8 Eigenwerte und Eigenvektoren

In Def. 6.41 haben wir eine quadratische Matrix B **ähnlich** zu einer quadratischen Matrix A genannt, wenn gilt

$$B = P^{-1}AP$$

für eine invertierbare Matrix P. In diesem Kapitel wollen wir untersuchen, unter welchen Bedingungen eine Matrix zu einer Diagonalmatrix ähnlich ist.

Wir werden in diesem Kapitel immer quadratische $n \times n$ Matrizen über einem Körper K betrachten, wenn nichts anderes explizit vereinbart wird.

Definition 8.1: *Sei A eine $n \times n$ Matrix und sei λ ein Skalar ($\in K$). Wenn es einen vom Nullvektor verschiedenen Spaltenvektor x gibt, sodass*

$$Ax \;=\; \lambda x \;,$$

dann ist λ ein **Eigenwert** *der Matrix A und x ein zum Eigenwert λ gehöriger* **Eigenvektor**. *Das* **Spektrum** *von A, geschrieben $\sigma(A)$, ist die Menge der Eigenwerte von A.*
□

Satz 8.2: *Ein Skalar λ ist ein Eigenwert der Matrix A genau dann, wenn*

$$|A - \lambda I_n| \;=\; 0 \;.$$

Korollar: *Ähnliche Matrizen haben dieselben Eigenwerte.*

Schreiben wir $A = (a_{ij})$, so sehen wir, dass offensichtlich

$$|A - \lambda I_n| \;=\; \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

ein Polynom vom Grad n in λ ist.

Definition 8.3: *Das Polynom $|A - \lambda I_n|$ vom Grad n in der Variablen λ heisst das* **charakteristische Polynom** *der Matrix A, geschrieben als $c_A(\lambda)$. Unter der* **charakteristischen Gleichung** *von A verstehen wir die Gleichung*

$$c_A(\lambda) \;=\; |A - \lambda I_n| \;=\; 0 \;.$$
□

Somit könnten wir Satz 8.2 auch wie folgendermassen ausdrücken: *Die Eigenwerte von A sind die Nullstellen des charakteristischen Polynoms von A.* Ist der Grundkörper etwa $K = \mathbb{C}$ (oder ein anderer algebraisch abgeschlossener Körper), so hat jedes univariate Polynom mit Koeffizienten in K auch (mindestens) eine Lösung in K. Insbesondere sieht man daraus, dass jede Matrix einen Eigenwert hat. Weiters lässt sich das charakteristische Polynom über einem algebraisch abgeschlossenen Körper in n Linearfaktoren zerlegen, also

$$c_A(\lambda) = (-1)^n(\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k} \;, \qquad (*)$$

wobei $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von A sind und $r_1 + \cdots r_k = n$.

Definition 8.4: *Die Exponenten $r_1, \ldots, r_k$ in der Gleichung $(*)$ heissen die zu den Eigenwerten $\lambda_1, \ldots, \lambda_k$ von A gehörigen* **algebraischen Vielfachheiten**.

Beispiel 8.5: (a) Wir betrachten die Matrix

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

in $\mathrm{Mat}_{2\times 2}(\mathbb{R})$. Das charakteristische Polynom von A ist

$$c_A(\lambda) = |A - \lambda I_2| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = (\lambda + i)(\lambda - i) \ .$$

Über dem Körper $\mathbb{R}$ hat also A keinen Eigenwert. Aber natürlich können wir die Matrix A auch über dem Grundkörper $\mathbb{C}$ betrachten, und dann hat A die Eigenwerte $\pm i$, jeden mit algebraischer Vielfachheit 1.
(b) Wir betrachten die Matrix

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

in $\mathrm{Mat}_{3\times 3}(\mathbb{R})$. Das charakteristische Polynom von B ist

$$c_B(\lambda) = |B - \lambda I_3| = \begin{vmatrix} -3-\lambda & 1 & -1 \\ -7 & 5-\lambda & -1 \\ -6 & 6 & -2-\lambda \end{vmatrix} = -(\lambda+2)^2(\lambda-4) \ .$$

B hat also die reellen Eigenwerte -2 und 4, und zwar mit den algebraischen Vielfachheiten 2 bzw. 1. □

Lemma 8.6: *Sei $A \in \mathrm{Mat}_{n\times n}(K)$ und sei λ ein Eigenwert von A. Dann ist die Menge (von Spaltenvektoren)*

$$E_{A,\lambda} = \{x \in \mathrm{Mat}_{n\times 1}(K) \mid Ax = \lambda x\}$$

(also die Menge der zu A und λ gehörigen Eigenvektoren zusammen mit dem Nullvektor) ein linearer Unterraum von $\mathrm{Mat}_{n\times 1}(K)$.

