Prof. Dr. Michael Eisermann • Lineare Algebra

Kapitel M

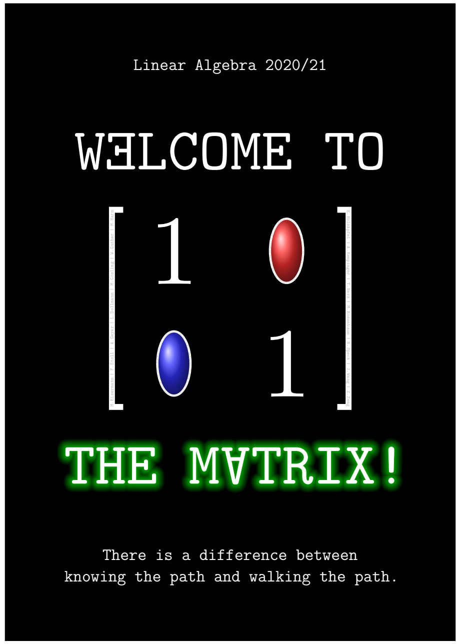
Eigenvektoren und Diagonalisierung

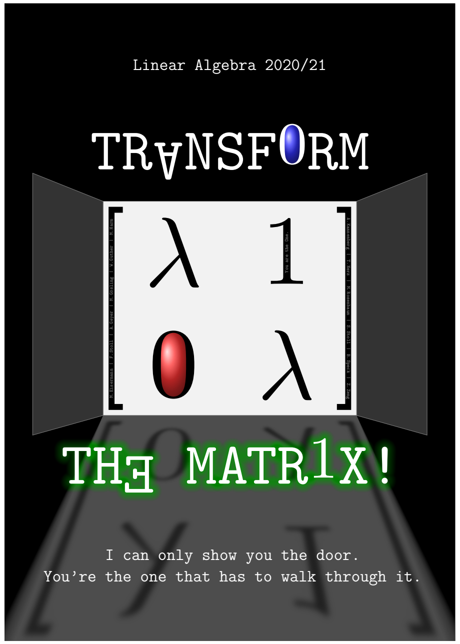
Live as if you were to die tomorrow.
Learn as if you were to live forever.
Mahatma Gandhi (1869–1948)

Vollversion • eiserm.de/lehre/LinA • 05.03.2022

Inhalt dieses Kapitels M

Willkommen zur Fortsetzung der Linearen Algebra!

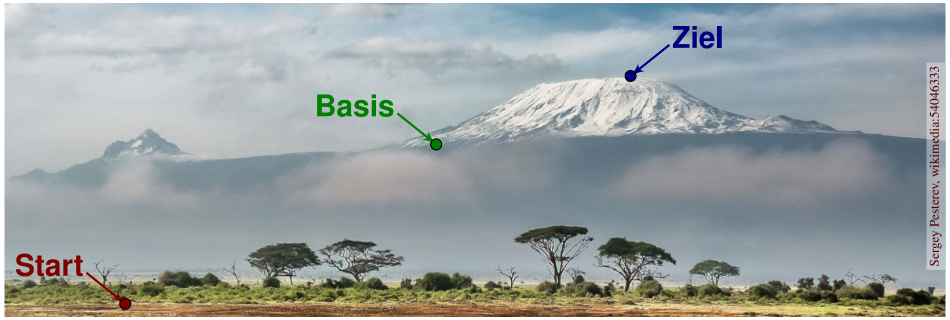




Das Ziel und der Weg

Mathematik ist schön und nützlich, zwar anstrengend doch lohnend!

Five percent of the people think; ten percent of the people think they think; and the other eighty-five percent would rather die than think.“
Thomas A. Edison (1847–1931)



„Because in the end, you won't remember the time you spent working in the office or mowing your lawn. Climb that goddamn mountain!“
Jack Kerouac (1922–1969)

Rückblick auf das Wintersemester: die Grundlagen

Mathematische Grundlagen	**Algebraische Grundlagen**	**Lineare Strukturen**
Mathematische Logik und Beweistechniken	Monoide und Gruppen	Lineare Räume und lineare Abbildungen
Mengen und Abbildungen	Ringe und Körper	Basis und Dimension
Kombinatorik und Quotienten	Polynomringe	Darstellung linearer Abbildungen durch Matrizen
Ordnungsrelationen und Kardinalität	Matrixringe	Signatur und Determinante

---

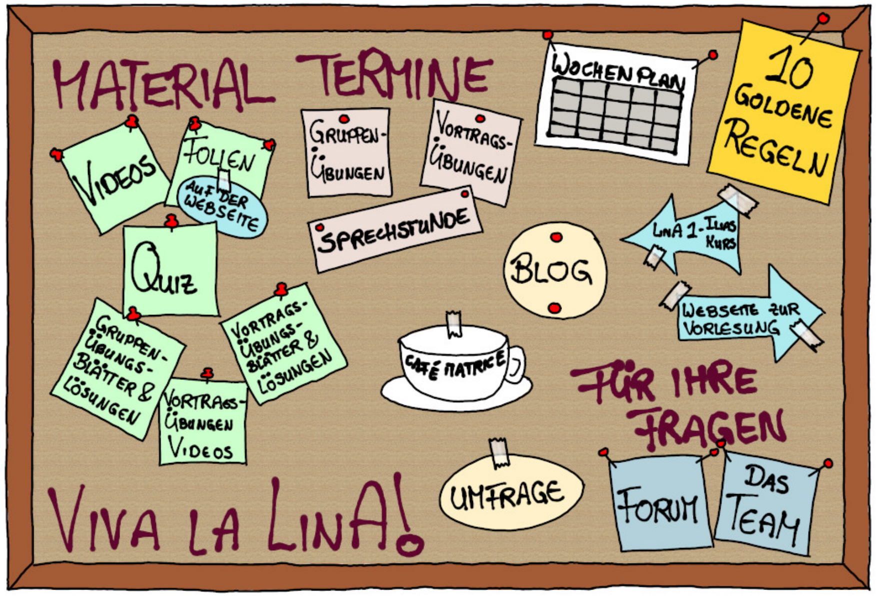
Ausblick auf das Sommersemester: weiterer Aufbau

Normalformen für Endos	**Bilineare Algebra**	**Euklidische Geometrie**	**Multilineare Algebra**
Diagonalisierung	Bilinearformen	Skalarprodukte	Dualität
Jordanisierung	Quadriken	Spektralsatz	Tensorprodukt

Im Verlauf dieses Sommersemesters werden auch diese neuen Themen für Sie konkrete Gestalt annehmen. Weiterhin gilt: Die Mathematik ist wunderschön und nützlich, darauf dürfen Sie sich freuen!

Wo finden Sie Materialien und Unterstützung?

Alle Angebote finden Sie in unserem liebevoll gestalteten Ilias-Kurs.



Ausblick auf das zweite Semester

Wir unterstützen Sie auch digital bestmöglich beim Lernen.

- Lehrvideos mit Skript, ergänzend Lehrbücher
- Gut abgestimmte Vorlesung und Übungen
- Ein erfahrenes und hochmotiviertes Team

Sprechen Sie mit uns! Nutzen Sie die vielfältigen Kontaktmöglichkeiten!

Tipps zum aktiven Lernen mit Vorlesungsvideos:

- Nutzen Sie die Pausetaste. Justieren Sie Ihre Geschwindigkeit.
- Halten Sie Stift und Papier bereit. Machen Sie sich Notizen.
- Führen Sie Nebenrechnungen aus nach Ihrem Bedarf.

Studieren bereitet Freude und erfordert Disziplin!

Kanonische Darstellung einer linearen Abbildung

Ziel: Wir wollen lineare Abbildungen möglichst einfach darstellen.

Sei K ein Körper und $m, n, r \in \mathbb{N}$ mit $r \le \min\{m, n\}$.
Die **Modellmatrix** der Größe $m \times n$ vom Rang r ist

$$D = D^r_{m\times n} := \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 1_{r\times r} & 0_{r\times n'} \\ 0_{m'\times r} & 0_{m'\times n'} \end{bmatrix}.$$

Die zugehörige K–lineare **Modellabbildung** ist

$$f_D \;:\; K^n \to K^m \;:\; (x_1, \dots, x_r, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0).$$

Daran lesen wir insbesondere Bild und Kern ab:

$$\begin{aligned} \operatorname{im}(f_D) &= \operatorname{im}(D^r_{m\times n}) = \langle\, e_1, \dots, e_r \,\rangle^!_K \le K^m, \\ \ker(f_D) &= \ker(D^r_{m\times n}) = \langle\, e_{r+1}, \dots, e_n \,\rangle^!_K \le K^n. \end{aligned}$$

Kanonische Darstellung einer linearen Abbildung

In geeigneten Basen sieht jede lineare Abbildung genau so aus:

$$\begin{tikzcd} K^n \arrow[rr, "{f_D : (x_1, \dots, x_r, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0)}"'] & & K^m \\ V \arrow[u, phantom] \arrow[rr, "{f : \begin{cases} v_i \mapsto w_i & \text{für } i = 1, \dots, r, \\ v_i \mapsto 0 & \text{für } i = r+1, \dots, n \end{cases}}"] & & W \\ \end{tikzcd}$$

$$\begin{tikzcd} K^n \arrow[d, "\Phi_{\mathcal{B}}"', "\cong"] \arrow[r, "f_D"] & K^m \arrow[d, "\Phi_{\mathcal{C}}"', "\cong"] \\ V \arrow[r, "f"] & W \end{tikzcd}$$

♦ Satz K2C: kanonische Darstellung einer linearen Abbildung

Sei $f : V \to W$ eine lineare Abbildung von K–Vektorräumen endlicher Dimension $n := \dim_K(V)$ und $m := \dim_K(W)$ mit Rang $r := \operatorname{rang}_K(f)$.

Dann existieren Basen $\mathcal{B} = (v_1, \dots, v_n)$ von V und $\mathcal{C} = (w_1, \dots, w_m)$ von W mit $f(v_i) = w_i$ für $i = 1, \dots, r$ und $f(v_i) = 0$ für $i = r+1, \dots, n$.

Somit wird f dargestellt durch die Modellmatrix $D^r_{m\times n} = \mathrm{M}^{\mathcal{C}}_{\mathcal{B}}(f)$.

Diese Matrix ist so einfach und übersichtlich wie möglich. Das bringt uns zum allgemeinen Ziel dieses Kapitels: Wir wollen nun Endomorphismen $f : V \to V$ so einfach wie möglich durch eine „Normalform“ darstellen.

Gauß–Algorithmus zur kanonischen Darstellung

Ausprobieren
mit Gaël!

Mit Gauß wandeln wir jede Matrix $A \in K^{m\times n}$ zur Modellmatrix $D^r_{m\times n}$:

$$\begin{tikzcd}[column sep=huge, row sep=large] \begin{array}{c} A \\ \begin{bmatrix} 3 & 6 & -6 & -6 & -1 \\ -2 & -4 & 3 & 1 & 5 \\ -1 & -2 & 3 & 5 & 4 \\ 1 & 2 & 1 & 7 & 1 \end{bmatrix} \end{array} \arrow[r, leftrightarrow, "\text{Zeilen-}", "\text{operationen}"'] \arrow[d, leftrightarrow, "\text{Spalten-}"', "\text{operationen}"] & \begin{array}{c} B = S^{-1}A \\ \begin{bmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \arrow[d, leftrightarrow, "\text{Spalten-}"', "\text{operationen}"] \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ * & * & * & 0 & 0 \end{bmatrix} \arrow[r, leftrightarrow, "\text{Zeilen-}", "\text{operationen}"'] & \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ D = S^{-1}AT \end{array} \end{tikzcd}$$

☺ Wir haben $\operatorname{im} D = \langle\, e_1, \dots, e_r \,\rangle^!_K$ und $\ker D = \langle\, e_{r+1}, \dots, e_n \,\rangle^!_K$; dank $SD = AT$ folgt $\operatorname{im} A = \langle\, Se_1, \dots, Se_r \,\rangle^!_K$ und $\ker A = \langle\, Te_{r+1}, \dots, Te_n \,\rangle^!_K$.

Gauß–Algorithmus zur kanonischen Darstellung

Ausführlich haben wir hierzu das folgende, allgemeine Verfahren:

♦ Satz K2F: Gauß–Algorithmus zur kanonischen Darstellung

Sei $A \in K^{m\times n}$ eine Matrix über dem Körper K.

(1) Der Gauß–Algorithmus K2F liefert hierzu invertierbare Matrizen $S, S^{-1} \in \mathrm{GL}_m(K)$ und $T, T^{-1} \in \mathrm{GL}_n(K)$, sodass $AT = SD^r_{m\times n}$ gilt.

(2) Daraus folgt $\operatorname{rang}(A) = r$ und $\operatorname{def}(A) = n - r$ und explizit

$$\boxed{\begin{aligned} \operatorname{im}(A) &= \langle\, Se_1, \dots, Se_r \,\rangle^!_K, \\ \ker(A) &= \langle\, Te_{r+1}, \dots, Te_n \,\rangle^!_K. \end{aligned}}$$

☺ Dies ist ein Basiswechsel: Wir lesen die Matrix $A \in K^{m\times n}$ in den richtigen Basen, und schon vereinfacht sich A zur Modellmatrix $D^r_{m\times n}$! Diese Gauß–Normalform (GNF) löst das Klassifikationsproblem K2H.

☺ Die Bestimmung von Bild $\operatorname{im}(A)$ und Kern $\ker(A)$ haben wir bereits zuvor in Satz J1P gelöst. Mit Satz K2F sehen Sie hier nun eine elegante Umformulierung; beide Algorithmen tun im Wesentlichen dasselbe.

Matrixdarstellung eines Endomorphismus

Wir betrachten einen Endomorphismus $f : V \to V$ über dem Körper K.

$$\begin{tikzcd}
K^n \arrow[r, "f_D"] \arrow[d, "\Phi_{\mathcal{B}}"', "\cong"] & K^n \arrow[d, "\Phi_{\mathcal{B}}"', "\cong"] \\
V \arrow[r, "f"] & V
\end{tikzcd}$$

Wir suchen eine Basis $\mathcal{B}$ von V, sodass die darstellende Matrix

$$D = \mathrm{M}_{\mathcal{B}}(f) := \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) \in K^{n \times n}$$

möglichst einfach wird. Die einfachsten Matrizen sind **diagonal**:

$$D = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Zum Beispiel können wir die Potenzen von D leicht berechnen:

$$D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k) = \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n^k \end{bmatrix}$$

Matrixdarstellung eines Endomorphismus

Strenger als im allgemeinen Fall $f : V \to W$ stimmen hier Startraum V und Zielraum W überein. Daher wollen wir statt zwei Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W nur eine Basis $\mathcal{B} = \mathcal{C}$ von $V = W$ verwenden.

Das klingt auf den ersten Takt trügerisch einfacher: Statt *zwei* Basen müssen wir nur *eine* Basis wählen. Tatsächlich ist es schwieriger: Statt zwei Basen *dürfen* wir nur eine Basis wählen.

Damit haben wir weniger Möglichkeiten zur Anpassung unserer Basis, weniger Freiheitsgrade zur Problemlösung, nämlich nur „halb“ so viele! Weniger ist mehr: Weniger Spielraum bedeutet mehr Herausforderung.

Die Vereinfachung auf Diagonalform wird dadurch tatsächlich spürbar erschwert, und sie gelingt nicht immer. Auch das Klassifikationsproblem wird dadurch kniffliger. Genau darum geht es in diesem Kapitel!

Diese Problemstellung der **Diagonalisierung** tritt sehr häufig auf, und ihre Lösung ist ein vielseitiges Werkzeug der Linearen Algebra.

Matrixdarstellung eines Endomorphismus

Diagonalisierung ist nicht immer möglich. In diesem Falle weicht man notgedrungen auf die „nächstbeste“ Möglichkeit aus und sucht eine Darstellung als Blockdiagonalmatrix mit möglichst einfachen Blöcken:

$$\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{bmatrix} B_1 & 0 & 0 & 0 \\ 0 & B_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & B_k \end{bmatrix}$$

Diagonalisierung entspricht $k = n$ Blöcken $B_1, B_2, \dots, B_k \in K = K^{1 \times 1}$. Das nächstbeste sind Jordan–Blöcke

$$B_i = J(n_i, \lambda) = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix} \in K^{n_i \times n_i}.$$

Die Größen addieren sich hierbei gemäß $n = n_1 + n_2 + \dots + n_k$. Diese Jordan–Normalform (JNF) diskutieren wir im nächsten Kapitel.

