposition 6.52: (Determinante erhält Produkt mit Elementarmatrizen.)
Sei E eine Elemenatmatrix, und A eine $n \times n$-Matrix. Dann gilt $\det(EA) = \det E \det A$.

BEWEIS. Das Multiplizieren mit E von links hat den Effekt, auf A die dazugehörige elementare Zeilenumformung auszuführen (Prop. 6.23), und wir kennen den Effekt dieser Zeilenumformungen auf die Determinante:

(i) Wenn $E = M_i(\lambda)$ eine Zeile mit $\lambda \in \mathbb{R}$, $\lambda \neq 0$ multipliziert, dann ist
$$\lambda \det A = \det(EA) = \det E \det A.$$

(ii) Wenn $E = T_{ij}$ zwei Zeilen tauscht, dann ist
$$-\det A = \det(EA) = \det E \det A.$$

(iii) Wenn $E = A_{ij}(\lambda)$ ein Vielfaches einer Zeile auf eine andere addiert, dann ist
$$\det A = \det(EA) = \det E \det A.$$
□

Wir sehen: wenn wir nur *elementare* Zeilenoperationen ausführen, haben wir Kontrolle über die Determinante. Daher können wir Schritte ähnlich zum Gauss-Algorithmus benutzen, um die Determinante einer Matrix zu berechnen.

Beispiel 6.53: Wir benutzen Zeilenumformungen, um $\det A$ für
$$A = \begin{pmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{pmatrix}$$
zu berechnen:
$$\det A \overset{\mathrm{I}\leftrightarrow\mathrm{II}}{=} - \begin{vmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix} \overset{\mathrm{I}/3}{=} -3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix}$$
$$\overset{\mathrm{III}-2\mathrm{I}}{=} -3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{vmatrix} \overset{\mathrm{III}-10\mathrm{II}}{=} -3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -55 \end{vmatrix} = (-3)(-55) = 165$$
Am Ende benutzen wir, dass die Determinante einer oberen Dreiecksmatrix das Produkt der Diagonaleinträge ist.

Übung 6.54: Berechnen Sie 3×3-Determinanten in Numbas, und probieren Sie dabei die verschiedenen Methoden aus.

H. Determinante eines Matrixprodukts

Wir haben gesehen, dass die Formeln für die Determinante und die Formel für Matrixmultiplikation beide sehr komplex sind. Das macht die folgende sehr schöne Verbindung zwischen diesen beiden Konzepten extrem unerwartet! Wir werden zeigen, dass für $n \times n$-Matrizen A und B gilt: $\det(AB) = \det A \det B$.
Wir wissen aus Proposition 6.52 schon, dass dies gilt, wenn eine der Matrizen eine Elementarmatrix ist. Mehrfache Anwendung gibt uns:

Korollar 6.55: *Wenn $E_1, \ldots, E_r \in \mathcal{M}_{n,n}$ Elementarmatrizen sind und $B \in \mathcal{M}_{n,n}$, dann ist*
$$\det(E_1 E_2 \ldots E_r B) = \det(E_1) \ldots \det(E_r) \det(B).$$

Dies benutzen wir in einer der wichtigsten Anwendungen der Determinante:

Theorem 6.56: *Eine quadratische Matrix A ist invertierbar genau dann wenn* $\det A \neq 0$.

BEWEIS. Sei R die reduzierte ZSF von A, und $E_1, \ldots, E_r$ die Elementarmatrizen, die $R = E_1 \ldots E_r A$ geben. Wegen Korrollar 6.55 haben wir $\det(R) = \det(E_1) \ldots \det(E_r) \det(A)$. Da die Determinante von Elementarmatrizen nicht Null ist, sehen wir, dass $\det R = 0$ genau dann wenn $\det A = 0$.
Wenn A invertierbar ist, dann ist $R = I_n$ und somit $\det A \neq 0$, weil $\det I_n = 1 \neq 0$.
Andersherum, wenn $\det A \neq 0$ dann ist $\det R \neq 0$, also kann R keine Nullzeile haben. Da R quadratisch ist, bedeutet dies $R = I_n$, und somit ist A invertierbar. □

Wir können dies also zu unserer Liste von Invertierbarkeitsbedingungen hinzufügen.

Theorem 6.57: (Invertierbarkeitsbedingungen)
Für eine quadratische Matrix A sind folgende Aussagen äquivalent:

(i) *A ist invertierbar.*
(ii) $\operatorname{Ker} A = \{0\}$.
(iii) *$Ax = 0$ hat nur die triviale Lösung $x = 0$.*
(iv) *Die reduzierte ZSF von A ist die Einheitsmatrix I.*
(v) *A ist das Produkt von Elementarmatrizen.*
(vi) $\det A \neq 0$.

Nun können wir das wichtigeste Resultat über Determinanten beweisen:

Theorem 6.58: (Determinante ist multiplikativ)
Wenn $A, B \in \mathcal{M}_{n,n}$ dann ist

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

BEWEIS. Wir betrachten zwei Fälle: ob A invertierbar ist oder nicht.
Wenn AB invertierbar, dann ist A auch intervierbar mit Inverser $B(AB)^{-1}$: $A \cdot B(AB)^{-1} = I$, was nach „zwei für eins für invertierbare Matrizen“ (Prop. 6.29) reicht.
Daraus folgt: wenn A nicht invertierbar, dann ist AB auch nicht invertierbar, also folgt aus Theorem 6.56 $\det(A) = \det(AB) = 0$. Dies zeigt das Theorem im Fall A nicht invertierbar.
Sei nun A invertierbar. Dann ist A das Produkt von Elementarmatrizen $E_1, \ldots, E_r$. Also ist

$$\det(AB) = \det(E_1 \ldots E_r B) = \det(E_1) \ldots \det(E_r) \det(B) = \det(A) \det(B).$$ □

Beispiel 6.59: Mit

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 3 \\ 5 & 8 \end{pmatrix} \qquad \text{und} \qquad AB = \begin{pmatrix} 2 & 17 \\ 3 & 14 \end{pmatrix},$$

haben wir

$$\begin{aligned} \det A &= 3 - 2 = 1, \\ \det B &= -8 - 15 = -23, \\ \det(AB) &= 28 - 51 = -23 = \det A \cdot \det B. \end{aligned}$$

Proposition 6.60: *Wenn A invertierbar ist, dann gilt*

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

BEWEIS. Da $A^{-1}A = I_n$ haben wir

$$\det(A^{-1})\det(A) = \det(I_n) = 1$$

was das Resultat beweist. □

I. Determinanten: Study guide

Konzeptübersicht.

- ◇ Bedeutung der Determinante für Invertierbarkeit und Volumen
- ◇ Determinante einer Matrix
- ◇ Entwicklungsformeln
- ◇ Gartenzaunregel und Vorraussetzungen für ihre Anwendung
- ◇ Eigenschaften der Determinante (inklusive Determinante der Transponierten)
- ◇ Determinante einer obere Dreiecksmatrix
- ◇ Effekt von elementaren Zeilenumformungen auf die Determinante
- ◇ Determinanten der drei Typen von Elementarmatrizen
- ◇ Erweiterte Invertierbarkeitsbedingungen
- ◇ Determinante eines Matrixprodukts
- ◇ Determinante der Inversen

Skills.

- ◇ Determinante einer quadratischen Matrix via Entwicklung in Zeile oder Spalte berechnen
- ◇ Determinante einer 2×2 oder 3×3-Matrix mit Gartenzaunregel berechnen
- ◇ Determinante einer Matrix mit Nullzeile oder Nullspalte bestimmen
- ◇ Determinanten von Dreiecksmatrizen (oder Diagonalmatrizen) durch Inspektion bestimmen
- ◇ Determinante durch Zeilenumformungen berechnen
- ◇ Obige Methoden zur Determinantenbestimmung vorteilhaft kombinieren
- ◇ Mit Hilfe der Determinante die Invertierbarkeit einer Matrix prüfen
- ◇ Determinante eines Matrixprodukts bestimmen
- ◇ Determinante einer inversen Matrix bestimmen

KAPITEL 7

Lineare Abbildungen als Matrizen

A. Matrixdarstellung von linearen Abbildungen

Mit der eindeutigen Basisdarstellung (Theorem 4.19) haben wir das Konzept der *Koordinatenvektoren* definiert, in Kapitel 4 (Lineare Unabhängigkeit und Basen). Dies können wir benutzen, um eine beliebige lineare Abbildung zwischen endlich-dimensionalen Vektorräumen als Matrix darzustellen.

Definition 7.1: Sei $f\colon V \longrightarrow W$ eine lineare Abbildung zwischen endlich-dimensonalen Vektorräumen, und seien $\mathcal{B}\colon v_1, \ldots, v_n$ eine Basis für V und $\mathcal{C}\colon w_1, \ldots, w_m$ eine Basis für W. Dann ist die **Matrix von f bezüglich Basen $\mathcal{B}$ and $\mathcal{C}$** (auch **Darstellungsmatrix**) die Matrix A, deren Spalten die Bilder der Basisvektoren aus V sind, jeweils geschrieben als Koordinatenvektoren bezüglich Basis $\mathcal{C}$ von W.

$$ {}_{\mathcal{C}}(f)_{\mathcal{B}} = \begin{pmatrix} {}_{\mathcal{C}}(f(v_1)) & {}_{\mathcal{C}}(f(v_2)) & \cdots & {}_{\mathcal{C}}(f(v_n)) \end{pmatrix} $$

Diese Matrix ist also eine $m \times n$-Matrix.

Lemma 7.2: (Lineare Abbildung als Matrixtransformation)
Für eine lineare Abbildung $f\colon V \longrightarrow W$ und Basen $\mathcal{B}\colon v_1, \ldots, v_n$ für V und $\mathcal{C}\colon w_1, \ldots, w_m$ für W gilt: für jedes $v \in V$ ist

$$ {}_{\mathcal{C}}(f(v)) = {}_{\mathcal{C}}(f)_{\mathcal{B}} \cdot {}_{\mathcal{B}}(v). $$

In Worten: Der Koordinatenvektor von $f(v)$ ist gegeben als Matrix von f mal Koordinatenvektor von v.
Die Reihenfolge und Position der Basen in der Notation soll helfen: die Basen müssen in der Notation zusammenpassen.

Beweis. Sei $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, also ist der Koordinatenvektor von v bezüglich Basis $\mathcal{B}$

$$ {}_{\mathcal{B}}(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. $$

Wegen Linearität ist dann $f(v) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n)$.
Wenn $A = {}_{\mathcal{C}}(f)_{\mathcal{B}}$ die Matrix von f bezüglich der gegebenen Basen ist, dann ist Spalte k dieser Matrix der Vektor $a_k = {}_{\mathcal{C}}(f(v_k))$. Da ${}_{\mathcal{C}}(-)\colon W \longrightarrow \mathbb{R}^m$ ein Isomorphismus und insbesondere linear ist (Beispiele 5.27), ist der Koordinatenvektor von $f(v)$ also

$$ {}_{\mathcal{C}}(f(v)) = \lambda_1 {}_{\mathcal{C}}(f(v_1)) + \cdots + \lambda_n {}_{\mathcal{C}}(f(v_n)) = \lambda_1 a_1 + \cdots + \lambda_n a_n = A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \qquad \square $$

Die Koordinatenvektoren der Basis von V sind

$$ {}_{\mathcal{B}}(v_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_1 \quad {}_{\mathcal{B}}(v_2) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_2 \quad \cdots \quad {}_{\mathcal{B}}(v_{n-1}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix} = e_{n-1} \quad {}_{\mathcal{B}}(v_n) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix} = e_n. $$

Dies kann Ihnen helfen, sich zu merken oder herzuleiten, was die Spalten der Matrix A darstellen: Da Matrixmultiplikation uns

$$Ae_1 = A\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_1, \qquad \text{die erste Spalte von } A,$$

gibt, und ebenso $Ae_k = a_k$, die $k$te Spalte von A, sieht man, dass die Spalten der Matrix die Bilder der Basisvektoren sind.
Hier ist noch ein Diagramm, was die Situation zeigt:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "{}_{\mathcal{B}}(-)"'] & W \arrow[d, "{}_{\mathcal{C}}(-)"] \\ \mathbb{R}^n \arrow[r, "f_A"'] & \mathbb{R}^m \end{tikzcd} \qquad \begin{tikzcd} v \arrow[r, mapsto, "f"] \arrow[d, mapsto, "{}_{\mathcal{B}}(-)"'] & f(v) \arrow[d, mapsto, "{}_{\mathcal{C}}(-)"] \\ {}_{\mathcal{B}}(v) \arrow[r, mapsto, "f_A"'] & A \cdot {}_{\mathcal{B}}(v) = {}_{\mathcal{C}}(f(v)) \end{tikzcd}$$

Bemerkung 7.3: Im Fall wenn $f\colon V \longrightarrow V$ von einem Vektorraum zu sich selbst geht, nehmen wir üblicherweise die selbe Basis auf beiden Seiten. Wir schreiben dann manchmal einfach $(f)_{\mathcal{B}}$ statt ${}_{\mathcal{B}}(f)_{\mathcal{B}}$.
Es gibt aber eine Ausnahme, wenn wir über Basiswechselmatrizen sprechen.

Beispiele 7.4: a) Die Identität $\mathrm{id}\colon V \longrightarrow V$ wird bezüglich jeder Basis durch die Einheitsmatrix dargestellt, solange wir die selbe Basis für Quelle und Ziel wählen.

b) Die Nullabbildung $0\colon V \longrightarrow W$ wird bezüglich jeder Basen für V und W als die Nullmatrix dargestellt.

c) Die Streckung $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mit $f(v) = rv$ hat Matrix

$$ {}_{\mathcal{B}}(f)_{\mathcal{B}} = \begin{pmatrix} r & 0 & 0 & \cdots & 0 \\ 0 & r & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & r & 0 \\ 0 & \cdots & 0 & 0 & r \end{pmatrix} = rI$$

mit r in den Diagonaleinträgen und sonst 0. Dies gilt für jede Basis von $\mathbb{R}^n$, solange wir die selbe Basis für Quelle und Ziel wählen.

d) Die Projektion $\pi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit $\pi\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)\right) = x_1$ hat Matrix

$$ {}_{\mathcal{E}_1}(\pi)_{\mathcal{E}_2} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

bezüglich Standardbasis $\mathcal{E}_2\colon e_1, e_2$ für $\mathbb{R}^2$ und Standardbasis $\mathcal{E}_1\colon e_1 = 1$ für $\mathbb{R}$.

Matrixdarstellung einer linearen Abbildung bestimmen
Sei $f\colon V \longrightarrow W$ eine lineare Abbildung, $v_1, \ldots, v_n$ Basis $\mathcal{B}$ für V und $w_1, \ldots, w_m$ Basis $\mathcal{C}$ für W. Wir bestimmen die Matrix für f bezüglich dieser Basen mit folgenden Schritten:

- Schreibe die Definition

$$ {}_{\mathcal{C}}(f)_{\mathcal{B}} = \begin{pmatrix} {}_{\mathcal{C}}(f(v_1)) & {}_{\mathcal{C}}(f(v_2)) & \cdots & {}_{\mathcal{C}}(f(v_n)) \end{pmatrix}$$

  auf, und setze am besten schon die konkreten Vektoren für $v_1, \ldots, v_n$ ein. Dies hilft uns, damit wir nicht alles im Kopf behalten können. Wir sehen dann, was wir nach und nach ausrechnen müssen.
- Berechne die Bilder der Basisvektoren:

$$f(v_1) = \ldots, \quad f(v_2) = \ldots, \quad \cdots, \quad f(v_n) = \ldots$$

  Also wende die Funktionsvorschrift von f auf die gegebenen Basisvektoren von V an.

