2.2 Kern und Bild; Basiswechsel

Jede lineare Abbildung definiert charakteristische Unterräume, sowohl im Ausgangsraum als auch im Bildraum.

2.2.1 Satz *Sei $L: V \to W$ eine lineare Abbildung. Dann gilt:*

1. *Das* $\mathrm{Bild}(L) := \{L(v) \mid v \in V\}$ *ist ein linearer Unterraum von W.*

2. *Der* $\mathrm{Kern}(L) := \{v \in V \mid L(v) = 0\}$*, also das Urbild der* 0 *in W, ist ein linearer Unterraum von V.*

3. *L ist genau dann injektiv, wenn* $\mathrm{Kern}(L) = \{0\}$.

4. *Liegt w im Bild von L, dann ist sein Urbild $L^{-1}(w) = \{v \in V \mid L(v) = w\}$ in V ein affiner Unterraum, genauer eine parallel verschobene Kopie von* $\mathrm{Kern}(L)$.

Beweis. Zu 1. Das Bild von L ist nichtleer, denn wegen $L(0) = 0$, enthält es zumindest den Nullvektor von W. Nehmen wir jetzt an $w_1, w_2 \in \mathrm{Bild}(L)$. Dann gibt es Vektoren $v_1, v_2 \in V$ mit $L(v_1) = w_1$ und $L(v_2) = w_2$. Aus der Linearität von L folgt $L(v_1 + v_2) = L(v_1) + L(v_2) = w_1 + w_2$. Also ist auch $w_1 + w_2$ im Bild von L enthalten. Schliesslich gilt für alle $\lambda \in \mathbb{R}$: $L(\lambda v_1) = \lambda L(v_1) = \lambda w_1 \in \mathrm{Bild}(L)$.

zu 2. Diesen Beweisteil lassen wir als Übung.

zu 3. Ist L injektiv, so ist $L(v) = 0$ nur für $v = 0$ möglich. Das heisst $\mathrm{Kern}(L) = \{0\}$. Sei jetzt umgekehrt $\mathrm{Kern}(L) = \{0\}$ und nehmen wir an, es sei $L(v_1) = L(v_2)$ für $v_1, v_2 \in V$. Dann folgt $L(v_1 - v_2) = L(v_1) - L(v_2) = 0$ und daher $v_1 - v_2 \in \mathrm{Kern}(L)$. Also muss $v_1 = v_2$ sein. q.e.d.

2.2.2 Beispiel Sei $L: \mathbb{R}^3 \to \mathbb{R}^3$ die lineare Abbildung, gegeben durch $L\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} (x+y)/2 \\ (x+y)/2 \\ z \end{pmatrix}$. Es handelt sich hier um eine Projektion auf die Ebene E in $\mathbb{R}^3$, die von der z-Achse und der Winkelhalbierenden in der x–y–Ebene aufgespannt wird. Hier ist das Bild die Ebene E, und der Kern ist die Gerade g, erzeugt vom Vektor $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. Der Punkt $w = \begin{pmatrix} a \\ a \\ 1 \end{pmatrix}$ hat das Urbild $L^{-1}(w) = \left\{ \begin{pmatrix} a+t \\ a-t \\ 1 \end{pmatrix} \mid x \in \mathbb{R} \right\}$. Es handelt sich also um eine zu g parallele Gerade durch w.

2.2.3 Bemerkung *Sei $L_A: \mathbb{R}^n \to \mathbb{R}^m$ diejenige lineare Abbildung, die durch die Multiplikation mit einer festen $m \times n$-Matrix A gegeben ist. Dann ist*

$$\mathrm{Kern}(L_A) = \{v \in \mathbb{R}^n \mid Av = 0\}\,,$$

also nichts anderes als Lösungsmenge $\mathbb{L}$ des durch die Matrix A definierten linearen Gleichungssystems. Das Urbild eines Vektors $b \in \mathbb{R}^m$ stimmt überein mit der Lösungsmenge des zugehörigen inhomogenen linearen Gleichungssystems

$$L^{-1}(b) = \{v \in \mathbb{R}^n \mid Av = b\}\,.$$

Bekanntlich ist diese Menge ein affiner Unterraum, genauer eine parallel verschobene Kopie von $\mathbb{L}$. Bezeichnen wir die Spaltenvektoren von A mit $w_1, \ldots, w_n$, dann ist

$$\mathrm{Bild}(L_A) = \mathrm{lin}(w_1, \ldots, w_n)\,,$$

d.h., das Bild von L_A ist derjenige lineare Unterraum von $\mathbb{R}^m$, der von den Spaltenvektoren von A erzeugt wird.

