Crashkurs: Diagonalisieren

Wenn man gefragt wird, welches Thema der Linearen Algebra am wichtigsten ist, dann kann man eigentlich nur eine Antwort geben: **Diagonalisierung**.
Diagonalisierung spielt die zentrale Rolle in immens vielen Anwendungen, die kaum zu überblicken sind, und sich über alle Bereiche der Mathematik und deren Anwendungsgebiete verteilen.
Hier versuche ich mich möglichst strikt an die Notation vom Vorabskript zur Vorlesung Lineare Algebra I und II von Herrn Prof. Dr. Maier [Maier12] zu halten. Was ich aber nicht benutzen möchte sind die Pfeile über Vektoren. Das ist nicht üblich und (das muss ich zugeben) wäre mir zu viel Arbeit. Es kann aber auch passieren, dass sich auch andere (mir vertrautere) Notationen einschleichen.

Inhaltsverzeichnis

1 Motivation

Gegeben sei im folgenden immer ein endlich-dimensionaler K-Vektorraum V von der Dimension $n = \dim_K V$. Auf diesem sei eine lineare Abbildung

$$\varphi \colon V \to V$$

gegeben. Wir wollen nun eine derartige Basis $\mathcal{B}$ finden, sodass die Darstellungsmatrix $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ eine möglichst einfache Gestalt hat. Doch was hätte man davon, wenn einem dies gelingt? Der Vorteil ist, dass man beispielsweise Potenzen φ^m der Abbildung (oder auch Polynomauswertungen $p(\varphi)$, oder wie hier für die Vorlesung wichtig die Exponentialfunktion $\exp(\varphi)$) einfach angeben kann. In der Tat ist nach [Maier12, Satz 3.5.7.]

$$M\left(\varphi^m; \mathcal{B}, \mathcal{B}\right) = M\left(\varphi; \mathcal{B}, \mathcal{B}\right)^m .$$

Ist nun $D = M\left(\varphi; \mathcal{B}, \mathcal{B}\right) = \operatorname{diag}\left(\lambda_1, ..., \lambda_n\right)$ eine Diagonalmatrix, dann ist

$$M\left(\varphi^m; \mathcal{B}, \mathcal{B}\right) = D^m = \operatorname{diag}\left(\lambda_1^m, ..., \lambda_n^m\right).$$

Ist doch recht angenehm damit zu arbeiten? Wir wollen das ganze an einem Beispiel verfolgen. Es wird sich immer um dasselbe Beispiel handeln. Wir wollen das aus verschiedenen Blickwinkeln betrachten.

Unser Standardbeispiel. *Es sei $V = \mathbb{R}^2$. Wir haben eine Abbildung $\varphi\colon V \to V$ gegeben. Diese sei eine Spiegelung an der Gerade $\langle e_1 + e_2 \rangle$. Es ist natürlich die Basis $\mathcal{B} = \{e_1 + e_2, e_1 - e_2\}$ zu benutzen. Dann ist die Darstellungsmatrix*

$$M\left(\varphi; \mathcal{B}, \mathcal{B}\right) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Wir haben also zum Beispiel

$$M\left(\varphi^{100}; \mathcal{B}, \mathcal{B}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

und damit also $\varphi^{100} = \mathrm{id}_{\mathbb{R}^2}$.

Die Frage ist nun, was bekannt ist über die Abbildung φ. Entweder die Abbildung ist konkret gegeben, oder man kennt die Darstellungsmatrix in einer anderen Basis. In der Anwendung, die uns in der Vorlesung begegnet, ist oft eine Basis $\tilde{\mathcal{B}}$ gegeben (ist zum Beispiel $V = K^n$, dann hat man auch oft als Basis die kanonische Basis aus den Einheitsvektoren). Bekannt ist dann die Darstellungsmatrix

$$A = \mathcal{M}\left(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}\right).$$

Die Frage ist nun, ob man eine Basis $\mathcal{B}$ finden kann, sodass

$$D = \mathcal{M}\left(\varphi; \mathcal{B}, \mathcal{B}\right)$$

eine Diagonalmatrix ist. Und dies wollen wir allein durch die Matrix A entscheiden. Durch [Maier12, Satz 3.5.4.] hat man dann die folgende Beziehung von A und D:

$$A = \mathcal{X} D \mathcal{X}^{-1} \text{ mit } \mathcal{X} = \mathcal{M}\left(\mathrm{id}; \tilde{\mathcal{B}}, \mathcal{B}\right).$$