- Bestimme die Koordinatenvektoren der gerade berechneten Bilder. Also finde $\lambda_1, \dots, \lambda_m$ sodass $f(v_1) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_m w_m$, durch Lösen eines inhomogenen LGS (mit Gauss-Jordan, wenn die Antwort nicht offensichtlich ist). Wiederhole für alle $f(v_i)$.
  Tipp: Man kann mit Gauss-Jordan alle gleichzeitig berechnen, indem man mehrere Erweiterungen an das selbe LGS hängt.
- Dies gibt die Koordinatenvektoren $_{\mathcal{C}}(f(v_1)) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$, und so weiter. Fasse diese in der Matrix zusammen:
$$_{\mathcal{C}}(f)_{\mathcal{B}} = \Big({_{\mathcal{C}}(f(v_1))} \quad {_{\mathcal{C}}(f(v_2))} \quad \cdots \quad {_{\mathcal{C}}(f(v_n))} \Big).$$

Beispiel 7.5: Sei $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ die Abbildung mit
$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1+x_2 \\ x_1-x_3 \end{pmatrix}.$$
Bezüglich Standardbasen für $\mathbb{R}^3$ und $\mathbb{R}^2$ ist die Matrix
$$A = {_{\mathcal{E}_2}(f)_{\mathcal{E}_3}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

Probe: $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_3 \end{pmatrix}$

Wir bestimmen die Matrix bezüglich der Basen $\mathcal{B}\colon v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ für $\mathbb{R}^3$ und $\mathcal{C}\colon w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ für $\mathbb{R}^2$.
Wir suchen
$$_{\mathcal{C}}(f)_{\mathcal{B}} = \left({_{\mathcal{C}}\left(f\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right)\right)} \quad {_{\mathcal{C}}\left(f\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right)\right)} \quad {_{\mathcal{C}}\left(f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)\right)} \right).$$
Also berechnen wir erst
$$f\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad f\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \qquad f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$
Nun finden wir die Koordinatenvektoren bezüglich Basis $\mathcal{C}$ indem wir drei inhomogene LGS gleichzeitig lösen:
$$\left(\begin{array}{cc|c|c|c} 1 & 1 & 2 & 2 & 1 \\ 1 & 2 & 0 & 1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c|c|c} 1 & 1 & 2 & 2 & 1 \\ 0 & 1 & -2 & -1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c|c|c} 1 & 0 & 4 & 3 & 1 \\ 0 & 1 & -2 & -1 & 0 \end{array}\right)$$
Also ist
$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{also} \qquad {_{\mathcal{C}}\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right)} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$
$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{also} \qquad {_{\mathcal{C}}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right)} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$
$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{also} \qquad {_{\mathcal{C}}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
Dies kommt alles in der Matrix zusammen:
$$_{\mathcal{C}}(f)_{\mathcal{B}} = \begin{pmatrix} 4 & 3 & 1 \\ -2 & -1 & 0 \end{pmatrix}$$

Die Probe ist hier nicht so leicht wie die Probe für eine Matrix in Standardbasis. Die Zeilen, in denen wir die Bildvektoren als Linearkombination der Basis des Ziels schreiben kann unsere Probe sein, oder:
Warum gibt diese Matrix die selbe lineare Abbildung? Wir müssen hier Koordinatenvektoren bezüglich der neuen Basen betrachten:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (x_2 - x_3) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (x_1 - x_2) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad {}_{\mathcal{B}}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_2 - x_3 \\ x_1 - x_2 \end{pmatrix}$$

$$\Rightarrow \quad {}_{\mathcal{C}}\left(f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) \right) = \begin{pmatrix} 4 & 3 & 1 \\ -2 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_3 \\ x_2 - x_3 \\ x_1 - x_2 \end{pmatrix} = \begin{pmatrix} 4x_3 + 3(x_2 - x_3) + (x_1 - x_2) \\ -2x_3 - (x_2 - x_3) \end{pmatrix}$$

$$= \begin{pmatrix} x_1 + 2x_2 + x_3 \\ -x_2 - x_3 \end{pmatrix}$$

$$\Rightarrow \quad f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = (x_1 + 2x_2 + x_3) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-x_2 - x_3) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_3 \end{pmatrix}$$

Beispiel 7.6: Aus Zeitgründen nicht in VL, hier als Angebot, falls diese Fragen aufkommen.
Um zu sehen, warum wir bei einer linearen Abbildung $f\colon V \longrightarrow V$ auf beiden Seiten die selbe Basis haben wollen, betrachten wir die Streckung $f(v) = 3v$ auf $\mathbb{R}^3$. Die Matrix für f bezüglich Standardbasis ist

$$A = {}_{\mathcal{E}_3}(f)_{\mathcal{E}_3} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = 3I,$$

was sehr offensichtlich aussieht wie „Multiplikation mit 3“. Wenn wir statt dessen Basis $\mathcal{B}\colon v_1 = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right), v_2 = \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right), v_3 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$ auf beiden Seiten nehmen, hat f wieder die Matrix

$$A = {}_{\mathcal{B}}(f)_{\mathcal{B}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = 3I,$$

weil $f(v_1) = 3v_1$, usw. Aber wenn wir die Matrix für f bezüglich Basis $\mathcal{B}$ für die Quelle und Standardbasis für das Ziel bestimmen, bekommen wir

$$C = {}_{\mathcal{E}_3}(f)_{\mathcal{B}} = \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 0 \end{pmatrix}$$

weil $f(v_1) = \left(\begin{smallmatrix} 3 \\ 3 \\ 3 \end{smallmatrix}\right)$, $f(v_2) = \left(\begin{smallmatrix} 3 \\ 3 \\ 0 \end{smallmatrix}\right)$ und $f(v_3) = \left(\begin{smallmatrix} 3 \\ 0 \\ 0 \end{smallmatrix}\right)$. Diese Matrix ist nicht so nützlich, weil wir hier nicht sehen, dass f einfach „Multiplikation mit 3“ ist.
Wir können sogar die Matrix

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

bekommen, wenn wir die Standardbasis für die Quelle und jeweils 3 mal die Standardbasisvektoren für das Ziel nehmen. Das sagt uns nicht, was die Abbildung tut, nur, dass sie ein Isomorphismus ist.

Das ist der Grund, warum wir für eine Abbildung $f: V \longrightarrow V$ (fast immer) die *selbe* Basis für Quelle und Ziel benutzen wollen.

Übung 7.7: Üben Sie in Numbas:
- ◇ Matrix einer linearen Abbildung bezüglich Standardbasis finden (einführend)
- ◇ Matrix einer linearen Abbildung $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ finden (schöne Basen)
- ◇ Matrix einer linearen Abbildung finden (etwas mehr randomisiert)
- ◇ Matrix einer linearen Abbildung finden, die durch Bilder einer Basis gegeben ist

Proposition 7.8: (Matrix einer Komposition ist Produkt der Matrizen.)
Wenn $f: U \longrightarrow V$ und $g: V \longrightarrow W$ linear sind und $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ Basen für U, V, W, dann gilt:

$$_{\mathcal{B}_3}(g \circ f)_{\mathcal{B}_1} = {}_{\mathcal{B}_3}(g)_{\mathcal{B}_2}\ {}_{\mathcal{B}_2}(f)_{\mathcal{B}_1}$$

In Worten: Komposition von linearen Abbildungen entspricht Matrixmultiplikation der Darstellungsmatrizen. (Die Position der Basen ist wieder hilfreich, um zu sehen, wann ich die Darstellungsmatrizen so multiplizieren kann.)

Beweis. Wir benutzen die Definition der Darstellungsmatrix.

Übung: Beweisen Sie dies selbst (beste Lernmöglichkeit, die Bedeutung der Definition zu durchdenken), oder lesen Sie den Beweis unten mit der Selsterklärungsmethode.

Seien $u_1, \ldots, u_l$ die Vektoren in der Basis $\mathcal{B}_1$ für U. Als Kurzform schreiben wir $A = {}_{\mathcal{B}_2}(f)_{\mathcal{B}_1}$, $B = {}_{\mathcal{B}_3}(g)_{\mathcal{B}_2}$ und $C = {}_{\mathcal{B}_3}(g \circ f)_{\mathcal{B}_1}$. Wir wollen also zeigen, dass $C = BA$.
Spalte k der Matrix $C = {}_{\mathcal{B}_3}(g \circ f)_{\mathcal{B}_1}$ ist $c_k = {}_{\mathcal{B}_3}(g(f(u_k)))$, nach Definition der Darstellungsmatrix, Def. 7.1.
Spalte k des Matrixprodukts BA ist Ba_k, B mal Spalte k von A (siehe Definition der Matrixmultiplikation, Def. 1.45). Da $a_k = {}_{\mathcal{B}_2}(f(u_k))$ (nach Definition der Darstellungsmatrix, Def. 7.1), und B die Matrix der Abbildung g ist, folgt $Ba_k = {}_{\mathcal{B}_3}(g)_{\mathcal{B}_2} \cdot {}_{\mathcal{B}_2}(f(u_k)) = {}_{\mathcal{B}_3}(g(f(u_k)))$, nach Lemma 7.2.
Also ist die $k$te Spalte von C gleich der $k$ten Spalte von BA. Da k beliebig war, folgt $C = BA$. □

Proposition 7.9: (Isomorphismen haben invertierbare Matrizen.)
Sei $f: V \longrightarrow W$ linear und $\mathcal{B}$ eine Basis für V, $\mathcal{C}$ eine Basis für W. Die Abbildung f ist ein Isomorphismus genau dann wenn die Matrix ${}_{\mathcal{C}}(f)_{\mathcal{B}}$ invertierbar ist. In diesem Fall ist

$$_{\mathcal{B}}(f^{-1})_{\mathcal{C}} = ({}_{\mathcal{C}}(f)_{\mathcal{B}})^{-1}.$$

Beweis. Übung.
Eine Richtung ist leicht: wenn f Isomorphismus, dann ist die Matrix von f invertierbar (zeigen Sie dies unter Verwendung von „Matrix einer Komposition ist Produkt der Matrizen", Prop 7.8).
Die andere Richtung ist schwerer: Wenn die Matrix $A = {}_{\mathcal{C}}(f)_{\mathcal{B}}$ invertierbar ist, dann müssen wir aus der Matrix A^{-1} eine Abbildung $g: W \longrightarrow V$ basteln, und zeigen, dass g die inverse Abbildung zu f ist.
Alternativ können wir einen Zusammenhang zwischen $\operatorname{Ker} f$ und $\operatorname{Ker} A$ einerseits und $\operatorname{Im} f$ und $\operatorname{SR}(A) = \operatorname{Im} f_A$ andererseits herstellen, und die Dimensionsformel benutzen, um zu zeigen, dass f bijektiv und somit ein Isomorphismus ist (ohne die inverse Abbildung anzugeben).

Hier ist ein zusammenfassendes Diagramm:

$$
\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "{}_{\mathcal{B}}(-)"'] \arrow[rr, bend left, "\mathrm{id}_V"] & W \arrow[r, "f^{-1}"] \arrow[d, "{}_{\mathcal{E}}(-)"] & V \arrow[d, "{}_{\mathcal{B}}(-)"] \\
\mathbb{R}^n \arrow[r, "A"'] \arrow[rr, bend right, "I_n"'] & \mathbb{R}^n \arrow[r, "A^{-1}"'] & \mathbb{R}^n
\end{tikzcd}
\qquad \text{und} \qquad
\begin{tikzcd}
W \arrow[r, "f^{-1}"] \arrow[d, "{}_{\mathcal{E}}(-)"'] \arrow[rr, bend left, "\mathrm{id}_W"] & V \arrow[r, "f"] \arrow[d, "{}_{\mathcal{B}}(-)"] & W \arrow[d, "{}_{\mathcal{E}}(-)"] \\
\mathbb{R}^n \arrow[r, "A^{-1}"'] \arrow[rr, bend right, "I_n"'] & \mathbb{R}^n \arrow[r, "A"'] & \mathbb{R}^n
\end{tikzcd}
$$

□

B. Basiswechsel

Wir betrachten jetzt, wie man zwischen Koordinatenvektoren bezüglich verschiedener Basen wechseln kann.

Definition 7.10: Sei V endlich-dimensional, und seien $\mathcal{B}_1$ und $\mathcal{B}_2$ Basen von V. Die Matrix der Identität $\mathrm{id}\colon V \longrightarrow V$ bezüglich $\mathcal{B}_1$ für die Quelle und $\mathcal{B}_2$ für das Ziel,

$$\underset{\mathcal{B}_2\leftarrow\mathcal{B}_1}{P} = {}_{\mathcal{B}_2}P_{\mathcal{B}_1} = {}_{\mathcal{B}_2}(\mathrm{id})_{\mathcal{B}_1}$$

heißt **Basiswechselmatrix** oder **Transformationsmatrix** von Basis $\mathcal{B}_1$ zu Basis $\mathcal{B}_2$.

Somit ist die Basiswechselmatrix automatisch invertierbar, weil die Identität ein Isomorphismus ist (vergleiche „Isomorphismen haben invertierbare Matrizen“, Prop. 7.9).
Wir definieren diese Basiswechselmatrizen, weil wir durch Multiplikation mit solch einer Matrix die betreffenden Koordinatenvektoren ineinander überführen können:

Proposition 7.11: (Basiswechsel)
Sei V endlich-dimensional, und seien $\mathcal{B}_1$ und $\mathcal{B}_2$ Basen von V. Dann gilt für jedes $v \in V$:

$${}_{\mathcal{B}_2}(v) = {}_{\mathcal{B}_2}P_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}(v).$$

Beweis. Wir benutzen Lemma 7.2 (lineare Abbildung als Matrixtransformation) und die Definition der Basiswechselmatrix:

$${}_{\mathcal{B}_2}(v) = {}_{\mathcal{B}_2}(\mathrm{id}(v)) = {}_{\mathcal{B}_2}(\mathrm{id})_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}(v) = {}_{\mathcal{B}_2}P_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}(v)$$

□

Es kann helfen, die Gleichungskette von rechts nach links zu lesen.

Beispiel 7.12: Wir betrachten die Standardbasis $\mathcal{E}\colon e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ für $\mathbb{R}^2$, und die Basis $\mathcal{B}\colon v_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right), v_2 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$. Die Basiswechselmatrix ${}_{\mathcal{E}}P_{\mathcal{B}}$ hat als erste Spalte den ersten Basisvektor von Basis $\mathcal{B}$, geschrieben als Koordinatenvektor in der Standardbasis $\mathcal{E}$. (Ebenso für die zweite Spalte). Also ist

$${}_{\mathcal{E}}P_{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Wenn $v = {}_{\mathcal{E}}(v) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, dann ist ${}_{\mathcal{B}}(v) = \begin{pmatrix} x_2 \\ x_1 - x_2 \end{pmatrix}$, weil $x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (x_1 - x_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.
Wir prüfen:

$${}_{\mathcal{E}}P_{\mathcal{B}}\,{}_{\mathcal{B}}(v) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 - x_2 \end{pmatrix} = \begin{pmatrix} x_2 + (x_1 - x_2) \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = {}_{\mathcal{E}}(v).$$

Basiswechsel *zur* Standardbasis ist leicht: Die Basiswechselmatrix hat als Spalten die Vektoren der alten Basis.

Das ist, weil wir in diesem Fall die Koordinatenvektoren bezüglich Standardbasis brauchen, und das ist immer leicht.
Für den Basiswechsel in die andere Richtung müssen wir die Standardbasis als Koordinatenvektoren bezüglich Basis $\mathcal{B}$ schreiben:

$$e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) = v_2 \qquad \text{also ist} \qquad {}_{\mathcal{B}}(e_1) = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$$
$$e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) = v_1 - v_2 \qquad \text{also ist} \qquad {}_{\mathcal{B}}(e_2) = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$$

und die Basiswechselmatrix ist

$${}_{\mathcal{B}}P_{\mathcal{E}} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Und als Probe: es gilt wirklich

$${}_{\mathcal{B}}P_{\mathcal{E}}\,\mathcal{E}(v) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 - x_2 \end{pmatrix} = {}_{\mathcal{B}}(v).$$

Wir bemerken im obigen Beispiel, dass ${}_{\mathcal{B}}P_{\mathcal{E}} \cdot {}_{\mathcal{E}}P_{\mathcal{B}} = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) = I_2$ und ${}_{\mathcal{E}}P_{\mathcal{B}} \cdot {}_{\mathcal{B}}P_{\mathcal{E}} = I_2$, also ist ${}_{\mathcal{B}}P_{\mathcal{E}} = ({}_{\mathcal{E}}P_{\mathcal{B}})^{-1}$. Dies ist immer der Fall.

Lemma 7.13: (Basiswechsel in umgekehrter Richtung hat inverse Matrix.)
Sei V endlich-dimensional, und seien $\mathcal{B}_1$ und $\mathcal{B}_2$ Basen von V. Dann gilt

$${}_{\mathcal{B}_1}P_{\mathcal{B}_2} = ({}_{\mathcal{B}_2}P_{\mathcal{B}_1})^{-1}.$$

Beweis. Sei $\dim V = n$. Wir benutzen, dass die Matrix der Komposition das Produkt der Darstellungsmatrizen ist (Prop. 7.8):

$${}_{\mathcal{B}_1}P_{\mathcal{B}_2}\,{}_{\mathcal{B}_2}P_{\mathcal{B}_1} = {}_{\mathcal{B}_1}(\mathrm{id})_{\mathcal{B}_2}\;{}_{\mathcal{B}_2}(\mathrm{id})_{\mathcal{B}_1} = {}_{\mathcal{B}_1}(\mathrm{id}\circ\mathrm{id})_{\mathcal{B}_1} = I_n$$

und genauso

$${}_{\mathcal{B}_2}P_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}P_{\mathcal{B}_2} = I_n,$$

also ist ${}_{\mathcal{B}_1}P_{\mathcal{B}_2} = ({}_{\mathcal{B}_2}P_{\mathcal{B}_1})^{-1}$ wie behauptet. □

Das kann uns helfen, Basiswechselmatrizen für kompliziertere Basen zu bestimmen.

Beispiel 7.14: Wir betrachten die Basis $\mathcal{B}\colon v_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right), v_2 = \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)$ von $\mathbb{R}^2$. Die Basiswechselmatrix ${}_{\mathcal{B}}P_{\mathcal{E}}$ können wir hier nicht einfach ablesen: wir sehen nicht sofort, wie wir e_1 und e_2 als Linearkombination der neuen Basisvektoren v_1 und v_2 schreiben können. Aber wir haben vorher gesehen, dass ${}_{\mathcal{E}}P_{\mathcal{B}}$ leicht ist:

$${}_{\mathcal{E}}P_{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Und jetzt wissen wir, dass die Basiswechselmatrix in die andere Richtung einfach die Inverse dieser Matrix ist, die wir (im 2×2-Fall sehr leicht) bestimmen können. Also ist

$${}_{\mathcal{B}}P_{\mathcal{E}} = {}_{\mathcal{E}}P_{\mathcal{B}}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix},$$

und wir können in der Tat nachprüfen:

$$e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) = 2\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right) - \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) = 2v_1 - v_2$$
$$e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) = -\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right) + \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) = -v_1 + v_2.$$

Übung 7.15: Bestimmen Sie Basiswechselmatrizen in Numbas. Dort heißt es Koordinatenwechselmatrix, das ist aber das selbe.

