Eigenwerttheorie

Wiederholungsvortrag zur Klausurvorbereitung an der
TU Darmstadt von Michael Thürauf und Florian Hopf

21.08.2009

Inhaltsverzeichnis

1 Grundlagen und Diagonalisierung

1.1 Grundlagen

In den folgenden Kapiteln beschränken wir uns auf Endomorphismen $T : V \to V$ mit derselben Basis in Bild und Urbildraum. Wir möchten nun eine Basis $\mathcal{B}$ des Vektoraums V finden, so dass die Matrix $M_T^{\mathcal{B},\mathcal{B}}$ möglichst einfach wird.

Definition: Eigenwert, Eigenvektor, Eigenraum, Spektrum
Sei $T : V \to V$ ein Endomorphismus, dann heißt $\lambda \in \mathbb{K}$ *Eigenwert* (EW) von T, falls es $\vec{x} \in V \setminus \{0\}$ gibt mit $T\vec{x} = \lambda\vec{x}$.
$\vec{x} \neq 0$ heißt dann *Eigenvektor* (EV) zum Eigenwert λ

Des weiteren heißt

$$E_\lambda^T = \{\vec{x} \in V : T\vec{x} = \lambda\vec{x}\} = \{\vec{x} \in V : (T - \lambda \cdot \mathbb{1})\, \vec{x} = 0\} = \ker (T - \lambda \cdot \mathbb{1})$$

Eigenraum (ER) von T zum Eigenwert λ.

Sei $\lambda \in \mathbb{K}$ Eigenwert von T, dann ist

$$\sigma (T) = \{\lambda \in \mathbb{K} : \lambda \text{ ist EW von } T\}$$

das *Spektrum* von T

Beispiel: Spiegelung in $\mathbb{R}^2$ an $y = x$

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} \text{ ist EV zum EW } \lambda = 1$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ ist EV zum EW } \lambda = -1$$

Bemerkungen zum Eigenraum:

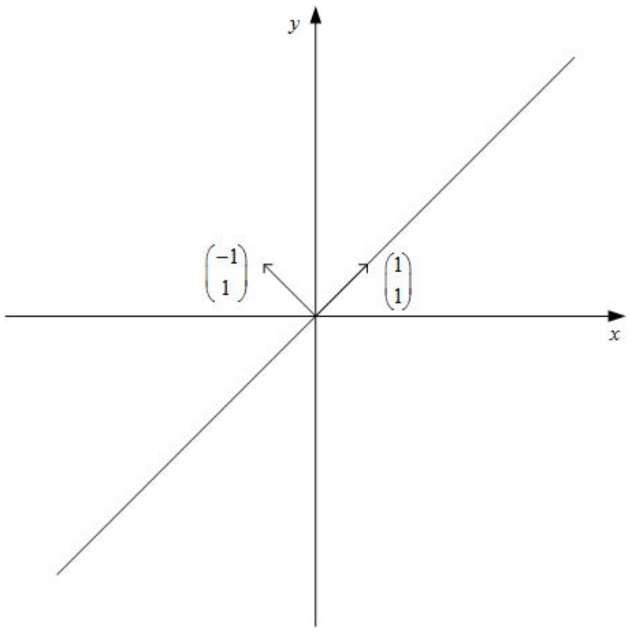

Abbildung 1: Spiegelung an der Winkelhalbierenden

E_λ^T ist linearer Teilraum von V. Gilt $E_\lambda^T = \{0\}$, so ist λ kein Eigenwert. Der Eigenraum ist also mindestens eindimendional.

Definition: geometische Vielfachheit
Die Dimension des Eigenraums heißt *geometrische Vielfachheit* (gV) des Eigenwertes λ. Sei die geometrische Vielfachheit eines Eigenraumes k, so ist λ k-fach entartet.

$$k = \mathrm{gV}\,(\lambda) := \dim E_\lambda^T$$

Definition: Diagonalisierbarkeit
T heißt *diagonalisierbar*, falls es eine Basis aus Eigenvektoren gibt

1.2 Das charakteristische Polynom

Um eine Matrix zu diagonalisieren, also sie möglichst einfach zu machen, braucht man die Eigenwerte und Eigenvektoren einer linearen Abbildung. Das Werkzeug zum Berechnen der Eigenwerte ist das charakteristische Polynom. Sind die Eigenwerte λ_i bekannt, so erhält man durch Lösen des LGS $(T - \lambda_i \cdot \mathbb{1})\,\vec{x} = 0$ die Eigenvektoren zum EW λ_i.
Bevor man sich aber mit dem charakteristischen Polynom beschäftigt, sollte

man sich folgende Sätze anschauen, die hier nur angegeben und nicht bewiesen werden, da das den Rahmen sprengen würde.