2.2.4 Beispiele

- Sei $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 3 & -2 \end{pmatrix}$. Das Bild der Abbildung $L_A\colon \mathbb{R}^2 \to \mathbb{R}^3$, definiert durch Multiplikation mit der Matrix A, ist diejenige Ebene in $\mathbb{R}^3$, die von den beiden Spalten von A erzeugt wird. Der Kern besteht hier nur aus dem Nullvektor.
- Sei $A = \begin{pmatrix} 2 & -6 & 1 \\ 1 & -3 & -1 \end{pmatrix}$. Hier definiert die Multiplikation mit A eine Abbildung $L_A\colon \mathbb{R}^3 \to \mathbb{R}^2$. Das Bild ist ganz $\mathbb{R}^2$ und der Kern besteht aus allen Vektoren in $\mathbb{R}^3$ der Form $\begin{pmatrix} 3t \\ t \\ 0 \end{pmatrix}$ $(t \in \mathbb{R})$.

Betrachten wir schliesslich noch ein Beispiel für eine lineare Abbildung zwischen Funktionenräumen.

2.2.5 Beispiel Sei $\lambda > 0$ vorgegeben. Durch die Vorschrift $L(y) = y'' - \lambda^2 y$ wird eine lineare Abbildung von dem Raum der beliebig oft stetig differenzierbaren Funktionen $V = C^\infty(\mathbb{R})$ in sich definiert. Hier ist

$$\mathrm{Kern}(L) = \{c_1 e^{\lambda t} + c_2 e^{-\lambda t} \mid c_1, c_2 \in \mathbb{R}\} \quad \text{und} \quad \mathrm{Bild}(L) = C^\infty(\mathbb{R})\,.$$

Das Urbild einer Funktion $b \in V$ unter L ist nichts anderes als die Lösungsmenge der inhomogenen Differentialgleichung

$$y'' - \lambda^2 y = b\,.$$

Wir hatten bereits gesehen, dass es sich dabei um eine parallel verschobene Kopie des Lösungsraums der entsprechenden homogenen Differentialgleichung, also des Kerns von L, handelt.

2.2.6 Satz *(Dimensionsformel) Sei $L\colon V \to W$ linear und* $\dim V = n$*. Dann gilt:*

$$\dim \mathrm{Kern}(L) + \dim \mathrm{Bild}(L) = n\,.$$

Wir wollen diese Aussage zunächst für Matrizen interpretieren. Sei also A eine $m \times n$-Matrix und L_A wie oben. Das Bild von L_A ist, wie gerade festgehalten, derjenige Unterraum von $\mathbb{R}^m$, der von den Spaltenvektoren von A erzeugt wird. Die Dimension dieses Unterraums stimmt überein mit der maximalen Anzahl linear unabhängiger Spalten von A, man nennt diese Zahl auch den *Spaltenrang* $\mathrm{Rang}(A)$.

Die Dimensionsformel liefert jetzt folgende Beziehung:

$$\dim \mathrm{Kern}(L_A) = n - \mathrm{Rang}(A)\,.$$

Bereits im ersten Paragraphen hatten wir im Zusammenhang mit dem Gaussschen Eliminationsverfahren eine ähnliche Beziehung für die Dimension des Lösungsraumes $\mathbb{L} = \mathrm{Kern}(L_A)$ gefunden, nämlich $\dim \mathbb{L} = n - r$, wobei r der Rang der durch elementare Zeilenumformungen auf Zeilenstufenform transformierten Matrix A' war. Der Rang einer Matrix in Zeilenstufenform gibt die Anzahl der Nichtnullzeilen an und stimmt überein mit der maximalen Anzahl linear unabhängiger Zeilen, also dem *Zeilenrang* von A'. Nun bleibt der Zeilenrang einer Matrix bei elementaren Zeilenumformungen aber unverändert. Wir erhalten also folgendes Ergebnis:

2.2.7 Folgerung *Der Zeilen- und der Spaltenrang einer Matrix stimmen miteinander überein. Man spricht deshalb kurz vom Rang einer Matrix. Der Rang gibt sowohl die Anzahl linear unabhängiger Spalten als auch die Anzahl linear unabhängiger Zeilen der Matrix an.*