Proposition 7.16: (Basiswechsel für Matrizen)
Sei $f\colon V \longrightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen, und seien $\mathcal{B}_1$, $\mathcal{B}_2$ Basen für V, $\mathcal{C}_1, \mathcal{C}_2$ Basen für W. Dann ist

$$ {}_{\mathcal{C}_2}(f)_{\mathcal{B}_2} = {}_{\mathcal{C}_1}P_{\mathcal{C}_2}^{-1} \cdot {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}P_{\mathcal{B}_2}. $$

BEWEIS. Wir haben

$$ {}_{\mathcal{C}_1}P_{\mathcal{C}_2}^{-1} \cdot {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}P_{\mathcal{B}_2} = {}_{\mathcal{C}_2}(\mathrm{id}_W)_{\mathcal{C}_1} \cdot {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}(\mathrm{id}_V)_{\mathcal{B}_2} = {}_{\mathcal{C}_2}(\mathrm{id}_W \circ f \circ \mathrm{id}_V)_{\mathcal{B}_2} = {}_{\mathcal{C}_2}(f)_{\mathcal{B}_2}. \qquad \square $$

Dies beweist, *dass* es so ist, sagt uns aber nicht, wie man darauf kommt und warum es hilfreich ist. Das wollen wir noch untersuchen.

$$ \begin{tikzcd} \text{Basis } B_1 & V \arrow[r, "{}_{\mathcal{C}_1}(f)_{\mathcal{B}_1}"] \arrow[d, "{}_{\mathcal{B}_2}P_{\mathcal{B}_1}"'] & W \arrow[d, "{}_{\mathcal{C}_2}P_{\mathcal{C}_1}"] & \text{Basis } C_1 \\ \text{Basis } B_2 & V \arrow[r, "{}_{\mathcal{C}_2}(f)_{\mathcal{B}_2}"'] & W & \text{Basis } C_2 \end{tikzcd} $$

Für $v \in V$ haben wir

$$ {}_{\mathcal{C}_1}(f(v)) = {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}(v) \qquad \text{und} \qquad {}_{\mathcal{C}_2}(f(v)) = {}_{\mathcal{C}_2}(f)_{\mathcal{B}_2} \cdot {}_{\mathcal{B}_2}(v). $$

Ebenso haben wir, mit Verwendung von Basiswechselmatrizen,

$$ {}_{\mathcal{B}_1}(v) = {}_{\mathcal{B}_1}P_{\mathcal{B}_2} \cdot {}_{\mathcal{B}_2}(v) \qquad \text{und} \qquad {}_{\mathcal{C}_1}(f(v)) = {}_{\mathcal{C}_1}P_{\mathcal{C}_2} \cdot {}_{\mathcal{C}_2}(f(v)). $$

Wenn wir diese zusammen fügen, bekommen wir zwei Ausdrücke für ${}_{\mathcal{C}_2}(f(v))$: einerseits ist

$$ \begin{aligned} {}_{\mathcal{C}_1}P_{\mathcal{C}_2} \cdot {}_{\mathcal{C}_2}(f(v)) &= {}_{\mathcal{C}_1}(f(v)) \\ &= {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}(v) \\ &= {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}P_{\mathcal{B}_2} \cdot {}_{\mathcal{B}_2}(v) \end{aligned} $$

was uns durch Umstellung den Ausdruck

$$ {}_{\mathcal{C}_2}(f(v)) = {}_{\mathcal{C}_1}P_{\mathcal{C}_2}^{-1}\ {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}P_{\mathcal{B}_2} \cdot {}_{\mathcal{B}_2}(v) $$

gibt. Andererseits haben wir auch

$$ {}_{\mathcal{C}_2}(f(v)) = {}_{\mathcal{C}_2}(f)_{\mathcal{B}_2} \cdot {}_{\mathcal{B}_2}(v). $$

Da dies für alle $v \in V$ gilt, bekommen wir die Matrixgleichheit ${}_{C_2}(f)_{B_2} = {}_{\mathcal{C}_1}P_{\mathcal{C}_2}^{-1} \cdot {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}P_{\mathcal{B}_2}$.

Wenn wir mit Matrizen für lineare Abbildungen arbeiten, müssen wir immer sehr sorgfältig schauen, auf welche Kooordinatenvektoren wir sie anwenden dürfen. Die Basen müssen immer zusammenpassen. Zum Beispiel können sowohl ${}_{\mathcal{C}_2}(f)_{\mathcal{B}_2}$ als auch ${}_{\mathcal{C}_1}P_{\mathcal{C}_2}^{-1} \cdot {}_{\mathcal{C}_1}(f)_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}P_{\mathcal{B}_2}$ auf einen Koordinatenvektor bezüglich Basis $\mathcal{B}_2$ angewendet werden, und geben als Ergebnis einen Vektor in W, der bezüglich Basis $\mathcal{C}_2$ geschrieben ist. Die Matrix ${}_{\mathcal{B}_1}P_{\mathcal{B}_2}$ verwandelt den Koordinatenvektor von v bezüglich $\mathcal{B}_2$ in den Koordinatenvektor bezüglich $\mathcal{B}_1$, welcher dann von der Matrix ${}_{\mathcal{C}_1}(f)_{\mathcal{B}_1}$ verarbeitet werden kann. Das Zwischenergebnis ist dann $f(v)$ bezüglich Basis $\mathcal{C}_1$, und die Matrix ${}_{\mathcal{C}_1}P_{\mathcal{C}_2}^{-1}$ verwandelt diesen Koordinatenvektor von $f(v)$ dann in den Koordinatenvektor bezüglich Basis $\mathcal{C}_2$.
Wenn wir verschiedene Vektorräume als Quelle und Ziel der linearen Abbildung haben, dann haben wir notwendigerweise auch verschiedene Basen für Quelle und Ziel, und somit verschiedene Basiswechselmatrizen. Der Basiswechsel der Darstellungsmatrizen sieht also abgekürzt so aus:

$$ A_2 = Q^{-1} A_1 P $$

Beispiel 7.17: In Beispiel 7.5 haben wir für $f\left(\left(\begin{smallmatrix}x_1\\x_2\\x_3\end{smallmatrix}\right)\right) = \left(\begin{smallmatrix}x_1+x_2\\x_1-x_3\end{smallmatrix}\right)$ zwei Matrizen bestimmt:

$$A_1 = {}_{\mathcal{E}_3}(f)_{\mathcal{E}_2} = \begin{pmatrix}1 & 1 & 0\\1 & 0 & -1\end{pmatrix} \qquad \text{und} \qquad A_2 = {}_{\mathcal{C}}(f)_{\mathcal{B}} = \begin{pmatrix}4 & 3 & 1\\-2 & -1 & 0\end{pmatrix}$$

für Basen $\mathcal{B}\colon v_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}, v_2 = \begin{pmatrix}1\\1\\0\end{pmatrix}, v_3 = \begin{pmatrix}1\\0\\0\end{pmatrix}$ für $\mathbb{R}^3$ und $\mathcal{C}\colon w_1 = \begin{pmatrix}1\\1\end{pmatrix}, w_2 = \begin{pmatrix}1\\2\end{pmatrix}$ für $\mathbb{R}^2$.
Wir haben Basiswechselmatrizen

$${}_{\mathcal{E}_3}P_{\mathcal{B}} = \begin{pmatrix}1 & 1 & 1\\1 & 1 & 0\\1 & 0 & 0\end{pmatrix} \qquad \text{und} \qquad {}_{\mathcal{E}_2}P_{\mathcal{C}} = \begin{pmatrix}1 & 1\\1 & 2\end{pmatrix}, \qquad \text{also} \qquad {}_{\mathcal{C}}P_{\mathcal{E}_2} = \begin{pmatrix}2 & -1\\-1 & 1\end{pmatrix}$$

Wir können also schauen:

$$\begin{aligned}{}_{\mathcal{E}_2}P_{\mathcal{C}}^{-1}\,{}_{\mathcal{E}_2}(f)_{\mathcal{E}_3}\,{}_{\mathcal{E}_3}P_{\mathcal{B}} = {}_{\mathcal{C}}P_{\mathcal{E}_2}\,{}_{\mathcal{E}_2}(f)_{\mathcal{E}_3}\,{}_{\mathcal{E}_3}P_{\mathcal{B}} &= \begin{pmatrix}2 & -1\\-1 & 1\end{pmatrix}\begin{pmatrix}1 & 1 & 0\\1 & 0 & -1\end{pmatrix}\begin{pmatrix}1 & 1 & 1\\1 & 1 & 0\\1 & 0 & 0\end{pmatrix}\\ &= \begin{pmatrix}1 & 2 & 1\\0 & -1 & -1\end{pmatrix}\begin{pmatrix}1 & 1 & 1\\1 & 1 & 0\\1 & 0 & 0\end{pmatrix} = \begin{pmatrix}4 & 3 & 1\\-2 & -1 & 0\end{pmatrix} = {}_{\mathcal{C}}(f)_{\mathcal{B}}\end{aligned}$$

Korollar 7.18: (Basiswechsel für quadratische Matrizen)
Wenn $f\colon V \longrightarrow V$ linear und $\mathcal{B}_1, \mathcal{B}_2$ Basen für V, dann ist

$${}_{\mathcal{B}_2}(f)_{\mathcal{B}_2} = {}_{\mathcal{B}_1}P_{\mathcal{B}_2}^{-1}\,{}_{\mathcal{B}_2}(f)_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}P_{\mathcal{B}_2}.$$

$$A_2 = P^{-1}A_1P.$$

Hier benutzen wir *dieselbe* Basiswechselmatrix für Quelle und Ziel.
Dieses Verhältnis von Matrizen ist sehr wichtig, deswegen bekommt es einen Namen:

Definition 7.19: Zwei quadratische Matrizen A und B sind **ähnlich** wenn sie dieselbe lineare Abbildung bezüglich verschiedener Basen darstellen, oder (äquivalent dazu) wenn es eine invertierbare Matrix P gibt mit $A = P^{-1}BP$.

Beispiel 7.20: Die lineare Abbildung $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ mit $f\left(\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)\right) = \left(\begin{smallmatrix}x_1+x_2\\-2x_1+4x_2\end{smallmatrix}\right)$ hat Darstellungsmatrizen

$$A = {}_{\mathcal{E}_2}(f)_{\mathcal{E}_2} = \begin{pmatrix}1 & 1\\-2 & 4\end{pmatrix} \qquad \text{bezüglich Standardbasis, und}$$

$$D = {}_{\mathcal{B}}(f)_{\mathcal{B}} = \begin{pmatrix}2 & 0\\0 & 3\end{pmatrix} \qquad \text{bezüglich Basis } \mathcal{B}\colon v_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right), v_2 = \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right).$$

Die relevante Basiswechselmatrix ist

$$P = {}_{\mathcal{E}_2}P_{\mathcal{B}} = \begin{pmatrix}1 & 1\\1 & 2\end{pmatrix} \qquad \text{mit} \qquad P^{-1} = \begin{pmatrix}2 & -1\\-1 & 1\end{pmatrix}$$

und wir haben in der Tat

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = D \end{aligned}$$

Also sind A und D ähnliche Matrizen.

Ähnliche Matrizen haben einige gemeinsame Eigenschaften.

Proposition 7.21: *Ähnliche Matrizen haben die selbe Determinante und dieselbe Spur.*

BEWEIS. Wenn $A = P^{-1}BP$ mit invertierbarem P, dann ist

$$\det A = \det(P^{-1}BP) = \det P^{-1} \det B \det P = \det(P^{-1}P) \det B = \det I_n \det B = \det B.$$

Übung: Beweisen Sie Spur $A = \mathrm{Spur}(P^{-1}BP)$. Benutzen Sie dazu $\mathrm{Spur}(CD) = \mathrm{Spur}(DC)$, was Sie zuerst beweisen können, mit Hilfe der Einträge der Matrizen.

□

Beispiel 7.22: Die ähnlichen Matrizen oben haben

$$\det A = \begin{vmatrix} 1 & 1 \\ -2 & 4 \end{vmatrix} = 6 \qquad \det D = \begin{vmatrix} 2 & 0 \\ 0 & 3 \end{vmatrix} = 6$$

$$\mathrm{Spur}\, A = 1 + 4 = 5 \qquad \mathrm{Spur}\, D = 2 + 3 = 5.$$

Da also jede Matrix einer linearen Abbildung dieselbe Determinante hat, können wir definieren:

Definition 7.23: Sei $f\colon V \longrightarrow V$ eine lineare Abbildung eines Vektorraums auf sich selbst. Dann ist die **Determinante von** f die Determinante einer Darstellungsmatrix von f bezügliche einer Basis $\mathcal{B}$. Die **Spur von** f ist die Spur einer Darstellungsmatrix von f.

Dies sind wohl-definierte Konzepte, die nicht von der Wahl der Basis abhängen, wie das vorherige Resultat zeigt. Sie sind nur für lineare Abbildungen von V *zu sich selbst* definiert, so wie Determinante und Spur nur für *quadratische* Matrizen definiert sind.

C. Lineare Abbildungen als Matrizen: Study guide

Konzeptübersicht.

- ◇ Darstellungsmatrix
- ◇ Lineare Abbildung als Matrixtransformation
- ◇ Wann man verschiedene Basen für Quelle und Ziel nimmt, und wann man dieselbe Basis nimmt
- ◇ Matrix von Identität und Streckung unabhängig der Basis
- ◇ Matrix von Komposition und Matrix einer inversen Abbildung
- ◇ Basiswechselmatrix und deren Inverse
- ◇ Basiswechsel zur Standardbasis
- ◇ Basiswechsel für Matrizen
- ◇ Änliche Matrizen
- ◇ Determinante und Spur einer linearen Abbildung

Skills.

- ◇ Die Matrix einer linearen Abbildung bezüglich gegebener Basen ermitteln
- ◇ Koordinatenvektor des Bildes eines Vektors unter einer linearen Abbildung mit Hilfe der relevanten Darstellungsmatrix berechnen
- ◇ Basiswechselmatrix ermitteln
- ◇ Korrekte Gleichung finden, die zwei Matrizen der gleichen linearen Abbildung und relevante Basiswechselmatrizen in Verbindung bringt
- ◇ Determinante und Spur einer linearen Abbildung ermittlen

KAPITEL 8

Eigenwerte und Eigenvektoren

Wir haben gesehen, dass eine lineare Abbildung bezüglich verschiedener Basen mit verschiedenen Matrizen dargestellt werden kann. In diesem Kapitel suchen wir nun jeweils eine besonders schöne Basis für eine gegebene lineare Abbildung.
Im ganzen Kapitel betrachten wir nur lineare Abbildungen $f\colon V \longrightarrow V$, also von einem Vektorraum auf sich selbst, korrespondierend zu quadratischen Matrizen.

A. Definitionen

Wenn eine lineare Abbildung einen bestimmten Vektor nur streckt und die Richtung beibelässt, dann ist dies besonders schön und wird für uns hilfreich sein.

Definition 8.1: Sei $f\colon V \longrightarrow V$ linear. Ein Vektor $v \neq 0 \in V$ heißt **Eigenvektor** mit **Eigenwert** λ wenn $f(v) = \lambda v$.

Da f linear ist, sind alle Vielfachen eines Eigenvektors auch Eigenvektoren mit demselben Eigenwert: $f(\mu v) = \mu f(v) = \mu\lambda v = \lambda(\mu v)$.

⚠ VORSICHT: Wir zählen $v = 0$ nicht als Eigenvektor, weil $f(0) = 0 = \lambda 0$ für jedes λ gilt. Somit sagt uns das nichts aus. Also nicht vergessen, Eigenvektoren müssen von Null verschieden sein.