Satz:
Sei $T : V \to V$ und $\mathcal{B} = \left\{\vec{b}_1, ..., \vec{b}_n\right\}$ Basis von V, dann sind äquivalent

i) $\mathcal{B}$ besteht aus Eigenvektoren.

ii) $M_T^{\mathcal{B},\mathcal{B}}$ ist diagonal und es gilt

$$M_T^{\mathcal{B},\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

wobei λ_i die EW von T sind

Bemerkung:

- Hat $M_T^{\mathcal{B},\mathcal{B}}$ die obige Form, dann gilt

$$M_{T^{-1}}^{\mathcal{B},\mathcal{B}} = \begin{pmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_n} \end{pmatrix}$$

- die Determinante von T ist $\det T = \lambda_1 \cdot ... \cdot \lambda_n$

Satz:
Seien $\lambda_1, ...\lambda_k \in \mathbb{K}$ verschiedene Eigenwerte mit Eigenvektoren $\{\vec{v}_1, ..., \vec{v}_k\}$, dann sind die Vektoren linear unabhängig.

Satz:
Sei $\sigma(T) = \{\lambda_1, ...\lambda_k\}$, $\dim E_{\lambda_1}^T = n_1$ mit der Basis $\mathcal{B}_1 = \left\{\vec{v}_1^1, ..., \vec{v}_{n_1}^1\right\}$ und $\dim E_{\lambda_k}^T = n_k$ mit der Basis $\mathcal{B}_k = \left\{\vec{v}_1^k, ..., \vec{v}_{n_1}^k\right\}$, dann ist

$$\mathcal{B}_1 \bigcup ... \bigcup \mathcal{B}_k = \bigcup_{i=1}^{k} \mathcal{B}_i = \left\{\vec{v}_1^1, ..., \vec{v}_{n_1}^1, ..., \vec{v}_1^k, ..., \vec{v}_{n_k}^k\right\}$$

linear unabhängig.

Satz:
$T : V \to V$ ist diagonalisierbar $\Leftrightarrow$ für die Zahlen n_i aus obigem Satz gilt

$$\sum_{i=1}^{k} n_i = n = \dim V$$

Nun können wir uns der Bestimmung der Eigenwerte zuwenden. Wie in der Einleitung zu diesem Kapitel schon erwähnt, geschieht dies über das charakteristische Polynom. Zunächst muss geklärt werden, was das charakteristische Polyom überhaupt ist.

Definition: charakteristisches Polynom
Der Ausdruck $\chi_T(\lambda) = \det(T - \lambda \cdot \mathbb{1})$ heißt *charakteristisches Polynom.*

Satz:
λ ist genau dann Eigenwert von T, wenn $\chi_T(\lambda) = 0$. Das heißt die Eigenwerte von T sind die Nullstellen des charakteristischen Polynoms von T. An dieser Stelle sei darauf hingewiesen, dass das charakteristische Polynom unabhängig von der gewählten Basis ist.

Definition: algebraische Vielfachheit
Ist λ eine Nullstelle k-ter Ordnung von $\chi_A(\lambda)$, dann ist k die *algebraische Vielfachheit* (aV) von λ

Korollar:

- Ist $\mathbb{K} = \mathbb{C}$, so hat $\chi_T(\lambda)$ eine Nullstelle (Hauptsatz der Algebra). Daraus folgt, dass ein Eigenwert und ein Eigenvektor existiert.
- Hat $\chi_T(\lambda)$ n verschiedene Nullstellen, dann ist T diagonalisierbar.
  n verschiedene Nullstellen $\Rightarrow$ n veschiedene EW $\Rightarrow$ n lin. unabhängige EV.

