1. PIRATENMANUSKRIPT

Titel des Manuskripts

Das Kreuzprodukt im $\mathbb{R}^n/\mathbb{C}^n$ und wie man damit die Lineare Algebra entert

Verfasser

anonymer Pirat

angestrebter Grad

Pirate Captain (PCpt.)

Einsame Insel, im Monat März 2018

Kennzahl lt. Piratenkodex:	424242
Richtung lt. Piratenkompass:	Nord-Nordwest in Richtung Lineare Algebra
Betreuer:	ao. PCpt. Nerviger Papagei

Rechtliche Information

Ich, Forty-Moo (Pseudonym), bin alleiniger Urheber und Autor dieser Arbeit. Meine Website ist erreichbar unter `forty-moo.github.io`.

Es ist (ohne meine ausdrückliche Einwilligung) untersagt …

- diese Arbeit auf jegliche Art und Weise zu modifizieren oder zu verändern
- die Arbeit an anderer Stelle hochzuladen oder anders zur Verfügung zu stellen
- die Arbeit zu verkaufen oder auf eine andere Art und Weise kommerziell zu vermarkten
- neue Ideen, Beweise bzw. Entdeckungen in dieser Arbeit als die eigenen auszugeben,

sowie jegliche andere Tätigkeiten, die mein Urheberrecht verletzen könnten. Diese Punkte gelten auch für Teile der Arbeit.

Ich garantiere weder die Vollständigkeit der Arbeit, noch ihre Fehlerfreiheit. Ich hafte nicht für Schäden, die durch diese Arbeit (z.B. durch die Anwendung der Methoden) entstehen. Teile der Beispiele (und ein paar nicht-mathematische Bilder) in dieser Arbeit stammen nicht von mir, sondern aus verschiedenen Quellen. Die Lösungen sind jedoch von mir selbst erstellt und durchdacht worden.

Der Leser/Verwender/Endbenutzer dieser Arbeit, stimmt Obigem – zusätzlich zu den allgemein geltenden Gesetzen betreffend Urheberrecht etc. – zu.

Sollte hiergegen verstoßen werden, sehe ich mich – je nach Fall und Situation – auch gezwungen, rechtliche Schritte einzugehen (z.B. um eine gewisse Tätigkeit abzustellen).

Inhaltsverzeichnis

1 Vorwort

Ahoi, ihr Landratten! Ich bin ein unbekannter Pirat, dessen Name keine Rolle spielt. Ihr braucht nur so viel zu wissen: Ich habe viele Jahre auf der Linearen Algebra gedient, dem Schiff der universitär-mathematischen Marine. Doch irgendwann hat mich dieses Pack aus nicht weiter wichtigen Gründen auf einer einsamen Insel ausgesetzt und dem Tod überlassen. Doch ich lebe noch und habe mich entschlossen, von nun an ein Pirat zu sein!

Mit mir zusammen haben sie auch einen halb verhungerten, auf einem Auge blinden Papagei auf der Insel ausgesetzt, den sie halb tot irgendwo in der hintersten Ecke ihres Lagerraumes gefunden haben. Doch das Vieh ist zäher, als sie gedacht haben und leistet mir von nun an Gesellschaft.

Und da sitz ich nun eines Abends auf meiner Insel unter einer Palme und überlege schon, das Tier über meinem Feuer zu grillen, da fängt das nervige Vieh plötzlich mit mir zu reden an! Und ich kann meinen Ohren kaum trauen, was es mir da zugeflüstert hat: Dinge über die Lineare Algebra, die ich noch nicht wusste, über das Kreuzprodukt, über Polynome und wie man damit Eigenwerte und Eigenvektoren von kleinen Matrizen schneller berechnen kann. Und natürlich, wie man das alles auf nxn Matrizen verallgemeinert, wenn gewünscht. Probleme, für die ich vorher eine lange Zeit benötigte, konnte ich nun in Minuten lösen.

Jetzt habe ich etwas Konkretes gegen die Lineare Algebra in der Hand. Würde ich nun noch von dieser Insel herunterkommen, könnte ich dieses Wissen als Druckmittel verwenden und mit der richtigen Strategie die Lineare Algebra vielleicht sogar entern. Sollte ich es nicht schaffen, habe ich alles Wichtige mit meinem eigenen Blut auf einige getrocknete Fischhäute geschrieben. Das Manuskript habe ich in einem ausgehöhlten Felsen in der Mitte der Insel versteckt. Vielleicht wird es irgendwann einmal jemandem nützen, sollte er der Linearen Algebra gegenüberstehen.

Ich werde jetzt meine Flucht planen, um von dieser verdammten Insel herunterzukommen. Ich habe da mal von einem Piraten gehört, der sich an Schildkröten gebunden hat. Vielleicht sollte ich das probieren...

Abbildung 1: Anonymer Pirat mit Papagei (Portrait)

2 Vorbereitung

Bevor sie mich auf dieser einsamen Insel ausgesetzt haben, habe ich noch alles mögliche von der Linearen Algebra mitgehen lassen. Ich habe meine Taschen vollgestopft, bis nichts mehr hineingegangen ist. Das nachfolgende Zeug ist möglicherweise wichtig, um wirklich alles im Manuskript zu verstehen. Andererseits muss man auch nicht alles verstehen, wenn man das Wissen ausschließlich anwenden will. Wir sind hier schließlich nicht mehr an Bord der Linearen Algebra...

2.1 Das Standardskalarprodukt im $\mathbb{R}^n$

Zuerst müssen wir uns ums Setting Gedanken machen. Wir werden im weiteren Verlauf des Textes den $\mathbb{R}$-Vektorraum $V = \mathbb{R}^n$, zusammen mit dem Standardskalarprodukt

$$
\begin{aligned}
\cdot : \qquad \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \quad \mathbb{R} \\
(\ \vec{u}, \vec{v}\ ) &\longmapsto \ \vec{u} \cdot \vec{v} \ := \ \sum_{i=0}^{n} u_i \cdot v_i
\end{aligned} \tag{2.1}
$$

betrachten. Hierbei sind u_i und v_i die Komponenten der Vektoren $\vec{u}$ bzw. $\vec{v}$. Man spricht auch von einem *Euklidischen Vektorraum*. Das ist ein $\mathbb{R}$-Vektorraum zusammen mit einem Skalarprodukt. Für das obige Standardskalarprodukt gelten ähnliche Rechenregeln, wie für das ganz normale Produkt in z.B. den ganzen oder reellen Zahlen. Deshalb habe ich es auch freundlicherweise mit dem Zeichen $\cdot$ eingeführt/notiert. Vom ganz normalen Produkt, z.B. in den reellen Zahlen, ist es leicht zu unterscheiden, da ich Vektoren immer mit einem Pfeil kennzeichnen werde - und das Standardskalarprodukt benötigt immer zwei Vektoren.
Allgemein ist ein Skalarprodukt übrigens eine symmetrische, positiv definite Bilinearform. Das bedeutet

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
2. $\vec{u} \cdot \vec{u} \geq 0 \quad \text{und} \quad \vec{u} \cdot \vec{u} = 0 \quad \Leftrightarrow \quad \vec{u} = \vec{0}$
3. $(\vec{u} + \lambda \vec{w}) \cdot \vec{v} = \vec{u} \cdot \vec{v} + \lambda\ \vec{w} \cdot \vec{v} \quad \text{und} \quad \vec{u} \cdot (\vec{v} + \lambda \vec{w}) = \vec{u} \cdot \vec{v} + \lambda\ \vec{u} \cdot \vec{w}$

für beliebige $\vec{u}, \vec{v}, \vec{w} \in V$ und Skalare $\lambda \in \mathbb{R}$.
Das ist aber nicht weiter wichtig. Vorerst reicht es aus, zu wissen, dass man mit dem obigen Skalarprodukt ähnlich rechnen kann, wie mit dem Produkt z.B. in den reellen Zahlen.

2.2 Das orthogonale Komplement $U^\perp$ eines Unterraumes U

Wir werden nun definieren, was es bedeutet, dass zwei Vektoren orthogonal, das heißt im 90 Grad Winkel aufeinander, stehen.

Definition 2.2.1. Zwei Vektoren $\vec{u}$ und $\vec{v}$ in V stehen orthogonal bzw. normal aufeinander, falls ihr Skalarprodukt $= 0$ wird, das heißt

$$\vec{u} \cdot \vec{v} = 0 \tag{2.2}$$

Haben wir nun einen Unterraum $U \leqslant V$ gegeben, das ist eine Teilmenge des Vektorraumes, die selbst ein Vektorraum ist, dann können wir alle Vektoren, die auf die Vektoren in U orthogonal stehen, in eine Menge packen. Diese Menge nennen wir orthogonales Komplement.

Definition 2.2.2. Das orthogonale Komplement $U^\perp$ sind all die Vektoren, die auf den Unterraum $U \leqslant V$ orthogonal stehen, das heißt

$$U^\perp = \left\{ \vec{v} \in V \,\middle|\, \vec{v} \cdot \vec{u} = 0 \;\; \forall\, \vec{u} \in U \right\} \tag{2.3}$$

Die Bedingung $\vec{v} \cdot \vec{u} = 0$ für jedes $\vec{u}$ im Vektorraum U zu überprüfen, kann schnell aufwändig werden. Glücklicherweise reicht es aus, die Bedingung nur für alle Vektoren $\vec{e}$ in einem Erzeugendensystem E oder einer Basis B zu überprüfen, denn

$$U^\perp = \left\{ \vec{v} \in V \,\middle|\, \vec{v} \cdot \vec{e} = 0 \;\; \forall\, \vec{e} \in E \right\} \quad \text{wobei} \quad U = \mathcal{L}\{E\}.$$

Warum nennt man es eigentlich ein Komplement? Das ist schnell erklärt. Summieren wir zu den Vektoren von U die Vektoren des orthogonalen Komplements $U^\perp$ hinzu, erhalten wir den (gesamten) Vektorraum V. Es gilt sogar: Jeder Vektor im Vektorraum V kann eindeutig als Summe eines Vektors aus dem Unterraum U und seines orthogonalen Komplements $U^\perp$ dargestellt werden. Wegen dieser Eindeutigkeit sagt und schreibt man auch

Satz 2.2.3. *Der endlichdimensionale, euklidische Vektorraum V ist die direkte Summe seines Unterraumes U und seines orthogonalen Komplements $U^\perp$ oder in Symbolen*

$$V = U \oplus U^\perp. \tag{2.4}$$

Insbesondere gilt somit auch

$$dim(V) = dim(U) + dim(U^\perp) \tag{2.5}$$

Beispiel 1. Dieses Beispiel soll dazu dienen, sich das Ganze mit der Orthogonalität und dem Komplement besser vorstellen zu können. Wir betrachten als Vektorraum $V = \mathbb{R}^2$ und in diesem Vektorraum einen Vektor $\vec{u} \neq 0$. Als Unterraum U nehmen wir die lineare Hülle dieses Vektors $\vec{u}$, das heißt $U = \mathcal{L}\{\vec{u}\}$. Zeichnet man das, beschreibt U eine Gerade durch den Nullpunkt in Richtung $\vec{u}$.
Unsere Aufgabe besteht nun darin, das orthogonale Komplement $U^\perp$ zu berechnen. Anders ausgedrückt, wir wollen all jene Vektoren $\vec{n} = (x, y)$ finden, die auf den gegebenen Vektor $\vec{u} = (u_1, u_2)$ orthogonal stehen. Orthogonalität bedeutet, dass das Skalarprodukt gleich Null

wird. Wir müssen also

$$\vec{u} \cdot \vec{n} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = u_1 x + u_2 y \stackrel{!}{=} 0 \tag{2.6}$$

für x und y lösen, um den orthogonalen Vektor $\vec{n}$ zu erhalten.
Diese Gleichung hat keine eindeutige Lösung, wir haben einen freien Parameter. O.B.d.A. sei nun $u_1 \neq 0$. Wir wählen y als frei, also $y = u_1 \cdot t$ für irgendein $t \in \mathbb{R}$. Jetzt können wir obige Gleichung nach x auflösen:

$$\begin{aligned} u_1 x + u_2 y &= 0 \\ u_1 x &= -u_2 y \\ x &= -\frac{u_2}{u_1} y = -\frac{u_2}{u_1} \cdot (u_1 \cdot t) \\ &= -u_2 \cdot t. \end{aligned}$$

Wir erhalten also als Lösung für $\vec{n}$

$$\vec{n} = t \cdot \begin{pmatrix} -u_2 \\ u_1 \end{pmatrix}, \quad \text{wobei} \quad t \in \mathbb{R}. \tag{2.7}$$

Wir sehen, dass ganz schön viele Vektoren auf $\vec{u}$ normal stehen, da t über $\mathbb{R}$ läuft – und zwar alle skalaren Vielfachen des Vektors $(-u_2, u_1)$. Packen wir alle diese Vektoren in eine Menge, erhalten wir das orthogonale Komplement, abkürzend können wir gleich $U^\perp = \mathcal{L}\{(-u_2, u_1)\}$ schreiben.

In obigen Beispiel – und auch allgemein im $\mathbb{R}^n$ mit Standardskalarprodukt – stimmen die mathematische Vorstellung von Orthogonalität mit unserer realen Vorstellung überein. Man kann also sagen: Die mathematische Definition der Orthogonalität über das Skalarprodukt ist sinnvoll.
Sehr schön sehen kann man das, wenn man die beiden Geraden U und $U^\perp$ aus obigen Beispiel in einen Graph einzeichnet. Hier ist das Bild:

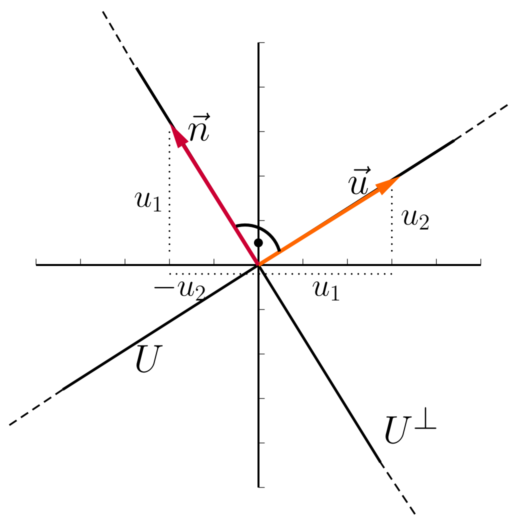


Abbildung 2: $U^\perp$ steht tatsächlich normal auf U...

Eine Anmerkung zum Beispiel: Wir wissen, dass wegen $dim(\mathbb{R}^2) = dim(U) + dim(U^\perp)$, die Dimension von $U^\perp$ Eins sein muss. Deshalb spannt bereits irgendein Vektor $\neq 0$ in $U^\perp$ diesen Unterraum auf. Somit hätten wir auch durch genaues Hinschauen ohne Rechnung sehen können, dass der Vektor $(-u_2, u_1)$ ungleich 0 ist und orthogonal auf $\vec{u}$ steht, und daher $U^\perp$ aufspannt. Fertig!

3 Das Kreuzprodukt im $\mathbb{R}^n$

Im letzten Kapitel haben wir das orthogonale Komplement $U^\perp$ eines Unterraumes U im Raum $\mathbb{R}^2$ berechnet. Nun stellt sich die Frage, wie man allgemein im $\mathbb{R}^n$ orthogonale Komplemente bzw. orthogonale Vektoren berechnen kann. Das verallgemeinerte Kreuzprodukt wird dabei ein nützliches Hilfsmittel sein – später dann auch beim Lösen von linearen Gleichungssystemen und Eigenwertproblemen.
Auch in diesem Abschnitt gilt wieder, dass man große Teile der Theorie überspringen kann. Wirklich relevant ist prinzipiell nur die Berechnungsformel des Kreuzprodukt am Ende des Abschnitts.

Das Setting ist hier wieder $V = \mathbb{R}^n$ zusammen mit dem Standardskalarprodukt. Das Kreuzprodukt macht nun folgendes: Es nimmt $n-1$ (linear unabhängige) Vektoren und gibt einen Vektor zurück, der normal auf alle diese $n-1$ Vektoren steht. Hier ist die mathematische Definition:

Definition 3.1. Gegeben seien $V = \mathbb{R}^n$ zusammen mit dem Standardskalarprodukt. Außerdem seien $n-1$ Vektoren $\vec{a}_2, \dots, \vec{a}_n$ im $\mathbb{R}^n$ gegeben. Wir unterscheiden zwei Fälle.

Fall 1 Falls $\vec{a}_2, \dots, \vec{a}_n$ linear unabhängig sind, so ist das Kreuzprodukt der Vektoren $\vec{a}_2, \dots, \vec{a}_n$ definiert als

$$\vec{a}_2 \times \cdots \times \vec{a}_n := det\left(\vec{n}, \vec{a}_2, \dots, \vec{a}_n\right) \cdot \vec{n} \tag{3.1}$$

wobei der Vektor $\vec{n}$ folgendermaßen zu wählen ist:

- $\vec{n} \cdot \vec{a}_i = 0$ für $i = 2, \dots, n$, d.h. $\vec{n}$ steht normal auf alle $\vec{a}_i$
- $||\ \vec{n}\ ||_2 = \sqrt{\vec{n} \cdot \vec{n}} = 1$, d.h der Vektor $\vec{n}$ ist von der Länge Eins.

Fall 2 Falls $\vec{a}_2, \dots, \vec{a}_n$ linear abhängig sind, dann definieren wir das Kreuzprodukt der Vektoren als den Nullvektor:

$$\vec{a}_2 \times \cdots \times \vec{a}_n := \vec{0} \tag{3.2}$$

Anmerkung. **ad Fall 1** Die Definition ist irgendwie nicht wirklich befriedigend! Sie gibt noch keine Berechnungsformel für die Komponenten des Kreuzprodukts. Vielmehr muss bereits der Normalenvektor $\vec{n}$ zum Zeitpunkt der Berechnung bekannt sein. Das Kreuzprodukt ist dann dieser Vektor $\vec{n}$ „gestreckt“ um den Faktor $det\left(\vec{n}, \vec{a}_2, \dots, \vec{a}_n\right) \in \mathbb{R}$.
Diesen Normalenvektor $\vec{n}$ haben wir aber im Normalfall nicht zur Verfügung – der Sinn des Kreuzproduktes soll es ja gerade sein, einen Vektor zu liefern, der orthogonal auf alle $\vec{a}_i$ steht.

ad Fall 2 Angenommen $U = \mathcal{L}\left\{\vec{a}_2, \dots, \vec{a}_n\right\}$. Im Fall 1 ist dann $dim(U) = n-1$ und $dim(U^\perp) = 1$. Wegen $\vec{n} \cdot \vec{a}_i = 0$ muss $\vec{n}$ in $U^\perp$ sein. Da $U^\perp$ eindimensional ist und wegen der Bedingung $||\ \vec{n}\ ||_2 = 1$ gibt es genau zwei solcher Vektoren: $\vec{n}$ und $-\vec{n}$. Es ist egal, welchen dieser beiden Vektoren wir in der Definition 3.1 verwenden, wir erhalten das gleiche Ergebnis. Die Definition ist somit wohldefiniert.
Im Fall 2 hingegen ist $dim(U) < n-1$ und $dim(U^\perp) > 1$. Es gibt also sicher mehr als zwei Vektoren, die auf alle $\vec{a}_i$ orthogonal stehen. Wir wählen also einfach den Nullvektor, da dieser auf alle Vektoren normal steht, egal wie sie aussehen.

Trotz der recht sperrigen Definition des Kreuzprodukts, können wir daraus sehr einfach eine Berechnungsformel für die einzelnen Komponenten des Produkts herleiten. Wir verwenden dazu folgendes

Lemma 3.2. *Für jedes $\vec{v} \in \mathbb{R}^n$ gilt*

$$\vec{v} \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n) = det\,(\vec{v}, \vec{a}_2, \ldots, \vec{a}_n) \tag{3.3}$$

Wir nennen diese Eigenschaft die ***charakteristische Eigenschaft*** *des Kreuzprodukts.*

Beweis. Seien $\vec{a}_2, \ldots, \vec{a}_n$ fixe gegebene Vektoren im $\mathbb{R}^n$.

Angenommen $\vec{a}_2, \ldots, \vec{a}_n$ sind linear abhängig. In obiger Definition des Kreuzprodukts sind wir dann im **Fall 2**. Das Kreuzprodukt ist also der Nullvektor. Somit stimmt Gleichung 3.3 für jedes $\vec{v} \in \mathbb{R}^n$, denn links steht das Skalarprodukt mit dem Nullvektor, was Null ergibt. Rechts steht ebenfalls Null, da die Determinante genau dann Null wird, wenn die Vektoren linear abhängig sind. Und wir haben ja gerade angenommen, dass $\vec{a}_2, \ldots, \vec{a}_n$ linear abhängig sind. Im Fall 2 stimmt die Gleichung also!

Betrachten wir nun den **Fall 1**, das heißt wir nehmen an, dass $\vec{a}_2, \ldots, \vec{a}_n$ linear unabhängig sind. Das Kreuzprodukt ist dann

$$\vec{a}_2 \times \cdots \times \vec{a}_n = det(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \cdot \vec{n} \tag{3.4}$$

wobei $\vec{n}$ der Normalenvektor ist wie in der Definition des Kreuzprodukts. Da die Summe von Unterraum und seinem orthogonalen Komplement direkt ist, bildet nun die Menge

$$B = \{\vec{a}_2, \ldots, \vec{a}_n, \vec{n}\} \tag{3.5}$$

eine Basis des $\mathbb{R}^n$. Falls man das nicht glaubt, kann man diese Vektoren auf lineare Unabhängigkeit prüfen...
Wir definieren nun zwei lineare Abbildungen φ_1 und φ_2 von $\mathbb{R}^n$ nach $\mathbb{R}$ folgendermaßen

$$\varphi_1\,(\vec{v}) := \vec{v} \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n) \qquad \varphi_2\,(\vec{v}) := det\,(\vec{v}, \vec{a}_2, \ldots, \vec{a}_n) \tag{3.6}$$

Wir zeigen nun, dass diese beide Abbildungen übereinstimmen, denn dann stimmt offensichtlich auch Gleichung 3.3. Da es sich um lineare Abbildungen handelt, reicht es die Gleichheit auf einer Basis zu zeigen. Nehmen wir einfach unsere vorher definierte Basis B her:
Es gilt $\varphi_1(\vec{a_i}) = 0 = \varphi_2(\vec{a_i})$ für $i = 2, \ldots, n$ (Links: Das Kreuzprodukt steht normal auf alle $\vec{a}_i$. Rechts: Die Determinante ist alternierend). Außerdem ist

$$\begin{aligned} \varphi_1(\vec{n}) &= \vec{n} \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n) = \vec{n} \cdot (det(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \cdot \vec{n}) \\ &= det(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \cdot \vec{n} \cdot \vec{n} = det(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \\ &= \varphi_2(\vec{n}), \end{aligned}$$

da der Normalenvektor normiert ist und somit $\vec{n} \cdot \vec{n} = ||\vec{n}||_2^2 = 1$. □

3.1 Berechnungsformel für das reelle Kreuzprodukt

Mit Hilfe der charakteristischen Eigenschaft des Kreuzprodukts können wir nun ganz einfach eine Berechnungsformel angeben. Das funktioniert so: Wir berechnen die einzelnen Komponenten des Kreuzproduktvektors, indem wir das Skalarprodukt des Vektors mit den Standardbasisvektoren ermitteln. Dann können wir die charakteristische Eigenschaft für $\vec{v} = \vec{e}_i$ verwenden und die Determinante berechnen.

Theorem 3.1.1. *Sei $V = \mathbb{R}^n$ mit dem Standardskalarprodukt. Die i-te Komponente des Kreuzproduktvektors lässt sich folgendermaßen berechnen*

$$\left(\vec{a}_2 \times \cdots \times \vec{a}_n\right)_i = (-1)^{i+1} \cdot det\, \underbrace{\left(\vec{a}_2, \ldots, \vec{a}_n\right)}_{i\text{-te Zeile streichen}} \tag{3.7}$$

Beweis. Wir verwenden die charakteristische Eigenschaft, $\vec{e}_i$ beschreibt hierbei den i-ten Einheitsvektor:

$$\begin{aligned}\left(\vec{a}_2 \times \cdots \times \vec{a}_n\right)_i &= \vec{e}_i \cdot \left(\vec{a}_2 \times \cdots \times \vec{a}_n\right) \\ &\underset{3.3}{=} det\left(\vec{e}_i, \vec{a}_2, \ldots, \vec{a}_n\right)\end{aligned}$$

Die Determinante können wir nun mit Hilfe des Laplace'schen Entwicklungssatzes auswerten, indem wir bezüglich der ersten Spalte entwickeln. Dort steht nur in der i-ten Zeile eine 1 sonst überall 0. Wir erhalten also obige Formel. □

Überlegen wir uns das einmal für konkrete Werte für n.

n=2 Sei $\vec{u} \neq 0$ ein Vektor im $\mathbb{R}^2$, das heißt $\vec{u} = (u_1, u_2)$. Das Kreuzprodukt diese Vektors nennen wir einfach mal $\times\vec{u}$. Dann gilt nach obiger Berechnungsformel

$$\begin{aligned}(\times\vec{u})_1 &= (-1)^2 \cdot det(u_2) = u_2 \\ (\times\vec{u})_2 &= (-1)^3 \cdot det(u_1) = -u_1\end{aligned}$$

Wir erhalten also

$$\times\vec{u} = \begin{pmatrix} u_2 \\ -u_1 \end{pmatrix}$$

Dieser Vektor steht normal auf $\vec{u}$, was man leicht überprüfen kann. Multiplizieren wir $\times\vec{u}$ mit -1, so erhalten wir den Normalvektor aus Beispiel 1. Somit haben wir dieses Beispiel schon auf 3 verschiedene Arten gelöst.

n=3 Nehmen wir zwei Vektoren $\vec{u} = (u_1, u_2, u_3)$ und $\vec{v} = (v_1, v_2, v_3)$ im $\mathbb{R}^3$ her. Wir wollen das Kreuzprodukt $\vec{u} \times \vec{v}$ berechnen. Wieder nach obiger Berechnungsformel

$$\begin{aligned}(\vec{u} \times \vec{v})_1 &= (-1)^2 \det\begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} = u_2 \cdot v_3 - v_2 \cdot u_3 \\ (\vec{u} \times \vec{v})_2 &= (-1)^3 \det\begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix} = u_3 \cdot v_1 - v_3 \cdot u_1 \\ (\vec{u} \times \vec{v})_3 &= (-1)^4 \det\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} = u_1 \cdot v_2 - v_1 \cdot u_2\end{aligned}$$

Man beachte das negative Vorzeichen in der zweiten Komponente! In der Praxis sollte man diese Formeln für die einzelnen Komponenten nicht auswendiglernen. Es ist besser man merkt sie sich über die Determinante...

Und nun betrachten wir noch zwei konkrete Beispiele im $\mathbb{R}^3$.

Beispiel 2. Wir betrachten die beiden Vektoren

$$\vec{u} = (1,1,0) \quad \text{und} \quad \vec{v} = (0,1,1)$$

und wollen einen Vektor finden, der auf $\vec{u}$ und $\vec{v}$ orthogonal steht. Wir benutzen das Kreuzprodukt:

$$\begin{pmatrix}1\\1\\0\end{pmatrix} \times \begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}1\cdot 1 - 0\cdot 1\\0\cdot 0 - 1\cdot 1\\1\cdot 1 - 0\cdot 1\end{pmatrix} = \begin{pmatrix}1\\-1\\1\end{pmatrix}.$$

Beispiel 3. Wir betrachten den Vektor $\vec{u} = (1,2,1)$ und wollen zwei linear unabhängige Vektoren finden, die auf $\vec{u}$ normal stehen. Wir sollen das Kreuzprodukt verwenden. (Man könnte das auch durch Hinschauen lösen, indem man so ähnlich vorgeht, wie in unserem Beispiel im $\mathbb{R}^2$).
Das Problem hierbei ist, dass wir nur einen Vektor zur Verfügung haben, das Kreuzprodukt im $\mathbb{R}^3$ benötigt aber zwei! Das Problem umgehen wir folgendermaßen: Wir ergänzen den Vektor $\vec{u}$ zu einer Basis des $\mathbb{R}^3$, indem wir Standardbasisvektoren benutzen. Zum Beispiel so:

$$B = \{\vec{u}, \vec{e}_1, \vec{e}_2\}.$$

Das ist eine Basis, davon kann man sich leicht überzeugen. Nun berechnen wir das Kreuzprodukt des Vektors $\vec{u}$ mit diesen Standaradbasisvektoren:

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

Diese beiden Vektoren sind linear unabhängig und stehen auf $\vec{u}$ orthogonal.

3.2 Eigenschaften des Kreuzprodukts

Das Kreuzprodukt besitzt einige Eigenschaften, die mehr von theoretischer als praktischer Relevanz sind. Der Vollständigkeit halber und weil sie für gewisse Beweise benötigt werden, gebe ich sie hier trotzdem an:

Eigenschaften 3.2.1. Fassen wir das Kreuzprodukt im $\mathbb{R}^n$ als Abbildung

$$\begin{aligned} \times : \quad & \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n-1 \text{ mal}} \longrightarrow \mathbb{R}^n \\ & (\vec{a}_2, \ldots, \vec{a}_n) \longmapsto \vec{a}_2 \times \cdots \times \vec{a}_n \end{aligned}$$

auf, dann hat diese Abbildung folgende Eigenschaften:

1. Das Kreuzprodukt ist multilinear, das heißt linear in jeder Komponente.
2. Es ist alternierend und somit schiefsymmetrisch.
3. Das Kreuzprodukt erfüllt

$$\vec{a}_2 \times \cdots \times \vec{a}_n = \vec{0} \quad \Leftrightarrow \quad (\vec{a}_2, \ldots, \vec{a}_n) \text{ linear abhängig.}$$

4. $\vec{a}_2 \times \cdots \times \vec{a}_n$ steht orthogonal auf alle $\vec{a}_i$ für $i = 2, \ldots, n$.

Beweis. Die Eigenschaften $2 - 4$ folgen direkt aus der Definition des Kreuzprodukts. Eigenschaft 1 können wir mit der Berechnungsformel 3.7 zeigen. Für ein j zwischen 2 und n gilt

$$\begin{aligned} &(\vec{a}_2 \times \cdots \times \vec{a}_j + \lambda \cdot \vec{b}_j \times \cdots \times \vec{a}_n)_i = (-1)^{i+1} \cdot det(\vec{a}_2, \ldots, \vec{a}_j + \lambda \cdot \vec{b}_j, \ldots, \vec{a}_n) \\ &= (-1)^{i+1} \cdot det(\vec{a}_2, \ldots, \vec{a}_j, \ldots, \vec{a}_n) + \lambda \cdot (-1)^{i+1} \cdot det(\vec{a}_2, \ldots, \vec{b}_j, \ldots, \vec{a}_n) \\ &= (\vec{a}_2 \times \cdots \times \vec{a}_j \times \cdots \times \vec{a}_n)_i + \lambda \cdot (\vec{a}_2 \times \cdots \times \vec{b}_j \times \cdots \times \vec{a}_n)_i \end{aligned}$$

wobei in den Determinanten jeweils die i-te Zeile gestrichen wird. Die Multilinearität folgt nun daraus, dass zwei Vektoren genau dann übereinstimmen, wenn sie in all ihren Komponenten übereinstimmen. □

4 Das Kreuzprodukt im $\mathbb{C}^n$

Wir wollen ja das Kreuzprodukt verwenden, um lineare Gleichungssysteme zu lösen, bzw. die Eigenvektoren zu gewissen Eigenwerten einer Matrix zu bestimmen. Selbst, wenn man nur reelle Matrizen betrachtet, kann es vorkommen, dass komplexe Eigenwerte auftreten. Bei der Berechnung der Eigenvektoren sieht man sich dann mit komplexen Matrizen konfrontiert. Deshalb ist es sinnvoll, das Kreuzprodukt auch im $\mathbb{C}^n$ zu definieren. Dabei bleibt vieles aus dem $\mathbb{R}^n$ erhalten, manches muss aber abgeändert werden. Eine Sache, die leicht modifiziert werden muss, ist das bekannte Standardskalarprodukt.

Ist man nur an Reellem interessiert, kann dieses Kapitel natürlich übersprungen werden. Ich schätze aber, früher oder später werden einem auch komplexe Zahlen in der Linearen Algebra begegnen.

4.1 Das komplexe Skalarprodukt

Unser Setting ist der Vektorraum $V = \mathbb{C}^n$ über dem Körper $\mathbb{C}$. Das komplexe Standardskalarprodukt sieht folgendermaßen aus:

$$\begin{aligned} \cdot : \qquad \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \quad \mathbb{C} \\ (\ \vec{u}, \vec{v}\ ) &\longmapsto\ \vec{u} \cdot \vec{v}\ :=\ \sum_{i=0}^{n} u_i \cdot \overline{v_i} \end{aligned} \tag{4.1}$$

Der einzige Unterschied zum Standardskalarprodukt im $\mathbb{R}^n$ ist die komplexe Konjugation der Komponenten v_i.

Warum die komplexe Konjugation?

Zu jedem Skalarprodukt ist durch Setzung von $||\vec{x}|| := \sqrt{\vec{x} \cdot \vec{x}}$ eine Norm gegeben. Eine Norm nimmt wiederum nur **reelle** Werte an. Dazu muss nun $\vec{x} \cdot \vec{x} \geqslant 0$ sein (so ähnlich wie x^2 in z.B. den ganzen Zahlen $\geqslant 0$ ist). Würden wir nun auf das Konjugieren verzichten und einfach das Standardskalarprodukt im $\mathbb{R}^n$ hernehmen, hätten wir schon bei dem Vektor $\vec{x} = (i, 0)$ ein Problem: Denn $\vec{x} \cdot_{\mathbb{R}} \vec{x} = i^2 = -1$.

Wichtige Unterschiede zum Skalarprodukt im $\mathbb{R}^n$

Hier noch zwei Unterschiede des komplexen Skalarprodukts zum reellen, auf die ich sicherheitshalber explizit hinweisen möchte.

1. Das komplexe Skalarprodukt ist linear im ersten Argument, jedoch **sesquilinear** im zweiten Argument, das heißt für $\lambda \in \mathbb{C}$

$$(\lambda\vec{x}) \cdot \vec{y} = \lambda(\vec{x} \cdot \vec{y}) \quad \text{aber} \quad \vec{x} \cdot (\lambda\vec{y}) = \overline{\lambda}(\vec{x} \cdot \vec{y}).$$

2. Das komplexe Skalarprodukt ist nur 'fast' kommutativ. Es erfüllt die **Hermite'sche Symmetrie**

$$\vec{x} \cdot \vec{y} = \overline{\vec{y} \cdot \vec{x}}$$

Ich möchte noch auf folgenden Zusammenhang zwischen dem Standardskalarprodukt im Reellen und dem komplexen Skalarprodukt hinweisen, da wir ihn später ohnehin brauchen werden:

Lemma 4.1.1. *Seien $\vec{x}, \vec{y}$ komplexe Vektoren. Mit $\cdot_{\mathbb{R}}$ sei das reelle, mit $\cdot_{\mathbb{C}}$ das komplexe Standardskalarprodukt notiert. Es gilt folgender Zusammenhang*

$$\vec{x} \cdot_{\mathbb{R}} \vec{y} = \vec{x} \cdot_{\mathbb{C}} \overline{\vec{y}}. \tag{4.2}$$

Beweis. Es gilt

$$\vec{x} \cdot_{\mathbb{R}} \vec{y} = \sum_{i=1}^{n} x_i \cdot y_i = \sum_{i=1}^{n} x_i \cdot \overline{\overline{y_i}} = \vec{x} \cdot_{\mathbb{C}} \overline{\vec{y}}.$$

□

Ansonsten übertragen sich die Begriffe Orthogonalität und Komplement, sowie das Resultat mit der direkten Summe 1 zu 1 auf den $\mathbb{C}^n$ ausgestattet mit dem obigen komplexen Standardskalarprodukt.

4.2 Definition des komplexen Kreuzproduktes

Auch die Definition des Kreuzprodukts ändert sich ein wenig. Hier ist sie:

Definition 4.2.1. Gegeben seien $V = \mathbb{C}^n$ zusammen mit dem komplexen Standardskalarprodukt. Außerdem seien $n-1$ Vektoren $\vec{a}_2, \dots, \vec{a}_n$ im $\mathbb{C}^n$ gegeben. Wir unterscheiden zwei Fälle.

Fall 1 Falls $\vec{a}_2, \dots, \vec{a}_n$ linear unabhängig sind, so ist das Kreuzprodukt der Vektoren $\vec{a}_2, \dots, \vec{a}_n$ definiert als

$$\vec{a}_2 \times \cdots \times \vec{a}_n := \overline{det\,(\vec{n}, \vec{a}_2, \dots, \vec{a}_n)} \cdot \vec{n} \tag{4.3}$$

wobei der Vektor $\vec{n}$ folgendermaßen zu wählen ist:

- $\vec{n} \cdot \vec{a}_i = 0$ für $i = 2, \dots, n$, d.h. $\vec{n}$ steht normal auf alle $\vec{a}_i$
- $||\ \vec{n}\ ||_2 = \sqrt{\vec{n} \cdot \vec{n}} = 1$, d.h der Vektor $\vec{n}$ ist von der Länge Eins.

Fall 2 Falls $\vec{a}_2, \dots, \vec{a}_n$ linear abhängig sind, dann definieren wir das Kreuzprodukt der Vektoren als den Nullvektor:

$$\vec{a}_2 \times \cdots \times \vec{a}_n := \vec{0} \tag{4.4}$$

Man beachte die komplexe Konjugation in 4.3!

Warum wird komplex konjugiert?

Es geht (mal wieder) um die Wohldefiniertheit des Kreuzprodukts. Da wir den Vektor $\vec{n}$ wählen dürfen, darf der Wert des Kreuzprodukts nicht von dieser Wahl abhängen. Im Reellen war das folgendermaßen: Es hat nur zwei normierte Vektoren gegeben, die auf alle Vektoren $\vec{a}_i$ orthogonal stehen – und zwar $\vec{n}$ und $-\vec{n}$. Und es war völlig egal, welchen der beiden Vektoren wir wählen, das reelle Kreuzprodukt hat für beide dasselbe Ergebnis geliefert. Das war die Wohldefiniertheit im **Reellen**.

Wie sieht das Ganze nun im **Komplexen** aus?

Wir nehmen wieder an, dass $U := \mathcal{L}\{\vec{a}_2, \dots, \vec{a}_n\}$. Dann hat $U^\perp$ die Dimension 1. Wählen wir nun einen Vektor $\vec{n} \neq \vec{0}$ aus $U^\perp$ und normieren ihn. Dann erfüllt der Vektor $\vec{n}$ die beiden Bedingungen in obiger Definition und somit wäre

$$\vec{a}_2 \times \cdots \times \vec{a}_n = \overline{det\,(\vec{n}, \vec{a}_2, \dots, \vec{a}_n)} \cdot \vec{n}.$$

Außerdem gilt $U^\perp = \mathcal{L}\{\vec{n}\}$.
Wir bestimmen nun alle Vektoren $\vec{x}$, die die beiden Bedingungen in obiger Definition erfüllen und zeigen dann, dass es nicht auf die Wahl ankommt. Die Bedingung 1 bedeutet nichts anderes, als $\vec{x} \in U^\perp = \mathcal{L}\{\vec{n}\}$. Das heißt, es gibt ein $\lambda \in \mathbb{C}$, sodass

$$\vec{x} = \lambda \cdot \vec{n} \quad \text{und} \quad ||\vec{x}|| = 1$$

wegen der zweiten Bedingung. Folglich gilt

$$1 = ||\vec{x}|| = ||\lambda \cdot \vec{n}|| = |\lambda| \cdot ||\vec{n}|| = |\lambda| \cdot 1 = |\lambda|$$

Wären wir im Reellen, würde nun $\lambda = \pm 1$ folgen. Da aber λ eine komplexe Zahl ist, bedeutet $|\lambda| = 1$, dass λ auf dem Einheitskreis in der komplexen Ebene liegt. Oder mathematischer

$$\lambda = e^{i\varphi} \quad \text{für ein} \quad \varphi \in [0, 2\pi)\,.$$

Somit gilt $\vec{x} = e^{i\varphi} \cdot \vec{n}$.
Benutzen wir diese $\vec{x}$ nun in der obigen Definition für das komplexe Kreuzprodukt, so erhalten wir wieder

$$\begin{aligned}
\vec{a}_2 \times \cdots \times \vec{a}_n &= \overline{det\left(\vec{x}, \vec{a}_2, \ldots, \vec{a}_n\right)} \cdot \vec{x} \\
&= \overline{det\left(e^{i\varphi} \cdot \vec{n}, \vec{a}_2, \ldots, \vec{a}_n\right)} \cdot e^{i\varphi} \cdot \vec{n} \\
&= e^{-i\varphi} \cdot \overline{det\left(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n\right)} \cdot e^{i\varphi} \cdot \vec{n} \\
&= \overline{det\left(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n\right)} \cdot \vec{n}.
\end{aligned}$$

Es spielt also tatsächlich keine Rolle, was für einen normierten Normalenvektor wir verwenden. Alle liefern dasselbe Ergebnis. Die Definition ist somit wohldefiniert.