Beweis der Dimensionsformel. Weil der Kern von L ein Unterraum von V ist, gilt sicher $k := \dim(\mathrm{Kern}(L)) \leq n$. Ausserdem können wir eine Basis $(v_1, \dots, v_k)$ von $\mathrm{Kern}(L)$ wählen und zu einer Basis $(v_1, \dots, v_k, v_{k+1}, \dots, v_n)$ von V ergänzen. Es reicht jetzt, folgende Behauptung zu beweisen:

$$(L(v_{k+1}), \dots, L(v_n)) \quad \text{ist eine Basis für das Bild von } L.$$

Dazu zeigen wir zunächst, dass die Menge ein Erzeugendensystem für das Bild ist. Sei also $w \in \mathrm{Bild}(L)$. Dann gibt es ein $v \in V$ mit $L(v) = w$. Wir schreiben v als Linearkombination der Basiselemente in der Form $v = \sum_{i=1}^n \alpha_i v_i$ ($\alpha_i \in \mathbb{R}$). Dann folgt $w = L(v) = \sum_{i=1}^n \alpha_i L(v_i) = \sum_{k+1}^n \alpha_i L(v_i)$, weil $L(v_i) = 0$ für alle $i \leq k$. Also liegt w in der linearen Hülle von $(L(v_{k+1}), \dots, L(v_n))$.

Im zweiten Schritt zeigen wir jetzt, dass die Menge linear unabhängig ist. Angenommen

$$0 = \sum_{i=k+1}^{n} \alpha_i L(v_i) = L(\sum_{i=k+1}^{n} \alpha_i v_i)\,.$$

Das bedeutet, der Vektor $u := \sum_{i=k+1}^n \alpha_i v_i$ liegt im Kern der Abbildung L, lässt sich also in der Basis $(v_1, \dots, v_k)$ schreiben. Das heisst, es gibt Zahlen $\beta_1, \dots, \beta_k$ mit

$$u = \sum_{i=1}^{k} \beta_i v_i = \sum_{i=k+1}^{n} \alpha_i v_i\,.$$

Daraus folgt die Relation $\beta_1 v_1 + \cdots + \beta_k v_k - \alpha_{k+1} v_{k+1} - \cdots - \alpha_n v_n = 0$. Da $(v_1, \dots, v_n)$ linear unabhängig gewählt waren, folgt $\beta_1 = \cdots = \beta_k = \alpha_{k+1} = \cdots = \alpha_n = 0$. q.e.d.

2.2.8 Folgerung *Seien V, W endlichdimensionale Vektorräume und $L: V \to W$ linear. Dann gilt: L ist genau dann bijektiv, wenn* $\dim V = \dim W$ *und* $\mathrm{Kern}(L) = \{0\}$*. In diesem Fall ist auch die Umkehrabbildung von L linear und man bezeichnet L als Vektorraumisomorphismus.*

2.2.9 Beispiel Sei V ein Vektorraum der Dimension n und $\mathcal{A}$ eine Basis für V. Dann ist die Zuordnung

$$V \to \mathbb{R}^n, \quad v \mapsto \mathrm{Koeff}_{\mathcal{A}}(v)$$

ein Vektorraumisomorphismus. Das bedeutet, jeder endlichdimensionale Vektorraum ist zu einem der Räume $\mathbb{R}^n$ $(n \in \mathbb{N}_0)$ isomorph.

2.2.10 Beispiel Eine quadratische Matrix A definiert genau dann eine bijektive lineare Abbildung $L_A: \mathbb{K}^n \to \mathbb{K}^n$, wenn A invertierbar ist, wenn also $\det A \neq 0$ ist. Ist dies der Fall, wird die Umkehrabbildung durch die Multiplikation mit der inversen Matrix A^{-1} beschrieben.