1.3 Eigenvektoren und Diagonalisierung

Mit dem charakteristischen Polynom haben wir ein Werkzeug gefunden, mit dem wir die Eigenwerte einer Abbildung T bestimmen können. Dies ist allerdings nur der erste Schritt auf dem Weg zu einer diagonalen Matrix von T. Als nächstes müssen wir die Eigenvektoren berechnen. Dazu setzten wir die berechneten Eigenwerte wieder in $T\vec{x} = \lambda\vec{x}$ ein. Durch Umstellen erhält man:

$$\begin{aligned} & T\vec{x} = \lambda\vec{x} \\ \Leftrightarrow & T\vec{x} - \lambda\vec{x} = 0 \\ \Leftrightarrow & (T - \lambda \cdot \mathbb{1})\,\vec{x} = 0 \end{aligned}$$

Das heißt wir müssen ein homogenes lineares Gleichungssystem lösen, dessen Lösung die Eigenvektoren zum EW λ sind.

Als nächstes Vergleicht man die algebraische Vielfachheit von λ mit der geometrischen Vielfachheit von λ. Stimmen diese für alle Eigenwerte überein, so ist T diagonalisierbar. Zur Probe kann man eine Transformationsmatrix S erstellen, deren Spalten aus dessen Eigenvektoren bestehen. Berechnet man nun $S^{-1}AS$, wobei A die Matrix von T in nicht diagonalform ist, erhält man die diagonale Form. Also gilt:

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \operatorname{diag}(\lambda_1, ..., \lambda_n)$$

1.4 Beispiele

Es wird nun Zeit das obige Verfahren auch auf Beispiele anzuwenden. Dabei sind folgende Punkte abzuarbeiten:

1. Berechne χ_A
2. Bestimme die Nullstellen λ_i von χ_A und $\operatorname{aV}(\lambda_i)$
3. Bestimme für jedes λ_i eine Basis und die Dimension von $\ker(T - \lambda \cdot \mathbb{1})$. Die Dimension gibt die $\operatorname{gV}(\lambda_i)$ an.
4. Überprüfe, ob $\operatorname{aV}(\lambda_i) = \operatorname{gV}(\lambda_i)$ für alle $\lambda_i \in \sigma(T)$ gilt.
5. Falls 4. gilt: Erstelle die Matrix S aus den EV
6. Rechne die Probe $S^{-1}AS$. Das Ergebnis muss die diagonale Matrix sein.

1. Beispiel $T : \mathbb{R}^2 \to \mathbb{R}^2 : \vec{x} \mapsto A\vec{x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \vec{x}$

1. $\chi_T(\lambda) = \lambda^2 - 1$
2. $\chi_T(\lambda) = 0 \Rightarrow \sigma(T) = \{\pm 1\}$ und $\operatorname{aV}(\lambda_i) = 1$
3. $\lambda_1 = 1 :$

$$(A - 1 \cdot \mathbb{1})\vec{x} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_1 = x_2$$

$\Rightarrow \vec{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist EV zum EW $\lambda = 1 \Rightarrow \mathrm{gV}(\lambda_1) = 1$
$\lambda_2 = -1$:

$$(A + 1 \cdot \mathbb{1})\,\vec{x} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_1 = -x_2$$

$\Rightarrow \vec{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist EV zum EW $\lambda = -1 \Rightarrow \mathrm{gV}(\lambda_2) = 1$

4. $\mathrm{gV}(\lambda_1) = 1 = \mathrm{aV}(\lambda_1)$
   $\mathrm{gV}(\lambda_2) = 1 = \mathrm{aV}(\lambda_2)$

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

$$\begin{aligned} S^{-1} &= \frac{1}{\det S} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ S^{-1}AS &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

2. Beispiel $T : \mathbb{R}^3 \to \mathbb{R}^3 : \vec{x} \mapsto A\vec{x} = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} \vec{x}$

1. $\chi_T(\lambda) = -\lambda^3 + 12\lambda + 16$

2. $\chi_T(\lambda) = 0 \Rightarrow \sigma(T) = \{-2, 4\}$ und $\mathrm{aV}(-2) = 2$, $\mathrm{aV}(4) = 1$

3. $\lambda_1 = -2$:

$$(A + 2 \cdot \mathbb{1})\,\vec{x} = \begin{pmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_1 - x_2 + x_3 = 0$$