Beispiele 8.2: a) Für jeden Skalar λ hat die Abbildung $f = \lambda \mathrm{id}_V$ mit $f(v) = \lambda v$ jeden (von Null verschiedenen) Vektor $v \in V$ als Eigenvektor mit Eigenwert λ.

b) In Beispiel 7.20 hatten wir die Abbildung $f\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} x_1+x_2 \\ -2x_1+4x_2 \end{smallmatrix}\right)$. Diese erfüllt

$$f\left(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} 2 \\ 2 \end{smallmatrix}\right) = 2\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \qquad \text{und} \qquad f\left(\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} 3 \\ 6 \end{smallmatrix}\right) = 3\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right).$$

Wir haben also Eigenvektor $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ mit Eigenwert 2 und Eigenvektor $\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$ mit Eigenwert 3.

c) (**Kern und Eigenwert** 0) Für jedes lineare $f\colon V \longrightarrow V$ gilt: wenn $v \in \mathrm{Ker}(f)$, dann ist $f(v) = 0 = 0v$, also ist jeder von Null verschiedene Vektor im Kern von f ein Eigenvektor mit Eigenwert 0. Also ist 0 ein Eigenwert von f genau dann wenn f einen nicht-trivialen Kern hat, genau dann wenn f nicht injektiv ist.

d) Wir betrachten die lineare Abbildung

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} -2x_3 \\ x_1+2x_2+x_3 \\ x_1+3x_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

und die Vektoren

$$v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Wir berechnen

$$f(v_1) = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} = 2v_1 \qquad f(v_2) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = 2v_2 \qquad f(v_3) = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = v_3$$

also hat f Eigenvektoren v_1 und v_2 mit Eigenwert 2, und v_3 mit Eigenwert 1. Es ist auch jede Linearkombination von v_1 und v_2 ein Eigenvektor mit Eigenwert 2:

$$f(av_1 + bv_2) = af(v_1) + bf(v_2) = 2av_1 + 2bv_2 = 2(av_1 + bv_2).$$

Dies gibt uns ein allgemeines Resultat:

Proposition 8.3: (Eigenraum)
Wenn $f\colon V \longrightarrow V$ linear mit Eigenwert λ, dann ist die Menge aller Eigenvektoren für λ, zusammen mit dem Nullvector, ein Unterraum $\mathrm{Eig}(f,\lambda) = \mathrm{Eig}(\lambda)$ von V, nämlich

$$\mathrm{Eig}(f,\lambda) = \mathrm{Ker}(f - \lambda \mathrm{id}_V).$$

BEWEIS. λ ist fest gegeben. Für $v \in V$ gilt

$$v \in \mathrm{Eig}(f,\lambda) \;\Leftrightarrow\; f(v) = \lambda v \;\Leftrightarrow\; f(v) - \lambda v = 0 \;\Leftrightarrow\; (f - \mathrm{id}_V)(v) = 0 \;\Leftrightarrow\; v \in \mathrm{Ker}(f - \lambda \mathrm{id}_V).$$

Also ist $\mathrm{Eig}(f,\lambda) = \mathrm{Ker}(f - \lambda \mathrm{id}_V)$. Wir wissen, dass jeder Kern ein Unterraum ist, also ist $\mathrm{Eig}(f,\lambda)$ auch ein Unterraum. □

Definition 8.4: Der Unterraum $\mathrm{Eig}(f,\lambda) = \mathrm{Ker}(f - \lambda \mathrm{id}_V)$ heißt **Eigenraum** von f.

⚠ Vorsicht! Es gibt einen separaten Eigenraum pro Eigenwert! Nicht einen Eigenraum für alle Eigenwerte zusammen. Dies ist ein häufiger Fehler.

Beispiele 8.5: In den gleichen Beispielen wie oben haben wir

a) $f\colon V \longrightarrow V$ mit $f(v) = \lambda v$ hat $\mathrm{Eig}(\lambda) = V$: der ganze Vektorraum V ist ein Eigenraum.
b) $f\left(\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} x_1 + x_2 \\ -2x_1 + 4x_2 \end{smallmatrix}\right)$ hat Eigenräume $\mathrm{Eig}(2) = \langle \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \rangle$ und $\mathrm{Eig}(3) = \langle \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right) \rangle$.
c) **(Kern ist 0-Eigenraum)** Für beliebiges $f\colon V \longrightarrow V$ haben wir $\mathrm{Eig}(f,0) = \mathrm{Ker}(f)$.
d) Die Abbildung $f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} -2x_3 \\ x_1 + 2x_2 + x_3 \\ x_1 + 3x_3 \end{pmatrix}$ hat Eigenräume

$$\mathrm{Eig}(2) = \langle v_1, v_2 \rangle \qquad \mathrm{Eig}(1) = \langle v_3 \rangle \qquad \text{mit } v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Korollar 8.6: (Isos via Eigenwerte)
$f\colon V \longrightarrow V$ ist ein Isomorphismus genau dann wenn 0 kein Eigenwert von f ist.

BEWEIS. Wir wissen von „zwei für eins bei Isos“ (Prop. 5.29), dass f ein Isomorphismus ist genau dann wenn $\mathrm{Ker}(f) = \{0\}$. Da $\mathrm{Eig}(f,0) = \mathrm{Ker}(f)$, ist $\mathrm{Ker}(f) = \{0\}$ genau dann wenn 0 kein Eigenwert von f ist. □

B. Eigenwerte finden: das charakteristische Polynom

Wie finden wir Eigenwerte und Eigenvektoren? Die Determinante hilft uns dabei:

Lemma 8.7: (Determinante findet Isos.)
Wenn V endlich-dimensional ist, dann ist die lineare Abbildung $f\colon V \longrightarrow V$ ein Isomorphismus genau dann wenn $\det(f) \neq 0$.
Anders formuliert, wenn A eine Darstellungsmatrix von f ist, dann ist f ein Isomorphismus genau dann wenn $\det(A) \neq 0$.

BEWEIS. Dies ist eine Umformulierung von Theorem 6.56 unter Benutzung von Prop. 7.21. □

Wir setzen diese Resultate zusammen zu:

Theorem 8.8: (Eigenwerte sind genau die Nullstellen des charakt. Polynoms.)
Für eine lineare Abbildung $f\colon V \longrightarrow V$ gilt:
λ ist ein Eigenwert von f genau dann wenn $\det(\lambda \cdot \mathrm{id}_V - f) = 0$.
Anders formuliert, wenn A eine Darstellungsmatrix für f ist, dann ist λ ein Eigenwert von f genau dann wenn $\det(\lambda I - A) = 0$.

BEWEIS. Wenn v ein Eigenvektor für f mit Eigenwert λ ist, dann ist v auch ein Eigenvektor von $\lambda\mathrm{id} - f$ mit Eigenwert 0, weil $f(v) = \lambda v \Leftrightarrow \lambda v - f(v) = 0 \Leftrightarrow (\lambda\mathrm{id} - f)(v) = 0 \cdot v$.
Nach „Isos via Eigenwerte" (Kor. 8.6) ist $\lambda \cdot \mathrm{id}_V - f$ ein Isomorphismus genau dann wenn 0 *kein* Eigenwert von $\lambda \cdot \mathrm{id}_V - f$ ist, und nach „Determinante findet Isos" (Lemma 8.7) ist $\lambda \cdot \mathrm{id}_V - f$ ein Isomorphismus genau dann wenn $\det(\lambda \cdot \mathrm{id}_V - f) \neq 0$.
Zusammengesetzt heißt das:
λ ist ein Eigenwert von f $\quad\Leftrightarrow\quad$ 0 ist ein Eigenwert von $\lambda \cdot \mathrm{id}_V - f$ $\quad\Leftrightarrow\quad$ $\det(\lambda \cdot \mathrm{id}_V - f) = 0$.
Die zweite Aussage folgt, weil ähnliche Matrizen die gleiche Determinante haben (Prop. 7.21), und die Determinante einer linearen Abbildung als Determinante einer beliebigen Darstellungsmatrix definiert ist. Wenn $A = {}_{\mathcal{B}}(f)_{\mathcal{B}}$, dann ist auch $\lambda I - A = {}_{\mathcal{B}}(\lambda \cdot \mathrm{id}_V - f)_{\mathcal{B}}$ (Übung). □

Definition 8.9: Für eine lineare Abbildung $f\colon V \longrightarrow V$ ist das **charakteristische Polynom von** f das Polynom $\chi_f(t) = \det(t \cdot \mathrm{id}_V - f)$, oder mit einer Darstellungsmatrix A von f ist $\chi_f(t) = \chi_A(t) = \det(tI - A)$.

Das Theorem oben zeigt also, dass λ ein Eigenwert von f ist genau dann wenn λ eine Nullstelle des charakteristischen Polynoms von f ist.
Also haben wir:

Die Eigenwerte sind *genau* die Nullstellen des charakteristischen Polynoms.

Warum ist $\det(tI - A)$ ein Polynom?

Beispiel 8.10: Wenn $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine 2×2-Matrix ist, dann ist

$$\det(tI - A) = \begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix} = (t-a)(t-d) - bc = t^2 - (a+d)t + ad - bc$$

ein Polynom von Grad 2.

Da wir Determinanten rekursiv als Linearkombinationen kleinerer Determinanten definiert haben, können wir so per Induktion zeigen, dass $\det(tI - A)$ ein Polynom in t von Grad n ist, wobei A eine $n \times n$-Matrix ist.
Also hat eine $n \times n$-Matrix höchstens n Eigenwerte.

Bei manchen Matrizen kann man die Eigenwerte einfach von der Diagonale ablesen.

Proposition 8.11: (Eigenwerte von Diagonal- und Dreiecksmatrizen)
Die Eigenwerte einer Diagonalmatrix und einer oberen oder unteren Dreiecksmatrix sind genau die Diagonaleinträge der Matrix.

BEWEIS. Nach Prop. 6.49 ist die Determinante einer unteren oder oberen Dreiecksmatrix das Produkt der Diagonaleinträge. Wenn die Matrix A Diagonaleinträge a_{ii} hat, dann hat $tI - A$ Diagonaleinträge $(t - a_{ii})$, also ist $\chi_A(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$. Die Nullstellen dieses Polynoms sind genau die Diagonaleinträge von A, und dies sind also genau die Eigenwerte. □

Übung 8.12: Berechnen Sie einführende charakteristische Polynome in Numbas.

Hier sind ein paar Beispiele, wie man allgemein Eigenwerte berechnet:

Beispiele 8.13: a) Für $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ ist

$$\begin{aligned} \chi_A(t) = \det(tI - A) &= \begin{vmatrix} t-1 & -1 \\ +2 & t-4 \end{vmatrix} \\ &= (t-1)(t-4) + 2 = t^2 - 5t + 6 = (t-2)(t-3) \end{aligned}$$

Also sind 2 und 3 die Eigenwerte von A, was wir in einem vorherigen Beispiel schon gesehen haben.

b) Wir betrachten $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, zunächst geometrisch: Als Abbildung $f_R \colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ schickt f_R den ersten Standardbasisvektor $e_1 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ auf $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = e_2$, und e_2 auf $-e_1$. Dies ist also eine Drehung in $\mathbb{R}^2$ um $90°$ gegen den Uhrzeigersinn. Wir erwarten also, dass es keine Eigenwerte gibt, weil kein Vektor die Richtung beibehält.
Wir berechnen das charakteristische Polynom:

$$\chi_R(t) = \det(tI - R) = \begin{vmatrix} t & 1 \\ -1 & t \end{vmatrix} = t^2 + 1$$

Dies hat keine reellen Nullstellen, also in der Tat keine reellen Eigenwerte.

⚠ **Vorsicht**: Vergessen Sie nicht, die Vorzeichen der Einträge zu ändern, wenn Sie die Einträge von $\det(tI - A)$ aufschreiben.

Bei einem quadratischen Polynom (also von einer 2×2-Matrix) können wir die Nullstellen leicht finden, zum Beispiel mit p, q-Formel. Wie findet man die Nullstellen eines Polynoms von höherem Grad? Dazu benutzen wir Polynomdivision.

Beispiel 8.14: (Polynomdivision)
Polynomdivision lässt sich besser verbal erklären. Einige kennen es bestimmt aus der Schule. Ich schreibe hier ein paar Beispiele auf, erkläre sie in der Vorlesung verbal, und mache (wenn hoffentlich die Zeit reicht) ein Video dazu.

Wir faktorisieren $t^4 - 10t^3 + 25t^2 - 50t + 24$. Normalerweise rät man erst eine Nullstelle (z.B. 1, -1, 2 etc), aber wir zeigen hier einfach etwas Polynomdivision mit verschiedenen Beispielen.

$$\begin{array}{ccccccccc}
(t^4 & -10t^3 & +35t^2 & -50t & +24) & : & (t-2) & = & t^3 - 8t^2 + 19t - 12 \\
t^4 & -2t^3 & & & & & & & \\ \hline
0 & -8t^3 & & & & & & & \\
 & -8t^3 & +16t^2 & & & & & & \\ \hline
 & 0 & +19t^2 & & & & & & \\
 & & 19t^2 & -38t & & & & & \\ \hline
 & & 0 & -12t & & & & & \\
 & & & -12t & 24 & & & & \\ \hline
 & & & 0 & +0 & & & &
\end{array}$$

Diesen Faktor von Grad 3 können wir weiter faktorisieren:

$$\begin{array}{ccccccccc}
(t^3 & -8t^2 & +19t & -12) & : & (t-4) & = & t^2-4t+3 \\
t^3 & -4t^2 & & & & & & \\
\hline
0 & -4t^2 & & & & & & \\
 & -4t^2 & +16tt & & & & & \\
\hline
 & 0 & 3t & & & & & \\
 & & 3t & -12 & & & & \\
\hline
 & & 0 & +0 & & & &
\end{array}$$

Also bekommen wir

$$\begin{aligned}
t^4 - 10t^3 + 25t^2 - 50t + 24 &= (t-2)(t^3 - 8t^2 + 19t - 12) \\
&= (t-2)(t-4)(t^2 - 4t + 3) = (t-2)(t-4)(t-1)(t-3).
\end{aligned}$$

Hier sind noch zwei Beispiele:

$$\begin{array}{ccccccccc}
(t^3 & & -3t & +2) & : & (t-1) & = & t^2+t-2 \\
t^3 & -t^2 & & & & & & \\
\hline
0 & t^2 & & & & & & \\
 & t^2 & -t & & & & & \\
\hline
 & 0 & -2t & & & & & \\
 & & -2t & +2 & & & & \\
\hline
 & & 0 & +0 & & & &
\end{array}$$

also ist $t^3 - 3t + 2) = (t-1)(t^2 + t - 2) = (t-1)(t-1)(t+2)$.
Und dann

$$\begin{array}{ccccccccc}
(t^3 & -t^2 & +t & -1) & : & (t-1) & = & t^2+1 \\
t^3 & -t^2 & & & & & & \\
\hline
0 & 0 & & & & & & \\
 & & t & -1 & & & & \\
\hline
 & & 0 & +0 & & & &
\end{array}$$

also ist $t^3 - t^2 + t - 1 = (t-1)(t^2+1)$. Das hätte man übrigens auch anders faktorisieren können:

$$t^3 - t^2 + t - 1 = t^3 + t - t^2 - 1 = t(t^2+1) + (-1)(t^2+1) = (t-1)(t^2+1)$$

Übung 8.15: Faktorisieren Sie Polynome in Numbas.

Eine Alternative zu Polynomdivision gibt es manchmal, wenn es schöne Nulleinträge gibt, so dass man durch Entwicklung schon einen Faktor bekommt.

Beispiel 8.16: Sei $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$. Zur Berechnung des charakterischen Polynoms entwickeln wir in der zweiten Spalte.

$$\begin{aligned}
\chi_A(t) &= \begin{vmatrix} t & 0 & 2 \\ -1 & t-2 & -1 \\ -1 & 0 & t-3 \end{vmatrix} = (t-2)\begin{vmatrix} t & 2 \\ -1 & t-3 \end{vmatrix} \\
&= (t-2)(t(t-3)+2) = (t-2)(t^2-3t+2) = (t-2)(t-2)(t-1)
\end{aligned}$$

Also gibt es Eigenwerte 2 und 1.

Das Theorem 8.8 zeigt auch:

Proposition 8.17: *Ähnliche Matrizen haben die gleichen Eigenwerte und das gleiche charakteristische Polynom.*

BEWEIS. Übung. (Alle Zutaten sind im Beweis von 8.8 schon vorhanden, Sie müssen sie nur zusammensetzen.) □

C. Eigenvektoren finden

Nun können wir Eigenwerte finden, und wollen als nächstes Eigenvektoren finden.
Wir haben oben schon gesehen, dass $\mathrm{Eig}(f,\lambda) = \mathrm{Ker}(f - \lambda\mathrm{id})$, und wir wissen schon, wie man einen Kern berechnet und dafür eine Basis findet.

Um **Eigenwerte und Eigenvektoren einer Matrix A zu finden**,
- ⋄ berechne das charakteristische Polynom $\chi_A(t) = \det(tI - A)$;
- ⋄ faktorisiere das charakt. Polynom, um die Nullstellen zu finden: dies sind die Eigenwerte.
- ⋄ Für jeden Eigenwert λ separat finde eine Basis von $\mathrm{Ker}(A - \lambda I)$.

Beispiele 8.18:

a) Wir betrachten die Matrix $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$ für die wir schon das charakteristiche Polynom $\chi_A(t) = (t-2)(t-2)(t-1)$ berechnet haben, also sind die Eigenwerte 2 und 1.
Eigenvektoren für $\lambda = 2$:
$$A - 2I = \begin{pmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$
Also ist eine Basis von $\mathrm{Eig}(2)$ gegeben durch $v_1 = \left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right)$. Also hat ein beliebiger Eigenvektor zu Eigenwert 2 die Form $\left(\begin{smallmatrix} s \\ t \\ -s \end{smallmatrix}\right)$.
Eigenvektoren für $\lambda = 1$:
$$A - I = \begin{pmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$
Also ist eine Basis von $\mathrm{Eig}(1)$ gegeben durch $v_3 = \left(\begin{smallmatrix} -2 \\ 1 \\ 1 \end{smallmatrix}\right)$.

b) Eigenwerte müssen keine ganzen Zahlen sein – wir haben sie bloß in Übungen oft als ganze Zahlen, damit es leichter zu rechnen ist, oder zum Beispiel leichter in Numbas einzugeben. Die meisten Matrizen haben keine ganzzahligen Eigenwerte.
Sei $A = \begin{pmatrix} 0 & 3 \\ 4 & 0 \end{pmatrix}$, dann ist $\chi_A(t) = \begin{vmatrix} t & -3 \\ -4 & t \end{vmatrix} = t^2 - 12 = (t - 2\sqrt{3})(t + 2\sqrt{3})$, also sind die Eigenwerte $2\sqrt{3}$ und $-2\sqrt{3}$.
Nun finden wir die Eigenvektoren: für $\lambda = 2\sqrt{3}$ berechnen wir den Kern von
$$A - 2\sqrt{3}I = \begin{pmatrix} -2\sqrt{3} & 3 \\ 4 & -2\sqrt{3} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -\frac{\sqrt{3}}{2} \\ 0 & 0 \end{pmatrix}$$

Also ist der zugehörige Eigenvektor $\begin{pmatrix} \frac{\sqrt{3}}{2} \\ 1 \end{pmatrix}$ (oder jedes Vielfache davon). Zum Beispiel, wenn wir keine Brüche wollen, können wir auch $\begin{pmatrix} \sqrt{3} \\ 2 \end{pmatrix}$ nehmen. Genauer gesagt: $\begin{pmatrix} \sqrt{3} \\ 2 \end{pmatrix}$ ist eine Basis des Eigenraums $\mathrm{Eig}(2\sqrt{3})$.
Für $\lambda = -2\sqrt{3}$ berechnen wir den Kern von

$$A + 2\sqrt{3}I = \begin{pmatrix} 2\sqrt{3} & 3 \\ 4 & 2\sqrt{3} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ 0 & 0 \end{pmatrix}$$

Also ist $\begin{pmatrix} -\sqrt{3} \\ 2 \end{pmatrix}$ eine Basis des Eigenraums $\mathrm{Eig}(-2\sqrt{3})$.