Matrixdarstellung eines Endomorphismus

Der allgemeinste Fall ist die Frobenius–Normalform (FNF) mit Blöcken

$$B_i = C(P_i) = \begin{bmatrix} 0 & 0 & \dots & 0 & -p_{n_i} \\ 1 & 0 & \dots & 0 & -p_{n_i-1} \\ 0 & 1 & \ddots & \vdots & -p_{n_i-2} \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -p_1 \end{bmatrix} \in K^{n_i \times n_i}.$$

Diese Matrix ist nicht ganz so simpel wie zuvor, doch von übersichtlicher Struktur und mit erfreulich vielen Nullen. Dies ist die Begleitmatrix des Polynoms $P_i = X^{n_i} + p_1 X^{n_i - 1} + \dots + p_{n_i} X^0 \in K[X]^1_{n_i}$, siehe M2Q.

Jede lineare Abbildung $f : V \to V$ eines endlich-dimensionalen Vektorraums V lässt sich so möglichst übersichtlich darstellen. Das ist zwar nicht so schön und einfach wie eine Diagonalmatrix, aber wie gesagt das nächstbeste und dafür universell einsetzbar.

Nach diesem kurzen Überblick beschäftigen wir uns nun in diesem Kapitel mit dem schönsten und einfachsten Fall: der Diagonalisierung.

Diagonalisierung eines Endomorphismus

Definition M1A: Diagonalisierung eines Endomorphismus

(1) Sei $f : V \to V$ linear über K. Eine **diagonalisierende Basis** zu f ist eine Basis $\mathcal{B}$ von V, für die die darstellende Matrix von f diagonal ist:

$$\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Hierzu sagen wir später kurz **Eigenbasis** in der Sprechweise von M1B. Existiert eine solche Basis $\mathcal{B}$ von V, so nennen wir f **diagonalisierbar**.

(2) Sei $A \in K^{n \times n}$. Ein **diagonalisierender Basiswechsel** zu A über K ist eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$, so dass $T^{-1}AT$ diagonal ist:

$$T^{-1}AT = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Existiert eine solche Matrix T, so nennen wir A **diagonalisierbar**.

Diagonalisierung eines Endomorphismus

(1) Wir interessieren uns besonders für den endlich-dimensionalen Fall $\dim_K(V) = n < \infty$. Dann hat jede Basis $\mathcal{B}$ von V Länge n, und wir können f durch eine Matrix darstellen, wenn möglich diagonal.

Im Allgemeinen suchen wir zu $f : V \to V$ eine Basis $\mathcal{B} = (v_i)_{i \in I}$ von V, mit der Eigenschaft $f(v_i) = \lambda_i v_i$ und $\lambda_i \in K$ für jeden Index $i \in I$. Dies nennen wir eine diagonalisierende Basis zu f.

Je nach Anwendung darf diese Basis durchaus auch unendlich sein, also $\dim_K(V) = \sharp I = \infty$. Speziell für den endlichen Fall haben wir eine besonders schöne Theorie, Sätze und Techniken. Der unendliche Fall findet später in der Funktionalanalysis eine umfassende Behandlung.

(2) Die Spaltenvektoren der Transformationsmatrix $T = (v_1, \dots, v_n)$ bilden eine Basis von K^n, und diese diagonalisiert die lineare Abbildung

$$f_A \;:\; K^n \to K^n \;:\; v \mapsto Av.$$

Aus $T^{-1}AT = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ folgt nämlich $T^{-1}ATe_i = \lambda_i e_i$, somit $A(Te_i) = \lambda_i(Te_i)$. Für $v_i = Te_i$ gilt also $Av_i = \lambda v_i$, wie gewünscht.

Diagonalisierung eines Endomorphismus

☺ Die obige Definition M1A erklärt zunächst das angestrebte Ziel: Wir wollen einen Endomorphismus $f : V \to V$ über K bzw. eine quadratische Matrix $A \in K^{n \times n}$ über K diagonalisieren.

Im Folgenden erarbeiten wir uns nun die nötigen Werkzeuge: präzise Begriffe (Definitionen) und wirksame Methoden (Sätze). Wir wiederholen zunächst ein einfaches, aber illustratives Beispiel, an dem Sie alle Techniken und das Vorgehen schon erkennen können.

Dieses unorthodoxe Vorgehen – Anwendung vor Theorie – hilft Ihnen zur Orientierung, zumindest möchte ich es so anbieten. Anschließend führen wir die Techniken sorgsam aus, also Definitionen und Beispiele, Sätze und Beweise, und schließlich schöne Anwendungsbeispiele, so wie Sie es im mathematisch-logischen Aufbau erwarten.

Diagonalisierung eines Endomorphismus

☺ Beachten Sie die logische Trennung von Ziel und Weg. Es lohnt sich, zunächst das Ziel klar zu benennen, dann mögliche Wege zu suchen. Es gibt im Allgemeinen mehrere alternative Lösungsmöglichkeiten, die sollten Sie kennen, und darüber Ziel und Weg nicht verwechseln.

Manche wünschen sich sofort am Anfang ein fertiges Rezept wie das Standardverfahren zur Diagonalisierung (M2I). Das erscheint zunächst verlockend als entlastende Abkürzung, erweist sich anschließend jedoch als verwirrend und ineffizient, als erschwerender Umweg:

Allein durch Auswendiglernen versteht man weder den Sinn noch die Herleitung, weder nützliche Zusammenhänge noch korrekte Nutzung. In realistischen Anwendungen benötigen Sie jedoch genau dies! Daher lohnt sich ein gründlicher, umsichtiger Aufbau.

Erinnerung und Vorschau zur Diagonalisierung

Zu $v = \left[\begin{smallmatrix}-1\\1\end{smallmatrix}\right] \in \mathbb{R}^2$ betrachten wir die $\mathbb{R}$–lineare Abbildung

$$f \,:\, \mathbb{R}^2 \to \mathbb{R}^2 \,:\, x \mapsto x - v \cdot v^{\intercal} \cdot x.$$

Aufgabe: (1) Finden Sie alle $\lambda \in \mathbb{R}$ mit nicht-trivialem **Eigenraum**

$$\mathrm{E}(\lambda) := \left\{\, x \in \mathbb{R}^2 \;\middle|\; f(x) = \lambda x \,\right\} = \ker(f - \lambda\,\mathrm{id}).$$

(2) Bestimmen Sie die Eigenräume. (3) Diagonalisieren Sie f.

Lösung: (1) Zur Bestimmung von λ nutzen wir die **Determinante**:

$$\ker(f - \lambda\,\mathrm{id}) \neq \{0\} \quad \overset{\text{L3D}}{\iff} \quad \det(f - \lambda\,\mathrm{id}) = 0$$

Zur Berechnung wählen wir eine Basis, etwa $\mathcal{E} = (e_1, e_2)$, und finden

$$A = \mathrm{M}^{\mathcal{E}}_{\mathcal{E}}(f) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1.$$

Dies ist das **charakteristische Polynom** von A bzw. von f.
Seine Nullstellen sind in diesem Beispiel $\lambda = +1$ und $\lambda = -1$.
Dies sind die **Eigenwerte** der Matrix A bzw. der Abbildung f.

Erinnerung und Vorschau zur Diagonalisierung

(2) Wir bestimmen die Eigenräume $\mathrm{E}(+1)$ und $\mathrm{E}(-1)$ wie folgt:

$$\mathrm{E}(+1) = \ker(f - \mathrm{id}) = \ker \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \mathbb{R}\, b_1 \quad \text{mit} \quad b_1 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathrm{E}(-1) = \ker(f + \mathrm{id}) = \ker \begin{bmatrix} +1 & 1 \\ 1 & +1 \end{bmatrix} = \mathbb{R}\, b_2 \quad \text{mit} \quad b_2 := \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Für alle weiteren $\lambda \in \mathbb{R} \smallsetminus \{\pm 1\}$ gilt $\mathrm{E}(\lambda) = \{0\}$ dank Satz L3D, denn hier ist $\det(A - \lambda I) \neq 0$, also $A - \lambda I$ in $\mathbb{R}^{2\times 2}$ invertierbar.

(3) Wir erhalten $\mathbb{R}^2 = \mathrm{E}(+1) \oplus \mathrm{E}(-1)$ und die Basis $\mathcal{B} = (b_1, b_2)$ von $\mathbb{R}^2$. Nach Konstruktion gilt $f(b_1) = +1 \cdot b_1$ und $f(b_2) = -1 \cdot b_2$, und somit

$$\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix}.$$

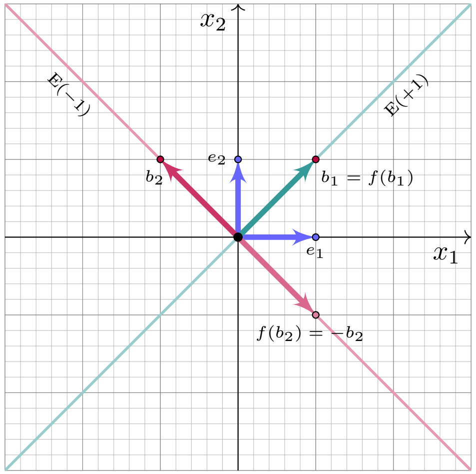
Auf Seite K237 haben wir bereits die Basiswechselmatrizen bestimmt:

$$\mathrm{T}^{\mathcal{E}}_{\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{und} \quad \mathrm{T}^{\mathcal{B}}_{\mathcal{E}} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Erinnerung und Vorschau zur Diagonalisierung

☺ In der angepassten Basis $\mathcal{B} = (b_1, b_2)$ können wir die Abbildung f besonders einfach darstellen… Den Eigenräumen sei Dank!

$$\mathrm{M}^{\mathcal{E}}_{\mathcal{E}}(f) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



$$\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ ist die Spiegelung an der Hauptdiagonalen.

Erinnerung und Vorschau zur Diagonalisierung

Eine solcherart angepasste Basis $\mathcal{B}$ zu f ist etwas ganz Besonderes: Bezüglich $\mathcal{B}$ wird f durch eine Diagonalmatrix dargestellt, wie erhofft.

Wir nennen dies eine **diagonalisierende Basis** zu f, wie oben in Definition M1A vereinbart, oder auch kurz eine **Eigenbasis** zu f.

☺ Auf jedem Eigenraum $\mathrm{E}(\lambda)$ ist die Abbildung f besonders einfach: Sie skaliert jeden Eigenvektor $v \in \mathrm{E}(\lambda)$ um den Eigenwert λ. Glücklicherweise erhalten wir hier $\mathbb{R}^2 = \mathrm{E}(+1) \oplus \mathrm{E}(-1)$, wie erhofft.

Die Eigenräume unserer Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ sind etwas Natürliches. Die Wahl der jeweiligen Basen $b_1 \in \mathrm{E}(+1)$ und $b_2 \in \mathrm{E}(-1)$ ist hingegen etwas willkürlich; wir können auch Vielfache dieser Vektoren wählen.

☺ Die so gewonnene Eigenbasis $\mathcal{B}$ zeigt uns, was f eigentlich tut: Hier offenbart die Abbildung f ihr wahres Wesen, ihren Charakter, ihr Wirken, hier erkennen wir f sofort als Spiegelung.

Eigenwerte und Eigenvektoren eines Endomorphismus

Definition M1B: Eigenwerte und Eigenvektoren

Sei $f : V \to V$ eine lineare Abbildung über dem Körper K.

(1) Zu jedem Skalar $\lambda \in K$ definieren wir den zugehörigen **Eigenraum**

$$\mathrm{E}(\lambda) = \mathrm{Eig}(f,\lambda) := \{\, v \in V \mid f(v) = \lambda v \,\} = \ker(f - \lambda\,\mathrm{id}_V) \le V.$$

Im nicht-trivialen Fall $\mathrm{E}(\lambda) \ne \{0\}$ nennen wir λ einen **Eigenwert** von f. Die Dimension $\dim_K \mathrm{E}(\lambda) \ge 1$ heißt **geometrische Vielfachheit** von λ.

(2) Die Menge aller Eigenwerte nennen wir das **(Eigenwert)Spektrum**

$$\sigma(f) = \sigma(f;K) := \{\, \lambda \in K \mid \exists v \in V \smallsetminus \{0\} : f(v) = \lambda v \,\}.$$

(3) Jedes Element $v \in \mathrm{E}(\lambda) \smallsetminus \{0\}$ heißt **Eigenvektor** zum Eigenwert λ.

(4) Eine **Eigenbasis** $\mathcal{B}$ ist eine Basis von V aus Eigenvektoren von f. Im endlichen Fall heißt das, die darstellende Matrix $\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f)$ ist diagonal.

Diese Begriffe nutzen wir ebenso für jede Matrix $A \in K^{n\times n}$ vermöge der zugehörigen K–linearen Abbildung $f = f_A : K^n \to K^n : v \mapsto Av$.

Eigenwerte und Eigenvektoren eines Endomorphismus

Zu $f \in \mathrm{End}_K(V)$ und $\lambda \in K$ definieren wir den zugehörigen Eigenraum

$$\mathrm{E}(\lambda) = \mathrm{Eig}(f,\lambda) := \{\, v \in V \mid f(v) = \lambda v \,\} = \ker(f - \lambda\,\mathrm{id}_V).$$

Die meisten Skalare $\lambda \in K$ erweisen sich dabei als uninteressant, denn meist ist der zugehörige Eigenraum trivial, also $\mathrm{E}(\lambda) = \{0\}$.

Wir betrachten daher nur die interessanten Werte $\lambda \in K$ mit $\mathrm{E}(\lambda) \ne \{0\}$. Einen solchen Skalar $\lambda \in K$ mit $\mathrm{E}(\lambda) \ne \{0\}$ nennen wir zur Betonung einen **Eigenwert** von f, manche sagen **charakteristischer Wert**, engl. *eigenvalue* oder *proper value* oder *characteristic value*.

Aus demselben Grund verlangen wir von einem Eigenvektor stets $v \ne 0$.

⚠ Der Nullvektor erfüllt $f(0) = 0 = \lambda 0$ für alle $\lambda \in K$; diese Gleichung gilt immer und ist ohne jedes Interesse. Anders gesagt: Der Nullvektor gehört zu jedem Eigenraum $\mathrm{E}(\lambda)$, ist aber niemals ein Eigenvektor.

⚠ Der Skalar $\lambda = 0$ kann durchaus ein Eigenwert sein. Dies geschieht genau dann, wenn f nicht injektiv ist, denn es gilt $\mathrm{E}(0) = \ker(f)$. Bitte lesen Sie die Definition aufmerksam durch und laut vor.

Eigenwerte und Eigenvektoren eines Endomorphismus

Abkürzungen: EW = EWert = Eigenwert, EV = EVektor = Eigenvektor, ebenso ER = ERaum = Eigenraum, EB = EBasis = Eigenbasis, etc.

Beispiel: Sei $V \ne \{0\}$ und $f = \mathrm{id}_V$. Dann ist $\lambda = 1$ der einzige EW, $\mathrm{Eig}(f,1) = V$ ist der ER, die geometrische Vielfachheit ist $\dim_K(V)$. Jeder Vektor $v \in V \smallsetminus \{0\}$ ist ein EV. Jede Basis von V ist eine EB zu f.

Beispiel M1C: eine Diagonalmatrix

Wir betrachten eine **Diagonalmatrix**

$$A = \mathrm{diag}(\lambda_1,\dots,\lambda_n) = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \in K^{n\times n}.$$

Hier gilt $\ker(A - \lambda I) \ne \{0\}$ genau dann, wenn $\lambda \in \{\lambda_1,\dots,\lambda_n\} = \sigma(A)$. Der Vektor e_k ist EV zum EW λ_k, ebenso jedes Vielfache $v \in Ke_k \smallsetminus \{0\}$. Die Standardbasis $(e_1,\dots,e_n)$ des Raums K^n ist eine Eigenbasis zu A.