Man beachte, dass $\mathcal{X}^{-1} = \mathcal{M}\left(\mathrm{id}; \mathcal{B}, \tilde{\mathcal{B}}\right)$ wieder nach [Maier12, Satz 3.5.4.] gilt. Man kann noch einmal sehen, dass die (eventuell) komplizierte Berechnung von A^m bzw. $f(A)$ (für eine geeignete Funktion z.B. ein Polynom, oder exp) sich zu der Bestimmung von $\mathcal{X}$ und der Berechung von D reduziert.

$$A^m = \mathcal{X} D^m \mathcal{X}^{-1} \text{ bzw. } f(A) = \mathcal{X} f(D) \mathcal{X}^{-1}$$

Also nur einmal die Arbeit $\mathcal{X}$ und D zu bestimmen und dafür alle Potenzen und die Exponentialfunktion von A schnell ausrechnen können. Ist das nicht super?

Unser Standardbeispiel. *Nehmen wir $\tilde{\mathcal{B}} = \{e_1, e_2\}$ als eine Basis von $\mathbb{R}^2$. Und sei φ wie oben. Wegen*

$$\varphi\left(e_1\right) = \varphi(e_1) = e_2,$$
$$\varphi\left(e_2\right) = \varphi(e_2) = e_1$$

ist dann

$$A := \mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In der Tat ist in dem Fall

$$A = \mathcal{X} D \mathcal{X}^{-1}$$

mit

$$\mathcal{X} = \mathcal{M}\left(\varphi; \tilde{\mathcal{B}}, \mathcal{B}\right) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

wie wir behauptet haben. Wir halten nochmal fest, dass wir die Potenzen berechnen können. Es ist

$$A^m = \begin{cases} A & \text{, falls } m \text{ ungerade} \\ E_2 & \text{, sonst.} \end{cases}$$

Aus A ist nun nicht mehr offensichtlich, was D war. Wir versuchen natürlich aus dem Wissen von A die Diagonalmatrix D und die Basiswechselmatrix $\mathcal{X}$ zu berechnen. Dies werden wir im Folgenden tun.

Wir machen uns hier die Situation etwas einfacher. Wir nehmen immer an, dass $K = \mathbb{C}$ ist, was für die meisten Anwendungen mehr als ausreichend ist.

2 Was ist diagonalisieren? Welche Begriffe spielen eine Rolle?

Wir halten das, was wir in der Motivation gemacht haben in einer Definition fest.

Definition. *Eine lineare Abbildung $\varphi \colon V \to V$ heißt* ***diagonalisierbar****, wenn es eine Basis $\mathcal{B}$ von V gibt derart, dass $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ eine Diagonalmatrix ist.*

In der Motivation haben wir das auch umformuliert in Matrizen. Dazu machen wir folgende Definition.

Definition. *Eine quadratische Matrix $A \in \mathbb{C}^{(n,n)}$ heißt **diagonalisierbar**, wenn es eine Matrix $\mathcal{X} \in \mathrm{GL}(n, \mathbb{C})$ gibt mit*

$$A = \mathcal{X} D \mathcal{X}^{-1}.$$

Dabei sei D eine Diagonalmatrix.

Die Verbindung zwischen den zwei Definitionen ist recht einfach. Gegeben eine feste Basis $\tilde{\mathcal{B}}$ von V. Eine lineare Abbildung $\varphi \colon V \to V$ ist genau dann diagonalisierbar, wenn die Matrix $\mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}})$ diagonalisierbar ist.
Ist $v \in \mathcal{B}$ ein Element der Basis, dann gilt offensichtlich

$$\varphi(v) = \lambda v \tag{1}$$

für ein λ, welches auf der Diagonalen von D steht.
Wir beschäftigen uns nun weiter mit dem Matrixfall. Die Gleichung

$$A = \mathcal{X} D \mathcal{X}^{-1}$$

können wir auch so schreiben:

$$A\mathcal{X} = \mathcal{X} D.$$

Ist $v \in \mathbb{C}^n$ die i-te Spalte in der Matrix $\mathcal{X}$ ($v \neq 0$ gilt automatisch), so gilt

$$Av = \lambda_i v, \tag{2}$$

wobei $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ ist.

Unser Standardbeispiel. *Oben hatten wir*

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

und

$$\mathcal{X} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

In der Tat ist

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

und

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

wie wir oben behauptet haben.

Wir sollten uns also auf Gleichungen der Form (2) und (1) konzentrieren. Deshalb die folgende Definition.