Übung 8.19: Üben Sie in Numbas:
- ◇ Charakteristisches Polynom, Eigenwerte, Eigenvektoren von 2×2-Matrizen
- ◇ Charakteristisches Polynom, Eigenwerte, Eigenvektoren von 3×3-Matrizen

D. Diagonalisierung

Wozu sind nun Eigenvektoren gut? Wir haben am Anfang des Kapitels gesagt, dass wir eine besonders schöne Basis suchen, bezüglich der wir eine gegebene lineare Abbildung darstellen können.

Proposition 8.20: (Basis aus Eigenvektoren gibt Diagonalmatrix.)
Wenn $\mathcal{B}\colon v_1, v_2, \ldots, v_n$ eine Basis von V ist, die aus Eigenvektoren von $f\colon V \longrightarrow V$ besteht, dann ist die Darstellungsmatrix ${}_{\mathcal{B}}(f)_{\mathcal{B}}$ bezüglich dieser Basis diagonal.

BEWEIS. Angenommen $f(v_k) = \lambda_k v_k$ (wobei die λ_k nicht unbedingt alle verschieden sind). Also ist der Koordinatenvektor ${}_{\mathcal{B}}(f(v_k)) = \lambda_k e_k$. Die Definition der Darstellungsmatrix gibt uns dann

$$A = \Big({}_{\mathcal{B}}(f(v_1)) \quad \cdots \quad {}_{\mathcal{B}}(f(v_k)) \quad \cdots \quad {}_{\mathcal{B}}(f(v_n)) \Big) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & 0 \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}$$

was eine Diagonalmatrix ist. □

Dies ist die best-mögliche Form, die wir erwarten können: sie sagt uns genau, in welche Richtungen die lineare Abbildung um wieviel streckt.
Allerdings können wir nicht immer so eine schöne Form erwarten. Wir hatten schon ein Beispiel, wo es gar keine reellen Eigenwerte gab (Drehung). Dies ist aber nicht das einzige Problem.

Beispiel 8.21: Wir betrachten die lineare Abbildung $f_A\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ mit Matrix

$$A = \begin{pmatrix} 1 & -3 & 2 \\ -1 & -5 & 6 \\ 2 & -2 & 0 \end{pmatrix}$$

Dann ist
$$\chi_A(t) = \begin{vmatrix} t-1 & 3 & -2 \\ 1 & t+5 & -6 \\ -2 & 2 & t \end{vmatrix} = t^2(t+4).$$
Also sind 0 und -4 die Eigenwerte. Wir berechnen die Eigenvektoren, zuerst für $\lambda = 0$:
$$\begin{pmatrix} 1 & -3 & 2 \\ -1 & -5 & 6 \\ 2 & -2 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -3 & 2 \\ 0 & -8 & 8 \\ 0 & 4 & -4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$
Also ist eine Basis von $\mathrm{Eig}(0)$ gegeben durch $\left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$.
Für $\lambda = -4$:
$$A + 4I = \begin{pmatrix} 5 & -3 & 2 \\ -1 & -1 & 6 \\ 2 & -2 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -6 \\ 0 & -8 & 32 \\ 0 & -4 & 16 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -6 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{pmatrix}$$
Also ist eine Basis von $\mathrm{Eig}(-4)$ gegeben durch $\left(\begin{smallmatrix} 2 \\ 4 \\ 1 \end{smallmatrix}\right)$.
Wir haben hier also nur 2 linear unabhängige Eigenvektoren, und können aus diesen keine Basis aus Eigenvektoren für $\mathbb{R}^3$ bekommen.

Das Problem kommt von der doppelten Nullstelle 0, für die es trotzdem nur einen ein-dimensionalen Eigenraum gibt.

Definition 8.22: Die **algebraische Vielfachheit** eines Eigenwerts λ von f ist die Potenz des Linearfaktors $(t-\lambda)$ im charakteristischen Polynom, und die **geometrische Vielfachheit** von λ ist die Dimension des Eigenraums $\mathrm{Eig}(\lambda)$.

Beispiel 8.23: Bei Matrix $A = \begin{pmatrix} 1 & -3 & 2 \\ -1 & -5 & 6 \\ 2 & -2 & 0 \end{pmatrix}$ hat der Eigenwert 0 algebraische Vielfachheit 2 und geometrische Vielfachheit 1, und Eigenwert -4 hat algebraische und geometrische Vielffachheit 1.

Es ist übrigens umgangssprachlicher Usus, dass man sagt: „λ_1 hat nur einen Eigenvektor“ oder „λ_2 hat zwei Eigenvektoren“. Was damit gemeint ist: Der Eigenraum von λ_1 hat eine Basis aus einem Element, also $\dim \mathrm{Eig}(\lambda_1) = 1$, und der Eigenraum von λ_2 hat eine Basis mit zwei Elementen, also $\dim \mathrm{Eig}(\lambda_2) = 2$. Die Information einer Basis des Eigenraums reicht ja schon aus, um den ganzen Eigenraum zu kennen. Aber wenn man es ganz genau nimmt, hat natürlich jeder Eigenwert unendlich viele Eigenvektoren, nämlich auch alle Vielfachen des einen angegebenen Eigenvektors, beziehungsweise jede Linearkombination der zwei angegebenen Eigenvektoren, wenn „es zwei gibt“.

Nach der Definition ist sofort klar, dass

Wenn λ ein Eigenwert von f ist, dann ist
- ⋄ die algebraische Vielfachheit $\geqslant 1$ und
- ⋄ die geometrische Vielfachheit $\geqslant 1$.

Sonst wäre λ ja gar kein Eigenwert.
Außerdem ist für jeden Eigenwert λ die geometrische Vielfachheit höchstens so groß wie die algebraische. Das kann man als zusammenfassen als:

Für jeden Eigenwert λ gilt:
$$1 \leqslant \text{geometrische Vielfachheit} \leqslant \text{algebraische Vielfachheit}$$

Das kann man auch beweisen, wir tun es hier aus Zeitgründen nicht (vielleicht auf einem Übungsblatt).

Wir definieren jetzt also erst einmal, was wir gerne untersuchen möchten.

Definition 8.24: Eine lineare Abbildung $f: V \longrightarrow V$ ist **diagonalisierbar** wenn es eine Basis von V gibt, bezüglich derer die Matrix von f diagonal ist.
Eine $n \times n$-Matrix A ist **diagonalisierbar** wenn sie zu einer Diagonalmatrix ähnlich ist.

Dies heißt, dass es eine invertierbar Matrix P (eine Basiswechselmatrix) gibt, sodass $A = P^{-1}DP$, wobei D diagonal ist. Man kann Diagonalisierbarkeit auch so ausdrücken:

Theorem 8.25: *Eine lineare Abbildung $f: V \longrightarrow V$ ist diagonalisierbar genau dann wenn es eine Basis von V gibt, die aus Eigenvektoren von f besteht.*

BEWEIS. Wir haben schon gesehen (Prop. 8.20), dass eine Basis aus Eigenvektoren eine diagonale Darstellungsmatrix gibt. Umgekehrt, wenn $\mathcal{B}: v_1, \ldots, v_n$ eine Basis von V ist, die eine diagonale Darstellungsmatrix ${}_{\mathcal{B}}(f)_{\mathcal{B}}$ mit Diagonaleinträgen $\lambda_1, \ldots, \lambda_k$ gibt, dann ist $f(v_k) = \lambda_k v_k$ nach Definition der Darstellungsmatrix. Also ist jedes v_k ein Eigenvektor von f. □

Wir haben ja oben schon gesehen, dass jeder Eigenwert mindestens einen ein-dimensionalen Eigenraum hat. Und wenn jede algebraische Vielfachheit gleich 1 ist, dann *muss* auch jede geometrische Vielfachheit genau 1 sein. Es ist aber nicht ganz offensichtlich, dass zum Beispiel drei Eigenvektoren von drei verschiedenen Eigenwerten wirklich linear unabhängig sind (obwohl wir das vielleicht intuitiv auch schon vermuten).

Proposition 8.26: (Eigenvektoren für verschiedene Eigenwerte)
Sei $f: V \longrightarrow V$ linear. Wenn $v_1, \ldots, v_k$ Eigenvektoren mit paarweise verschiedenen Eigenwerten sind, dann sind sie linear unabhängig.

BEWEIS. Für Interessierte, nicht offizieller Teil des Kurses.
Wir haben $f(v_i) = \lambda_i v_i$, wobei alle λ_i paarweise verschieden sind. Angenommen $v_1, \ldots, v_k$ sind linear abhängig (wir wollen also einen Widerspruch erzeugen). Da ein einzelner Vektor v_1 linear unabhängig ist (weil $v_1 \neq 0$), gibt es ein größtes m, sodass $v_1, \ldots, v_m$ linear unabhängig, aber $v_1, \ldots, v_m, v_{m+1}$ linear abhängig sind. Nach unserer Annahme ist $m + 1 \leqslant k$. Dann gibt es Skalare μ_i nicht alle Null mit

$$\mu_1 v_1 + \cdots + \mu_m v_m + \mu_{m+1} v_{m+1} = 0.$$

Wenn wir f auf beiden Seiten der Gleichung anwenden, bekommen wir

$$\mu_1 \lambda_1 v_1 + \cdots + \mu_m \lambda_m v_m + \mu_{m+1} \lambda_{m+1} v_{m+1} = 0.$$

Wenn wir nun λ_{m+1} mal die erste Gleichung von dieser zweiten Gleichung abziehen, bekommen wir

$$\mu_1(\lambda_1 - \lambda_{m+1})v_1 + \cdots + \mu_m(\lambda_m - \lambda_{m+1})v_m + 0v_{m+1} = 0.$$

Da all λ_i verschieden sind und $v_1, \ldots, v_m$ linear unabhängig sind, folgt hieraus $\mu_1 = \cdots = \mu_m = 0$. Dann ist aber auch $\mu_{m+1} = 0$, was unserer Annahme widerspricht, dass $v_1, \ldots, v_{m+1}$ linear unabhängig sind.
Also sind $v_1, \ldots, v_k$ linear abhängig. □

Beispiele 8.27:

a) $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ hat Eigenwerte 2 und 3, und somit sind die dazugehörigen Eigenvektoren $\binom{1}{1}$ und $\binom{1}{2}$ linear unabhängig.

b) $A = \begin{pmatrix} 5 & 1 & 3 \\ 0 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$ hat Eigenwerte -1 mit Eigenvektor $\left(\begin{smallmatrix} 0 \\ 3 \\ -1 \end{smallmatrix}\right)$, 2 mit Eigenvektor $\left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right)$ und 5 mit Eigenvektor $\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$. Das Resultat sagt uns also, dass diese drei Vektoren zusammen linear unabhängig sind. Da es auch drei Vektoren in $\mathbb{R}^3$ sind, bilden sie also eine Basis von $\mathbb{R}^3$ (nach „zwei für eins bei Basen“, Prop. 5.11).

c) $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$ hat Eigenwert 2 mit Eigenvektoren $v_1 = \left(\begin{smallmatrix} -1 \\ 0 \\ 1 \end{smallmatrix}\right)$ und $v_2 = \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right)$, und Eigenwert 1 mit Eigenvektor $v_3 = \left(\begin{smallmatrix} -2 \\ 1 \\ 1 \end{smallmatrix}\right)$. Also wissen wir nach dem Resultat, dass v_1, v_3 linear unabhängig sind, und auch v_2, v_3. Dieses Resultat sagt uns aber nichts über alle drei zusammen, weil v_1 und v_2 nicht unterschiedliche Eigenwerte haben.

Korollar 8.28: *Wenn $\dim V = n$ und $f\colon V \longrightarrow V$ genau n verschiedene Eigenwerte hat, dann ist f diagonalisierbar.*

Beispiele 8.29: Beispiele 8.27(a) und (b) haben jeweils n verschiedene Eigenwerte und sind damit diagonalisierbar.

⚠ Vorsicht! Dies ist aber nur ein Resultat in eine Richtung. Es gibt durchaus auch Abbildungen, die weniger Eigenwerte haben und doch diagonalisierbar sind.

Beispiel 8.30: Matrix A aus Beispiel 8.27(c) hat nur zwei Eigenwerte, 1 mit algebraischer Vielfachheit 1 und 2 mit algebraischer Vielfachheit 2. Aber Eigenwert 2 hat auch geometrische Vielfachheit 2 (weil $\mathrm{Eig}(2) = \langle v_1, v_2 \rangle$ Dimension 2 hat), und die dort angegebenen Vektoren v_1, v_2, v_3 bilden eine Basis aus Eigenvektoren, und somit ist A diagonalisierbar.

Bei einer Diagonalmatrix kann man leicht die algebraischen und geometrischen Vielfachheiten bestimmen:

Beispiel 8.31: Eine Diagonalmatrix D mit Diagonaleinträgen $d_1, \ldots, d_n$, nicht unbedingt verschieden, hat Eigenwerte $d_1, \ldots, d_n$, weil $\chi_D(t) = (t - d_1) \cdots (t - d_n)$ ist. Die algebraische Vielfachheit eines Eigenwerts λ ist also die Anzahl von Diagonaleinträgen, die gleich λ sind. Da jeder Vektor der Basis, bezüglich der D geschrieben ist, ein Eigenvektor ist, ist die geometrische Vielfachheit von λ auch die Anzahl von Diagonaleinträgen, die gleich λ sind.

Zum Beispiel hat $f\colon \mathbb{R}^5 \longrightarrow \mathbb{R}^5$ mit ${}_{\mathcal{B}}(f)_{\mathcal{B}} = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{smallmatrix}\right)$ hat Eigenwert 1 mit algebraischer Vielfachheit 3 und $\dim(\mathrm{Eig}(1)) = 3$, und Eigenwert 2 mit algebraischer Vielfachheit 2 und $\dim(\mathrm{Eig}(2)) = 2$.

Wir haben gesehen, dass ähnliche Matrizen die gleichen Eigenwerte und charakaterischen Polynome haben. Wir können nicht erwarten, dass sie die gleichen Eigenvektoren haben: selbst wenn wir die beiden Matrizen als verschiedene Darstellungsmatrizen der gleichen Abbildung betrachten, haben wir ja durch den Basiswechsel nun andere Spaltenvektoren. Aber die *Dimension* der Eigenräume ist hier gleich. Und das zusammen mit der Überlegung über diagonale Matrizen hilft uns dann, ein Kriterium für Diagonalisierbarkeit zu formulieren.

Proposition 8.32: *Ähnliche Matrizen haben die gleichen Eigenraumdimensionen, also die gleichen geometrischen Vielfachheiten (der gleichen Eigenwerte).*

BEWEIS. Angenommen $A = P^{-1}BP$ für eine invertierbare Matrix P. Wir haben schon gezeigt, dass A und B die selben Eigenwerte haben (Prop. 8.17). Sei $\mathrm{Eig}(A,\lambda)$ der λ-Eigenraum von A, und $\mathrm{Eig}(B,\lambda)$ der λ-Eigenraum von B. Dann gilt

$$Av = \lambda v \qquad \Leftrightarrow \qquad P^{-1}BPv = \lambda v \qquad \Leftrightarrow \qquad B(Pv) = \lambda(Pv).$$

Also ist $v \in \mathrm{Eig}(A,\lambda)$ genau dann wenn $Pv \in \mathrm{Eig}(B,\lambda)$. Wenn also $v_1,\ldots,v_k$ eine Basis von $\mathrm{Eig}(A,\lambda)$ ist, dann sind $Pv_1,\ldots,Pv_k$ alle in $\mathrm{Eig}(B,\lambda)$. Da P invertierbar ist, sind diese Vektoren immer noch linear unabhängig. („Versteckte Übung“, vergleiche Prop. 4.32.)
Also ist $\dim \mathrm{Eig}(B,\lambda) \geqslant \dim \mathrm{Eig}(A,\lambda)$. Das selbe Argument funktioniert für eine Basis $w_1,\ldots,w_l$ für $\mathrm{Eig}(B,\lambda)$ und P^{-1}, also gilt a uch $\dim \mathrm{Eig}(A,\lambda) \geqslant \dim \mathrm{Eig}(B,\lambda)$. Also sind die Dimensionen gleich. □
Hintergrund hier: wenn wir den Isomorphismus f_P auf $\mathrm{Eig}(A,\lambda)$ einschränken, bekommen wir einen Isomorphismus $P'\colon \mathrm{Eig}(A,\lambda) \longrightarrow \mathrm{Eig}(B,\lambda)$.