Definition 8.7: *Seien A und λ wie in Lemma 8.6. Der lineare Raum $E_{A,\lambda}$ heisst der zu A und λ gehörige* **Eigenraum**. $\dim(E_{A,\lambda})$ *heisst die* **geometrische Vielfachheit** *des Eigenwertes λ.*

Ist A aus dem Kontext klar, so schreiben wir statt $E_{A,\lambda}$ auch einfach E_λ.

Beispiel 8.8: Sei B die 3×3 Matrix aus Beispiel 8.5(b). Der Eigenraum zum Eigenvektor -2, also $E_{B,-2} = E_{-2}$ besteht aus den Lösungen des SLG

$$\begin{pmatrix} -1 & 1 & -1 \\ -7 & 7 & -1 \\ -6 & 6 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \ .$$

Man prüft leicht nach, dass

$$E_{-2} \;=\; \operatorname{span}((1,1,0)^T)\ .$$

Die geometrische Vielfachheit des Eigenwertes -2 ist also 1 und somit kleiner als die algebraische Vielfachheit.

Der Eigenraum zum Eigenvektor 4, also E_4, besteht aus den Lösungen des SLG

$$\begin{pmatrix} -7 & 1 & -1 \\ -7 & 1 & -1 \\ -6 & 6 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\ .$$

Man prüft leicht nach, dass

$$E_4 \;=\; \operatorname{span}((0,1,1)^T)\ .$$

Die geometrische Vielfachheit des Eigenwertes 4 ist also 1 und somit gleich wie die algebraische Vielfachheit.

Wir führen einige dieser Berechnungen in Maple 9.5 aus:

> **with(Linear Algebra):**

> **A := Matrix([[-3,1,-1],[-7,5,-1],[-6,6,-2]]);**

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

> **Eigenvalues(A);**

$$\begin{bmatrix} 4 \\ -2 \\ -2 \end{bmatrix}$$

> **CharacteristicPolynomial(A,x);**

$$x^3 - 12x - 16$$

> **factor(%);**

$$(x-4)(x+2)^2$$

> **Eigenvectors(A);**

$$\begin{bmatrix} 4 \\ -2 \\ -2 \end{bmatrix},\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Dabei ist die i-te Spalte des zweiten Teils des Ergebnisses ein Element des Eigenraums des i-ten Eigenwertes im ersten Teil des Ergebnisses, und diese Vektoren zusammen spannen den jeweiligen Eigenraum auf. □

Satz 8.9: *Sind λ_1, λ_2 zwei verschiedene Eigenwerte von A, dann gilt $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$.*

Wir wissen aus Kapitel 6, dass sich jede lineare Abbildung zwischen endlichdimensionalen Vektorräumen darstellen lässt als Multiplikation mit einer Matrix (Darstellungsmatrix). Andererseits ist natürlich jede Abbildung der Form

$$x \longrightarrow Ax$$

linear. Dieser Zusammenhang führt auf natürliche Weise zum Begriff des Eigenwertes einer linearen Abbildung.

Definition 8.10: *Sei V ein linearer Teilraum von W. Sei $f : V \to W$ eine lineare Abbildung und sei λ ein Skalar im Grundkörper K. Wenn es einen Vektor $x \in V \setminus \{0\}$ gibt mit*

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

dann ist λ ein **Eigenwert** *von f und x ein zum Eigenwert λ gehöriger* **Eigenvektor**. *Das* **Spektrum** *von f, geschrieben $\sigma(f)$, ist die Menge der Eigenwerte von f.* □

Beispiel 8.11: (a) Sei f die lineare Abbildung

$$f: \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (y+z, x+z, x+y) \end{array} .$$

Bezüglich der natürlichen geordneten Basis von $\mathbb{R}^3$ hat f die Darstellungsmatrix

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

Die Matrix A hat das charakteristische Polynom

$$c_A(\lambda) = -(\lambda+1)^2(\lambda-2) \ .$$

Somit hat A und daher auch f die Eigenwerte -1 und 2, und zwar mit algebraischer Vielfachheit 2 bzw. 1.
(b) Sei $\mathrm{Diff}(\mathbb{R}, \mathbb{R})$ der Vektorraum aller differentierbaren reellen Funktionen (über $\mathbb{R}$). Die Differentiationsabbildung

$$D : \mathrm{Diff}(\mathbb{R}, \mathbb{R}) \longrightarrow \mathrm{Fun}(\mathbb{R}, \mathbb{R})$$

ist linear. Ein Eigenvektor von D ist eine von der Nullfunktion verschiedene differentierbare Funktion $f(x)$, sodass für ein $\lambda \in \mathbb{R}$ gilt: $D(f) = \lambda f$. Aus der Analysis und der Theorie der Differentialgleichung ergibt sich, dass die Eigenvektoren von D die Funktionen der Form

$$f(x) = ke^{\lambda x}$$

sind. Die lineare Abbildung Diff hat also unendlich viele Eigenwerte $\lambda \in \mathbb{R} \setminus \{0\}$. □