4.3 Charakteristische Eigenschaft reloaded

Die charakteristische Eigenschaft 3.3 bleibt im Komplexen 1 zu 1 erhalten.

Lemma 4.3.1. *Für jedes $\vec{v} \in \mathbb{C}^n$ gilt*

$$\vec{v} \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n) = det\,(\vec{v}, \vec{a}_2, \ldots, \vec{a}_n) \tag{4.5}$$

*Wir nennen diese Eigenschaft die **charakteristische Eigenschaft** des Kreuzprodukts.*

Beweis. Im Falle der linearen Abhängigkeit von $\vec{a}_2, \ldots, \vec{a}_n$ steht links des Gleichheitszeichens 0, sowie auch rechts davon. In diesem Fall stimmt die Gleichheit also.
Im Falle der linearen Unabhängigkeit nehmen wir wieder die Basis $B := \{\vec{n}, \vec{a}_2, \ldots, \vec{a}_n\}$ her und betrachten die beiden Abbildungen

$$\varphi_1(\vec{v}) := \vec{v} \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n) \quad \text{und} \quad \varphi_2(\vec{v}) := det\,(\vec{v}, \vec{a}_2, \ldots, \vec{a}_n)\,.$$

Das sind Abbildungen von $\mathbb{C}^n$ nach $\mathbb{C}$. Die Abbildung φ_1 ist $\mathbb{C}$-linear, da das komplexe Skalarprodukt im ersten Argument linear ist. Abbildung φ_2 ist linear, da Determinantenfunktionen sowieso immer linear in jeder Komponente sind (gilt sogar in allgemeinen K-Vektorräumen). Wir überprüfen nun die Gleichheit dieser beiden Abbildungen auf der Basis B. Wie im Reellen gilt $\varphi_1(\vec{a}_i) = \varphi_2(\vec{a}_i) = 0$, wobei $i = 2, \ldots, n$. Für den Vektor $\vec{n}$ folgt:

$$\begin{aligned}
\varphi_1(\vec{n}) &= \vec{n} \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n) \\
&= \vec{n} \cdot \left(\overline{det\,(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n)} \cdot \vec{n}\right) \\
&= det\,(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \cdot (\vec{n} \cdot \vec{n}) \\
&= det\,(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \cdot \underset{\uparrow\ =1}{||\vec{n}||^2} \\
&= det\,(\vec{n}, \vec{a}_2, \ldots, \vec{a}_n) \\
&= \varphi_2(\vec{n}).
\end{aligned}$$

Beim dritten Gleichheitszeichen haben wir die Sesquilinearität im zweiten Argument benutzt. Somit gilt $\varphi_1 \equiv \varphi_2$, die beiden Abbildungen sind gleich. Folglich haben wir die charakteristische Eigenschaft für alle $\vec{v} \in \mathbb{C}^n$ gezeigt. $\square$

4.4 Eine Berechnungsformel für das komplexe Kreuzprodukt

Wir können nun wie im Reellen Vorgehen. Wir benutzen die charakteristische Eigenschaft, um eine Berechnungsformel für das komplexe Kreuzprodukt herzuleiten. Bei der komplexen Berechnungsformel gibt es ebenfalls einen kleinen Unterschied zur reellen. Dieser ensteht durch die Sesquilinearität im zweiten Argument des komplexen Standardskalarprodukts.

Theorem 4.4.1. *Sei $V = \mathbb{C}^n$ mit dem komplexen Standardskalarprodukt. Die i-te Komponente des komplexen Kreuzproduktvektors lässt sich folgendermaßen berechnen*

$$(\vec{a}_2 \times \cdots \times \vec{a}_n)_i = (-1)^{i+1} \cdot \overline{det\,\underbrace{(\vec{a}_2, \ldots, \vec{a}_n)}_{i\text{-te Zeile streichen}}} \tag{4.6}$$

Bei der Berechnung des Kreuzproduktvektors gehen wir also prinzipiell zuerst einmal so vor, wie im Reellen. Sind alle Komponenten des Kreuzprodukts berechnet, dürfen wir dann aber nicht vergessen, den gesamten Ergebnisvektor komplex zu konjugieren. **Obige Formel unterscheidet sich von der reellen nur durch die komplexe Konjugation!**

Beweis. Sei $\vec{e}_i$ der i-te Einheitsvektor. Es gilt

$$\begin{aligned}\overline{(\vec{a}_2 \times \cdots \times \vec{a}_n)_i} &= \vec{e}_i \cdot (\vec{a}_2 \times \cdots \times \vec{a}_n)\\ &\underset{4.6}{=} det(\vec{e}_i, \vec{a}_2, \ldots, \vec{a}_n)\end{aligned}$$

Die Konjugation kommt von der Definition des komplexen Standardskalarprodukts! Die Komponenten des zweiten Arguments im Standardskalarprodukt fließen ja in die Summe konjugiert ein.
Die rechte Determinante lösen wir nun wieder mit Laplace, das heißt wir entwickeln bezüglich der ersten Spalte. Somit erhalten wir

$$\overline{(\vec{a}_2 \times \cdots \times \vec{a}_n)_i} = (-1)^{i+1} \cdot det\ \underbrace{(\vec{a}_2, \ldots, \vec{a}_n)}_{i\text{-te Zeile streichen}}$$

Konjugation auf beiden Seiten liefert schließlich das gewünschte Ergebnis. □

Die Formeln für das Kreuzprodukt, die wir im Reellen für $n = 2$ und $n = 3$ hergeleitet haben, bleiben somit erhalten. Wir müssen lediglich jede Komponente des Ergebnisses komplex konjugieren.
Betrachten wir nun konkrete Beispiele:

Beispiel 4. Betrachten wir die beiden komplexen Vektoren

$$\vec{x} = \begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ \mathrm{i} \end{pmatrix} \quad \text{und} \quad \vec{y} = \begin{pmatrix} \mathrm{i}+1 \\ \mathrm{i}+2 \\ \mathrm{i}+3 \end{pmatrix}$$

Wir suchen einen Vektor, der bezüglich dem komplexen Standardskalarprodukt orthogonal auf die beiden Vektoren $\vec{x}$ und $\vec{y}$ steht. Wir benutzen das (komplexe) Kreuzprodukt:

$$\begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ \mathrm{i} \end{pmatrix} \times \begin{pmatrix} \mathrm{i}+1 \\ \mathrm{i}+2 \\ \mathrm{i}+3 \end{pmatrix} = \overline{\begin{pmatrix} \mathrm{i}\cdot(\mathrm{i}+3) - \mathrm{i}\cdot(\mathrm{i}+2) \\ -\left(\mathrm{i}\cdot(\mathrm{i}+3) - \mathrm{i}\cdot(\mathrm{i}+1)\right) \\ \mathrm{i}\cdot(\mathrm{i}+2) - \mathrm{i}\cdot(\mathrm{i}+1) \end{pmatrix}} = \overline{\begin{pmatrix} \mathrm{i} \\ -2\mathrm{i} \\ \mathrm{i} \end{pmatrix}}$$

$$= \begin{pmatrix} -\mathrm{i} \\ 2\mathrm{i} \\ -\mathrm{i} \end{pmatrix}$$

Prüfen wir die Orthogonalität dieses Vektors zur Kontrolle nach:

$$\begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ \mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} -\mathrm{i} \\ 2\mathrm{i} \\ -\mathrm{i} \end{pmatrix} = \mathrm{i}\cdot\mathrm{i} + \mathrm{i}\cdot(-2\mathrm{i}) + \mathrm{i}\cdot\mathrm{i} = \mathrm{i}^2 - 2\mathrm{i}^2 + \mathrm{i}^2 = 0$$

$$\begin{pmatrix} \mathrm{i}+1 \\ \mathrm{i}+2 \\ \mathrm{i}+3 \end{pmatrix} \cdot \begin{pmatrix} -\mathrm{i} \\ 2\mathrm{i} \\ -\mathrm{i} \end{pmatrix} = (\mathrm{i}+1)\cdot\mathrm{i} + (\mathrm{i}+2)\cdot(-2\mathrm{i}) + (\mathrm{i}+3)\cdot\mathrm{i}$$

$$= \mathrm{i}^2 - 2\mathrm{i}^2 + \mathrm{i}^2 + \mathrm{i} - 4\mathrm{i} + 3\mathrm{i} = 0$$

Beispiel 5. Gegeben seien die beiden Vektoren

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{y} = \begin{pmatrix} 0 \\ \mathrm{i}+1 \\ \mathrm{i}+2 \end{pmatrix}$$

Wir berechnen wieder das Kreuzprodukt

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ \mathrm{i}+1 \\ \mathrm{i}+2 \end{pmatrix} = \overline{\begin{pmatrix} \mathrm{i}+2 \\ -\mathrm{i}-2 \\ \mathrm{i}+1 \end{pmatrix}} = \begin{pmatrix} 2-\mathrm{i} \\ -2+\mathrm{i} \\ 1-\mathrm{i} \end{pmatrix} =: \vec{z}$$

und machen die Probe

$$\begin{pmatrix} 2-\mathrm{i} \\ -2+\mathrm{i} \\ 1-\mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = (2-\mathrm{i}) + (-2+\mathrm{i}) = 0$$

$$\begin{pmatrix} 2-\mathrm{i} \\ -2+\mathrm{i} \\ 1-\mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \mathrm{i}+1 \\ \mathrm{i}+2 \end{pmatrix} = (-2+\mathrm{i})\cdot(1-\mathrm{i}) + (1-\mathrm{i})\cdot(2-\mathrm{i})$$

$$= (1-\mathrm{i})\cdot(-2+2+\mathrm{i}-\mathrm{i}) = 0$$

Achtung. Man beachte, dass es unbedingt von Nöten ist, während der Berechnung des Kreuzproduktvektors $\vec{z}$, komplex zu konjugieren, damit der Vektor schließlich orthogonal bezüglich des komplexen Standardskalarprodukts steht. Würden wir auf dieses Konjugieren vergessen, also $\overline{\vec{z}}$, statt $\vec{z}$ als Ergebnis hernehmen, würde folgendes gelten:

$$\overline{\vec{z}} \cdot \vec{y} = (-2 - \mathrm{i}) \cdot (1 - \mathrm{i}) + (1 + \mathrm{i}) \cdot (2 - \mathrm{i}) = 2\mathrm{i} \neq 0$$

Wir können also nicht auf das komplexe Konjugieren verzichten!

Aber: $\overline{\vec{z}}$ steht bezüglich des reellen Standardskalarporduktes $\cdot_{\mathbb{R}}$ orthogonal auf $\vec{x}$ und $\vec{y}$, denn

$$\begin{aligned} \overline{\vec{z}} \cdot_{\mathbb{R}} \vec{x} &= (\mathrm{i} + 2) + (-\mathrm{i} - 2) = 0 \quad \text{sowie} \\ \overline{\vec{z}} \cdot_{\mathbb{R}} \vec{y} &= (-2 - \mathrm{i}) \cdot (\mathrm{i} + 1) + (\mathrm{i} + 1) \cdot (\mathrm{i} + 2) \\ &= (\mathrm{i} + 1) \cdot (\mathrm{i} - \mathrm{i} + 2 - 2) = 0. \end{aligned}$$

Das gilt auch allgemein und folgt aus Zusammenhang 4.2.

4.5 Eigenschaften des komplexen Kreuzprodukts

Bei den Eigenschaften ändert sich ebenfalls ein wenig beim komplexen Kreuzprodukt. Die Multilinearität geht verloren, dafür erhalten wir etwas anderes.

Eigenschaften 4.5.1. Fassen wir das Kreuzprodukt im $\mathbb{C}^n$ als Abbildung

$$\begin{aligned} \times : \quad \underbrace{\mathbb{C}^n \times \cdots \times \mathbb{C}^n}_{n-1 \text{ mal}} &\longrightarrow \mathbb{C}^n \\ (\vec{a}_2, \ldots, \vec{a}_n) &\longmapsto \vec{a}_2 \times \cdots \times \vec{a}_n \end{aligned}$$

auf, dann hat diese Abbildung folgende Eigenschaften:

1. Das Kreuzprodukt ist multi-**sesqui**linear, das heißt **sesqui**linear in jeder Komponente.
2. Es ist alternierend und somit schiefsymmetrisch.
3. Das Kreuzprodukt erfüllt
$$\vec{a}_2 \times \cdots \times \vec{a}_n = \vec{0} \quad \Leftrightarrow \quad (\vec{a}_2, \ldots, \vec{a}_n) \text{ linear abhängig.}$$
4. $\vec{a}_2 \times \cdots \times \vec{a}_n$ steht orthogonal auf alle $\vec{a}_i$ für $i = 2, \ldots, n$ (bezüglich des komplexen Standardskalarprodukts).

Beweis. Die Eigenschaften $2-4$ folgen direkt aus der Definition des Kreuzprodukts. Eigenschaft 1 können wir mit der Berechnungsformel 4.6 zeigen. Für ein j zwischen 2 und n gilt

$$\begin{aligned} &(\vec{a}_2 \times \cdots \times \vec{a}_j + \lambda \cdot \vec{b}_j \times \cdots \times \vec{a}_n)_i = \overline{(-1)^{i+1} \cdot det(\vec{a}_2, \ldots, \vec{a}_j + \lambda \cdot \vec{b}_j, \ldots, \vec{a}_n)} \\ &= \overline{(-1)^{i+1} \cdot det(\vec{a}_2, \ldots, \vec{a}_j, \ldots, \vec{a}_n) + \lambda \cdot (-1)^{i+1} \cdot det(\vec{a}_2, \ldots, \vec{b}_j, \ldots, \vec{a}_n)} \\ &= (-1)^{i+1} \cdot \overline{det(\vec{a}_2, \ldots, \vec{a}_j, \ldots, \vec{a}_n)} + \overline{\lambda} \cdot (-1)^{i+1} \cdot \overline{det(\vec{a}_2, \ldots, \vec{b}_j, \ldots, \vec{a}_n)} \\ &= (\vec{a}_2 \times \cdots \times \vec{a}_j \times \cdots \times \vec{a}_n)_i + \overline{\lambda} \cdot (\vec{a}_2 \times \cdots \times \vec{b}_j \times \cdots \times \vec{a}_n)_i \end{aligned}$$

wobei in den Determinanten jeweils die i-te Zeile gestrichen wird. Die Multisesquilinearität folgt nun daraus, dass zwei Vektoren genau dann übereinstimmen, wenn sie in all ihren Komponenten übereinstimmen. □

5 Das Lösen homogener linearer Gleichungssysteme mit dem Kreuzprodukt

Ziel dieses Abschnittes ist es, alle $\vec{x}$ in $\mathbb{R}^n$ oder $\mathbb{C}^n$ zu finden, die $A\vec{x} = \vec{0}$ erfüllen. A ist hierbei eine gegebene reelle oder komplexe Matrix.
Vorerst betrachten wir nur „kleine“ $n \times n$ Matrizen über $\mathbb{R}$. Mit „klein“ meine ich hierbei bis maximal $n = 5$. Bei größeren Matrizen ist die Methode mit dem Kreuzprodukt ohnehin nicht effizient. Da greift man besser auf den Gauß-Algo oder noch besser auf ein CAS zurück.
Wenn wir komplexe Zahlen hinzuziehen, können wir genauso vorgehen wie im Reellen - warum, werde ich später kurz erklären.
Außerdem setzen wir voraus, dass der Rang der Matrix bekannt ist. Manchmal ist dieser leicht an der Struktur der Matrix abzulesen, manchmal etwas schwieriger. Kann man den Rang nicht „erkennen“, kann man das LGS ebenfalls mit dem Kreuzprodukt lösen. Es kann dann eine Reduktionsmethode angewendet werden. Dazu später mehr.
Außerdem werde ich einen kleinen Trick angeben, der bei der Rangbestimmung möglicherweise helfen könnte.

In den folgenden Abschnitten werde ich explizit zeigen, wie bei $n = 2, 3, 4$ vorzugehen ist, um $A\vec{x} = 0$ zu lösen und Beispiele angeben. Dazu gehe ich nach dem Rang der Matrizen vor, von dem wir ja vorausgesetzt haben, dass er bekannt ist.
Die beiden Fälle $rg(A) = n$, d. h. $\det(A) \neq 0$ und $rg(A) = 0$, d. h. A ist die Nullmatrix, werde ich immer auslassen. Denn im ersten Fall ist $\vec{x} = \vec{0}$ die einzige Lösung und im zweiten Fall ist der Lösungsraum der gesamte $\mathbb{R}^n$.
Außerdem werden wir die vorgestellte Methode in einem der nächsten Abschnitte verwenden, um die Eigenvektoren von Matrizen zu ermitteln. Und da ist $\det(A - \lambda I_n)$ ohnehin immer $= 0$ und demnach der Rang niemals $= n$. Noch ein Grund, warum dieser Fall uninteressant ist.

5.1 Grundsätzliche Idee

Also, wir wollen alle Lösungen von $A\vec{x} = \vec{0}$ ermitteln, oder anders gesagt: eine Basis des Lösungsraumes finden. A sei, wie vorher erwähnt, eine reelle Matrix. Außerdem sei $rg(A) = k$. Wir schreiben nun die Matrix A folgendermaßen

$$A = \begin{pmatrix} \text{———} & \vec{a}_1 & \text{———} \\ & \vdots & \\ \text{———} & \vec{a}_k & \text{———} \\ & \vdots & \\ \text{———} & \vec{a}_n & \text{———} \end{pmatrix},$$

wobei, $\vec{a}_1, \ldots, \vec{a}_n$ die Zeilen der Matrix beschreiben. Außerdem können wir o. B. d. A. annehmen, dass $\vec{a}_1, \ldots, \vec{a}_k$ linear unabhängig sind. Denn: $rg(A) = k$, d. h. es gibt schon einmal k-linear unabhängige Zeilen. Liegen diese nicht in den ersten k-Zeilen der Matrix, vertauschen wir einfach Zeilen, bis es der Fall ist. Zeilenvertauschungen ändern ja nichts am Lösungsraum von $A\vec{x} = \vec{0}$.
Was heißt „Finde $\vec{x}$, sodass $A\vec{x} = \vec{0}$“ eigentlich? Es bedeutet: Finde $\vec{x}$, sodass

$$A\vec{x} = \begin{pmatrix} \vec{a}_1 & * & \vec{x} \\ & \vdots & \\ \vec{a}_k & * & \vec{x} \\ & \vdots & \\ \vec{a}_n & * & \vec{x} \end{pmatrix} = \vec{0}.$$

Hierbei beschreibt $*$ die ganz normale Matrixmultiplikation (hier nur so notiert, um sie vom Standardskalarprodukt abzugrenzen).
Formulieren wir es noch etwas anders, dieses Mal mit unserem reellen Standardskalarprodukt $\cdot$: Finde $\vec{x}$, sodass

$$\vec{a}_1^{\,T} \cdot \vec{x} = 0 \quad , \quad \ldots \quad , \quad \vec{a}_k^{\,T} \cdot \vec{x} = 0 \quad , \quad \ldots \quad , \quad \vec{a}_n^{\,T} \cdot \vec{x} = 0.$$

Das ^{T} beschreibt die Transposition einer Matrix. Hier in diesem Fall bedeutet es einfach, dass wir die Zeilenvektoren als Spaltenvektoren auffassen.
Da $\vec{a}_1^{\,T}, \ldots, \vec{a}_k^{\,T}$ wegen ihrer linearer Unabhängigkeit und $rg(A) = k$ eine Basis des Zeilenraumes bilden, können wir $\vec{a}_{k+1}^{\,T}, \ldots, \vec{a}_n^{\,T}$ als Linearkombination dieser Basisvektoren schreiben.
Es reicht also diese $\vec{x}$ zu finden, die

$$\vec{a}_1^{\,T} \cdot \vec{x} = 0 \quad , \quad \ldots \quad , \quad \vec{a}_k^{\,T} \cdot \vec{x} = 0$$

erfüllen.
Das bedeutet aber gerade, dass wir all die Vektoren $\vec{x}$ suchen, die auf $\vec{a}_1^{\,T}, \ldots, \vec{a}_k^{\,T}$ **orthogonal** stehen.
Wir wollen also das orthogonale Komplement von $U = \mathcal{L}\{\vec{a}_1^{\,T}, \ldots, \vec{a}_k^{\,T}\}$ berechnen. Da $dim(U) = k$, ist $dim(U^{\perp}) = n - k$. Das orthogonale Komplement enthält somit $n - k$ linear unabhängige Vektoren, die wegen obiger Überlegung Lösungen des Gleichungssystems $A\vec{x} = 0$ sind.
Und das sind auch schon alle - diese Vektoren spannen unseren Lösungsraum auf. Denn aus dem Dimensionssatz folgt

$$dim(kerA) = dim(\mathbb{R}^n) - rg(A) = n - k.$$

Somit wissen wir: $n-k$ linear unabhängige Lösungsvektoren bilden bereits eine Basis des Lösungsraumes.

Unsere Hauptaufgabe lässt sich also darauf reduzieren, das orthogonale Komplement von Unterräumen halbwegs effizient zu bestimmen. Dafür haben wir aber ein Hilfsmittel: Das Kreuzprodukt!
Wie das konkret für verschiedene Größen und Ränge von Matrizen funktioniert, werden wir in den nachfolgenden Abschnitten sehen. Es ist wirklich nicht so kompliziert, wie es auf den ersten Blick scheint!

5.2 Der Fall n=2

Fangen wir einfach an und betrachten eine 2×2 Matrix A über $\mathbb{R}$. Wir bestimmen zuerst den Rang dieser Matrix A. Viele Möglichkeiten haben wir da nicht: Die Matrix kann Rang 0 haben, dann ist aber A die Nullmatrix – wie schon einmal gesagt, ein uninteressanter Fall. Bleiben also noch die Fälle, dass die Matrix Rang 1 oder 2 besitzt.

Wie sieht die Matrix aus, wenn sie Rang 1 hat? Rang 1 bedeutet, dass es einen Zeilenvektor $\vec{a}_1 \neq 0$ gibt, der den Zeilenraum aufspannt! Nehmen wir einfach mal an, dass dieser Vektor $\vec{a}_1$ in der ersten Zeile der Matrix liegt (wenn nicht vertauschen wir einfach Zeilen). Die Matrix hat dann folgende Struktur:

$$A = \begin{pmatrix} — \vec{a}_1 — \\ — \lambda\, \vec{a}_1 — \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ \lambda a_{11} & \lambda a_{12} \end{pmatrix},$$

mit einem $\lambda \in \mathbb{R}$. Es ist also leicht, zu erkennen, wann eine Matrix $\neq 0$ Rang 1 besitzt. Die Zeilen (bzw. Spalten) sind einfach skalare Vielfache voneinander. Ist das nicht der Fall, hat sie Rang 2, und die einzige Lösung des Systems $A\vec{x} = \vec{0}$ ist der Nullvektor.

Was ist die Lösung im Falle, dass A Rang 1 annimmt. Wir können die Lösung ohne Rechnung angeben:

Satz 5.2.1. *Sei A eine reelle 2×2 Matrix. Der Rang der Matrix A sei 1. Das heißt die Matrix ist von der Struktur*

$$A = \begin{pmatrix} — \vec{a}_1 — \\ — \lambda\, \vec{a}_1 — \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ \lambda a_{11} & \lambda a_{12} \end{pmatrix},$$

wobei $\lambda \in \mathbb{R}$ und o. B. d. A. $\vec{a}_1 \neq 0$ (sonst Zeilenvertauschung).
Der Lösungsraum L des homogenen linearen Gleichungssystems $A\vec{x} = \vec{0}$ ist dann gegeben durch

$$L = \mathcal{L}\left\{\begin{pmatrix} -a_{12} \\ a_{11} \end{pmatrix}\right\}. \tag{5.1}$$

Beweis. Nach der Überlegung des vorigen Abschnittes wissen wir, dass es ausreicht, einen Vektor $\vec{n}$ zu finden, der normal auf $\vec{a}_1{}^T$ steht. Dieser spannt bereits den Lösungsraum auf. Diesen Vektor haben wir aber im $\mathbb{R}^2$ schon mehrfach berechnet. Siehe z.B. hier. Somit gilt

$$\vec{n} = \begin{pmatrix} -a_{12} \\ a_{11} \end{pmatrix}$$

und wir sind fertig. Man kann sich auch leicht durch Einsetzen davon überzeugen, dass $\vec{n}$ tatsächlich eine (linear unabhängige) Lösung ist. □

Beispiel 6. Gegeben sei die Matrix

$$A = \begin{pmatrix} \sqrt{2} & \sqrt{3} \\ 2 & \sqrt{6} \end{pmatrix}.$$

Wir suchen die Lösung des linearen Gleichungssystems $A\vec{x} = \vec{0}$.

Wir sehen, dass die zweite Zeile der Matrix durch Skalarmultiplikation der ersten Zeile mit $\sqrt{2}$ zu Stande kommt. Die Matrix hat also Rang 1. Alternativ hätten wir auch die Determinante zu Hilfe nehmen können: $\det(A) = 0$, und wir erhalten wieder $rg(A) = 1$.
Der Lösungsraum ist also nach obigem Satz gegeben durch

$$L = \mathcal{L}\left\{\begin{pmatrix} -\sqrt{3} \\ \sqrt{2} \end{pmatrix}\right\},$$

wovon man sich leicht durch Einsetzen überzeugen kann.

5.3 Der Fall n=3

Für 3×3 Matrizen A sieht die Sache ähnlich einfach aus. Fangen wir damit an, den Fall zu betrachten, wo A Rang 1 besitzt.

Vorgehen bei rg(A)=1

Wir haben folgendes Resultat

Satz 5.3.1. *Sei A eine reelle 3×3 Matrix mit $rg(A) = 1$. Das heißt sie hat folgende Gestalt*

$$A = \begin{pmatrix} - \ \vec{a}_1 \ - \\ - \ \lambda\, \vec{a}_1 \ - \\ - \ \mu\, \vec{a}_1 \ - \end{pmatrix},$$

wobei o. B. d. A. $\vec{a}_1 \neq \vec{0}$ (sonst Zeilenvertauschung) und λ, μ reelle Zahlen sind.

Die Lösung des LGS $A\vec{x} = \vec{0}$ hängt nun von der Gestalt des Vektors $\vec{a}_1$ ab:

Fall 1: *$\vec{a}_1$ besitzt genau einen Eintrag $\neq 0$ und zwei Einträge $= 0$. Dann wird der Lösungsraum von zwei Einheitsvektoren aufgespannt.*

Z. B. falls $\vec{a}_1 = (a_{11}, 0, 0)$ mit $a_{11} \neq 0$, so ist

$$L = \mathcal{L}\{\vec{e}_2, \vec{e}_3\} \tag{5.2}$$

Fall 2 *$\vec{a}_1$ besitzt mindestens zwei Einträge $\neq 0$. Dann können zwei l. u. Lösungsvektoren durch „Vertauschen und Vorzeichenwechsel“ der Einträge von $\vec{a}_1$ gefunden werden.*

Z. B. falls $\vec{a}_1 = (a_{11}, a_{12}, a_{13})$ mit $a_{11}, a_{12} \neq 0$, dann ist der Lösungsraum gegeben durch

$$L = \mathcal{L}\left\{ \begin{pmatrix} -a_{12} \\ a_{11} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -a_{13} \\ a_{12} \end{pmatrix} \right\}. \tag{5.3}$$

Beweis. Im Fall 1 ist offensichtlich, dass die Einheitsvektoren orthogonal auf $\vec{a}_1$ stehen und linear unabhängig sind. Es ist in diesem Fall also nichts zu zeigen.

Wenn wir im Fall 2 den Vektor $\vec{a}_1 = (a_{11}, a_{12}, a_{13})$ mit $a_{11}, a_{12} \neq 0$ hernehmen, ist ersichtlich, dass die oben angegebenen Vektoren, die den Lösungsraum aufspannen sollen, orthogonal auf $\vec{a}_1$ stehen. Somit sind sie tatsächlich Lösungen.
Sie sind auch l. u., denn aus

$$s_1 \cdot \begin{pmatrix} -a_{12} \\ a_{11} \\ 0 \end{pmatrix} + s_2 \cdot \begin{pmatrix} 0 \\ -a_{13} \\ a_{12} \end{pmatrix} = 0 \quad \text{folgt}$$

$$a_{12} \cdot s_1 = 0 \quad \text{und} \quad a_{12} \cdot s_2 = 0.$$

Und da $a_{12} \neq 0$ gilt folglich $s_1, s_2 = 0$.
Hat $\vec{a}_1$ eine andere Struktur, d.h. zwei andere Einträge sind $\neq 0$, müssen die Lösungsvektoren leicht angepasst werden. Der Beweis funktioniert analog. $\square$

Am Besten wir schauen uns das an Beispielen an. Dadurch sollte das Ganze viel leichter zu verstehen sein. Es ist nämlich wirklich nicht schwer!

Beispiel 7. Gegeben seien die Matrizen

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 42 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 42 & 0 \end{pmatrix}$$

Wir sehen, dass die zweite und dritte Zeile skalare Vielfache der ersten Zeile sind. Außerdem ist in der ersten Zeile genau ein Eintrag $\neq 0$. Somit sind wir im Fall 1. Folglich wird der Lösungsraum von Einheitsvektoren aufgespannt.
Bei A_1 ist der Lösungsraum durch $L = \mathcal{L}\{\vec{e}_2, \vec{e}_3\}$ gegeben. Bei A_2 durch $L = \mathcal{L}\{\vec{e}_1, \vec{e}_3\}$.

Beispiel 8. Gegeben seien die Matrizen

$$A_1 = \begin{pmatrix} \sqrt{3} & 0 & \sqrt{2} \\ 2\sqrt{3} & 0 & 2\sqrt{2} \\ 3 & 0 & \sqrt{6} \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 4 & 8 & 12 \end{pmatrix}$$

Auch hier sind die zweite und dritte Zeile skalare Vielfache der ersten. Der Rang der Matrizen ist somit 1. Hier sind aber zwei Einträge in der ersten Zeile $\neq 0$. Somit sind wir im Fall 2. Wir lösen das durch „Vertauschen und Vorzeichenwechsel“. Der Lösungsraum für A_1 ist

$$L = \mathcal{L}\left\{ \begin{pmatrix} -\sqrt{2} \\ 0 \\ \sqrt{3} \end{pmatrix}, \begin{pmatrix} 0 \\ -\sqrt{2} \\ 0 \end{pmatrix} \right\}$$

Man beachte, dass wir zuerst die Einträge, die ungleich Null sind, bearbeitet haben.
Der Lösungsraum von A_2 ist z.B.

$$L = \mathcal{L}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix} \right\} = \mathcal{L}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\} = \mathcal{L}\left\{ \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix} \right\}$$

Wie man sieht haben wir hier mehrere Möglichkeiten zu vertauschen und jede liefert das richtige Ergebnis. Ein Vektorraum kann ja ganz viele Basen haben!

Vorgehen bei rg(A)=2

Hat die Matrix Rang 2 ist das Ganze noch einfacher. Wir brauchen nur das Kreuzprodukt berechnen und sind fertig. Genaueres steht hier:

Satz 5.3.2. *Sei A eine reelle 3×3 Matrix mit $rg(A) = 2$. Das heißt sie hat o. B. d. A. (sonst Zeilenvertauschung) folgende Gestalt*

$$A = \begin{pmatrix} — & \vec{a}_1 & — \\ — & \vec{a}_2 & — \\ — & \lambda\, \vec{a}_1 + \mu\, \vec{a}_2 & — \end{pmatrix},$$

wobei $\vec{a}_1$ und $\vec{a}_2$ linear unabhängig und λ, μ reelle Zahlen sind.

Der Lösungsraum des homogenen LGS $A\vec{x} = 0$ ist dann gegeben durch

$$L = \mathcal{L}\{\vec{a}_1 \times \vec{a}_2\} \tag{5.4}$$

Beweis. Der Vektor $\vec{a}_1 \times \vec{a}_2$ steht sowohl orthogonal auf $\vec{a}_1$, als auch auf $\vec{a}_2$. Somit ist er eine Lösung des LGS. Außerdem folgt aus der Definition des Kreuzprodukts, dass $\vec{a}_1 \times \vec{a}_2 \neq \vec{0}$, denn $\vec{a}_1$ und $\vec{a}_2$ sind linear unabhängig. □

Beispiel 9. Gegeben sei die folgende Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -2 & -1 & 0 \end{pmatrix}$$

Die Matrix hat nicht Rang 1, denn dann wären die Zeilen skalare Vielfache voneinander, was offensichtlich nicht der Fall ist.
Sie hat auch nicht Rang 3. Das kann man z.B. über die Determinante überprüfen. Diese ist $= 0$. Ich würde aber empfehlen, einfach einmal zwei linear unabhängige Vektoren herzunehmen und das Kreuzprodukt zu berechnen. Dann macht man die Probe beim letzten verbleibenden Vektor: Passt die Probe, bedeutet das Rang 2, passt sie nicht, dann bedeutet das Rang 3 und $\vec{x} = \vec{0}$ ist die einzige Lösung.
Machen wir das einfach einmal mit der 1. und der 3. Zeile, die sind linear unabhängig. Das Kreuzprodukt dieser beiden Vektoren ist

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \\ 3 \end{pmatrix}$$

Und jetzt machen wir die Probe beim verbleibenden Vektor. Dieser Vektor ist die 2. Zeile der Matrix:

$$4 \cdot 3 + 5 \cdot (-6) + 6 \cdot 3 = 12 - 30 + 18 = 0.$$

Passt also alles! Somit hat die Matrix Rang 2 und obiger Vektor spannt den Lösungsraum auf.

Betrachten wir noch ein Beispiel, wo wir sehen, was bei einer Rang 3 Matrix passiert.

Beispiel 10. Dieses Mal sei

$$A = \begin{pmatrix} \sqrt{42} & \sin(42) & \cos(42) \\ 1 & 1 & 2 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \end{pmatrix}$$

Offensichtlich hat die Matrix nicht Rang 1. Wir wählen die 2. und 3. Zeile, diese sind linear unabhängig, da sie keine skalaren Vielfachen voneinander sind und berechnen das Kreuzprodukt:

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \\ 0 \end{pmatrix} \Big|\Big| \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Das Symbol || bedeutet, dass wir den Vektor skaliert haben (um den Faktor $1/\sqrt{2}$). Die Länge des Vektor ändert ja nichts an der Orthogonalität.
Nun machen wir die Probe beim verbleibenden Vektor, das ist der erste Zeilenvektor:

$$-\underbrace{\sqrt{42}}_{\in(6,7)} + \underbrace{\sin(42)}_{\in(-1,1)} \neq 0$$

Das passt nicht! Somit ist $rg(A) = 3$ und der Nullvektor ist die einzige Lösung.

Wir haben nun gesehen, dass wir im Fall $n = 3$ prinzipiell komplett auf den Gauß-Algo verzichten können. Matrizen mit Rang 1 sind leicht zu erkennen und die Lösung ist schnell angegeben. Zwischen Rang 2 und 3 können wir mit dem Kreuzprodukt unterscheiden. Die Lösung kann ebenfalls damit berechnet werden.

5.4 Der Fall n=4

Nun wird es etwas interessanter. Im Gegensatz zum Fall $n = 3$, wo der Rang entweder leicht ersichtlich ist (falls $rg(A) = 1$) oder im Laufe der Rechnung feststeht (im Fall $rg(A) > 1$), kann es bei 4×4 Matrizen schwieriger sein, den Rang zu erkennen. Außerdem ist es aufwändiger orthogonale Vektoren zu berechnen, da wir im Extremfall vier 3×3 Determinanten für das Kreuzprodukt berechnen müssen. Jeder muss also selbst entscheiden, mit welcher Methode er schneller ist: Mit der hier vorgestellten oder dem altbekannten Gauß-Algorithmus.

Vorgehen bei rg(A)=1 oder rg(A)=3

Hier gibt es nicht viel zu erklären, da diese beiden Fälle im Grunde schon bei 3×3 Matrizen vorgekommen sind.

Im Falle, dass $rg(A) = 1$ ist, wird der Zeilenraum wieder von einem Vektor $\vec{a}_1$ aufgespannt. Die linear unabhängigen Lösungsvektoren von $A\vec{x} = \vec{0}$, können dann wieder durch „Vertauschen und Vorzeichenwechsel" der Einträge von $\vec{a}_1$ ermittelt werden.

Bei $rg(A) = 3$ (und auch bei $rg(A) = 4$) haben wir drei linear unabhängige Zeilenvektoren zur Verfügung. Wir berechnen das Kreuzprodukt (im $\mathbb{R}^4$) dieser drei Vektoren. Anschließend machen wir mit diesem berechneten Vektor noch die Probe beim verbleibenden Zeilenvektor. Passt alles, haben wir $rg(A) = 3$ und der Kreuzproduktvektor spannt den Lösungsraum auf. Stimmt die Probe nicht, dann ist $rg(A) = 4$.

Beispiel 11. Betrachten wir die Matrix

$$A = \begin{pmatrix} 1 & 2 & 2 & 2 \\ 3 & 1 & 1 & 1 \\ 0 & 2 & 1 & 2 \\ 4 & 7 & 5 & 7 \end{pmatrix}$$

Wir vermuten einfach mal, dass die ersten drei Zeilen l. u. sind, somit $rg(A) \geqslant 3$. Berechnen wir das Kreuzprodukt dieser drei Vektoren:

$$\begin{pmatrix} 1 \\ 2 \\ 2 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ 1 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2+0-2 \\ -(1-12+6) \\ 0-12+12 \\ -(-1-6+12) \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 0 \\ -5 \end{pmatrix} \parallel \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

Wie haben wir das Kreuzprodukt berechnet? Gehen wir auf die Berechnung des zweiten Eintrags im Detail ein! Für den zweiten Eintrag muss die zweite Zeile gestrichen werden und folgendes berechnet werden:

$$(-1)^{2+1} \cdot \det \begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

Für die Determinante kann man Sarrus oder Laplace verwenden. Man kann aber auch wieder das Kreuzprodukt benutzen. das funktioniert mit 3.3:

$$\det\begin{pmatrix}1&3&0\\2&1&1\\2&1&2\end{pmatrix} = \begin{pmatrix}1\\2\\2\end{pmatrix}\cdot\begin{pmatrix}3\\1\\1\end{pmatrix}\times\begin{pmatrix}0\\1\\2\end{pmatrix} = \begin{array}{l}1\cdot(1\cdot 2-1\cdot 1)+\\ 2\cdot(0\cdot 1-3\cdot 2)+\\ 2\cdot(3\cdot 1-1\cdot 0)\end{array} = 1-12+6.$$

Da der Kreuzproduktvektor $\neq \vec{0}$ ist, sind die ersten drei Zeilen tatsächlich linear unabhängig und somit hat unsere Annahme $rg(A) \geqslant 3$ gestimmt.
Wenn wir nun den Kreuzproduktvektor in die letzte Zeile einsetzen, sehen wir, dass ebenfalls Null herauskommt. Somit $rg(A) = 3$ und der Kreuzproduktvektor ist l. u. Lösung des homogenen GLS $A\vec{x} = \vec{0}$.

Tipp 5.4.1. Wie berechne ich das Kreuzprodukt am geschicktesten und ohne Fehler zu machen?
Arbeitet man auf einem Blatt Papier ist folgendes zu empfehlen: Zuerst schreibt man sich die Vektoren, die im Kreuzprodukt involviert sind schön und groß nebeneinander auf.
Dann zeichnet man die erste Klammer für den Lösungsvektor und notiert sich die Vorzeichen, damit man sie später nicht vergisst. In etwa so:

$$\left(\begin{array}{c}+\\-\\+\\\vdots\end{array}\right.$$

Für die Berechnung des ersten Eintrages, deckt man sich die erste Zeile z.B. mit einem Bleistift oder Kugelschreiber ab. Von dem, was noch sichtbar ist, berechnet man die Determinante, z.B. so, wie im obigen Beispiel. Zwischenergebnisse kann man sich zur Sicherheit irgendwo notieren.
Ist der erste Eintrag berechnet, fährt man mit dem zweiten fort: diesmal die zweite Zeile abdecken, Determinante berechnen, Ergebnis in den Lösungsvektor eintragen.
So macht man weiter bis alle Einträge berechnet sind.