Also, da diagonale Matrizen für jeden Eigenwert jeweils gleiche algebraische und geometrische Vielfachheiten haben, und diese Vielfachheiten unter Ähnlichkeit gleich bleiben, muss dies auch eine Bedingung für diagonalisierbare Matrizen sein. So ganz reicht dies nicht alleine: wir brauchen auch „genug Eigenwerte“, um die Diagonale zu füllen. Wir formulieren also verschiedene Diagonalisierbarkeitskriterien.

Theorem 8.33: (Diagonalisierbarkeitskriterien)
Eine lineare Abbildung $f\colon V \longrightarrow V$ mit $\dim V = n$ ist diagonalisierbar genau dann wenn

- ◇ *V eine Basis aus Eigenvektoren von f hat, genau dann wenn*
- ◇ *die Summe der geometrischen Vielfachheiten der verschiedenen Eigenwerte gleich n ist, genau dann wenn*
- ◇ *die Summe der algebraischen Vielfachheiten der verschiedenen Eigenwerte gleich n ist und für jeden Eigenwert die geometrische Vielfachheit gleich der algebraischen Vielfachheit ist, genau dann wenn*
- ◇ *das charakteristische Polynom in Linearfaktoren zerfällt und für jeden Eigenwert die geometrische Vielfachheit gleich der algebraischen Vielfachheit ist.*

Wir können die Diagonalisierbarkeitskriterien noch einmal von der andere Richtung beleuchten:

Hürden zur Diagonalisierbarkeit
Es gibt zwei Situationen, in denen eine lineare Abbildung *nicht* diagonalisierbar ist:

- ◇ Es gibt nicht genug Eigen*werte*: das charakteristische Polynom hat nicht genug Nullstellen, oder anders ausgedrückt, die Summe der algebraischen Vielfachheiten ist zu klein.
- ◇ Es gibt nicht genug Eigen*vektoren*: selbst wenn es genug Eigenwerte gibt, kann es sein, dass ein Eigenwert mit algebraischer Vielfachheit > 1 nicht genug Eigenvektoren hat, also dass für einen Eigenwert die geometrische Vielfachheit kleiner als die algebraische Vielfachheit ist. Anders ausgedrückt: die Summe der geometrischen Vielfachheiten ist zu klein.

Die Diagonalisierbarkeitskriterien stellen also sicher, dass beide diese Hürden überwunden werden. Wenn also eine Matrix oder lineare Abbildung diagonalisierbar ist, wie sieht dann die Diagonalform aus? Und welche Basiswechselmatrix gibt sie uns?

Theorem 8.34: (Diagonale Form und Eigenvektorbasiswechsel)
Wenn $f\colon V \longrightarrow V$ diagonalisierbar ist, dann hat die Diagonalform D alle Eigenwerte von f auf der Diagonale, entsprechend der algebraischen Vielfachheiten.
Des Weiteren gilt für $V = \mathbb{R}^n$: wenn $A = {}_{\mathcal{E}}(f)_{\mathcal{E}}$ die Darstellungsmatrix von f bezüglich Standardbasis ist und $P = {}_{\mathcal{E}}P_{\mathcal{B}}$ die Matrix, deren Spalten die Eigenvektoren von f sind (in der Reihenfolge entsprechend der Diagonaleinträge von D), dann ist $D = P^{-1}AP$.

BEWEIS. Dies folgt direkt aus Theorem 8.25: f ist diagonalisierbar genau dann wenn es eine Basis aus Eigenvektoren gibt. □

Beispiele 8.35: Jeweils aus Beispiele 8.27:

a) $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ hat Eigenvektorbasiswechselmatrix $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, und dann ist

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = D = P^{-1}AP.$$

b) $A = \begin{pmatrix} 5 & 1 & 3 \\ 0 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$ hat $P = \begin{pmatrix} 0 & 1 & 1 \\ 3 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix}$, und dann ist $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} = D = P^{-1}AP$.

c) $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$ hat $P = \begin{pmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, und dann ist $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = D = P^{-1}AP$.

Am Ende dieses Kapitels geben wir noch eine Zusammenfassung von Eigenschaften, die unter Ähnlichkeit invariant sind. Man kann auch sagen, dies sind Eigenschaften einer linearen Abbildung und nicht nur einer Matrix.

Eigenschaften einer linearen Abbildung (also Invarianten unter Ähnlichkeit)

Eigenschaft	**Kommentar**
Dimension des Bildes	Basiswechsel ändert Dimension nicht
Dimension des Kerns	
Invertierbarkeit	z.B. via Dimensionen von Kern und Bild
Determinante	wegen $\det(P^{-1}AP) = \det A$
Spur	wegen $\mathrm{Spur}(P^{-1}AP) = \mathrm{Spur}\, A$
Eigenwerte	via Determinante und charakteristisches Polynom
Charakteristisches Polynom	also auch algebraische Vfh der Eigenwerte
Dimension der Eigenräume	also geometrische Vfh der Eigenwerte

E. Eigenwerte und Eigenvektoren: Study guide

Konzeptübersicht.

- ◇ Eigenwert, Eigenvektor, Eigenraum
- ◇ Verbindung von Kern und Eigenräumen
- ◇ Verbindung von Isomorphismen und Eigenwerten
- ◇ Characteristisches Polynom und Bedeutung für Eigenwerte
- ◇ Eigenwerte von Diagonal- und Dreiecksmatrizen
- ◇ Verbindung von Eigenvektorbasis und Diagonalmatrix
- ◇ Algebraische und geometrische Vielfachheiten von Eigenwerten, und deren Beziehungen
- ◇ Diagonalisierbarkeit, verschiedene Diagonalisierbarkeitskriterien
- ◇ Diagonale Form und Eigenvektorbasiswechsel
- ◇ Eigenschaften, die unter Ähnlichkeit erhalten bleiben

Skills.

- ◇ Charakteristisches Polynom einer Abbildung/Matrix finden
- ◇ Eigenwerte und Eigenvektoren einer Abbildung/Matrix finden
- ◇ Algebraische und geometrische Vielfachheiten von Eigenwerten bestimmen
- ◇ Bestimmen, ob eine Abbildung/Matrix diagonalisierbar ist
- ◇ Diagonalform und dazugehörigen Basiswechsel für diagonalisierbare Abbildung/Matrix finden

KAPITEL 9

Abstände und Winkel

Wir haben bisher in Vektorräumen gearbeitet, ohne ein Konzept von Längen oder Winkeln zu haben. Damit konnten wir auch schon einiges an Geometrie machen (Darstellung von Geraden und Ebenen, Schnittmengen solcher, geometrische Interpretation bestimmter Abbildungen, insbesondere geometrische Bedeutung von Eigenvektoren und Eigenwerten). Aber wenn wir nun Abstände und Winkel betrachten wollen, brauchen wir eine Extrastruktur auf dem Vektorraum, die heißt Skalarprodukt.
Wir werden uns in diesem Kapitel auf die Vektorräume $\mathbb{R}^n$ beschränken. Man kann das Konzept des Skalarprodukts auch allgemeiner definieren und viele verschiedene Beispiele betrachten, aber wir schauen uns nur dieses eine, das wichtigste, an.

A. Standardskalarprodukt

Wir kennen vielleicht aus der Schule im $\mathbb{R}^2$ oder $\mathbb{R}^3$ das Konzept des *Skalarprodukts* und der *Länge* von Vektoren:

- ◇ Für $x = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$ und $y = \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)$ ist „$x \cdot y$“ $= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1y_1 + x_2y_2$.
- ◇ Die Länge ist $\|x\| = \sqrt{x_1^2 + x_2^2}$ (nach Pythagoras).

Wir haben hier das Skalarprodukt als Matrixmultiplikation geschrieben: Vektoren kann man ja nicht einfach miteinander multiplizieren, deswegen werden wir auch nicht mehr „$x \cdot y$“ schreiben, aber wir können einen Zeilenvektor mit einem Spaltenvektor per Matrixmultiplikation multiplizieren und bekommen dann eine 1×1-Matrix, also eine reelle Zahl.
Mit dieser Schreibweise können wir auch sehr leicht nach $\mathbb{R}^n$ verallgemeinern:

Definition 9.1: Das **Standardskalarprodukt** (auch euklidisches Skalarprodukt) auf $\mathbb{R}^n$ ist die Funktion $\langle -, - \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ definiert durch

$$\begin{aligned} \langle v, w \rangle = v^T w &= \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \\ &= v_1w_1 + v_2w_2 + \cdots + v_nw_n = \sum_{i=1}^{n} v_iw_i. \end{aligned}$$

Die **Norm** (oder Länge) eines Vektors in $\mathbb{R}^n$ ist definiert als

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Beispiel 9.2:

- ◇ Vektor $v = \left(\begin{smallmatrix} 2 \\ -1 \\ 3 \\ -5 \end{smallmatrix}\right)$ hat $\|v\| = \sqrt{2^2 + (-1)^2 + 3^2 + (-5)^2)} = \sqrt{39}$.
- ◇ Für v wie oben und $w = \left(\begin{smallmatrix} -3 \\ -4 \\ 1 \\ 0 \end{smallmatrix}\right)$ ist $\langle v, w \rangle = 2 \cdot (-3) + (-1) \cdot (-4) + 3 \cdot 1 + (-5) \cdot 0 = 2$.

Es bleibt Ihnen überlassen, ob Sie lieber $\langle v, w\rangle$ oder $v^T w$ schreiben.
⚠ Aber wir schreiben nun nicht $v \cdot w$, da dies zu Verwirrung führen kann. Die Matrixmultiplikationsschreibweise $v^T w$ ist ja wirklich nicht viel länger, und führt zur korrekten Verwendung.

Übung 9.3: Üben Sie in Numbas: Einführend: Skalarprodukt und Norm

Damit diese Norm wirklich eine Länge darstellt, sollte sie bestimmte Eigenschaften besitzen:

- Die Norm eines Vektors ist nicht-negativ.
- Der Nullvektor ist der einzige Vektor mit Norm 0.
- Multiplikation eines Vektors mit einem Skalar multpiliziert die Norm des Vektors mit dem Betrag des Skalars.

Wir erwarten vielleicht auch so etwas wie eine Dreiecksungleichung.

Wir werden auch einige Eigenschaften des Skalarprodukts sehen, nämlich: Das Skalarprodukt

- ist linear in jedem Eintrag.
- ist symmetrisch.
- kodiert Orthogonalität und Winkel.

Da wir die Norm über das Skalarprodukt definiert haben, fangen wir mit den Skalarprodukteigenschaften an.

Theorem 9.4: (Eigenschaften des Skalarprodukts)
Das Standardskalarprodukt $\langle -, -\rangle\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ erfüllt:
Für alle $u, v, w \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$ gilt

(i) $\langle \lambda u + \mu v, w\rangle = \lambda\langle u, w\rangle + \mu\langle v, w\rangle$ *und* $\langle v, \lambda u + \mu w\rangle = \lambda\langle v, u\rangle + \mu\langle v, w\rangle$ ***(bilinear)***
(ii) $\langle v, w\rangle = \langle w, v\rangle$ ***(symmetrisch)***
(iii) $\langle v, v\rangle \geqslant 0$, *und* $\langle v, v\rangle = 0 \Leftrightarrow v = 0$ ***(positiv definit)***

Beweis. (i) Matrixmultiplikation ist linear (Prop. 1.63). Das gibt

$$\langle v, \lambda u + \mu w\rangle = v^T(\lambda u + \mu w) = \lambda v^T u + \mu v^T w = \lambda\langle v, u\rangle + \mu\langle v, w\rangle.$$

Wir bekommen die zweite Aussage, nachdem wir Symmetrie gezeigt haben.

(ii) $\langle v, w\rangle = v^T w = (v^T w)^T = w^T v = \langle w, v\rangle$. Wir benutzen hier: wenn wir eine 1×1-Matrix (also eine Zahl) transponieren, ändert sie sich nicht.
Wir könnten es auch explizit ausschreiben als $\langle v, w\rangle = \sum_{i=1}^n v_i w_i = \sum_{i=1}^n w_i v_i = \langle w, v\rangle$.

(iii) $\langle v, v\rangle = \sum_{i=1}^n v_i v_i = \sum_{i=1}^n (v_i)^2$. Quadrate sind immer nicht-negativ, und die Summe von nicht-negativen Termen ist nicht-negativ. Also ist $\langle v, v\rangle \geqslant 0$.
Weiterhin ist $\langle v, v\rangle = v_1^2 + \cdots + v_n^2 = 0$ genau dann wenn jeder Summand einzeln gleich Null ist, weil ja alle nicht-negativ sind (also kann sich nichts gegenseitig aufheben). Dies ist der Fall genau dann wenn $v = 0$. □

Diese Eigenschaften des Skalarprodukts geben sofort die Eigenschaften der Norm:

Korollar 9.5: (Eigenschaften der Norm)
Die Norm auf $\mathbb{R}^n$ erfüllt: Für alle $v \in V$, $\lambda \in \mathbb{R}$ gilt
(i) $\|v\| \geqslant 0$. *(ii)* $\|v\| = 0 \Leftrightarrow v = 0$. *(iii)* $\|\lambda v\| = |\lambda|\|v\|$.

Beweis. Übung. Benutzen Sie die Eigenschaften des Skalarprodukts. □

Folgende leichte Umstellung der Definition der Norm ist super hilfreich:

$$\|v\|^2 = \langle v, v\rangle.$$

Wir können mit dieser Definition von Länge also jetzt auch Vektoren auf eine einheitliche Länge, zum Beispiel Länge 1, normieren.

Definition 9.6: Ein **Einheitsvektor** oder **normierter Vektor** ist ein Vektor mit Norm 1. Jeden Vektor $v \in \mathbb{R}^n$, $v \neq 0$ kann man zu einem Einheitsvektor $\frac{1}{\|v\|}v$ **normieren**.

Beispiele 9.7: ◇ Alle Vektoren der Standardbasis sind Einheitsvektoren.
◇ Der Vektor $v = \begin{pmatrix} 2 \\ -1 \\ 3 \\ -5 \end{pmatrix}$ aus Beispiel 9.2 hat normierten Vektor $u = \frac{1}{\|v\|}v = \frac{1}{\sqrt{39}}\begin{pmatrix} 2 \\ -1 \\ 3 \\ -5 \end{pmatrix}$.
◇ Der normierte Vektor von $v = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$ ist $\frac{1}{\sqrt{n}}v$.

Übung 9.8: Üben Sie in Numbas: Vektoren normieren.

Es wäre schön, wenn wir auch die Dreiecksungleichung für diese Norm zeigen könnten. Dazu brauchen wir erst:

Proposition 9.9: (Skalarprodukt findet 0)
Für ein $v \in \mathbb{R}^n$ gilt $\langle v, w \rangle = 0$ für alle $w \in \mathbb{R}^n$ genau dann wenn $v = 0$.

BEWEIS. Da $\langle v, w \rangle = 0$ *für alle* $w \in \mathbb{R}^n$, ist auch $\langle v, v \rangle = 0$, und somit $v = 0$. Die andere Richtung folgt, weil das Skalarprodukt linear im ersten Eintrag ist („Eigenschaften des Skalarprodukts", Theorem 9.4), und lineare Abbildungen immer Null auf Null schicken. □

Hiermit können wir jetzt beweisen:

Theorem 9.10: (Cauchy-Schwarz Ungleichung)
Für $u, v \in \mathbb{R}^n$ gilt
$$\langle u, v \rangle^2 \leqslant \|u\|^2 \|v\|^2$$
($\Rightarrow |\langle u, v \rangle| \leqslant \|u\|\|v\|$ und $\langle u, v \rangle \leqslant \|u\|\|v\|$).
Gleichheit gilt genau dann wenn $u = \mu v$ für ein $\mu \in \mathbb{R}$. (Also genau dann wenn u, v kollinear sind.)

BEWEIS. Wir betrachten $\langle u + tv, u + tv \rangle \geqslant 0$, ein quadratisches Polynom in t, welches nicht-negativ ist, weil das Skalarprodukt positiv definit ist (siehe Theorem 9.4). Dies ist die Hauptidee und der super schöne Trick dieses Beweises.
Wir haben
$$\langle u + tv, u + tv \rangle = \langle u, u \rangle + t\langle u, v \rangle + t\langle v, u \rangle + t^2\langle v, v \rangle = \|u\|^2 + 2t\langle u, v \rangle + \|v\|^2 t^2.$$
Die Nullstellen dieses quadratischen Polynoms sind
$$\frac{-2\langle u, v \rangle \pm \sqrt{(2\langle u, v \rangle)^2 - 4\|u\|^2\|v\|^2}}{2\|v\|^2}.$$
Da wir wissen, dass das Polynom nicht-negativ ist, hat es genau eine oder keine Nullstellen. Also muss die Diskriminante (der Ausdruck unter der Wurzel) nicht-positiv sein. (Diskriminante Null: genau eine Nullstelle, Diskriminante negativ: keine reellen Nullstellen) Also ist
$$4\langle u, v \rangle^2 - 4\|u\|^2\|v\|^2 \leqslant 0$$
was wir umformen können zu
$$\langle u, v \rangle^2 \leqslant \|u\|^2\|v\|^2.$$
Wenn wir hier auf beiden Seiten die Wurzel ziehen, bekommen wir $|\langle u, v \rangle| \leqslant \|u\|\|v\|$. Aber da $\langle u, v \rangle \leqslant |\langle u, v \rangle|$, können wir auch sagen
$$\langle u, v \rangle \leqslant \|u\|\|v\|.$$

Um bei der Diskriminante eine Gleichheit zu bekommen, brauchen wir $\langle u+tv, u+tv\rangle = 0$. Dies geschieht genau dann wenn $u+tv=0$, also $u=-tv$ für ein t. $\square$

Dies ist schon für sich eine schöne Aussage. Wir können sie auch benutzen für:

Korollar 9.11: (Dreiecksungleichung für Norm)
Für $v, w \in \mathbb{R}^n$ gilt $\|v+w\| \leqslant \|v\| + \|w\|$.