$\Rightarrow \vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ und $\vec{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ sind EVen zum EW $\lambda = -2$
$\Rightarrow \mathrm{gV}(\lambda_1) = 2$

$\lambda_2 = 4:$

$$(A - 4\cdot \mathbb{1})\,\vec{x} = \begin{pmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_1 = x_2, x_3 = 2x_1$$

$\Rightarrow \vec{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ ist EV zum EW $\lambda = 4 \Rightarrow \mathrm{gV}(\lambda_2) = 1$

4. $\mathrm{gV}(\lambda_1) = 2 = \mathrm{aV}(\lambda_1)$
   $\mathrm{gV}(\lambda_2) = 1 = \mathrm{aV}(\lambda_2)$

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

$$S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$S^{-1}AS = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

3. Beispiel
Interessant ist noch die Diagonalisierung der Drehmatrix

$$T : \mathbb{R}^2 \to \mathbb{R}^2 : \vec{x} \mapsto D\vec{x} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \vec{x}$$

1. $\chi_T(\lambda) = \lambda^2 - 2\lambda\cos\varphi + 1$

2. $\chi_T(\lambda) = 0 \Rightarrow \sigma(T) = \{\cos\varphi \pm i\sin\varphi\}$ und $\mathrm{aV}(\lambda_i) = 1$

3. $(D - (\cos\varphi \pm i\sin\varphi)\cdot\mathbb{1}) = \begin{pmatrix} \pm i\sin\varphi & -\sin\varphi \\ \sin\varphi & \pm i\sin\varphi \end{pmatrix} = \sin\varphi \begin{pmatrix} \pm i & -1 \\ 1 & \pm i \end{pmatrix}$
   $\lambda_1 = \cos\varphi + i\sin\varphi$:
   $ix_1 - x_2 = 0 \Rightarrow \vec{x}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ ist EV zum EW $\lambda = \cos\varphi + i\sin\varphi$
   $\Rightarrow \mathrm{gV}(\lambda_1) = 1$
   $\lambda_2 = \cos\varphi - i\sin\varphi$:
   $-ix_1 - x_2 = 0 \Rightarrow \vec{x}_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ ist EV zum EW $\lambda = \cos\varphi - i\sin\varphi$
   $\Rightarrow \mathrm{gV}(\lambda_2) = 1$

4. $\mathrm{gV}(\lambda_1) = 1 = \mathrm{aV}(\lambda_1)$
   $\mathrm{gV}(\lambda_2) = 1 = \mathrm{aV}(\lambda_2)$

5. $S = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$

$$S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$
$$S^{-1}AS = \begin{pmatrix} \cos\varphi + i\sin\varphi & 0 \\ 0 & \cos\varphi - i\sin\varphi \end{pmatrix}$$

2 Minimalpolynom

Betrachte eine beliebige $n \times n$ Matrix $A \in M_{n,n}(\mathbb{K})$ mit Koeffizienten aus $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Aufgrund der Tatsache, dass man Potenzen quadratischer Matrizen bilden kann, eröffnet sich die Möglichkeit, Matrizen in Polynome einzusetzen. Für ein Polynom $p(X) = \sum_{i=0}^{n} a_i X^i$ sei dann

$$p(A) := \sum_{i=0}^{n} a_i A^i = a_0 \cdot \mathbb{1} + a_1 \cdot A + ... + a_n \cdot A^n$$

mit der Vereinbarung $A^0 := \mathbb{1}$. Das Ergebnis einer solchen polynomiellen Auswertung ist wieder eine $n \times n$-Matrix. Handelt es sich hierbei um die 0-Matrix, so sagt man, die Matrix A wird von p ***annuliert***. Die Menge aller Polynome, welche eine gegebene Matrix A annulieren, hat besondere Eigenschaften und bekommt einen eigenen Namen:

Definition:
Sei $A \in M_{n,n}(\mathbb{K})$ eine $n \times n$-Matrix über $\mathbb{K}$. Dann heißt

$$\mathcal{I}_A := \{p \in \mathbb{K}[X] : p(A) = 0\}$$

das ***Ideal von*** A.