Zum Skalar $\lambda \in K$ gehört der Eigenraum $\mathrm{Eig}(f,\lambda) = \langle\, e_i \mid \lambda_i = \lambda \,\rangle_K^!$. Gilt $\lambda_i \ne \lambda_j$ für $i \ne j$, so haben wir n Eigenräume $\mathrm{Eig}(f,\lambda_i) = \langle\, e_i \,\rangle_K^!$.

Eigenwerte und Eigenvektoren eines Endomorphismus

Beispiel M1D: Jordan–Blöcke sind nicht-diagonalisierbar.

In jeder Dimension $n \ge 2$ betrachten wir den **Jordan–Block**:

$$B = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix} \in K^{n\times n}, \qquad B - \lambda I = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hier gilt $\mathrm{Eig}(B,\lambda) = \langle\, e_1 \,\rangle_K^!$ und $\mathrm{Eig}(B,\mu) = \{0\}$ für $\mu \ne \lambda$, also $\sigma(B) = \{\lambda\}$. Somit existiert zu B keine Eigenbasis.

Beispiel M1E: reelle Matrix ohne reelle Eigenwerte

Zu $a, b \in \mathbb{R}$ mit $b \ne 0$ betrachten wir die Matrix

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{und} \quad \begin{vmatrix} a-\lambda & -b \\ b & a-\lambda \end{vmatrix} = (a-\lambda)^2 + b^2.$$

Für alle Skalare $\lambda \in \mathbb{R}$ gilt somit $\ker(C - \lambda I) = \{0\}$, kurz $\sigma(C;\mathbb{R}) = \emptyset$. Hingegen ist $\lambda = a \pm ib \in \mathbb{C}$ ein Eigenwert, genauer $\sigma(C;\mathbb{C}) = \{\, a \pm ib \,\}$.

Lineare Unabhängigkeit von Eigenvektoren

Satz M1F: lineare Unabhängigkeit von Eigenvektoren

Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Ausführlich: Sei $f : V \to V$ eine lineare Abbildung über dem Körper K.

1. Seien $v_0, \dots, v_r \in V \smallsetminus \{0\}$ mit $f(v_i) = \lambda_i v_i$ und $\lambda_i \neq \lambda_j$ für $i \neq j$.
2. Es gelte $\mu_0 v_0 + \dots + \mu_r v_r = 0$ mit Koeffizienten $\mu_0, \dots, \mu_r \in K$.

Dann folgt $\mu_0 = \dots = \mu_r = 0$.

Beweis per Induktion über $r \in \mathbb{N}$:
Für $r = 0$ ist die Aussage klar, da wir $v_0 \neq 0$ voraussetzen (I1D).
Sei nun $r \geq 1$. Auf Gleichung (2) wenden wir $f - \lambda_0$ an und nutzen (1):

$$\mu_0(\lambda_0 - \lambda_0)v_0 + \mu_1(\lambda_1 - \lambda_0)v_1 + \dots + \mu_r(\lambda_r - \lambda_0)v_r = 0$$

Nach Induktionsvoraussetzung folgt $\mu_i(\lambda_i - \lambda_0) = 0$ für $i = 1, \dots, r$.
Dank der Voraussetzung $\lambda_i \neq \lambda_0$ folgt $\mu_i = 0$ für alle $i = 1, \dots, r$.
Von Gleichung (2) bleibt schließlich nur $\mu_0 v_0 = 0$, also $\mu_0 = 0$. $\boxed{\text{QED}}$

Lineare Unabhängigkeit von Eigenvektoren

Beweis mit dem Vandermonde–Trick: Gegeben sind die Vektoren $u_i = \mu_i v_i \in \operatorname{Eig}(f, \lambda_i)$ mit $u_0 + u_1 + \dots + u_r = 0$. Wir wenden f^k an:

$$\lambda_0^k u_0 + \lambda_1^k u_1 + \dots + \lambda_r^k u_r = 0.$$

Für $k = 0, 1, \dots, r$ erhalten wir so das folgende Gleichungssystem:

$$\begin{bmatrix} \lambda_0^0 & \lambda_1^0 & \dots & \lambda_r^0 \\ \lambda_0^1 & \lambda_1^1 & \dots & \lambda_r^1 \\ \vdots & & \vdots & \\ \lambda_0^r & \lambda_1^r & \dots & \lambda_r^r \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Die Vandermonde–Matrix ist invertierbar dank Satz B3A:

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_r \end{bmatrix} = \begin{bmatrix} \lambda_0^0 & \lambda_1^0 & \dots & \lambda_r^0 \\ \lambda_0^1 & \lambda_1^1 & \dots & \lambda_r^1 \\ \vdots & & \vdots & \\ \lambda_0^r & \lambda_1^r & \dots & \lambda_r^r \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Daraus folgt $u_0 = u_1 = \dots = u_r = 0$, wie behauptet. $\boxed{\text{QED}}$

Lineare Unabhängigkeit von Eigenvektoren

Zur Illustration nenne ich zwei einfache doch grundlegende Beispiele, die Lineare Algebra und Analysis elegant verbinden: die Eigenfolgen der Verschiebung (M1G) und die Eigenfunktionen der Ableitung (M1H).

Beispiel M1G: Eigenfolgen der Verschiebung

Über dem Körper K betrachten wir den Vektorraum $K^{\mathbb{N}}$ aller Folgen $f : \mathbb{N} \to K : n \mapsto f_n$, kurz $f = (f_n)_{n \in \mathbb{N}}$, mit dem Verschiebeoperator

$$\begin{aligned} s \;:\; K^{\mathbb{N}} \to K^{\mathbb{N}} \;:\; & (f_n)_{n\in\mathbb{N}} \mapsto (f_{n+1})_{n\in\mathbb{N}}, \\ & (f_0, f_1, f_2, \dots) \mapsto (f_1, f_2, f_3, \dots). \end{aligned}$$

Zu jeder Konstanten $\lambda \in K$ haben wir die Folge $e_\lambda = (\lambda^n)_{n\in\mathbb{N}}$.
Sie erfüllt $s(e_\lambda) = \lambda e_\lambda$, ist also eine Eigenfolge der Verschiebung.
Insbesondere ist die Familie $(e_\lambda)_{\lambda \in K}$ in $K^{\mathbb{N}}$ somit linear unabhängig.

Die lineare Unabhängigkeit ist nicht ganz leicht zu beweisen.
Alternativ gelingt dies direkt mit der Vandermonde–Matrix B3A.
Als Eigenfolgen erhalten wir Unabhängigkeit gratis dank Satz M1F.

Lineare Unabhängigkeit von Eigenvektoren

Beispiel M1H: Eigenfunktionen der Ableitung

Sei $I = \left]a, b\right[\subseteq \mathbb{R}$ ein offenes Intervall der reellen Zahlen, etwa $I = \mathbb{R}$.
Über $\mathbb{K} = \mathbb{R}, \mathbb{C}$ betrachten wir den Vektorraum $\mathscr{C}^\infty(I, \mathbb{K})$ aller beliebig oft differenzierbaren Funktionen $f : I \to \mathbb{K}$ mit dem Ableitungsoperator

$$\partial \;:\; \mathscr{C}^\infty \to \mathscr{C}^\infty \;:\; f \mapsto f'.$$

Zu jeder Konstanten $\lambda \in \mathbb{K}$ haben wir die Exponentialfunktion

$$e_\lambda \;:\; I \to \mathbb{K} \;:\; t \mapsto \mathrm{e}^{\lambda t}.$$

Sie erfüllt $\partial e_\lambda = \lambda e_\lambda$, ist also eine Eigenfunktion der Ableitung.
Insbesondere ist die Familie $(e_\lambda)_{\lambda \in \mathbb{K}}$ in $\mathscr{C}^\infty$ somit linear unabhängig.

Als Eigenfunktionen erhalten wir Unabhängigkeit gratis dank Satz M1F.
Versuchen Sie alternative Beweise zu formulieren (siehe etwa K143).

Dieses Beispiel ist eine nützliche Beobachtung, die wir später bei der Lösung von linearen Differentialgleichungen nutzen und weiterführen.

Eigenraumzerlegung und Diagonalisierung

Wie können wir effizient feststellen, ob $f:V\to V$ diagonalisierbar ist?

Satz M1I: Eigenraumzerlegung und Diagonalisierung

Vorgelegt sei eine lineare Abbildung $f:V\to V$ über dem Körper K.

(1) Die Summe $\mathrm{E} := \sum_{\lambda\in K} \mathrm{E}(\lambda) \le V$ aller Eigenräume ist direkt, also

$$\mathrm{E} = \bigoplus_{\lambda\in K} \mathrm{E}(\lambda) = \bigoplus_{\lambda\in\sigma(f)} \mathrm{Eig}(f,\lambda).$$

(2) Genau dann ist $f:V\to V$ diagonalisierbar, wenn $\mathrm{E}=V$ gilt, also

$$\boxed{V = \bigoplus_{\lambda\in\sigma(f)} \mathrm{Eig}(f,\lambda).}$$

(3) Im Falle $\dim_K V < \infty$ ist dies äquivalent zur Dimensionsgleichung

$$\dim_K V = \sum_{\lambda\in\sigma(f)} \dim_K \mathrm{Eig}(f,\lambda).$$

Beweis: (1) Dies folgt aus der linearen Unabhängigkeit (Satz M1F).
(2) Es gibt genügend Eigenvektoren, um eine Eigenbasis zu bilden.
(3) Dies folgt dank Additivität der Dimension (Satz J2K und J2O). QED

Eigenraumzerlegung und Diagonalisierung

Aufgabe: Führen Sie den Beweis zu (1) detailliert aus.

Lösung: (1) Die Definition I2J zur direkten Summe verlangt:

$$\mathrm{E}(\lambda_0) \cap \left(\sum_{\lambda\ne\lambda_0} \mathrm{E}(\lambda)\right) = \{0\}$$

Gegeben seien also Vektoren $v_i\in\mathrm{E}(\lambda_i)$ mit $\lambda_1,\dots,\lambda_r \in K \smallsetminus \{\lambda_0\}$.

Behauptung: Aus $v_0 = v_1 + \dots + v_r$ folgt $v_0 = 0$.

Beweis: Wir führen Induktion über r. Für $r=0$ ist die Aussage klar. Sei nun $r\ge 1$. Wir wenden $f-\lambda_r$ an und erhalten nach Kürzung:

$$v_0 = \frac{\lambda_1-\lambda_r}{\lambda_0-\lambda_r}\, v_1 + \dots + \frac{\lambda_{r-1}-\lambda_r}{\lambda_0-\lambda_r}\, v_{r-1} + \frac{\lambda_r-\lambda_r}{\lambda_0-\lambda_r}\, v_r$$

Nach Induktionsvoraussetzung für $r-1$ folgt daraus $v_0 = 0$.

Zur Deutlichkeit wiederhole ich hier den raffinierten Induktionsbeweis von Satz M1F. Alternativ können Sie den Vandermonde–Trick nutzen oder auch direkt die Aussage von M1F anwenden. Sehen Sie wie? Wir werden in Satz M3V einen weiteren Beweis kennenlernen.

Eigenraumzerlegung und Diagonalisierung

☺ Diagonalisierbarkeit bedeutet anschaulich vereinfacht formuliert: Es gibt genügend Eigenvektoren, um eine Eigenbasis zu bilden.

Hierzu untersuchen wir auch die zweite Aussage noch etwas genauer:
(2) Genau dann ist f diagonalisierbar, wenn $V = \bigoplus_{\lambda\in K} \mathrm{Eig}(f,\lambda)$ gilt.

Aufgabe: Führen Sie den Beweis zu (2) detailliert aus.
Hierzu benötigen Sie keine Voraussetzung zur Dimension.

Lösung: Die Äquivalenz beweisen wir durch akribische Buchführung.

„⇐“: Zu jedem $\mathrm{E}(\lambda)\le V$ wählen wir eine Basis $\mathcal{B}_\lambda = (v_i)_{i\in I_\lambda}$ (J2B). Wir können $I_\lambda \cap I_\mu = \emptyset$ für $\lambda\ne\mu$ annehmen und setzen $I = \bigsqcup_{\lambda\in K} I_\lambda$. Dank $V = \bigoplus_{\lambda\in K} \mathrm{E}(\lambda)$ erhalten wir eine Basis $\mathcal{B} = (v_i)_{i\in I}$ zu V (J2O). Diese diagonalisiert $f:V\to V$, denn es gilt $f(v_i) = \lambda v_i$ für $i\in I_\lambda$.

Eigenraumzerlegung und Diagonalisierung

„⇒“: Sei $\mathcal{B} = (v_i)_{i\in I}$ eine diagonalisierende Basis zu f, das heißt $f(v_i) = \lambda_i v_i$ mit $\lambda_i\in K$ für alle $i\in I$. Wir zerlegen die Indexmenge $I = \bigsqcup_{\lambda\in K} I_\lambda$ in $I_\lambda = \{\, i\in I \mid \lambda_i = \lambda \,\}$. Dann gilt $\mathrm{E}(\lambda) = \langle\, v_i \mid i\in I_\lambda \,\rangle_K^{!}$.

(a) Die Inklusion „⊇“ ist klar: Für jede Linearkombination $v = \sum_{i\in I_\lambda} \mu_i v_i$ mit Koeffizienten $\mu_i\in K$ gilt $f(v) = \sum_{i\in I_\lambda} \mu_i f(v_i) = \sum_{i\in I_\lambda} \mu_i \lambda v_i = \lambda v$.

(b) Zum Beweis der Umkehrung „⊆“ sei $v\in\mathrm{E}(\lambda)\le V$. Wir haben also $v = \sum_{i\in I} \mu_i v_i$ mit $\mu\in K^{(I)}$ und $0 = f(v) - \lambda v = \sum_{i\in I} (\mu_i\lambda_i - \lambda\mu_i) v_i$. Da $\mathcal{B}$ linear unabhängig ist, folgt $(\lambda_i - \lambda)\mu_i = 0$ für jeden Index $i\in I$. Daraus folgt $\lambda_i = \lambda$ oder $\mu_i = 0$ dank I1D, also $v \in \langle\, v_i \mid i\in I_\lambda \,\rangle_K^{!}$.

Wir erhalten somit $V = \bigoplus_{\lambda\in K} \langle\, v_i \mid i\in I_\lambda \,\rangle_K^{!} = \bigoplus_{\lambda\in K} \mathrm{E}(\lambda)$.
Das beweist die Äquivalenz (2) der Eigenraumzerlegung M1I.

Beispiel: Von der Eigenraumzerlegung zur Matrix

Aufgabe: (1) Sei $v_1 = \begin{bmatrix}2\\1\end{bmatrix}$ und $v_2 = \begin{bmatrix}-1\\1\end{bmatrix}$. Finden Sie alle $A \in \mathbb{R}^{2\times 2}$ mit

$$\operatorname{Eig}(A,-1) = \langle\, v_1 \,\rangle \quad \text{und} \quad \operatorname{Eig}(A,2) = \langle\, v_2 \,\rangle.$$

(2) Hat A noch weitere Eigenwerte? (3) Bestimmen Sie $\sigma(A)$.

Lösung: (1) Wir wissen $Av_1 = -v_1$ und $Av_2 = 2v_2$. Dank dem Prinzip der linearen Fortsetzung (K1B) gibt es genau eine solche Matrix A. Da $\mathcal{B} = (v_1, v_2)$ eine Basis von $\mathbb{R}^2$ ist, wird A durch $\mathcal{B}$ diagonalisiert:

$$T^{-1}AT = D = \begin{bmatrix}-1 & 0\\ 0 & 2\end{bmatrix} \in \mathbb{R}^{2\times 2} \text{ mit}$$

$$T = \mathrm{T}^{\mathcal{E}}_{\mathcal{B}} = \begin{bmatrix}2 & -1\\ 1 & 1\end{bmatrix} \quad \text{und} \quad T^{-1} = \mathrm{T}^{\mathcal{B}}_{\mathcal{E}} = \frac{1}{3}\begin{bmatrix}1 & 1\\ -1 & 2\end{bmatrix}, \quad \text{also}$$

$$A = TDT^{-1} = \frac{1}{3}\begin{bmatrix}2 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 2\end{bmatrix} = \begin{bmatrix}0 & -2\\ -1 & 1\end{bmatrix}.$$

(2) Wir haben $\mathbb{R}^2 = \mathrm{E}(-1) \oplus \mathrm{E}(2)$. Für $\lambda \in \mathbb{R} \smallsetminus \{-1, 2\}$ folgt dank M1I:

$$\mathrm{E}(\lambda) = \mathrm{E}(\lambda) \cap \mathbb{R}^2 = \mathrm{E}(\lambda) \cap \big(\mathrm{E}(-1) \oplus \mathrm{E}(2)\big) = \{0\}$$

(3) Das Spektrum der Matrix A ist demnach $\sigma(A) = \{-1, 2\}$.