Definition (Eigenwerte und Eigenvektoren)**.** *Gegeben eine Matrix* $A \in \mathbb{C}^{(n,n)}$ *(bzw. eine lineare Abbildung* $\varphi\colon V \to V$*). Man nennt* $\lambda \in \mathbb{C}$ *einen* ***Eigenwert*** *von* A *(bzw.* φ*), falls es ein* $v \in \mathbb{C}^n \backslash \{0\}$ *(bzw. ein* $v \in V \backslash \{0\}$*) gibt mit*

$$Av = \lambda v \ \textit{(bzw.}\varphi(v) = \lambda v\textit{)}.$$

Jedes solche v *nennt man einen* ***Eigenvektor*** *zum Eigenwert* λ*.*

Man kann nun die Definition von Diagonalisierbarkeit mit dem neuen Begriff nun recht schön formulieren.

Lemma. *Eine Matrix* $A \in \mathbb{C}^{(n,n)}$ *(bzw. eine lineare Abbildung* $\varphi\colon V \to V$*) ist genau dann diagonalisierbar, wenn es eine Basis bestehend aus Eigenvektoren gibt.*

Man kann auch, und das ist sehr wichtig, die Matrix $\mathcal{X}$ nun mit den Eigenvektoren angeben. Es sei $v_1, ..., v_n$ eine Basis (von $\mathbb{C}^n$) von Eigenvektoren zu den Eigenwerten $\lambda_1, ..., \lambda_n$. Dann setzt man $\mathcal{X} = (v_1 \ v_2 \ ... \ v_n)$. Dies ist nun eine invertierbare Matrix und es gilt

$$A\mathcal{X} = \mathcal{X}D$$

für $D = \operatorname{diag}(\lambda_1, ..., \lambda_n)$. Dazu muss man einfach die Gleichung spaltenweise verifizieren.

Unser Standardbeispiel. *Wir erinnern, dass wir*

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

betrachtet hatten. In der Tat ist

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

eine Basis von Eigenvektoren. Hier bekommen wir als Basiswechselmatrix unser altes $\mathcal{X}$*. Es ist aber auch*

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

eine Basis von Eigenvektoren. Dann bekommen wir eine andere Basiswechselmatrix, aber dieselbe Diagonalmatrix. Man kann schon bemerken, dass die Basiswechselmatrix nicht eindeutig ist, die Diagonalmatrix D *aber schon (warum? - das kann man leicht argumentieren, wenn man das charakteristische Polynom und den Determinantenmultiplikationssatz benutzt - siehe unten).*

Analog besteht im Fall einer linearen Abbildung die Basis $\mathcal{B}$ genau aus der Basis der Eigenvektoren. Die Matrix D hat auf der Diagonalen wieder die Eigenwerte.
Noch ist völlig unklar, wie man sowas überhaupt ausrechnet. Dazu betrachten wir wieder den Fall einer linearen Abbildung φ. Es ist λ per Definition genau dann ein Eigenwert, wenn $\varphi - \lambda \operatorname{id}$ nicht injektiv ist. Dies ist nach [Maier12, Satz 3.2.6.] äquivalent dazu, dass

$\varphi - \lambda\,\mathrm{id}$ nicht bijektiv ist. Das kann man wieder nach [Maier12, Satz 5.4.2.] ausdrücken über

$$\det(\varphi - \lambda\,\mathrm{id}) = 0.$$

Es macht also Sinn sich das Polynom

$$P_\varphi(\lambda) = \det(\varphi - \lambda\,\mathrm{id})$$

anzusehen.

Definition (charakteristisches Polynom). *Es sei $A \in \mathbb{C}^{(n,n)}$ (bzw. eine lineare Abbildung $\varphi\colon V \to V$) gegeben. Man nennt*

$$P_A(\lambda) = \det(A - \lambda E_n) \text{ (bzw. } P_\varphi(\lambda) = \det(\varphi - \lambda\,\mathrm{id}))$$

*das **charakteristische Polynom** von A (bzw. φ).*

Wir haben oben die Begründung gegeben, warum man das betrachten sollte. Wir halten das nochmal fest. Wir berechnen das mal in unserem Beispiel.

Unser Standardbeispiel. *Es war*

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

Dann bekommen wir

$$P_A = \det(A - \lambda E_2) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1.$$

Die Nullstellen sind in der Tat die Eigenwerte und diese sind genau die Diagonalelemente der Diagonalmatrix D, die wir am Anfang gewählt haben.

Wie hier im Besipiel gilt ganz allgemein, dass man die Eigenwerte als Nullstellen des charakteristischen Polynoms bekommt.