Beweis.

$$\begin{aligned}
\|v+w\|^2 &= \langle v+w, v+w\rangle \\
&= \|v\|^2 + 2\langle v,w\rangle + \|w\|^2 \\
&\leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 && \text{Cauchy-Schwarz} \\
&= (\|v\|+\|w\|)^2
\end{aligned}$$

Dies gibt

$$\|v+w\| \leqslant \|v\| + \|w\|$$

da beide Seiten nicht-negativ sind. $\square$

B. Orthogonalität und Winkel

Wie ist das nun mit den Winkeln? Mit einem Diagram am Einheitskreis in $\mathbb{R}^2$ können wir sehen: Für die zwei Einheitsvektoren $x = \binom{1}{0}$ und $y = \binom{y_1}{y_2}$ haben wir

$$x_1y_1 + x_2y_2 = y_1 = \cos\theta.$$

Hier ist θ der Winkel von y zur positiven x-Achse, oder auch der Winkel zwischen x und y. Wenn nun $x = \binom{x_1}{x_2}$ ein beliebiger Einheitsvektor ist, können wir das verallgemeinern:
Wenn θ_1 der Winkel von x zur positive x-Achse ist und θ_2 der Winkel von y zur positiven x-Achse, dann ist der Winkel zwischen x und y der Winkel $\theta = (\pm)(\theta_1 - \theta_2)$ (mit $0 \leqslant \theta \leqslant 180°$). Weiterhin sind $x = \binom{\cos\theta_1}{\sin\theta_1}$, $y = \binom{\cos\theta_2}{\sin\theta_2}$, und

$$x_1y_1 + x_2y_2 = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 = \cos(\theta_1-\theta_2) = \cos\theta.$$

Wir bekommen so also nur den kleineren Winkel zwischen x und y.
Insbesondere sehen wir, dass

$$x_1y_1 + x_2y_2 = 0 \Leftrightarrow x \text{ und } y \text{ sind senkrecht zueinander.}$$

Das geht nun auch alles in $\mathbb{R}^n$:

Definition 9.12: Zwei Vektoren $v, w \in \mathbb{R}^n$ heißen **orthogonal** genau dann wenn $\langle v, w\rangle = 0$. Der Winkel $\theta \in [0, 180°)$ zwischen $v, w \in \mathbb{R}^n$ ist definiert via

$$\cos\theta = \frac{\langle v,w\rangle}{\|v\|\|w\|}.$$

Nach Cauchy-Schwarz ist der Ausdruck $\frac{\langle v,w\rangle}{\|v\|\|w\|}$ immer zwischen -1 und 1, und er ist ± 1 genau dann wenn v und w kollinear sind. Dies entspricht den Winkeln $0°$ und $180°$.

Beispiele 9.13: ◇ Die Standardbasisvektoren sind paarweise zueinander orthogonal.
◇ Wir berechnen den Winkel zwischen $u_1 = \binom{3}{0}$ und $u_2 = \binom{2}{2}$:

$$\cos\theta = \frac{6+0}{\sqrt{9}\sqrt{8}} = \frac{6}{6\sqrt{2}} = \frac{1}{\sqrt{2}}, \text{ also } \theta = 45° \text{ oder } \frac{\pi}{4}.$$

◇ In $\mathbb{R}^4$ sind zum Beispiel $v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 2 \\ -1 \\ 6 \\ 2 \end{pmatrix}$ zueinander orthogonal:

$$\langle v_1, v_2\rangle = 2 - 2 - 6 + 6 = 0$$

Übung 9.14: Üben Sie in Numbas: Orthogonalität testen.

Wir können uns fragen, welche Vektoren zu einem gegebenen Vektor orthogonal sind. Dies gibt uns ein homogenes LGS mit genau einer Gleichung.

Proposition 9.15: *Sei $u \in \mathbb{R}^n$, $u \neq 0$. Die Menge $\{v \in \mathbb{R}^n \mid \langle u, v\rangle = 0\}$ aller Vektoren orthogonal zu u ist ein Unterraum von $\mathbb{R}^n$ von Dimension $n-1$. Wir nennen sie eine* ***Hyperebene****.*

BEWEIS. Seien $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ und $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Dann wird $\langle u, v\rangle = 0$ zu $u_1x_1 + \cdots + u_nx_n = 0$, eine lineare Gleichung mit n Unbekannten. Solange $u \neq 0$, hat die reduzierte ZSF dieses LGS genau eine führende Eins, also $n-1$ freie Variablen in der Lösung. Daher ist die Lösungsmenge dieses homogenen Systems ein $n-1$-dimensionaler Unterraum. □

Beispiele 9.16: ◇ Für $n = 2$ haben wir zum Beispiel für $u = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ die Gleichung

$$\langle u, v\rangle = 1 \cdot x_1 + 1 \cdot x_2 = 0 \quad \text{mit Lösungsmenge} \quad \operatorname{Span}\left(\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\right).$$

Dies ist eine Gerade durch den Ursprung, senkrecht zu u.

◇ Für $n = 3$ gibt $u = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$ die Gleichung $x_1 + x_2 + x_3 = 0$. Die Lösung dieses LGS ist

$$\operatorname{Span}\left(\left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ -1 \end{smallmatrix}\right)\right),$$

die Ebene, die orthogonal zu $\left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$ ist. Wir können also sagen, dass u ein Normalenvektor zu dieser Ebene ist.

◇ Für $n = 4$ bekommen wir dann mit $u = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \\ 1 \end{smallmatrix}\right)$ die Gleichung $x_1 + x_2 + x_3 + x_4 = 0$. Die Vektoren orthogonal zu u bilden also die 3-dimensionale Hyperebene mit Basis

$$\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \\ -1 \end{smallmatrix}\right).$$

◇ Dies kann man zu $\mathbb{R}^n$ verallgemeinern.

Übung 9.17: Finden Sie alle Vektoren, die zu $u = \left(\begin{smallmatrix} 1 \\ 3 \\ 2 \\ -1 \\ 4 \end{smallmatrix}\right) \in \mathbb{R}^5$ orthogonal sind.

Wir sehen also, dass eine Normalendarstellung einer Ebene nur in $\mathbb{R}^3$ funktioniert: in $\mathbb{R}^n$ bekommen wir so nicht eine Ebene, sondern eine Hyperebene, also einen Unterraum der Dimension $n-1$. In $\mathbb{R}^2$ zum Beispiel eine Gerade. Und diese Normalendarstellung ist eigentlich nichts anderes als ein LGS aus einer einzigen Gleichung.
Wir können dieses Konzept noch verallgemeinern:

Definition 9.18: Sei U ein Unterraum von $\mathbb{R}^n$. Der **Orthogonalraum**, auch das **orthogonale Komplement**, von U ist die Menge aller Vektoren in $\mathbb{R}^n$, die zu allen Vektoren in U orthogonal sind.

$$U^\perp = \{v \in \mathbb{R}^n \mid \text{für alle } u \in U \text{ gilt } \langle u, v\rangle = 0\}$$

Proposition 9.19: *Jeder Orthogonalraum ist ein Unterraum von $\mathbb{R}^n$.*

BEWEIS. Übung. □

Beispiele 9.20: Aus dem vorherigen Beispiel:
- $\operatorname{Span}\left(\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\right)^\perp = \operatorname{Span}\left(\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)\right)$.
- $\operatorname{Span}\left(\left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right)\right)^\perp = \operatorname{Span}\left(\left(\begin{smallmatrix}1\\0\\-1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\\-1\end{smallmatrix}\right)\right)$.
- $\operatorname{Span}\left(\left(\begin{smallmatrix}1\\1\\1\\1\end{smallmatrix}\right)\right)^\perp = \operatorname{Span}\left(\left(\begin{smallmatrix}1\\0\\0\\-1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\\0\\-1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\0\\1\\-1\end{smallmatrix}\right)\right)$.

Was ist dann das orthogonale Komplement von ganz $\mathbb{R}^n$? „Skalarprodukt findet Null“ sagt uns:

Beispiel 9.21: Wir haben $(\mathbb{R}^n)^\perp = \{0\}$. Wir haben auch $\{0\}^\perp = \mathbb{R}^n$.

Was ist die Dimension eines Orthogonalraums?

Proposition 9.22: *Sei U ein Unterraum von $\mathbb{R}^n$. Dann gilt* $\dim U + \dim U^\perp = n$.

BEWEIS. (Sketch). Sei $\dim U = k$ und seien $u_1, \dots, u_k$ eine Basis von U. Dann ist $U^\perp$ die Menge von Vektoren, die zu jedem Basisvektor orthogonal sind. Das gibt uns k Gleichungen mit n Unbekannten. Da die Basis linear unabhängig ist, bekommen wir k führende Einsen, und somit $n - k$ Spalten ohne führende Eins. Somit hat die Lösungsmenge $U^\perp$ die Dimension $n - k$. □

Eigentlich kommt dieser Dimensionszusammenhang von

$$\mathbb{R}^n = U \oplus U^\perp$$

(Siehe direkte Summe in Abschnitt 2D.) Wir beweisen es hier nicht.

C. Orthonormalbasen

Im Kontext des Skalarprodukts gibt es eine Sorte von Basis, die uns vieles erleichtert.

Definition 9.23: Eine Menge oder Liste von Vektoren heißt **orthogonal** wenn die Vektoren paarweise orthogonal sind.
Eine Menge oder Liste von Vektoren heißt **orthonormal** wenn sie orthogonal ist und alle Vektoren Einheitsvektoren sind (also Norm 1 haben). Eine Basis, die orthonormal ist, heißt **Orthonormalbasis** (ONB).

Beispiele 9.24:
- Die Standardbasis in $\mathbb{R}^n$ ist eine ONB.
- $v_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ und $v_2 = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$ sind orthogonal, aber haben beide Norm $\sqrt{2}$. Also ist

$$u_1 = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}, u_2 = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\end{pmatrix}$$

  eine ONB von $\mathbb{R}^2$.
- $v_1 = \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix}1\\-1\\0\end{smallmatrix}\right)$, $v_3 = \left(\begin{smallmatrix}1\\1\\-2\end{smallmatrix}\right)$ sind eine orthogonale Basis von $\mathbb{R}^3$. Wir normieren sie zu einer ONB

$$u_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}, u_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}.$$

  Man kann auch die Reihenfolge der Einträge etwas mischen, bei allen gleichzeitig, und bekommt weitere ONB von $\mathbb{R}^3$.

◇ Nicht in Vorlesung. Man kann obige Idee erweitern auf $\mathbb{R}^n$:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ -4 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 4 \\ -8 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix},$$

$$\ldots, \quad v_k = \begin{pmatrix} 1 \\ 2^0 \\ 2^1 \\ 2^2 \\ \vdots \\ 2^{k-3} \\ -2^{k-2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad v_{n-1} = \begin{pmatrix} 1 \\ 2^0 \\ 2^1 \\ 2^2 \\ 2^3 \\ 2^4 \\ \vdots \\ 2^{n-4} \\ -2^{n-3} \\ 0 \end{pmatrix}, \quad v_n = \begin{pmatrix} 1 \\ 2^0 \\ 2^1 \\ 2^2 \\ 2^3 \\ 2^4 \\ \vdots \\ 2^{n-4} \\ 2^{n-3} \\ -2^{n-2} \end{pmatrix}$$

Diese Vektoren kann man jetzt noch normalisieren, um eine ONB zu erhalten. (Machen wir hier nicht.)

Übung 9.25: Üben Sie in Numbas: ONB oder nicht?

Man kann in den Beispielen leicht prüfen, dass sie paarweise orthogonal sind und jeweils Norm 1 haben. Aber sind es dann auch automatisch Basen? Schönerweise ja.

Proposition 9.26: (Orthogonale Vektoren sind linear unabhängig.)
Wenn $v_1, \ldots, v_k$ orthogonal sind und alle $v_i \neq 0$, dann sind die Vektoren linear unabhängig.

Beweis. Wir betrachten

$$\lambda_1 v_1 + \cdots \lambda_k v_k = 0$$

und wenden auf beiden Seiten $\langle v_1, - \rangle$ an:

$$\begin{aligned} \Rightarrow \qquad \lambda_1 \langle v_1, v_1 \rangle + \lambda_2 \langle v_1, v_2 \rangle + \cdots + \lambda_n \langle v_1, v_k \rangle &= 0 \\ \Rightarrow \qquad \lambda_1 \langle v_1, v_1 \rangle + 0 &= 0 \\ \Rightarrow \qquad \lambda_1 &= 0 \end{aligned}$$

weil $\langle v_1, v_1 \rangle = \|v_1\|^2 \neq 0$.
Genauso gibt Anwendung von $\langle v_i, - \rangle$ für alle $i = 2, \ldots, k$, dass alle $\lambda_i = 0$. Also haben wir linear unabhängige Vektoren. □

Wenn wir also genau n orthogonale (oder orthonormale) Vektoren haben, dann haben wir eine Basis von $\mathbb{R}^n$ (mit zwei für eins bei Basen).
Das schöne an einer ONB ist, dass wir die Koordinaten bezüglich der ONB super leicht berechnen können.

Proposition 9.27: (Koordinaten bezüglich ONB)
Wenn $u_1, \ldots, u_n$ eine Orthonormalbasis bilden, dann gilt für jedes $v \in \mathbb{R}^n$:

$$v = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2 + \cdots + \langle v, u_n\rangle u_n.$$

BEWEIS. Da $u_1, \ldots, u_n$ eine Basis bilden, gibt es Skalare λ_i mit $v = \lambda_1 u_1 + \cdots + \lambda_n u_n$. Dann

$$\begin{aligned}\langle v, u_i\rangle &= \langle \lambda_1 u_1 + \cdots + \lambda_n u_n, u_i\rangle \\ &= \lambda_1\langle u_1, u_i\rangle + \cdots + \lambda_n\langle u_n, u_i\rangle \\ &= \lambda_1 0 + \cdots + 0 + \lambda_i\langle u_i, u_i\rangle + 0 + \cdots + \lambda_n 0 \\ &= \lambda_i\end{aligned}$$

da $\langle u_j, u_i\rangle = 0$ für $j \neq i$ und $\langle u_i, u_i\rangle = 1$. □

Dies ist ⚠ *sehr* anders zum Finden von Koordinaten bei einer gewöhlichen Basis! Dort müssen wir ein LGS lösen und alle Koordinaten zusammen finden, während wir hier jeden einzelnen Koeffizienten separat finden können. Das ist super hilfreich.

Beispiel 9.28: Wir hatten oben die ONB

$$u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

von $\mathbb{R}^2$. Um die Koordinaten von $v = \left(\begin{smallmatrix} -4 \\ 13 \end{smallmatrix}\right)$ bezüglich dieser Basis zu finden, berechnen wir:

- $\lambda_1 = \langle v, u_1\rangle = \frac{1}{\sqrt{2}}((-4)\cdot 1 + 13\cdot 1) = \frac{9}{\sqrt{2}}$
- $\lambda_2 = \langle v, u_2\rangle = \frac{1}{\sqrt{2}}(-4)\cdot 1 + 13\cdot(-1)) = \frac{-17}{\sqrt{2}}$

Und als Probe: es ist in der Tat

$$\frac{9}{\sqrt{2}}u_1 + \frac{-17}{\sqrt{2}}u_2 = \frac{9}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{-17}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 9-17 \\ 9+17 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -8 \\ 26 \end{pmatrix} = \begin{pmatrix} -4 \\ 13 \end{pmatrix}$$

Übung 9.29: Üben Sie in Numbas: Koordinaten bzgl. ONB bestimmen.

Dies hat eine schöne Anwendung, die uns später bei Abständen sehr helfen wird.

Beispiel 9.30: Wir betrachten die Gerade Span$\left(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right)$, und den Vektor $v = \left(\begin{smallmatrix} -4 \\ 13 \end{smallmatrix}\right)$. Was ist der Abstand von v zur Gerade (also der kleinste Abstand von v zu einem Punkt auf der Geraden)?
Wir wissen schon, dass es der senkrechte Abstand ist. Aber wie finden wir den Punkt auf der Gerade, der uns diesen senkrechten Abstand gibt?
Wir haben

$$\begin{pmatrix} -4 \\ 13 \end{pmatrix} = \frac{9}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{-17}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

wobei $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \perp \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$. Der Teil $u = \frac{9}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ ist also der Punkt auf der Geraden, der sozusagen „senkrecht unter v" liegt. Also ist der Abstand von v zur Geraden die Norm von $v - u$, der Abstand von v zum Punkt u. $\frac{-17}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$, der orthogonal zur Geraden ist.

Wir projezieren also v auf U und berechnen dann den Abstand von v zu diesem Punkt.