Die Matrix, die eine lineare Abbildung beschreibt, hängt wesentlich von der Wahl der Basen — also der Koordinatensysteme — ab! Hierzu das schon erwähnte Beispiel der Spiegelung L der Ebene an einer Geraden g durch den Nullpunkt, die mit der x–Achse den Winkel α bildet. Die Matrix von L bezüglich der kanonischen Basis lautet, wie schon früher festgestellt:

$$M_{(e_1,e_2)}(L) = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} .$$

Wählt man dagegen als Basisvektoren einen Vektor $v_1 \neq 0$ in der Richtung von g und einen dazu senkrechten Vektor $v_2 \neq 0$, dann ist $L(v_1) = v_1$ und $L(v_2) = -v_2$ und daher lautet die zugehörige Matrix:

$$M_{(v_1,v_2)}(L) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Man kann durch Wahl einer günstigen Basis versuchen, die Abbildung durch eine möglichst einfache Matrix zu beschreiben, an der sich wichtige Eigenschaften direkt ablesen lassen. Dazu sei hier noch beschrieben, wie sich die Matrix einer linearen Selbstabbildung von V bei einem Basiswechsel ändert.

2.2.11 Satz *Sei $L: V \to V$ eine lineare Selbstabbildung des Vektorraums V. Seien weiter $\mathcal{A}$, $\mathcal{B}$ Basen von V und seien $A := M_{\mathcal{A}}(L)$ und $B := M_{\mathcal{B}}(L)$ die zugehörigen Matrizen. Dann gilt:*

$$B = T^{-1}AT\,, \quad \text{wobei}$$

die Transformationsmatrix T den Basiswechsel von $\mathcal{B}$ nach $\mathcal{A}$ beschreibt, das heisst, die Spalten von T sind die Koeffizientenvektoren der Vektoren aus $\mathcal{B}$, ausgedrückt in der Basis $\mathcal{A}$. Ist speziell $V = \mathbb{K}^n$ und $\mathcal{A}$ die kanonische Basis, erhält man T einfach, indem man die Elemente von $\mathcal{B}$ als Spalten zu einer Matrix zusammenfügt.

Beweis. Nach Wahl der Transformationsmatrix gilt für jedes $v \in V$:

$$T \cdot \mathrm{Koeff}_{\mathcal{B}}(v) = \mathrm{Koeff}_{\mathcal{A}}(v)\,.$$

Daraus folgt $\quad AT\,\mathrm{Koeff}_{\mathcal{B}}(v) = \mathrm{Koeff}_{\mathcal{A}}(Lv)$, und das liefert, wie behauptet

$$T^{-1}AT \cdot \mathrm{Koeff}_{\mathcal{B}}(v) = \mathrm{Koeff}_{\mathcal{B}}(Lv)\,.$$

Hier sind die Zusammenhänge nochmals in einem Diagramm dargestellt:

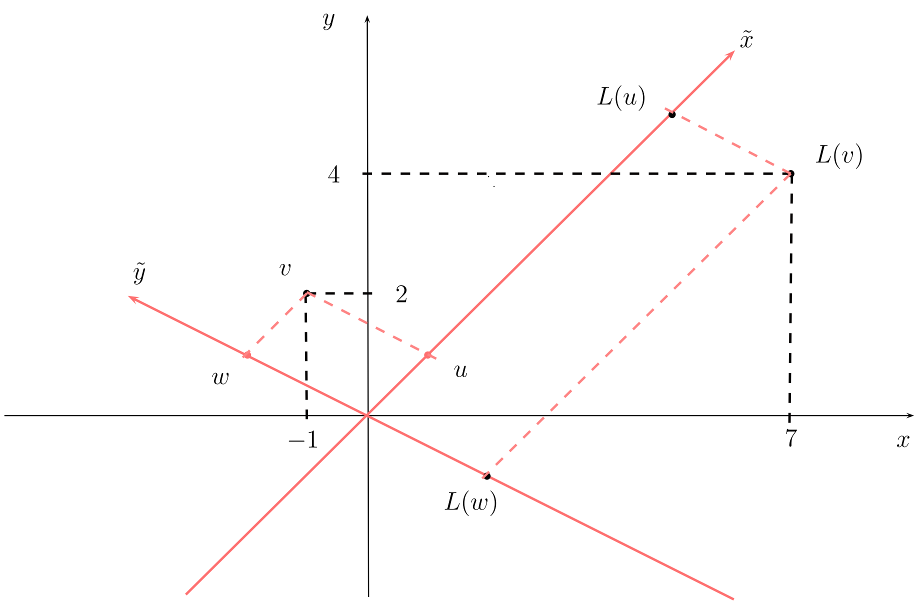
$$\begin{tikzcd}
\mathbb{R}^n \arrow[rrr, "L_B"] \arrow[dd, "T"'] & & & \mathbb{R}^n \\
& V \arrow[ul, "\mathcal{B}"'] \arrow[dl, "\mathcal{A}"] \arrow[r, "L"] & V \arrow[ur, "\mathcal{B}"] \arrow[dr, "\mathcal{A}"'] & \\
\mathbb{R}^n \arrow[rrr, "L_A"] & & & \mathbb{R}^n \arrow[uu, "T^{-1}"']
\end{tikzcd}$$

q.e.d.