Satz 8.12: *Sei $f : V \to V$ eine lineare Abbildung, seien $\lambda_1, \ldots, \lambda_k$ verschiedene Eigenwerte von f und sei für jedes $i \in \{1, \ldots, k\}$ der Vektor x_i ein Eigenvektor zum Eigenwert λ_i. Dann sind die Vektoren $x_1, \ldots, x_k$ linear unabhängig.*

Definition 8.13: *Die lineare Abbildung $f : V \to V$ auf einem endlichdimensionalen Vektorraum V heisst* **diagonalisierbar**, *wenn es eine geordnete Basis $(v_1, \ldots, v_n)$ von V gibt, bezüglich derer die Darstellungsmatrix von f eine Diagonalmatrix ist.*
Ebenso nennen wir eine quadratische Matrix A **diagonalisierbar**, *wenn sie ähnlich ist zu einer Diagonalmatrix.*

Satz 8.14: (a) *Eine lineare Abbildung $f : V \to V$ auf einem endlichdimensionalen Vektorraum V ist diagonalisierbar g.d.w. V eine Basis besitzt, welche aus Eigenvektoren von f besteht.*

(b) *Eine quadratische $n \times n$ Matrix A ist diagonalisierbar g.d.w. sie n linear unabhängige Eigenvektoren besitzt.*

Beweis: (a) Laut Definition ist f diagonalisierbar g.d.w. es eine geordnete Basis $(v_1, \ldots, v_n)$ gibt, sodass

$$\begin{array}{rcllll} f(v_1) & = & \lambda_1 v_1 & & & \\ f(v_2) & = & & \lambda_2 v_2 & & \\ & \vdots & & & \ddots & \\ f(v_n) & = & & & & \lambda_n v_n \ . \end{array}$$

Das ist offensichtlich genau dann der Fall, wenn die λ_i Eigenwerte mit zugehörigen Eigenvektoren v_i sind.
(b) Wir interpretieren die Matrix A als Darstellungsmatrix der linearen Abbildung

$$\begin{array}{rccc} f_A: & K^n & \longrightarrow & K^n \\ & x & \mapsto & Ax \ . \end{array}$$

Die Matrix A ist diagonalisierbar g.d.w.
es eine reguläre Matrix P und eine Diagonalmatrix D gibt, sodass $A = P^{-1}DP$. Diese Matrix P ist natürlich eine Basistransformationsmatrix. Das ist also der Fall g.d.w.
die lineare Abbildung f_A diagonalisierbar ist, also (wegen (a)) g.d.w.
es eine geordnete Basis $(v_1, \ldots, v_n)$ von K^n und $\lambda_1, \ldots, \lambda_n$ gibt, sodass $Av_i = f_A(v_i) = \lambda_i v_i$ für alle i. Das aber heisst nichts anderes als dass A n linear unabhängige Eigenvektoren besitzt. □

Bemerkung: Laut Definition ist A diagonalisierbar g.d.w. es eine invertierbare Matrix P und eine Diagonalmatrix D gibt, sodass $A = P^{-1}DP$. A und D haben dieselben Eigenwerte. Sei

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} .$$

Die kanonische Basis von K^n ist eine Basis aus Eigenvektoren von D: $De_i = \lambda_i e_i$. Die Vektoren $P^{-1}e_1, \ldots, P^{-1}e_n$ sind Eigenvektoren von A, denn

$$A(P^{-1}e_i) = P^{-1}DP(P^{-1}e_i) = P^{-1}De_i = P^{-1}\lambda_i e_i = \lambda_i(P^{-1}e_i) \ .$$

Ausserdem sind die Vektoren $P^{-1}e_1, \ldots, P^{-1}e_n$ linear unabhängig, denn

$$0 = \sum \mu_i(P^{-1}e_i) = P^{-1}(\sum \mu_i e_i) \iff \sum \mu_i e_i = 0 \iff m_i = 0 \ \forall i.$$

Somit gibt es n linear unabhängige Eigenvektoren von A.
Umgekehrt seien $x_1, \ldots, x_n$ linear unabhängig und $\lambda_1, \ldots, \lambda_n$ so, dass $Ax_i = \lambda_i x_i$ für alle i gilt. Sei P die (Basistransformations-) Matrix, für welche gilt

$$e_i = Px_i \quad \forall i \ .$$

Dann gilt für $D := PAP^{-1}$:

$$De_i = PAP^{-1}e_i = \lambda_i e_i \quad \forall i \ .$$

Also ist A ähnlich zur Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

und somit diagonalisierbar. □

Beispiel: Sei

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

die Darstellungsmatrix einer linearen Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$. A ist diagonalisierbar, denn für