Lemma. *Es sei $A \in \mathbb{C}^{(n,n)}$ (bzw. eine lineare Abbildung $\varphi\colon V \to V$) gegeben. Dann sind die Eigenwerte genau die Nullstellen des charakteristischen Polynoms.*

Jetzt sind wir endlich in der Lage die Eigenwerte zu berechnen. Dazu müssen wir das charakteristische Polynom berechnen. Wie kommen wir dann auf die Eigenwerte? Das sind ganz einfach die Nullstellen des eben berechneten Polynoms. Was ist aber mit den Eigenvektoren zum Eigenwert λ? Die brauchen wir doch auch, um die Basiswechselmatrix $\mathcal{X}$ zu bestimmen. Das ist nun wieder ganz einfach. Um die Eigenvektoren zu A (bzw. φ) zu berechnen, muss man einfach nur den Kern von

$$A - \lambda E_n \text{ (bzw. } \varphi - \lambda\,\mathrm{id}\,)$$

berechnen. Jeder Vektor $\neq 0$ in dem Kern ist ein Eigenvektor. Und das kennt man ja, dass man den Kern über den Algorithmus von Gauß berechnen kann. Wir zeigen dies wieder an unserem Beispiel auf.

Unser Standardbeispiel. *Wir hatten bereits berechnet, dass*

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

die Eigenwerte 1 *und* -1 *hat. Wir wollen jetzt zu jedem Eigenwert einen Eigenvektor finden (Allgemein muss man immer möglichst viele linear unabhängige Eigenvektoren zu einem festen Eigenwert finden, das muss nicht immer nur einer sein). Eigenvektor zum Eigenwert* 1*: Wir müssen die Gleichung*

$$(A - E_2)v = 0$$

lösen. Diese Lösungsmenge ist ganz explizit

$$\operatorname{Kern}\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle.$$

Dann ist zum Beispiel

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

ein Eigenvektor zum Eigenwert 1*. Aber auch*

$$-2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

ist ein Eigenvektor. Man beachte, dass aber 0 *(obwohl Lösung der Gleichung* $(A-E_2)v = 0$*) kein Eigenvektor nach Definition ist. Analog erhält man einen Eigenvektor zum Eigenwert* -1*, wenn man die Gleichung*

$$(A + E_2)v = 0$$

löst. Man kommt dann auf die Lösungsmenge

$$\operatorname{Kern}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle,$$

wie auch zu erwarten war.

In dem Zusammenhang gibt es noch weitere Definitionen, die man kennen sollte.

Definition (Eigenraum und Vielfachheiten)**.** *Es sei* $A \in \mathbb{C}^{(n,n)}$ *(bzw. eine lineare Abbildung* $\varphi\colon V \to V$*) gegeben. Es sei* λ *ein Eigenwert. Den Vektorraum*

$$U_\lambda = \operatorname{Kern}(A - \lambda E_n) \text{ (bzw. } U_\lambda = \operatorname{Kern}(\varphi - \lambda \operatorname{id})\text{)}$$

*nennt man den **Eigenraum** zu* λ*.*
Man nennt

$$\dim_{\mathbb{C}} U_\lambda$$

*die **geometrische Vielfachheit** und die Vielfachheit der Nullstelle* λ *vom charakteristischen Polynom die **algebraische Vielfachheit** von* λ*.*

Es ist nun nicht ganz elementar, dass folgendes gilt.

Satz. *Es sei $A \in \mathbb{C}^{(n,n)}$ (bzw. eine lineare Abbildung $\varphi\colon V \to V$) gegeben. Dann gilt*

(a) Die geometrische Vielfachheit eines Eigenwertes ist nach unten durch 1 *und nach oben durch die algebraische Vielfachheit beschränkt.*

(b) Die Eigenräume zu verschiedenen Eigenwerten haben nur die Null gemeinsam.

(c) A (bzw. φ) ist genau dann diagonalisierbar, wenn die Summe der geometrischen Vielfachheiten n ergibt.

(d) A (bzw. φ) ist genau dann diagonalisierbar, wenn die geometrischen und algebraischen Vielfachheiten übereinstimmen.

Unser Standardbeispiel. *In unserem Beispiel zeigte sich in der letzten Rechung, dass die geometrischen Vielfachheiten (= Dimension der Eigenräume) Eins waren. Die algebraischen Vielfachheiten waren ebenfalls Eins, wie man am charakteristischen Polynom ablesen kann. Dies spiegelt die Tatsache wieder, dass A diagonalisierbar ist. Auch die Aussage (b) aus dem Satz lässt sich leicht am Beispiel verifizieren.*

Es gibt noch viel über Diagonalisierbarkeit und Eigenwerte zu sagen. Wir wollen es vorest damit bewenden lassen. Wir kommen aber nochmal darauf zurück, wenn wir einfache Kriterien für Diagonalisierbarkeit betrachten wollen.