2.2.12 Beispiel Sei wiederum $V = \mathbb{R}^2$, $\mathcal{A} = (e_1, e_2)$, und sei L die lineare Abbildung, festgelegt durch $L(e_1) = e_1 + 2e_2$ und $L(e_2) = 4e_1 + 3e_2$. Dann ist $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$. Sei weiter $\mathcal{B}$ die Basis, gebildet aus den Vektoren $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $w = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$. Also ist hier $T = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$ und $T^{-1} = \frac{1}{\det(T)} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$. Die Matrix von L bezogen auf die Basis $\mathcal{B}$ lautet daher:

$$B = T^{-1}AT = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}\,.$$

Das bedeutet, wenn wir einen Vektor durch die Basis $\mathcal{B}$ ausdrücken in der Form $v = \tilde{x}u + \tilde{y}w$, dann ist $L(v) = 5\tilde{x}u - \tilde{y}w$.



Hier ist noch ein weiteres Beispiel:

2.2.13 Beispiel Sei $V = \mathbb{R}^3$, $\mathcal{A} = (e_1, e_2, e_3)$ und $\mathcal{B} = (e_3, e_1, e_2)$. Die räumliche Drehung L um die e_1–Achse um den Winkel α wird, bezogen auf die Basis $\mathcal{A}$, beschrieben durch die Matrix $M_{\mathcal{A}}(L) = A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$. Die Transformationsmatrix, die den Basiswechsel von $\mathcal{B}$ nach $\mathcal{A}$ angibt, lautet hier $T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.
Also ist

$$M_{\mathcal{B}}(L) = B = T^{-1}AT = \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix}.$$

In den Spalten dieser Matrix stehen die Bilder der Vektoren e_3, e_1, e_2, jeweils ausgedrückt in Koordinaten bezogen auf die Basis $\mathcal{B}$.

2.3 Eigenwerte und Eigenvektoren

Lineare Abbildungen werden je nach Basiswahl durch unterschiedliche Matrizen beschrieben. Besonders einfach ist die Diagonalform. Wir werden in diesem Abschnitt der Frage nachgehen, welche linearen Abbildungen sich durch eine Diagonalmatrix darstellen lassen.

Schauen wir uns zunächst genauer an, welche Wirkung eine durch eine Diagonalmatrix definierte lineare Abbildung hat.

2.3.1 Beispiel Sei $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ die durch Multiplikation mit der Matrix $A = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{pmatrix}$ definierte Abbildung. Dann gilt: $L(e_1) = Ae_1 = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} = \frac{1}{2}e_1$. Der Vektor e_1 behält also unter der Abbildung seine Richtung, aber seine Länge wird halbiert. Dasselbe gilt für alle Vektoren, die in der x-Achse liegen. Die x-Achse als Ganzes bleibt also stabil.

Weiter ist $L(e_2) = Ae_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 3e_2$. Der Vektor e_2 wird also von der Abbildung L um den Faktor 3 gestreckt, ebenso wie alle Vielfachen von e_2. Die y-Achse bleibt also ebenfalls stabil unter L.

2.3.2 Definition Eine Zahl $\lambda \in \mathbb{R}$ ist ein *Eigenwert* einer linearen Abbildung $L\colon V \to V$, falls ein Vektor $0 \neq v \in V$ existiert mit $L(v) = \lambda \cdot v$. In diesem Fall bezeichnet man v als einen zum Eigenwert λ gehörigen *Eigenvektor*. Jedes Vielfache $w = \alpha v \neq 0$ von v ist ebenfalls ein Eigenvektor zum Eigenwert λ. Denn $L(w) = \alpha L(v) = \alpha\lambda v = \lambda w$. Die von v aufgespannte Gerade $\mathrm{lin}(v)$ bleibt unter der Abbildung L stabil, es ist eine *Eigenrichtung* von L.