Vorgehen bei rg(A)=2

Dieser Fall erfordert ein etwas anderes Vorgehen als bisher. Da wir nur zwei linear unabhängige Zeilenvektoren zur Verfügung haben, können wir das Kreuzprodukt (noch) nicht berechnen. Für das Kreuzprodukt benötigen wir drei linear unabhängige Vektoren. Das Zauberwort lautet „Basisergänzung“.

Satz 5.4.2. *Sei A eine reelle 4×4 Matrix mit $rg(A) = 2$. Das bedeutet sie hat o. B. d. A. folgende Gestalt*

$$\begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \lambda_1 \cdot \vec{a}_1 + \mu_1 \cdot \vec{a}_2 \\ \lambda_2 \cdot \vec{a}_1 + \mu_2 \cdot \vec{a}_2 \end{pmatrix}$$

wobei $\vec{a}_1$ und $\vec{a}_2$ linear unabhängig seien und $\lambda_1, \lambda_2, \mu_1, \mu_2$ reelle Skalare.

Weiterhin sei $B = \{\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2\}$ eine Basis des $\mathbb{R}^4$, die durch Basisergänzung der linear unabhängigen Vektoren $\vec{a}_1$ und $\vec{a}_2$ zu Stande kommt.

Dann ist der Lösungsraum des homogenen LGS $A\vec{x} = \vec{0}$ gegeben durch

$$L = \mathcal{L}\left\{ \ \vec{a}_1 \times \vec{a}_2 \times \vec{b}_1 \ , \ \vec{a}_1 \times \vec{a}_2 \times \vec{b}_2 \ \right\}. \tag{5.5}$$

Beweis. Die Vektoren $\vec{a}_1 \times \vec{a}_2 \times \vec{b}_i$ sind für $i = 1, 2$ nicht der Nullvektor, da die involvierten Vektoren l. u. sind. Außerdem stehen diese beiden Kreuzproduktvektoren orthogonal auf $\vec{a}_1$ und $\vec{a}_2$. Somit sind sie tatsächlich Lösungen des homgenen LGS.
Wir müssen nun noch zeigen, dass diese beiden Lösungsvektoren linear unabhängig sind. Dann wären wir fertig, denn wir wissen, dass zwei l. u. Lösungen ausreichen, um den Lösungsraum aufzuspannen (Stichwort: Dimensionsformel).
Machen wir das indirekt: Angenommen es gibt zwei Skalare s_1, s_2 nicht beide Null (hier sogar $s_1, s_2 \neq 0$), sodass

$$s_1 \cdot (\vec{a}_1 \times \vec{a}_2 \times \vec{b}_1) \ + \ s_2 \cdot (\vec{a}_1 \times \vec{a}_2 \times \vec{b}_2) = \vec{0}.$$

Wir benutzen die Linearität des Kreuzprodukts im letzten Argument und erhalten

$$\vec{a}_1 \times \vec{a}_2 \times (s_1 \cdot \vec{b}_1 + s_2 \cdot \vec{b}_2) = \vec{0}.$$

Das Kreuzprodukt wird definitionsgemäß genau dann Null, wenn die darin involvierten Vektoren linear abhängig sind. Das wiederum bedeutet, dass es Skalare r_1, r_2, r_3 nicht alle $= 0$ gibt, sodass

$$r_1 \cdot \vec{a}_1 \ + \ r_2 \cdot \vec{a}_2 \ + \ r_3 \cdot s_1 \cdot \vec{b}_1 \ + \ r_3 \cdot s_2 \cdot \vec{b}_2 \ = \ \vec{0}.$$

Da es sich bei den Vektoren $\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2$ um Basisvektoren handelt, sind sie linear unabhängig. Daraus folgt

$$r_1 = 0 \quad , \quad r_2 = 0 \quad , \quad r_3 \cdot s_1 = 0 \quad , \quad r_3 \cdot s_2 = 0.$$

Da nicht alle r_i gleich Null sein dürfen und r_1, r_2 bereits Null sind, muss $r_3 \neq 0$ sein. Nun gilt aber $s_1, s_2 \neq 0$ und somit $r_3 \cdot s_1, \ r_3 \cdot s_2 \neq 0$. Das ist ein Widerspruch! □

Beispiel 12. Gegeben seien die zwei linear unabhängigen Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} \quad , \quad \vec{a}_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} .$$

Wir fassen diese beiden Vektoren nun als die ersten zwei Zeilen einer Matrix auf. Die dritte und vierte Zeile entstehen dadurch, dass wir die vorherigen beiden Zeilen miteinander addieren (so ähnlich wie bei den Fibonacci- Zahlen, nur eben mit Vektoren). Die entstehende Matrix hat dann sicher Rang 2. Sie sieht folgendermaßen aus:

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 2 \\ 2 & 3 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{pmatrix}$$

Nun ergänzen wir $\vec{a}_1$ und $\vec{a}_2$ zu einer Basis des $\mathbb{R}^4$. Man sollte zur Ergänzung immer Standardbasisvektoren verwenden. Unsere Basis sieht so aus: $B = \{\vec{a}_1, \vec{a}_2, \vec{e}_3, \vec{e}_4\}$. Jetzt berechnen wir

$$\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

und

$$\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} +3 \\ -1 \\ -1 \\ 0 \end{pmatrix} .$$

Das sind die zwei linear unabhängigen Lösungen. Sie spannen den Lösungsraum auf.

Tipp 5.4.3. Es ist unbedingt zu empfehlen, die Vektoren mit Standardbasisvektoren zu einer Basis zu ergänzen. Denn dann braucht man bei der Berechnung des Kreuzprodukts nur 2×2 Unterdeterminanten berechnen und ist wesentlich schneller fertig. Warum das? Exemplarisch anhand der Berechnung des ersten Eintrages: Die erste Zeile wird gestrichen, dann muss die 3×3 Unterdeterminante berechnet werden. Diese berechnen wir mit Laplace: Wir entwickeln bezüglich der letzten Spalte. Dort sind im Normalfall zwei Nullen und ein Einser. Die Zeile mit der Eins (und die Spalte) werden wegen Laplace gestrichen. Übrig bleibt eine 2×2 Matrix, deren Determinante wir berechnen müssen.
Wir müssen eigentlich nur auf das Vorzeichen aufpassen: Sowohl bei der Berechnung des Kreuzprodukts, als auch bei der Laplace Entwicklung!

5.5 Rangbestimmung mit der Determinante

Wie kann man mit Hilfe der Determinante den Rang einer reellen $n \times n$ Matrix A bestimmen? Folgendes Resultat ist manchmal hilfreich:

Satz 5.5.1. *Sei $k \leqslant n$. Seien $\vec{a}_1, \ldots, \vec{a}_k$ Vektoren im $\mathbb{R}^n$, z.B. können das die Zeilen einer Matrix sein. Außerdem sei A die Matrix, die die Vektoren $\vec{a}_1, \ldots, \vec{a}_k$ als Spalten besitzt. Dann gilt:*

Die Vektoren $\vec{a}_1, \ldots, \vec{a}_k$ sind genau dann linear unabhängig, falls es in A eine $k \times k$ Untermatrix A_k gibt – das ist eine Matrix, die durch streichen von $n-k$ Zeilen von A zu Stande kommt – mit $\det(A_k) \neq 0$.

Beweis. $(\Leftarrow)$ Angenommen es gibt in A eine $k \times k$ Untermatrix A_k mit $\det(A_k) \neq 0$. Das heißt A_k ist eine reguläre Matrix, hat also vollen Rang. Sei $\vec{s} = (s_1, \ldots, s_k)^T \in \mathbb{R}^k$.
Wir müssen zeigen, dass

$$A\vec{s} = s_1 \cdot \vec{a}_1 + \cdots + s_k \cdot \vec{a}_k = \vec{0}$$

nur für $\vec{s} = \vec{0}$ erfüllt ist.
Schreiben wir das Gleichungssystem $A\vec{s} = \vec{0}$ etwas anders auf: $A\vec{s} = \vec{0}$ bedeutet

$$A_k\vec{s} = \vec{0} \quad \text{und} \quad R_{n-k}\vec{s} = \vec{0},$$

wobei R_{n-k} die $(n-k) \times k$ Matrix ist, die die Zeilen von A enthält, die nicht schon in A_k vorhanden sind. R_{n-k} enthält also die gestrichenen Zeilen.
Die Bedingung $A_k\vec{s} = \vec{0}$ ist aber nur für $\vec{s} = \vec{0}$ erfüllt, da A_k regulär ist. Die Richtung $(\Leftarrow)$ ist somit gezeigt.

$(\Rightarrow)$ Wir gehen indirekt vor. Angenommen $\vec{a}_1, \ldots, \vec{a}_k$ seien l. u., das heißt $rg(A) = k$, aber für jede $k \times k$ Untermatrix in A ist die Determinante gleich Null. Das bedeutet nichts anderes, als dass jede Auswahl von k Zeilen in A linear abhängig sein muss. Dann kann A aber nicht Rang k haben. Falls nämlich $rg(A) = k$ müsste es k l. u. Zeilen geben. Somit haben wir einen Widerspruch erzeugt und auch diese Richtung ist gezeigt. ☐

Beispiel 13. Betrachten wir die beiden Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \\ \sqrt{\pi} \\ \sqrt{19} \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 1 \\ 1 \\ \sqrt{10} \\ \sqrt{17} \end{pmatrix}.$$

Die Vektoren sind linear unabhängig, da

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

Mit Hilfe des obigen Resultats können wir nun folgenden Satz zeigen, der nützlich ist, wenn zu wenig Vektoren vorhanden sind, um das Kreuzprodukt zu berechnen.

Satz 5.5.2. *Sei $k < n$. Je $\vec{a}_1, \ldots, \vec{a}_k$ linear unabhängige Vektoren im $\mathbb{R}^n$ können mit Standardbasisvektoren zu einer Basis des $\mathbb{R}^n$ ergänzt werden.*

Beweis. Da $\vec{a}_1, \ldots, \vec{a}_k$ linear unabhängig sind, gibt es nach dem vorigen Resultat eine Untermatrix A_k mit $\det(A_k) \neq 0$. Nehmen wir einfach einmal an, diese Untermatrix geht aus den ersten k Zeilen der Matrix

$$A = \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_k \end{pmatrix}$$

hervor. Wir wollen zeigen, dass dann $B = \{\vec{a}_1, \ldots, \vec{a}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n\}$ eine Basis des $\mathbb{R}^n$ ist. Dazu betrachten wir die Matrix, die die Vektoren in der Basis B als Spalten besitzt. Nennen wir diese Matrix C. Dann gilt

$$\det(C) = \det \begin{pmatrix} A_k & 0 \\ * & I_{n-k} \end{pmatrix} = \det(A_k) \cdot \det(I_{n-k}) = \det(A_k) \neq 0.$$

Die Matrix C haben wir als Blockmatrix aufgefasst: I_{n-k} beschreibt die Einheitsmatrix der Größe $n - k$ und * sind die restlichen (für die Determinante uninteressanten) Einträge der Matrix A. Da $\det(C) \neq 0$ sind die Vektoren in B linear unabhängig und somit ist B eine Basis des $\mathbb{R}^n$.

Befindet sich die Untermatrix A_k nicht in den ersten k Zeilen der Matrix A, müssen wir andere Standardbasisvektoren in unserer Basis B verwenden. Und zwar ergänzen wir diese Standardbasisvektoren, die genau dort die Eins stehen haben, wo wir in A eine Zeile gestrichen haben. Dann betrachten wir wieder die Determinante. Wir vertauschen in der Determinante solange Zeilen, bis wir wieder auf die Blockmatrix

$$\begin{pmatrix} A_k & 0 \\ * & I_{n-k} \end{pmatrix}$$

kommen. Durch das Vertauschen der Zeilen ändert sich nur das Vorzeichen der Determinante, ihr Wert ist weiterhin $\neq 0$. □

5.6 Der allgemeine Fall

In diesem Abschnitt geht es darum, die Methode zum Lösen von homogenen Gleichungssystemen mit dem Kreuzprodukt auf $n \times n$ Matrizen zu verallgemeinern. Dabei wird vor allem wichtig sein, dass k linear unabhängige Vektoren immer mit **Einheitsvektoren** zu einer Basis des $\mathbb{R}^n$ ergänzt werden können.

Theorem 5.6.1. *Sei A eine $n \times n$ Matrix mit $rg(A) = k < n$. Außerdem seien $\vec{a}_1, \dots, \vec{a}_k$ die linear unabhängigen Vektoren, die den Zeilenraum von A aufspannen. Die Menge B sei eine Basis des $\mathbb{R}^n$, die durch Basisergänzung der Vektoren $\vec{a}_1, \dots, \vec{a}_k$ mit Standardbasisvektoren entsteht.*

Dann spannt die Menge

$$L := \left\{ \vec{a}_1 \times \cdots \times \vec{a}_k \times \vec{b}_{k+1} \times \cdots \times \vec{b}_{n-1} \;\middle|\; \vec{b}_{k+1}, \dots, \vec{b}_{n-1} \text{ Auswahl an } n-k-1 \text{ Einheitsvektoren aus } B \right\} \tag{5.6}$$

den Lösungsraum des Systems $A\vec{x} = \vec{0}$ auf.

Beweis. Die Vektoren in L stehen orthogonal auf $\vec{a}_1, \dots, \vec{a}_k$, es handelt sich somit um Lösungen des Systems $A\vec{x} = \vec{0}$. Wir müssen uns also nur um zwei Dinge Gedanken machen: Anzahl und lineare Unabhängigkeit der Vektoren in L.

Anzahl Da der Rang der Matrix k beträgt, gibt es $n - k$ l.u. Lösungen des Systems $A\vec{x} = \vec{0}$. Es müssten folglich $n - k$ Vektoren in L liegen. In der Menge B befinden sich $n - k$ Einheitsvektoren. Um $|L|$ zu berechnen, ermitteln wir, wie viele Möglichkeiten es gibt, $n-k-1$ Standardbasisvektoren aus den $n-k$ Standardbasisvektoren in B, auszuwählen. Wir erhalten

$$|L| = \binom{n-k}{n-k-1} = \frac{(n-k)!}{1! \cdot (n-k-1)!} = n - k.$$

Also stimmt zumindest schon einmal die Anzahl der Vektoren in L.

Unabhängigkeit Nun zum schwierigen Teil des Beweises. Wir wollen zeigen, dass die Vektoren in L linear unabhängig sind. Da $\vec{a}_1, \dots, \vec{a}_k$ linear unabhängig sind, findet sich eine $k \times k$ Untermatrix A_k mit $\det(A_k) \neq 0$. Durch Zeilenvertauschung ($\vec{a}_1, \dots, \vec{a}_k$ sollen in den ersten k Zeilen von A liegen), sowie Variablenumbenennung (d.h. Spaltenvertauschung), können wir erreichen, dass die Untermatrix A_k im linken oberen Eck von A angesiedelt ist. Oder anders formuliert: Es gilt o.B.d.A.

$$(\vec{a}_1, \dots, \vec{a}_k) = \begin{pmatrix} A_k \\ * \end{pmatrix}.$$

Die zugehörige Basis B sieht dann folgendermaßen aus:

$$B = \{\vec{a}_1, \dots, \vec{a}_k, \vec{e}_{k+1}, \dots, \vec{e}_n\}$$

Berechnen wir nun den ersten Vektor in L, nennen wir ihn $\vec{l}_1$. Dazu müssen wir zuerst $n-k-1$ Einheitsvektoren aus B auswählen. Nehmen wir $\vec{e}_{k+2}, \ldots, \vec{e}_n$. Folglich ist

$$
\begin{aligned}
(\vec{l}_1)_i &= (\vec{a}_1 \times \cdots \times \vec{a}_k \times \vec{e}_{k+2} \times \cdots \times \vec{e}_n)_i \\
&= (-1)^{i+1} \cdot \det \underbrace{\begin{pmatrix} A_k & 0 & 0 & \cdots & 0 \\ * & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & & \vdots \\ * & 0 & 1 & & \\ * & \vdots & \vdots & \ddots & 0 \\ * & 0 & 0 & \cdots & 1 \end{pmatrix}}_{i\text{-te Zeile steichen}} .
\end{aligned}
$$

Demnach weist $\vec{l}_1$ folgende Struktur auf:

$$
\vec{l}_1 = \begin{pmatrix} * \\ \pm \det(A_k) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \to \text{erste } k \text{ Einträge} \\ \\ \\ \\ \\ \end{matrix} .
$$

Die ersten k Einträge $*$ von $\vec{l}_1$ sind irgendwelche (für den Beweis unwichtige) Einträge. Wird die $(k+1)$-te Zeile gestrichen, steht links oben A_k und rechts unten die Einheitsmatrix. Das ergibt $\det(A_k)$ (mit einem Vorzeichen). Für die Berechnung des $(k+2)$-ten und den nachfolgenden Einträgen von $\vec{l}_1$ wird in der Determinante eine Zeile mit einer 1 gestrichen. Das generiert immer eine Nullspalte. Die Determinante wird $= 0$.

Nun zur Berechnung von $\vec{l}_2$. Wir wählen eine andere Auswahl an $n-k-1$ Einheitsvektoren. Nehmen wir $\vec{e}_{k+1}, \vec{e}_{k+3}, \ldots, \vec{e}_n$. Folglich ist

$$
\begin{aligned}
(\vec{l}_2)_i &= (\vec{a}_1 \times \cdots \times \vec{a}_k \times \vec{e}_{k+1} \times \vec{e}_{k+3} \times \cdots \times \vec{e}_n)_i \\
&= (-1)^{i+1} \cdot \det \underbrace{\begin{pmatrix} A_k & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & 0 & 0 & & \vdots \\ * & 0 & 1 & & \\ * & \vdots & \vdots & \ddots & 0 \\ * & 0 & 0 & \cdots & 1 \end{pmatrix}}_{i\text{-te Zeile streichen}} .
\end{aligned}
$$

Demnach weist $\vec{l}_2$ folgende Struktur auf:

$$
\vec{l}_2 = \begin{pmatrix} * \\ 0 \\ \pm \det(A_k) \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \to \text{erste } k \text{ Einträge} \\ \\ \\ \\ \\ \end{matrix} .
$$

Wir bemerken: $\det(A_k)$ ist einen Eintrag nach unten gewandert.

Fahren wir so fort, erhalten wir schlussendlich

$$L = \left\{ \begin{pmatrix} * \\ \pm\det(A_k) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} * \\ 0 \\ \pm\det(A_k) \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} * \\ 0 \\ \vdots \\ \pm\det(A_k) \\ 0 \end{pmatrix}, \begin{pmatrix} * \\ 0 \\ \vdots \\ \\ \pm\det(A_k) \end{pmatrix} \right\}.$$

Die $(n-k) \times (n-k)$ Untermatrix

$$\begin{pmatrix} \pm\det(A_k) & & \\ & \ddots & \\ & & \pm\det(A_k) \end{pmatrix}$$

weist wegen $\det(A_k) \neq 0$ vollen Rang auf. Daher sind die Vektoren in L linear unabhängig.

□

Beispiel 14. Gegeben seien die beiden l.u. Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 1 \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \\ 1 \end{pmatrix}.$$

Wir betrachten die Matrix A, deren Zeilenraum von diesen beiden Vektoren aufgespannt wird. Gesucht sind die l.u. Lösungen von $A\vec{x} = \vec{0}$.

Zuerst ergänzen wir die Vektoren $\vec{a}_1, \vec{a}_2$ mit Standardbasisvektoren zu einer Basis B des $\mathbb{R}^5$:

$$B = \{\vec{a}_1, \vec{a}_2, \vec{e}_3, \vec{e}_4, \vec{e}_5\}.$$

Jetzt berechnen wir folgende Kreuzprodukte: Man beachte, dass wir wegen der Laplace-

Entwicklung nur Determinanten von 2×2 Matrizen ermitteln müssen.

$$\vec{a}_1 \times \vec{a}_2 \times \vec{e}_4 \times \vec{e}_5 = \begin{pmatrix}1\\2\\3\\1\\1\end{pmatrix} \times \begin{pmatrix}1\\-1\\2\\1\\1\end{pmatrix} \times \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix} \times \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}7\\1\\-3\\0\\0\end{pmatrix}$$

$$\vec{a}_1 \times \vec{a}_2 \times \vec{e}_3 \times \vec{e}_5 = \begin{pmatrix}1\\2\\3\\1\\1\end{pmatrix} \times \begin{pmatrix}1\\-1\\2\\1\\1\end{pmatrix} \times \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} \times \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}-3\\0\\0\\3\\0\end{pmatrix}$$

$$\vec{a}_1 \times \vec{a}_2 \times \vec{e}_3 \times \vec{e}_4 = \begin{pmatrix}1\\2\\3\\1\\1\end{pmatrix} \times \begin{pmatrix}1\\-1\\2\\1\\1\end{pmatrix} \times \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} \times \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix} = \begin{pmatrix}3\\0\\0\\0\\-3\end{pmatrix}$$

Diese drei berechneten Vektoren spannen den Lösungsraum des Systems $A\vec{x} = \vec{0}$ auf.

5.6.1 Schnelle Kreuzproduktberechnung bei Ergänzung mit Einheitsvektoren

Ich werde nun eine kompakte Notation einführen, um Kreuzprodukte, wie sie etwa im Beispiel 14 vorkommen, effizienter und mit weniger Schreibarbeit zu berechnen. Die Idee ist, anstatt mehrere einzelne Produkte zu berechnen, den gesamten Rechenaufwand geschickt auf einmal zu erledigen.

Gegeben linear unabhängige Vektoren $\vec{a}_1, \ldots, \vec{a}_k$
Gesucht $n-k$ Vektoren, die orthogonal auf die gegebenen Vektoren stehen

Schritt 1 Ergänze $\vec{a}_1, \ldots, \vec{a}_k$ mit Einheitsvektoren zu einer Basis B. Wir nehmen wieder an

$$B = \{\vec{a}_1, \ldots, \vec{a}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n\}\,.$$

Schritt 2 Fasse die Vektoren in B als die Spalten einer Matrix auf. Nennen wir diese Matrix (passenderweise) ebenfalls B. Es gilt also

$$B = (\vec{a}_1, \ldots, \vec{a}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n) = \begin{pmatrix} A_k & 0 \\ * & I_{n-k} \end{pmatrix},$$

wobei $\det(A_k) \neq 0$.

Schritt 3 Bereite eine Ergebnismatrix X vor. Wähle $n-k-1$ verschiedene Diagonaleinträge in der Einheitsmatrix I_{n-k}. Streiche nun in B (gedanklich oder durch Abdecken) die gesamten Zeilen und Spalten, in denen sich diese Einträge befinden. Trage in der ersten

Spalte von X genau in den Zeilen eine Null ein, die in B vorher gestrichen wurden. Streiche nun (bzw. ignoriere) die Spalte in B mit der letzten verbleibenden Eins von I_{n-k}. Berechne nun das Kreuzprodukt der noch sichtbaren Spaltenvektoren von B im $\mathbb{R}^{k+1}$. Trage das Ergebnis in die freien Einträge der ersten Spalte von X ein.

Schritt 4 Wiederhole Schritt 3 mit einer anderen Auswahl an $n-k-1$ Diagonalelementen aus I_{n-k}. Benutze für die Ergebnisse die zweite Spalte (allgemein: j-te Spalte) von X. Wiederhole solange, bis alle $n-k$ Möglichkeiten $n-k-1$ Diagonalelemente aus I_{n-k} auszuwählen durchgegangen sind. In der Ergebnismatrix X stehen nun die $n-k$ Vektoren, die orthogonal auf $\vec{a}_1, \ldots, \vec{a}_k$ sind.

Hinweis: Müssen andere Standardbasisvektoren als die oben angegebenen zur Ergänzung von $\vec{a}_1, \ldots, \vec{a}_k$ herangezogen werden, kann der Algorithmus nach einer naheliegenden Anpassung ebenfalls angewendet werden.

Achtung: Die Spalten der Ergebnismatrix X können sich von den tatsächlichen Kreuzprodukten um ein Vorzeichen unterscheiden. Das ist aber weder für die Orthogonalität, noch für die Lösung von homogenen linearen Gleichungssystemen von Bedeutung.

Der obige Algorithmus wirkt auf den ersten Blick etwas sperrig. Das liegt daran, dass man solche Dinge besser mit Bildern erklären kann, als mit Worten bzw. Text. Deshalb verwenden wir nun die Methode, um nochmals Beispiel 14 zu lösen. Dort finden sich auch Abbildungen, die das Ganze viel verständlicher machen sollten.

Beispiel 15. Betrachten wir nochmal die Vektoren

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 1 \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \\ 1 \end{pmatrix}.$$

Gesucht sind 3 l.u. Vektoren, die auf diese beiden Vektoren orthogonal stehen.

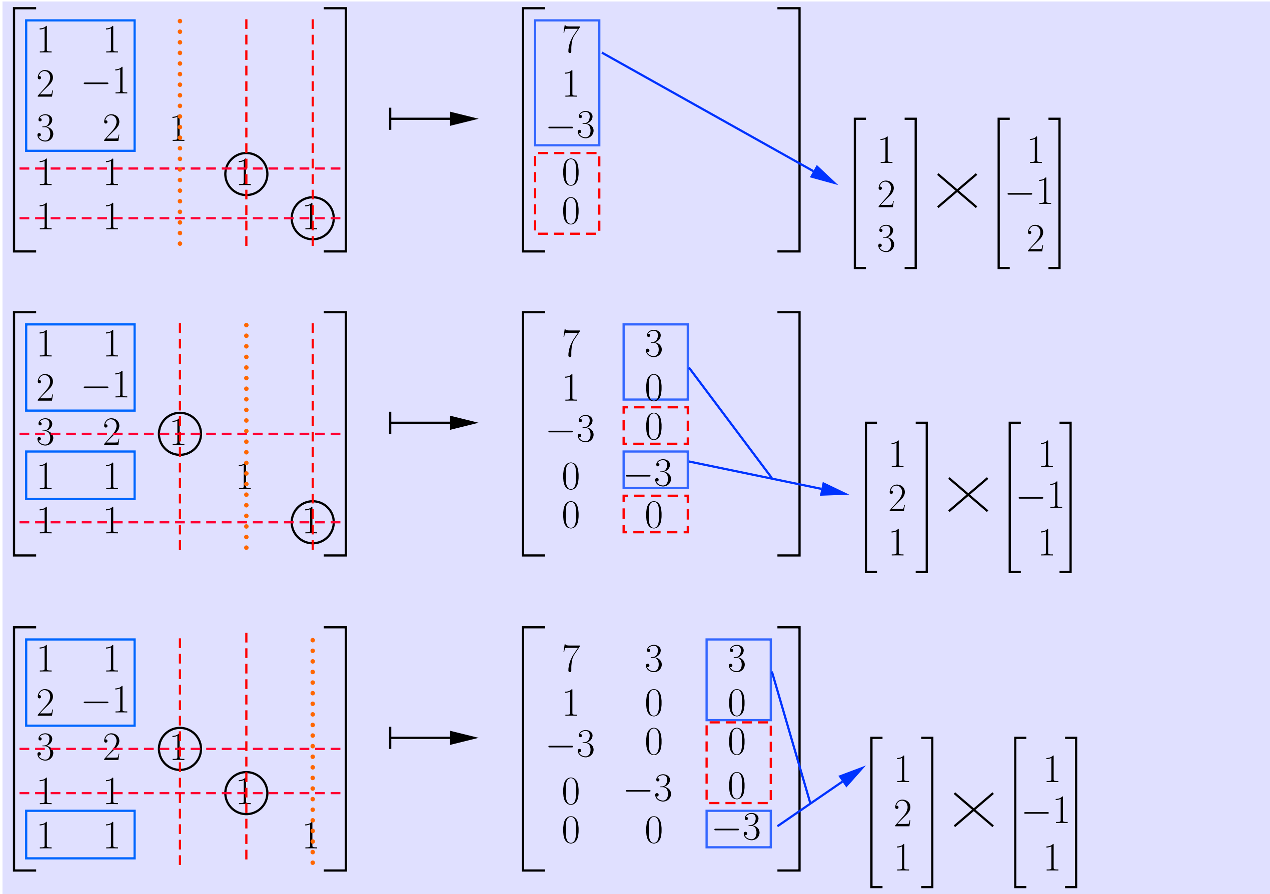
Als Basis B nehmen wir wieder

$$B = \{\vec{a}_1, \vec{a}_2, \vec{e}_3, \vec{e}_4, \vec{e}_5\}.$$

Damit wäre Schritt 1 auch schon abgehakt.
Schritt 2 ist ebenfalls sehr einfach. Für unsere Matrix B erhalten wir

$$B = \begin{pmatrix} 1 & 1 & & & \\ 2 & -1 & & & \\ 3 & 2 & 1 & & \\ 1 & 1 & & 1 & \\ 1 & 1 & & & 1 \end{pmatrix}.$$

Nun zu Schritt 3 und 4. Dazu werfe man einen Blick auf Folgendes:



Somit ist

$$X = \begin{pmatrix} 7 & 3 & 3 \\ 1 & 0 & 0 \\ -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

und die Spalten sind die gesuchten orthogonalen Vektoren.

5.7 Die Reduktionsmethode

Bisher haben wir immer angenommen, dass der Rang der Matrix A bekannt ist. Im Normalfall haben wir jedoch keinerlei Informationen über den Rang zur Verfügung. Um den Rang von A zu bestimmen, sowie die Lösung von $A\vec{x} = \vec{0}$ zu ermitteln, kann das Gleichungssystem reduziert werden. Außerdem ist diese Methode praktisch, falls die Determinanten für die Berechnung der Einträge des Kreuzproduktes zu groß werden. Durch Reduktion können prinzipiell auch höherdimensionale Gleichungssysteme mit dem Kreuzprodukt gelöst werden – falls man das möchte.

5.7.1 Die Idee

Im Grunde modifizieren wir nur unsere bekannte Methode ein bisschen. Im Unterschied zur bekannten Methode, ist der Rang der Matrix A noch nicht bekannt. Nennen wir diesen (vorerst unbekannten) Rang einfach k, das heißt $rg(A) = k$.

Ziel ist es, die l.u. Lösungen von $A\vec{x} = \vec{0}$ zu ermitteln. Dazu betrachten wir die Zeilen der Matrix A und suchen l.u. Zeilen. Da wir den Rang noch nicht kennen, wissen wir noch nicht genau, wie viele l.u. Zeilen es maximal gibt. Wir suchen uns einfach so viele l.u. Zeilen $\vec{a}_i$ der Matrix aus, wie wir finden können bzw. möchten. Ihre Anzahl sei j. Dann gilt klarerweise $j \leqslant k$. Außerdem seien diese j Zeilen o.B.d.A. die ersten j Zeilen der Matrix (ansonsten vertauschen wir einfach Zeilen bis sie es sind).

Betrachten wir nun den Unterraum

$$U := \mathcal{L}\left\{\vec{a}_1, \dots, \vec{a}_j\right\}.$$

Wir wissen bereits aus den früheren Abschnitten, dass JEDE Lösung von $A\vec{x} = \vec{0}$, orthogonal auf die Zeilen der Matrix A stehen muss. Somit auch auf unsere oben gewählten Zeilen $\vec{a}_1, \dots, \vec{a}_j$. Das heißt aber nichts anderes, als dass jede Lösung $\vec{x}$ im orthogonalen Komplement $U^{\perp}$ des Unterraumes U liegt.
Eine Basis dieses orthogonalen Komplements können wir nun mit Hilfe des Kreuzprodukts und der im letzten Abschnitt vorgestellten Methode ermitteln. Bis hierhin ist noch nichts Neues passiert. Falls nun $j = k$ ist, spannen diese Vektoren bereits den Lösungsraum auf. Eingesetzt in die restlichen Zeilen der Matrix, kommt immer $= 0$ heraus. Auch das sollte uns schon bekannt sein.

Und nun zu etwas Neuem: Falls $j < k$ ist, lösen diese berechneten Vektoren – nennen wir sie $\vec{x}_{j+1}, \dots, \vec{x}_n$ – klarerweise nicht alle die restlichen Zeilen der Matrix A. Denn $rg(A)$ ist gleich k und nicht gleich j. Was aber weiterhin gilt: Jede Lösung $\vec{x}$ liegt in $U^{\perp}$, weshalb $\vec{x}$ als Linearkombination der $\vec{x}_{j+1}, \dots, \vec{x}_n$ geschrieben werden kann. Packen wir die Vektoren $\vec{x}_{j+1}, \dots, \vec{x}_n$ spaltenweise in eine Matrix X, erfüllt also jede Lösung $\vec{x}$ von $A\vec{x} = \vec{0}$

$$\vec{x} = X\ \vec{s} \quad \text{mit irgenwelchen Skalaren } \vec{s} = (s_{j+1}, \dots, s_n)^T.$$

Diesen Vektor $\vec{s}$ können wir bestimmen, indem wir die „restliche Information“ der Matrix A nutzen. Es gilt ja

$$A\ \vec{x} = A\ (X\ \vec{s}) = AX\ \vec{s} = \vec{0}.$$

Das LGS $AX\ \vec{s} = \vec{0}$ ist das reduzierte Gleichungssystem, da aufgrund der Konstruktion die ersten j Zeilen von AX Nullzeilen darstellen. Dieses reduzierte System müssen wir nun lösen.

Sind wir fertig, setzen wir die l.u. Lösungsvektoren des Gleichungssystems $AX\ \vec{s} = \vec{0}$ in die Matrix X ein, um die l.u. Lösungsvektoren von $A\,\vec{x} = \vec{0}$ zu erhalten (l.u. deshalb, da X injektiv ist und daher l.u. Teilmengen auf l.u. Teilmengen abbildet).

5.7.2 Beispiele

Beispiel 16. Gegeben sei die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 3 & -2 \\ 3 & 2 & 5 & -2 \end{pmatrix}.$$

Gesucht sind alle Lösungen des Gleichungssystems $A\vec{x} = \vec{0}$.
Die ersten zwei Zeilen der Matrix A sind l.u., das sieht man z.B. indem man eine passende Unterdeterminante der Ordnung 2 auswählt. Wir können sie also für unsere Reduktion heranziehen. Ergänzen wir diese beiden Zeilen mit zwei Einheitsvektoren zu einer Basis und berechnen wir die Kreuzprodukte:

$$\begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ 1 & 1 & 1 & \\ 1 & -1 & & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} 0 & 2 \\ 2 & 0 \\ 0 & -2 \\ -2 & 0 \end{bmatrix} \Bigg|\Bigg| \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & -1 \\ -1 & 0 \end{bmatrix} = X.$$

Dabei sind wir folgendermaßen vorgegangen: In der Matrix links vom Pfeil rechts unten eine 1 auswählen, gesamte Zeile streichen und das Kreuzprodukt der sichtbaren Vektoren im $\mathbb{R}^3$ berechnen. Das Ergebnis rechts eintragen. Das Gleiche nochmal, nur mit der zweiten 1. Einträge, die im Ergebnis frei sind, mit Nullen füllen (die kommen in die Zeile, wo vorher die 1 gestrichen wurde …). Die Ergebnisvektoren können sich vom „echten" Kreuzprodukt um ein Vorzeichen unterscheiden (was egal ist, weil wir hier nur an orthogonalen Vektoren interessiert sind).
Jede Lösung $\vec{x}$ steht sowohl orthogonal auf die erste, wie auch auf die zweite Zeile der Matrix A. Somit gilt $\vec{x} = X\ \vec{s}$ mit irgendeinem $\vec{s}$. Um $\vec{s}$ zu bestimmen, setzen wir $\vec{x}$ ins Gleichungssystem $A\vec{x} = \vec{0}$ ein. Dabei können wir auch gleich überprüfen, ob wir die Normalenvektoren korrekt berechnet haben: Die ersten zwei Zeilen von AX müssen Nullzeilen sein!

$$A\vec{x} = AX\ \vec{s} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 4 & -2 \\ 4 & -2 \end{pmatrix}\ \vec{s} = \vec{0}.$$

Für die unteren Einträge von AX kann man einfach die Spalten von X in die letzten zwei Zeilen von A einsetzen und die Ergebnisse eintragen.

Dieses Gleichungssystem kann nun sehr schnell mit „Vertauschen und Vorzeichenwechsel“ gelöst werden. Es gilt

$$\vec{s} = t \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Den Vektor $(1, 2)$ müssen wir jetzt nur noch in X einsetzen. Wir erhalten als allgemeine Lösung von $A\vec{x} = \vec{0}$

$$\vec{x} = t \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix} \quad \text{wobei} \quad t \in \mathbb{R}.$$

Somit gilt $rg(A) = 3$.

Nun machen wir das ganze Beispiel nochmal. Dieses Mal wählen wir aber zu Beginn nicht zwei Zeilen der Matrix A, sondern nur eine.

Beispiel 17. Gegeben sei nochmals die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 3 & -2 \\ 3 & 2 & 5 & -2 \end{pmatrix}.$$

Gesucht sind alle Lösungen des Gleichungssystems $A\vec{x} = \vec{0}$.
Wir ziehen nur die erste Zeile für unsere Reduktion heran. Um die Vektoren zu bestimmen, die orthogonal auf die erste Zeile stehen, verwenden wir „Vertauschen und Vorzeichenwechsel“. Somit gilt

$$X = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Da jede Lösung $\vec{x}$ orthogonal auf die erste Zeile steht, folgt $\vec{x} = X\ \vec{s}$ für irgendein $\vec{s}$. Um $\vec{s}$ zu bestimmen, setzen wir in das Gleichungssystem $A\vec{x} = \vec{0}$ ein.

$$A\vec{x} = AX\ \vec{s} = \begin{pmatrix} 0 & 0 & 0 \\ -2 & 2 & -2 \\ 1 & 1 & -5 \\ -1 & 3 & -7 \end{pmatrix} \vec{s} = \vec{0}.$$

Um das zu lösen, verwenden wir das Kreuzprodukt. Zweite und dritte Zeile von AX sind l.u., folglich

$$\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix} = \begin{pmatrix} -4 \\ -6 \\ -2 \end{pmatrix} \mathbin{||} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

Da dieser Vektor auch die letzte Zeile von $AX\ \vec{s} = \vec{0}$ löst (im Grunde ist das noch ein, wenn auch einfacher, Reduktionsschritt), folgt

$$\vec{s} = t \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \quad \text{mit} \quad t \in \mathbb{R}.$$

Jetzt nur noch den Vektor (2, 3, 1) in die Matrix X einsetzen. Wir erhalten

$$\vec{x} = t \cdot \begin{pmatrix} -2 \\ -1 \\ 2 \\ 1 \end{pmatrix} \quad \text{mit} \quad t \in \mathbb{R}.$$

Man sieht: Beide Varianten liefern schlussendlich denselben Lösungsraum. Welche Variante einem sympathischer ist, kann jeder für sich selbst entscheiden.

5.8 Reduktionsmethode mit einer bzw. zwei Zeilen

Das ist im Grunde nichts Neues, denn wir haben diese Methoden bereits in den letzten beiden Beispielen angewendet. Deshalb werde ich an dieser Stelle nicht weiter auf das allgemeine Vorgehen zum Lösen von $A\vec{x} = \vec{0}$ eingehen, wenn man mit einer bzw. zwei Zeilen reduziert.

Sollte man dennoch interessiert sein, kann man einen Blick in meine optimierte Kurzfassung dieser Arbeit werfen. Dort wird mit illustrierenden Abbildungen und mit einfachen „Kochrezepten“ beschrieben, wie die Reduktion mit einer bzw. zwei Zeilen angewendet wird. Es wird auch gezeigt, wie man anhand eines Schemas vorgehen kann, um die Matrix X schneller aufzustellen. Und natürlich finden sich in der Kurzfassung noch viele weitere nützliche Dinge!

Für weiterführende Infos zur optimierten Kurzfassung besuche einfach meine Website unter dem Link

`Forty-Moo.github.io/Kategorien/mathe/mathe_texte/kreuzprodukt_kurz.html`

5.9 Vorgehen im Komplexen

Bisher haben wir nur Vektoren und Matrizen mit reellen Werten betrachtet. Wie müssen wir im Komplexen vorgehen? Für diejenigen, die den Abschnitt überspringen wollen, werde ich die Antwort spoilern. Wir können genauso vorgehen wie im Reellen – wir benutzen die reelle Formel für das Kreuzprodukt, kein komplexes Konjugieren.