Beispiel: Von der Eigenraumzerlegung zur Matrix

☺ In (2) nutzen wir geschickt die Eigenraumzerlegung aus Satz M1I. Für weitere nicht-triviale Eigenräume ist daher in $\mathbb{R}^2$ kein Platz mehr.

Das Eigenwertspektrum der Matrix A ist demnach $\sigma(A) = \{-1, 2\}$: Es gilt „$\supseteq$“ nach Konstruktion, und „$\subseteq$“ dank Satz M1I.

☹ Alternativ (aber umständlich) können Sie dies für jeden weiteren Kandidaten $\lambda \in \mathbb{R} \smallsetminus \{-1, 2\}$ einzeln explizit ausrechnen, indem Sie jeweils $\ker(A - \lambda I) = \{0\}$ bestimmen. Das ist möglich, aber mühselig.

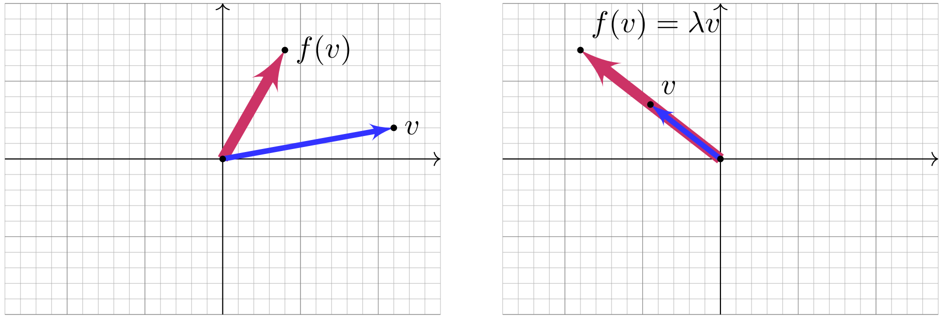
Zur Übung können Sie dies gerne ausprobieren und selbst spüren.

☺ Am besten sparen Sie sich unnötige Arbeit mit der zugehörigen Theorie, hier der Eigenraumzerlegung aus Satz M1I. Das kostet anfangs eine gewisse Investition, doch zahlt sich rasch aus.

Mit den passenden Werkzeugen arbeiten Sie effizienter.

Geometrische Bedeutung

Sei $f : V \to V$ linear über K. Was bedeuten EV und EW geometrisch?



Wir vergleichen einen Vektor $v \in V \smallsetminus \{0\}$ mit seinem Bild $w = f(v) \in V$. Im Allgemeinen besteht keine Relation, beide sind linear unabhängig.

Im Fall $w = \lambda v$ mit $\lambda \in K$ sind sie linear abhängig, also parallel. Der zugehörige Eigenwert λ ist der Streckfaktor von v zu w.

Algebraische Bedeutung

Für jeden Eigenvektor v mit $f(v) = \lambda v$ gilt

$$\boxed{f^n(v) = \lambda^n v.}$$

Dies folgt sofort per Induktion über $n \in \mathbb{N}$:

$$f^n(v) = f(f^{n-1}(v)) = f(\lambda^{n-1}v) = \lambda^{n-1}f(v) = \lambda^{n-1}\lambda v = \lambda^n v$$

Für jedes Polynom $P = \sum_{i=0}^{n} a_i X^i \in K[X]$ gilt demnach

$$\boxed{P(f)(v) = P(\lambda)\, v.}$$

Wir nutzen hier den Einsetzungshomomorphismus G3E:

$$\begin{aligned} P(f)(v) &= (a_0 + a_1 f + \dots + a_n f^n)(v) \\ &= a_0 v + a_1 f(v) + \dots + a_n f^n(v) \\ &= a_0 v + a_1 \lambda v + \dots + a_n \lambda^n v \\ &= (a_0 v + a_1 \lambda + \dots + a_n \lambda^n) v = P(\lambda)\, v \end{aligned}$$

Beispiel: Von der Matrix zur Eigenraumzerlegung

Ausprobieren mit Gaël!

Aufgabe: Vorgelegt sei die Matrix

$$A = \begin{bmatrix} 0 & 4 & 2 \\ -2 & 6 & 2 \\ 4 & -8 & -2 \end{bmatrix}.$$

(1) Bestimmen Sie die Eigenräume $\mathrm{E}(\lambda)$ für $\lambda = 0, 1, 2, 3$.
(2) Ist A diagonalisierbar? Falls ja, diagonalisieren Sie A.
(3) Hat A noch weitere Eigenwerte? Bestimmen Sie $\sigma(A)$.

Lösung: (1) Für $\lambda = 0$ bestimmen wir $\mathrm{E}(0) = \ker(A - 0I) = \ker(A)$:

$$A = \begin{bmatrix} 0 & 4 & 2 \\ -2 & 6 & 2 \\ 4 & -8 & -2 \end{bmatrix} \xrightarrow[\text{RZSF}]{\text{Gauß}} \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \implies \mathrm{E}(0) = \left\langle \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\rangle^{!}$$

Für $\lambda = 1$ bestimmen wir $\mathrm{E}(1) = \ker(A - 1I)$ und finden:

$$A - 1I = \begin{bmatrix} -1 & 4 & 2 \\ -2 & 5 & 2 \\ 4 & -8 & -3 \end{bmatrix} \xrightarrow[\text{RZSF}]{\text{Gauß}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \implies \mathrm{E}(1) = \{0\}$$

Beispiel: Von der Matrix zur Eigenraumzerlegung

Ausprobieren mit Gaël!

Für $\lambda = 2$ bestimmen wir $\mathrm{E}(2) = \ker(A - 2I)$ und finden:

$$\begin{bmatrix} -2 & 4 & 2 \\ -2 & 4 & 2 \\ 4 & -8 & -4 \end{bmatrix} \xrightarrow[\text{RZSF}]{\text{Gauß}} \begin{bmatrix} 1 & -2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \mathrm{E}(2) = \left\langle \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle^{!}$$

Für $\lambda = 3$ wissen wir schon $\mathrm{E}(3) = \ker(A - 3I) = \{0\}$ dank M1I.

(2) Die Matrix A ist diagonalisierbar, etwa wie oben mit der Eigenbasis

$$\mathcal{B} = \left(\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right).$$

Daraus erhalten wir den zugehörigen Basiswechsel zur Diagonalmatrix:

$$T = \mathrm{T}^{\mathcal{E}}_{\mathcal{B}} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \implies T^{-1}AT = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Beispiel: Von der Matrix zur Eigenraumzerlegung

Erläuterung

☺ In diesem Beispiel hat der erste Eigenwert $\lambda = 0$ die algebraische Vielfachheit $\dim_{\mathbb{R}} \mathrm{E}(0) = 1$. Hingegen hat der zweite Eigenwert $\lambda = 2$ die algebraische Vielfachheit $\dim_{\mathbb{R}} \mathrm{E}(2) = 2$. Auch das kommt vor.

(3) Wir haben $\mathbb{R}^3 = \mathrm{E}(0) \oplus \mathrm{E}(2)$. Für $\lambda \in \mathbb{R} \smallsetminus \{0, 2\}$ folgt dank M1I:

$$\mathrm{E}(\lambda) = \mathrm{E}(\lambda) \cap \mathbb{R}^3 = \mathrm{E}(\lambda) \cap \big(\mathrm{E}(0) \oplus \mathrm{E}(2)\big) = \{0\}$$

☺ Für weitere nicht-triviale Eigenräume ist daher in $\mathbb{R}^3$ kein Platz. Das Eigenwertspektrum der Matrix A ist demnach $\sigma(A) = \{0, 2\}$. Es gilt „$\supseteq$“ nach Rechnung, und „$\subseteq$“ dank Satz M1I.

☺ Alternativ können Sie $\lambda = 3$ und weitere Beispiele explizit ausrechnen. Am besten sparen Sie sich unnötige Arbeit mit der passenden Theorie, hier der Eigenraumzerlegung aus Satz M1I.

Diese Aufgabe hat als mögliche Eigenwerte die Kandidaten $\lambda = 0, 1, 2, 3$ vorgegeben. Diese enge Fragestellung dient hier als didaktischer Kniff, um Ihre Aufmerksamkeit auf die grundlegenden Definitionen M1A und M1B sowie die omnipräsente Eigenraumzerlegung M1F zu fokussieren.

Beispiel: Von der Matrix zur Eigenraumzerlegung

Erläuterung

Die Vorgabe von Eigenwerten kann manchmal durchaus sinnvoll sein, doch in den allermeisten Anwendungen ist sie eher unrealistisch. Oft haben Sie keine solche Anhaltspunkte oder Vorgaben. Dann müssen die möglichen Eigenwerte selbst finden.

Für das allgemeine Problem benötigen Sie daher weitere Werkzeuge. Dies gelingt uns im Folgenden mit dem charakteristischen Polynom.

⚠ Definition M1B und Satz M1F gelten unabhängig von der Dimension, egal ob endlich oder unendlich. Da wir nun Matrizen und Determinanten nutzen wollen, werden wir uns auf endliche Dimension konzentrieren. Das ist eine Weggabelung: Weniger Allgemeinheit, stärkere Werkzeuge.

Ausblick: Die Funktionalanalysis zweigt hier in die andere Richtung ab und betrachtet unendlich-dimensionale Vektorräume mit ihren eigenen, raffinierten Werkzeugen. Einfache Beispiele können wir jetzt bereits bewundern, in Ihrem Studium dürfen Sie sich auf noch viel mehr freuen.

Was bedeutet „Diagonalisieren Sie A“?

🙂 Diese Standardaufgabe ist typisch für Übungen und Klausuren. Dabei stellt sich jeweils die Frage, was genau gefordert ist: Welche Daten gehören zu einer vollständigen Antwort?

⚠ Je nach Aufgabe bzw. Anwendung sind verschiedene Stufen der Ausführung denkbar. Beginnen wir mit der maximalen Ausbaustufe:

Gegeben ist eine Matrix $A \in K^{n\times n}$ über dem Körper K.

Gesucht ist im Sinne einer vollständigen Analyse:

1. die Menge $\sigma(A) = \sigma(A; K)$ aller Eigenwerte,
2. die Dimension jedes Eigenraums $\operatorname{Eig}(A; \lambda)$ für M1ı,
3. eine Basis $\mathcal{B}_\lambda = (v_i)_{i\in I_\lambda}$ für jeden Eigenraum $\operatorname{Eig}(A; \lambda)$,
4. die daraus gebildete Basiswechselmatrix $T = (v_1, \dots, v_n) \in \mathrm{GL}_n K$,
5. die hierzu inverse Matrix T^{-1}, mit der Eigenschaft / Probe $T^{-1}T = I$,
6. die so gewonnene Diagonalmatrix $D = T^{-1}AT = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, mit der Eigenschaft / Probe $D = T^{-1}AT$ bzw. $AT = TD$.

Verschiedene Härtegrade

Neben der vollständigen Analyse gibt es mehrere kleinere **Varianten**:

Ist nur die **Diagonalisierbarkeit** von A gefragt, so genügen (1) und (2). Daran lässt sich die Diagonalisierbarkeit entscheiden und (6) D ablesen. Die fehlenden Daten (3,4,5) lassen sich anschließend ergänzen…

Ist ein **diagonalisierender Basiswechsel** gefragt, so genügen (4) T und (5) T^{-1} und (6) D. Damit lassen sich die Eigenschaften $T^{-1}T = I$ und $T^{-1}AT = D$ prüfen und die Daten (1,2,3) ablesen.

Ist nur eine **Eigenbasis** zu A gefragt, so genügen (4) T und (6) D. Damit lässt sich $AT = TD$ prüfen und die Daten (1,2,3) ablesen.

In Klausuren geben wir die Fragen kleinschrittig vor und sagen genau, was jeweils gefragt ist. Erfahrungsgemäß bereitet das keine Probleme.

Bei einer mündlichen Präsentation, etwa in Ihrer Gruppenübung, haben Sie mehr Freiheiten, Sie können nachfragen und ergänzen. In solchen Fällen ist eine offene Fragestellung meist sinnvoller.

Ergebnis und Zertifikat

Die algorithmische Sichtweise ist oft hilfreich. Hier geht es speziell um die Frage der Spezifikation: Was ist die Eingabe? Was ist die Ausgabe? Welche Eigenschaften müssen garantiert bzw. überprüft werden?

Spezifikation: Eigenbasis einer Matrix

Eingabe: eine Matrix $A \in K^{n\times n}$ über dem Körper K
Ausgabe: $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ und $T \in \mathrm{GL}_n(K)$ mit $AT = TD$
oder notfalls die Antwort „A ist nicht diagonalisierbar“

Allgemein gilt: Je mehr Daten die Spezifikation als Ausgabe verlangt, desto aufwändiger die Berechnung, doch umso leichter ist die Prüfung.

Umgekehrt: Wird von der Eingabe mehr verlangt, so ist dies schwerer für den Auftraggeber, doch umso leichter für den Auftragnehmer.

Spezifikation: Eigenbasis einer diagonalisierbaren Matrix

Eingabe: eine diagonalisierbare Matrix $A \in K^{n\times n}$ über K
Ausgabe: $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ und $T \in \mathrm{GL}_n(K)$ mit $AT = TD$

Was bedeutet „Diagonalisieren Sie f“?

Welche Daten gehören zu einer vollständigen Antwort?

⚠ Je nach Aufgabe bzw. Anwendung sind verschiedene Stufen der Ausführung denkbar. Beginnen wir mit der maximalen Ausbaustufe:

Gegeben ist $f \in \operatorname{End}_K(V)$ und $n = \dim_K(V) < \infty$.

Gesucht ist im Sinne einer vollständigen Analyse:

1. die Menge $\sigma(f) = \sigma(f; K)$ aller Eigenwerte,
2. die Dimension jedes Eigenraums $\operatorname{Eig}(f; \lambda)$ für M1ı,
3. eine Basis $\mathcal{B}_\lambda = (v_i)_{i\in I_\lambda}$ für jeden Eigenraum $\operatorname{Eig}(f; \lambda)$,
4. die durch $I = \bigsqcup_\lambda I_\lambda$ gebildete Eigenbasis $\mathcal{B} = (v_i)_{i\in I}$ von V,
5. die so gewonnene Diagonalmatrix $D = \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) \in K^{n\times n}$.

Zur **Diagonalisierbarkeit** von f genügen (1) und (2). Daran lässt sich (5) ablesen und die fehlenden Daten (3,4) anschließend ergänzen…

Für eine **Eigenbasis** zu f genügen (4) und (5), daraus folgen (1,2,3). Selbst im Falle $\dim_K(V) = \infty$ bleiben (1,3,4) ebenso sinnvoll.

Wie finden wir alle Eigenwerte?

Wie finden wir alle Eigenwerte einer vorgelegten Matrix $A \in K^{n\times n}$?

Satz M2A: Eigenwerte und Determinante

Sei K ein Körper, $A \in K^{n\times n}$ eine Matrix und $\lambda \in K$ ein Skalar. Genau dann ist λ ein Eigenwert von A, wenn $\det(A - \lambda I) = 0$ gilt.