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

gilt $A = P^{-1}DP$.
Die Diagonalmatrix D hat offensichtlich die Eigenwerte 1 und 2, und das sind wegen der Ähnlichkeit auch die Eigenwerte von A.
$(1,0)^T$ ist ein Eigenvektor von D zum Eigenwert 1, und

$$P^{-1} \cdot (1,0)^T = (0,1)^T$$

ist ein Eigenvektor von A zum Eigenwert 1.
$(0,1)^T$ ist ein Eigenvektor von D zum Eigenwert 2, und

$$P^{-1} \cdot (0,1)^T = (1,-1)^T$$

ist ein Eigenvektor von A zum Eigenwert 2. □

Nun wollen wir zeigen, dass eine lineare Abbildung $f : V \to V$ bzw. die zugehörige Matrix diagonalisierbar ist genau dann, wenn für jeden Eigenwert die algebraische und die geometrische Vielfachheit übereinstimmen. Dazu brauchen wir zunächst einige Vorbereitung.

Satz 8.15: *Sei $f : V \to V$ eine lineare Abbildung auf dem n-dimensionalen Vektorraum V. Seien $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von f, und $d_1, \ldots, d_k$ die zugehörigen geometrischen Vielfachheiten. Dann gilt*

$$d_1 + \ldots + d_k \leq n. \quad (*)$$

In $(*)$ gilt die Gleichheit g.d.w. f diagonalisierbar ist.

Satz 8.16: *Für jeden Eigenwert λ einer linearen Abbildung $f : V \to V$ ist die algebraische Vielfachheit mindestens so gross wie die geometrische Viellfachheit.*

Satz 8.17: *Sei $f : V \to V$ eine lineare Abbildung. Die folgenden Aussagen sind äquivalent:*

(i) f ist diagonalisierbar;

(ii) für jeden Eigenwert λ von f ist die geometrische Vielfachheit gleich der algebraischen Vielfachheit.

Wir haben also nun ein Kriterium für die Diagonalisierbarkeit einer linearen Abbildung bzw. einer Matrix A (welche ja einer linearen Abbildung entspricht). Wie aber bestimmt man eine invertierbare Matrix P, sodass $P^{-1}AP$ eine Diagonalmatrix ist?

Satz 8.18: *Sei A eine diagonalisierbare $n \times n$ Matrix. Seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A (dabei komme in dieser Liste jeder Eigenwert so oft vor, wie seine algebraische Vielfachheit beträgt), und seien $p_1, \dots, p_n \in K^n$ so, dass für jeden Index i der Vektor p_i ein Eigenvektor zum Eigenwert λ_i ist, und dass die Vektoren $p_1, \dots, p_n$ überdies eine Basis von K^n bilden. Sei P diejenige $n \times n$ Matrix, deren i-te Spalte p_i ist. Dann gilt*

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 & \\ & \lambda_2 & & \\ & & \ddots & \\ & 0 & & \lambda_n \end{pmatrix} .$$

Beispiel 8.19: Wir betrachten die Matrix

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

Das charakteristische Polynom von A ist

$$c_A(\lambda) = -(\lambda - 4)(\lambda + 2)^2.$$

Die Matrix A hat also den Eigenwert 4 mit algebraischer Vielfachheit 1 und den Eigenwert -2 mit algebraischer Vielfachheit 2. So wie im Satz 8.18 seien also

$$\lambda_1 = 4, \quad \lambda_2 = -2, \quad \lambda_3 = -2 .$$

Als Eigenvektoren wählen wir (so wie im Satz 8.18 angegeben)

$$p_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} .$$

Dann gilt für die Matrix $P = (p_1, p_2, p_3)$

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

Wir erreichen also die Diagonalisierung mittels einer Matrix, deren Spalten eine Basis von $\mathbb{C}^3$ sind bestehend aus Eigenvektoren von A. □

Einige Anwendungsbeispiel:

Beispiel 8.20 (Äquilibrium in einem System): Wir betrachten die Entwicklung der Arbeitslosigkeit in einer Volkswirtschaft. Dazu nehmen wir an, dass 75% der Personen, welche am Beginn eines Jahres arbeitslos sind, im Laufe dieses Jahres Arbeit finden, und dass 5% der Personen, welche am Beginn eines Jahres Arbeit haben, im Laufe dieses Jahres arbeitslos werden. Wir betrachten die Entwicklung der Arbeitslosenrate über die Jahre hinweg. Am Beginn der Studie, also am Ende des 0-ten Jahres, sei L_0 der Anteil der Bevölkerung, welche keine Arbeit hat, und $M_0 = 1 - L_0$ der Anteil der Bevölkerung mit Arbeit. In analoger Weise für die Folgejahre. Also