3 Anwendungsbeispiele*

Dieser Abschnitt spielte keine Rolle im Crashkurs. Es ist aber doch schön mal zu sehen, was man alles mit Eigenvektoren machen kann und wo man diese in der Anwendung finden kann. Diese kleine Liste ist -und muss es bei der Fülle der Anwendungen auch sein- lückenhaft. Außerdem habe ich Dinge rausgesucht, die natürlich für meine Begriffe besonders schön und wichtig (und mir gerade eingefallen) sind. Damit treffe ich natürlich nicht jeden Geschmack und jedes Gebiet.

- Als erstes möchte ich -weil natürlich für die Vorlesung von besonderem Interesse- nennen, dass man mittels Diagonalisierung eine **lineare gewöhnliche Differentialgleichung** lösen kann. Ich verliere hier mal kein weiteres Wort dazu, weil das auch Stoff der Vorlesung sein soll.

- Der **PageRank**-Algorithmus ist ein von Lawrence Edward Page und Sergei Michailowitsch Brin vorgestelltes Verfahren zur Berechung der Bedeutung von Webseiten. Dieses patentierte Verfahren bildet noch heute die Grundlage für ihre Suchmaschiene Google. Der sicherlich ein wenig modifizierte Algorithmus ist in seinen Details aber geheim. Klar ist aber auf jeden Fall, dass Eigenvektoren eine Rolle spielen. Die Information über die Bedeutung der Webseiten liegt in einem bestimmten Eigenvektor verborgen. Dieser wird auch Googles Eigenvektor genannt.

- Die **Hauptachsentransformation** ist ein recht allgemeines Verfahren zur Reduktion von Datenmengen und Vereinfachung von Quadriken. Dieses Verfahren ist Grundlage für viele Anwendungen in der Wirtschaft und zählt dort zu den Standardwerkzeugen. Von jedem Absolventen der Mathematik wird erwartet, dass er dieses Verfahren versteht und anwenden kann. Oft verwendet man den Namen Hauptachsentransformation auch einfach für die Diagonalisierung von symmetrischen Matrizen, weil dies genau das ist, was bei der Hauptachsentransformation passiert.

- **Lineare Modelle** in der Statistik sind in ihrer Grundform auch nur eine Hauptachsentransformation.

- Der **Strukturtensor** ist ein Begriff aus der Bildverarbeitung. Aus einem lokalen Ausschnitt berechnet man eine Matrix, die man Strukturtensor nennt. Aus der Größe der Eigenwerte dieser Matrix kann man ablesen, ob es sich bei dem lokalen Ausschnitt um eine Kante, eine Ecke, oder eine Fläche handelt (Bei Videosequenzen kann man auch viel Information über Bewegung erhalten, z.B. Gleichförmigkeit der Bewegung und Geschwindigkeit). Auch dies ist eine Anwendung der Hauptachsentransformation.

- Ein weiteres Beispiel für die Anwendung der Hauptachsentransformation ist die **Gesichtserkennung**. In dem Zusammenhang spricht man statt von Eigenvektoren auch von Eigengesichtern (engl. eigenfaces).

- Lösungsmethode für **lineare Rekursionsgleichungen**. Dies nichts weiter als das diskrete Analogon für die Lösungsmethode der linearen gewöhnlichen Differentialgleichungen. Dies ist nur etwas elementarer und wird oft als Anwendung in der linearen Algebra diskutiert.

- ...

4 Rechenbeispiele

Man lernt sowas eigentlich nur, wenn man selber etwas damit arbeitet. Deshalb werden wir im Crashkurs auch darauf Wert legen, dass Sie selber einige Aufgaben rechnen.