2.3.3 Beispiele 1. Die durch die Matrix $A = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{pmatrix}$ definierte Abbildung hat den Eigenvektor e_1 zum Eigenwert $\frac{1}{2}$ und e_2 zum Eigenwert 3.

2. Sei $L: \mathbb{R}^2 \to \mathbb{R}^2$ die Multiplikation mit der Matrix $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ wie in Beispiel 2.2.12. Die Vektoren $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $w = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ sind Eigenvektoren von L zum Eigenwert 5 bzw. -1. Denn $L(u) = 5u$ und $L(w) = -w$.

2.3.4 Definition Eine lineare Abbildung $L: V \to V$ heisst *diagonalisierbar*, wenn es eine Basis $\mathcal{B}$ von V gibt, so dass $M_{\mathcal{B}}(L)$ eine Diagonalmatrix ist.

Beispielsweise sind Spiegelungen diagonalisierbar, Drehungen dagegen nicht.

2.3.5 Satz *Die Abbildung L ist genau dann diagonalisierbar, wenn sie n linear unabhängige Eigenvektoren in V besitzt (für $n = \dim V$).*

Beweis. Für eine Basis $\mathcal{B}$ von V gilt genau dann

$$M_{\mathcal{B}}(L) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wenn $L(v_j) = \lambda_j v_j$ für $j = 1, \ldots, n$. Das bedeutet aber gerade, dass die Basis $\mathcal{B}$ nur aus Eigenvektoren von L besteht. q.e.d.

Man kann Eigenwerte, Eigenvektoren und Diagonalisierbarkeit auch für Matrizen definieren.

2.3.6 Definition Eine Zahl $\lambda \in \mathbb{R}$ ist ein *Eigenwert* einer $n \times n$-Matrix A, falls ein Vektor $0 \neq v \in \mathbb{R}^n$ existiert mit $Av = \lambda \cdot v$. In diesem Fall bezeichnet man v als einen zum Eigenwert λ gehörigen *Eigenvektor*. Die Matrix A heisst *diagonalisierbar*, wenn es eine invertierbare $n \times n$-Matrix T gibt, so dass $T^{-1}AT$ Diagonalform hat.

Nun gilt wieder der entsprechende Satz über Diagonalisierbarkeit:

2.3.7 Satz *Eine $n \times n$-Matrix A ist genau dann diagonalisierbar, wenn sie n linear unabhängige Eigenvektoren in $\mathbb{R}^n$ besitzt.*

Beweis. Nehmen wir an, $v_1, \ldots, v_n$ seien linear unabhängige Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Wir bilden aus den Spalten $v_1, \ldots, v_n$ eine Matrix T. Diese Matrix ist dann invertierbar und es gilt einerseits:

$$AT = (\, Av_1 \quad \ldots \quad Av_n \,) = (\, \lambda_1 v_1 \quad \ldots \quad \lambda_n v_n \,)\ .$$

Andererseits ist

$$T \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (v_1 \ldots v_n) \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (\, \lambda_1 v_1 \quad \ldots \quad \lambda_n v_n \,)\ .$$

Also hat $T^{-1}AT$ Diagonalform. Das heisst, A ist diagonalisierbar. Diese Argumentation lässt sich auch umkehren. q.e.d.

Nun wollen wir die Frage behandeln, wie man Eigenwerte und Eigenvektoren einer vorgegebenen Matrix finden kann. Nehmen wir zunächst an, $v \in \mathbb{R}^n$ sei ein Eigenvektor der Matrix A zum Eigenwert λ. Dann ist $v \neq 0$ und es gilt $\lambda v - Av = (\lambda E - A)v = 0$. Das Gleichungssystem $(\lambda E - A)x = 0$ hat also zusätzlich zu der trivialen Lösung $x = 0$ noch eine Lösung $v \neq 0$ in $\mathbb{R}^n$. Daraus folgt

$$\det(\lambda E - A) = 0\,.$$

Betrachtet man jetzt λ als Unbekannte, so gilt: Die Eigenwerte von A sind gerade die Lösungen der Gleichung $\det(\lambda E - A) = 0$.