Beweis: Wir setzen die Definition ein und formen sorgsam um:

$$
\begin{aligned}
& && \operatorname{Eig}(A,\lambda) \neq \{0\} \\
\underset{\text{M1B}}{\overset{\text{Def}}{\iff}} \quad & \exists v \in K^n,\ v \neq 0: && Av = \lambda v \\
\underset{\text{B1A}}{\overset{\text{Lin}}{\iff}} \quad & \exists v \in K^n,\ v \neq 0: && (A - \lambda I)v = 0 \\
\underset{\text{I1R}}{\overset{\text{Def}}{\iff}} \quad & \exists v \in K^n,\ v \neq 0: && v \in \ker(A - \lambda I) \\
\underset{\text{I1R}}{\overset{\text{Def}}{\iff}} \quad & && \ker(A - \lambda I) \neq \{0\} \\
\underset{\text{L2Q}}{\overset{\text{Det}}{\iff}} \quad & && \det(A - \lambda I) = 0
\end{aligned}
$$

Wie finden wir alle Eigenwerte?

🙂 Die Eigenwerte der Matrix $A \in K^{n\times n}$ sind demnach die Nullstellen der **charakteristischen Funktion** $K \to K : \lambda \mapsto \det(A - \lambda I)$.

Sie ordnet jedem Skalar $\lambda \in K$ die Determinante $\det(A - \lambda I) \in K$ zu und zeigt an, ob der Kern von $A - \lambda I$ trivial ist oder nicht (Satz L2Q).

Diese Funktion wollen wir nun genauer untersuchen! Insbesondere werden wir sehen, dass dies eine Polynomfunktion ist. Somit können wir all unsere Werkzeuge zu Polynomen hier nutzbringend anwenden!

⚠ In Definition M1B haben wir zunächst erklärt, was Eigenwerte sind. Hier nun lernen Sie eine Rechenmethode mit der Determinante kennen. Um Matrizen und Determinanten überhaupt einsetzen zu können, werden wir uns auf endliche Dimension konzentrieren.

Bitte unterscheiden Sie Definition und Methoden, also Ziel und Wege! Eigenwerte und Eigenräume lassen sich auch in vielen Situationen erklären und nutzen, in denen die Determinante nicht anwendbar ist, insbesondere auch für unendlich-dimensionale Vektorräume.

Wie finden wir alle Eigenwerte?

Wie finden wir alle Eigenwerte eines Endomorphismus $f : V \to V$?

$$\begin{tikzcd}
K^n \arrow[r, "A : x \mapsto Ax"] \arrow[d, "\Phi_{\mathcal{B}}"', "\cong"] & K^n \arrow[d, "\Phi_{\mathcal{B}}"', "\cong"] \\
V \arrow[r, "f : v \mapsto f(v)"] & V
\end{tikzcd}$$

Im Falle $\dim_K V = n < \infty$ wählen wir eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V und stellen f dar durch die zugehörige Matrix $A = \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) \in K^{n\times n}$.

Jeden Vektor $v \in V$ stellen wir eindeutig dar durch seine Koordinaten $x \in K^n$ mit $v = x_1 v_1 + \dots + x_n v_n$. Aus obigem Diagramm lesen wir ab:

Bemerkung: Genau dann gilt $f(v) = \lambda v$, wenn $Ax = \lambda x$.

⚠ Die entscheidende Voraussetzung ist hier $\dim_K V = n < \infty$. In diesem Falle können wir Matrizen und die Determinante nutzen!

Auf unendlichen Vektorräumen ist die Sachlage schwieriger… und weitaus interessanter (siehe Funktionalanalysis).

Wie finden wir alle Eigenwerte?

Damit wird das Problem berechenbar dank der Determinante und der charakteristischen Funktion von f bzw. A:

$$
\begin{aligned}
& \operatorname{Eig}(f,\lambda) \neq \{0\} \\
\iff\quad & \operatorname{Eig}(A,\lambda) \neq \{0\} \\
\iff\quad & \det(A - \lambda I) = 0 \\
\iff\quad & \det(f - \lambda \operatorname{id}_V) = 0
\end{aligned}
$$

Zur Determinante eines Endomorphismus $f : V \to V$ siehe L3D: Dies gelingt in einer Basis durch Darstellung als Matrix $A \in K^{n\times n}$.

Meist interessiert uns eigentlich die lineare Abbildung $f : V \to V$, doch mit der Matrix $A \in K^{n\times n}$ können wir besonders gut rechnen.

Notation: Hier ist $I = E = 1_{n\times n} = \operatorname{diag}(1, \dots, 1)$ die Einheitsmatrix. Da E schon für Eigenräume genutzt wird, schreibe ich hier lieber I.

Für $A - \lambda I$ schreiben wir kurz $A - \lambda$, und für $f - \lambda \operatorname{id}_V$ ebenso $f - \lambda$. Diese Kurzschreibweise ist etwas nachlässig, aber manchmal bequem.

Das charakteristische Polynom

Definition M2B: das charakteristische Polynom einer Matrix

Zur Matrix $A \in K^{n\times n}$ definieren wir das **charakteristische Polynom**

$$\tilde{P}_A(X) = \tilde{\chi}_A(X) := \det(A - XI) \in K[X].$$

Alternativ nutzen wir das **normierte charakteristische Polynom**

$$P_A(X) = \chi_A(X) := \det(XI - A) = (-1)^n\,\tilde{\chi}_A(X).$$

Dies ist tatsächlich ein Polynom, wie wir in Satz M2C nachrechnen.

⚠ Beide Konventionen sind in der Literatur üblich, aber uneinheitlich. Das normierte Polynom hat immer Leitkoeffizient 1, das ist schöner.

Die unnormierte Variante ist in Rechnungen oft bequemer, da die Matrix A bereits vorliegt, und nur auf der Diagonalen jeweils X subtrahiert wird: Das ist per Hand leichter zu schreiben und vermeidet Vorzeichenfehler.

☺ Beide Varianten unterscheiden sich nur in ungerader Dimension. Für die Nullstellen (also Eigenwerte) spielt das Vorzeichen keine Rolle. Auch für den Eigenraum gilt $\mathrm{Eig}(A,\lambda) = \ker(A - \lambda I) = \ker(\lambda I - A)$.

Das charakteristische Polynom

Zur Illustration betrachten wir ein einfaches Zahlenbeispiel:

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

Der Übergang von A zu $A - XI$ ist etwas leichter auszuschreiben:

$$A - XI = \begin{bmatrix} 1-X & 4 & 5 \\ 0 & 0-X & 1 \\ 0 & 2 & 3-X \end{bmatrix}$$

Der Übergang von A zu $XI - A$ erfordert etwas mehr Änderungen:

$$XI - A = \begin{bmatrix} X-1 & -4 & -5 \\ -0 & X-0 & -1 \\ -0 & -2 & X-3 \end{bmatrix}$$

Für einen Computer ist der Unterschied unerheblich. Menschen jedoch machen (Schreib-)Fehler, daher biete ich Ihnen hier beide Varianten an.

Anwendungsbeispiel zum charakteristischen Polynom

Aufgabe: (1) Berechnen Sie das charakteristische Polynom der Matrix

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 0 & 1 \\ 0 & 2 & 3 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

(2) Finden Sie damit alle Eigenwerte, also das Spektrum $\sigma(A;\mathbb{R})$.

(3) Ist A über $\mathbb{R}$ diagonalisierbar? (4) Falls ja, diagonalisieren Sie A.

Lösung: (1) Wir folgen der Definition und berechnen die Determinante:

$$\begin{aligned} \tilde{\chi}_A(X) = \begin{vmatrix} 1-X & 4 & 5 \\ 0 & -X & 1 \\ 0 & 2 & 3-X \end{vmatrix} &= (1-X)\big[-X(3-X) - 1\cdot 2\big] \\ &= (1-X)(X^2 - 3X - 2) \\ &= -X^3 + 4X^2 - X - 2 \end{aligned}$$

(2) Die Eigenwerte von A sind genau die Nullstellen von $\chi_A = -\tilde{\chi}_A$:

$$\sigma(A) = \big\{\, \lambda_1 = 1,\ \lambda_2 = \tfrac{1}{2}(3 - \sqrt{17}),\ \lambda_3 = \tfrac{1}{2}(3 + \sqrt{17}) \,\big\}$$

Anwendungsbeispiel zum charakteristischen Polynom

Summe von Monomen vs **Produkt von Linearfaktoren**

$$X^3 - 4X^2 + X + 2 \ \longleftarrow\ (X-1)(X^2-3X-2) \ \longrightarrow\ (X-1)(X-\lambda_2)(X-\lambda_3)$$

Das Ausmultiplizieren (hier nach links) ist eine leichte Routineübung. Das Faktorisieren (nach rechts) gelingt routiniert in kleinen Graden, ist im Allgemeinen jedoch eine schwieriges Problem.

In Frage (1) wollen wir das charakteristische Polynom bestimmen. Polynome können wir jedoch verschieden darstellen; implizit gemeint ist meist die Monomform, dazu multiplizieren wir alles geduldig aus.

Für die Frage (2) hingegen ist die Monomform nicht die beste Wahl! Wir suchen hier Nullstellen unseres Polynoms, also Linearfaktoren. Dazu geht das Ausmultiplizieren genau in die falsche Richtung; es ist schlauer, die bereits vorliegende partielle Faktorisierung zu nutzen.

⚠ In Klausuren ist das ein häufiger Fehler! Unnötige Umwege kosten Zeit und erhöhen die Wahrscheinlichkeit von Rechenfehlern. Sie sollten genau wissen, was Sie wollen und wie Sie es am besten erreichen.

Grad und Koeffizienten des char. Polynoms

Satz M2C: Grad und Koeffizienten des char. Polynoms

Zu $A \in K^{n\times n}$ ist $\tilde\chi_A(X) = \det(A - XI)$ ein Polynom von Grad n:

$$\tilde\chi_A(X) = a_0 - a_1 X \pm \dots + (-1)^{n-1} a_{n-1} X^{n-1} + (-1)^n X^n$$

Normiert erhalten wir für $\chi_A(X) = \det(XI - A)$ demnach das Polynom:

$$\chi_A(X) = X^n - a_{n-1} X^{n-1} \pm \dots + (-1)^{n-1} a_1 + (-1)^n a_0$$

Die extremen Koeffizienten erkennen wir wieder: $a_0 = \det(A)$ ist die **Determinante** und $a_{n-1} = \sum_{i=1}^n a_{ii} =: \operatorname{tr}(A)$ ist die **Spur** der Matrix A.

Beweis: Jeder Eintrag der Matrix $C = XI - A \in K[X]^{n\times n}$ hat Grad ≤ 1. Dank der Leibniz–Formel (L2G) hat $\chi_A = \det C$ den Grad $\le n$ (G3B):

$$\chi_A = \det C = \sum\nolimits_{\tau\in S_n} \operatorname{sign}(\tau)\cdot c_{\tau(1),1}\cdot c_{\tau(2),2}\cdots c_{\tau(n),n}$$

Der Summand $(X - a_{11})(X - a_{22})\cdots(X - a_{nn}) = X^n - X^{n-1}\operatorname{tr} A + \dots$ für $\tau = \mathrm{id}$ hat Grad n. Alle weiteren Summanden haben Grad $\le n-2$, denn jede Permutation $\tau \ne \mathrm{id}$ hat höchstens $n-2$ Fixpunkte. $\boxed{\text{QED}}$

Grad und Koeffizienten des char. Polynoms

Die Koeffizienten $a_0 = \det A$ und $a_{n-1} = \operatorname{tr} A$ haben eine besondere, klare Bedeutung, die anderen haben keine so leichte Interpretation.

⚠ Die charakteristische Matrix $XI - A \in K[X]^{n\times n}$ hat Koeffizienten im Polynomring $K[X]$. Wir rechnen also nicht mehr über dem Grundring K, sondern im Ring $K[X]$ der Polynome in der Variablen X über K.

☺ Dazu ist alles vorbereitet, denn wir haben die Determinante wohlweislich gleich allgemein über kommutativen Ringen erklärt.

⚠ Manche Autoren arbeiten zunächst nur über einem Körper K, insbesondere für Matrizen und Determinanten, doch spätestens beim charakteristischen Polynom genügt dies genau genommen nicht mehr.

☺ Man kann sich mit verschiedenen Tricks aus dieser misslichen Lage befreien, zum Beispiel können wir den Polynomring $K[X]$ einbetten in seinen Bruchkörper $K(X)$ der gebrochen-rationalen Funktionen (siehe Seite A139 und Satz E3L) und über diesem größeren Körper arbeiten.

Das Natürlichste scheint mir jedoch die allgemeine Betrachtung über kommutativen Ringen. Genau dafür haben wir alles sorgsam vorbereitet.

Konjugation und Ähnlichkeit von Matrizen

Wie ändert sich das charakteristische Polynom bei Basiswechsel?

Definition M2D: Konjugation und Ähnlichkeit von Matrizen

Sei K ein Ring und $n \in \mathbb{N}_{\ge 1}$. Die allgemeine lineare Gruppe $\mathrm{GL}_n(K)$ operiert auf dem Matrizenring $K^{n\times n}$ durch Konjugation gemäß

$$K^{n\times n} \times \mathrm{GL}_n(K) \to K^{n\times n} \;:\; (A, T) \mapsto B = T^{-1} A T.$$

Zwei Matrizen $A, B \in K^{n\times n}$ heißen **ähnlich**, oder $\mathrm{GL}_n(K)$–konjugiert, falls es eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$ mit $B = T^{-1} A T$ gibt.

Der Übergang von A zu B heißt auch **Ähnlichkeitstransformation** oder **Konjugation** vermöge T. Dies entspricht einem Basiswechsel.

Übung: Dies ist eine Äquivalenzrelation auf der Menge $K^{n\times n}$.

☺ Diese Äquivalenzrelation wollen wir nutzen, um Matrizen soweit wie möglich zu vereinfachen: Wir wollen die vorgegebene Matrix A durch eine ähnliche, aber schönere Matrix B ersetzen. Im aktuellen Kontext geht es uns zunächst darum, dass B diagonal wird, falls möglich.

Konjugation und Ähnlichkeit von Matrizen

Lemma M2E: Ähnlichkeit und Basiswechsel

Seien $A, B \in K^{n\times n}$. Die folgenden Aussagen sind gleichbedeutend:

1. Die Matrizen A und B sind ähnlich, also $\mathrm{GL}_n(K)$–konjugiert.
2. Beide sind darstellende Matrizen eines Endomorphismus $f \in \mathrm{End}_K(V)$ bezüglich geeigneter Basis $\mathcal{A}$ und $\mathcal{B}$ von V.

$$\begin{tikzcd}
K^n \arrow[rrr, "A=\mathrm{M}^{\mathcal{A}}_{\mathcal{A}}(f)"'] \arrow[dr, "\Phi_{\mathcal{A}}", "\cong"'] \arrow[dd, shift right, "\mathrm{T}^{\mathcal{B}}_{\mathcal{A}}"'] & & & K^n \arrow[dl, "\Phi_{\mathcal{A}}"', "\cong"] \arrow[dd, shift right, "\mathrm{T}^{\mathcal{B}}_{\mathcal{A}}"'] \\
& V \arrow[r, "f"] & V & \\
K^n \arrow[uu, shift right, "\mathrm{T}^{\mathcal{A}}_{\mathcal{B}}"', "\cong"] \arrow[ur, "\Phi_{\mathcal{B}}"', "\cong"] \arrow[rrr, "B=\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f)"] & & & K^n \arrow[ul, "\Phi_{\mathcal{B}}", "\cong"'] \arrow[uu, shift right, "\mathrm{T}^{\mathcal{A}}_{\mathcal{B}}"', "\cong"]
\end{tikzcd}$$

In $K^{n\times n}$ haben wir $B = T^{-1}\cdot A \cdot T$ mit $T = \mathrm{T}^{\mathcal{A}}_{\mathcal{B}}$ und $T^{-1} = \mathrm{T}^{\mathcal{B}}_{\mathcal{A}}$.

Das charakteristische Polynom eines Endomorphismus

Satz M2F: Invarianz und Wohldefiniertheit des char. Polynoms

Sei K ein Körper.