Sei A also eine komplexe $n \times n$ Matrix mit $rg(A) = k$. Seien $\vec{a}_1, \dots, \vec{a}_k$ linear unabhängige Zeilen der Matrix A. Um den Lösungsraum des homogenen Gleichungssystems $A\vec{x} = \vec{0}$ zu ermitteln, müssen wir wieder all jene linear unabhängigen Vektoren finden, die orthogonal auf $\vec{a}_1, \dots, \vec{a}_k$ stehen, bezüglich des **reellen** Standardskalarprodukts $\cdot_{\mathbb{R}}$. Symbolisch

$$\vec{a}_1 \cdot_{\mathbb{R}} \vec{x} = 0, \quad \dots \quad , \vec{a}_k \cdot_{\mathbb{R}} \vec{x} = 0.$$

Wir haben aber die Begriffe Orthogonalität und Kreuzprodukt für das **komplexe** Standardskalarprodukt $\cdot_{\mathbb{C}}$ eingeführt! Das ist aber nicht wirklich ein Problem. Wir benutzen einfach den Zusammenhang 4.2 zwischen reellem und komplexen Standardskalarprodukt.

Wie müssen wir also vorgehen? Wir suchen zuerst die Vektoren, die orthogonal auf $\vec{a}_1, \dots, \vec{a}_k$ bezüglich $\cdot_{\mathbb{C}}$ stehen. Das heißt sie erfüllen

$$\vec{a}_1 \cdot_{\mathbb{C}} \vec{x} = 0, \quad \dots \quad , \vec{a}_k \cdot_{\mathbb{C}} \vec{x} = 0.$$

Dafür können wir das komplexe Kreuzprodukt zu Hilfe nehmen. Dann konjugieren wir diese Vektoren und erhalten $\overline{\vec{x}}$. Wegen Zusammenhang 4.2 gilt nun

$$\vec{a}_1 \cdot_{\mathbb{R}} \overline{\vec{x}} = \vec{a}_1 \cdot_{\mathbb{C}} \vec{x} = 0, \quad \dots \quad , \vec{a}_k \cdot_{\mathbb{R}} \overline{\vec{x}} = \vec{a}_k \cdot_{\mathbb{C}} \vec{x} = 0.$$

Die Vektoren $\overline{\vec{x}}$ erfüllen also $A\overline{\vec{x}} = 0$, sind somit Lösungen.

Angenommen die Vektoren $\vec{x}_{k+1}, \dots, \vec{x}_n$ sind linear unabhängige Vektoren, die orthogonal auf alle $\vec{a}_i$ stehen bezüglich $\cdot_{\mathbb{C}}$. Dann brauchen wir diese Vektoren nur komplex zu konjugieren und haben eine Basis des Lösungsraumes gefunden.

Nun wissen wir aber, dass sich das komplexe Kreuzprodukt vom reellen nur durch komplexe Konjugation unterscheidet. Berechnen wir also $\vec{x}_{k+1}, \dots, \vec{x}_n$ mit Hilfe des komplexen Kreuzprodukts und konjugieren diese Vektoren anschließend, damit sie Lösungen von $A\vec{x} = 0$ sind, kommen wir wieder auf die reelle Berechnungsformel für das Kreuzprodukt!

In den früheren Sätzen und Resultaten können wir also überall dort, wo $\mathbb{R}^n$ steht, auch $\mathbb{C}^n$ schreiben. Dann müssen die Lösungsvektoren in den Sätzen, die nun durch das komplexe Kreuzprodukt berechnet werden, noch konjugiert werden. Ansonsten ändert sich an den Sätzen nichts – und auch diese Änderungen sind prinzipiell theoretischer Natur. In der Praxis ändert sich gar nichts wie wir in folgendem Beispiel sehen werden:

Beispiel 18. Gegeben sei die komplexe Matrix

$$\begin{pmatrix} \mathrm{i}+1 & \mathrm{i} & 2 \\ \mathrm{i} & 1 & \mathrm{i} \\ 1-\mathrm{i} & \mathrm{i}-2 & 2-2\mathrm{i} \end{pmatrix}.$$

Wir interessieren uns an der Lösung von $A\vec{x} = 0$.
Wir sehen, dass die ersten zwei Zeilen der Matrix linear unabhängig sind. Dafür können wir z. B. das Resultat mit der Unterdeterminante verwenden (siehe 5.5.1). Wir berechnen einfach einmal das reelle Kreuzprodukt dieser zwei Zeilen, kennzeichnen wir es mit $\times_{\mathbb{R}}$

$$\begin{pmatrix} \mathrm{i}+1 \\ \mathrm{i} \\ 2 \end{pmatrix} \times_{\mathbb{R}} \begin{pmatrix} \mathrm{i} \\ 1 \\ \mathrm{i} \end{pmatrix} = \begin{pmatrix} -3 \\ \mathrm{i}+1 \\ \mathrm{i}+2 \end{pmatrix}$$

Um abzuschätzen, ob die Matrix Rang 2 oder 3 besitzt, machen wir die Probe mit der letzten Zeile:

$$(1-\mathrm{i}) \cdot (-3) + (\mathrm{i}-2) \cdot (\mathrm{i}+1) + (2-2\mathrm{i}) \cdot (\mathrm{i}+2) = 0.$$

Somit hat die Matrix Rang 2 und obiger Kreuzproduktvektor spannt den Lösungsraum auf.

Man beachte, dass das komplexe Kreuzprodukt zu keiner Lösung des Problems führt, denn

$$\begin{pmatrix} \mathrm{i}+1 \\ \mathrm{i} \\ 2 \end{pmatrix} \times_{\mathbb{C}} \begin{pmatrix} \mathrm{i} \\ 1 \\ \mathrm{i} \end{pmatrix} = \overline{\begin{pmatrix} -3 \\ \mathrm{i}+1 \\ \mathrm{i}+2 \end{pmatrix}} = \begin{pmatrix} -3 \\ -\mathrm{i}+1 \\ -\mathrm{i}+2 \end{pmatrix}.$$

Setzen wir das in die zweite Zeile ein, folgt

$$\mathrm{i} \cdot (-3) + 1 \cdot (-\mathrm{i}+1) + \mathrm{i} \cdot (-\mathrm{i}+2) = 2 - 2\mathrm{i} \neq 0.$$

6 Das Lösen inhomogener LGS mit dem Kreuzprodukt

In diesem Abschnitt wollen wir die Lösungsmenge des inhomogenen LGS $A\vec{x} = \vec{b}$ ermitteln, wobei A eine komplexe oder reelle Matrix und $\vec{b} \neq \vec{0}$ einen komplexen oder reellen Vektor beschreibe. Auch hier kann das Kreuzprodukt zu Hilfe genommen werden. Wir beschränken uns wieder auf $n \times n$ Matrizen, die Theorie sollte aber auch auf nicht quadratische Matrizen anwendbar sein.

6.1 Grundsätzliches - Affine Abbildungen

Beginnen wir damit die Bedingung $A\vec{x} = \vec{b}$ umzuschreiben: Wir suchen all jene $\vec{x}$, sodass

$$\varphi(\vec{x}) := A\vec{x} - \vec{b} = \vec{0}.$$

Wäre φ nun eine lineare Abbildung, hätte wir keine Probleme. Wir könnten so vorgehen wie im vorigen Abschnitt bei homogenen System.

Die Abbildung φ ist aber nicht linear, wovon man sich leicht überzeugen kann (z.B. wird der Nullvektor von φ nicht wieder auf den Nullvektor abgebildet). Es handelt sich bei φ um eine **affine Abbildung**. Das ist allgemein eine Abbildung, die aus einem linearen Anteil und einem Translationsanteil besteht. Bei obiger Abbildung φ ist $A\vec{x}$ der lineare Anteil. Dann wird das Ganze noch um den fixen Vektor $-\vec{b} \neq \vec{0}$ verschoben – somit ist $-\vec{b}$ der Translationsanteil.

Eigentlich ist obige Fragestellung kein Problem der Linearen Algebra, es ist ein Problem der **Affinen Geometrie**. In der linearen Algebra waren Unterräume von zentraler Wichtigkeit. In der affinen Geometrie betrachten wir hingegen **affine Unterräume**. Das ist ein linearer Unterraum um einen Vektor $\vec{v}$ verschoben. Ein affiner Unterraum H ist also von der Form $H = \vec{v} + U$, wobei U ein linearer Unterraum ist.

Am besten man stellt sich das im $\mathbb{R}^3$ vor: Angenommen U ist ein 2-dimensionaler linearer Unterraum des $\mathbb{R}^3$. Dieser Unterraum beschreibt eine Ebene durch den Ursprung. Wir können nun jeden Punkt auf dieser Ebene hernehmen und ihn um 1 nach oben verschieben. Die Menge H die dabei entsteht ist klarerweise wieder eine Ebene, für die $H = \vec{e}_3 + U$ gilt. Die Ebene H muss nun aber nicht mehr unbedingt durch den Ursprung verlaufen. Der lineare Unterraum U wurde um den Vektor $\vec{e}_3$ verschoben. Die Menge H beschreibt einen affinen Unterraum.

Etwas ungenau aber einprägsam: **„Affin“ ist immer etwas Lineares plus eine Translation (= Verschiebung).**

6.2 Grundsätzliches - Lösbarkeit

Bei homogenen System $A\vec{x} = \vec{0}$ haben wir uns über die Lösbarkeit keine Gedanken gemacht. Der Nullvektor war ja immer triviale Lösung dieses Systems. Somit hat es auch immer mindestens eine Lösung gegeben.

Bei inhomogenen Systemen $A\vec{x} = \vec{b}$ sieht es da schon etwas anders aus. Wir wissen nicht, ob der Vektor $\vec{b}$ von der linearen Abbildung A überhaupt angenommen wird. Es könnte also durchaus passieren, dass dieses System gar keine Lösung besitzt.

Die Matrix $(A \mid \vec{b}\,)$ beschreibe im Folgenden die um den Vektor $\vec{b}$ erweiterte Matrix. Das heißt, wenn A eine Matrix ist und $\vec{b}$ ein Vektor, der genau so viele Einträge hat, wie A Zeilen, dann ist

$$(A \mid \vec{b}\,) := (A \quad \vec{b}\,).$$

Wir fügen also zu A einfach eine weitere Spalte mit den Einträgen von $\vec{b}$ hinzu.

Folgendes (bekanntes) Resultat gibt Auskunft über die Lösbarkeit von inhomogenen LGS und die Struktur der Lösung.

Satz 6.2.1. *Das inhomogene LGS $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn*

$$rg(A) = rg((\ A \mid \vec{b}\ )). \tag{6.1}$$

Im Falle der Lösbarkeit ist der Lösungsraum L des Systems $A\vec{x} = \vec{b}$ gegeben durch

$$L = \vec{x}_p + U_{hom} \tag{6.2}$$

Der Vektor $\vec{x}_p$ beschreibt hierbei eine beliebige Partikulärlösung. Das ist ein Vektor der $A\vec{x}_p = \vec{b}$ erfüllt.

Und U_{hom} beschreibt den Lösungsraum des homogenen Systems $A\vec{x} = \vec{0}$.

Die Menge L beschreibt also einen affinen Unterraum. Der lineare Lösungsraum U_{hom} des homogenen LGS $A\vec{x} = 0$ wird um (irgendeine) Partikulärlösung $\vec{x}_p$ von $A\vec{x} = \vec{b}$ verschoben.

Im Falle der Lösbarkeit des System $A\vec{x} = \vec{b}$ gehen wir wie folgt vor: Wir ermitteln zuerst den Lösungsraum U_{hom} des homogenen Systems $A\vec{x} = \vec{0}$. Wie man das mit dem Kreuzprodukt macht haben wir schon ausführlich behandelt.
Dann ermitteln wir einen Vektor $\vec{x}_p$, der $A\vec{x}_p = \vec{b}$ erfüllt. Hier wissen wir noch nicht, wie wir diesen Vektor mit Hilfe des Kreuzprodukts berechnen sollen.

Unser Ziel im nächsten Abschnitt wird es somit sein, einen Weg zu finden, mit dem Kreuzprodukt einen Vektor $\vec{x}_p$ zu ermitteln, der $A\vec{x}_p = \vec{b}$ erfüllt.

6.3 Das Ermitteln einer Partikulärlösung mit dem Kreuzprodukt

Wir unterscheiden zwei Fälle: Der Fall, wo A vollen Rang hat und der Fall, wo A nicht vollen Rang hat.

6.3.1 Vorgehen, falls A vollen Rang besitzt

Falls A vollen Rang besitzt, ist das System $A\vec{x} = \vec{b}$ immer lösbar. Mit Hilfe des folgenden Resultats können wir eine Partikulärlösung (hier: DIE Lösung) des Problems $A\vec{x} = \vec{b}$ ermitteln:

Satz 6.3.1. *Sei A eine reelle oder komplexe $n \times n$ Matrix mit $rg(A) = n$. Sei $\vec{b}_1 \neq \vec{0}$. Weiters seien*

$$U = \mathcal{L}\ \{\vec{b}_1\} \quad \text{und} \quad U^\perp = \mathcal{L}\ \{\vec{b}_2^\perp, \ldots, \vec{b}_n^\perp\}.$$

Für jeden Vektor $\vec{x}_p$, der $A\vec{x}_p = \vec{b}_1$ erfüllt, gilt

$$\vec{x}_p \in \mathcal{L}\ \left\{A^*\vec{b}_2^\perp \times \cdots \times A^*\vec{b}_n^\perp\right\}, \tag{6.3}$$

wobei $A^* := \overline{A}^T$ (falls A nur reelle Einträge hat, dann $A^* = A^T$).

Die Partikulärlösung $\vec{x}_p$ ist somit immer (irgend)ein skalares Vielfaches des Vektors

$$A^*\vec{b}_2^\perp \times \cdots \times A^*\vec{b}_n^\perp.$$

Beweis. Sei $\vec{x}_p$ eine Partikulärlösung des Systems $A\vec{x} = \vec{b}_1$, das heißt $A\vec{x}_p = \vec{b}_1$. Es gilt

$$\vec{b}_1 \cdot \vec{b}_2^\perp = 0, \quad \ldots \quad, \vec{b}_1 \cdot \vec{b}_n^\perp = 0$$

wegen der Orthogonalität.
Somit folgt

$$(A\vec{x}_p) \cdot \vec{b}_i^\perp = (A\vec{x}_p)^T * \overline{\vec{b}_i^\perp} = \vec{x}_p^T * A^T * \overline{\vec{b}_i^\perp} = \vec{x}_p^T * \overline{\overline{A}^T * \vec{b}_i^\perp} = \vec{x}_p \cdot (A^*\vec{b}_i^\perp)$$
$$= \vec{b}_1 \cdot \vec{b}_i^\perp = 0 \quad \text{für} \quad i = 2, \ldots, n,$$

wobei $*$ das ganz normale Matrizenprodukt ist.
Die Matrix A^* beschreibt einen Isomorphismus, da $rg(A^*) = rg(A) = n$. Da Isomorphismen linear unabhängige Teilmengen wieder auf linear unabhängige Teilmengen abbilden, ist die Menge

$$M := \left\{A^*\vec{b}_2^\perp, \ldots, A^*\vec{b}_n^\perp\right\}$$

linear unabhängig.
Die Bedingung $\vec{x}_p \cdot (A^*\vec{b}_i^\perp) = 0$ bedeutet, dass $\vec{x}_p \in \mathcal{L}\{M\}^\perp$. Außerdem ist $\mathcal{L}\{M\}^\perp$ 1-dimensional, das heißt dieser Raum wird bereits von einem Vektor $\vec{0} \neq \vec{n} \in \mathcal{L}\{M\}^\perp$ aufgespannt. Das Kreuzprodukt der Vektoren $A^*\vec{b}_i^\perp$, $i = 2, \ldots, n$, liegt in $\mathcal{L}\{M\}^\perp$ und ist $\neq \vec{0}$, da die involvierten Vektoren l. u. sind. □

Algorithmisches Vorgehen bei vollem Rang

Gegeben: reelle oder komplexe $n \times n$ Matrix A mit $rg(A) = n$; $\vec{b_1} \neq \vec{0}$
Ziel: einen (hier sogar DEN) Vektor $\vec{x}_p$ finden, der $A\vec{x}_p = \vec{b_1}$ erfüllt.

Schritt 1 Ermittle $n-1$ Vektoren $\vec{b}_2^{\perp}, \ldots, \vec{b}_n^{\perp}$, die orthogonal auf $\vec{b}_1$ stehen. Dazu verwende man Varianten von „Vertauschen und Vorzeichenwechsel“ der Einträge von $\vec{b}_1$.

Schritt 2 Berechne $A^*\vec{b}_2^{\perp}, \ldots, A^*\vec{b}_n^{\perp}$. Um die Einträge von $A^*\vec{b}_i^{\perp}$ zu berechnen, kann man z.B. einfach das reelle Skalarprodukt der Vektoren $\vec{b}_i^{\perp}$ mit den Spalten von $\overline{A}$ ausrechnen.

Schritt 3 Berechne das Kreuzprodukt der in Schritt 2 berechneten Vektoren:

$$A^*\vec{b}_2^{\perp} \times \cdots \times A^*\vec{b}_n^{\perp}$$

Achtung: Sind komplexe Werte involviert, muss das komplexe Kreuzprodukt verwendet werden (Konjugieren!).

Schritt 4 Die Partikulärlösung $\vec{x}_p$ ist ein skalares Vielfaches des in Schritt 4 berechneten Kreuzproduktsvektors, das heißt

$$\vec{x}_p = t \cdot \left(A^*\vec{b}_2^{\perp} \times \cdots \times A^*\vec{b}_n^{\perp} \right) \quad \text{wobei} \quad t \in \mathbb{R}/\mathbb{C}.$$

Schritt 5 Den freien Parameter t ermittelt man, indem man $\vec{x}_p$ in eine Zeile des Gleichungssystems $A\vec{x} = \vec{b}_1$ einsetzt und nach t auflöst.

Hinweis: Da $rg(A) = n$ ist, ist A ein Isomorphismus und deshalb ist $\vec{x}_p$ die eindeutige Lösung des inhomogenen LGS. Das heißt $\vec{x}_p$ ist die einzige Lösung des Gleichungssystems $A\vec{x} = \vec{b}_1$ im Falle von A quadratisch und $rg(A) = n$.

Beispiel 19. Betrachten wir das System

$$A\vec{x} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 1 & 1 \end{pmatrix} \vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \vec{b}_1.$$

Wir suchen alle Lösungen dieses Systems.
Da die Determinante von A nicht Null ist, hat A vollen Rang und es gibt nur eine Lösung $\vec{x}_p$. Mit Hilfe von „Vertauschen und Vorzeichenwechsel“ ermitteln wir die beiden Normalenvektoren zu $\vec{b}_1$:

$$\vec{b}_2^{\perp} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \qquad \vec{b}_3^{\perp} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

Nun erhalten wir die beiden Vektoren

$$A^T\vec{b}_2^{\perp} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \qquad A^T\vec{b}_3^{\perp} = \begin{pmatrix} -1 \\ -4 \\ -5 \end{pmatrix}.$$

Wir berechnen das Kreuzprodukt dieser beiden Vektoren

$$\begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \times \begin{pmatrix} -1 \\ -4 \\ -5 \end{pmatrix} = \begin{pmatrix} -3 \\ 12 \\ -9 \end{pmatrix}.$$

Und somit ist

$$\vec{x}_p = t \cdot \begin{pmatrix} -3 \\ 12 \\ -9 \end{pmatrix} \quad \text{für ein} \quad t \in \mathbb{R}.$$

Setzen wir $\vec{x}_p$ in die letzte Zeile des inhomogenen LGS ein, erhalten wir

$$3 \cdot (-3t) + 1 \cdot 12t + 1 \cdot (-9t) = 1$$

$$-6t = 1 \quad \Rightarrow \quad t = -\frac{1}{6}$$

und wir haben die eindeutig bestimmte Lösung $\vec{x}_p$ des Systems ermittelt!

Beispiel 20. Betrachten wir nun

$$A = \begin{pmatrix} \mathrm{i} & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 2\mathrm{i} & \mathrm{i} \\ 1 & -2 & 1 & 1 \end{pmatrix} \quad \text{und} \quad \vec{b}_1 = \begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ 1 \\ 1 \end{pmatrix}$$

Die Determinante von A kann schnell berechnet werden, da sie Blockgestalt besitzt. So sieht man leicht, dass die Matrix vollen Rang besitzt. Wir berechnen die orthogonalen Vektoren zu $\vec{b}_1$ mit „Vertauschen und Vorzeichenwechsel“. Wir dürfen dabei nicht auf das Konjugieren vergessen, da wir mit dem komplexen Skalarprodukt arbeiten.

$$\vec{b}_2^{\perp} = \begin{pmatrix} \mathrm{i} \\ -\mathrm{i} \\ 0 \\ 0 \end{pmatrix} \quad \vec{b}_3^{\perp} = \begin{pmatrix} 0 \\ -1 \\ -\mathrm{i} \\ 0 \end{pmatrix} \quad \vec{b}_4^{\perp} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

Wir erhalten

$$A^*\vec{b}_2^{\perp} = \begin{pmatrix} 1-\mathrm{i} \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad A^*\vec{b}_3^{\perp} = \begin{pmatrix} -1-3\mathrm{i} \\ -1-2\mathrm{i} \\ -2 \\ -1 \end{pmatrix} \quad A^*\vec{b}_4^{\perp} = \begin{pmatrix} -2 \\ -4 \\ 2\mathrm{i}+1 \\ \mathrm{i}+1 \end{pmatrix}$$

und berechnen

$$\begin{pmatrix} 1-\mathrm{i} \\ 0 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1-3\mathrm{i} \\ -1-2\mathrm{i} \\ -2 \\ -1 \end{pmatrix} \times \begin{pmatrix} -2 \\ -4 \\ 2\mathrm{i}+1 \\ \mathrm{i}+1 \end{pmatrix} = \overline{\begin{pmatrix} +0 \\ -((1-\mathrm{i})\cdot(-1)) \\ +(-6) \\ -(\mathrm{i}-9) \end{pmatrix}} = \begin{pmatrix} 0 \\ 1+\mathrm{i} \\ -6 \\ 9+\mathrm{i} \end{pmatrix}$$

Die Lösung $\vec{x}_p$ ist skalares Vielfaches dieses Vektors, das heißt

$$\vec{x}_p = t \cdot \begin{pmatrix} 0 \\ 1+\mathrm{i} \\ -6 \\ 9+\mathrm{i} \end{pmatrix} \quad \text{für ein} \quad t \in \mathbb{C}.$$

Den freien Parameter t ermitteln wir, indem wir z.B. in die zweite Zeile des Gleichungssystems einsetzen:

$$t \cdot (\mathrm{i}+1) = \mathrm{i} \quad \Rightarrow \quad t = \frac{\mathrm{i}}{\mathrm{i}+1} = \frac{1}{2}(1+\mathrm{i}).$$

Und somit

$$\vec{x}_p = \begin{pmatrix} 0 \\ \mathrm{i} \\ -3-3\mathrm{i} \\ 4+5\mathrm{i} \end{pmatrix}.$$

Alternatives Vorgehen bei vollem Rang - Reduktion

Will man keine größeren Unterdeterminanten berechnen oder macht man zu viele Rechenfehler beim Kopfrechnen, kann das Gleichungssystem „reduziert“ werden. Wirklich nutzen tut das nur etwas, wenn die Matrix größer als 3×3 ist. Ich werde eine mögliche Herangehensweise an einem Beispiel demonstrieren, da es sich im Prinzip nur um eine Variante der schon bekannten Methode handelt.

Beispiel 21. Seien R_1, R_2, R_3, R_4, R_i und E positive, reelle Zahlen. Folgendes inhomogenes LGS muss gelöst werden, um die Stromstärken $x_1, \ldots, x_4$ in einem elektrischen Netzwerk zu bestimmen. Da es sich um ein physikalisches Beispiel handelt, erwarten wir eine eindeutige Lösung.

$$\begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ R_1 & R_2 & 0 & R_4 + R_i \\ 0 & R_2 & -R_3 & 0 \end{pmatrix} \cdot \vec{x} = \begin{pmatrix} 0 \\ 0 \\ E \\ 0 \end{pmatrix}$$

Wie immer beginnen wir damit, die orthogonalen Vektoren zu $\vec{b}_1$ zu bestimmen. Anstatt alle drei anzugeben, ermitteln wir aber dieses Mal nur zwei:

$$\vec{b}_2^{\perp} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \vec{b}_3^{\perp} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Somit erhalten wir

$$A^T \vec{b}_2^{\perp} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \quad A^T \vec{b}_3^{\perp} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

Nun wissen wir, dass die Lösung $\vec{x}_p$ des Problems im orthogonalen Komplement des von diesen zwei Vektoren aufgespannten Unterraums liegt. Berechnen wir dieses orthogonale Komplement mit Hilfe des Kreuzprodukts. Hierbei brauchen wir nur 2×2 Unterdeterminanten berechnen:

$$\begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} +(-1) \\ -1 \\ +0 \\ -1 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} +0 \\ -1 \\ +1 \\ -0 \end{pmatrix}$$

Somit gilt

$$\vec{x}_p = s_1 \cdot \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \end{pmatrix} + s_2 \cdot \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} = X \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \quad \text{für} \quad s_1, s_2 \in \mathbb{R}.$$

Nun setzen wir $\vec{x}_p$ in das Gleichungssystem ein, um die freien Parameter zu ermitteln. Die

Matrix AX hat dabei Rang 2, da A vollen Rang besitzt.

$$A\vec{x}_p = AX\vec{s} = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ R_1 & R_2 & 0 & R_4+R_i \\ 0 & R_2 & -R_3 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -1 & -1 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} \vec{s} = \begin{pmatrix} 0 \\ 0 \\ E \\ 0 \end{pmatrix}$$

Wir erhalten

$$AX\vec{s} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -(R_1+R_2+R_4+R_i) & -R_2 \\ -R_2 & -(R_2+R_3) \end{pmatrix} \vec{s} = \begin{pmatrix} 0 \\ 0 \\ E \\ 0 \end{pmatrix}.$$

Diese Lösung ermitteln wir, indem wir die Methode nochmal auf die letzten zwei Zeilen anwenden:

$$\tilde{A} = \begin{pmatrix} -(R_1+R_2+R_4+R_i) & -R_2 \\ -R_2 & -(R_2+R_3) \end{pmatrix} \quad \tilde{\vec{b}} = \begin{pmatrix} E \\ 0 \end{pmatrix} \quad \tilde{\vec{b}}^{\perp} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \tilde{A}^T \tilde{\vec{b}}^{\perp} = \begin{pmatrix} -R_2 \\ -(R_2+R_3) \end{pmatrix}$$

Und somit

$$\vec{s} = t \cdot \begin{pmatrix} R_2+R_3 \\ -R_2 \end{pmatrix} \quad \text{für ein} \quad t \in \mathbb{R}.$$

Einsetzen in die erste Zeile liefert nun

$$t \cdot \left(-(R_2+R_3)\cdot(R_{ges}-R_3) + R_2^2\right) = E \quad \Rightarrow \quad t = \frac{E}{-(R_2+R_3)\cdot(R_{ges}-R_3)+R_2^2},$$

wobei $R_{ges} = R_1 + R_2 + R_3 + R_4 + R_i$.
Somit können wir $\vec{s}$ und anschließend $\vec{x}_p$ berechnen und haben das Problem gelöst.

Wie man sieht, ist das etwas mehr Schreibarbeit, führt aber ebenfalls zum Ziel. Jeder kann selbst entscheiden, ob er die Methode lieber mehrfach mit einfacheren Rechnungen oder nur einmal, dafür aber mit größeren Determinanten und mehr Kopfrechnen, durchführen möchte.

6.3.2 Vorgehen bei rg(A)<n

Betrachten wir nun komplexe oder reelle $n \times n$ Matrizen, die nicht mehr vollen Rang haben. Im Unterschied zu Matrizen mit vollem Rang müssen die Vektoren $A^*\vec{b}_2^\perp, \dots, A^*\vec{b}_n^\perp$ nicht unbedingt linear unabhängig sein, da die Abbildung nicht mehr injektiv ist.
Wir könnten nun so vorgehen wie bei dem Beispiel mit der Reduktion. Wir picken uns nur die l. u. Vektoren heraus und berechnen das orthogonale Komplement. Die Lösung liegt dann in diesem orthogonalen Komplement, das nun nicht notwendigerweise nur von einem Vektor aufgespannt wird. Die freien Parameter bestimmen wir dann durch Einsetzen der Partikulärlösung in das Gleichungssystem.
Wir können aber auch etwas anders vorgehen. Folgendes Resultat zeigt, wie:

Satz 6.3.2. *Sei A eine reelle oder komplexe $n \times n$ Matrix mit $rg(A) = k < n$. Außerdem sei das das System $A\vec{x} = \vec{b}$ mit $\vec{b} \neq \vec{0}$ lösbar, das heißt $rg(A) = rg(A \mid \vec{b})$.*

Dann gibt es eine $k \times k$ Untermatrix A_k in A mit vollem Rang. Das Teilsystem $A_k\vec{x} = \vec{b}_k$ hat dann eine eindeutige Lösung $\vec{x}_p^{\,k}$ mit k- Einträgen.

Sei nun $\vec{x_p}$ der Vektor mit n- Einträgen, der durch $\vec{x}_p^{\,k}$ entsteht, indem wir ihn an den passenden Variablenstellen mit Nullen füllen. Dieser Vektor $\vec{x}_p$ ist dann eine Partikulärlösung des System $A\vec{x} = \vec{b}$.

Beweis. Zuerst wählen wir k linear unabhängige Spalten in der Matrix A aus. Nach Resultat 5.5.1 gibt es innerhalb dieser k Spaltenvektoren eine $k \times k$ Matrix A_k mit vollem Rang. Durch Spaltenvertauschungen (Achtung: Variablenposition verändert sich) und Zeilenvertauschungen, bringen wir die Matrix in folgende Form

$$A = \begin{pmatrix} A_k & * \\ * & * \end{pmatrix}.$$

Anstatt einen Vektor $\vec{x}_p$ zu finden, der $A\vec{x}_p = \vec{b}$ erfüllt, können wir genauso gut einen Vektor $\vec{x} = \binom{\vec{x}_p}{-1}$ finden, der

$$(A \mid \vec{b}\,)\ \vec{x} = \vec{0}$$

erfüllt. Oder etwas anders formuliert:

$$\begin{pmatrix} A_k & * & \vec{b}_k \\ * & * & \vec{b}_{rest} \end{pmatrix} \vec{x} = \vec{0}.$$

Wir müssen also zeigen, dass

$$\vec{x} = \begin{pmatrix} \vec{x}_p^{\,k} & 0 & \dots & 0 & -1 \end{pmatrix}$$

obige Gleichung erfüllt.
Da $A_k\ \vec{x}_p^{\,k} = \vec{b}_k$ ist die Gleichheit schon einmal für die ersten k-Zeilen erfüllt. Die ersten k Zeilen (und Spalten) der Matrix $(A \mid \vec{b})$ sind linear unabhängig, da $\det(A_k) \neq 0$. Außerdem gilt ja $rg(A) = rg(A \mid \vec{b}) = k$. Somit spannen die ersten k Zeilen bereits den gesamten Zeilenraum der Matrix $(A \mid \vec{b})$ auf.
Damit ist die Gleichheit auch für die restlichen Zeilen von $A \mid \vec{b}$ erfüllt und $\vec{x}$ ist eine Partikulärlösung. $\square$

Im nachfolgenden Beispiel können wir alles anwenden, was wir bisher gelernt haben. Außerdem sehen wir dabei gleich, wie man die Methode bei nicht quadratischen Matrizen anwendet:

Beispiel 22. Man betrachte folgendes LGS

$$\begin{pmatrix} 0 & 0 & 2 & 1 & 2 \\ 3 & 1 & 2 & 1 & 2 \\ 3 & 1 & 6 & 3 & 6 \end{pmatrix} \vec{x} = \begin{pmatrix} 5 \\ 9 \\ 19 \end{pmatrix}$$

Wir sollen zuerst den Lösungsraum des homogenen LGS $A\vec{x} = \vec{0}$ ermitteln. Danach sollen wir entscheiden, ob obiges inhomogenes LGS lösbar ist und, wenn ja, den Lösungsraum angeben. Die ersten zwei Zeilen der Matrix sind l.u., ergänzen wir sie also zu einer Basis des $\mathbb{R}^5$:

$$B = \{\vec{a}_1, \vec{a}_2, \vec{e}_1, \vec{e}_3, \vec{e}_5\}$$

Nun berechnen wir die drei Kreuzprodukte. Dabei brauchen wir nur 2×2 Unterdeterminanten berechnen:

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

$$\begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} +0 \\ -0 \\ +(-1) \\ -(-2) \\ +0 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} +(-1) \\ -(-3) \\ +0 \\ -0 \\ +0 \end{pmatrix}$$

Die Vorzeichen hätte man wie immer auch erraten können, damit man etwas schneller ist.
Da diese 3 Vektoren auch die letzte Zeile von $A\vec{x} = \vec{0}$ lösen, spannen sie den homogenen Lösungsraum auf. Damit hat die Matrix Rang 2.
Kümmern wir uns nun um die Partikulärlösung. Dazu betrachten wir die Untermatrix

$$A_k = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad A_k\vec{x} = \begin{pmatrix} 5 \\ 9 \end{pmatrix}.$$

Es gilt

$$A_k^T \begin{pmatrix} -9 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \end{pmatrix} \quad \text{und somit} \quad \vec{x}_p{}^k = t \cdot \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

Einsetzen in die erste Zeile liefert $t = 1$. Da

$$\vec{x}_p = \begin{pmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{pmatrix}$$

auch die letzte Zeile von $A\vec{x} = \vec{b}$ löst, ist es eine Partikulärlösung. Das System ist somit lösbar.

Beispiel 23. Wir betrachten das LGS

$$\begin{pmatrix} c & -1 & 0 \\ -1 & c & -1 \\ 0 & -1 & c \end{pmatrix} \vec{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

wobei c irgendein reeller Wert ist. Wir wollen wieder zuerst die Lösungen des homogenen Systems ermitteln. Dann sollen wir entscheiden, für welche Parameter c das inhomogene System lösbar bzw. unlösbar ist.
Die zweite und dritte Zeile der Matrix sind immer l.u., das kann man über eine geeignete Unterdeterminante sehen. Berechnen wir das Kreuzprodukt dieser Zeilen:

$$\begin{pmatrix} -1 \\ c \\ -1 \end{pmatrix} \times \begin{pmatrix} 0 \\ -1 \\ c \end{pmatrix} = \begin{pmatrix} c^2 - 1 \\ c \\ 1 \end{pmatrix}$$

Setzen wir diesen Vektor nun in die erste Zeile von $A\vec{x} = \vec{0}$ ein, erhalten wir

$$c^3 - 2c = c \cdot (c^2 - 2) = 0 \quad \Rightarrow c \in \{0, \sqrt{2}, -\sqrt{2}\} =: C$$

Liegt der Parameter in der Menge C, hat die Matrix Rang 2 und obiger Kreuzproduktvektor spannt den Lösungsraum auf. Ansonsten hat die Matrix Rang 3 und $\vec{0}$ ist die einzige Lösung des System $A\vec{x} = \vec{0}$.
Betrachten wir nun das inhomogene System. Angenommen $c \in C$. Nehmen wir

$$A_k = \begin{pmatrix} -1 & 0 \\ c & -1 \end{pmatrix} \quad \text{daher} \quad A_k\vec{e}_2 = \begin{pmatrix} c \\ -1 \end{pmatrix}$$

und $\vec{x}_p{}^k = t \cdot \binom{1}{c}$ und $t = -1$. Der Vektor $\vec{x}_p = (0, -1, -c)$ erfüllt die letzte Zeile, falls $c \in \{-1, 1\}$. Somit ist das System für $c \in C$ nicht lösbar.
Liegt c nicht in der Menge C, dann hat die Matrix vollen Rang. Die Lösung ist von der Form

$$\vec{x}_p = t \cdot \begin{pmatrix} c^2 - 1 \\ c \\ 1 \end{pmatrix}.$$

Indem wir das in die erste Zeile einsetzen, erhalten wir

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

Im nachfolgenden Beispiel sehen wir noch, was passiert, wenn wir den Rang einer Matrix falsch einschätzen.

Beispiel 24. Betrachten wir

$$\begin{pmatrix} 1 & 1 & -2 & 1 & 3 \\ 2 & -1 & 2 & 2 & 6 \\ 3 & 2 & -4 & -3 & -9 \end{pmatrix} \vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Wir sehen sofort, dass die ersten zwei Zeilen linear unabhängig sind. Nehmen wir einfach einmal an, dass die Matrix Rang 2 hat und berechnen wir folgende Kreuzprodukte, um die Lösung des homogenen Systems zu ermitteln.
Um Zeit und Platz zu sparen, schreibe ich das Ganze in Matrixform. Rechts stehen die gelösten Kreuzprodukte, die sich möglicherweise um ein Vorzeichen von der tatsächlichen Lösung unterscheiden (das macht aber nichts – wie schon einmal gesagt).

$$\begin{bmatrix} 1 & 2 & & & \\ 1 & -1 & & & \\ -2 & 2 & 1 & & \\ 1 & 2 & & 1 & \\ 3 & 6 & & & 1 \end{bmatrix} \longmapsto \begin{bmatrix} 9 & 0 & 3 \\ 0 & -6 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \\ -3 & 0 & 0 \end{bmatrix} =: X$$

Leider lösen diese Vektoren nicht alle die letzte Zeile von $A\vec{x} = \vec{0}$. Die Matrix hat also Rang 3.
Wir müssen nun nicht wieder von vorne anfangen und nochmals Kreuzprodukte berechnen. Wir wissen nämlich, dass eine Lösung $\vec{x}$ des homogenen Systems eine Linearkombination obiger drei Kreuzproduktvektoren sein muss. Um die freien Parameter s_1, s_2, s_3 zu erhalten, setzen wir einfach in die letzte Zeile ein. Wir erhalten das reduzierte Gleichungssystem

$$\begin{pmatrix} 54 & 0 & 18 \end{pmatrix} \vec{s} = \vec{0}$$

Das ist schnell gelöst

$$\vec{s} = t_1 \cdot \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + t_2 \cdot \begin{pmatrix} -1 & 0 & 3 \end{pmatrix}.$$

Nun müssen wir nur noch diese Vektoren in die Matrix X einsetzen. Der Lösungsraum des homogenen Systems wird somit von den Vektoren

$$\vec{u}_1 := \begin{pmatrix} 0 & -6 & -3 & 0 & 0 \end{pmatrix} \quad \vec{u}_2 := \begin{pmatrix} 0 & 0 & 0 & -9 & 3 \end{pmatrix}$$

aufgespannt.
Obige Methode sagt uns auch, welche Untermatrix wir für die Partikulärlösung nehmen können. Da $18 \neq 0$ ist die Matrix, die durch Streichen der 3. und 5. Spalte von A entsteht, ein

passender Kandidat. Da wir schon geübt sind, geht das nun ganz schnell

$$\vec{x}_p{}^k = t \cdot \begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 7 \\ -12 \end{pmatrix} = t \cdot \begin{pmatrix} 36 \\ 0 \\ 0 \end{pmatrix}$$

Einsetzen in die Untermatrix liefert $t = \frac{1}{36}$ und somit ist

$$\vec{x}_p = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

eine Partikulärlösung.

6.4 Ein etwas direkterer Weg zur Lösung inhomogener LGS

Bisher sind wir folgendermaßen vorgegangen, um das System $A\vec{x} = \vec{b}$ zu lösen: Wir haben zuerst das homogene System $A\vec{x} = \vec{0}$ gelöst und somit auch den Rang k der Matrix ermittelt. Dann haben wir eine $k \times k$ Untermatrix gefunden und diese verwendet, um eine Partikulärlösung auszurechnen. Das ist vollkommen in Ordnung. Denn in den meisten Aufgaben ist ohnehin auch die Lösung des homogenen Systems verlangt. Außerdem kann bei kleinen Matrizen (z.B. 3×3) die homogene Lösung schnell ermittelt werden.

Manchmal möchte man aber vielleicht einen etwas direkteren Weg nehmen. Anstatt zuerst die Lösung des homogenen Systems zu ermitteln und dann eine Partikulärlösung zu berechnen, lieber gleich alles auf einmal erhalten. Ja, die Menschen sind gierig und ungeduldig. Und ja, es ist möglich so vorzugehen – sogar ganz ohne Gauß-Algo. Dazu benutzen wir die Matrixdarstellung affiner Abbildungen.