Für die Matrix $B = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$ zum Beispiel ist

$$\det(\lambda E - B) = \det \begin{pmatrix} \lambda - 3 & -2 \\ -1 & \lambda - 4 \end{pmatrix} = \lambda^2 - 7\lambda + 10\,.$$

Die Nullstellen dieses Polynoms, $\lambda_1 = 5$ und $\lambda_2 = 2$, sind die Eigenwerte der Matrix B. Allgemein gilt:

2.3.8 Satz *Sei $A \in M_{n\times n}$. Dann ist $p_A(\lambda) := \det(\lambda E - A)$ ein normiertes Polynom in λ von Grad n. Man nennt p_A das charakteristische Polynom von A. Die reellen Eigenwerte von A sind gerade die reellen Nullstellen von p_A. Deshalb hat A höchstens n verschiedene Eigenwerte.*

Beweis. Wir zeigen per Induktion über n, dass p_A ein Polynom von Grad n ist. Für $n = 1$ ist $A = (a)$ und $p_A = \lambda - a$. Für $n > 1$ entwickeln wir die Determinante von $\lambda E - A$ nach der ersten Spalte:

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & & \dots & \lambda - a_{nn} \end{pmatrix} \\ &= (\lambda - a_{11}) \det(\lambda E_{n-1} - A_{11}) + \sum_{k=2}^{n} (-1)^{k+1} a_{k1} \det((\lambda E - A)_{k1})\,. \end{aligned}$$

Per Induktion ist $\det(\lambda E_{n-1} - A_{11})$ ein Polynom von Grad $n-1$, und $\det((\lambda E - A)_{k1})$ sind Polynome von Grad $\leq n-2$ für alle $k \geq 2$. Denn $(\lambda E - A)_{k1}$ entsteht aus $\lambda E - A$ durch Streichung der ersten Spalte und der k-ten Zeile, die Variable λ kommt also nur noch in jeweils $n-2$ Zeilen vor. Also folgt, dass der Grad von p_A gleich n ist. q.e.d.

Schauen wir uns den Fall $n = 2$ genauer an. Das charakteristische Polynom einer 2×2-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ lautet $p_A(\lambda) = \det(\lambda E - A) = \det \begin{pmatrix} \lambda - a & -b \\ -c & \lambda - d \end{pmatrix} =$

$(\lambda - a)(\lambda - d) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$. Die Summe der Diagonaleinträge einer Matrix wird als ihre Spur bezeichnet. Damit gilt:

$$p_A(\lambda) = \lambda^2 - \operatorname{Spur}(A) \cdot \lambda + \det A\,.$$

Durch vollständige Induktion kann man zeigen, dass die Spur und die Determinante einer Matrix, wie im zweidimensionalen Fall schon nachgerechnet, immer als Koeffizienten des charakteristischen Polynoms auftreten. Genauer:

2.3.9 Satz *Für $n \times n$-Matrizen A gilt:*

$$p_A(\lambda) = \lambda^n - \operatorname{Spur}(A)\lambda^{n-1} + \cdots + (-1)^n \det(A)\,.$$

Bezeichnet man mit $\lambda_1, \ldots, \lambda_n$ sämtliche (möglicherweise auch komplexen) Nullstellen von p_A und zwar mit Vielfachheit gezählt, dann folgt:

$$\operatorname{Spur}(A) = \lambda_1 + \cdots + \lambda_n \quad \text{und} \quad \det(A) = \lambda_1 \cdot \cdots \cdot \lambda_n\,.$$

Um nun zu einem gegebenen Eigenwert λ von A die zugehörigen Eigenvektoren zu bestimmen, ist das lineare Gleichungssystem

$$(\lambda E - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

zu lösen. Den Lösungsraum dieses Gleichungssystem bezeichnet man als den *Eigenraum* zum Eigenwert λ. Wir schreiben dafür $\mathbb{L}_\lambda$. Der Raum $\mathbb{L}_\lambda$ besteht aus allen Eigenvektoren zum Eigenwert λ zusammen mit dem Nullvektor. Für jeden Eigenwert λ gilt:

$$\dim \mathbb{L}_\lambda = n - \operatorname{Rang}(\lambda E - A) \geq 1\,.$$

2.3.10 Beispiele 1. Das charakteristische Polynom der Matrix $B = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$ lautet $p_B(\lambda) = \lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)$. Um den Eigenraum zum Eigenwert $\lambda_1 = 5$ zu bestimmen, betrachten wir:

$$(5E - B)v = \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Der Lösungsraum dazu ist

$$\mathbb{L}_5 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid x = y \right\} = \left\{ \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