1. Sind A und B in $K^{n\times n}$ ähnlich, so folgt $\chi_A = \chi_B$ in $K[X]$.
2. Zu $f \in \mathrm{End}_K(V)$ wie oben ist $\chi_f := \chi_A = \chi_B$ wohldefiniert.

Beweis: (1) Wir haben $B = T^{-1}AT$ mit $T \in \mathrm{GL}_n(K)$. Daraus folgt:

$$\begin{aligned}
\chi_B(X) &\overset{\text{Def}}{=} \det(XI - B) \\
&\overset{\text{Vor}}{=} \det(XI - T^{-1}AT) \\
&\overset{\text{Com}}{=} \det\big[T^{-1}(XI - A)T\big] \\
&\overset{\text{L2N}}{=} \det(T^{-1}) \cdot \det(XI - A) \cdot \det(T) \\
&\overset{\text{Def}}{=} \chi_A(X)
\end{aligned}$$

(2) Für $A = \mathrm{M}^{\mathcal{A}}_{\mathcal{A}}(f)$ und $B = \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f)$ wie in M2E gilt $B = T^{-1}AT$. $\boxed{\text{QED}}$

☺ Die implizite Definition (2) formulieren wir nun explizit aus.

Das charakteristische Polynom eines Endomorphismus

Definition M2G: das char. Polynom eines Endomorphismus

Sei K ein Körper und V ein K–Vektorraum mit $\dim_K(V) = n < \infty$.

Zu jedem Endomorphismus $f : V \to V$ über K definieren wir sein **charakteristisches Polynom** $\chi_f \in K[X]$ wie folgt.

Wir wählen eine Basis $\mathcal{A}$ von V und stellen f dar durch die zugehörige Matrix $A = \mathrm{M}^{\mathcal{A}}_{\mathcal{A}}(f) \in K^{n\times n}$. Damit definieren wir das Polynom:

$$\boxed{\chi_f := \det(XI - A) \in K[X]}$$

Dank Satz M2F ist das Ergebnis χ_f wohldefiniert,
da unabhängig von der willkürlich gewählten Basis $\mathcal{A}$.

☺ Sie sehen hier sehr schön: Manchmal können wir eine Definition erst nach einem Satz aussprechen, der den Weg zum Begriff bereitet. Das geschieht immer dann, wenn die Wohldefiniertheit zu klären ist, bzw. Existenz und Eindeutigkeit des zu definierenden Gegenstands. So gesehen müsste ich auch Definition M2B und Satz M2C umdrehen.

Das char. Polynom ist keine vollständige Invariante.

Je zwei ähnliche Matrizen haben dasselbe charakteristische Polynom. Somit ist $K^{n\times n} \to K[X]^1_n : A \mapsto \chi_A$ eine Invariante unter Ähnlichkeit.

Sie ist jedoch noch keine vollständige Invariante. Es gibt Matrizen, die dasselbe charakteristische Polynom haben, aber nicht ähnlich sind.

Aufgabe: Sind die folgenden Matrizen ähnlich?

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \quad \text{vs} \quad B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

Lösung: (1) Die charakteristischen Polynome sind gleich:

$$\chi_A = \chi_B = (X - \lambda)^2$$

(2) Dennoch sind die Matrizen A und B nicht ähnlich:
A ist diagonal, doch B ist nicht diagonalisierbar (M1D).

Das char. Polynom ist keine vollständige Invariante.

☺ Aus $\chi_A \neq \chi_B$ schließen wir, dass A und B nicht ähnlich sind. Das ist der klare, einfache Fall und gilt allgemein für jede Invariante: Sind χ_A und χ_B verschieden, so können A und B nicht ähnlich sein.

⚠ Die Umkehrung gilt nicht: Aus $\chi_A = \chi_B$ können wir im Allgemeinen noch nicht schließen, dass A und B ähnlich sind. Hierzu sind weitere Untersuchungen notwendig. Genau das zeigt dieses Beispiel!

Bemerkung: Für $\lambda = 0$ ist dies besonders augenfällig:

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{vs} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Auch der Rang einer Matrix ist invariant unter Ähnlichkeit $A \mapsto T^{-1}AT$, noch allgemeiner sogar unter Äquivalenz $A \mapsto S^{-1}AT$, siehe Satz K2H.

Hier gilt $\operatorname{rang} A = 0$ und $\operatorname{rang} B = 1$, also sind A und B nicht ähnlich, nicht einmal äquivalent: $B \neq S^{-1}AT$ für alle $S, T \in \mathrm{GL}_2\,\mathbb{K}$.

Das charakteristische Polynom eines Produkts

⚠ Für das charakteristische Polynom gilt schon aus Gradgründen:

$$\chi_{A\cdot B} \neq \chi_A \cdot \chi_B$$

🙂 Die Determinante ist multiplikativ (Satz L2N)

$$\det(A \cdot B) = \det(A) \cdot \det(B),$$

☹ die Spur jedoch nicht:

$$\operatorname{tr}(A \cdot B) \neq \operatorname{tr}(A) \cdot \operatorname{tr}(B)$$

Ein besonders einfaches, wahrhaft minimalistisches Gegenbeispiel ist

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{mit} \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Übung: Um sich gegen solch naiven Irrglauben zu immunisieren, erfinden und berechnen Sie selbst geeignete Gegenbeispiele!

Das charakteristische Polynom eines Produkts

Für die Determinante $\det : K^{n\times n} \to K$ gilt dank Multiplikativität L2N:

$$\det(A \cdot B) = \det(A) \cdot \det(B) = \det(B) \cdot \det(A) = \det(A \cdot B)$$

für alle $A, B \in K^{n\times n}$. Die Spur $\operatorname{tr} : K^{n\times n} \to K$ ist nicht multiplikativ, doch dank Satz B1J ebenfalls invariant unter zyklischer Vertauschung:

$$\operatorname{tr}(A \cdot B) = \operatorname{tr}(B \cdot A).$$

Frage: Gilt dies sogar für das gesamte charakteristische Polynom?

$$\chi_{AB} \overset{?}{=} \chi_{BA}$$

Teilantwort: Ist eine der beiden Matrizen A oder B invertierbar, so folgt dies aus der Invarianz M2F unter Konjugation / Ähnlichkeit:

$$\chi_{AB} = \chi_{A^{-1}(AB)A} = \chi_{BA}$$

$$\chi_{AB} = \chi_{B(AB)B^{-1}} = \chi_{BA}$$

🙂 Satz M2H verallgemeinert dies auf nicht-invertierbare Matrizen.

Das charakteristische Polynom eines Produkts

Satz M2H: das charakteristische Polynom eines Produkts

(0) Für je zwei rechteckige Matrizen $A \in K^{m\times n}$ und $B \in K^{n\times m}$ gilt

$$\chi_{AB} \cdot X^n = \chi_{BA} \cdot X^m.$$

(1) Für je zwei quadratische Matrizen $A, B \in K^{n\times n}$ gilt

$$\chi_{AB} = \chi_{BA}.$$

Aus (0) folgt (1) dank $m = n$ und Kürzung des Faktors $X^m = X^n$. Umgekehrt folgt (0) aus (1) durch Auffüllen mit Nullzeilen/spalten.

Für Determinante und Spur ist die Aussage klar, wie zuvor erklärt. Dies sind beiden extremen Koeffizienten des Polynoms (Satz M2C).

Die Koeffizienten zwischen Spur und Determinante in $\chi_{AB} = \chi_{BA}$ haben keine so leichte Interpretation, doch auch sie sind invariant unter zyklischer Vertauschung der Matrixfaktoren von AB zu BA.

Das charakteristische Polynom eines Produkts

Beweis: (0) Wir berechnen zunächst folgende Matrixprodukte:

$$\begin{bmatrix} XI_m & -A \\ 0 & I_n \end{bmatrix} \cdot \begin{bmatrix} I_m & A \\ B & XI_n \end{bmatrix} \overset{(a)}{=} \begin{bmatrix} XI_m - AB & 0 \\ B & XI_n \end{bmatrix}$$

$$\begin{bmatrix} XI_m & 0 \\ -B & I_n \end{bmatrix} \cdot \begin{bmatrix} I_m & A \\ B & XI_n \end{bmatrix} \overset{(b)}{=} \begin{bmatrix} XI_m & XA \\ 0 & XI_n - BA \end{bmatrix}$$

Dank Multiplikativität der Determinante L2N und Satz L2V folgt:

$$\det(XI_m - AB)X^n \overset{(a)}{=} \begin{vmatrix} I_m & A \\ B & XI_n \end{vmatrix} X^m \overset{(b)}{=} \det(XI_n - BA)X^m$$

Das ist trickreich-raffiniert. Rechnen Sie es nach! $\boxed{\text{QED}}$

Das ist eine gefürchtete Übungsaufgabe (oder Klausuraufgabe?) aus Paul Halmos, *Finite dimensional vector spaces* (1958), §53, exercise 13.

🙂 Ohne den Trick oder einen Hinweis ist es schwer, mit ist es leicht. Liegen die Formeln erst einmal vor uns, so genügt Nachrechnen.

Das Standardverfahren zur Diagonalisierung

Verfahren M2I: Standardverfahren zur Diagonalisierung

Sei K ein Körper und V ein n–dimensionaler K–Vektorraum. Gegeben sei ein Endomorphismus $f : V \to V$ über K.

(1) Wähle eine Basis $\mathcal{A}$ von V und bestimme $A = \mathrm{M}_{\mathcal{A}}(f) \in K^{n \times n}$.

(2) Berechne das charakteristische Polynom $\chi_f = \det(XI - A) \in K[X]$.

(3) Bestimme alle Nullstellen, $\sigma(f) = \sigma(f; K) = \{\, \lambda \in K \mid \chi_f(\lambda) = 0 \,\}$.

(4) Zu $\lambda \in \sigma(f)$ bestimme den Eigenraum $\mathrm{Eig}(f, \lambda) = \ker(f - \lambda)$.

Genau dann ist f diagonalisierbar, wenn $V = \bigoplus_{\lambda \in \sigma(f)} \mathrm{Eig}(f, \lambda)$ gilt, also die Dimensionsgleichung $n = \sum_{\lambda \in \sigma(f)} \dim_K \mathrm{Eig}(f, \lambda)$ erfüllt ist.

(5) In diesem Falle wähle eine Basis $\mathcal{B}$ von V aus Eigenvektoren und erhalte die ersehnte Darstellung $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

☺ Schritte (1,2,4,5) sind Routineaufgaben der Linearen Algebra. Hierzu kennen Sie effiziente Algorithmen (mit Aufwand $\le \mathrm{const} \cdot n^4$).

☹ Allein die Suche nach Nullstellen (3) ist im Allgemeinen schwierig. Dieses Problem gehört zur (nicht-linearen!) Algebra bzw. Numerik.

Das Standardverfahren zur Diagonalisierung

Dieses Verfahren *können* Sie immer anwenden, insbesondere wenn keine weitere Information vorliegt. Zur Diagonalisierung *müssen* Sie nicht so vorgehen, manchmal sind auch geschickte Abkürzungen oder raffinierte Tricks möglich. Diese sollten Sie erkennen, um sich unnötige Arbeit und längliche Umwege zu ersparen, siehe Übungen (etwa M313).

Beispiel: Ist der Körper K klein doch die Matrix $A \in K^{n \times n}$ recht groß, etwa $\sharp K < n$, dann kann man Schritte (2) und (3) überspringen, und in (4) alle Kandidaten $\lambda \in K$ einsetzen und direkt $\ker(A - \lambda I)$ berechnen.

Das Standardverfahren verdeutlicht einige wichtige Eigenschaften: Alle Schritte (bis auf Punkt 3) sind Routineaufgaben der Linearen Algebra. Diese können von einem Computer übernommen werden. Der Rechenaufwand ist dabei polynomiell, höchstens $\mathrm{const} \cdot n^4$.

Kritisch ist und bleibt einzig der Punkt 3: die Bestimmung der Nullstellen. Wenn dieser gelingt, sei es durch günstige Umstände der gegebenen Daten oder durch zusätzliche Überlegungen zur vorliegenden Struktur, so ist das Diagonalisierungsproblem insgesamt gelöst.

Schnelle Berechnung des charakteristischen Polynoms

Wie schnell können wir das charakteristische Polynom berechnen?

Satz M2J: Berechnung des charakteristischen Polynoms

Sei K ein Körper mit Elementezahl $\sharp K \ge n$.

Zu jeder Matrix $A \in K^{n \times n}$ können wir das charakteristische Polynom $\chi_A \in K[X]^1_n$ berechnen mit $\le n^4$ arithmetischen Operationen in K.

Dies gelingt trickreich und genial-einfach durch Lagrange–Interpolation:

Algo M2J: Berechnung des charakteristischen Polynoms

```
Eingabe:  eine Matrix $A \in K^{n \times n}$ über $K$, wobei $\sharp K \ge n$
Ausgabe:  das charakteristische Polynom $\chi_A \in K[X]^1_n$
  Wähle $n$ verschiedene Stützstellen $x_1, \dots, x_n \in K$.
  Berechne die Werte $y_k = \det(x_k I - A)$ dank Gauß L2x.
  Rekonstruiere $\chi_A \in K[X]^1_n$ dank Lagrange–Interpolation B3A.
```

Übung: Führen Sie die Details aus und beweisen Sie den Satz.

Schnelle Berechnung des charakteristischen Polynoms

In Übungsaufgaben und Klausuren berechnen Sie charakteristische Polynome meist für kleine Matrizen, etwa $n = 2, 3, 4$. Es gelingt auch noch für große, strukturierte Matrizen, etwa Begleitmatrizen M2Q.

☺ Zur Behandlung von großen, unstrukturierten Matrizen nutzen Sie sinnvollerweise ein Computer-Algebra-System (CAS). Damit stellt sich die dringende Frage, auch für Sie als Nutzer/in, wie der Computer damit effizient umgehen soll, und welcher Rechenaufwand zu erwarten ist.

⚠ Die Leibniz–Formel L2G und die Laplace–Entwicklung L2Z erfordern schlimmstenfalls exponentiellen Aufwand, nämlich $n \cdot n!$ Operationen. Das überfordert selbst Supercomputer für moderate n, siehe L259.

☺ Die Berechnung mit n^4 Operationen ist zwar immer noch aufwändig für sehr große n, aber effizient genug für viele Anwendungen mittlerer Größe, genau so wie Gauß: Für typische Laufzeiten siehe B216.

☺ Der Algorithmus von Faddeev–LeVerrier erreicht dasselbe Ergebnis unter der Voraussetzung $\mathrm{char}\, K > n$. Für Details verweise ich auf `de.wikipedia.org/wiki/Algorithmus_von_Faddejew-Leverrier`.

Zerfällung von Polynomen in Linearfaktoren

Definition M2K: Zerfällung von Polynomen in Linearfaktoren

(0) Das Polynom $P \in K[X] \smallsetminus \{0\}$ **zerfällt** über K in Linearfaktoren, falls

$$P(X) = c(X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n) \quad \text{mit } \lambda_1, \lambda_2, \dots, \lambda_n \in K.$$

Wir sagen hierzu kurz, P zerfällt über K oder P ist K–zerfallend.

(1) Das Polynom P **zerfällt einfach** oder ist **einfach zerfallend**, falls alle Nullstellen einfach sind, also $\lambda_i \neq \lambda_j$ für alle $i \neq j$ gilt.

(2) Durch Zusammenfassen mehrfacher Nullstellen erhalten wir $P(X) = c(X-\mu_1)^{r_1}\cdots(X-\mu_k)^{r_k}$ mit Nullstellen $\mu_1, \dots, \mu_k \in K$ und Vielfachheiten $r_1, \dots, r_k \in \mathbb{N}_{\geq 1}$, wobei $\mu_i \neq \mu_j$ für $i \neq j$ gilt.

Der Exponent $r_i =: \operatorname{ord}(P, \mu_i)$ heißt (Nullstellen-)**Ordnung** oder (algebraische) **Vielfachheit** der Nullstelle μ_i im Polynom P.