Die Elemente des Ideals von A kann man relativ leicht charakterisieren. Man kann zeigen:

Satz:
Es gibt ein eindeutig bestimmtes Polynom $\mu_A \in \mathcal{I}_A$ mit den folgenden Eigenschaften

1. μ_a ist normiert, d.h. der Leitkoeffizient ist 1.

2. μ_a hat unter allen $p \in \mathcal{I}_A$ minimalen Grad.

3. Für jedes $p \in \mathcal{I}_A$ gibt es ein $q \in \mathbb{K}[X]$ mit $p = q \cdot \mu_A$

Das Polynom μ_A heißt ***Minimalpolynom von*** A.

Mit anderen Worten: Jedes Polynom im Ideal von A kann geschrieben werden als Vielfaches des Minimalpolynoms. Eine äquivalente Formulierung dieser Aussage ist, dass jedes Polynom im Ideal von A vom Minimalpolynom geteilt wird. Die Forderung des minimalen Grades und der Normiertheit sorgen hierbei für die Eindeutigkeit.

Beispiel: Wie erkennt man das Minimalpolynom?
Betrachte die Matrix

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

Behauptung: Das Minimalpolynom von A ist $\mu_A = X^2$.

Zunächst beobachten wir, dass

$$A^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

d.h. A wird in der Tat von X^2 annuliert, also ist $X^2 \in \mathcal{I}_A$. Weiter ist X^2 offensichtlich normiert. Es bleibt noch zu zeigen, dass X^2 minimalen Grad hat. Die möglichen kleineren Grade sind $n = 0$ und $n = 1$. Der Fall $n = 0$ kann ausgeschlossen werden, da jedes Polynom 0-ten Grade konstant ist, d.h. $p = c$. Die Forderung normiert zu sein impliziert dann $c = 1$. Dann ist aber $p(A) = \mathbb{1} \neq 0$, sodass $p \notin \mathcal{I}_A$. Ein normiertes Polynom vom Grad 1 hat die Form $p = X + c$ für ein $c \in \mathbb{K}$. Dann folgt aber

$$p(A) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} = \begin{pmatrix} c & 1 \\ 0 & c \end{pmatrix} \neq 0$$

sodass auch der Fall $n = 1$ ausgeschlossen werden kann. Folglich ist in der Tat $\mu_A = X^2$.

Im Allgemeinen kann man das Minimalpolynom einer Matrix nicht einfach erraten. Zwei theoretische Resultate erlauben jedoch eine systematische Herangehensweise an diese Problematik:

Satz von CAYLEY-HAMILTON:

Sei $A \in M_{n,n}(\mathbb{K})$ eine $n \times n$-Matrix über $\mathbb{K}$ und χ_A deren charakteristisches Polynom. Dann gilt

$$\chi_A(A) = 0$$

d.h. A wird von ihrem eigenen charakteristischen Polynom annuliert.

Insbesondere ist damit $\chi_A \in \mathcal{I}_A$ und folglich wird χ_A von μ_A geteilt. Das ist bereits eine sehr starke Einschränkung und würde es im Prinzip ermöglichen, das Minimalpolynom zu bestimmen. Man müsste einfach alle Teiler von χ_A aufschreiben, sie nach Grad sortieren und normieren und danach mit kleinstem Grad beginnend überprüfen, ob sie A annulieren. Dabei stellen sich jedoch zwei Probleme:

1. Um die Teiler von χ_A zu finden, muss man seine Nullstellen (also die Eigenwerte von A) kennen (s.u.).

2. Im Allgemeinen hat ein Polynom sehr viele Teiler, sodass der Rechenaufwand schnell unverhältnismäßig hoch wird.

Gegen das erste Problem ist kein Kraut gewachsen. Wenn man die Nullstellen von χ_A nicht kennt, kann man in der Regel nichts machen. Für das zweite Problem gibt es aber ein nützliches Werkzeug:

Satz:
Sei $A \in M_{n,n}(\mathbb{K})$ eine $n \times n$-Matrix über $\mathbb{K}$. Dann ist jede Nullstelle von χ_A auch Nullstelle von μ_A.

Jetzt können wir ein halbwegs brauchbares Verfahren angeben, mit dem man das Minimalpolynom von A mit erträglichem Aufwand errechnen kann.