L_i .... Anteil der Arbeitslosen am Ende des i-ten Jahres
M_i .... Anteil der Arbeitenden am Ende des i-ten Jahres

Der Übergang von einem Jahr zum nächsten drückt sich also aus als

$$\begin{pmatrix} L_{i+1} \\ M_{i+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{4} & \frac{1}{20} \\ \frac{3}{4} & \frac{19}{20} \end{pmatrix}}_{U} \begin{pmatrix} L_i \\ M_i \end{pmatrix} .$$

Ausgehend von den Startwerten gilt dann für ein Jahr k:

$$\begin{pmatrix} L_k \\ M_k \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{20} \\ \frac{3}{4} & \frac{19}{20} \end{pmatrix}^k \begin{pmatrix} L_0 \\ M_0 \end{pmatrix} .$$

Wegen

$$\lim_{k\to\infty} \begin{pmatrix} \frac{1}{4} & \frac{1}{20} \\ \frac{3}{4} & \frac{19}{20} \end{pmatrix}^k = \lim_{k\to\infty} \frac{1}{16} \begin{pmatrix} 1+\frac{15}{5^k} & 1-\frac{1}{5^k} \\ 15-\frac{15}{5^k} & 15+\frac{1}{5^k} \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 1 & 1 \\ 15 & 15 \end{pmatrix}$$

gilt also

$$\begin{pmatrix} L_k \\ M_k \end{pmatrix} \longrightarrow_{k\to\infty} \frac{1}{16} \begin{pmatrix} 1 & 1 \\ 15 & 15 \end{pmatrix} \begin{pmatrix} L_0 \\ M_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{16} \\ \frac{15}{16} \end{pmatrix} .$$

Dieses System konvergiert also unabhängig von den Ausgangswerten L_0, M_0 zu einem Äquilibrium.

Wie aber sind haben wir oben die k-te Potenz der Übergangsmatrix berechnen können? Diese Überlegung wollen wir nun nachholen. Hat man einmal die Vermutung für die Formel von U^k, so lässt sich die Korrektheit mittels Induktion leicht nachweisen. Wie aber kommt man zu dieser Vermutung?

Die Eigenwerte von U sind die Nullstellen des charakteristischen Polynoms

$$c_U(\lambda) = (\frac{1}{4} - \lambda)(\frac{19}{20} - \lambda) - \frac{3}{80} = (5\lambda - 1)(\lambda - 1) = 0 .$$

Die Eigenwerte sind also $\lambda_1 = \frac{1}{5}$ und $\lambda_2 = 1$. Zugehörige Eigenvektoren sind etwa $v_1 = (1, -1)^T$ und $v_2 = (1, 15)^T$. Wegen der Bemerkung nach Satz 8.14 bewirkt also die Matrix $P = (v_1, v_2)$ eine Ähnlichkeitstranformation in eine Diagonalmatrix, nämlich

$$\underbrace{\frac{1}{16} \begin{pmatrix} 15 & -1 \\ 1 & 1 \end{pmatrix}}_{P^{-1}} \cdot \underbrace{\frac{1}{20} \begin{pmatrix} 5 & 1 \\ 15 & 19 \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 15 \end{pmatrix}}_{P} = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix} .$$

Somit können wir für beliebiges $k \in \mathbb{N}$ die k-te Potenz der Übergangsmatrix U einfach berechnen:

$$\begin{aligned} U^k &= P \cdot \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix}^k \cdot P^{-1} \\ &= \tfrac{1}{16} \begin{pmatrix} 1 & 1 \\ -1 & 15 \end{pmatrix} \begin{pmatrix} \frac{1}{5^k} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 15 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \tfrac{1}{16} \begin{pmatrix} 1 + \frac{15}{5^k} & 1 - \frac{1}{5^k} \\ 15 - \frac{15}{5^k} & 15 + \frac{1}{5^k} \end{pmatrix} . \end{aligned}$$

Beispiel 8.21 (Fibonacci-Zahlen): Leonardo von Pisa (1175–1250), besser bekannt als Fibonacci (Sohn des Bonaccio, filius Bonacii) betrachtete die rekursive Folge

$$a_0 = 0,\ a_1 = 1,\ \text{und } a_{n+2} = a_{n+1} + a_n \text{ für } n \geq 0.$$

Mittels der Hilfsfolge $(b_n)_{n\geq 1}$ können wir das Bildungsgesetz der Fibonacci-Folge auch schreiben als ein System von Differenzengleichungen der Form

$$\begin{pmatrix} a_{n+2} \\ b_{n+2} \end{pmatrix} = \begin{pmatrix} a_{n+1} + b_{n+1} \\ a_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}}_{A} \begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} .$$