Definition 9.31: Sei U ein Unterraum von $\mathbb{R}^n$ und $u_1, \ldots, u_k$ eine Orthonormalbasis von U. Die **orthogonale Projektion** ist die lineare Abbildung $\mathrm{pr}_U : \mathbb{R}^n \longrightarrow U$ mit

$$\mathrm{pr}_U(v) = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2 + \cdots + \langle v, u_k\rangle u_k.$$

Wir sehen, das schöne hier ist, dass wir dazu nur die Basis von U brauchen. Wenn wir dies mit gewöhnlichen Basen machen würden, müssten wir erst mal zu einer Basis von ganz $\mathbb{R}^n$ erweitern, dann das LGS lösen, um die ganze Basisdarstellung von v zu bekommen, und dann nur den Teil nehmen, der in U liegt.

Beispiel 9.32: Sei $U = \operatorname{Span}\left(\widetilde{u_1} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \widetilde{u_2} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ -4 \end{pmatrix}\right)$ in $\mathbb{R}^4$. Sei $v = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$.
Um die orthogonale Projektion $\operatorname{pr}_U(v)$ von v auf U zu berechnen, brauchen wir zuerst eine ONB von U. Die Vektoren $\widetilde{u_1}$ und $\widetilde{u_2}$ bilden eine Basis von U und sind orthogonal (weil $\langle \widetilde{u_1}, \widetilde{u_2} \rangle = 1 - 1 + 0 + 0 = 0$), aber sie haben nicht Norm 1. Wir normieren sie also zu

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{22}} \begin{pmatrix} 1 \\ 1 \\ 2 \\ -4 \end{pmatrix}.$$

Dann ist

$$\begin{aligned} \operatorname{pr}_U(v) &= \langle u_1, v \rangle u_1 + \langle u_2, v \rangle u_2 \\ &= \frac{1}{\sqrt{2}}(1 - 2 + 0 + 0) \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{22}}(1 + 2 + 6 - 16) \frac{1}{\sqrt{22}} \begin{pmatrix} 1 \\ 1 \\ 2 \\ -4 \end{pmatrix} \\ &= -\frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} - \frac{7}{22} \begin{pmatrix} 1 \\ 1 \\ 2 \\ -4 \end{pmatrix} = \frac{1}{22} \begin{pmatrix} -18 \\ 4 \\ -14 \\ 28 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} -9 \\ 2 \\ -7 \\ 14 \end{pmatrix} \end{aligned}$$

D. Abstände

Wir können jetzt Abstände von einem Punkt zu einem Unterraum berechnen. Wir haben oben schon erwähnt, schreiben es aber noch mal auf:

Definition 9.33: Der **Abstand** $d(v, u)$ von Punkt v zu Punkt u ist $\|v - u\|$.

Proposition 9.34: *Der Abstand $d(v, U)$ von Punkt v zu einem Unterraum U von $\mathbb{R}^n$ ist $d(v, U) = \|v - \operatorname{pr}_U(v)\|$.*

BEWEIS. Wir müssen zeigen, dass $v - \operatorname{pr}_U(v)$ orthogonal zu jedem Vektor in U ist. Dies ist dann der senkrechte Abstand von v zu einem Punkt in U.
Sei $u_1, \ldots, u_k$ eine ONB von U. Also ist $\operatorname{pr}_U(v) = \langle u_1, v \rangle u_1 + \cdots + \langle u_k, v \rangle u_k$.
Wir haben für $u \in U$: $u = \lambda_1 u_1 + \cdots + \lambda_k u_k$, weil $u_1, \ldots, u_k$ eine Basis von U ist. Wegen Linearität im ersten Eintrag des Skalarprodukts reicht es also zu zeigen, dass $v - \operatorname{pr}_U(v)$ orthogonal zu jedem Basisvektor u_i ist. Wir haben

$$\begin{aligned} \langle u_i, v - \operatorname{pr}_U(v) \rangle &= \langle u_i, v \rangle - \langle u_i, \operatorname{pr}_U(v) \rangle \\ &= \langle u_i, v \rangle - \langle u_i, \langle u_1, v \rangle u_1 + \cdots + \langle u_k, v \rangle u_k \rangle \\ &= \langle u_i, v \rangle - \langle u_1, v \rangle \langle u_i, u_1 \rangle - \cdots - \langle u_i, v \rangle \langle u_i, u_i \rangle - \cdots - \langle u_k, v \rangle \langle u_i, u_k \rangle \\ &= \langle u_i, v \rangle - 0 - \cdots - \langle u_i, v \rangle \cdot 1 - \cdots - 0 \\ &= 0 \end{aligned}$$

Hier haben wir benutzt, dass $u_1, \ldots, u_k$ eine ONB sind, also ist $\langle u_i, u_j \rangle = 0$ für $i \neq j$, und $\langle u_i, u_i \rangle = 1$.
Somit ist $v - \operatorname{pr}_U(v)$ orthogonal zu jedem Element von U, und damit der kürzeste Abstand von v zu einem Punkt auf U. $\square$

Beispiel 9.35: Mit U und v wie in Beispiel 9.32 haben wir

$$\begin{aligned} d(v,U) = \|v - \mathrm{pr}_U(v)\| &= \left\| \begin{pmatrix} 1\\2\\3\\4 \end{pmatrix} - \frac{1}{11}\begin{pmatrix} -9\\2\\-7\\14 \end{pmatrix} \right\| = \left\| \frac{1}{11}\begin{pmatrix} 20\\20\\40\\30 \end{pmatrix} \right\| \\ &= \frac{10}{11}\sqrt{2^2+2^2+4^2+3^2} = \frac{10}{11}\sqrt{33} = \frac{10\sqrt{3}}{\sqrt{11}} \end{aligned}$$

Wie können wir jetzt den Abstand von einem Punkt zu einer Ebene (oder anderem Unterraum) berechnen, die nicht durch den Ursprung geht?

Beispiel 9.36: Sei $E = \{b + su_1 + tu_2 \mid s,t \in \mathbb{R}\}$ eine Ebene, die nicht durch den Ursprung geht, und sein $v \in \mathbb{R}^n$ ein weiterer Punkt. Dann ist $d(v,E) = d(v-b, E-b)$, und $E-b$ ist nun wieder eine Ebene durch den Ursprung. Wir verschieben einfach alles um b wieder auf den Nullpunkt.

E. Gram-Schmidt Orthogonalisierung

Wir haben gesehen, dass wir für Abstände und orthogonale Projektion eine Orthonormalbasis brauchen. In den bisherigen Beispielen hatten wir schon eine orthogonale Basis, die wir leicht normieren konnten. Aber was ist, wenn wir eine Basis gegeben haben, die nicht orthogonal ist?
Mit dem folgenden Algorithmus können wir sie in eine orthogonale Basis des selben Unterraums verwandeln. Der Algorithmus ist am besten an einem Beispiel veständlich, deswegen fangen wir damit an.

Beispiel 9.37: (Gram-Schmidt Orthogonalisierung)
Seien $v_1 = \begin{pmatrix} 1\\1\\1 \end{pmatrix}, v_2 = \begin{pmatrix} 1\\1\\0 \end{pmatrix}, v_3 = \begin{pmatrix} 1\\0\\0 \end{pmatrix}$, welche eine Basis von $\mathbb{R}^3$ bilden. Wir werden diese Basis Schritt für Schritt in eine Orthonormalbasis von $\mathbb{R}^3$ umwandeln.

- Wir fangen an mit $\widetilde{u_1} = v_1 = \begin{pmatrix} 1\\1\\1 \end{pmatrix}$.
- Nun benutzen wir die Methode, Koeffizienten bzgl. einer ONB zu bestimmen (Prop. 9.27), um den Anteil von v_2 abzuziehen, der in Richtung $\widetilde{u_1}$ geht:
$$\widetilde{u_2} = v_2 - \frac{\langle v_2, \widetilde{u_1}\rangle}{\|\widetilde{u_1}\|^2}\widetilde{u_1} = \begin{pmatrix} 1\\1\\0 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1\\1\\1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3}\\ \frac{1}{3}\\ -\frac{2}{3} \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1\\1\\-2 \end{pmatrix}$$
(Wir brauchen hier $\frac{1}{\|\widetilde{u_1}\|^2}$ weil wir $\widetilde{u_1}$ noch nicht normiert haben.)
- Für den dritten Vektor ziehen wir die Komponenten in Richtung beider vorherigen Vektoren ab:
$$\widetilde{u_3} = v_3 - \frac{\langle v_3, \widetilde{u_1}\rangle}{\|\widetilde{u_1}\|^2}\widetilde{u_1} - \frac{\langle v_3, \widetilde{u_2}\rangle}{\|\widetilde{u_2}\|^2}\widetilde{u_2} = \begin{pmatrix} 1\\0\\0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1\\1\\1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \frac{1}{3}\\ \frac{1}{3}\\ -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\\ -\frac{1}{2}\\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1\\-1\\0 \end{pmatrix}$$

Nun können wir prüfen
- $\langle \widetilde{u_1}, \widetilde{u_2}\rangle = 0$
- $\langle \widetilde{u_1}, \widetilde{u_3}\rangle = 0$
- $\langle \widetilde{u_2}, \widetilde{u_3}\rangle = 0$

(Vergleichen Sie mit dem Beispiel in $\mathbb{R}^3$ aus Beispiel 9.24.)

Nun können wir alle Vektoren normieren und bekommen

$$\begin{aligned} u_1 &= \frac{1}{\|\widetilde{u_1}\|}\widetilde{u_1} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1\\1\\1 \end{pmatrix}, \\ u_2 &= \frac{1}{\|\widetilde{u_2}\|}\widetilde{u_2} = \frac{1}{-\frac{\sqrt{6}}{3}}\begin{pmatrix} \frac{1}{3}\\ \frac{1}{3}\\ -\frac{2}{3} \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1\\1\\-2 \end{pmatrix}, \\ u_3 &= \frac{1}{\|\widetilde{u_3}\|}\widetilde{u_3} = \frac{1}{\frac{1}{\sqrt{2}}}\begin{pmatrix} \frac{1}{2}\\ -\frac{1}{2}\\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1\\-1\\0 \end{pmatrix} \end{aligned}$$

eine Orthonormalbasis.

Wir bemerken etwas sehr besonderes über diese neue Basis:

$\mathrm{Span}(u_1) = \mathrm{Span}(v_1)$, $\quad \mathrm{Span}(u_1, u_2) = \mathrm{Span}(v_1, v_2)$, $\quad \mathrm{Span}(u_1, u_2, u_3) = \mathrm{Span}(v_1, v_2, v_3)$.

Dies ist sehr schön, weil wir zum Beispiel eine besondere Richtung, etwa eine Eigenvektorrichtung, beibehalten können, und wir können auch eine Basis eines Unterraums orthogonalisieren und bekommen wieder eine Basis des Unterraums.

Hier ist also der allgemeine Algorithmus.

Theorem 9.38: (Gram-Schmidt Orthogonalisierung)
Aus einer Basis $v_1, v_2, \ldots, v_n$ von $\mathbb{R}^n$ kann mit folgenden Schritten eine Orthonormalbasis $u_1, \cdots, u_n$ konstruiert werden, mit der Eigenschaft $\mathrm{Span}(v_1, \cdots, v_k) = \mathrm{Span}(u_1, \cdots, u_k)$ für jedes $k = 1, \cdots, n$:

- $\widetilde{u_1} = v_1$
- $\widetilde{u_2} = v_2 - \frac{\langle v_2, \widetilde{u_1}\rangle}{\|\widetilde{u_1}\|^2}\widetilde{u_1}$ *(subtrahiere den Anteil von v_2 der in Richtung $\widetilde{u_1}$ geht)*
- $\widetilde{u_3} = v_3 - \frac{\langle v_3, \widetilde{u_1}\rangle}{\|\widetilde{u_1}\|^2}\widetilde{u_1} - \frac{\langle v_3, \widetilde{u_2}\rangle}{\|\widetilde{u_2}\|^2}\widetilde{u_2}$ *(subtrahiere Anteile in Richtungen $\widetilde{u_1}$ und $\widetilde{u_2}$)*
- $\vdots$
- $\widetilde{u_n} = v_n - \sum_{i=1}^{n-1} \frac{\langle v_n, \widetilde{u_i}\rangle}{\|\widetilde{u_i}\|^2}\widetilde{u_i}$ *(subrahiere Anteile in alle vorherigen Ricthungen)*
- *Für jedes k, $u_k = \frac{1}{\|\widetilde{u_k}\|}\widetilde{u_k}$.* *(normiere alle Vektoren)*

BEWEIS. Für Interessierte
Wir prüfen zuerst, dass $\widetilde{u_k}, \widetilde{u_l}$ orthogonal sind, für $k \neq l$. Wir beginnen mit 1 und 2:

$$\begin{aligned}
\langle \widetilde{u_1}, \widetilde{u_2}\rangle &= \langle \widetilde{u_1}, (v_2 - \frac{\langle v_2, \widetilde{u_1}\rangle}{\|\widetilde{u_1}\|^2}\widetilde{u_1})\rangle \\
&= \langle \widetilde{u_1}, v_2\rangle - \frac{\langle v_2, \widetilde{u_1}\rangle}{\|\widetilde{u_1}\|^2}\langle \widetilde{u_1}, \widetilde{u_1}\rangle && \text{(linear im zweiten Eintrag)} \\
&= \langle \widetilde{u_1}, v_2\rangle - \langle v_2, \widetilde{u_1}\rangle && \text{(Def der Norm)} \\
&= 0 && \text{(symmetrisch)}
\end{aligned}$$

Nun nehmen wir an, dass für ein bestimmtes $k \leqslant n$ die Vektoren $\widetilde{u_1}, \ldots, \widetilde{u_k}$ alle orthogonal zueinander sind, also $\langle \tilde{u}_i, \tilde{u}_j\rangle = 0$ für alle $i \neq j \leqslant k$. Dann betrachten wir $\widetilde{u_{k+1}}$:

$$\begin{aligned}
\langle \tilde{u}_l, \widetilde{u_{k+1}}\rangle &= \left\langle \tilde{u}_l, \left(v_{k+1} - \sum_{i=1}^{k} \frac{\langle v_{k+1}, \tilde{u}_i\rangle}{\|\tilde{u}_i\|^2}\tilde{u}_i\right)\right\rangle \\
&= \langle \tilde{u}_l, v_{k+1}\rangle - \sum_{i=1}^{k} \frac{\langle v_{k+1}, \tilde{u}_i\rangle}{\|\tilde{u}_i\|^2}\langle \tilde{u}_l, \tilde{u}_i\rangle && \text{(linear im zweiten Eintrag)} \\
&= \langle \tilde{u}_l, v_{k+1}\rangle - \frac{\langle v_{k+1}, \tilde{u}_l\rangle}{\|\tilde{u}_l\|^2}\langle \tilde{u}_l, \tilde{u}_l\rangle && \text{(nach Annahme sind die anderen 0)} \\
&= 0
\end{aligned}$$

Also folgt mit vollständiger Induktion, dass alle $\tilde{u}_i$ paarweise orthogonal sind. Daher bilden die normierten Vektoren u_i eine orthonormale Liste.
Aus der Konstruktion ist klar, dass $\mathrm{Span}(v_1, \ldots, v_k) = \mathrm{Span}(u_1, \ldots, u_k)$ für jedes $k = 1, \ldots, n$. $\square$

Übung 9.39: Sei $U = \mathrm{Span}\left(v_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, v_2 = \begin{pmatrix}2\\0\\1\\1\end{pmatrix}, v_3 = \begin{pmatrix}3\\1\\-1\\1\end{pmatrix}\right)$ ein Unterraum von $\mathbb{R}^4$.
Bestimmen Sie mit Gram-Schmidt eine ONB für U.

F. Abstände und Winkel: Study guide

Konzeptübersicht.

- ◇ Standardskalarprodukt, Norm, und Verbindung zwischen beiden
- ◇ Eigenschaften des Skalarprodukts: bilinear, symmetrisch, positiv definit.
- ◇ Eigenschaften der Norm, die wir von einer „Länge“ erwarten
- ◇ Einheitsvektoren, Vektoren normieren
- ◇ Skalarprodukt findet 0
- ◇ Cauchy-Schwarz Ungleichung
- ◇ Dreiecksungleichung der Norm
- ◇ Orthogonale Vektoren, Winkel zwischen zwei Vektoren
- ◇ Hyperebene orthogonal zu einem Vektor
- ◇ Orthogonalraum, Zusammenspiel der Dimensionen
- ◇ Orthonormalbasis (ONB)
- ◇ Liste von orthogonalen Vektoren ist linear unabhängig
- ◇ Spezielle Form der Koordinaten bzgl. ONB
- ◇ Orthogonale Projektion
- ◇ Abstand von Punkt zu Punkt, von Punkt zu Unterraum
- ◇ Gram-Schmidt Orthogonalisierung

Skills.

- ◇ Skalarprodukt von zwei Vektoren bestimmen
- ◇ Norm eines Vektors bestimmen
- ◇ Winkel zwischen zwei Vektoren bestimmen
- ◇ Eigenschaften des Skalarprodukts für Umformungen benutzen
- ◇ Vektor zu Einheitsvektor normieren
- ◇ Bestimmen, ob zwei Vektoren orthogonal sind
- ◇ Bestimmen, ob eine Liste von Vektoren orthogonal oder orthonormal ist
- ◇ Koordinaten bzgl. einer ONB finden
- ◇ Orthogonale Projektion eines Vektors berechnen
- ◇ Abstand von einem Vektor zu einem Unterraum berechnen
- ◇ Mit Gram-Schmidt eine ONB aus einer Basis berechnen