1. Bestimme χ_A und finde seine Nullstellen $\lambda_1, ..., \lambda_n$. Mit anderen Worten, finde seine Linearfaktorzerlegung

$$\chi_A(\lambda) = c \cdot (\lambda_1 - \lambda)^{a_1} \cdot ... \cdot (\lambda_n - \lambda)^{a_n}$$

2. Nach dem Satz von CAYLEY-HAMILTON wird χ_A von μ_A geteilt. Der obige Satz sagt jedoch zusätzlich aus, dass jede Linearfaktor von χ_A auch wenigstens einmal in μ_A auftauchen muss. Die möglichen normierten Polynome, die diese Bedingungen erfüllen, sind daher von der Form

$$p(\lambda) = (\lambda_1 - \lambda)^{b_1} \cdot ... \cdot (\lambda_n - \lambda)^{b_n} \quad \text{mit } 1 \leq b_1 \leq a_1, ..., 1 \leq b_n \leq a_n$$

3. Sortiere die in Schritt 2 gefundenen Polynome nach ihrem Grad und berechne $p(A)$, beginnend mit dem kleinsten Grad. Gilt für ein solches Polynom $p(A) = 0$, so ist es das Minimalpolynom.

Beispiel: Wie berechnet man das Minimalpolynom?
Wir möchten das Minimalpolynom von

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

bestimmen. Es gilt

$$\chi_A(\lambda) = \det(A - \lambda \cdot \mathbb{1}) = (2-\lambda)^3(3-\lambda)$$

Die möglichen Minimalpolynome sind daher

$$\begin{array}{rcl} p_1(\lambda) & = & (2-\lambda)(3-\lambda) \\ p_2(\lambda) & = & (2-\lambda)^2(3-\lambda) \\ p_3(\lambda) & = & (2-\lambda)^3(2-\lambda) \end{array}$$

Für das erste Polynom gilt

$$p_1(A) = (2 \cdot \mathbb{1} - A)(3 \cdot \mathbb{1} - A) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Es kann also nicht das Minimalpolynom sein. Das zweite Polynom ergibt dagegen

$$p_2(A) = (2 \cdot \mathbb{1} - A)^2(3 \cdot \mathbb{1} - A) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Folglich gilt $\mu_A = p_2$

3 Jordannormalform

Wie wir gesehen haben, ist eine Matrix A genau dann diagonalisierbar, wenn zum einen ihr charakteristisches Polynom in Linearfaktoren zerfällt (was über

$\mathbb{C}$ immer der Fall ist) und zum anderen die geometrische Vielfachheit eines jeden Eigenwertes gleich seiner algebraischen ist, also

$$gV(\lambda_i) = aV(\lambda_i) \quad \text{für alle Eigenwert } \lambda_i \text{ von } A$$

Um die Diagonalisierung analytisch durchführen zu können, müssen die Eigenwerte darüber hinaus exakt bekannt sein. Wir gehen im Folgenden davon aus, dass das charakteristische Polynom in Linearfaktoren zerfällt und wir seine Eigenwerte kennen. Wie einfachste Beispiele zeigen, ist nicht jede Matrix (nicht einmal über $\mathbb{C}$!) diagonalisierbar:

Beispiel: Nicht alles ist diagonalisierbar
Sei

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

Dann gilt

$$\chi_A(\lambda) = \lambda^2$$

Wir haben also nur den zweifachen Eigenwert $\lambda = 0$, d.h. $aV(0) = 2$. Zur Bestimmung des Eigenraumes zum Eigenwert 0 betrachten wir das LGS

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

Die Lösung ist unmittelbar abzulesen. Wir finden $y = 0$ und $x \in \mathbb{K}$ ist beliebig. Also erhalten wir als Eigenraum

$$Eig(A, 0) = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{K}^2 : x \in \mathbb{K} \right\} = \left\{ x \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{K}^2 : x \in \mathbb{K} \right\} = lin \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

Dieser Raum ist offensichtlich eindimensional, sodass $gV(\lambda) = 1 \neq 2 = aV(\lambda)$. Daher ist A nicht diagonalisierbar.