Wir bestimmen die Eigenwerte von A:

$$c_A(\lambda) = \lambda^2 - \lambda - 1,$$

also

$$\lambda_1 = \frac{1}{2}(1 + \sqrt{5}), \quad \lambda_2 = \frac{1}{2}(1 - \sqrt{5}).$$

Zugehörige Eivenvektoren sind

$$\text{zu } \lambda_1 : \begin{pmatrix} 1 \\ -\lambda_2 \end{pmatrix}, \quad \text{zu } \lambda_2 : \begin{pmatrix} 1 \\ -\lambda_1 \end{pmatrix}.$$

Mittels der Matrix P bestehend aus diesen Eigenvektoren, also

$$P = \begin{pmatrix} 1 & 1 \\ -\lambda_2 & -\lambda_1 \end{pmatrix}, \quad P^{-1} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} -\lambda_1 & -1 \\ \lambda_2 & 1 \end{pmatrix},$$

können wir also A diagonalisieren:

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} .$$

Die Potenzen A^n der Matrix A sind also leicht berechenbar. Wir benutzen dazu noch den Zusammenhang $\lambda_1\lambda_2 = -1$. Für beliebiges $n \in \mathbb{N}$ gilt also:

$$\begin{aligned} A^n &= \tfrac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} 1 & 1 \\ -\lambda_2 & -\lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} -\lambda_1 & -1 \\ \lambda_2 & 1 \end{pmatrix} \\ &= \tfrac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2^{n+1} - \lambda_1^{n+1} & \lambda_2^n - \lambda_1^n \\ \lambda_2^n - \lambda_1^n & \lambda_2^{n-1} - \lambda_1^{n-1} \end{pmatrix} . \end{aligned}$$

Aus den Anfangswerten $b_1 = a_0 = 0$ und $a_1 = 1$ ergibt sich somit

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = A^n \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2^{n+1} - \lambda_1^{n+1} \\ \lambda_2^n - \lambda_1^n \end{pmatrix} ,$$

und schliesslich

$$a_n = \frac{1}{\lambda_2 - \lambda_1}(\lambda_2^n - \lambda_1^n) = \frac{1}{2^n\sqrt{5}}\Big((1+\sqrt{5})^n - (1-\sqrt{5})^n\Big).$$

Somit haben wir eine explizite Formel für die Fibonacci-Folge ermittelt. □

Beispiel 8.22 (Systeme gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten): Seien $y_1(t), \ldots, y_n(t)$ differenzierbare Funktionen von $\mathbb{R}$ nach $\mathbb{R}$, bzw. von $\mathbb{C}$ nach $\mathbb{C}^n$. Wir schreiben diese Funktionen als Vektor

$$y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} ,$$

$y(t)$ ist also eine Funktion von $\mathbb{R}$ nach $\mathbb{R}^n$, bzw. von $\mathbb{C}$ nach $\mathbb{C}^n$. Die Ableitung $y'(t) = \frac{dy}{dt}(t)$ von y nach t ist komponentenweise definiert als

$$y'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_n'(t) \end{pmatrix} .$$

Für eine gegebene Matrix A in $\mathrm{Mat}_{n\times n}(\mathbb{R})$ bzw. in $\mathrm{Mat}_{n\times n}(\mathbb{C})$ ist

$$y'(t) = Ay(t) \quad (*)$$

ein **homogenes System linearer gewöhnlicher Differentialgleichungen (hSLGDG) mit konstanten Koeffizienten**. Eine Lösung dieses Systems ist eine differenzierbare Funktion, welche $(*)$ erfüllt. Die Menge aller Lösungen eines solchen hSLGDG ist ein linearer Teilraum des Vektorraums aller differenzierbaren Funktionen von $\mathbb{R}$ nach $\mathbb{R}^n$, bzw. von $\mathbb{C}$ nach $\mathbb{C}^n$. Man kann zeigen, dass dieser Lösungsraum die Dimension n hat. Eine Basis

$$(\varphi_1(t), \ldots, \varphi_n(t))$$

des Lösungsraumes nennt man auch ein **Fundamentalsystem** des hSLGDG $(*)$. Es gilt die folgende Beziehung:

$$\begin{gathered} n \text{ Lösungen } (\varphi_1, \ldots, \varphi_n) \text{ von } (*) \text{ bilden ein Fundamentalsystem} \\ \Longleftrightarrow \\ \det(\varphi_1(t_0), \ldots, \varphi_n(t_0)) \neq 0 \text{ für ein } t_0. \end{gathered}$$