Sei dazu $\vec{x}$ eine Lösung des Systems $A\vec{x} = \vec{b}$. Das heißt, der Vektor $\vec{x}$ erfüllt auch folgendes

$$\tilde{A}\,\vec{y} := \begin{pmatrix} A & -\vec{b} \\ 0_{1\times n} & 1 \end{pmatrix} \begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ 1 \end{pmatrix}.$$

Die Matrix $\tilde{A}$ ist eine $(n+1)\times(n+1)$ Matrix. Und jede Lösung $\vec{y} = \left(\begin{smallmatrix} \vec{x} \\ 1 \end{smallmatrix}\right)$ steht orthogonal auf die ersten n Zeilen der Matrix $\tilde{A}$. Somit kann $\vec{y}$ mit Hilfe des Kreuzprodukts berechnet werden. Auch hier kann das System – wenn nötig – reduziert werden, wenn man das Berechnen großer Determinante vermeiden will oder der Rang der Matrix nicht ersichtlich ist.

Am besten wir schauen uns das an einem Beispiel an.

Beispiel 25. Betrachten wir nochmals Beispiel 6 aus dem letzten Abschnitt:

$$\begin{pmatrix} 1 & 1 & -2 & 1 & 3 \\ 2 & -1 & 2 & 2 & 6 \\ 3 & 2 & -4 & -3 & -9 \end{pmatrix} \vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Anstatt das obige zu lösen, können wir auch folgendes lösen

$$\tilde{A}\,\vec{y} = \begin{pmatrix} 1 & 1 & -2 & 1 & 3 & -1 \\ 2 & -1 & 2 & 2 & 6 & -2 \\ 3 & 2 & -4 & -3 & -9 & -3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Berechnen wir die orthogonalen Vektoren zu den ersten zwei Zeilen von $\tilde{A}$.

$$\begin{bmatrix} 1 & 2 & & & & \\ 1 & -1 & & & & \\ -2 & 2 & 1 & & & \\ 1 & 2 & & 1 & & \\ 3 & 6 & & & 1 & \\ -1 & -2 & & & & 1 \end{bmatrix} \longmapsto \begin{bmatrix} 9 & 0 & 3 & -3 \\ 0 & -6 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} =: X$$

Man sieht: die ersten drei Spalten von X sind identisch zu den berechneten in Beispiel 6. Nur die letzte Spalte ist neu. Berechnet wird sie wie die anderen: In der linken Matrix 3 Zeilen mit einer 1 streichen (hier: die 3., 4. und 5. Zeile; diese liefern in X einen Nulleintrag) und das Kreuzprodukt der noch sichtbaren Vektoren im $\mathbb{R}^3$ berechnen.
Somit ist

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ 1 \end{pmatrix} = s_1 \cdot \begin{pmatrix} 9 \\ 0 \\ 0 \\ 0 \\ -3 \\ 0 \end{pmatrix} + s_2 \cdot \begin{pmatrix} 0 \\ -6 \\ -3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_3 \cdot \begin{pmatrix} 3 \\ 0 \\ 0 \\ -3 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 0 \\ -3 \end{pmatrix}$$

Aus der letzten Zeile folgt direkt $t = -1/3$. Die freien Parameter ermitteln wir, indem wir das Vektor für Vektor in die 3. Zeile von $\tilde{A}$ einsetzen. Wir erhalten

$$54\ s_1 + 0\ s_2 + 18\ s_3 + 0 = 0.$$

Somit

$$\vec{s} = t_1 \cdot \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + t_2 \cdot \begin{pmatrix} -1 & 0 & 3 \end{pmatrix}.$$

Das setzen wir jetzt noch oben ein (Stichwort: Matrizenmultiplikation). Den Parameter t kennen wir bereits. Abschließend erhalten wir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ 1 \end{pmatrix} = t_1 \cdot \begin{pmatrix} 0 \\ -6 \\ -3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ -9 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Das ist die gleiche Lösung, die wir auch im letzten Abschnitt erhalten haben!

6.5 Noch ein direkter Weg: Die allgemeine Reduktionsmethode

Bisher sind wir so vorgegangen, um $A\vec{x} = \vec{b}$ zu lösen - vorausgesetzt, dass dieses System überhaupt lösbar ist: Wir haben zuerst das homogene System gelöst und dann mit Hilfe einer passenden quadratischen Untermatrix A_k eine Partikulärlösung ermittelt. Das lieferte die allgemeine Lösung.

Im letzten Abschnitt haben wir gesehen, dass wir die erweiterte Matrix verwenden können, um gleich die allgemeine Lösung des Systems zu ermitteln. Der Nachteil: Die erweiterte Matrix ist um eine Dimension größer als die ursprüngliche Matrix.

Es gibt noch einen dritten Weg, die Lösung von $A\vec{x} = \vec{b}$ zu ermitteln. Dieser Weg liefert ebenfalls direkt die allgemeine Lösung dieses Systems. Anstatt unsere Methode auf eine passende Untermatrix A_k anzuwenden und nur eine Partikulärlösung zu berechnen, wenden wir sie gleich auf die Matrix A an: Das heißt wir multiplizieren $A\vec{x} = \vec{b}$ von rechts mit orthogonalen Vektoren auf $\vec{b}$ und ermitteln mit Hilfe des Kreuzprodukts eine notwendige Bedingung für die allgemeine Lösung $\vec{x}$. Freie Parameter bestimmen wir durch Rückeinsetzen, indem wir ein reduziertes System lösen.

Im Grunde ist das Ganze nur eine Kombination aus bereits bekannten Resultaten. Deshalb werde ich an dieser Stelle nicht weiter auf die Methode eingehen. Interessierten empfehle ich meine optimierte Kurzfassung zu dieser Arbeit. Dort findet sich ein „Kochrezept" wie diese Methode gut anzuwenden ist. Man sieht dann auch, dass man im Grunde so ähnlich vorgehen kann, wie bei der Reduktion mit einer bzw. zwei Zeilen zum Lösen homogener Systeme. Außerdem beinhaltet die optimierte Kurzfassung noch viel mehr Nützliches.

Weitere Infos zur Kurzfassung der Arbeit finden sich auf meiner Webseite

`Forty-Moo.github.io/Kategorien/mathe/mathe_texte/kreuzprodukt_kurz.html`

6.6 Das Invertieren von Matrizen mit dem Kreuzprodukt

Wollen wir eine quadratische Matrix A mit vollem Rang invertieren, dann suchen wir eine Matrix X, sodass

$$A\ X = E_n,$$

wobei E_n die $n \times n$ Einheitsmatrix beschreibt. Das ist äquivalent dazu, Vektoren $x_1, \ldots, x_n$ zu finden, sodass

$$A\vec{x}_1 = \vec{e}_1 \quad \ldots \quad A\vec{x}_n = \vec{e}_n.$$

Das sind n inhomogene Gleichungssysteme, die wir mit der Methode aus dem letzten Abschnitt lösen könnten. Diese Gleichungssysteme haben sogar eine besonders freundliche Form, da rechts immer ein Einheitsvektor steht.
Wirklich sinnvoll ist diese Methode aber nur für Matrizen, die nicht größer als 3×3 sind. Denn sonst wird der Aufwand zum Berechnen der Kreuzprodukte einfach zu groß. Man sollte dann den Gauß-Algo verwenden. Dieser bietet nämlich den Vorteil, obige Gleichungssysteme simultan zu lösen.

Folgendes Beispiel zeigt, wie man vorgehen kann, um eine 3×3 Matrix (relativ elegant) mit Hilfe des Kreuzproduktes zu invertieren:

Beispiel 26. Aufgabe ist es folgende Matrix zu invertieren

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

Im Grunde brauchen wir nur alle möglichen Kreuzprodukte der Zeilen dieser Matrix bilden. Machen wir das:

$$\begin{bmatrix} 2 & 0 & 4 \\ 1 & 3 & 1 \\ 2 & -1 & 1 \end{bmatrix} \mapsto \begin{bmatrix} 4 & -1 & -7 \\ -4 & 6 & 2 \\ -12 & -2 & 6 \end{bmatrix} =: X.$$

Dabei sind wir folgendermaßen vorgegangen: Um die i-te Spalte von X zu berechnen, haben wir die i-te Spalte von A^T (die linke Matrix) gestrichen und das Kreuzprodukt der Vektoren berechnet, die übriggeblieben sind.
Nun setzen wir die erste Spalte von X in die erste Zeile von A ein, da ja $A\vec{x}_1 = \vec{e}_1$ gelten muss. Wir erhalten -20 (das ist übrigens wegen der charakteristischen Eigenschaft des Kreuzprodukts 3.3 gleich der Determinante von A).
Das bedeutet, die 1. Spalte von X muss mit dem Faktor $t_1 = -1/20$ multipliziert werden. Der Faktor der zweiten Spalte von X muss wegen der charakteristischen Eigenschaft und der Schiefsymmetrie des Kreuzprodukts $t_2 = 1/20$ sein. Ebenfalls ohne zusätzliche Rechnung muss $t_3 = -1/20$ werden.
Um die inverse Matrix A^{-1} zu erhalten, müssen wir also X mit dem Faktor $-1/20$ multiplizieren und die Vorzeichen in der zweiten Spalte verändern (bitte nicht vergessen!). Somit

ist

$$A^{-1} = -\frac{1}{20} \cdot \begin{pmatrix} 4 & 1 & -7 \\ -4 & -6 & 2 \\ -12 & 2 & 6 \end{pmatrix}.$$

7 Eigenwerte und Eigenvektoren mit dem Kreuzprodukt

Jetzt sind wir endlich soweit, die bisherige Theorie auch anzuwenden. Und wir werden sehen, wie nützlich das Kreuzprodukt ist, um die Eigenvektoren von Matrizen zu berechnen. Ein Vorteil ist, dass bei der Eigenvektorberechnung nur homogene Systeme zu lösen sind. Über das Ermitteln von Partikulärlösungen brauchen wir uns (vorerst) also keine Gedanken mehr machen. Ein weiterer Vorteil ist, dass wegen der gegebenen Problemstellung die in den homogenen Systemen involvierten Matrizen niemals vollen Rang besitzen.

7.1 Grundlegendes - Eigenwertproblem

Die Begriffe Eigenwert und Eigenvektor sind schnell erklärt. Ein reeller oder komplexer Wert λ (allgemein: ein Wert in einem Körper) ist ein Eigenwert einer reellen oder komplexen $n \times n$ Matrix A (allgemein: einer Matrix über einem Körper K), falls es ein $\vec{v} \neq \vec{0}$ gibt, sodass

$$A\ \vec{v} = \lambda \cdot \vec{v} \tag{7.1}$$

Den Vektor $\vec{v} \neq \vec{0}$ nennen wir dann (einen) Eigenvektor zum Eigenwert λ. Da die Bedingung $\neq \vec{0}$ oft vergessen wird, möchte ich noch einmal darauf hinweisen.

Wie ermittle ich Eigenwerte und Eigenvektoren? Auch das kann schnell beantwortet werden. Dazu formen wir Bedingung 7.1 einfach um:

$$A\ \vec{v} - \lambda \cdot \vec{v} = (A - \lambda \cdot E_n)\ \vec{v} = A_\lambda\ \vec{v} = \vec{0}$$
$$\text{wobei} \quad A_\lambda := A - \lambda \cdot E_n.$$

Die Bedingung 7.1 kann also folgendermaßen umformuliert werden: Gesucht ist ein Wert λ und ein $\vec{v} \neq \vec{0}$, sodass $A_\lambda\ \vec{v} = \vec{0}$.
Solch ein $\vec{v} \neq \vec{0}$ kann es aber nur geben, falls A_λ singulär ist, d.h. nicht vollen Rang besitzt. Wir müssen also λ so wählen, dass A_λ singulär wird. Oder anders ausgedrückt: Finde λ, sodass $\det(A_\lambda) = 0$.
Um die Eigenwerte zu ermitteln, brauchen wir also nichts weiter zu tun, als die Nullstellen der Funktion

$$p_A(x) := \det(A_x) = \det(A - x \cdot E_n)$$

auszurechnen. Bei p_A handelt es sich um ein Polynom vom Grad n. Das kann man z.B. über den Satz von Laplace sehen, indem man sich die Determinante von A_x genauer anschaut. Wir nennen p_A das charakteristische Polynom der Matrix A.
Sind die Nullstellen des Polynoms p_A berechnet, erhält man die zugehörigen Eigenvektoren, indem man die homogenen Systeme $A_\lambda\ \vec{v} = \vec{0}$ für jeden Eigenwert λ löst. Klarerweise gilt $rg(A_\lambda) < n$. Zum Lösen können wir das Kreuzprodukt zu Hilfe nehmen.

7.2 Grundlegendes - Eigenraum, geometrische und algebraische Vielfachheit

Wir benötigen noch ein paar Begriffe. Packen wir alle Eigenvektoren zu einem Eigenwert λ in eine Menge, erhalten wir den Eigenraum U_λ zu einem Eigenwert λ:

$$U_\lambda := ker(A_\lambda) = \left\{ \vec{v} \in V \mid A_\lambda \, \vec{v} = \vec{0} \right\}.$$

Die Dimension dieses Raumes U_λ nennen wir die geometrische Vielfachheit g_λ des Eigenwertes λ

$$g_\lambda := dim(U_\lambda).$$

Die Zahl g_λ beschreibt also die Maximalanzahl linear unabhängiger Eigenvektoren zu einem Eigenwert λ. Die geometrische Vielfachheit ist immer > 0, da es zu jedem Eigenwert per definitionem eine Eigenvektor $\vec{v} \neq \vec{0}$ gibt. Somit $U_\lambda \neq \{\vec{0}\}$.
Es gibt noch eine weitere Kennzahl für einen Eigenwert λ einer Matrix A: die algebraische Vielfachheit. Wir wissen bereits, dass die Eigenwerte die Nullstellen des charakteristischen Polynoms p_A der Matrix A sind. Nun kann es aber vorkommen, dass der Eigenwert λ eine mehrfache Nullstelle des Polynoms p_A ist. Genauso wie z.B. der Wert 0 zweifache Nullstelle des Polynoms $q(x) = x^2$ ist.
Die Vielfachheit des Eigenwertes λ als Nullstelle im Polynom p_A bezeichnen wir als algebraische Vielfachheit a_λ des Eigenwerts λ. Also

$$a_\lambda := \text{„Vielfachheit von } \lambda \text{ in } p_A\text{“}.$$

Es gilt folgender Zusammenhang zwischen algebraischer und geometrischer Vielfachheit eines Eigenwertes λ

$$0 < g_\lambda \leqslant a_\lambda.$$

Daraus wird auch ersichtlich: Ist λ ein einfacher Eigenwert, das heißt $a_\lambda = 1$, dann gibt es nur einen l. u. Eigenvektor zu λ. Dieser spannt dann den Eigenraum U_λ auf.

Folgender Spezialfall ist besonders wichtig: Gilt

1. $g_\lambda = a_\lambda$ für jeden Eigenwert λ von A und
2. die Matrix A hat genau n Eigenwerte (mit ihren Vielfachheiten gezählt),

dann sagt man, die Matrix A ist diagonalisierbar. Denn es gibt dann eine Basis B, bestehend aus den l. u. Eigenvektoren, die die verschiedenen Eigenräume U_λ aufspannen, bezüglich der die lineare Abbildung A Diagonalgestalt besitzt. Das heißt in diesem Fall kann man die Abbildung A auf eine besonders einfache Gestalt bringen.

Berechnen wir nun die Eigenwerte und Eigenvektoren einiger Matrizen. Hier ist das charakteristische Polynom noch einfach zu berechnen und wir müssen zur Ermittlung der Eigenwerte nicht auf „Tricks“ zurückgreifen.

Beispiel 27. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

Gesucht sind Eigenwerte und Eigenvektoren. Beginnen wir damit, das charakteristische Polynom zu berechnen.

$$p_A(x) = \det(A - x \cdot E_n) = \det \begin{pmatrix} 1-x & 2 & 3 \\ 0 & 4-x & 5 \\ 0 & 0 & 6-x \end{pmatrix} = (1-x) \cdot (4-x) \cdot (6-x)$$

Die Nullstellen von p_A sind $\lambda_1 = 1, \lambda_2 = 4$ und $\lambda_3 = 6$. Somit haben wir die Eigenwerte berechnet.
Kümmern wir uns nun um die Eigenvektoren. Wir wollen zuerst

$$A_{\lambda_1}\ \vec{v} = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 3 & 5 \\ 0 & 0 & 5 \end{pmatrix}\ \vec{v} = \vec{0}$$

lösen. Die Matrix A_{λ_1} hat Rang 2, da λ_1 algebraische Vielfachheit 1 besitzt. Um die Lösung zu ermitteln, berechnen wir das Kreuzprodukt zweier linear unabhängiger Zeilen, z.B. von Zeile 2 und 3.

$$\vec{v}_1 := \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 15 \\ 0 \\ 0 \end{pmatrix} \mid\mid \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Dieser Vektor $\vec{v}_1$ spannt nun den Eigenraum zu λ_1 auf.
Analog gehen wir bei λ_2 vor. Wir wollen

$$A_{\lambda_2}\ \vec{v} = \begin{pmatrix} -3 & 2 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 2 \end{pmatrix}\ \vec{v} = \vec{0}$$

lösen und berechnen das Kreuzprodukt der ersten und letzten Zeile von A_{λ_2}. Somit ist

$$\vec{v}_2 := \begin{pmatrix} -3 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 0 \end{pmatrix} \mid\mid \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

Nun müssen wir nur noch

$$A_{\lambda_3}\ \vec{v} = \begin{pmatrix} -5 & 2 & 3 \\ 0 & -2 & 5 \\ 0 & 0 & 0 \end{pmatrix}\ \vec{v} = \vec{0}$$

lösen. Wir erhalten

$$\vec{v}_3 := \begin{pmatrix}-5\\2\\3\end{pmatrix} \times \begin{pmatrix}0\\-2\\5\end{pmatrix} = \begin{pmatrix}16\\25\\10\end{pmatrix}$$

Im nachfolgenden Beispiel sehen wir, wie wir uns eine Menge Schreibarbeit ersparen können, indem wir unsere Notation anpassen.

Beispiel 28. Wir betrachten die Matrix

$$A = \begin{pmatrix}3 & 1 & 0\\0 & 1 & 0\\0 & 0 & 4\end{pmatrix}$$

Gesucht sind Eigenwerte und Eigenvektoren. Beginnen wir damit, das charakteristische Polynom zu berechnen.

$$p_A(x) = \det(A - x \cdot E_n) = \det \begin{pmatrix}3-x & 1 & 0\\0 & 1-x & 0\\0 & 0 & 4-x\end{pmatrix} = (3-x)\cdot(1-x)\cdot(4-x)$$

Die Nullstellen von p_A sind $\lambda_1 = 3, \lambda_2 = 1$ und $\lambda_3 = 4$. Somit haben wir die Eigenwerte berechnet.
Die Eigenvektoren berechnen wir mit dem Kreuzprodukt. Machen wir das in Kurzschreibweise, da sich für jedes λ_i für die Berechnung von A_{λ_i} nur die Diagonalelemente von A ändern:

$$A^T_{\lambda_i} = \begin{pmatrix}0 \mid 2 \mid -1 & 0 & 0\\1 & -2 \mid 0 \mid -3 & 0\\0 & 0 & 1 \mid 3 \mid 0\end{pmatrix} \quad \mapsto \quad \begin{bmatrix}1 & 3 & 0\\0 & -6 & 0\\0 & 0 & 3\end{bmatrix} =: X$$

Links haben wir das Kreuzprodukt zweier linear unabhängiger Spalten berechnet und das Ergebnis rechts eingetragen. Da die Eigenwerte einfach sind, wissen wir, dass $A^T_{\lambda_i}$ Rang 2 besitzt. Es muss diese zwei l. u. Spalten also geben.
Die erste Spalte von X ist der Eigenvektor, der den Eigenraum zu λ_1 aufspannt. Die zweite Spalte ist der Eigenvektor, der den Eigenraum zu λ_2 aufspannt. Und die dritte Spalte ist der Vektor, der den Eigenraum zu λ_3 aufspannt.

Wir sehen, dass wir uns tatsächlich mit dieser abgekürzten Notation für die $A^T_{\lambda_i}$ viel Schreibarbeit ersparen. Vor allem bei kleinen Matrizen sind wir so wesentlich schneller fertig.
Durch diese abgekürzte Notation können aber auch leichter Fehler bei der Kreuzproduktberechnung passieren. Man könnte z.B. ein falsches Diagonalelement für die Berechnung heranziehen und erhält dann ein falsches Ergebnis. Deshalb würde ich für die Praxis folgendes Vorgehen empfehlen: Man schreibt zuerst die Matrix A^T hin, wobei man die Diagonaleinträge frei lässt und dort genügend Platz „reserviert“. Danach trägt man zunächst die Diagonalelemente von A_{λ_1} ein und berechnet das Kreuzprodukt bzw. die orthogonalen Vektoren. Dann streicht man die vorher eingetragenen Diagonalelemente durch und trägt die neuen Diagonalelemente von A_{λ_2} ein. Anschließend berechnet man wieder das Kreuzprodukt usw.

Mit dieser Methode sollte man relativ platzsparend und schnell Eigenvektoren berechnen können. Zumindest bei 3×3 Matrizen sogar völlig ohne irgendwelche Nebenrechnungen. Bei größeren Matrizen könnte eine Nebenrechnung für die Kreuzproduktberechnung nötig werden.

7.3 Eigenwertberechnung ohne Kenntnis des charakteristischen Polynoms

Eigenvektoren von kleinen Matrizen (bis 4×4) sollten wir nun schon relativ effizient ermitteln können. Unangenehm kann nur noch die Berechnung des charakteristischen Polynoms werden. Wer schon einmal das charakteristische Polynom einer voll besetzten 3×3 Matrix per Hand berechnen musste, weiß, wovon ich rede. Das ist eine äußerst lästige und fehleranfällige Aufgabe. Und dann müssen auch noch die Nullstellen dieses Polynoms berechnet werden. Das heißt Linearfaktor abspalten und pq-Formel bedienen. Und schon ist das Blatt voll. Bei Prüfungen hat man diese Zeit einfach nicht!
Was viele nicht wissen: Zumindest bei 3×3 Matrizen (und teils auch bei 4×4) kann man sich diese ganze Arbeit meistens ersparen. Im Idealfall braucht man nur die Determinante und Spur einer 3×3 Matrix berechnen und kann danach in eine Formel einsetzen. Das sind ein bis zwei Zeilen Arbeit!

7.3.1 Die Struktur des charakteristischen Polynoms

Wir benutzen im Folgenden eine etwas abgeänderte Definition des charakteristischen Polynoms

$$p_A(x) := \det\left(x \cdot E_n - A\right)$$

Die Definition unterscheidet sich von der vorigen nur um ein Vorzeichen. Ansonsten ändert sich nichts. Der einzige Grund dieser Änderung: Das charakteristische Polynom ist jetzt normiert (Leitkoeffizient $= 1$), was angenehm ist.

Die Koeffizienten des charakteristischen Polynoms können explizit ausgerechnet werden. Folgendes Resultat zeigt wie:

Satz 7.3.1. *Sei A eine $n \times n$ Matrix. Das charakteristische Polynom der Matrix A sei*

$$p_A(x) = x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0.$$

Die Koeffizienten dieses Polynoms lassen sich dann folgendermaßen berechnen

$$a_k = (-1)^{n-k} \cdot tr(\Lambda^{n-k}\ A), \tag{7.2}$$

wobei $tr(\Lambda^{n-k}\ A)$ die Summe aller Hauptminoren der Ordnung $n-k$ von A beschreibt.

Ein Hauptminor der Ordnung $n-k$ ist hierbei die Determinante einer $(n-k) \times (n-k)$ Untermatrix von A, die durch Streichungen von Zeilen und Spalten derselben Nummern entsteht.

Beweis. Es sei

$$A = \begin{pmatrix} | & & | \\ \vec{a}_1 & \ldots & \vec{a}_n \\ | & & | \end{pmatrix} \quad \text{und} \quad p_A(x) = \det(x \cdot E_n - A).$$

Folglich ist

$$p_A(x) = \det(x \cdot \vec{e}_1 - \vec{a}_1, \ldots, x \cdot \vec{e}_n - \vec{a}_n).$$

Um den Koeffizienten a_k vor x^k des Polynoms p_A zu bestimmen, benutzen wir nun die Multilinearität der Determinante. Um auf x^k zu kommen, müssen wir in obiger Determinante in k der Argumente den Ausdruck mit x wählen und in den restlichen $n-k$ Argumenten, die übrigbleiben, den jeweiligen „$\vec{a}_i$ -Vektor“.
Machen wir das einmal exemplarisch, indem wir die x-Ausdrücke in den ersten k Argumenten der Determinante auswählen. Wir erhalten

$$\begin{aligned} a_k \cdot x^k &= \det(x \cdot \vec{e}_1, \ldots, x \cdot \vec{e}_k, -\vec{a}_{k+1}, \ldots, -\vec{a}_n) + (\ldots)\ x^k \\ &= (-1)^{n-k}\ \det(\vec{e}_1, \ldots, \vec{e}_k, \vec{a}_{k+1}, \ldots, \vec{a}_n)\ x^k + (\ldots)\ x^k \end{aligned}$$

Den Ausdruck $\det(\vec{e}_1, \ldots, \vec{e}_k, \vec{a}_{k+1}, \ldots, \vec{a}_n)$ berechnen wir nun mit Laplace-Entwicklung (nacheinander bezüglich der ersten k-Spalten). Bei diesem Vorgang werden die ersten k-Zeilen (und Spalten) gestrichen. Was übrigbleibt ist der rechte untere Hauptminor der Ordnung $n-k$ von A.
Das war EINE Möglichkeit k Ausdrücke mit x, aus n Argumenten zu wählen. Um den Koeffizienten vor x^k zu berechnen, müssen wir ALLE Möglichkeiten durchgehen. Insgesamt gibt es „n über k“ Möglichkeiten aus n Argumenten k-mal x auszuwählen. Bei jeder dieser Möglichkeiten erhalten wir einen anderen Hauptminor der Ordnung $n-k$ (da bei jeder Möglichkeit andere Einheitsvektoren involviert sind) mit Vorzeichen $(-1)^{n-k}$.
Schlussendlich werden alle auf diese Weise berechneten Werte addiert (wegen Multilinearität) und das Ergebnis liefert den Koeffizienten vor x^k. Somit haben wir gezeigt, dass a_k die Summe der Hauptminoren der Ordnung $n-k$ ist, mit dem Vorzeichen $(-1)^{n-k}$.
Hat man Probleme mit dem Beweis, wäre es sicherlich eine gute Übung, den Beweis Schritt für Schritt für die Werte $n = 3, 4$ und $k = 2$ durchzugehen, d.h. a_2 mit obiger Methode zu berechnen. □

Hieraus erhalten wir direkt folgendes Korollar

Korollar 7.3.2. *Es gilt immer*

$$a_{n-1} = -tr(A) \quad \text{und} \quad a_0 = (-1)^n \cdot \det(A). \tag{7.3}$$

Beweis. Wir benutzen 7.2. Es folgt

$$a_{n-1} = (-1)^1 \cdot tr(\Lambda^1\ A).$$

Die Hauptminoren der Ordnung 1 sind aber gerade die Diagonalelemente der Matrix A. Ihre Summe ergibt die Spur der Matrix A.
Außerdem erhalten wir

$$a_0 = (-1)^n \cdot tr(\Lambda^n\ A).$$

Es gibt nur einen Hauptminor der Ordnung n, da es nur eine $n \times n$ Untermatrix in A gibt. Diese Untermatrix ist die Matrix A selbst. Ihre Determinante liefert den Hauptminor. □

Außerdem erhalten wir folgendes, oftmals nützliches, Resultat:

Korollar 7.3.3. *Zerfällt das charakteristische Polynom p_A in Linearfaktoren (über $\mathbb{C}$ tut es das z.B. immer), das heißt*

$$p_A(x) = \prod_{i=1}^{n} (x - \lambda_i),$$

wobei die λ_i nicht notwendigerweise verschieden sein müssen, dann gilt

$$tr(A) = \sum_{i=1}^{n} \lambda_i \quad \textit{sowie} \quad \det(A) = \prod_{i=1}^{n} \lambda_i. \tag{7.4}$$

Beweis. Aus dem letzten Korollar wissen wir, dass

$$a_0 = (-1)^n \cdot \det(A)$$

gilt. Außerdem gilt

$$a_0 = p_A(0) = \prod_{i=1}^{n}(-\lambda_i) = (-1)^n \cdot \prod_{i=1}^{n} \lambda_i.$$

Somit ist die Gleichheit für die Determinante bewiesen.
Ebenso wissen wir, dass

$$a_{n-1} = -tr(A)$$

gilt. Wir wollen also den Koeffizienten vor x^{n-1} in

$$p_A(x) = (x - \lambda_1) \cdots (x - \lambda_n)$$

bestimmen. Dazu benutzen wir das Distributivgesetz: Wir multiplizieren eines der $-\lambda_i$ mit $n-1$ x–Termen. Das müssen wir für jedes der $-\lambda_i$ machen und die Ergebnisse addieren. Für den Koeffizient vor x^{n-1} erhalten wir also

$$a_{n-1} = -(\lambda_1 + \cdot + \lambda_n).$$

Somit stimmt auch die Gleichheit für die Spur. □

7.3.2 Eigenwertberechnung für 2×2 Matrizen

Für komplexe oder reelle 2×2 Matrizen ist es leicht, die Eigenwerte zu bestimmen. Denn das charakteristische Polynom nimmt eine besonders einfache Form an.

Korollar 7.3.4. *Sei A eine 2×2 Matrix. Das charakteristische Polynom hat dann die Form*

$$p_A(x) = x^2 - tr(A) \cdot x + \det(A). \tag{7.5}$$

Die Eigenwerte λ_{12} berechnen sich mit Hilfe der pq-Formel zu

$$\lambda_{12} = \frac{tr(A)}{2} \pm \sqrt{\frac{tr(A)^2}{4} - \det A}. \tag{7.6}$$

Beweis. Das folgt direkt aus 7.3.2, wenn man $n = 2$ setzt:

$$a_1 = -tr(A) \quad \text{sowie} \quad a_0 = (-1)^2 \cdot \det(A) = \det(A).$$

□

Hinweis 7.3.5. Oftmals ist es nicht einmal nötig, in die pq-Formel einzusetzen. Falls wir auch komplexe Eigenwerte zulassen, gilt ja wegen 7.3.3

$$\begin{aligned} \lambda_1 + \lambda_2 &= tr(A) \\ \lambda_1 \cdot \lambda_2 &= \det(A) \end{aligned} \tag{7.7}$$

Jeder Eigenwert erfüllt obiges System. Ebenso sind zwei Werte, die dieses System erfüllen, bereits Eigenwerte.
Man könnte also versuchen, die Eigenwerte über das System 7.7 einfach zu erraten. Ein Versuch schadet nicht! Sollte das Erraten nicht möglich sein, kann man ja immer noch in die pq-Formel einsetzen.

Beispiel 29. Betrachten wir die Matrix

$$A = \begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix}.$$

Gesucht sind Eigenwerte und Eigenvektoren. Es gilt $tr(A) = 2$ sowie $\det(A) = -24$. Aus $\lambda_1 + \lambda_2 = 2$ und $\lambda_1 \cdot \lambda_2 = -24$ könnten wir nun direkt $\lambda_1 = 6$ und $\lambda_2 = -4$ erraten. Ansonsten setzt man eben in die Formel ein, was dasselbe Ergebnis liefert:

$$\lambda_{21} = \frac{tr(A)}{2} \pm \sqrt{\frac{tr(A)^2}{4} - \det A} = 1 \pm \sqrt{1 + 24} = 1 \pm 5.$$

Für die Eigenvektoren brauchen wir nur den orthogonalen Vektor zur ersten Spalte von $A_{\lambda_i}^T$

ermitteln, was sehr einfach ist:

$$A_{\lambda_i}^T = \begin{bmatrix} -5 \mid 5 & 5 \\ 5 & -5 \mid 5 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} -5 & -5 \\ -5 & 5 \end{bmatrix} =: X$$

Die Spalten von X sind die gesuchten Eigenvektoren, die die jeweiligen Eigenräume aufspannen.

7.3.3 Eigenwertberechnung für 3×3 Matrizen

Wenden wir uns nun der Eigenwertberechnung von 3×3 Matrizen zu. Im Grunde ist das nicht viel schwieriger als bei 2×2 Matrizen. Nur die Berechnung der Determinante ist etwas aufwändiger. Bei voll besetzten Matrizen, bei denen das charakteristische Polynom nicht sofort ersichtlich ist, kann die nachfolgende Methode einen beträchtlichen Geschwindigkeitsvorteil bei der Eigenwertberechnung liefern.
Fangen wir wieder mit der Gestalt des charakteristischen Polynoms an:

Korollar 7.3.6. *Sei A eine 3×3 Matrix. Das charakteristische Polynom p_A hat dann folgende Gestalt*

$$p_A(x) = x^3 - tr(A) \cdot x^2 + \big(\det(A^{11}) + \det(A^{22}) + \det(A^{33})\big) \cdot x - \det(A), \tag{7.8}$$

wobei A^{ij} die Matrix ist, die durch Streichen der i-ten Zeile und j-ten Spalte von A zu Stande kommt.

Beweis. Folgt direkt aus Formel 7.2 und Korollar 7.3.2 für $n = 3$:

$$a_0 = (-1)^3 \cdot \det(A) = -\det(A) \quad a_2 = -tr(A) \quad a_1 = (-1)^2 \cdot tr(\Lambda^2 \ A),$$

wobei $tr(\Lambda^2 \ A)$ die Summe der 2×2 Hauptminoren von A beschreibt. □

Bei einer voll besetzten 3×3 Matrix können wir also das charakteristische Polynom mit dem obigen Resultat bestimmen. Das weitere Vorgehen sieht dann meistens so aus: Man errät eine Nullstelle des Polynoms dritten Grades und spaltet einen Linearfaktor ab ... Auch dieses Vorgehen können wir noch optimieren!
Zuerst aber ein kurzes Lemma zu Nullstellen und Linearfaktoren, bevor wir zum wirklich wichtigen Resultat kommen.

Lemma 7.3.7. *Sei K ein Körper (z.B. $K = \mathbb{R}$ oder $K = \mathbb{C}$). Für ein Polynom p mit Koeffizienten im Körper K gilt dann*

$$x_0 \text{ ist Nullstelle von } p \quad \Longleftrightarrow \quad (x - x_0) \text{ teilt } p, \tag{7.9}$$

wobei die Aussage „$(x - x_0)$ teilt p“ folgendes bedeutet: es gibt ein Polynom q, sodass

$$p(x) = q(x) \cdot (x - x_0).$$

Beweis. ($\Leftarrow$) Hier ist nichts zu zeigen: Der Wert x_0 ist Nullstelle, da $(x_0 - x_0) = 0$.

($\Rightarrow$) Sei x_0 Nullstelle von p, das heißt $p(x_0) = 0$. Da K ein Körper ist, haben wir die Polynomdivision zur Verfügung: Wir teilen nun p durch $x - x_0$ mit Polynomrest r: Es gibt also ein Polynom q, sodass

$$p(x) = q(x) \cdot (x - x_0) + r(x).$$

Da $(x - x_0)$ als Polynom Grad 1 hat und bei der Polynomdivision der Grad des Polynomrestes r immer abnimmt, muss r ein Polynom vom Grad 0 sein. Das heißt r ist ein konstantes Polynom. Setzen wir nun x_0 in die obige Gleichheit ein:

$$0 = p(x_0) = q(x_0) \cdot (x - x_0) + r(x_0) = r(x_0)$$

und da r ein konstantes Polynom ist, folgt $r \equiv 0$. □

Haben wir also eine Nullstelle eines Polynoms gefunden, kann immer ein Linearfaktor abgespalten werden. Wie hilft uns dieses (abstrakte) Resultat weiter? Ganz einfach! Wir können uns nun die Polynomdivision völlig ersparen!

Satz 7.3.8. *Sei u ein normiertes Polynom vom Grad 3, das heißt, es hat folgende Gestalt*

$$u(x) = x^3 + a_2 \cdot x^2 + a_1 \cdot x + a_0,$$

wobei die Koeffizienten aus einem Körper stammen (z.B. aus $\mathbb{R}$ oder $\mathbb{C}$). Außerdem sei x_0 eine Nullstelle von u, das heißt u lässt sich folgendermaßen darstellen

$$u(x) = (x^2 + p \cdot x + q) \cdot (x - x_0).$$

Dann gilt für die Werte p und q:

$$\textbf{\textit{falls}}\ \mathbf{x_0 \neq 0}: \qquad p = a_2 + x_0 \quad \textit{sowie} \quad q = -\frac{a_0}{x_0} \tag{7.10}$$

$$\textbf{\textit{falls}}\ \mathbf{x_0 = 0}: \qquad p = a_2 \quad \textit{sowie} \quad q = a_1 \tag{7.11}$$

Beweis. **falls $\mathbf{x_0 = 0}$:** Dann gilt

$$0 = u(0) = a_0.$$

Somit ist

$$u(x) = x^3 + a_2 \cdot x^2 + a_1 \cdot x = (x^2 + a_2 \cdot x + a_1) \cdot x$$

und die Aussage ist für $x_0 = 0$ gezeigt.

falls $\mathbf{x_0 \neq 0}$: Wir ermitteln p und q durch Koeffizientenvergleich:

$$x^3 + a_2 \cdot x^2 + a_1 \cdot x + a_0 = (x^2 + p \cdot x + q) \cdot (x - x_0).$$

Zuerst vergleichen wir die konstanten Koeffizienten. Wir erhalten

$$-q \cdot x_0 = a_0 \quad \Rightarrow \quad q = -\frac{a_0}{x_0}.$$

Nun vergleichen wir noch die Koeffizienten vor x^2. Daraus folgt

$$p - x_0 = a_2 \quad \Rightarrow \quad p = a_2 + x_0.$$

□

Mit Hilfe des obigen Resultats können wir die restlichen beiden Eigenwerte von komplexen oder reellen 3×3 Matrizen bestimmen, unter der Voraussetzung, dass ein Eigenwert bereits bekannt ist (bzw. erraten wurde).

Korollar 7.3.9. *Sei p_A das charakteristische Polynom einer 3×3 Matrix A. Außerdem sei λ_1 ein bekannter Eigenwert der Matrix A. Dann lassen sich die restlichen beiden Nullstellen*

λ_{23} mit der pq - Formel berechnen:

$$\lambda_{23} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \tag{7.12}$$

Die Werte p und q lassen sich hierbei folgendermaßen berechnen:

falls $\lambda_1 \neq 0$:

$$p = \lambda_1 - tr(A) \quad , \quad q = \frac{\det(A)}{\lambda_1} \tag{7.13}$$

falls $\lambda_1 = 0$:

$$p = -tr(A) \quad , \quad q = \sum_{i=1}^{3} \det(A^{ii}). \tag{7.14}$$

Beweis. Wir kombinieren einfach 7.3.6 mit den Formeln für p und q aus 7.3.8. Das Resultat folgt direkt. □

Die Eigenwertberechnung von 3×3 Matrizen wird also auf folgendes hinauslaufen:

Punkt 1 Ermittle Spur und Determinante der Matrix A. Wenn nötig, berechne auch die Summe der Hauptminoren der Ordnung 2.

Punkt 2 Errate einen Eigenwert λ_1.

Punkt 3 Setze in die Formeln für p, q ein und ermittle die restlichen beiden Eigenwerte mit der pq - Formel.

Bevor wir uns das anhand von Beispielen näher anschauen, noch ein nützlicher Hinweis, der sich auf Punkt 2 bezieht:

Hinweis 7.3.10. Natürlich könnte man nun so vorgehen, dass man via 7.3.6 das charakteristische Polynom explizit aufstellt und dann durch willkürliches Einsetzen eine Nullstelle errät.
Oftmals ist solch ein Vorgehen aber gar nicht nötig. Ist z.B. die Determinante der Matrix $= 0$ bzw. ist es an der Gestalt der Matrix ersichtlich, dass sie singulär ist, ist $\lambda_1 = 0$ automatisch ein Eigenwert.
Selbst, wenn die Matrix nicht singulär ist, ist es bei vielen Aufgaben möglich, sehr einfach einen Eigenwert zu ermitteln. Man muss ja nur dafür sorgen, dass A_λ singulär wird. Dazu kann man sich ein wenig mit den Diagonaleinträgen der Matrix A „spielen". Vielleicht schafft man es ja, dadurch eine doppelte Zeile oder Spalte zu generieren. Ein Versuch schadet nicht!
Trägt das keine Früchte, kann man immer noch das charakteristische Polynom explizit aufstellen und versuchen, eine Nullstelle durch Einsetzen zu ermitteln.
Später werden wir sehen, dass die Determinante von A ein Hilfsmittel ist, um schneller Nullstellen bzw. Eigenwerte zu erraten.