4.1 Zusammenfassung der bisherigen Beobachtungen

Hier eine Warnung. Alleiniges Lernen von Rechenverfahren bringt Ihnen nichts. Sie müssen diese auch verstehen. Wenn Sie ein Verfahren verstanden haben, können Sie dies deutlich länger anwenden und im Notfall auch herleiten. Wenn man Dinge verstanden hat, dann muss man nicht viel in der Mathematik auswendig lernen. Lernen Sie dagegen immer nur die Verfahren, dann findet das kein Ende. Sollten sie einen Schritt nicht verstehen, dann sollte man sich nochmal die obigen Bemerkungen anschauen (, Kommilitonen fragen, oder andere Quellen benutzen), damit man das versteht. Außerdem kann es sein, dass im

gegebenen Fall dies deutlich einfacher geht. So passiert es auch hier gelegentlich. Wenn man nur stupide auswendig lernt, dann wird man in den Fällen mehr rechnen müssen und das kostet wertvolle Zeit.
Wenn wir eine Matrix $A \in \mathbb{C}^{(n,n)}$ diagonalisieren wollen, dann kann man so vorgehen:

1. Man bestimme das charakteristische Polynom $P_A(\lambda) = \det(A - \lambda E_n)$.

2. Man bestimme die Eigenwerte von A, also die Nullstellen von P_A.

3. Man berechne mit dem Algorithmus von Gauß eine Basis der Eigenräume für jeden Eigenwert.

4. Sind die Vektoren aus dem Schritt 3 zusammengenommen genau n viele, dann ist die Matrix diagonalisierbar. Wenn nicht, dann ist die Matrix nicht diagonalisierbar. Für den letzten Schritt kann man dann davon ausgehen, dass es n Vektoren sind.

5. Seien $v_1, ..., v_n$ die Vektoren aus dem Schritt 4 und $\lambda_1, ..., \lambda_n$ die zugehörigen (nicht notwendig verschiedenen) Eigenwerte zu den Eigenvektoren. Man setze dann

$$\mathcal{X} = (v_1 \; v_2 \; ... \; v_n)\,, \; D = \operatorname{diag}(\lambda_1, ..., \lambda_n).$$

   Mit diesen Matrizen gilt $\mathcal{X} \in \mathrm{GL}(n, \mathbb{C})$ und

$$A = \mathcal{X} D \mathcal{X}^{-1}.$$

4.2 Beispiel 1 - Eine Tafelaufgabe

Diese Aufgaben wurden an der Tafel vorgerechnet. Die anderen Aufgaben dagegen von euch selbständig erarbeitet. Hier gebe ich auch nur für die Tafelaufgaben Lösungen. Die Lösungen der anderen Aufgaben werden eventuell kurz in dem Crashkurs besprochen.
Welche der folgenden Matrizen sind diagonalisierbar? Man finde die Basiswechselmatrix $\mathcal{X}$ gegebenenfalls. Man bestimme von einer der Matrizen auch A^{1002}.

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & -1 \\ -3 & -1 \end{pmatrix}$$

Wir beginnen mit A_1 und gehen mal die Berechnungsschritte durch

1. P_{A_1} bestimmen:

$$P_{A_1} = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

2. Die Eigenwerte von A_1 sind die Nullstellen von P_{A_1} und damit ist 1 der einzige Eigenwert von A_1.

3. Der Eigenraum zum Eigenwert 1: Wir müssen lösen

$$(A_1 - E_n)v = 0.$$

   Der Lösungsraum ist aber genau $\langle e_1 \rangle$. Eine Basis davon ist $\{e_1\}$.

4. Wir haben zu wenig linear unabhängige Eigenvektoren. Wir brauchen also nicht weitermachen. Die Matrix A_1 ist nicht diagonalisierbar.

Machen wir weiter mit A_2:

1. P_{A_2} bestimmen:

$$P_{A_2} = \det \begin{pmatrix} 1-\lambda & -1 \\ -3 & -1-\lambda \end{pmatrix} = (1-\lambda)(-1-\lambda) - 3 = \lambda^2 - 4.$$

2. Die Eigenwerte von A_2 sind die Nullstellen von P_{A_2} und damit sind ± 2 die einzigen Eigenwerte von A_2.

3. a) Der Eigenraum zum Eigenwert 2: Wir müssen lösen

$$(A_2 - 2E_n)v = \begin{pmatrix} -1 & -1 \\ -3 & -3 \end{pmatrix} v = 0.$$

Der Lösungsraum ist aber genau

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

Eine Basis ist etwa gegeben durch

$$\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

b) Der Eigenraum zum Eigenwert -2: Wir müssen lösen

$$(A_2 + 2E_n)v = \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix} v = 0.$$

Der Lösungsraum ist aber genau

$$\left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\rangle.$$

Eine Basis ist etwa gegeben durch

$$\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}.$$

4. Wir haben genug Eigenvektoren. Die Matrix A_2 ist also diagonalisierbar und wir können weitermachen.

5. Wenn wir setzen

$$\mathcal{X} = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}, \ D = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix},$$

dann erhalten wir

$$A_2 = \mathcal{X} D \mathcal{X}^{-1}.$$

Wir wollen jetzt noch A_2^{1002} berechnen. Wir wissen bereits

$$A_2^{1001} = \mathcal{X} D^{1002} \mathcal{X}^{-1} = 2^{501} E_2,$$

weil $D^{1002} = 2^{501} E_2$ ist.