Beispiele: $P = X^2 - 1 \in \mathbb{Q}[X]$ zerfällt gemäß $P = (X-1)(X+1)$.
$X^2 - 2 \in \mathbb{Q}[X]$ zerfällt nicht über $\mathbb{Q}$, aber über $\mathbb{R}$ in $(X-\sqrt{2})(X+\sqrt{2})$.
$X^2 + 1 \in \mathbb{Q}[X]$ zerfällt nicht über $\mathbb{R}$, aber über $\mathbb{C}$ in $(X-\mathrm{i})(X+\mathrm{i})$.

Zerfällung von Polynomen in Linearfaktoren

Alle bisherigen Begriffe und Verfahren der Linearen Algebra sind tatsächlich linear (im Sinne der mathematischen Struktur) und daher algorithmisch recht einfach (im Sinne der informatischen Komplexität).

Die Faktorisierung von Polynomen und speziell die Nullstellensuche verhalten sich dagegen spürbar anders, mathematisch und informatisch: Dieses Problem gehört zur (nicht-linearen!) Algebra bzw. Numerik.

Dieser Frage können wir nun nicht länger ausweichen. Alles weitere ist wieder originärer Bestandteil der Linearen Algebra, doch um überhaupt weiter arbeiten zu können, müssen wir uns um Nullstellen kümmern.

Definition M2K klärt zunächst die nötigen Begriffe und benennt das zu erreichende Ziel: Wir wollen jedes Polynom in Linearfaktoren zerlegen.

☺ Über den komplexen Zahlen $\mathbb{C}$ lässt sich dies immer erreichen, genau dies ist die Aussage des Fundamentalsatzes der Algebra.

Algebraisch abgeschlossene Körper

◆ Satz A3C: Fundamentalsatz der Algebra

Jedes komplexe Polynom $P \in \mathbb{C}[X]$ zerfällt über $\mathbb{C}$. Ausführlich:

Zu jedem Polynom $P(X) = X^n + c_1 X^{n-1} + \dots + c_n$ mit $c_1, \dots, c_n \in \mathbb{C}$ existieren Nullstellen $z_1, \dots, z_n \in \mathbb{C}$ sodass $P(X) = (X-z_1)\cdots(X-z_n)$.

Definition M2L: algebraischer Abschluss

Seien $C \geq K$ Körper, zum Beispiel $\mathbb{C} \geq \mathbb{R}$. Wir nennen C **algebraisch abgeschlossen**, falls jedes Polynom $P \in C[X]$ über C zerfällt.

Die Körpererweiterung $C \geq K$ ist **algebraisch**, falls jedes Element $a \in C$ Nullstelle eines geeigneten Polynoms $P \in K[X]$ ist.

Wir nennen $C \geq K$ einen **algebraischen Abschluss** der Körpers K, falls C algebraisch ist über K und zudem algebraisch abgeschlossen.

Beispiel: Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen, $\mathbb{R}$ jedoch nicht. Die Erweiterung $\mathbb{C} \geq \mathbb{R}$ ist ein algebraischer Abschluss des Körpers $\mathbb{R}$.

Algebraisch abgeschlossene Körper

☺ Daher arbeiten wir meist lieber mit dem Körper $\mathbb{C}$ statt mit $\mathbb{R}$. Glücklicherweise ist solch ein Abschluss über jedem Körper möglich:

Satz M2M: Existenz und Eindeutigkeit

Zu jedem Körper K existiert ein algebraischer Abschluss $C \geq K$.
Je zwei algebraische Abschlüsse C und C' von K sind isomorph.

Ich zitiere dies hier vor allem zur Beruhigung und als schönen Ausblick: Diese und weitere Konstruktionen lernen Sie in der Vorlesung *Algebra*.

☺ Es besteht also kein grundsätzliches mathematisches Problem: Wir können prinzipiell jedes Polynom in Linearfaktoren zerlegen, notfalls nutzen wir hierzu eine geeignete Körpererweiterung.

⚠ Die algorithmischen Fragen hingegen sind überaus anspruchsvoll und führen zu zwei wichtigen, modernen und tiefliegenden Themen: einerseits die Computeralgebra für exaktes, symbolisches Rechnen, andererseits die Numerik für effiziente Näherungslösungen.
Beide gehören zum mathematischen Werkzeugkasten.

Einfache Zerfällung impliziert Diagonalisierbarkeit.

Der folgende Spezialfall ist besonders einfach und sympathisch:

Satz M2N: Einfache Zerfällung impliziert Diagonalisierbarkeit.

Angenommen $f \in \operatorname{End}_K(V)$ hat ein einfach zerfallendes Polynom, also

$$\chi_f(X) = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n)$$

mit n Nullstellen $\lambda_1, \lambda_2, \dots, \lambda_n \in K$, wobei $\lambda_i \ne \lambda_j$ für alle $i \ne j$.

(1) In diesem Falle ist f diagonalisierbar, denn es gilt

$$V = \operatorname{Eig}(f,\lambda_1) \oplus \operatorname{Eig}(f,\lambda_2) \oplus \cdots \oplus \operatorname{Eig}(f,\lambda_n).$$

(2) Wir erhalten eine Eigenbasis $\mathcal{B} = (v_1, v_2, \dots, v_n)$ durch Wahl von Eigenvektoren $v_i \in \operatorname{Eig}(f,\lambda_i) \smallsetminus \{0\}$. Somit wird f dargestellt durch

$$\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Beweis: Wegen $\dim_K \mathrm{E}(\lambda_i) \ge 1$ hat $\mathrm{E} = \bigoplus_{i=1}^n \mathrm{E}(\lambda_i) \le V$ mindestens die Dimension n. Andererseits gilt $\dim_K V = n$, also $\mathrm{E} = V$. $\boxed{\text{QED}}$

Einfache Zerfällung impliziert Diagonalisierbarkeit.

Im allgemeinen Fall muss das char. Polynom $\chi_f \in K[X]$ nicht zerfallen, siehe etwa das Beispiel M1E. Selbst wenn $\chi_f \in K[X]$ über K zerfällt, kann es mehrfache Nullstellen haben, siehe Beispiele M1C und M1D. In diesem Falle ist zur Diagonalisierung noch kein Urteil möglich: f kann diagonalisierbar sein (M1C) oder auch nicht (M1D).

Im Falle mehrfacher Nullstellen hilft letztlich nur die Bestimmung der Eigenräume und insbesondere ihrer Dimension (Satz M1I): Genau dann ist f diagonalisierbar, wenn $V = \bigoplus_{\lambda\in\sigma(f)} \operatorname{Eig}(f,\lambda)$ gilt, also die Dimensionsgleichung $n = \sum_{\lambda\in\sigma(f)} \dim_K \operatorname{Eig}(f,\lambda)$ erfüllt ist. Das ist genau dann der Fall, wenn zu jedem Eigenwert $\lambda \in K$ die geometrische und die algebraische Vielfachheit übereinstimmen (M2O).

Hierzu zeigen wir die folgende nützliche Ungleichung: Die geometrische Vielfachheit ist immer kleiner oder gleich der algebraischen Vielfachheit. Das ist eine recht einfache, aber ebenso grundlegende Beobachtung, die wir anschließend in Theorie und Praxis überall nutzen werden.

Geometrische und algebraische Vielfachheit

Satz M2O: geometrische und algebraische Vielfachheit

(1) Für jeden Eigenwert $\lambda \in \sigma(f)$ gilt

$$\boxed{1 \le \dim_K \operatorname{Eig}(f,\lambda) \le \operatorname{ord}(\chi_f,\lambda) \le \dim V}$$

(2) Genau dann ist f diagonalisierbar, wenn (a) χ_f über K zerfällt und (b) für jeden Eigenwert $\lambda \in \sigma(f)$ gilt $\dim_K \operatorname{Eig}(f,\lambda) = \operatorname{ord}(\chi_f,\lambda)$:

„Die geometrische Vielfachheit erreicht die algebraische.“

Beweis: (1) Wir ergänzen eine Basis $\mathcal{B} = (v_1, \dots, v_k)$ von $\operatorname{Eig}(f,\lambda)$ zu einer Basis $\mathcal{A} = (v_1, \dots, v_k, \dots, v_n)$ von V und erhalten die Darstellung

$$A = \mathrm{M}^{\mathcal{A}}_{\mathcal{A}}(f) = \begin{bmatrix} B & * \\ 0 & C \end{bmatrix} \quad \text{mit } B = \lambda \cdot I_k.$$

Somit gilt $\chi_A = \chi_B \cdot \chi_C = (X-\lambda)^k Q$ und $\dim_K \operatorname{Eig}(f,\lambda) \le \operatorname{ord}(\chi_f,\lambda)$.

(2) Dies folgt aus der Eigenraumzerlegung M1I, dank der Gleichung $\sum_{\lambda\in\sigma(f)} \dim_K \operatorname{Eig}(f,\lambda) \overset{(b)}{=} \sum_{\lambda\in\sigma(f)} \operatorname{ord}(\chi_f,\lambda) \overset{(a)}{=} n = \dim V$. $\boxed{\text{QED}}$

Geometrische und algebraische Vielfachheit

(2a) Die rechte Gleichung $\sum_{\lambda\in\sigma(f)} \operatorname{ord}(\chi_f,\lambda) = n$ ist genau dann erfüllt, wenn das charakteristische Polynom χ_f zerfällt. Andernfalls enthält

$$\chi_f = (X-\lambda_1)^{r_1} \cdots (X-\lambda_k)^{r_k} Q$$

einen nicht-trivialen Faktor $Q \in K[X]$ ohne Nullstellen in K. Aus

$$\textstyle\sum_{\lambda\in\sigma(f)} \operatorname{ord}(\chi_f,\lambda) + \deg Q = n$$

mit $\deg Q > 0$ folgt somit $\sum_{\lambda\in\sigma(f)} \operatorname{ord}(\chi_f,\lambda) < n$.

(2b) Die linke Gleichung

$$\textstyle\sum_{\lambda\in\sigma(f)} \dim_K \operatorname{Eig}(f,\lambda) = \sum_{\lambda\in\sigma(f)} \operatorname{ord}(\chi_f,\lambda)$$

ist genau dann erfüllt, wenn summandenweise

$$\dim \operatorname{Eig}(f,\lambda) = \operatorname{ord}(\chi_f,\lambda)$$

für jeden Eigenwert λ gilt. Dank (1) gilt „$\le$“ in jedem Summanden; gilt hierbei auch nur einmal die strikte Ungleichung „$<$“, so kann dies durch die anderen Summanden nicht mehr ausgeglichen werden.

Anwendungsbeispiel zur Diagonalisierung

Aufgabe: Ist A über $\mathbb{R}$ diagonalisierbar? Falls ja, diagonalisieren Sie A!

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 0 & 1 \\ 0 & 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 0 & 2 \\ 0 & 2 & 3 \end{bmatrix}$$

Lösung: (1a) Wir bestimmen das charakteristische Polynom:

$$\begin{aligned} \tilde{\chi}_A(X) = \begin{vmatrix} 1-X & 4 & 5 \\ 0 & -X & 1 \\ 0 & 2 & 3-X \end{vmatrix} &= (1-X)\big[-X(3-X) - 1 \cdot 2\big] \\ &= (1-X)(X^2 - 3X - 2) \\ &= -X^3 + 4X^2 - X - 2 \end{aligned}$$

(1b) Die Nullstellen von $\chi_A = -\tilde{\chi}_A$ sind die Eigenwerte von A:

$$\sigma(A) = \left\{\, \lambda_1 = 1,\ \lambda_2 = \tfrac{1}{2}(3 - \sqrt{17}),\ \lambda_3 = \tfrac{1}{2}(3 + \sqrt{17}) \,\right\}$$

(1c) Da χ_A einfach zerfällt, schließen wir, dass A diagonalisierbar ist.

Anwendungsbeispiel zur Diagonalisierung

(1d) Die Diagonalisierung verläuft weiter nach Standardverfahren M2ı:

$$D = T^{-1}AT = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}(3-\sqrt{17}) & 0 \\ 0 & 0 & \frac{1}{2}(3+\sqrt{17}) \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & \frac{1}{8}(15-\sqrt{17}) & \frac{1}{8}(15+\sqrt{17}) \\ 0 & \frac{1}{4}(-3-\sqrt{17}) & \frac{1}{4}(-3+\sqrt{17}) \\ 0 & 1 & 1 \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{9}{4} \\ 0 & -\frac{2}{\sqrt{17}} & \frac{1}{2} - \frac{3}{2\sqrt{17}} \\ 0 & +\frac{2}{\sqrt{17}} & \frac{1}{2} + \frac{3}{2\sqrt{17}} \end{bmatrix}$$

Die Rechnungen sind Routine, etwa für ein Computer-Algebra-System (CAS), doch für die Handrechnung nicht ganz kurz. Obwohl die Matrix A über $\mathbb{Q}$ gegeben ist, sind Rechnungen in der Körpererweiterung $\mathbb{Q}[\sqrt{17}]$ nötig. Auch das gelingt letztlich problemlos, wie wir in Kapitel A bereits gesehen haben. **Übung:** Machen Sie die Probe! Was ist zu prüfen?

Anwendungsbeispiel zur Diagonalisierung

Wir führen das Standardverfahren ebenso für die Matrix B aus.

Lösung: (2a) Wir bestimmen das charakteristische Polynom:

$$\begin{aligned} \tilde{\chi}_B(X) = \begin{vmatrix} 1-X & 4 & 5 \\ 0 & -X & 2 \\ 0 & 2 & 3-X \end{vmatrix} &= (1-X)\big[-X(3-X) - 2 \cdot 2\big] \\ &= (1-X)(X^2 - 3X - 4) \\ &= -X^3 + 4X^2 + X - 4 \end{aligned}$$

(2b) Die Nullstellen von $\chi_B = -\tilde{\chi}_B$ sind die Eigenwerte von B:

$$\sigma(B) = \left\{\, \lambda_1 = 1,\ \lambda_2 = -1,\ \lambda_3 = 4 \,\right\}$$

(2c) Da χ_B einfach zerfällt, schließen wir, dass B diagonalisierbar ist.

☺ Die Matrizen A und B unterscheiden sich nur an einer Stelle. Die Rechnungen für B sind dennoch spürbar einfacher.

Anwendungsbeispiel zur Diagonalisierung

(2d) Die Diagonalisierung verläuft weiter nach Standardverfahren M2ı:

$$D = T^{-1}BT = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 3 & 14 \\ 0 & -4 & 3 \\ 0 & 2 & 6 \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{16}{3} \\ 0 & -\frac{1}{5} & -\frac{1}{10} \\ 0 & -\frac{1}{15} & -\frac{2}{15} \end{bmatrix}$$

☺ Alle Eigenwerte liegen in $\mathbb{Q}$, das vereinfacht die Rechnung spürbar!

Übung: Machen Sie auch hier die Probe! Was ist zu prüfen?

Beispiel: symmetrische reelle Matrizen

Beispiel M2P: symmetrische reelle 2×2–Matrix

Zu $a, b, c \in \mathbb{R}$ betrachten wir die symmetrische Matrix

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Ist A über $\mathbb{R}$ diagonalisierbar? Ihr charakteristisches Polynom ist

$$\chi_A = \begin{vmatrix} a-X & b \\ b & c-X \end{vmatrix} = X^2 - (a+c)X + (ac-b^2).$$

Die Diskriminante ist $D = (a+c)^2 - 4(ac-b^2) = (a-c) + 4b^2 \geq 0$.

Somit zerfällt $\chi_A = (X-\lambda_1)(X-\lambda_2)$ mit $\lambda_{1/2} = \frac{1}{2}(a+c\pm\sqrt{D}) \in \mathbb{R}$.

- Bei $D > 0$ zerfällt χ_A einfach, daher ist A diagonalisierbar (M2N).
- Bei $D = 0$ gilt $b = 0$ und $a = c$, somit ist A bereits diagonal.

☺ Jede symmetrische Matrix $A \in R^{2\times 2}$ ist $\mathbb{R}$–diagonalisierbar.

Beispiel: symmetrische reelle Matrizen

☺ Das ist ein sehr schönes, allgemeines und nützliches Ergebnis:
Wir betrachten eine symmetrische reelle Matrix $A \in \mathbb{R}^{n\times n}$, also $A^{\intercal} = A$.

Der erste interessante Fall ist die Dimension $n = 2$:
Das obige Beispiel zeigt, dass A diagonalisierbar ist.