Übrigens besteht auch jederzeit die Möglichkeit, (verbleibende) Eigenwerte über die Bedingungen

$$\lambda_1 + \lambda_2 + \lambda_3 = tr(A)$$

$$\lambda_1 \cdot \lambda_2 + \lambda_1 \cdot \lambda_3 + \lambda_2 \cdot \lambda_3 = \sum_{i=1}^{3} \det(A^{ii})$$

$$\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = \det(A)$$

zu ermitteln. Die Eigenwerte sind dadurch eindeutig bestimmt.
Besonders nützlich ist das, wenn man es schafft, zwei Eigenwerte zu erraten. Der dritte folgt dann nämlich direkt aus der ersten Gleichheit!

Beispiel 30. Wir betrachten die Matrix

$$A = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$$

Gesucht sind Eigenwerte und Eigenräume.
Berechnen wir zuerst Spur und Determinante dieser Matrix.

$$tr(A) = 0 \quad , \quad \det(A) = (-3) \cdot (-4) + (-7) \cdot (-4) + (-6) \cdot 4 = 16.$$

Erraten wir nun einen Eigenwert λ_1. Nach kurzem Probieren erkennt man: Addiert man 2 zu den Diagonalelementen, sind die ersten zwei Spalten der Matrix, bis auf ein Vorzeichen, identisch. Somit ist $\lambda_1 = -2$.
Hieraus folgt

$$p = \lambda_1 - tr(A) = -2 \quad , \quad q = \frac{\det(A)}{\lambda_1} = -8$$

und somit

$$\lambda_{23} = 1 \pm \sqrt{1+8} = 1 \pm 3.$$

Jetzt berechnen wir noch die Eigenvektoren zu diesen Eigenwerten mit dem Kreuzprodukt. Verwenden wir wieder die Kurzschreibweise:

$$A_{\lambda_i}^T = \begin{bmatrix} -1 \mid -7 & -7 & -6 \\ 1 & 7 \mid 1 & 6 \\ -1 & -1 & 0 \mid -6 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} 6 & 0 \\ 6 & -36 \\ 0 & -36 \end{bmatrix} \Bigg|\Bigg| \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & -1 \end{bmatrix}.$$

Aufgrund der sehr einfachen Struktur der Matrizen A_{λ_i}, hätten wir hier nicht einmal das Kreuzprodukt benötigt. Wir hätten die Eigenvektoren durch „Hinschauen“ ermitteln können.

Manchmal ist es klug, zuerst zu versuchen, einen Eigenwert zu erraten, bevor man die Determinante der Matrix berechnet. Im nachfolgenden Beispiel sehen wir, warum.

Beispiel 31. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$$

Gesucht sind Eigenwerte und Eigenräume.
Addieren wir zu den Diagonaleinträgen den Wert 2, sehen wir, dass die zweite und dritte Spalte skalare Vielfache der ersten sind. Der Wert $\lambda_1 = -2$ ist somit nicht nur einfacher Eigenwert, sondern mindestens ein doppelter. Das folgt daraus, dass es wegen $rg(A_{-2}) = 1$, zwei linear unabhängige Eigenvektoren zum Eigenwert -2 gibt. Somit ist die geometrische Vielfachheit dieses Wertes gleich 2. Und wir wissen, dass die algebraische Vielfachheit eines Eigenwertes immer größer gleich der geometrischen ist. Somit muss $\lambda_1 = -2$ mindestens ein doppelter Eigenwert sein.
Das ist praktisch, denn wir können jetzt den potentiellen dritten Eigenwert aus der Bedingung

$$\lambda_1 + \lambda_2 + \lambda_3 = 2 \cdot \lambda_1 + \lambda_3 = tr(A) = 0$$

ermitteln. Hieraus folgt $\lambda_3 = 4$.
Die Eigenvektoren zu λ_1 berechnen wir durch „Vertauschen und Vorzeichenwechsel". Somit erhalten wir

$$\vec{v}_1 = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \quad , \quad \vec{v}_2 = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}.$$

Den Eigenvektor zu λ_3 berechnen wir mit dem Kreuzprodukt:

$$\vec{v}_3 = \begin{pmatrix} -3 \\ -3 \\ 3 \end{pmatrix} \times \begin{pmatrix} 6 \\ -6 \\ 0 \end{pmatrix} = \begin{pmatrix} 18 \\ 18 \\ 36 \end{pmatrix} \Bigg|\Bigg| \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Machen wir noch ein letztes Beispiel, in dem der Eigenwert 0 auftritt.

Beispiel 32. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & -3 & 5 \\ 4 & -1 & 8 \end{pmatrix}$$

Gesucht sind die Eigenwerte.
Entweder über die Determinante oder durch genaues Hinsehen (die dritte Zeile ist die Summe der ersten beiden), erkennen wir, dass die Matrix singulär ist. Somit ist $\lambda_1 = 0$ ein Eigenwert. Es gilt $tr(A) = 6$. Jetzt benötigen wir noch die Summe der Hauptminoren:

$$tr(\Lambda^2\ A) = \det \begin{pmatrix} -3 & 5 \\ -1 & 8 \end{pmatrix} + \det \begin{pmatrix} 1 & 3 \\ 4 & 8 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 \\ 3 & -3 \end{pmatrix} = -32.$$

Somit ist $p = -6$ und $q = -32$. Wir benutzen die pq-Formel und erhalten

$$\lambda_{23} = 3 \pm \sqrt{9 + 32} = 3 \pm \sqrt{41}.$$

Wenn gewünscht, könnten wir nun wieder die 3 linear unabhängigen Eigenvektoren mit dem Kreuzprodukt berechnen.

7.3.4 Eigenwertberechnung für 4×4 Matrizen

Aufgaben mit 3×3 Matrizen können wir nun schon relativ schnell lösen. Bei 4×4 Matrizen gestaltet sich das Ermitteln der Eigenwerte etwas schwieriger. Glücklicherweise treten in Aufgaben selten 4×4 Matrizen auf, da das Berechnen des charakteristischen Polynoms zur Qual werden kann. Treten sie doch einmal auf, haben sie meist eine recht angenehme Form: Sie zerfallen z.B. in Blockmatrizen und das charakteristische Polynom kann einfach bestimmt werden. Hat die Matrix hingegen keine freundliche Gestalt, kann man sich trotzdem noch behelfen.
Wir starten wieder mit der Gestalt des charakteristischen Polynoms:

Korollar 7.3.11. *Sei A eine 4×4 Matrix. Das charakteristische Polynom hat dann die Form*

$$p_A(x) = x^4 - tr(A) \cdot x^3 + tr(\Lambda^2\ A) \cdot x^2 - tr(\Lambda^3\ A) \cdot x + \det(A). \tag{7.15}$$

Beweis. Folgt wieder direkt aus 7.2. □

Auch hier können wieder Linearfaktoren ohne Polynmomdivision abgespalten werden:

Satz 7.3.12. *Sei u ein normiertes Polynom vom Grad 4, das heißt, es hat folgende Gestalt*

$$u(x) = x^4 + a_3 \cdot x^3 + a_2 \cdot x^2 + a_1 \cdot x + a_0,$$

wobei die Koeffizienten aus einem Körper stammen (z.B. aus $\mathbb{R}$ oder $\mathbb{C}$). Außerdem sei x_0 eine Nullstelle von u, das heißt u lässt sich folgendermaßen darstellen

$$u(x) = (x^3 + b_2 \cdot x^2 + b_1 \cdot x + b_0) \cdot (x - x_0).$$

Dann gilt für die Werte b_2, b_1 und b_0:

$$\textbf{\textit{falls}}\ \mathbf{x_0 \neq 0}: \qquad b_2 = a_3 + x_0 \quad , \quad b_1 = a_2 + b_2 \cdot x_0 \quad \textit{und} \quad b_0 = -\frac{a_0}{x_0} \tag{7.16}$$

$$\textbf{\textit{falls}}\ \mathbf{x_0 = 0}: \qquad b_2 = a_3 \qquad , \quad b_1 = a_2 \quad \textit{und} \quad b_0 = a_1. \tag{7.17}$$

Beweis. Siehe den Beweis von 7.3.8. Hier funktioniert es analog. Wir benötigen nur einen zusätzlichen Koeffizientenvergleich. □

Für die a_i, $i = 0, \ldots, 3$, können wir nun die Koeffizienten des charakteristischen Polynoms p_A einsetzen. Diese müssen wir natürlich erst einmal berechnen. Neben Spur und Determinante benötigen wir hier im Falle $x_0 \neq 0$ auch die Summe der Hauptminoren der Ordnung 2. Das ist die Summe der Determinanten von 6 Untermatrizen der Größe 2×2.
Haben wir schließlich b_2, b_1 und b_0 berechnet, müssen wir die Nullstellen von

$$\tilde{u}(x) := x^3 + b_2 \cdot x^2 + b_1 \cdot x + b_0$$

ermitteln. Dazu können wir nun die Formeln für p, q aus 7.3.8 auf das normierte Polynom $\tilde{u}$ anwenden. Wir erraten also eine Nullstelle x_1 von $\tilde{u}$, und berechnen p und q. Die restlichen beiden Nullstellen ermitteln wir dann mit der pq-Formel.

Soviel also zum allgemeinen Vorgehen. Ganz schön aufwändig! Sollte man sich aber überhaupt nicht behelfen können, kann man wie oben beschrieben vorgehen. Aber wie gesagt: Die meisten 4×4 Matrizen in Aufgaben haben eine angenehme bzw. sehr einfache Gestalt.

Hinweis 7.3.13. Wesentlich einfacher wird die Sache, wenn man zu Beginn anstatt eines Eigenwertes zwei oder mehr erraten kann. Wenn man Glück hat, ist der erratene Eigenwert λ_1 ein doppelter Eigenwert. Das kann z.B. über den Rang von A_{λ_1} bestimmt werden. Diesen Rang erhält man sowieso bei der Eigenvektorberechnung zu λ_1 „geschenkt“. Es könnte also durchaus günstig sein, die Eigenvektorberechnung zu λ_1 noch vor der Berechnung der restlichen Eigenwerte durchzuführen, damit man den Rang von A_{λ_1} kennt. Man macht sich keine zusätzliche Arbeit, könnte sich aber Arbeit ersparen!

Hat man zwei verschiedene (oder einen doppelten) Eigenwert gefunden, können die restlichen beiden durch die Bedingungen

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= tr(A) \\ \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdot \lambda_4 &= \det(A) \end{aligned}$$

eindeutig bestimmt werden.

Beispiel 33. Betrachten wir die Matrix

$$A = \begin{pmatrix} 4 & 1 & 0 & 0 \\ -3 & 1 & 1 & 1 \\ 0 & 0 & 4 & 1 \\ 1 & 0 & -3 & 0 \end{pmatrix}.$$

Gesucht sind Eigenwerte und Eigenvektoren.
Wenn wir 3 von den Diagonalelementen der Matrix abziehen, sind die letzten beiden Spalten identisch. Somit haben wir einen Eigenwert $\lambda_1 = 3$ gefunden. Leider können wir nichts über die (algebraische) Vielfachheit dieses Eigenwertes aussagen, denn $rg(A_{\lambda_1}) = 3$.
Wir berechnen also die Determinante, die Spur und die Summe der 2×2 Hauptminoren der Matrix A. Wir erhalten

$$\begin{aligned} \det(A) &= 4 \cdot 1 \cdot 3 - 1 \cdot ((-3) \cdot 3 + 1 \cdot (-3)) = 24 \\ tr(A) &= 4 + 1 + 4 + 0 = 9 \\ tr(\Lambda^2 \ A) &= 7 + 3 + 4 + 0 + 16 + 0 = 30. \end{aligned}$$

Wir ermitteln

$$b_2 = 3 - tr(A) = -6$$
$$b_1 = tr(\Lambda^2\ A) + b_2 \cdot 3 = 12$$
$$b_0 = -\det(A) \cdot \frac{1}{3} = -8.$$

Somit müssen wir die Nullstellen von

$$\tilde{u}(x) = x^3 - 6 \cdot x^2 + 12 \cdot x - 8$$

berechnen. Durch Einsetzen erraten wir die Nullstelle $\lambda_2 = 2$. Nun wenden wir 7.3.8 auf $\tilde{u}$ an:

$$p = b_2 + \lambda_2 = -4 \quad , \quad q = -\frac{b_0}{\lambda_2} = 4.$$

Jetzt nur noch die pq- Formel anwenden:

$$\lambda_{34} = 2 \pm \sqrt{4-4} = 2.$$

Jetzt haben wir alle Eigenwerte und müssen uns um die Eigenvektoren kümmern.
Wir könnten die Kurschreibweise für die A_λ verwenden, bei größeren Matrizen ist mir das aber zu unübersichtlich. Also lieber langsamer:

$$A_{\lambda_1} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -3 & -2 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & -3 & -3 \end{pmatrix} \quad \text{und} \quad A_{\lambda_2} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -3 & -1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 1 & 0 & -3 & -2 \end{pmatrix}.$$

Beide Matrizen haben Rang 3, was man leicht an einer passenden Unterdeterminante sehen kann (z.B. die linke untere 3×3 Untermatrix hat Determinante $\neq 0$). Wählen wir einfach 3 linear unabhängige Zeilen der Matrizen aus und berechnen das Kreuzprodukt:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ -3 \\ -3 \end{pmatrix} = \begin{pmatrix} +0 \\ -0 \\ +1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

$$\vec{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ -3 \\ -2 \end{pmatrix} = \begin{pmatrix} +(-1) \\ -(2 \cdot (-1)) \\ +1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 1 \\ -2 \end{pmatrix}$$

Aufgrund der vielen 0-Einträge sind die Kreuzprodukte schnell berechnet. Somit haben wir auch diese Aufgabe abgeschlossen.

Nun kommen wir noch zu einem Beispiel, das zwar auf den ersten Blick furchteinflößend wirkt, doch sehr schnell und einfach gelöst werden kann.

Beispiel 34. Betrachten wir die Matrix

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

wobei a und $b \neq 0$ irgendwelche komplexen oder reellen Werte sind. Gesucht sind Eigenwerte und Eigenvektoren.

Bei diesem Beispiel sollte man nun wirklich nicht mehr mit dem charakteristischen Polynom herumhantieren. Addiert man zu den Diagonalelementen von A den Wert $-a+b$ erhält man eine Matrix, in der alle Zeilen gleich sind. Somit ist $\lambda_1 = a - b$ ein Eigenwert. Und nicht nur ein einfacher, sondern mindestens ein dreifacher! Denn die Matrix A_{λ_1} besitzt Rang 1, somit ist die geometrische Vielfachheit von λ_1 gleich 3. Und die algebraische Vielfachheit ist ja bekanntlich immer größer oder gleich der geometrischen. Deshalb kommt der Wert $a - b$ als Eigenwert sicher schon einmal drei Mal vor.

Den letzten Eigenwert können wir nun über die Bedingung

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = tr(A) = 4 \cdot a$$

ermitteln. Wir erhalten also

$$3 \cdot (a - b) + \lambda_4 = 4 \cdot a \quad \Rightarrow \quad \lambda_4 = a + 3 \cdot b$$

Zwischen dem Wert λ_4 und λ_1 besteht Gleichheit, falls $b = 0$. Diesen Fall haben wir aber zu Beginn ausgeschlossen. Somit ist λ_4 tatsächlich ein neuer Eigenwert und der Eigenwert λ_1 kommt „nur“ dreimal vor.

Nun zu den Eigenvektoren. Da A_{λ_1} Rang 1 besitzt können wir die Eigenvektoren über „Vertauschen und Vorzeichenwechsel“ ermitteln. Dabei skalieren wir die Vektoren gleich um den Faktor $1/b$.

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Da λ_4 ein einfacher Eigenwert ist, hat A_{λ_4} den Rang 3. Skalieren wir diese Matrix um den Faktor $1/b$, das ändert ja nichts an der Lösung des homogenen LGS.

$$\frac{1}{b} \cdot A_{\lambda_4} = \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix},$$

Zur Berechnung des Eigenvektors könnte man nun wieder das Kreuzprodukt zu Hilfe nehmen. Besser und auch schneller: Man erkennt, dass

$$\vec{v}_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

eine Lösung ist.

7.4 Das Erraten von Polynomnullstellen

Im Extremfall können wir das charakteristische Polynom explizit aufstellen, indem wir die Koeffizienten berechnen. Das bringt uns aber nicht viel, wenn wir dann keine Nullstelle erraten können. Dieses Kapitel soll das Erraten von Polynomnullstellen etwas erleichtern.
Das nachfolgende Resultat zeigt, dass bei Polynomen mit ganzzahligen Koeffizienten nur ganz bestimmte rationale Nullstellen in Frage kommen. Anstatt willkürlich und ohne Idee in das Polynom einzusetzen, brauchen wir also nur zu überprüfen, ob es sich bei diesen speziellen rationalen Werten um Nullstellen handelt.

Satz 7.4.1. *Sei p ein Polynom vom Grad n mit **ganzzahligen** Koeffizienten. Das heißt das Polynom hat folgende Gestalt*

$$p(x) = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0,$$

wobei die a_i für $i = 1, \ldots, n$ ganzzahlig sind.

***Falls** eine rationale Nullstelle x_0 des Polynoms p existiert, dann befindet sie sich in folgender Menge Q:*

$$Q = \left\{ \frac{a}{b} \;\middle|\; a \text{ teilt } a_0,\; b \text{ teilt } a_n \right\}. \tag{7.18}$$

Beweis. Sei x_0 also eine rationale Nullstelle des Polynoms p. Durch Kürzen können wir diese rationale Nullstelle folgendermaßen darstellen

$$x_0 = \frac{c}{d} \quad \text{mit} \quad ggT(c, d) = 1.$$

Als Anschauungsbeispiel: Statt $\frac{2}{4}$ würden wir $\frac{1}{2}$ schreiben, was zu ersterem äquivalent ist und $ggT(1, 2) = 1$.
Nun gilt

$$0 = p(x_0) = \sum_{i=0}^{n} a_i \cdot {x_0}^i = \sum_{i=0}^{n} a_i \cdot \frac{c^i}{d^i}.$$

Diese Gleichheit multiplizieren wir auf beiden Seiten mit $d^n \neq 0$ und erhalten:

$$0 = \sum_{i=0}^{n} a_i \cdot c^i d^{n-i} = a_0 \cdot d^n + \sum_{i=1}^{n} a_i \cdot c^i d^{n-i} = \sum_{i=0}^{n-1} a_i \cdot c^i d^{n-i} + a_n \cdot c^n.$$

Bei den rechten beiden Gleichheitszeichen haben wir die Summe einfach umgeschrieben, indem wir einmal den ersten Term und beim zweiten Mal den letzten Term „extrahiert“ haben. Wir formen um und erhalten die beiden Gleichheiten

$$-a_0 \cdot d^n = c \cdot \sum_{i=1}^{n} a_i \cdot c^{i-1} d^{n-i} \quad \text{und} \quad -a_n \cdot c^n = d \cdot \sum_{i=0}^{n-1} a_i \cdot c^i d^{n-i-1}.$$

Um den Beweis zu verstehen, ist nun folgendes wichtig: Die Werte c^{i-1} sind im Bereich $i = 1, \ldots, n$ sicher ganzzahlig, gleiches gilt für die Werte d^{n-i-1} im Bereich $i = 0, \ldots, n-1$. Somit sind auch die beiden Summen ganze Zahlen. Hieraus folgt jetzt

$$c \quad \text{teilt} \quad a_0 \cdot d^n \quad \text{sowie} \quad d \quad \text{teilt} \quad a_n \cdot c^n.$$

Da $ggT(c, d) = 1$ dürfen die beiden ganzen Zahlen c und d keine gemeinsamen Primfaktoren besitzen. Wegen den obigen Teilbarkeitsaussagen sind somit die Primfaktoren von c in a_0 enthalten und die Primfaktoren von d in a_n. Und daraus folgt nun

$$c \quad \text{teilt} \quad a_0 \quad \text{sowie} \quad d \quad \text{teilt} \quad a_n$$

und folglich $x_0 \in Q$. □

Beispiel 35. Wir betrachten das folgende Polynom

$$p(x) = 3 \cdot x^4 - 6 \cdot x^3 - 21 \cdot x^2 - 11 \cdot x - 4$$

und wollen alle rationalen Nullstellen bestimmen. Da das Polynom ganzzahlige Koeffizienten besitzt, können wir das obige Resultat anwenden. Falls überhaupt eine rationale Nullstelle existiert, liegt sie in der Menge

$$Q = \left\{ \frac{a}{b} \;\middle|\; a \text{ teilt } a_0,\; b \text{ teilt } a_n \right\}.$$

Berechnen wir also die Menge Q. Wir benötigen dazu alle Teiler von $a_0 = -4$ und $a_n = 3$. Man beachte, dass wir positive und **negative** Teiler betrachten müssen. Somit erhalten wir

$$a \in \{\pm 1, \pm 2, \pm 4\} \quad \text{und} \quad b \in \{\pm 1, \pm 3\}.$$

Folglich ist

$$Q = \left\{ \pm 1, \pm 2, \pm 4, \pm\frac{1}{3}, \pm\frac{2}{3}, \pm\frac{4}{3} \right\}.$$

Das sind jetzt unsere Kandidaten für rationale Nullstellen. Diese Werte setzen wir nun in das Polynom ein und überprüfen, ob es Null wird. Wir sehen, dass nur $p(4) = 0$ gilt. Somit ist 4 die einzige rationale Nullstelle.

Besonders freundlich wird das Resultat, wenn ein normiertes Polynom vorliegt, das heißt, falls der Leitkoeffizient gleich 1 ist.

Korollar 7.4.2. *Sei p ein* ***normiertes*** *Polynom vom Grad n mit* ***ganzzahligen*** *Koeffizienten. Das heißt das Polynom hat folgende Gestalt*

$$p(x) = x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0,$$

Falls *eine rationale Nullstelle x_0 des Polynoms p existiert, dann befindet sie sich in folgender Menge Q:*

$$Q = \left\{ a \;\middle|\; a \text{ teilt } a_0 \right\}. \tag{7.19}$$

Insbesondere ist jede rationale Nullstelle eine ganze Zahl.

Beispiel 36. Gegeben sei das Polynom

$$p(x) = x^5 + x + 2.$$

Offensichtlich ist das Polynom normiert und besitzt ganzzahlige Koeffizienten. Falls eine rationale Nullstelle existieren sollte, ist sie ganzzahlig und ein Teiler von $a_0 = 2$. Die Teiler von 2 sind schnell bestimmt und somit liegen die Kandidaten für ganzzahlige Nullstellen in der Menge

$$Q = \{\pm 1, \pm 2\}.$$

Setzen wir die Werte aus Q in das Polynom ein, sehen wir, dass nur $p(-1) = 0$ gilt. Daher ist -1 die einzige ganzzahlige Nullstelle.
Es gilt sogar: Der Wert -1 ist die einzige reelle Nullstelle. Denn wegen

$$p'(x) = 5 \cdot x^4 + 1 > 0$$

ist p eine streng monoton wachsende Funktion. Sie schneidet die x-Achse daher nur ein einziges Mal und tendiert dann gegen $+\infty$.

Was soll ich tun, wenn es keine rationalen bzw. ganzzahligen Nullstellen gibt?

Das ist eine berechtigte Frage! Es kann nämlich durchaus passieren, dass man alle Werte aus Q in das Polynom einsetzt und trotzdem keine Nullstelle dabei war. Das bedeutet dann, dass es keine rationalen Nullstellen gibt! Denn gäbe es sie, würden sie ja in Q liegen.
Man muss sich dann also im Irrationalen oder Komplexen nach Nullstellen umsehen. Häufige Kandidaten für Nullstellen sind aus meiner Erfahrung Werte wie $\pm\sqrt{2}, \pm\sqrt{3}, \pm\mathrm{i}$. Die kann man ja einmal ausprobieren ...

Man braucht sich aber nun nicht zu fürchten. Die meisten Beispiele, die einem in Übungen und Prüfungen begegnen, sind so konstruiert, dass sie „schön" lösbar sind. Das heißt, es gibt zumindest eine rationale bzw. ganzzahlige Nullstelle, die z.B. mit obiger Methode auffindbar ist. Es macht nämlich keinen Sinn, Leute nach irgendwelchen irrationalen Werten suchen zu lassen – da kann man nämlich (ohne System) sehr lange suchen!
Außerdem sind wir hier nicht in der Numerik. Dort gibt es Methoden, die Nullstellen von Polynomen und anderen Funktionen zu ermitteln. Die wendet aber im Normalfall der Computer an, der ist nämlich schneller und effizienter als wir Menschen.

Das Problem ist nur: Der Mensch macht Fehler. Es kann durch hastiges Erstellen von Angaben passieren, dass sich Fehler in die Angabe einschleichen. Aus einer 1 wird eine 7 usw. Blöd nur, wenn diese fehlerhaften Werte das Polynom beeinflussen. Dann sind die Nullstellen keine „schönen" Werte mehr.
Natürlich kann es auch sein, dass der Ersteller der Angabe ein Sadist ist und es den Leuten nicht zu leicht machen will. So etwas kann man nie ausschließen.

Aus diesem Grund ist es wohl keine schlechte Idee, sich anhand eines Beispieles anzuschauen, wie man vorgehen könnte, wenn keine rationalen Nullstellen existieren.

Beispiel 37. Wir betrachten das Polynom

$$p(x) = x^3 - x + 2.$$

Mögliche rationale Nullstellen liegen in der Menge

$$Q = \{\pm 1, \pm 2\}.$$

Nun gilt aber (leider) $p(x) \neq 0$ für jedes $x \in Q$. Das bedeutet, dass das Polynom p keine rationalen Nullstellen besitzt.
Da das Polynom Grad 3 aufweist, gibt es sicher eine reelle Nullstelle. Machen wir uns also auf die Suche nach einer irrationalen Nullstelle. Oder besser: Wir versuchen sie anzunähern.
Wir verwenden dazu das Newton-Verfahren: Hierfür müssen wir folgende Iteration genügend oft ausführen

$$x_{n+1} = x_n - \frac{p(x_n)}{p'(x_n)}.$$

Den Startwert x_0 sollten wir nah an der zu ermittelnden Nullstelle wählen, damit garantiert ist, dass das Verfahren auch wirklich gegen die Nullstelle konvergiert. Um x_0 zu wählen, müssen wir also ungefähr wissen, wo eine Nullstelle der Funktion p liegt.
Das ist aber nicht schwer! Denn es gilt

$$p(-2) = -4 \quad \text{sowie} \quad p(-1) = 2.$$

Da die Funktion stetig ist, das heißt keine Sprünge aufweist, muss sie im Bereich $[-2, -1]$ die x-Achse schneiden. Denn p ist zuerst negativ und dann positiv. Und auf ihrem Weg vom Negativen ins Positive muss sie die x-Achse schneiden. Somit liegt im Bereich $[-2, -1]$ eine Nullstelle.
Versuchen wir als Startwert also einfach einmal $x_0 = -1$. Für das Verfahren benötigen wir nun noch die Ableitung von p:

$$p'(x) = 3 \cdot x^2 - 1 \quad \text{und daher} \quad p'(-1) = 2.$$

Setzen wir nun in die Iterationsvorschrift ein

$$x_1 = x_0 - \frac{p(x_0)}{p'(x_0)} = -1 - \frac{2}{2} = -2.$$

Wegen $p(x_1) = -4$ und $p'(x_1) = 11$ gilt

$$x_2 = x_1 - \frac{p(x_1)}{p'(x_1)} = -2 - \frac{(-4)}{11} = -\frac{18}{11}$$

Machen wir noch einen Iterationsschritt. Es gilt $p(x_2) = -\frac{992}{11^3}$ und $p'(x_2) = \frac{851}{11^2}$. Einsetzen liefert

$$x_3 = x_2 - \frac{p(x_2)}{p'(x_2)} = -\frac{18}{11} + \frac{992}{11 \cdot 851} = -\frac{14326}{9361} \approx -1,53$$

Diesen angenäherten Wert benutzen wir nun, um p, q zu ermitteln

$$p = a_2 + x_3 = x_3 = -1,53 \quad \text{und} \quad q = -\frac{a_0}{x_3} = \frac{2}{1,53} \approx 1,31.$$

Die pq-Formel liefert für die restlichen beiden Nullstellen u_{12}:

$$u_{12} = 0,765 \pm \sqrt{-0,725} = 0,765 \pm 0,851 \text{ i}.$$

Ein Computerprogramm ermittelt für die Nullstellen des Polynoms

$$u_{12} = 0,760 \pm 0,858 \text{ i} \quad \text{und} \quad u_3 = -1,52.$$

Wie man sieht kann das Verfahren ziemlich aufwändig werden, wenn man per Hand rechnet. Hat man dagegen einen einfachen Taschenrechner zur Verfügung, kann man Nullstellen relativ schnell annähern. Es kann z.B. folgendermaßen vorgegangen werden

1. Schritt Wähle einen Startwert x_0, der „nahe“ an der zu ermittelnden Nullstelle liegt.

2. Schritt Lege x_0 in den Zwischenspeicher des Taschenrechners

$$\text{Ans} \quad \longleftarrow \quad x_0$$

3. Schritt Berechne die Ableitung f' und tippe folgendes in den Taschenrechner

$$\text{Ans} - \frac{f(\text{Ans})}{f'(\text{Ans})}$$

4. Schritt Betätige mehrfach die = Taste auf dem Taschenrechner. Nähern sich die Ergebnisse einem Wert an, dann ist das die gesuchte Nullstelle x^*.
Falls sie sich keinem Wert nähern, kehre zu Schritt 1 zurück und wähle einen neuen Startwert x_0, der näher an der zu ermittelnden Nullstelle liegt.

7.5 Das Erraten von Eigenwerten über die Determinante

Kehren wir aus der Algebra zurück zum Eigenwertproblem. Wie hilft uns das Teilbarkeitskriterium aus dem letzten Abschnitt beim Erraten von Nullstellen weiter?

Satz 7.5.1. *Sei A eine $n \times n$ Matrix mit* ***ganzzahligen*** *Einträgen.* ***Falls*** *ein rationaler Eigenwert λ existiert, befindet er sich in der folgenden Menge Q:*

$$Q = \left\{ \, a \;\middle|\; a \text{ teilt } \det(A) \right\}. \tag{7.20}$$

Insbesondere ist jeder rationale Eigenwert λ bereits ganzzahlig.

Beweis. Wir wenden einfach Korollar 7.4.2 auf das charakteristische Polynom p_A an. Die Eigenwerte sind ja die Nullstellen von p_A. Zuerst müssen wir natürlich prüfen, ob die Voraussetzungen erfüllt sind, um das Korollar anzuwenden.
Das charakteristische Polynom hat ganzzahlige Koeffizienten. Das folgt aus der Formel 7.2 für die Koeffizienten des charakteristischen Polynoms. Da die Matrix nur ganzzahlige Einträge besitzt und Spur und Determinanten von solchen Matrizen wieder ganze Zahlen sind, folgt, dass die Koeffizienten ganzzahlig sind. Ebenso ist ersichtlich, dass das charakteristische Polynom normiert ist.
Wir können also das Korollar anwenden. Aus 7.3.2 erhalten wir für den konstanten Term a_0 von p_A

$$a_0 = (-1)^n \cdot \det(A).$$

Somit ist jede rationale Nullstelle von p_A eine ganze Zahl und teilt $\det(A)$. □

Wie schon einmal kurz erwähnt, haben die meisten Eigenwertprobleme, die einem in Form von Aufgaben begegnen, eine „schöne“ Lösung. Und mit „schön“ meine ich, dass es einen ganzzahligen Eigenwert gibt. Bei Matrizen mit ganzzahligen Einträgen, kommen neben irrationalen und komplexen Eigenwerten nur noch ganzzahlige in Frage. Und diese teilen die Determinante der Matrix.
Außerdem tritt bei 3×3 Matrizen immer mindestens ein reeller Eigenwert auf. Dieser Eigenwert ist dann entweder ein irrationaler Wert oder ganzzahlig. Meistens wird er ganzzahlig sein. Außerdem ist bei vielen Aufgaben neben der Eigenwertberechnung auch die Eigenvektorberechnung gefragt. Dabei sind irrationale Werte als Eigenwerte auch nicht unbedingt förderlich. Wir halten also fest: In den meisten Fällen existiert (zumindest ein) ganzzahliger Eigenwert. Die Determinante der Matrix ist dabei ein Hilfsmittel, um diesen zu finden bzw. zu erraten.

Beispiel 38. Gesucht sind die Eigenwerte der Matrix

$$A = \begin{pmatrix} -2 & 1 & 3 \\ 1 & -2 & -3 \\ 2 & 1 & 0 \end{pmatrix}.$$

Berechnen wir Spur und Determinante der Matrix:

$$tr(A) = -4 \qquad \det(A) = 2 \cdot 3 + (-1) \cdot 3 = 3.$$

Die möglichen ganzzahligen Eigenwerte teilen die Determinante der Matrix. Das heißt sie liegen in der Menge

$$Q = \{\pm 1, \pm 3\}.$$

Addieren wir 1 zu den Diagonalelementen der Matrix A sind die ersten zwei Zeilen bis auf ein Vorzeichen gleich. Somit ist $\lambda_1 = -1$ ein ganzzahliger Eigenwert.
Berechnen wir p und q:

$$p = -1 + 4 = 3 \qquad q = -\frac{3}{1} = -3.$$

Die restlichen beiden Eigenwerte berchnen wir mit der pq-Formel

$$\lambda_{23} = -\frac{3}{2} \pm \sqrt{\frac{9}{4} + \frac{12}{4}} = \frac{-3 \pm \sqrt{21}}{2}.$$

Was tun, falls die Matrix Brüche als Einträge aufweist?

Das obige Resultat kann dann nicht angewendet werden. Außerdem ist es meistens unangenehm und fehleranfällig mit Brüchen zu rechnen. Glücklicherweise müssen wir das auch nicht! Wir skalieren die Matrix einfach mit einem passenden Faktor und führen dann alle Rechnungen in den ganzen Zahlen durch.
Sei also A eine Matrix mit rationalen Einträgen. Ist ein Eintrag eine ganze Zahl, fassen wir ihn ebenfalls als einen Bruch auf, der im Nenner eine 1 stehen hat. Der Wert n_{ij} beschreibe den Nenner des ij-ten Eintrages der Matrix A. Außerdem sei

$$r = kgV\,(n_{ij} \mid 1 \leqslant i, j \leqslant n)$$

das kleinste gemeinsame Vielfache aller dieser Nenner.
Die Matrix

$$B := r \cdot A$$

hat dann ganzzahlige Einträge. Es gilt folgender Zusammenhang zwischen den Eigenwerten und Eigenvektoren von B und A:

Satz 7.5.2. *Seien A und B wie oben. Ist λ_B ein Eigenwert von B zum Eigenvektor $\vec{v}_B \neq \vec{0}$, dann ist $\lambda_A = \frac{1}{r} \cdot \lambda_B$ ein Eigenwert von A zum Eigenvektor $\vec{v}_A = r \cdot \vec{v}_B$.*

Beweis. Sei also λ_B ein Eigenwert von B zum Eigenvektor $\vec{v}_B$. Es gilt also

$$B\,\vec{v}_B = \lambda_B \cdot \vec{v}_B.$$

Folglich ist

$$A\,\vec{v}_A = r \cdot A\,\vec{v}_B = B\,\vec{v}_B = \lambda_B \cdot \vec{v}_B = \frac{1}{r} \cdot \lambda_B \cdot r \cdot \vec{v}_B = \lambda_A \cdot \vec{v}_A.$$

□

Beispiel 39. Gesucht sind die Eigenwerte der Matrix

$$A = \begin{pmatrix} 1/3 & 1/4 \\ 1/4 & 1/3 \end{pmatrix}.$$

Somit ist

$$r = kgV(3,4) = 12 \quad \text{und} \quad B = r \cdot A = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$$

und $\det(B) = 7$. Die ganzzahligen Eigenwert von B liegen also in

$$Q = \{\pm 1, \pm 7\}.$$

Offensichtlich sind sowohl 1, als auch 7 Eigenwerte von B. Die Eigenvektoren von B berechnen wir mit „Vertauschen und Vorzeichenwechsel“:

$$\vec{v}_B^{\,1} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad \vec{v}_B^{\,2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Folglich sind $1/12$ und $7/12$ die Eigenwerte der Matrix A zu den Eigenvektoren

$$\vec{v}_A^{\,1} = 12 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \Big|\Big| \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \vec{v}_B^{\,1} \qquad \vec{v}_A^{\,2} = 12 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Big|\Big| \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \vec{v}_B^{\,2}.$$

8 Hauptvektoren und Jordan-Normalform

In diesem Abschnitt betrachten wir als zugrunde liegenden Körper immer die komplexen Zahlen. Das machen wir, damit das charakteristische Polynom in Linearfaktoren zerfällt.

8.1 Hauptvektoren

Worum geht es nun? Wie schon einmal zu Beginn des letzten Abschnittes kurz erwähnt, gilt folgender Zusammenhang zwischen geometrischer und algebraischer Vielfachheit eines Eigenwertes λ:

$$1 \leqslant g_\lambda \leqslant a_\lambda.$$

Oder anders formuliert: Zu jedem Eigenwert λ gibt es zumindest einen Eigenvektor. Die Anzahl der linear unabhängigen Eigenvektoren zu einem Eigenwert ist immer kleiner oder gleich der Anzahl, wie oft λ als Eigenwert vorkommt.

Besonders interessant ist der Fall, wo $g_\lambda = a_\lambda$ für jeden Eigenwert λ, gilt. Dann gibt es n linear unabhängige Eigenvektoren, die eine Basis B von V bilden. Man spricht auch von Diagonalisierbarkeit (der Ausdruck kommt daher, dass die Basisdarstellung der linearen Abbildung A bezüglich dieser Basis B Diagonalgestalt aufweist, wobei die Eigenwerte auf der Diagonalen stehen).

Nun kann es aber durchaus passieren, dass die Bedingung $g_\lambda = a_\lambda$ nicht für jeden Eigenwert erfüllt ist. Es kann z.B. sein, dass es zu einem dreifachen Eigenwert λ nur einen linear unabhängigen Eigenvektor gibt. Am Ende haben wir dann weniger als n linear unabhängige Eigenvektoren zur Verfügung und diese bilden sicher keine Basis von V mehr.

Hier kommt der Begriff **Hauptvektor** ins Spiel! Ein Hauptvektor ist ein geschickt gewählter Ersatz für einen fehlenden Eigenvektor, sodass die linear unabhängigen Eigenvektoren zusammen mit den Hauptvektoren doch noch eine Basis B von V bilden. Bei unserem oben gewählten Beispiel zum dreifachen Eigenwert λ würde es neben dem einen Eigenvektor, noch zwei Hauptvektoren geben. Somit gibt es genau so viele linear unabhängige Vektoren, wie der Eigenwert λ vorkommt!

Wie sind diese Hauptvektoren zu wählen? So, dass sie „fast“ Eigenvektoren sind! Sei $\vec{v}$ dazu ein Eigenvektor zu einem Eigenwert λ. Der erste Hauptvektor $\vec{h_1}$ zu $\vec{v}$ erfüllt dann

$$A \cdot \vec{h}_1 = \lambda \cdot \vec{h}_1 + \vec{v} \qquad \text{oder} \qquad A_\lambda \cdot \vec{h}_1 = \vec{v}.$$

Dieses inhomogene GLS ist für den ersten Hauptvektor zu lösen. Der Eigenvektor $\vec{v}$ hängt hierbei von freien Parametern ab. Diese müssen so gewählt werden, dass das System lösbar ist!
Der zweite Hauptvektor $\vec{h}_2$ muss so gewählt werden, dass

$$A \cdot \vec{h}_2 = \lambda \cdot \vec{h}_2 + \vec{h}_1 \qquad \text{oder} \qquad A_\lambda \cdot \vec{h}_2 = \vec{h}_1.$$

Freie Parameter sind wieder so zu wählen, dass das System lösbar ist.