4.3 Beispiel 2 - Diagonalisieren von Matrizen

Wir wollen nun die folgenden Matrizen diagonalisieren:

$$A_1 = \begin{pmatrix} i & 1 \\ 0 & -i \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

Hier sollte man A_2 und A_4 auf alle Fälle mal versuchen!

4.4 Beispiel 3 - Man kann nicht immer diagonalisieren

Welche der folgenden Matrizen sind diagonalisierbar?

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

4.5 Beispiel 4 - Potenzen einer Matrix berechnen

Wir wollen die Potenzen der Matrix

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

berechnen. Speziell zum Beispiel A^{1000}.

4.6 Beispiel 5 - Rechenintensiveres Diagonalisieren, oder doch nicht?

Welche der folgenden Matrizen kann man diagonalisieren? Man führe gegebenenfalls eine Diagonalisierung durch.

$$A_1 = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 3 & 1 \\ 2 & 0 & 2 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 1 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 0 & 1 & -1 \\ 0 & -2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

$$A_5 = \begin{pmatrix} 1 & i & 25 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \qquad A_6 = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 3 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

Wenn man nur stur nach Schmema F vorgeht, dann sind manche der Aufgaben schwieriger zu lösen. Kleine Überlegungen, geschicktes ausrechnen der Determinante und ähnliches erleitert die Berechnung ungemein.

4.7 Beispiel 6 - Fragen, die man beantworten kann, wenn man schneller fertig ist

Richtig oder falsch (alle Gegenbeispiele, die man benötigt sind in den anderen Aufgaben zu finden)?

(a) Ist A diagonalisierbar, dann ist die Diagonalmatrix D eindeutig bestimmt.

(b) Ist A diagonalisierbar, dann ist die Basiswechselmatrix eindeutig bestimmt.

(c) Ist A diagonalisierbar, dann ist die Basiswechselmatrix bis auf Multiplikation der Spalten mit Skalaren eindeutig bestimmt.

(d) Ist A eine reelle Matrix, dann hat A nur reelle Eigenwerte.

(e) Ist A eine reelle Matrix und $\lambda \in \mathbb{C}$ ein Eigenwert mit Eigenvektor $v \in \mathbb{C}^n$. Dann ist $\overline{v}$ ein Eigenvektor zum Eigenwert $\overline{\lambda}$.

(f) Hat eine Matrix nur einen Eigenwert λ und ist die Matrix diagonalisierbar, dann ist $A = \lambda E_n$.

(g) Hat eine Matrix nur einen Eigenwert λ, dann ist $A = \lambda E_n$.

5 Wann kann man diagonalisieren? - Einfache Kriterien für Diagonalisierbarkeit

Wir haben jetzt einige Beispiele gerechnet. Außerdem haben wir gesehen, wie man entscheiden kann, ob eine Matrix diagonalisierbar ist. In vielen Fällen kann man das aber bereits ohne Rechnen vorhersagen. Neben den allgemeinen Kriterien vom zweiten Abschnitt, findet man auch viele nützliche Kriterien, die zwar notwendig, aber nicht hinreichend sind.
Um solche Kriterien soll es hier gehen.

Satz. *Hat eine $n \times n$-Matrix A genau n* ***verschiedene*** *Eigenwerte, dann ist A diagonalisierbar.*

Definition. *Eine Matrix $A \in \mathbb{C}^{(n,n)}$ heißt*

(a) ***symmetrisch****, falls $A \in \mathbb{R}^{(n,n)}$ und $A = A^T$ gilt.*

(b) ***selbstadjungiert*** *oder* ***hermitesch****, falls $A = \overline{A}^T$ gilt.*

(c) ***orthogonal****, falls $A \in \mathbb{R}^{(n,n)}$ und $A^T A = E_n$ gilt.*

(d) ***unitär****, falls $\overline{A}^T A = E_n$ gilt.*

(e) ***normal****, falls $A\overline{A}^T = \overline{A}^T A$ gilt.*

Man nennt weiter A ***unitär diagonalisierbar****, wenn man eine unitäre Matrix $\mathcal{X}$ findet mit*

$$A = \mathcal{X} D \mathcal{X}^{-1}.$$

Wenn man $\mathcal{X}$ sogar orthogonal wählen kann, dann nennt man die Matrix A ***orthogonal diagonalisierbar****. Insbesondere ist in diesen Fällen A also diagonalisierbar.*

Im Gegensatz zur Diagonalisierbarkeit von Matrizen, die man nicht einfach ablesen kann, ist dies bei der unitären Diagonalisierbarkeit ganz anders. Dies zeigt der folgende Satz.