Hat nun die Matrix A den Eigenwert λ mit zugehörigem Eigenvektor x, dann ist

$$\varphi(t) = xe^{\lambda t}$$

eine Lösung des hSLGDG $(*)$, denn

$$\varphi'(t) = \lambda x e^{\lambda t} = Axe^{\lambda t} = A\varphi(t).$$

Tatsächlich gilt folgender Zusammenhang (der typischerweise in Vorlesungen zu Differentialgleichungen bewiesen wird):
Satz: *Besitzt die Matrix A eine Basis von Eigenvektoren $x_1, \ldots, x_n$ zu den Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, so bilden die Funktionen $\varphi_k(t) = x_k e^{\lambda_k t}$ für $1 \leq k \leq n$ ein Fundamentalsystem des hSLGDG $(*)$.*

Konkret betrachten wir das Differentialgleichungssystem

$$y'(t) = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{A} y(t)\ .$$

Die Matrix A hat das charakteristische Polynom

$$c_A(\lambda) = \lambda^2(\lambda - 3).$$

Der Eigenwert $\lambda_1 = \lambda_2 = 0$ hat also die algebraische Vielfachheit 2, der Eigenwert $\lambda_3 = 3$ hat die algebraische Vielfachheit 1. Eigenvektoren x_i zu den Eigenwerten λ_i sind

$$x_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Diese Eigenvektoren (x_1, x_2, x_3) bilden offensichtlich eine Basis von $\mathbb{R}^3$. Wegen des obigen Satzen erhalten wir daraus das folgende Fundamentalsystem für den Lösungsraum:

$$\varphi_1(t) = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{0t} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \varphi_2(t) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} e^{0t} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \varphi_3(t) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{3t}\ .$$

Die allgemeine Lösung des Differentialgleichungssystems lautet also

$$y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = a \cdot \varphi_1(t) + b \cdot \varphi_2(t) + c \cdot \varphi_3(t) = \begin{pmatrix} a + b + ce^{3t} \\ -b + ce^{3t} \\ -a + ce^{3t} \end{pmatrix},$$

wobei a, b, c beliebige Konstante sind.

Solche homogenen Systeme linearer gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten lassen sich also vollständig lösen mittels Methoden der linearen Algebra.

Wir betrachten noch ein zweites konkretes Differentialgleichungssystem, nämlich

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}\ .$$

Die Koeffizientenmatrix dieses Systems hat die Eigenwerte

$$\lambda_1 = i, \quad \lambda_2 = -i$$

mit den zugehörigen Eigenvektoren

$$x_1 = \begin{pmatrix} 1 \\ \frac{1}{2} - \frac{1}{2}i \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ \frac{1}{2} + \frac{1}{2}i \end{pmatrix}.$$

Als Fundamentalsystem für den Lösungsraum können wir also etwa wählen

$$\varphi_1(t) = \begin{pmatrix} 1 \\ \frac{1}{2} - \frac{1}{2}i \end{pmatrix} e^{it}, \qquad \varphi_2(t) = \begin{pmatrix} 1 \\ \frac{1}{2} + \frac{1}{2}i \end{pmatrix} e^{-it} \ .$$

Natürlich ist dann auch

$$(\psi_1(t), \psi_2(t)) = (-\frac{i}{2}\varphi_1(t) + \frac{i}{2}\varphi_2(t),\ \frac{1}{2}\varphi_1(t) + \frac{1}{2}\varphi_2(t))$$

ein Fundamentalsystem, welches wir unter Verwendung der Zusammenhänge

$$\begin{array}{rcl} e^{it} & = & \cos t + i \sin t \\ e^{-it} & = & \cos t - i \sin t \end{array} \quad \text{bzw.} \quad \begin{array}{rcl} \cos t & = & \frac{1}{2}\Big(e^{it} + e^{-it}\Big) \\ \sin t & = & \frac{-i}{2}\Big(e^{it} - e^{-it}\Big) \end{array}$$

auch schreiben können als

$$\psi_1(t) = \begin{pmatrix} \sin t \\ -\frac{1}{2}\cos t + \frac{1}{2}\sin t \end{pmatrix}, \qquad \psi_2(t) = \begin{pmatrix} \cos t \\ \frac{1}{2}\cos t + \frac{1}{2}\sin t \end{pmatrix} \ .$$