Das wiederholt man solange, bis die Anzahl der Eigenvektoren plus die Anzahl der Hauptvektoren genauso groß ist, wie die Vielfachheit des Eigenwertes λ.

Das Einzige, das also zur Berechnung von Hauptvektoren zu tun ist, ist inhomogene GLS zu lösen. Dafür können wir z.B. das Kreuzprodukt zu Hilfe nehmen.

Beispiel 40. Gesucht sind Eigenwerte, Eigenvektoren und Hauptvektoren zur Matrix

$$A = \begin{pmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Die Eigenwerte sind schnell berechnet, denn die Matrix hat eine freundliche Blockgestalt. Der erste Eigenwert ist $\lambda_1 = 2$. Die restlichen beiden erraten wir über die Bedingungen

$$\lambda_2 + \lambda_3 = 8$$

$$\lambda_2 \cdot \lambda_3 = 16$$

Daraus erhalten wir $\lambda_2 = \lambda_3 = 4$.
Es gilt

$$A_{\lambda_2} = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Wir berechnen einen Eigenvektor, indem wir das Kreuzprodukt der letzten beiden Zeilen ausrechnen. Wir erhalten

$$\vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Der Eigenwert 4 ist ein doppelter Eigenwert. Es gibt aber nur einen l.u. Eigenvektor zu diesem Eigenwert. Somit muss es einen Hauptvektor $\vec{h}_1^{\,2}$ geben.
Wir müssen also

$$A_{\lambda_2} \cdot \vec{h}_1^2 = t \cdot \vec{v}_2$$

lösen, wobei t so zu wählen ist, dass das System lösbar wird. Wir müssen also folgendes lösen

$$\begin{pmatrix} 2 & 2 & 2 \\ -2 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \cdot \vec{h}_1^{\,2} = t \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Da die Matrix Rang 2 aufweist, brauchen wir für eine Partikulärlösung nur eine passende 2×2 Untermatrix finden. Die rechte untere Untermatrix ist ein guter Kandidat:

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \cdot \vec{\tilde{x}}_p = \begin{pmatrix} -t \\ 0 \end{pmatrix}$$

Deshalb ist eine Partikulärlösung gegeben durch

$$\vec{x}_p = \begin{pmatrix} 0 \\ t/2 \\ 0 \end{pmatrix}$$

Den Parameter t ermittelt man, indem man das in die erste Zeile von A_{λ_2} einsetzen. So bemerken wir, dass wir t beliebig wählen können. Nehmen wir also $t = 2$. Daher ist

$$\vec{h}_1^{\,2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{Hauptvektor zu} \quad \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}.$$

Wir brauchen noch den Eigenvektor $\vec{v}_1$ zum einfachen Eigenwert $\lambda_1 = 2$. Wir benutzen das Kreuzprodukt und erhalten

$$\vec{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Wie man an dem letzten Beispiel sieht, ist es für kleine Matrizen leicht, die Hauptvektoren zu ermitteln. In diesem Beispiel war es egal, wie der freie Parameter t gewählt wird. Das System $A_{\lambda_2} \cdot \vec{h}_1^{\,2} = t \cdot \vec{v}_2$ war immer lösbar. Das ist nicht immer der Fall. Machen wir noch ein Beispiel, bei dem die Parameter nicht beliebig gewählt werden können.

Beispiel 41. Gegeben sei die Matrix

$$A = \begin{pmatrix} 3 & 0 & 0 & -1 \\ 1 & 3 & -1 & 0 \\ 2 & 0 & 2 & -2 \\ 1 & -1 & 1 & 0 \end{pmatrix}.$$

Gesucht sind Eigen- und Hauptvektoren.

Man erkennt sehr schnell, dass $\lambda_1 = 2$ ein Eigenwert der Matrix ist. Es gilt

$$A_{\lambda_1} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 2 & 0 & 0 & -2 \\ 1 & -1 & 1 & -2 \end{pmatrix}.$$

Berechnen wir mit Hilfe des Kreuzprodukts die Eigenvektoren. Benutzen wir dazu die ersten zwei Zeilen der Matrix A_{λ_1}, sie sind linear unabhängig.

$$\begin{bmatrix} 1 & 1 & & \\ 0 & 1 & & \\ 0 & -1 & 1 & \\ -1 & 0 & & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Jeder Eigenvektor ist Linearkombination dieser beiden berechneten Vektoren. Setzen wir die beiden Vektoren in die letzten zwei Zeilen der Matrix A_{λ_1} ein, so erhalten wir Null. Somit sind die beiden berechneten Vektoren bereits die l.u. Eigenvektoren $\vec{v}_1, \vec{v}_2$ zu λ_1. Außerdem ist λ_1 mindestens ein doppelter Eigenwert, d.h. $\lambda_1 = \lambda_2 = 2$.
Die restlichen beiden Eigenwerte berechnen wir über die Bedingungen

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= tr(A) \\ \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdot \lambda_4 &= \det(A). \end{aligned}$$

Es gilt
$$tr(A) = 8 \quad \text{und} \quad \det(A) = 12 + 4 = 16.$$

Folglich ist $\lambda_1 = 2$ ein vierfacher Eigenwert und es muss zwei Hauptvektoren geben.

Sei $\vec{v}$ nun irgendeine Eigenvektor von A, d.h. es gilt $\vec{v} = t \cdot \vec{v}_1 + s \cdot \vec{v}_2$. Gesucht ist ein Hauptvektor $\vec{h}_1$, sodass
$$A_{\lambda_1} \, \vec{h}_1 = \vec{v}.$$
erfüllt ist. Wir müssen also ein Partikulärlösung von

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 2 & 0 & 0 & -2 \\ 1 & -1 & 1 & -2 \end{pmatrix} \cdot \vec{h}_1 = \begin{pmatrix} t \\ s-t \\ s \\ t \end{pmatrix}$$

ermitteln, wobei s, t so zu wählen sind, dass auch eine Lösung existiert. Da die Matrix A_{λ_1} Rang 2 aufweist, suchen wir uns eine reguläre 2×2 Untermatrix aus. Wir können z.B.

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \vec{x} = \begin{pmatrix} t \\ s-t \end{pmatrix}$$

lösen. Es folgt $x_2 = s - t$ und $x_4 = -t$. Somit gilt

$$\vec{x}_p = (0, s-t, 0, -t).$$

Um Lösbarkeit zu garantieren, MUSS dieser Vektor $\vec{x}_p$ auch die letzten beiden Zeilen des obigen inhomogenen Systems erfüllen! Hieraus erhalten wir die Bedingung $2t = s$.

Wir können also z.B. $t = 1, s = 2$ wählen und erhalten

$$\vec{v} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{mit Hauptvektor} \quad \vec{h}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} + k \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + r \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Man beachte: Wir haben die allgemeine Form des Hauptvektors (Partikulärlösung plus homogene Lösung) verwendet. Das ist unbedingt notwendig, denn es gibt ja noch einen zweiten Hauptvektor $\vec{h}_2$. Dieser wird bestimmt über die Bedingung

$$A_{\lambda_1} \vec{h}_2 = \vec{h}_1.$$

Die freien Parameter k und r müssen so gewählt werden, dass das obige System lösbar ist. Machen wir das:

Gesucht ist also ein Vektor $\vec{h}_2$, der

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 2 & 0 & 0 & -2 \\ 1 & -1 & 1 & -2 \end{pmatrix} \cdot \vec{h}_2 = \begin{pmatrix} k \\ 1 + r - k \\ r \\ k - 1 \end{pmatrix}$$

erfüllt. Wir ermitteln wieder eine Partikulärlösung mit Hilfe einer regulären Untermatrix. Um es uns leicht zu machen, nehmen wir die gleiche Untermatrix wie vorher. Dieses mal erhalten wir $x_2 = 1 + r - k$ und $x_4 = -k$ und deshalb

$$\vec{x_p} = (0, 1 + r - k, 0, -k).$$

Auch hier muss dieser Vektor die letzten beiden Zeilen erfüllen, um Lösbarkeit zu garantieren. Wir erhalten die Bedingung $2k = r$. Wir können also $k = 1$ und $r = 2$ wählen. Als abschließendes Ergebnis erhalten wir somit

$$\vec{v} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{mit Hauptvektoren} \quad \vec{h}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{h}_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ -1 \end{pmatrix}.$$

Jetzt haben wir einen Eigenvektor und zwei Hauptvektoren. Hauptvektoren gibt es also keine mehr. Dafür noch einen Eigenvektor, der l.u. zu $\vec{v}$ ist, da die geometrische Vielfachheit von λ_1 gleich 2 ist. Nehmen wir dafür einfach

$$\vec{v}_2 = (0, 1, 1, 0)$$

her, der ist sicher l.u. zu $\vec{v}$. Jetzt haben wir zwei Eigenvektoren und zwei Hauptvektoren und sind fertig.

Und hier haben wir nun gesehen, dass das Berechnen von Hauptvektoren bei etwas größeren

Matrizen zu einer langwierigen und lästigen Aufgabe werden kann.

8.2 Die Jordan-Normalform

Ist eine lineare Abbildung A diagonalisierbar, so gibt es eine Basis B bestehend aus Eigenvektoren. Das ist genau dann der Fall, wenn $g_\lambda = a_\lambda$ für jeden Eigenwert λ erfüllt ist. Wie schon einmal kurz am Rande erwähnt, kommt der Begriff „Diagonalisierbar“ daher, dass die Basisdarstellung der Abbildung A bezüglich der Basis B Diagonalgestalt aufweist. Wer mit Basisdarstellungen nichts am Hut hat, der merke sich folgendes zur Diagonalisierbarkeit: Diagonalisierbarkeit bedeutet im Grunde, dass es eine invertierbare $n \times n$ Matrix X gibt, sodass

$$X^{-1}\, A\, X = D \tag{8.1}$$

erfüllt ist. Die Matrix D ist hierbei die Diagonalmatrix, die die Eigenwerte von A auf der Diagonalen stehen hat (mit ihren Vielfachheiten gezählt). Auch die Matrix X können wir (im Falle der Diagonalisierbarkeit von A) explizit angeben. Die Spalten von X sind die n linear unabhängigen Eigenvektoren von A. Sollte also jemand nach dieser Matrix X fragen, die die Matrix A auf Diagonalgestalt transformiert, brauchen wir nur die Eigenvektoren in eine Matrix packen und sind fertig.

Das Problem ist nun: Nicht jede Matrix ist diagonalisierbar, das heißt es gibt Matrizen, die nicht auf Diagonalgestalt transformiert werden können. Das ist dann der Fall, wenn man keine n linear unabhängigen Eigenvektoren zur Verfügung hat. Oder anders formuliert, falls $g_\lambda < a_\lambda$ für einen Eigenwert λ.

Die Mathematiker haben sich also folgendes überlegt: „Falls wir die Matrix schon nicht auf eine Diagonalgestalt transformieren können, dann zumindest auf eine Gestalt, die fast diagonal ist!“ Und das ist auch schon die grundsätzliche Idee hinter der Jordan-Normalform.

Aber wie ist nun die Transformationsmatrix X zu wählen, falls A nicht diagonalisierbar ist, das heißt die Eigenvektoren nicht ausreichen? Und was bedeutet „fast diagonal“ ? Die erste Frage ist schnell beantwortet, die zweite wird uns noch etwas beschäftigen.

Die Transformationsmatrix X

Im Falle der Nicht-Diagonalisierbarkeit der Matrix A haben wir keine n linear unabhängige Eigenvektoren zur Verfügung. Wir brauchen für die fehlenden Eigenvektoren einen Ersatz! Und hier kommen die Hauptvektoren ins Spiel. Denn zusammen mit den Hauptvektor haben wir wieder n linear unabhängige Vektoren zur Verfügung! Da es also nicht möglich ist, die Matrix X vollständig mit Eigenvektoren zu „füllen“, packen wir neben den Eigenvektoren auch noch ihre zugehörigen Hauptvektoren in die Matrix X. Dabei sollten wir eine gewisse Reihenfolge zwischen Eigen- und Hauptvektoren einhalten: Es kommen immer zuerst die Eigenvektoren, dann die zugehörigen Hauptvektoren, falls es welche gibt. Zur besseren Veranschaulichung, findet sich unten noch ein Bild, wie die Matrix X korrekt zu befüllen ist:

$$X = \begin{bmatrix} | & | & | & & | & | & & | & | & | & \\ \vec{v}_1 & \vec{h}_1{}^1 & \vec{h}_2{}^1 & \bullet\bullet\bullet & \vec{v}_i & \vec{h}_1{}^i & \bullet\bullet\bullet & \vec{v}_k & \vec{h}_1{}^k & \vec{h}_2{}^k & \bullet\bullet\bullet \\ | & | & | & & | & | & & | & | & | & \end{bmatrix}$$

$\vec{v}_i \quad i = 1, \ldots, k \quad$ sind die k l.u. Eigenvektoren

$\vec{h}_j{}^i \quad$ ihre zugehörigen Hauptvektoren.

Abbildung 3: Zur Befüllungsreihenfolge: Zuerst der Eigenvektor, dann seine zugehörigen Hauptvektoren.

Transformieren wir die Matrix A mit Hilfe der Matrix X, erhalten wir eine Matrix J:

$$J = X^{-1}\ A\ X. \tag{8.2}$$

Diese Matrix J ist die gesuchte „fast diagonale“ Matrix. Sie wird auch **Jordan-Normalform** der Matrix A genannt. Wie diese genau aussieht und wie wir sie effizient berechnen können, werden wir später sehen.
Im Moment halten wir nur fest: Ist der zugrunde liegende Körper der komplexe Zahlenkörper, so kann jede $n \times n$ Matrix auf Jordan-Normalform transformiert werden!

Beispiel (Fortsetzung von 40) 42. Die zwei Eigenvektoren und den einen Hauptvektor haben wir bereits berechnet:

$$\vec{v}_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} \quad \vec{h}_1{}^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Gesucht ist nun die Transformationsmatrix X, die unsere Matrix A auf Normalform bringen soll. Wir packen also die berechneten Vektoren in eine Matrix und beachten dabei die Reihenfolge der Vektoren: Nach einem Eigenvektor sollten immer seine zugehörigen Hauptvektoren stehen.

$$X = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \vec{h}_1{}^2 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ 1 & -2 & 1 \\ -1 & 0 & 0 \end{pmatrix}.$$

Invertieren wir die Matrix mit Hilfe des Kreuzprodukts:

$$X^T = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \mapsto \quad -\frac{1}{2} \cdot \begin{bmatrix} 0 & 0 & 2 \\ -1 & 0 & 0 \\ -2 & -2 & -2 \end{bmatrix} = X^{-1}$$

Sollte man nicht mehr genau wissen, wie das geht, kann man sich den jeweiligen Abschnitt nochmal durchlesen.
Um nun die Jordan-Normalform J zu erhalten, müssen wir folgendes Matrixprodukt berechnen

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

Das ist also die vorher angesprochene „fast diagonale“ Matrix. Wir erinnern uns: Die Eigenwerte der Matrix waren $\lambda_1 = 2$ und $\lambda_2 = \lambda_3 = 4$. Diese Werte stehen auf der Diagonalen von J. Das Einzige, das J von einer Diagonalmatrix unterscheidet, ist die 1 auf der Nebendiagonalen.

Allgemeine Struktur der Jordan–Normalform J

Dass die Matrix J im Beispiel so aussieht, wie sie aussieht, ist kein Zufall. Die Jordan–Normalform einer beliebigen Matrix A hat eine solche Struktur: Auf der Hauptdiagonalen stehen die Eigenwerte (mit ihren Vielfachheiten gezählt) und auf der oberen Nebendiagonalen steht entweder 1 oder 0. Alle anderen Einträge der Matrix J sind $= 0$.
Schauen wir uns den allgemeinen Aufbau einer solchen Jordan–Normalform etwas genauer an. Dieser weist mehrere „Level“ auf.

Level 1 Die Jordan–Normalform J zerfällt in sogenannte **Jordankästchen** K_{λ_i}. Von diesen gibt es genau so viele, wie es (verschiedene!) Eigenwerte gibt. Seien also $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte der Matrix A. Dann gilt

$$J = \begin{pmatrix} K_{\lambda_1} & & & \\ & K_{\lambda_2} & & \\ & & \ddots & \\ & & & K_{\lambda_k} \end{pmatrix}. \tag{8.3}$$

Ein Jordankästchen K_{λ_i} ist hierbei eine $a_{\lambda_i} \times a_{\lambda_i}$ Matrix. Das heißt, das Kästchen ist genauso groß, wie die Anzahl, die λ_i als Eigenwert vorkommt.

Level 2 (Aufbau eines Jordankästchens) Die Jordankästchen K_{λ_i} zerfallen in sogenannte **Jordanblöcke** $B_{\lambda_i}^{l_j}$. Es gilt

$$K_{\lambda_i} = \begin{pmatrix} B_{\lambda_i}^{l_1} & & & \\ & B_{\lambda_i}^{l_2} & & \\ & & \ddots & \\ & & & B_{\lambda_i}^{l_{g_i}} \end{pmatrix}. \tag{8.4}$$

Die Zahl $l_j \geqslant 1$ beschreibt die Länge des jeweiligen Jordanblockes $B_{\lambda_i}^{l_j}$, das heißt $B_{\lambda_i}^{l_j}$ ist eine $l_j \times l_j$ Matrix.
Für den Index j gilt $j = 1, \ldots, g_i$. Der Wert g_i ist hierbei die geometrische Vielfachheit des Eigenwertes λ_i. Oder etwas verständlicher formuliert: Die Anzahl der Blöcke in einem Kästchen entspricht der Anzahl an linear unabhängigen Eigenvektoren zum Eigenwert λ_i.
Die Blöcke werden (normalerweise) der Größe nach absteigend im Kästchen angeordnet. Das bedeutet

$$l_1 \geqslant l_2 \geqslant \cdots \geqslant l_{g_i} \geqslant 1.$$

Level 3 (Struktur eines Jordanblockes) Die Jordanblöcke $B_{\lambda_i}^{l_j}$ sind, wie schon erwähnt, $l_j \times l_j$ Matrizen. Sie weisen immer folgende Gestalt auf:

$$B_{\lambda_i}^{l_j} = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}. \tag{8.5}$$

Auf den ersten Blick wirkt das sicher sehr kompliziert, vor allem, wegen den vielen neuen Begriffen. Tatsächlich ist es aber recht simpel. Wir betrachten ein Beispiel:

Beispiel 43. Es sei die Jordan-Normalform J irgendeiner Matrix A gegeben. Sie hat folgende Form:

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

Wir sollen Jordan-Kästchen und Blöcke, sowie algebraische und geometrische Vielfachheiten der Eigenwerte von A, bestimmen. Die Eigenwerte kann man einfach aus der Matrix J ablesen – sie stehen immer auf der Hauptdiagonalen. Somit sind $\lambda_1 = 3$ und $\lambda_2 = 2$ die Eigenwerte.

Die Kästchen sind auch schnell in der Matrix J gefunden, denn sie haben den gleichen Eigenwert auf der Diagonalen stehen. Somit sind

$$K_{\lambda_1} = \begin{pmatrix} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & 0 & \\ & & & 3 & 1 \\ & & & & 3 \end{pmatrix} \quad \text{und} \quad K_{\lambda_2} = \begin{pmatrix} 2 & 0 \\ & 2 \end{pmatrix}$$

die Jordan-Kästchen. Die Größe dieser Kästchen bestimmt die algebraische Vielfachheit des jeweiligen Eigenwertes. Folglich ist 3 ein fünffacher und 2 ein doppelter Eigenwert.

Nach den Jordan-Blöcken müssen wir uns in den Kästchen auf die Suche machen. Starten wir mit K_{λ_1} und betrachten wir die Nebendiagonale. Eine 0 beschreibt immer eine „Trennlinie“ zwischen den Jordan-Blöcken. Also sind die gesuchten Blöcke

$$B_{\lambda_1}^{l_1} = \begin{pmatrix} 3 & 1 & \\ & 3 & 1 \\ & & 3 \end{pmatrix} \quad \text{und} \quad B_{\lambda_1}^{l_2} = \begin{pmatrix} 3 & 1 \\ & 3 \end{pmatrix}.$$

Die Matrix $B_{\lambda_1}^{l_1}$ ist ein Block der Größe 3, somit ist $l_1 = 3$. Und der Block $B_{\lambda_1}^{l_2}$ hat Länge 2, also ist $l_2 = 2$. Man beachte auch: Die Blöcke wurden (per Konvention) so angeordnet, dass sie der Größe nach absteigend im Kästchen vorkommen (es gilt $l_1 \geqslant l_2$).
Im ersten Kästchen haben wir also zwei Jordanblöcke! Die Anzahl der Jordanblöcke im Kästchen, entspricht der geometrischen Vielfachheit g_1 des ersten Eigenwertes. Folglich ist $g_1 = 2$ oder anders formuliert: Es gibt zwei l.u. Eigenvektoren zu $\lambda_1 = 3$.
Nun bestimmen wir die Blöcke von K_{λ_2}. Eine 0 in der (oberen) Nebendiagonalen trennt die Blöcke. Also sind die Blöcke gegeben durch

$$B_{\lambda_2}^{l_1} = \begin{pmatrix} 2 \end{pmatrix} \quad \text{und} \quad B_{\lambda_2}^{l_2} = \begin{pmatrix} 2 \end{pmatrix}.$$

Beide Blöcke sind von der Größe 1, hier gilt folglich $l_1 = l_2 = 1$. Wir haben zwei Blöcke im zweiten Kästchen, somit ist die geometrische Vilfachheit g_2 des zweiten Eigenwertes $\lambda_2 = 2$ gleich $g_2 = 2$. Es gibt zwei l.u. Eigenvektoren zu diesem Eigenwert.

8.3 Das Ermitteln der Jordan-Normalform

Eine Methode zur Ermittlung der Jordan-Normalform J einer Matrix A haben wir im Grunde schon kennengelernt: Wir berechnen Eigen- und Hauptvektoren der Matrix und stellen dann die Transformationsmatrix X auf. Wir invertieren X und berechnen

$$J = X^{-1}\ A\ X.$$

Das Problem: Das dauert zu lange. Außerdem ist bei vielen Aufgaben die Ermittlung der Eigen- und Hauptvektoren gar nicht gefragt. Es ist nur die Normalform J gesucht. Warum also Dinge berechnen, die überhaupt nicht verlangt sind?

8.3.1 Vorgehen bei Kästchengröße $\leqslant 3$

Wir haben im letzten Abschnitt den allgemeinen Aufbau einer Jordan-Normalform untersucht. Auf der Diagonale stehen immer die Eigenwerte. Darüber Nullen oder Einsen. Stehen auf der Diagonale gleiche Eigenwerte, bildet diese Untermatrix ein Kästchen. Innerhalb dieser Kästchen finden sich wiederum Blöcke, die nur Einser auf der Nebendiagonale stehen haben. Sie sind der Größe nach absteigend angeordnet. Die einzelnen Blöcke im Kästchen werden von Nullen auf der Nebendiagonale getrennt. Auf eine Sache möchte ich nochmals ganz besonders hinweisen:

Merkregel 8.3.1. Die Anzahl der Jordan Blöcke in einem Kästchen K_λ entspricht der geometrischen Vielfachheit g_λ des jeweiligen Eigenwertes λ.

Somit kennen wir bereits die Anzahl der Blöcke in einem Kästchen. Wir wissen auch, dass sie der Größe nach absteigend angeordnet werden. Was wir nicht wissen, ist die genaue Länge der einzelnen Blöcke im Kästchen. Können wir diese bestimmen, so können wir die die Normalform J aufstellen. Im Falle von $dim(K_\lambda) \leqslant 3$ ist das ganz besonders einfach:

Satz 8.3.2. *Sei K_λ ein Kästchen, das eine Größe $\leqslant 3$ aufweist. Das heißt die algebraische Vielfachheit a_λ des Eigenwertes λ erfüllt $a_\lambda \leqslant 3$.*
Dann ist der Aufbau des Kästchens eindeutig durch die geometrische Vielfachheit g_λ des Eigenwertes λ bestimmt.

Insbesondere gilt: Ist A eine Matrix der Größe 2 oder 3, so ist die Jordan-Normalform J eindeutig durch die geometrischen Vielfachheiten der Eigenwerte bestimmt. Das folgt daraus, dass in solche Matrizen nur Kästchen der Größe $\leqslant 3$ hineinpassen.

Beweis. Wir wissen, dass immer

$$1 \leqslant g_\lambda \leqslant a_\lambda = dim(K_\lambda) \leqslant 3$$

erfüllt ist. Betrachten wir die wichtigen Fälle

K_λ ist von der Größe 2 Hier haben wir nicht sehr viele Möglichkeiten. Ist $g_\lambda = 1$, so befindet sich nur ein Block im Kästchen. Das Kästchen hat also folgende Gestalt

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

Ist dagegen $g_\lambda = 2$, so finden sich zwei Blöcke im Kästchen und

$$K_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

K_λ ist von der Größe 3 Hier haben wir 3 Möglichkeiten: Ist $g_\lambda = 1$, so gibt es nur einen Block im Kästchen und

$$K_\lambda = \begin{pmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}.$$

Ist hingegen $g_\lambda = 3$, so gibt es 3 Blöcke. Es folgt

$$K_\lambda = \begin{pmatrix} \lambda & 0 & \\ & \lambda & 0 \\ & & \lambda \end{pmatrix}.$$

Die letzte Möglichkeit ist $g_\lambda = 2$. Da gibt es zwei Blöcke im Kästchen. Hier nutzen wir den Fakt, dass die Blöcke der Größe nach absteigend angeordnet werden. Das Kästchen kann nur so aussehen:

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

□

Beispiel 44. Betrachten wir die Matrix

$$A = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}.$$

Gesucht ist die Jordan-Normalform.
Wie man leicht durch Probieren erkennt, sind $\lambda_1 = -2$ und $\lambda_2 = 4$ Eigenwerte. Es gilt

$$\lambda_1 + \lambda_2 + \lambda_3 = 2 + \lambda_3 = tr(A) = 0.$$

Hieraus folgt $\lambda_3 = -2$ und -2 ist ein doppelter Eigenwert.
Mit passenden 2×2 Unterdeterminanten sieht man schnell, dass $rg(A_{\lambda_1}) = rg(A_{\lambda_2}) = 2$ gilt.
Somit ist $g_{\lambda_1} = g_{\lambda_2} = 1$. Es gibt also jeweils einen Block im jeweiligen Kästchen.
Da λ_1 ein doppelter Eigenwert ist, hat das Kästchen K_{λ_1} die Länge 2. Das zweite Kästchen hat Länge 1, da es sich bei λ_2 um einen einfachen Eigenwert handelt.
Folglich erhalten wir für die Jordan-Normalform

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

8.3.2 Der Spezialfall 4×4

Ist A eine 4×4 Matrix, so kann es erstmals passieren, dass ein Kästchen der Größe 4 auftritt. Das ist genau dann der Fall, wenn es nur einen Eigenwert λ gibt. Dieser tritt viermal auf, das heißt $a_\lambda = 4$.
Betrachten wir zuerst die freundlichen Fälle. Ist $g_\lambda = 1$, tritt ein Block im Kästchen auf. Folglich ist

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}.$$

Ist $g_\lambda = 3$ ist es ebenfalls einfach. Es kann nur

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

gelten. Im Fall $g_\lambda = 4$ gibt es 4 Blöcke, somit ist J eine Diagonalmatrix.

Und nun zum interessanten Fall $g_\lambda = 2$. Es existieren dann zwei Blöcke im Kästchen. Und hier haben wir erstmals das Problem, dass zwei Matrizen als Jordan-Normalform in Frage kommen können. Die erste Möglichkeit sind zwei Blöcke der Länge 2. Die zweite sind ein Block der Länge 3 und ein Block der Länge 1. Irgendwie müssen wir also entscheiden, welche der beiden Möglichkeiten für unser Matrix A die richtige ist!

Satz 8.3.3. *Sei A eine 4×4 Matrix mit einem vierfachen Eigenwert λ. Außerdem habe der Eigenwert die geometrische Vielfachheit $g_\lambda = 2$.*

Gilt $A_\lambda^2 = 0$, so ist die Jordan-Normalform J von A gegeben durch

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

Gilt hingegen $A_\lambda^2 \neq 0$, so ist $A_\lambda^3 = 0$ und

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

Beweis. Es gibt eine Matrix X, sodass

$$X^{-1}\, A\, X = J.$$

Daraus folgt

$$\begin{aligned} X^{-1}\ A_\lambda\ X &= X^{-1}\ (A - \lambda \cdot E_4)\ X \\ &= X^{-1}\ A\ X - \lambda \cdot E_4 \\ &= J - \lambda \cdot E_4 = J_\lambda. \end{aligned}$$

Angenommen $A_\lambda^2 = 0$. Dann gilt

$$(X^{-1}\ A_\lambda\ X)^2 = X^{-1}\ A_\lambda^2\ X = 0 = J_\lambda^2.$$

Nehmen wir nun an, die Jordan Normalform J wäre von der Form

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 0 \\ & & & \lambda \end{pmatrix} \quad \text{,somit ist} \quad J_\lambda = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}.$$

Es folgt

$$J_\lambda^2 = \begin{pmatrix} 0 & 0 & 1 & \\ & 0 & 0 & \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}.$$

Beim Quadrieren solcher Matrizen „wandern“ die Einser auf der Nebendiagonalen innerhalb eines Blockes um eine Stelle nach rechts.
Nun gilt aber $J_\lambda^2 \neq 0$, was ein Widerspruch ist.
Die zweite Möglichkeit für J haben wir damit ausgeschlossen, bleibt im Fall $A_\lambda^2 = 0$ nur noch die erste.

Ist $A_\lambda^2 \neq 0$ geht man analog vor. Das überlasse ich dem Leser. □

Beispiel 45. Gegeben sei die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ 1 & -1 & 0 & 3 \end{pmatrix}.$$

Gesucht ist die Jordan-Normalform.
Mit Laplace-Entwicklung (2. Zeile, 3. Spalte) folgt für das charakteristische Polynom

$$p(x) = (2-x)^2 \cdot \det \begin{pmatrix} 1-x & -1 \\ 1 & 3-x \end{pmatrix}$$

Somit ist $\lambda = 2$ mindestens ein doppelter Eigenwert. Die restlichen beiden Eigenwerte erfüllen

$$\begin{aligned} \lambda_2 + \lambda_3 &= 4 \\ \lambda_2 \cdot \lambda_3 &= 4 \end{aligned}$$

und deshalb $\lambda_2 = \lambda_3 = 2$. Der Eigenwert $\lambda = 2$ ist somit ein vierfacher Eigenwert.

Die Matrix

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

besitzt Rang 2. Folglich ist $g_\lambda = 2$.
Die genau Gestalt der Jordan-Normalform können wir durch Potenzieren der Matrix A_λ feststellen. Es gilt

$$A_\lambda^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq 0 \quad \text{und} \quad A_\lambda^3 = 0.$$

Folglich ist

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

8.3.3 Noch größere Matrizen

Es ist eher unwahrscheinlich, dass in Aufgaben die Jordan-Normalform von Matrizen zu bestimmen ist, die größer als 4×4 sind. Sollte man doch einmal über so eine Aufgabe stolpern, gibt es eine Formel, um die Anzahl der Jordan-Blöcke einer bestimmten Länge im Kästchen zu bestimmen:

Satz 8.3.4. *Sei A eine $n \times n$ Matrix. Die Zahl $N(\lambda, l)$ beschreibe die Anzahl der Jordan-Blöcke der Länge l zum Eigenwert λ der Matrix A. Dann ist*

$$N(\lambda, l) = rg\left({A_\lambda}^{l+1}\right) + rg\left({A_\lambda}^{l-1}\right) - 2 \cdot rg\left({A_\lambda}^{l}\right). \tag{8.6}$$

Beweis. **Vorbereitung**
Zuerst will ich nochmals auf Folgendes hinweisen:

$$\left({B_0}^l\right)^2 = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 0 & \cdots & & & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \cdots & & 0 & 1 \\ 0 & \cdots & & & 0 \\ 0 & \cdots & & & 0 \end{pmatrix}$$

Durch Quadrieren solcher Matrizen „wandert“ also die Diagonale mit den Einsen um einen Schritt nach rechts. Allgemein gilt: In $\left({B_0}^l\right)^k$ ist die Diagonale um $k-1$ Schritte verschoben. Außerdem ist $\left({B_0}^l\right)^{l-1} \neq 0$, aber $\left({B_0}^l\right)^l = 0$.
Weiters gilt $rg(\left({B_\lambda}^l\right)^k) = l$ für $\lambda \neq 0$, da $det\left({B_\lambda}^l\right) = \lambda^l \neq 0$.

Sei nun o.B.d.A λ der Eigenwert im ersten Kästchen der Jordan-Normalform J (Kästchen können ja angeordnet werden, wie wir wollen). Anstatt die Größe der Blöcke in J zu untersuchen, können wir genausogut die Größe der Blöcke in J_λ untersuchen. Außerdem gilt

$$X^{-1}\ A_\lambda\ X = J_\lambda \quad \text{und} \quad X^{-1}\ {A_\lambda}^l\ X = {J_\lambda}^l.$$

Oder anders formuliert: Die Matrizen ${A_\lambda}^l$ und ${J_\lambda}^l$ sind für $l \geqslant 1$ ähnlich und weisen deshalb den selben Rang auf.

Für $\mathbf{l = l_{\max}}$
Beweisen wir die Formel zuerst für $l_{\max}$, das ist die Länge des größten Jordan-Blockes im Kästchen. Wir wollen $rg({A_\lambda}^{l_{\max}})$ bestimmen. Es gilt

$${J_\lambda}^{l_{\max}} = \begin{pmatrix} K_0^{l_{\max}} & & & \\ 0 & K_{\lambda_2-\lambda}^{l_{\max}} & & \\ & & \ddots & \\ & & & K_{\lambda_k-\lambda}^{l_{\max}} \end{pmatrix}$$

Es gilt $\lambda_i - \lambda \neq 0$ für $i = 2, \ldots, k$, da die Eigenwerte in unterschiedlichen Kästchen verschieden sind. Folglich ist $\det(K_{\lambda_i-\lambda}) \neq 0$ und deshalb auch $\det({K_{\lambda_i-\lambda}}^l) \neq 0$ für beliebige l. Unabhängig vom genauen Wert von l sind also in ${J_\lambda}^l$ immer $n - a_\lambda$ linear unabhängige Spalten vorhanden (und zwar die „rechts von" ${K_0}^l$). Das heißt, es gilt immer

$$rg(A_\lambda^l) = rg({J_\lambda}^l) \geqslant n - a_\lambda \quad \text{für} \quad l = 1, 2, 3, \ldots$$

Der Rang von ${J_\lambda}^l$ hängt also nur vom Rang von ${K_0}^l$ ab!

Im Falle von $l = l_{\max}$ und auch $l = l_{\max+1}$ ist ${K_0}^l = 0$. Der größte Block im Kästchen hat ja gerade die Länge $l_{\max}$. Durch das Potenzieren verschieben sich die Eins-Diagonalen aller Blöcke weit genug nach rechts, dass nur noch Null-Blöcke übrig bleiben.
Es gilt also

$$rg({A_\lambda}^{l_{\max}}) = rg({A_\lambda}^{l_{\max}+1}) = n - a_\lambda.$$

Für $l = l_{\max} - 1$ verschwinden im Kästchen K_0 alle Blöcke der Länge $< l_{\max}$. Die Blöcke der Länge $l_{\max}$ nehmen durch das Potenzieren folgende Gestalt an:

$$\left({B_0}^{l_{\max}}\right)^{l_{\max}-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}.$$

Also ist

$$rg({A_\lambda}^{l_{\max}-1}) = n - a_\lambda + N(\lambda, l_{\max}).$$

Es folgt

$$\begin{aligned} & rg({A_\lambda}^{l_{\max}+1}) + rg({A_\lambda}^{l_{\max}-1}) - 2 \cdot rg({A_\lambda}^{l_{\max}}) \\ & = (n - a_\lambda) + (n - a_\lambda) + N(\lambda, l_{\max}) - 2 \cdot (n - a_\lambda) \\ & = N(\lambda, l_{\max}). \end{aligned}$$

Somit stimmt die Formel für $l = l_{\max}$.

Für beliebige l
Wählen wir $l < l_{\max}$ beliebig, gehen wir ähnlich vor. Wir schreiben $l = l_{\max} - m$ für ein passendes $m \geqslant 1$. Dann gilt

$$rg(A_\lambda^l) = rg({A_\lambda}^{l_{\max}-m}) = n - a_\lambda + \sum_{i=1}^{m} i \cdot N(\lambda, l_{\max} - (m - i))$$

Wie kommt diese Formel für den Rang zu Stande? Schauen wir uns das einmal für $m = 2$ an. Dann ist

$$rg({A_\lambda}^{l_{\max}-2}) = n - a_\lambda + N(\lambda, l_{\max} - 1) + 2 \cdot N(\lambda, l_{\max}).$$

Der zweite Teil nach $n - a_\lambda$ beschreibt den Rang von ${K_0}^{l_{\max}-2}$. Im Kästchen ${K_0}^{l_{\max}-2}$ verschwinden durchs Potenzieren alle Blöcke der Größe $< l_{\max} - 1$. Die Blöcke der Größe $l_{\max} - 1$ nehmen folgende Gestalt an:

$$\left({B_0}^{l_{\max}-1}\right)^{l_{\max}-2} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}.$$

Die Blöcke der Größe $l_{\max}$ schauen so aus:

$$\left(B_0{}^{l_{\max}}\right)^{l_{\max}-2} = \begin{pmatrix} 0 & \cdots & 1 & 0 \\ \vdots & & & 1 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}.$$

Die Anzahl der Blöcke der Länge $l_{\max} - 1$ geht also für den Rang von $K_0{}^{l_{\max}-2}$ einfach ein, die Anzahl der Blöcke der Länge $l_{\max}$ jedoch doppelt (zwei Einsen im Block)!
So kann man sich das auch für beliebige m überlegen.

Jetzt müssen wir nur noch einsetzen:

$$\begin{aligned}
& rg\left(A_\lambda{}^{l+1}\right) + rg\left(A_\lambda{}^{l-1}\right) - 2 \cdot rg\left(A_\lambda{}^{l}\right) \\
=& \sum_{i=1}^{m-1} i \cdot N(\lambda, l_{\max} - (m-1-i)) + \sum_{i=1}^{m+1} i \cdot N(\lambda, l_{\max} - (m+1-i)) \\
& - 2 \cdot \sum_{i=1}^{m} i \cdot N(\lambda, l_{\max} - (m-i)) \\
=& N(\lambda, l_{\max} - m) = N(\lambda, l).
\end{aligned}$$

Die restlichen Terme in den Summen heben sich weg. Somit haben wir die Formel bewiesen. □

Der Beweis ist ziemlich lang und teils schwer in schriftlicher Form anschaulich zu erklären. Und wie gesagt: Die Formel wird im Normalfall sowieso nicht benötigt. Man sollte sie wirklich nur dann verwenden, wenn man die Länge der einzelnen Blöcke nicht auf andere Art und Weise bestimmen kann. Machen wir trotzdem ein kurzes Beispiel.

Beispiel 46. Betrachten wir das letzte Beispiel. Es gilt

$$rg(A_\lambda) = 2 \quad rg(A_\lambda^2) = 1 \quad rg(A_\lambda^3) = rg(A_\lambda^4) = 0.$$

Somit ist

$$\begin{aligned}
N(\lambda, 3) &= 0 + 1 - 2 \cdot 0 = 1 \\
N(\lambda, 2) &= 0 + 2 - 2 \cdot 1 = 0 \\
N(\lambda, 1) &= 1 + 4 - 2 \cdot 2 = 1.
\end{aligned}$$

Im Kästchen gibt es also einen Block der Länge 3 und einen Block der Länge 1.

8.4 Ein abschließendes Beispiel

In diesem Abschnitt betrachten wir noch ein Beispiel zur Jordan-Normalform. Es ist etwas schwieriger als die bisherigen Beispiele und stammt von einem Übungsblatt. Außerdem dient dieser Abschnitt als Wiederholung und Zusammenfassung des bisher Gelernten. Wir werden in dem Beispiel so ziemlich alles anwenden, was wir bisher über das Eigenwertproblem wissen.

Beispiel 47. Gegeben sei die Matrix

$$A = \begin{pmatrix} 9 & -1 & 9 & -5 \\ 26 & -2 & 32 & -18 \\ -7 & 1 & -7 & 5 \\ -4 & 1 & -5 & 6 \end{pmatrix}.$$

Gesucht sind Eigenwerte, Eigenvektoren, Hauptvektoren, die Jordan-Normalform J, sowie die Transformationsmatrix X.