Satz. *Eine Matrix ist genau dann unitär diagonalisierbar, falls die Matrix normal ist.*

Wir sammeln nun einige Spezialfälle und notieren uns noch, was man in diesen Spezialfällen über die Eigenwerte weiß. Dazu beachte man, dass klarerweise jede unitäre, orthogonale, hermitsche und symmetrische Matrix normal ist.

Korollar. *Jede unitäre Matrix ist unitär diagonalisierbar. Die Eigenwerte haben zudem den Betrag Eins.*

Korollar. *Jede orthogonale Matrix ist unitär diagonalisierbar. Die Eigenwerte treten in Paaren $\lambda, \overline{\lambda}$ mit $|\lambda| = 1$ auf oder sind ± 1.*

Korollar. *Jede hermitesch Matrix ist unitär diagonalisierbar. Die Eigenwerte sind alle reell.*

Korollar. *Jede symmetrische Matrix ist othogonal diagonalisierbar. Die Eigenwerte sind alle reell.*

6 Ein paar letzte Beispiele

Welche der folgenden Matrizen sind diagonalisierbar?
Man schaue sich A_4 und A_7 besonders genau an.

$$A_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} \qquad A_2 = \begin{pmatrix}1 & i\\ -i & 0\end{pmatrix}$$

$$A_3 = \begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix} \qquad A_4 = \begin{pmatrix}1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & 2\end{pmatrix}$$

$$A_5 = \begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} \qquad A_6 = \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}$$

$$A_7 = \begin{pmatrix}2 & i\\ i & 0\end{pmatrix} \qquad A_8 = \begin{pmatrix}i & i\\ i & 0\end{pmatrix}$$

Man rechnet leicht nach, dass A_1 unitär (sogar orthogonal) ist. Weiter ist A_2 hermitesch. Also sind A_1 und A_2 diagonalisierbar (sogar unitär diagonalisierbar). Die Matrix A_3 hingegen ist nicht diagonalisierbar (die geometrische Vielfachheit von 1 ist 1, aber die algebraische ist 2). Die Matrix A_4 ist dagegen nicht unitär diagonalisierbar (da nicht normal), aber diagonalisierbar (weil sie 3 verschiedene Eigenwerte hat). Das kann also auch passieren. Die beiden Matrizen A_5 und A_6 (A_6 ist sogar symmetrisch!) sind normal und damit unitär diagonalisierbar. Bei der Matrix A_7 kann man leicht in die Falle laufen und sagen, dass $A_7 = A_7^T$ gilt und damit A_7 diagonalisierbar ist. Dies ist aber nicht richtig. A_7 ist nicht reell. Deshalb reicht dies nicht. Und in der Tat ist A_7 nicht diagonalisierbar. Der einzige Eigenwert ist 1 und der hat eine geometrische Vielfachheit von 1 (man bräuchte aber eine geometrische Vielfachheit von 2). A_8 dagegen ist wieder normal und damit diagonalisierbar.

7 Randbemerkung: Jordanform

Wir haben nun einige Besipiele gesehen von diagonalisierbaren Matrizen. Einige Kriterien gefunden, wann man diagonalisieren kann. Es kann aber auch passieren, dass man schlicht und einfach nicht diagonalisieren kann. Ein solches Beispiel ist

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

Man sieht leicht, dass der einzige Eigenwert 0 ist. Die Matrix ist aber nicht diagonalisierbar, da der einzige Eigenvektor e_1 ist. Man bekommt also keine Basis von Eigenvektoren, wie man es bräuchte.
Allgemeiner hat man die Matrix der Form (Einsen in der ersten Subdiagonalen und sonst

nur Nullen)

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

dann zeigt das gleiche Argument, dass diese Matrix auch nicht diagonalisierbar ist. Was soll man aber machen, wenn man Matrixpotenzen ausrechnen will oder muss?
Der Ausweg ist die **Jordanform**, die auch zum Grundrepertoire des Mathematikers gehört. Es stellt sich heraus, dass das hier gezeigte Beispiele (in einer noch zu bestimmenden Art) das schlimmste ist, was man befürchten muss.

Literatur

[Maier12] *Vorabskript zur Vorlesung Lineare Algebra I und II*, Prof. Dr. Helmut Maier, 2012