Im Falle von nicht diagonalisierbaren Matrizen gibt es aber eine recht gute Alternative, welche (unter der Bedingung, dass das charakteristische Polynom in Linearfaktoren zerfällt) immer funktioniert. Dies ist die Aussage des wohl bekanntesten (und schwersten) Theorems der elementaren linearen Algebra, für das wir zunächst noch eine Definition benötigen:

Definition:
Eine $d \times d$-Matrix J_d der Form

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

heißt ***Jordanmatrix der Länge*** d.

Nun können wir das oben angesprochene Resultat angeben:

Theorem:
Sei $\varphi : V \to V$ eine lineare Abbildung eines $\mathbb{K}$-Vektorraums V in sich selbst, deren charakteristisches Polynom in Linearfaktoren zerfällt. Dann gibt es eine Basis $\mathcal{B}$ von V, bezüglicher derer φ durch folgende blockdiagonale Matrix dargestellt wird:

$$M_\varphi^{\mathcal{B},\mathcal{B}} = \begin{pmatrix} \lambda_1 \cdot \mathbb{1} + N_1 & & \\ & \ddots & \\ & & \lambda_n \cdot \mathbb{1} + N_n \end{pmatrix}$$

Hierbei sind λ_i die Eigenwerte von φ und

$$N_i = \begin{pmatrix} J_{d_i} & & & & & & \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ & & & & & & J_1 \end{pmatrix}$$

Die Basis $\mathcal{B}$ nennt man eine zu φ gehörige ***Jordanbasis*** und $M_\varphi^{\mathcal{B},\mathcal{B}}$ heißt ***Jordannormalform*** von φ.

Diese etwas komplizierte Notation illustrieren wir an einem

Beispiel:

Wir betrachten die Matrix

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

Diese ist in Jordannormalform, denn wir können sie schreiben als

$$A = \begin{pmatrix} 2\cdot \mathbb{1} + N_1 & & \\ & 3\cdot \mathbb{1} + N_2 & \\ & & -5\cdot \mathbb{1} + N_3 \end{pmatrix}$$

mit

$$\begin{aligned} N_1 &= \begin{pmatrix} J_1 & 0 \\ 0 & J_1 \end{pmatrix} \\ N_2 &= \begin{pmatrix} J_2 & 0 \\ 0 & J_1 \end{pmatrix} \\ N_3 &= \begin{pmatrix} J_4 & 0 \\ 0 & J_1 \end{pmatrix} \end{aligned}$$

und

$$\begin{aligned} J_1 &= \begin{pmatrix} 0 \end{pmatrix} \\ J_2 &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ J_4 &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Die Theorie der Jordannormalform ist sehr umfangreich und übersteigt bei weitem den Umfang dieser Wiederholung. Wir begnügen uns damit, ihre wichtigsten Eigenschaften an der obigen Beispielmatrix zu illustrieren. Sie besteht im Wesentlichen aus Jordanmatrizen verschiedener Länge, entlang deren Hauptdiagonalen eine bestimmte Zahl λ hinzugefügt ist. Diese Zahlen werden sich später als Eigenwerte herausstellen. Handelt es sich dabei um eine $d \times d$-Matrix, so nennt man dies einen ***Jordanblock zum Eigenwert*** λ ***der Länge*** d. Die oben auftretenden Jordanblöcke sind

Jordanblock	λ	**Länge**
$\begin{pmatrix} 2 \end{pmatrix}$	2	1
$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$	3	2
$\begin{pmatrix} 3 \end{pmatrix}$	3	1
$\begin{pmatrix} -5 & 1 & 0 & 0 \\ 0 & -5 & 1 & 0 \\ 0 & 0 & -5 & 1 \\ 0 & 0 & 0 & -5 \end{pmatrix}$	-5	4
$\begin{pmatrix} -5 \end{pmatrix}$	-5	1

Hierbei ist eine interessante Beobachtung zu machen. Es besteht nämlich eine Eins-zu-Eins Beziehung von Eigenvektoren und Jordanblöcken. Das sieht man am obigen Beispiel. Es gilt

$$Ae_1 = 2e_1 \quad Ae_2 = 2e_2 \quad Ae_3 = 3e_3 \quad Ae_5 = 3e_5 \quad Ae_6 = -5e_6 \quad Ae_{10} = -5e_{10}$$

Die Diagonalelemente $a_{1,1}$, $a_{2,2}$, $a_{3,3}$, $a_{5,5}$, $a_{6,6}$ und $a_{10,10}$ sind aber gerade die linken oberen Ecken der auftretenden Jordanblöcke. Diese Beobachtung verallgmeinert sich im folgenden

Lemma:
Die Jordannormalform einer Matrix enthält genauso viele Jordanblöcke zu einem gegebenen Eigenwert λ, wie es linear unabhängige Eigenvektoren zu λ gibt. D.h. deren Anzahl ist gerade durch $gV(\lambda)$ gegeben.