Zuerst erraten wir einen Eigenwert λ. Subtrahieren wir von den Diagonalelementen der ersten und dritten Zeile den Wert 2, so stimmen diese beiden Zeilen bis auf ein Vorzeichen überein. Folglich ist $\lambda = 2$ ein Eigenwert.

Nun berechnen wir die Determinante der Matrix. Da die Matrix voll besetzt ist, ist es wohl am günstigsten, die Matrix mit Gauß zu bearbeiten:

$$\det(A) = \det \begin{pmatrix} 9 & -1 & 9 & -5 \\ 8 & 0 & 14 & -8 \\ 2 & 0 & 2 & 0 \\ 5 & 0 & 4 & 1 \end{pmatrix} = \det \begin{pmatrix} 8 & 14 & -8 \\ 2 & 2 & 0 \\ 5 & 4 & 1 \end{pmatrix} = 4.$$

Die Spur der Matrix ergibt sich zu $tr(A) = 6$ und die Summe der Hauptminoren der Ordnung 2 ist

$$tr(\Lambda^2\ A) = 8 - 17 - 18 + 34 + 0 + 6 = 13.$$

Nun führen wir die Polynomdivision durch:

$$b_2 = 2 - 6 = -4, \quad b_1 = 13 - 4 \cdot 2 = 5, \quad b_0 = -4/2 = -2.$$

Folglich

$$u(x) = x^3 - 4x^2 + 5x - 2.$$

Ganzzahlige Nullstellen sind Teiler von 2. Wir sehen schnell, dass $x = 1$ Nullstelle ist. Somit ist

$$p = 1 - 4 = -3, \quad q = 2$$

und daher

$$x_{34} = \frac{3}{2} \pm \sqrt{\frac{9}{4} - \frac{8}{4}} = 1, 2.$$

Deshalb sind $\lambda_1 = 1$ und $\lambda_2 = 2$ jeweils doppelte Eigenwerte.

Nun berechnen wir Eigen- und Hauptvektoren. Es gilt

$$A_{\lambda_1} = \begin{pmatrix} 8 & -1 & 9 & -5 \\ 26 & -3 & 32 & -18 \\ -7 & 1 & -8 & 5 \\ -4 & 1 & -5 & 5 \end{pmatrix}$$

Die letzten beiden Zeilen sind linear unabhängig. Berechnen wir also die Kreuzprodukte

$$\begin{bmatrix} -7 & -4 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -8 & -5 & 1 & 0 \\ 5 & 5 & 0 & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} 0 & 3 \\ 15 & -3 \\ 0 & -3 \\ -3 & 0 \end{bmatrix} \Bigg|\Bigg| \begin{bmatrix} 0 & 1 \\ 5 & -1 \\ 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Der Eigenvektor ist eine Linearkombination dieser beiden Vektoren. Setzen wir sie in die ersten beiden Zeilen von A_{λ_1} ein, so erhalten wir die Matrix:

$$\begin{pmatrix} 0 & 0 \\ 3 & -3 \end{pmatrix}.$$

Lösung ist $\vec{x} = (1, 1)$. Folglich ist

$$\vec{v}_1 = (1, 4, -1, -1)$$

der linear unabhängige Eigenvektor zu $\lambda_1 = 1$.
Nun zum zugehörigen Hauptvektor: Wir müssen die Lösung von

$$\begin{pmatrix} 26 & -3 & 32 \\ -7 & 1 & -8 \\ -4 & 1 & -5 \end{pmatrix} \cdot \vec{h_1}^1 = \begin{pmatrix} 4t \\ -t \\ -t \end{pmatrix}.$$

bestimmen. Wir berechnen

$$t^2 \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -3 \\ 0 \\ -3 \end{pmatrix} = t^2 \cdot \begin{pmatrix} -3 \\ -6 \\ 3 \end{pmatrix}$$

Die Lösung ist skalares Vielfaches dieses Vektors. Setzen wir in die letzte Zeile ein (=freien Parameter ermitteln), erhalten wir für den Hauptvektor

$$\vec{h_1}^1 = \frac{1}{9} t \cdot \begin{pmatrix} -3 \\ -6 \\ 3 \\ 0 \end{pmatrix}.$$

Setzen wir das wiederum in die erste Zeile von A_{λ_1} ein, sehen wir, dass es auf den Parameter t nicht ankommt (klar, er hebt sich ja auf beiden Seiten weg, wegen Lösbarkeit brauchen wir uns also keine Gedanken machen). Setzen wir also $t = 9$. Somit ist

$$\vec{v}_1 = (9, 36, -9, -9) \quad \text{mit} \quad \vec{h}_1^{\,1} = (-3, -6, 3, 0).$$

Jetzt müssen wir das Ganze noch für den zweiten Eigenwert machen. Es gilt

$$A_{\lambda_2} = \begin{pmatrix} 7 & -1 & 9 & -5 \\ 26 & -4 & 32 & -18 \\ -7 & 1 & -9 & 5 \\ -4 & 1 & -5 & 4 \end{pmatrix}.$$

Die letzten zwei Zeilen dieser Matrix sind linear unabhängig. Wir berechnen die Kreuzprodukte

$$\begin{bmatrix} -7 & -4 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -9 & -5 & 1 & 0 \\ 5 & 4 & 0 & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} -1 & 4 \\ 8 & 1 \\ 0 & -3 \\ -3 & 0 \end{bmatrix}$$

Die Eigenvektoren sind Linearkombination dieser beiden Vektoren. Setzen wir sie in die ersten beiden Zeilen der Matrix A_{λ_2} ein, so erhalten wir die Matrix

$$\begin{pmatrix} 0 & 0 \\ -4 & 4 \end{pmatrix}.$$

Lösung ist $\vec{x} = (1, 1)$. Folglich ist

$$\vec{v}_2 = (3, 9, -3, -3)$$

der l.u. Eigenvektor zu $\lambda_2 = 2$.
Für den zugehörigen Hauptvektor müssen wir

$$\begin{pmatrix} 26 & -4 & 32 \\ -7 & 1 & -9 \\ -4 & 1 & -5 \end{pmatrix} \vec{h}_1^{\,2} = t \cdot \begin{pmatrix} 9 \\ -3 \\ -3 \end{pmatrix}$$

lösen. Wir berechnen

$$t^2 \cdot \begin{pmatrix} 15 \\ -3 \\ 15 \end{pmatrix} \times \begin{pmatrix} -9 \\ 0 \\ -12 \end{pmatrix} = t^2 \cdot \begin{pmatrix} 36 \\ 45 \\ -27 \end{pmatrix}.$$

Die Lösung ist skalares Vielfaches dieses Vektors. Ermitteln wir den freien Parameter, erhalten wir

$$\vec{h}_1^{\,2} = -\frac{1}{12} t \cdot \begin{pmatrix} 36 \\ 45 \\ -27 \\ 0 \end{pmatrix} = \frac{1}{4} t \cdot \begin{pmatrix} -12 \\ -15 \\ 9 \\ 0 \end{pmatrix}.$$

Wieder kommt es auf den Parameter t nicht an. Wir wählen ihn also passenderweise $t = 4$. Somit ist

$$\vec{v}_2 = (12, 36, -12, -12) \quad \text{mit} \quad \vec{h}_1^{\,2} = (-12, -15, 9, 0).$$

Transformationsmatrix X und Jordan-Normalform J ergeben sich zu

$$X = \begin{pmatrix} \vec{v}_1 & \vec{h}_1^{\,1} & \vec{v}_2 & \vec{h}_1^{\,2} \end{pmatrix} \quad \text{und} \quad J = \begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}.$$

Beide Eigenwerte haben geometrische Vielfachheit gleich 1. Deshalb befindet sich jeweils ein Block im jeweiligen Kästchen.

Das letzte Beispiel war ziemlich lang und hat viel Kopfrechnen erfordert. Deshalb wird man solche Aufgaben wohl niemals in Tests und Prüfungen erblicken. Ihre Lösung ist einfach zu aufwändig. Ich habe schon Leute gesehen, die bei diesem Beispiel bereits bei der Eigenwertberechnung gescheitert sind, da sie das charakteristische Polynom nicht aufstellen konnten. Gut, dass das bei meiner Methode überhaupt nicht notwendig ist!

9 Beispiele

In diesem Abschnitt werden mit den in den früheren Abschnitten vorgestellten Methoden Beispiele gelöst. Die meisten dieser Aufgaben stammen aus Prüfungsangaben einer TU-Vorlesung in Linearer Algebra.

9.1 Lineare Gleichungssysteme

Aufgabe 1. Gegeben sei das lineare Gleichungssystem $A\vec{x} = \vec{b}$ mit

$$A = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ -1 \\ a \\ 0 \end{pmatrix}$$

(a) Geben sie an, für welche Werte a das System lösbar ist. Für diese Werte a gebe man die allgemeine Lösung von $A\vec{x} = \vec{b}$ an.

Wir betrachten das erweiterte Gleichungssystem

$$\tilde{A}\vec{y} = \begin{pmatrix} A & -\vec{b} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ 1 \end{pmatrix}.$$

Erste und vierte Zeile der erweiterten Matrix $\tilde{A}$ sind linear unabhängig. Wir berechnen die Kreuzprodukte:

$$\begin{bmatrix} 2 & 0 & 1 & & & \\ 2 & 0 & & & & \\ -1 & 1 & & & & \\ 0 & 1 & & 1 & & \\ 1 & 1 & & & 1 & \\ -1 & 0 & & & & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -2 & -1 & -2 & 1 \\ 0 & -2 & -2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = X$$

Es gilt $\vec{y} = X\vec{s}$. In das erweiterte Gleichungssystem $\tilde{A}\vec{y}$ eingesetzt (in die verbleibenden Zeilen) ergibt

$$\begin{pmatrix} 0 & -9 & 0 & 1 \\ 0 & 3 & 0 & 1-2a \\ 0 & 0 & 0 & 2 \end{pmatrix} \vec{s} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Es folgt $s_4 = 1/2$ und

$$-9s_2 = -1/2 \quad 3s_2 = a - 1/2.$$

Somit folgt

$$3a - 3/2 = 1/2 \quad \Rightarrow 3a = 2 \quad \Rightarrow a = 2/3.$$

Es gilt

$$\vec{s} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \vec{t} + \begin{pmatrix} 0 \\ 1/18 \\ 0 \\ 1/2 \end{pmatrix}$$

Einsetzen in X ergibt für die allgemeine Lösung im Fall von $a = 2/3$

$$\vec{x} = t_1 \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ -2 \\ -2 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 4/9 \\ -1/9 \\ 1/9 \\ 0 \end{pmatrix}.$$

(b) Bestimme Bild und Kern von A (den Lösungsraum von $A\vec{x} = \vec{0}$).

Für das Bild wähle man drei beliebige linear unabhängige Spalten von A. Sie spannen das Bild auf.
Den Kern haben wir in Punkt (a) schon mitberechnet. Der Vektor $\vec{x}$ ist Lösung von $A\vec{x} = \vec{0}$ genau dann, wenn

$$\vec{x} = t_1 \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ -2 \\ -2 \\ 0 \\ 2 \end{pmatrix}.$$

Aufgabe 2. Gegeben sei das lineare Gleichungssystem $A\vec{x} = \vec{b}$ mit

$$A = \begin{pmatrix} 1 & 5 & 4 & 0 \\ 0 & 1 & -1 & 3 \\ 1 & 7 & 2 & 6 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ a \\ 3 \end{pmatrix}, \quad a \in \mathbb{R}.$$

(a) Berechne den Lösungsraum von $A\vec{x} = \vec{0}$.

Die ersten zwei Zeilen von A sind linear unabhängig. Wir berechnen die Kreuzprodukte

$$\begin{bmatrix} 1 & 0 & & \\ 5 & 1 & & \\ 4 & -1 & 1 & \\ 0 & 3 & & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} 15 & -9 \\ -3 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = X$$

Diese zwei Spalten von X lösen auch die dritte Zeile von $A\vec{x} = \vec{0}$, somit spannen sie den gesuchten Lösungsraum auf.

(b) Für welche Werte a hat das System $A\vec{x} = \vec{b}$ eine Lösung?

Die Matrix hat Rang 2. Wir betrachten das Teilsystem

$$\begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix} \vec{y} = \begin{pmatrix} 1 \\ a \end{pmatrix}.$$

Es folgt

$$\vec{y} = \begin{pmatrix} 1-5a \\ a \end{pmatrix}.$$

Folglich ist die Partikulärlösung im Fall der Lösbarkeit gegeben durch

$$\vec{x}_p = \begin{pmatrix} 1-5a \\ a \\ 0 \\ 0 \end{pmatrix}.$$

Das System ist nur lösbar, falls der Vektor $\vec{x}_p$ auch die letzte Zeile von $A\vec{x} = \vec{b}$ löst. Durch Einsetzten erhalten wir

$$1 - 5a + 7a = 3 \quad \Rightarrow a = 1.$$

(c) Geben Sie die allgemeine Lösung des Gleichungssystems im Falle der Lösbarkeit an.

Das System ist nur für $a = 1$ lösbar. Die allgemeine Lösung $\vec{x}$ ist

$$\vec{x} = \vec{x}_p + X\,\vec{t}.$$

Aufgabe 3. Gegeben sei die Matrix

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

(a) Man berechne die allgemeine Lösung von $A\vec{x} = \vec{0}$.

Die ersten zwei Zeilen sind linear unabhängig. Wir berechnen die Kreuzprodukte

$$\begin{bmatrix} 1 & 0 & & & \\ 0 & 2 & & & \\ 1 & 0 & 1 & & \\ 0 & 2 & & 1 & \\ 1 & 0 & & & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} -2 & 0 & -2 \\ 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = X.$$

Einsetzen in die dritte Zeile von $A\vec{x} = \vec{0}$ (vierte und fünfte Zeile sind identisch mit den ersten beiden), liefert

$$(4\ 0\ 0)\ \vec{s} = 0.$$

Somit ist

$$\vec{s} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{t}.$$

Und folglich spannen zweite und dritte Spalte von X den Lösungsraum auf.

(b) Man bestimme eine Orthogonalbasis von Bild(A).

Eine Orthogonalbasis ist eine Basis, in der die Vektoren paarweise orthogonal aufeinander stehen. Bestimmen wir zuerst eine ganz normale Basis vom Bild. Die Matrix hat Rang 3. Deshalb bilden die ersten drei Spalten von A eine Basis des Bildes.
Wir bemerken: Die ersten beiden Spalten stehen bereits orthogonal aufeinander. Hier ist nichts zu erledigen. Nur der dritte Spaltenvektor muss modifiziert werden. Und mit „Modifizieren" meine ich, dass ein Bildvektor derart zu ihm addiert werden muss, dass das Resultat orthogonal auf die ersten beiden Spalten steht. Diesen Bildvektor müssen wir natürlich erst einmal bestimmen. Machen wir das:

Gesucht sind Parameter t_1, t_2, sodass

$$\vec{b}_3 = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} + t_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \\ 0 \end{pmatrix} = \vec{s}_3 + t_1\ \vec{s}_1 + t_2\ \vec{s}_2$$

orthogonal auf die ersten beiden Spalten $\vec{s}_1$, $\vec{s}_2$ von A steht. Das heißt, es muss

$$\vec{b}_3 \cdot \vec{s}_1 = \vec{s}_3 \cdot \vec{s}_1 + t_1\ \vec{s}_1 \cdot \vec{s}_1 + t_2\ \underbrace{\vec{s}_2 \cdot \vec{s}_1}_{=0} = 0$$

$$\vec{b}_3 \cdot \vec{s}_2 = \vec{s}_3 \cdot \vec{s}_2 + t_1\ \underbrace{\vec{s}_1 \cdot \vec{s}_2}_{=0} + t_2\ \vec{s}_2 \cdot \vec{s}_2 = 0$$

gelten. Nun können wir durch Umformen t_1 und t_2 berechnen. Aus der ersten Gleichheit folgt

$$t_1 = -\frac{\vec{s}_3 \cdot \vec{s}_1}{\vec{s}_1 \cdot \vec{s}_1} = -\frac{5}{3}$$

und aus der zweiten Gleichheit erhalten wir

$$t_2 = -\frac{\vec{s}_3 \cdot \vec{s}_2}{\vec{s}_2 \cdot \vec{s}_2} = 0.$$

Somit gilt

$$\vec{b}_3 = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} - \frac{5}{3} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 0 \\ 4/3 \\ 0 \\ -2/3 \end{pmatrix}$$

und die Menge $\{\vec{s}_1, \vec{s}_2, \vec{b}_3\}$ ist die gesuchte orthogonale Basis von Bild(A).

Aufgabe 4. Es sei P_2 der Vektorraum der Polynome mit reellen Koeffizienten vom Maximalgrad 2, und $q \in P_2$ fest gegeben durch $q(x) = 1 + x + x^2$.

(a) Man betrachte die lineare Abbildung $\varphi : \; P_2 \to \mathbb{R}$, definiert durch

$$\varphi(p) := \int_0^1 \; p'(x)q(x) \; dx.$$

Berechnen Sie Kern(φ), das heißt den Lösungsraum von $\varphi(p) = 0$.

Eine Basis B für P_2 ist gegeben durch $B = \{1, x, x^2\}$, folglich ist P_2 ein Vektorraum der Dimension 3. Der Vektorraum $\mathbb{R}$ ist von der Dimension 1 mit Basis $C = \{1\}$. Die Darstellungsmatrix von φ bezüglicher dieser beiden Basen ist also eine 1×3 Matrix. Bestimmen wir ihre Einträge:

$$\varphi(1) = 0 = 0 \cdot 1, \quad \varphi(x) = \int_0^1 \; q(x) \; dx = \frac{11}{6} \cdot 1, \quad \varphi(x^2) = \int_0^1 \; 2x \cdot q(x) \; dx = \frac{13}{6} \cdot 1.$$

Folglich gilt für die Darstellungsmatrix $[\varphi]_{C \leftarrow B}$

$$[\varphi]_{C \leftarrow B} = \begin{pmatrix} 0 & 11/6 & 13/6 \end{pmatrix}$$

Das Gleichungssystem $[\varphi]_{C \leftarrow B} \cdot \vec{y} = 0$ können wir mit „Vertauschen und Vorzeichenwechsel“ lösen. Wir erhalten als l.u. Lösungsvektoren

$$\vec{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{y}_2 = \begin{pmatrix} 0 \\ -13 \\ 11 \end{pmatrix}.$$

Genau genommen sind das nur die Koordinaten der Lösungen bzgl. der Basis B. Wollen wir die Lösung in Form von Polynomen p_1, p_2, müssen wir lediglich die Einträge obiger Vektoren als die Koeffizienten von Polynomen auffassen. Folglich gilt

$$p_1 = 1, \quad p_2 = -13x + 11x^2.$$

Diese beiden Polynome spannen den Lösungsraum auf.

(b) Durch

$$\phi(p) := \begin{pmatrix} \varphi(p) \\ p(0) \\ p'(0) \end{pmatrix}$$

ist eine lineare Abbildung $\phi: \ P_2 \to \mathbb{R}^3$ definiert.
Berechne Kern(ϕ).

Für P_2 wählen wir wieder unsere Basis $B = \{1, x, x^2\}$. Den Raum $\mathbb{R}^3$ statten wir mit der Standardbasis E aus. Somit ist die Darstellungsmatrix $[\phi]_{E\leftarrow B}$ eine 3×3 Matrix. Es gilt

$$\phi(1) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \phi(x) = \begin{pmatrix} 11/6 \\ 0 \\ 1 \end{pmatrix}, \quad \phi(x^2) = \begin{pmatrix} 13/6 \\ 0 \\ 0 \end{pmatrix}$$

und folglich

$$[\phi]_{E\leftarrow B} = \begin{pmatrix} 0 & 11/6 & 13/6 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Diese Matrix besitzt vollen Rang, deshalb ist Kern(ϕ)$= \{0\}$, wobei 0 das Nullpolynom bezeichnet.

Aufgabe 5. Man betrachte das Gleichungssystem $A\vec{x} = \vec{b}$, mit

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & a \\ 1 & a & 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \quad a \in \mathbb{R}.$$

Man untersuche das System in Abhängigkeit von a auf Lösbarkeit und gebe, wenn möglich, die allgemeine Lösung an.

Wir betrachten das erweiterte Gleichungssystem

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 2 & 3 & a & -3 \\ 1 & a & 3 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die ersten zwei Zeilen sind immer linear unabhängig. Wir berechnen die Kreuzprodukte

$$\begin{bmatrix} 1 & 2 & & \\ 1 & 3 & & \\ -1 & a & 1 & \\ -1 & -3 & & 1 \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} a+3 & 0 \\ -2-a & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = X.$$

Wir setzen diese zwei Spalten wieder in das Gleichungssystem ein und erhalten das reduzierte System

$$\begin{pmatrix} -a^2-a+6 & a-2 \\ 0 & 1 \end{pmatrix} \vec{s} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad (a^2+a-6)s_1 = a-2, \quad s_2 = 1$$

Es gilt

$$a^2+a-6=0 \quad \Leftrightarrow \quad a = 2, -3.$$

Im Fall $a=2$ erhalten wir

$$0 \cdot s_1 = 0 \quad \Rightarrow \quad s_1 = t, \quad t \in \mathbb{R}$$

und die allgemeine Lösung ist

$$\begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = X \begin{pmatrix} t \\ 1 \end{pmatrix}.$$

Im Fall $a=-3$ folgt

$$0 \cdot s_1 = -5 \quad \Rightarrow \quad 0 = -5.$$

In diesem Fall ist das System also nicht lösbar.
Falls $a \notin \{2, -3\}$, ist das System eindeutig lösbar. Für die Lösung erhalten wir

$$\begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = X \begin{pmatrix} s_1 \\ 1 \end{pmatrix}, \quad \text{wobei} \quad s_1 = \frac{1}{a+3}.$$

Aufgabe 6. Gegeben sei das lineare Gleichungssystem $A\vec{x} = \vec{b}$ mit

$$A = \begin{pmatrix} 8 & -2 & 4 & 3 & 1 \\ 0 & 1 & -4 & 1 & -2 \\ 2 & 1 & 0 & -4 & 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ -10 \\ 1 \end{pmatrix}.$$

(a) Berechnen sie den Lösungsraum von $A\vec{x} = \vec{0}$.

Die ersten zwei Zeilen der Matrix A sind linear unabhängig. Wir berechnen die Kreuzprodukte

$$\begin{bmatrix} 8 & 0 & & & \\ -2 & 1 & & & \\ 4 & -4 & 1 & & \\ 3 & 1 & & 1 & \\ 1 & -2 & & & 1 \end{bmatrix} \mapsto \begin{bmatrix} 4 & -5 & 3 \\ 32 & -8 & 16 \\ 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} \Bigg|\Bigg| \begin{bmatrix} 1 & -5 & 3 \\ 8 & -8 & 16 \\ 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} = X$$

Spaltenweise eingesetzt in die letzte Zeile von $A\vec{x} = \vec{0}$ liefert

$$\begin{pmatrix} 10 & -50 & 22 \end{pmatrix} \vec{s} = 0.$$

Die Lösungsvektoren dieses reduzierten Systems sind

$$\begin{pmatrix} -11 \\ 0 \\ 5 \end{pmatrix}, \quad \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix}.$$

Wenn wir diese beiden Vektoren in X einsetzen, erhalten wir die linear unabhängigen Lösungsvektoren $\vec{x}_1, \vec{x}_2$ von $A\vec{x} = \vec{0}$

$$\vec{x}_1 = \begin{pmatrix} 4 \\ -8 \\ -22 \\ 0 \\ 40 \end{pmatrix} \Big|\Big| \begin{pmatrix} 2 \\ -4 \\ -11 \\ 0 \\ 20 \end{pmatrix}, \quad \vec{x}_2 = \begin{pmatrix} 0 \\ 32 \\ 10 \\ 8 \\ 0 \end{pmatrix} \Big|\Big| \begin{pmatrix} 0 \\ 16 \\ 5 \\ 4 \\ 0 \end{pmatrix}.$$

(b) Berechne die allgemeine Lösung von $A\vec{x} = \vec{b}$.

Um eine Partikulärlösung $\vec{x}_p$ zu ermitteln, betrachten wir das Teilsystem

$$\begin{pmatrix} 8 & 4 & 1 \\ 0 & -4 & -2 \\ 2 & 0 & 0 \end{pmatrix} \vec{y} = \begin{pmatrix} 2 \\ -10 \\ 1 \end{pmatrix}.$$

Es folgt direkt $y_1 = 1/2$ und wir müssen

$$\begin{pmatrix} 4 & 1 \\ -4 & -2 \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -2 \\ -10 \end{pmatrix}$$

lösen. Unsere Methode zum Auffinden einer Partikulärlösung liefert

$$\begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = t \cdot \begin{pmatrix} -7 \\ 24 \end{pmatrix}.$$

Durch Einsetzen in die erste Zeile erhalten wir für $t = 1/2$.
Somit ist

$$\vec{x}_p = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 0 \\ -7 \\ 0 \\ 24 \end{pmatrix}.$$

Die allgemeine Lösung ist folglich.

$$\vec{x} = \vec{x}_p + t_1 \cdot \vec{x}_1 + t_2 \cdot \vec{x}_2.$$

9.2 Eigenwertproblem

Aufgabe 1. Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -2 & -3 \\ -1 & 3 & 3 \\ 1 & -2 & -2 \end{pmatrix}.$$

(a) Berechnen Sie die Eigenwerte der Matrix.

Subtrahieren wir den Wert 1 von den Diagonaleinträgen der Matrix A, entsteht eine Matrix vom Rang 1. Somit ist $\lambda_1 = 1$ mindestens ein doppelter Eigenwert. Aus der Bedingung

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 + 1 + \lambda_3 = tr(A) = 3$$

folgt $\lambda_3 = 1$. Somit ist $\lambda_1 = 1$ sogar ein dreifacher Eigenwert.

(b) Berechnen Sie Eigenvektoren und Hauptvektoren.

Es gilt

$$A_{\lambda_1} = \begin{pmatrix} 1 & -2 & -3 \\ -1 & 2 & 3 \\ 1 & -2 & -3 \end{pmatrix}.$$

Diese Matrix hat, wie schon einmal erwähnt, Rang 1. Wir ermitteln die beiden l.u. Eigenvektoren mit „Vertauschen und Vorzeichenwechsel". Wir erhalten

$$\vec{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Es gibt einen Hauptvektor $\vec{h}_1$, der

$$A_{\lambda_1}\,\vec{h}_1 = \begin{pmatrix} 1 & -2 & -3 \\ -1 & 2 & 3 \\ 1 & -2 & -3 \end{pmatrix}\vec{h}_1 = t_1 \cdot \vec{v}_1 + t_2 \cdot \vec{v}_2 = \begin{pmatrix} 2t_1 + 3t_2 \\ t_1 \\ t_2 \end{pmatrix}$$

erfüllen muss. Die Parameter t_1, t_2 sind so zu wählen, dass das obige System lösbar ist.
Für die Partikulärlösung (=Hauptvektor) arbeiten wir mit der linken unteren 1×1 Untermatrix von A_{λ_1} und erhalten

$$\vec{h}_1 = \begin{pmatrix} t_2 \\ 0 \\ 0 \end{pmatrix}.$$

Da wir das System als lösbar angenommen haben, muss der Vektor $\vec{h}_1$ auch die ersten zwei Zeilen von $A_{\lambda_1}\,\vec{h}_1 = t_1\vec{v}_1 + t_2\vec{v}_2$ erfüllen. Einsetzen liefert nun die Bedingung $t_1 = -t_2$. Wählen wir z.B. $t_1 = 1$ und $t_2 = -1$, dann ist

$$\vec{v} = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \quad \text{mit Hauptvektor} \quad \vec{h}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}.$$

Als einen von $\vec{v}$ linear unabhängigen Eigenvektor können wir z.B. $\vec{v}_2$ nehmen.

(c) Bestimmen Sie die Jordan-Normalform J von A und die Transfomationsmatrix X, die A in J überführt.

Da die geometrische Vielfachheit von $\lambda_1 = 1$ gleich 2 ist, finden sich zwei Blöcke im Kästchen. Die Normalform kann nur so aussehen:

$$J = \begin{pmatrix} 1 & 1 & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

Für die Matrix X erhalten wir

$$X = \begin{pmatrix} \vec{v} & \vec{h}_1 & \vec{v}_2 \end{pmatrix}.$$

Man beachte die Reihenfolge der Vektoren!

Aufgabe 2. Gegeben sei die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix}.$$

(a) Berechnen Sie die Eigenwerte und das charakteristische Polynom p_A.

Zieht man den Wert 2 von den Diagonalelementen der Matrix A ab, entsteht eine Matrix vom Rang 1. Folglich ist $\lambda = 2$ mindestens ein dreifacher Eigenwert. Außerdem gilt

$$3 \cdot \lambda + \lambda_4 = 6 + \lambda_4 = tr(A) = 8 \quad \Rightarrow \quad \lambda_4 = 2.$$

Also ist $\lambda = 2$ ein vierfacher Eigenwert und $p_A = (x-2)^4$.

(b) Bestimmen Sie Eigen- und Hauptvektoren.

Es gilt

$$A_\lambda = \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix}.$$

Somit sind die l.u. Eingenvektoren gegeben durch

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Es gibt einen Hautvektor $\vec{h}_1$, der so gewählt werden muss, dass

$$A_\lambda \, \vec{h}_1 = t_1\vec{v}_1 + t_2\vec{v}_2 + t_3\vec{v}_3,$$

wobei t_1, t_2, t_3 so sind, dass das System lösbar ist. Um eine Partikulärlösung (=Hauptvektor) zu bestimmen, arbeiten wir mit der linken unteren 1×1 Untermatrix von A_λ. Wir erhalten

$$\vec{h}_1 = \begin{pmatrix} -t_2 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Da wir das System als lösbar vorausgesetzt haben, soll dieser Vektor auch die ersten drei Zeilen des Systems lösen. Folglich gilt

$$\begin{aligned} &\text{1. Zeile} \quad && \Rightarrow t_2 = t_1 + t_2 \quad \Rightarrow \quad t_1 = 0 \\ &\text{3. Zeile} \quad && \Rightarrow t_2 = t_3. \end{aligned}$$

Setzen wir z.B. $t_1 = t_2 = 1$, dann ist

$$\vec{v} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{mit Hautvektor} \quad \vec{h}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Zwei zu $\vec{v}$ linear unabhängige Eigenvektoren sind z.B. $\vec{v}_1, \vec{v}_3$.

(c) Bestimme die Jordan-Normalform J und die Transformationsmatrix X.

Die geometrische Vielfachheit des Eigenwertes $\lambda = 2$ ist gleich 3. Folglich sind drei Blöcke im Kästchen. Die Matrix J kann demnach nur so aussehen

$$J = \begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}.$$

Für die Matrix X erhalten wir

$$X = \begin{pmatrix} \vec{v} & \vec{h}_1 & \vec{v}_1 & \vec{v}_3 \end{pmatrix}.$$

Aufgabe 3. Gegeben sei die Matrix

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

(a) Berechnen Sie die Eigenwerte von A.

Subtrahieren wir den Wert 1 von den Diagonalelemten, entsteht eine Matrix vom Rang 2. Somit ist $\lambda = 1$ mindestens ein doppelter Eigenwert. Aus

$$2 \cdot \lambda + \lambda_3 + \lambda_4 = tr(A) = 4$$

$$\lambda^2 \cdot \lambda_3 \cdot \lambda_4 = \det(A) = 1$$

folgt $\lambda_3 = \lambda_4 = \lambda = 1$. Also ist $\lambda = 1$ ein vierfacher Eigenwert.

(b) Bestimmen Sie die Jordan-Normalform J von A.

Es gilt

$$A_\lambda = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Diese Matrix hat, wie schon erwähnt, Rang 2. Deshalb gibt es zwei Blöcke im Kästchen und es existieren zwei Möglichkeiten für die Normalform. Um zu entscheiden, welche richtig ist, potenzieren wir A_λ. Gilt $A_\lambda^2 = 0$, so ist der größte Block von der Länge 2, ansonsten von der Länge 3.
Wegen

$$(A_\lambda^2)_{11} = 1 \quad \text{gilt} \quad A_\lambda^2 \neq 0,$$

folglich ist der größte Block im Kästchen von der Länge 3 und deshalb

$$J = \begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

(c) Bestimme Eigen- und Hauptvektoren.

Die Eigenvektoren können wir entweder mit dem Kreuzprodukt oder durch „genaues Hinschauen" auf die ersten zwei Zeilen von A_λ ermitteln.

Für die l.u. Eigenvektoren erhalten wir folglich

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

Es gibt zwei Hauptvektoren $\vec{h}_1, \vec{h}_2$. Der erste Hauptvektor erfüllt

$$A_\lambda \vec{h}_1 = t_1 \vec{v}_1 + t_2 \vec{v}_2,$$

wobei t_1, t_2 so gewählt werden, dass das System lösbar ist. Die homogene Lösung dieses Systems kennen wir bereits. Für die Partikulärlösung $\vec{h}_1^p$ benutzen wir die linke obere 2×2 Untermatrix von A_λ und erhalten

$$\vec{h}_1^p = \begin{pmatrix} -t_1 \\ -t_2 \\ 0 \\ 0 \end{pmatrix}.$$

Da wir Lösbarkeit fordern, soll dieser Vektor auch den letzten zwei Zeilen des obigen inhomogenen Gleichungssystems genügen. Durch Einsetzen erhalten wir die Bedingung $t_1 = -t_2$. Wir können also z.B. $t_1 = 1$ und $t_2 = -1$ wählen und erhalten

$$\vec{v} = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \quad \text{mit (allgemeinem) Hauptvektor} \quad \vec{h}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s_1 \vec{v}_1 + s_2 \vec{v}_2.$$

Die Parameter s_1, s_2 werden wieder so gewählt, dass das Gleichungssystem $A_\lambda \vec{h}_2 = \vec{h}_1$ lösbar ist. Für die Partikulärlösung (=Hauptvektor $\vec{h}_2$, da es danach keine weiteren Hauptvektoren zu berechnen gibt) benutzen wir wieder die linke obere 2×2 Untermatrix und erhalten

$$\vec{h}_2 = \begin{pmatrix} -1 - s_1 \\ -1 - s_2 \\ 0 \\ 0 \end{pmatrix}.$$

Wegen der Lösbarkeit soll dieser Vektor den letzten beiden Zeilen von $A_\lambda \vec{h}_2 = \vec{h}_1$ genügen und folglich gilt

$$s_1 + 1 = -s_2.$$

Wir können also z.B. $s_1 = -1, s_2 = 0$ wählen und erhalten somit

$$\vec{v} = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \quad \text{mit Hauptvektoren} \quad \vec{h}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{h}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Ein zu $\vec{v}$ l.u. Eingenvektor ist z.B. $\vec{v}_1$ (oder $\vec{v}_2$).

Aufgabe 4. Gegeben sei die Matrix

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

(a) Bestimme die Eigenwerte von A.

Offensichtlich ist $\lambda_1 = 1$ ein Eigenwert von A. Die restlichen beiden Eigenwerte bestimmen/erraten wir aus

$$\lambda_2 + \lambda_3 = 2 - 2 = 0, \quad \lambda_2 \cdot \lambda_3 = -4 + 4 = 0.$$

Wir erhalten $\lambda_2 = \lambda_3 = 0$.

(b) Berechne Eigen- und Hauptvektoren.

Es gilt

$$A_{\lambda_1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 4 & -3 \end{pmatrix}.$$

Da $\lambda_1 = 1$ ein einfacher Eigenwert ist, gibt es nur einen l.u. Eigenvektor $\vec{v}_1$. Ohne viel Rechnung kann man sehen:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

erfüllt $A_{\lambda_1}\vec{v}_1 = \vec{0}$.
Es gilt $A_{\lambda_2} = A$ und A besitzt Rang 2. Folglich ist

$$\vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}.$$

Es gibt einen Hauptvektor $\vec{h}_2$, der $A_{\lambda_2}\vec{h}_2 = t\vec{v}_2$ erfüllt. Der Parameter t wird so gewählt, dass das System lösbar ist. Für die Partikulärlösung (=Hauptvektor) dieses inhomogenen Systems, benutzen wir die linke obere 2×2 Untermatrix von $A_{\lambda_2} = A$ und erhalten

$$\vec{h}_2 = \begin{pmatrix} -2t \\ t/2 \\ 0 \end{pmatrix}.$$

Da der Vektor $\vec{h}_2$ auch die letzte Zeile von $A_{\lambda_2}\vec{h}_2 = t\vec{v}_2$ für beliebige t erfüllt, können wir z.B. $t = 2$ wählen.
Wir erhalten

$$\vec{v} = \begin{pmatrix} -4 \\ 2 \\ 4 \end{pmatrix} \quad \text{mit Hauptvektor} \quad \vec{h}_2 = \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix}.$$

(c) Man bestimme die Jordan-Normalform J von A und die Transformationsmatrix X, die A in J überführt.

Für die Matrix X gilt

$$X = \begin{pmatrix} \vec{v}_1 & \vec{v} & \vec{h}_2 \end{pmatrix}.$$

Die Matrix J hat folgende Gestalt

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

Aufgabe 5. Gegeben sei die Matrix

$$A = \begin{pmatrix} 1 & a^2 & 1 \\ 0 & 1 & a^2 \\ 0 & 1 & 1 \end{pmatrix}, \quad a \in \mathbb{R}.$$

(a) Berechne die Eigenwerte von A und weise nach, dass A für $a \neq 0$ diagonalisierbar ist.

Offensichtlich ist $\lambda_1 = 1$ ein Eigenwert. Betrachten wir die Bedingungen

$$\lambda_2 + \lambda_3 = 2$$

$$\lambda_2 \cdot \lambda_3 = 1 - a^2 = (1-a) \cdot (1+a)$$

sehen wir direkt, dass $\lambda_2 = 1 - a$ und $\lambda_3 = 1 + a$ gilt.
Für $a \neq 0$ sind diese drei Eigenwerte paarweise verschieden, folglich ist die Matrix diagonalisierbar.

(b) Bestimme die Jordan-Normalform von A im Fall $a = 0$.

Im Fall $a = 0$ ist $\lambda = 1$ ein dreifacher Eigenwert. Die Matrix A_λ hat Rang 2. Folglich ist die geometrische Vielfachheit von λ gleich 1. Es befindet sich also ein Block im Kästchen der Jordan-Normalform und somit gilt

$$J = \begin{pmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{pmatrix}.$$

Aufgabe 6. Gegeben sei die Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R}.$$

(a) Wann weist die Matrix A reelle Eigenwerte auf?

Für das charakteristische Polynom p_A von A gilt bekanntlich

$$p_A(x) = x^2 - tr(A)x + \det(A).$$

Folglich gilt für die beiden Eigenwerte $\lambda_{1,2}$

$$\lambda_{1,2} = \frac{tr(A)}{2} \pm \sqrt{\frac{tr(A)^2}{4} - \det(A)}.$$

Beide Eigenwerte sind reell, falls der Ausdruck unter der Wurzel nicht negativ ist, das heißt, falls

$$\frac{tr(A)^2}{4} \geqslant \det(A).$$

(b) Im Fall von reellen Eigenwerten: Wann sind die Eigenwerte paarweise verschieden? Was heißt das für die Matrix A?

Es gibt einen doppelten Eigenwert, falls der Ausdruck unter der Wurzel $= 0$ wird. Folglich sind die Eigenwerte verschieden, falls

$$\frac{tr(A)^2}{4} > \det(A).$$

Sind die Eigenwerte einer Matrix paarweise verschieden, so ist sie diagonalisierbar.

(c) Wann ist die Matrix im Fall eines doppelten Eigenwertes diagonalisierbar?

Es existiert ein doppelter Eigenwert, falls der Ausdruck unter der Wurzel Null wird. Dieser doppelte Eigenwert ist dann

$$\lambda = \frac{tr(A)}{2} = \frac{a+d}{2}.$$

Damit die Matrix im Falle eines doppelten Eigenwertes diagonalisierbar ist, müssen zwei linear unabhängige Eigenvektoren existieren. Das ist nur möglich, falls Rang(A_λ)$= 0$ gilt. Das heißt A_λ muss die Nullmatrix sein:

$$A_\lambda = \begin{pmatrix} (a-d)/2 & b \\ c & (d-a)/2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hieraus folgt: $b = c = 0$ und $a = d$. Die Matrix A ist dann eine Diagonalmatrix mit gleichen Diagonaleinträgen.