Somit hat die allgemeine Lösung des Differentialgleichungssystems die Gestalt

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} a \sin t + b \cos t \\ \frac{a}{2}(\sin t - \cos t) + \frac{b}{2}(\sin t + \cos t) \end{pmatrix} \ ,$$

wobei a, b beliebige Konstante sind. □

Das Minimalpolynom einer Matrix

Im Vektorraum $\mathrm{Mat}_{n\times n}(K)$ sei für alle $1 \leq i, j \leq n$ die Matrix E_{ij} so, dass sie überall 0 enthält, ausser in der i-ten Zeile und j-ten Spalte, wo der Eintrag 1 ist. Die Menge

$$\{E_{ij} \mid 1 \leq i, j \leq n\}$$

ist offensichtlich linear unabhängig und spannt ausserdem den ganzen Raum $\mathrm{Mat}_{n\times n}(K)$ auf. Der Vektorraum der $n \times n$ Matrizen hat also die Dimension n^2. Das bedeutet auch, dass jede Menge von $n^2 + 1$ Matrizen oder mehr linear abhängig ist. Insbesondere ist für eine beliebige $n \times n$ Matrix A die Menge der Potenzen

$$\{A^0 = I_n, A, A^2, \dots, A^n\} \;=\; \{A^k \mid 0 \leq k \leq n\}$$

linear abhängig, d.h. es gibt ein Polynom vom Grad n^2

$$p(x) = a_0 + a_1 x + \ldots + a_{n^2} x^{n^2} \;\in\; K[x],$$

sodass $p(A) = 0$. Analog gilt das natürlich auch für jede lineare Abbildung $f \in \mathrm{Hom}_K(V, V)$, wobei V ein Vektorraum der Dimension n ist; denn nach Satz 6.34 ist $\mathrm{Hom}_K(V, V) \cong \mathrm{Mat}_{n\times n}(K)$.

Diese Schranke für die lineare Abhängigkeit der Potenzen einer quadratischen Matrix kann aber deutlich verbessert werden. Es gibt nämlich ein Polynom $p(x)$ vom Grad höchstens n, sodass $p(A) = 0$. Dieser Satz geht zurück auf Cayley (Arthur Cayley, 1821–1895) und Hamilton (William R. Hamilton, 1805–1865).

Satz 8.23 (Cayley-Hamilton): *Jede quadratische Matrix A ist Nullstelle seines charakteristischen Polynoms, also $c_A(A) = 0$.*

Ein vom Nullpolynom verschiedenes Polynom $p(x)$ heisst **monisch**, wenn sein führender Koeffizient, also der Koeffizient des höchsten Terms x^k in p, 1 ist.

Satz 8.24: *Sei k der kleinste Grad eines Polynoms, welches A als Nullstelle hat. Dann gibt es genau ein monisches Polynom $m(x)$ vom Grad k, also von der Form $m(x) = x^k + m_{k-1}x^{k-1} + \cdots + m_0$.*

Definition 8.25: *Seit $A \in \mathrm{Mat}_{n\times n}(K)$. Das eindeutig bestimmte monische Polynom kleinsten Grades $m_A(x)$ mit $m_A(A) = 0$ heisst das* **Minimalpolynom** *der Matrix A.*

Satz 8.26: *Ist $p(x)$ ein Polynom mit $p(A) = 0$, dann teilt das Minimalpolynom $m_A(x)$ das Polynom $p(x)$.*

Korollar: *Das Minimalpolynom $m_A(x)$ ist ein Teiles des charakteristischen Polynoms $c_A(x)$.*

Satz 8.27: *Das Minimalpolynom $m_A(x)$ und das charakteristische Polynom $c_A(x)$ haben dieselben Nullstellen in K.*

Beispiel 8.28: (a) Das charakteristische Polynom der Matrix

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

ist $c_A(x) = (x-2)^3$. Sowohl $A-2I_3$ als auch $(A-2I_3)^2$ sind verschieden von der Nullmatrix. Somit ist das charakteristische Polynom gleich dem Minimalpolynom, $m_A(x) = c_A(x)$.
(b) Das charakteristische Polynom der Matrix

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

ist $c_B(x) = (x-1)(x-2)^2$. Wegen Satz 8.27 ist das Minimalpolynom von B entweder

$$(x-1)(x-2) \quad \text{oder} \quad (x-1)(x-2)^2 \ .$$

Es gilt $(A-I)(A-2I) = 0$, also ist $m_B(x) = (x-1)(x-2)$. □

Satz 8.29: *Eine quadratische Matrix ist invertierbar genau dann, wenn der konstante Term in ihrem charakteristischen Polynom verschieden von* 0 *ist.*

Beispiel 8.30: Die Matrix

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

hat das charakteristische Polynom $c_A(x) = (x-1)^3$. Aus dem Satz von Cayley-Hamilton sehen wir also

$$0 = (A-I_3)^3 = A^3 - 3A^2 + 3A - I_3 \ ,$$

und weiter

$$A(A^2 - 3A + 3I_3) = I_3 \ .$$

Also gilt

$$A^{-1} = A^2 - 3A + 3I_3 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

für die inverse Matrix. □