Einen weiteren wichtigen Zusammenhang erkennen wir, wenn wir das charakteristische Polynom unserer Beispielmatrix ausrechnen. Es ergibt sich

$$\chi_A = (2-\lambda)^2 \cdot (3-\lambda)^3 \cdot (5-\lambda)^5 \cdot$$

Vergleichen wir die hier auftretenden Vielfachheiten einer bestimmten Nullstelle (also eines Eigenwertes) mit der Anzahl der Stellen in der Matrix A, an der diese Nullstelle auftritt, so liegt folgende Vermutung nahe:

Lemma:
Die Summe der Längen der Jordanblöcke in der Jordannormalform einer Matrix zu einem bestimmten Eigenwert λ ist gleich der algebraischen Vielfachheit $aV(\lambda)$ des entsprechenden Eigenwerts.

Mithilfe dieser beiden Resultate ist es möglich, die Jordannormalform in einfachen Situationen zu berechnen.

Beispiel: Bestimmung der Jordannormalform aus aV und gV

Angenommen, wir hätten eine bestimmte Matrix A gegeben. Wir haben bereits das charakteristische Polynom und die Eigenräume (und damit die geometrischen Vielfachheiten) bestimmt und es ergab sich

λ	$aV(\lambda)$	$gV(\lambda)$
-2	1	1
-1	2	1
2	3	2

Wir möchten nun die Jordannormalform bestimmen. Die erste Zeile der Tabelle sagt uns, dass es einen Jordanblock zum Eigenwert -2 gibt. Dieser muss Länge 1 haben, da die Summe der Längen aller Jordanblöcke zum Eigenwert -2 gerade 1 ist. Zum Eigenwert -1 gibt es auch nur einen Jordanblock. Dieser hat wegen $aV(-2) = 2$ jetzt allerdings Länge 2. Aus der dritten Zeile folgt schließlich, dass zum Eigenwert 2 insgesamt 2 Jordanblöcke existieren. Die Summe der Längen dieser beiden ist 3. Folglich muss es sich um einen Block der Länge 2 und einen Block der Länge 1 handeln. Es ergibt sich damit die Jordannormalform zu

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

Beispiel: Was passiert wenn $gV(\lambda) = aV(\lambda)$ für jedes λ gilt?

Tritt für einen Eigenwert λ die Situation $gV(\lambda) = aV(\lambda)$ ein, so wissen wir, dass die Anzahl der Jordanblöcke zu λ gleich der Summe der Längen dieser Jordanblöcke ist. Dann müssen aber alle Jordanblöcke die Länge 1 haben, was bedeutet, dass sie alle von der Form

$$(\lambda)$$

sind. Gilt dies nun für alle Eigenwerte, so erhält man eine Matrix, auf deren Hauptdiagonale die Eigenwerte stehen und die ansonsten nur Nullen enthält, also eine Diagonalmatrix. Wir haben somit gezeigt, dass Diagonalisierung lediglich ein Spezialfall der Jordannormalform ist!

Zum Abschluss betrachten wir noch einen Satz, welcher die Jordannormalform und das Minimalpolynom auf bemerkenswerte Weise miteinander verknüpft

Satz:
Seien $\lambda_1, ..., \lambda_n$ die Eigenwerte einer Matrix A. Dann gilt für deren Minimalpolynom

$$\mu_A(\lambda) = (\lambda - \lambda_1)^{L_1} \cdot ... \cdot (\lambda - \lambda_n)^{L_n}$$

wobei L_i die Länge des längsten Jordanblocks zum Eigenwert λ_i ist.

In der Tat gilt für unsere Beispielmatrix

$$\mu_A = (\lambda - 2) \cdot (\lambda - 3)^2 \cdot (\lambda - (-5))^4$$

wie man leicht nachrechnen kann.