Für den Eigenwert $\lambda_2 = 2$ ergibt sich entsprechend:

$$(2E - B)v = \begin{pmatrix} -1 & -2 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Der Lösungsraum dazu ist

$$\mathbb{L}_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid x = -2y \right\} = \left\{ \alpha \begin{pmatrix} -2 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

2. Das charakteristische Polynom der Matrix $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ lautet

$$p_A(\lambda) = \det(\lambda E - A) = \det \begin{pmatrix} \lambda - 1 & -1 & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & \lambda - 2 \end{pmatrix} = (\lambda - 1)^2(\lambda - 2)\,.$$

Also hat A einen doppelten Eigenwert, nämlich $\lambda_1 = 1$ und einen einfachen Eigenwert $\lambda_2 = 2$. Es gilt $\mathbb{L}_2 = \mathrm{lin}(e_3)$. Bestimmen wir nun den Eigenraum zu dem doppelten Eigenwert. Dazu lösen wir das Gleichungssystem

$$(E - A)v = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0\,.$$

Der Lösungsraum $\mathbb{L}_1 = \mathrm{lin}(e_1)$ ist nur eindimensional. Die Matrix A kann deshalb nicht diagonalisierbar sein.

Eigenvektoren zu verschiedenen Eigenwerten sind automatisch linear unabhängig (siehe Übungsaufgabe). Deshalb gilt:

2.3.11 Satz *Eine $n \times n$-Matrix A ist genau dann diagonalisierbar, wenn gilt:*

$$\dim(\mathbb{L}_{\lambda_1}) + \ldots + \dim(\mathbb{L}_{\lambda_r}) = n\,.$$

Dabei bezeichnen $\lambda_1, \ldots, \lambda_r$ sämtliche verschiedenen Eigenwerte von A.

2.3.12 Folgerung *Hat eine $n \times n$-Matrix n verschiedene Eigenwerte, dann ist sie diagonalisierbar.*

2.3.13 Beispiel Das charakteristische Polynom der Matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ lautet

$$p_A(\lambda) = \det \begin{pmatrix} \lambda - 1 & -2 & -3 \\ -2 & \lambda - 3 & -1 \\ -3 & -1 & \lambda - 2 \end{pmatrix} = \lambda^3 - 6\lambda^2 - 3\lambda + 18\,.$$

Der Vektor $v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 6. Das heisst, p_A hat u.a. die Nullstelle 6. Wenn wir p_A durch $(\lambda - 6)$ teilen, erhalten wir: $p_A(\lambda) = (\lambda - 6)(\lambda^2 - 3)$. Also hat A die drei verschiedenen Eigenwerte $6, \sqrt{3}, -\sqrt{3}$, und ist daher diagonalisierbar.

Wir wollen nun zeigen, dass sich das charakteristische Polynom einer Matrix bei einem Basiswechsel nicht ändert. Dazu führen wir folgenden Begriff ein.

2.3.14 Definition Zwei Matrizen $A, B \in M_{n\times n}$ heissen *ähnlich*, wenn eine invertierbare $n \times n$-Matrix T existiert, so dass

$$B = T^{-1}AT\,.$$

2.3.15 Satz *Sind A, B ähnliche $n \times n$-Matrizen, so gilt:*

$$p_A = p_B \quad \text{und insbesondere} \quad \mathrm{Spur}(A) = \mathrm{Spur}(B) \quad \text{und} \quad \det(A) = \det(B)\,.$$

Beweis. Die charakteristischen Polynome stimmen überein, denn

$$\begin{aligned} p_B(\lambda) &= \det(\lambda E - B) = \det(\lambda T^{-1}T - T^{-1}AT) = \det(T^{-1}(\lambda E - A)T) \\ &= \det(\lambda E - A) = p_A(\lambda)\,. \qquad \text{q.e.d.} \end{aligned}$$

2.3.16 Definition Sei $L\colon V \to V$ eine lineare Abbildung, $\dim V = n$. Sei $A = M_{\mathcal{B}}(L)$ für eine Basis $\mathcal{B}$ von V. Dann nennt man $p(\lambda) := p_A(\lambda)$ auch das *charakteristische Polynom* von L. Das Polynom p hängt nicht von der Wahl der Basis ab. Denn Basiswechsel führen zu ähnlichen Matrizen.