Dies beweisen wir später allgemein in jeder Dimension $n \in \mathbb{N}$:
Jede symmetrische reelle Matrix ist über $\mathbb{R}$ diagonalisierbar!

Die Ausführung der Diagonalisierung erfordert wie immer die explizite Berechnung einer Eigenbasis. Im Falle $n = 2$ gelingt dies bereits jetzt ganz direkt. Im allgemeinen Fall werden wir dies später genauer noch untersuchen und hilfreiche Werkzeuge erarbeiten.

Übung: Diagonalisieren Sie die obige Matrix A im Falle $b > 0$.
(1) Untersuchen Sie einige konkrete Zahlenbeispiele.
(2) Entwickeln Sie eine allgemeine Formel.

Normiertes Polynom und seine Begleitmatrix

Tritt jedes normierte Polynom als charakteristisches Polynom auf? Ja!

Satz M2Q: normiertes Polynom und seine Begleitmatrix

Sei K ein Körper. Gegeben sei ein normiertes Polynom

$$P = X^n + p_1 X^{n-1} + \dots + p_n X^0 \in K[X]^1_n.$$

Hierzu definieren wir die **Begleitmatrix** (engl. *companion matrix*):

$$C = C(P) := \begin{bmatrix} 0 & 0 & \dots & 0 & -p_n \\ 1 & 0 & \dots & 0 & -p_{n-1} \\ 0 & 1 & \ddots & \vdots & -p_{n-2} \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -p_1 \end{bmatrix} \in K^{n\times n}$$

(1) Ihr charakteristisches Polynom ist P. (2) Genau dann ist C über K diagonalisierbar, wenn $P = (X-\lambda_1)\cdots(X-\lambda_n)$ einfach zerfällt.

In diesem Falle gelingt die Diagonalisierung $VCV^{-1} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ mit der Vandermonde–Matrix $V = (\lambda_i^j)^{j=0,\dots,n-1}_{i=1,\dots,n} \in \mathrm{GL}_n(K)$, siehe B3A.

Normiertes Polynom und Begleitmatrix

Einfache Zerfällung ist immer hinreichend für Diagonalisierbarkeit (M2N). Satz M2Q besagt speziell für die Begleitmatrix $C = C(P)$, dass einfache Zerfällung auch notwendig ist und gibt explizit eine Eigenbasis an.

Übung: (1) Berechnen Sie das Polynom, hier $\chi_C = P$, indem Sie die charakteristische Matrix $XI - C$ nach der letzten Spalte entwickeln.

(2a) Für jeden Skalar $\lambda \in K$ gilt $\dim \operatorname{Eig}(C, \lambda) \leq 1$,
denn die Matrix $\lambda I - C$ hat mindestens Rang $n - 1$.

(2b) Wie folgt daraus das Diagonalisierbarkeitskriterium?
Warum ist es hier nicht nur hinreichend, sondern auch notwendig?

(2c) Die Eigenvektoren finden wir auf ganz natürliche Weise in den folgenden Untersuchungen zur linearen Rekursionen. Sie können es jetzt schon direkt versuchen, oder anschließend darauf zurückkommen.

☺ Die inverse Matrix V^{-1} ist mühsam, wie so oft. Für die Probe ist sie gar nicht nötig: Es genügt sicherzustellen, dass V invertierbar ist (B3A), und statt $VCV^{-1} = D$ die äquivalente Gleichung $VC = DV$ zu prüfen.

Anwendungsbeispiel: die Fibonacci–Folge

Die Fibonacci–Folge $f:\mathbb{N}\to\mathbb{N}:n\mapsto f_n$ ist definiert durch die Startwerte $f_0=0$ und $f_1=1$ sowie für $n\in\mathbb{N}_{\ge 2}$ die **Rekursionsvorschrift**

$$\boxed{f_n = f_{n-1} + f_{n-2}.}$$

Die ersten Werte sind demnach:

n	0	1	2	3	4	5	6	7	8	9	$\dots$
f_n	0	1	1	2	3	5	8	13	21	34	$\dots$

Leonardo Fibonacci beschrieb damit im Jahr 1202 das Wachstum einer Kaninchenpopulation; in der Natur ist sie ein recht häufiges Muster.

Für alle $n\in\mathbb{N}$ gilt die phantastische **Binet–Formel**

$$\boxed{f_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].}$$

Übung: (1) Wie *beweisen* Sie die gegebene Formel? Induktion! (C443)
(2) Wie *finden* Sie eine solche Formel? Eigenwerte und Eigenvektoren!

Anwendungsbeispiel: die Fibonacci–Folge

Näherungswerte sind $\phi = \frac{1+\sqrt{5}}{2} \approx 1.618$ und $\psi = \frac{1-\sqrt{5}}{2} \approx -0.618$.

☺ Die geschlossene Formel für f_n zeigt das Wachstumsverhalten:

- Der zweite Summand $(\frac{1-\sqrt{5}}{2})^n$ konvergiert rasch gegen Null.
- Die Fibonacci–Folge f_n wächst exponentiell, wie $(\frac{1+\sqrt{5}}{2})^n$.

☺ Zunächst handelt die Fragestellung von Zahlenfolgen und Rekursion. Auf den ersten Blick hat das nichts mit Linearer Algebra, Matrizen oder gar Eigenvektoren zu tun. Es zeigt sich jedoch, dass diese universellen Werkzeuge auch hier wunderbar effizient angewendet werden können!

Das ist ein allgemeines und häufiges Phänomen: Die Problemstellung deutet meist noch nicht die möglichen oder nötigen Werkzeuge an.
Zur Problemlösung braucht es daher Erfahrung und Kreativität.
Man erblickt nur, was man schon weiß und versteht. (Goethe)

☺ Erfahrung sammeln Sie durch Übungen wie etwa der folgenden. Vielfältige Rechnungen und Erprobung der Werkzeuge übt die korrekte Anwendung und fördert Ihre Kreativität bei zukünftigen neuen Aufgaben.

Der Folgenraum $\mathbb{K}^{\mathbb{N}}$ mit Verschiebeoperator

Sei $\mathbb{K}$ ein Körper, etwa $\mathbb{K}=\mathbb{Q},\mathbb{R},\mathbb{C}$. Wir betrachten den Vektorraum $\mathbb{K}^{\mathbb{N}}$ aller Folgen $f:\mathbb{N}\to\mathbb{K}:n\mapsto f_n$ und hierauf den Verschiebeoperator

$$s\;:\;\mathbb{K}^{\mathbb{N}}\to\mathbb{K}^{\mathbb{N}}\;:\;(f_0,f_1,f_2,\dots)\mapsto(f_1,f_2,f_3,\dots).$$

Ausgeschrieben bedeutet das $s(f)=g$ mit $g_n=f_{n+1}$.
Dies ist offensichtlich eine $\mathbb{K}$–lineare Abbildung.

Aufgabe: Bestimmen Sie zu s alle Eigenwerte und Eigenräume.

Lösung: Sei $\lambda\in\mathbb{K}$. Die Gleichung $s(f)=\lambda f$ bedeutet ausgeschrieben

$$s(f) = (f_1,f_2,f_3,\dots) \overset{!}{=} (\lambda f_0,\lambda f_1,\lambda f_2,\dots) = \lambda f$$

Das entspricht der Rekursionsvorschrift $f_{n+1}=\lambda f_n$ für alle $n\in\mathbb{N}$.
Jede solche Folge $f:\mathbb{N}\to\mathbb{K}$ erfüllt somit $f_n=\lambda^n f_0$ für alle $n\in\mathbb{N}$.
Wir wählen den Eigenvektor $e_\lambda:\mathbb{N}\to\mathbb{K}:n\mapsto\lambda^n$ und erhalten:

$$\operatorname{Eig}(s,\lambda) = \langle\, e_\lambda\,\rangle_{\mathbb{K}}^{!}$$

Der Folgenraum $\mathbb{K}^{\mathbb{N}}$ mit Verschiebeoperator

☺ Hier ist *jeder* Skalar $\lambda\in\mathbb{K}$ ein Eigenwert, also $\sigma(s;\mathbb{K})=\mathbb{K}$.
Jeder Eigenraum ist eindimensional, das ist äußerst übersichtlich.
Die Eigenvektoren e_λ können wir bequem und explizit angeben.

⚠ Hier ist der Vektorraum $\mathbb{K}^{\mathbb{N}}$ unendlich-dimensional. Determinante und charakteristisches Polynom stehen uns daher nicht zur Verfügung. Dennoch können wir Eigenwerte und Eigenräume direkt berechnen.

Übung: Ist der Verschiebeoperator $s:\mathbb{K}^{\mathbb{N}}\to\mathbb{K}^{\mathbb{N}}$ diagonalisierbar? Auch wenn der Raum $\mathbb{K}^{\mathbb{N}}$ unendlich-dimensional ist, können wir uns dennoch eine diagonalisierende Basis wünschen. Ist dies hier erfüllbar? (Der endliche Fall $\sharp\mathbb{K}<\infty$ ist leicht, der allgemeine Fall etwas knifflig.)

Übung: Im Falle $\mathbb{K}=\mathbb{R},\mathbb{C}$ betrachten wir die Menge $U=\ell^\infty(\mathbb{N},\mathbb{K})$ der beschränkten Folgen. Dies ist ein Unterraum von $\mathbb{K}^{\mathbb{N}}$. Dieser Unterraum ist s–invariant, das heißt $s(U)\subseteq U$. Durch Einschränkung erhalten wir den Verschiebeoperator $s:U\to U$. Bestimmen Sie sein Spektrum.

Anwendungsbeispiel zu rekursiven Folgen

☺ Wir untersuchen nun die Fibonacci–Folge mit unseren Werkzeugen der Linearen Algebra, hier also mit Eigenwerten und Eigenvektoren.

Wir betrachten in $\mathbb{K}^{\mathbb{N}}$ den Kern von $s^2 - s - \mathrm{id}$, also den Unterraum

$$V = \left\{ f:\mathbb{N}\to\mathbb{K} \;\middle|\; \forall n\in\mathbb{N}: f_{n+2} = f_{n+1} + f_n \right\}.$$

Nach Konstruktion ist $V \le \mathbb{K}^{\mathbb{N}}$ zudem s–invariant, das heißt $s(V) \subseteq V$. Durch Einschränkung erhalten wir den Verschiebeoperator $s : V \to V$.

Aufgabe: (1) Welche Dimension hat V? (2) Stellen Sie s als Matrix dar. (3) Bestimmen Sie Eigenwerte & Eigenräume. (4) Diagonalisieren Sie s. (5) Linearkombinieren Sie die Fibonacci–Folge aus Eigenvektoren.

Lösung: (1) Wir betrachten die Projektion

$$q \;:\; \mathbb{K}^{\mathbb{N}} \ge V \to \mathbb{K}^2 \;:\; (f_0, f_1, f_2, \dots) \mapsto (f_0, f_1).$$

Zu beliebigen Startwerten $f_0, f_1 \in \mathbb{K}$ existiert genau eine Folge $f \in V$. Somit ist $q : V \xrightarrow{\sim} \mathbb{K}^2$ ein Isomorphismus, insbesondere $\dim_{\mathbb{K}} V = 2$.

Anwendungsbeispiel zu rekursiven Folgen

(2) Als eine Basis $\mathcal{A} = (a_1, a_2)$ von V wählen wir $a_1 = (1, 0, 1, 1, 2, 3, \dots)$ und $a_2 = (0, 1, 1, 2, 3, 5, \dots)$. Es gilt $s(a_1) = a_2$ und $s(a_2) = a_1 + a_2$. Also:

$$A = \mathrm{M}^{\mathcal{A}}_{\mathcal{A}}(s) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

(3) Aus dieser Matrix gewinnen wir das charakteristische Polynom:

$$\chi_A = \begin{vmatrix} -X & 1 \\ 1 & 1-X \end{vmatrix} = X^2 - X - 1$$

Die beiden Nullstellen sind $\phi = \frac{1}{2}(1+\sqrt{5})$ und $\psi = \frac{1}{2}(1-\sqrt{5})$. Als Eigenvektoren erhalten wir $u = (\phi^n)_{n\in\mathbb{N}}$ und $v = (\psi^n)_{n\in\mathbb{N}}$ in V.

(4) Damit erhalten wir eine Eigenbasis $\mathcal{B} = (u, v)$ von $s : V \to V$ sowie

$$D = T^{-1} A T = \begin{bmatrix} \phi & 0 \\ 0 & \psi \end{bmatrix} \quad \text{mit } T = \begin{bmatrix} 1 & 1 \\ \phi & \psi \end{bmatrix} \quad \text{und } T^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} -\psi & 1 \\ \phi & -1 \end{bmatrix}.$$

Obwohl die Matrix A über $\mathbb{Q}$ gegeben ist, sind Rechnungen in der Körpererweiterung $\mathbb{Q}[\sqrt{5}]$ nötig. **Übung:** Machen Sie die Probe!

Anwendungsbeispiel zu rekursiven Folgen

(5) Wir stellen die Fibonacci–Folge $f = (0, 1, 1, 2, 3, 5, 8, \dots) \in V$ dar als Linearkombination der Eigenvektoren $u = (\phi^n)_{n\in\mathbb{N}}$ und $v = (\psi^n)_{n\in\mathbb{N}}$:

$$f = au + bv \quad\Leftrightarrow\quad \begin{Bmatrix} 0 = a\phi^0 + b\psi^0 \\ 1 = a\phi^1 + b\psi^1 \end{Bmatrix} \quad\Leftrightarrow\quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = T \begin{bmatrix} a \\ b \end{bmatrix}$$

Wir finden so die einzige Lösung für die gesuchten Koeffizienten:

$$\begin{bmatrix} a \\ b \end{bmatrix} = T^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Explizit ausgeschrieben ergibt dies die Binet–Formel:

$$\boxed{f_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]}$$

☺ So lösen Sie Rekursionsgleichungen durch geschlossene Formeln. Dies gelingt mit den Werkzeugen der Linearen Algebra: Eigenvektoren!

Übung: Prüfen Sie die so gefundene Gleichung per Induktion (C443).

Anwendungsbeispiel zu rekursiven Folgen

Das hier gezeigte Verfahren gilt allgemein für rekursive Folgen!

$$V = \left\{ f:\mathbb{N}\to\mathbb{K} \;\middle|\; \forall k \ge n \;:\; f_k + p_1 f_{k-1} + \dots + p_n f_{k-n} = 0 \right\}$$

Dies entspricht der Rekursionsgleichung $P(s)(f) = 0$ mit dem Polynom

$$P = X^n + p_1 X^{n-1} + \dots + p_n X^0 \in \mathbb{K}[X]^1_n.$$

Im Fibonacci–Beispiel haben wir $P = X^2 - X - 1$ betrachtet.

Allgemein ist $P \in \mathbb{K}[X]^1_n$ gegeben, wir haben also $V = \ker P(s)$. Dieser Unterraum V in $\mathbb{K}^{\mathbb{N}}$ ist s–invariant, das heißt $s(V) \subseteq V$.

Wir wollen den Verschiebeoperator $s : V \to V$ diagonalisieren, also eine Basis aus Eigenvektoren der Form $(\lambda^n)_{n\in\mathbb{N}}$ bestimmen.

Falls P über $\mathbb{K}$ einfach zerfällt, so gelingt dies wörtlich wie zuvor: Zu beliebigen Anfangswerten finden wir so eine geschlossene Formel!

☺ Das ist das Standardverfahren M2I zur Diagonalisierung, und es funktioniert auch für Rekursionsgleichungen wunderbar.

Lineare Rekursion und ihre Eigenfolgen

Satz M2R: lineare Rekursion und ihre Eigenfolgen

Sei $\mathbb{K}$ ein Körper. Vorgelegt sei ein beliebiges normiertes Polynom

$$P = X^n + p_1 X^{n-1} + \dots + p_n X^0 \in \mathbb{K}[X]^1_n.$$

Wir betrachten den Folgenraum $\mathbb{K}^{\mathbb{N}}$ mit dem Verschiebeoperator

$$s \;:\; \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}} \;:\; (f_0, f_1, f_2, \dots) \mapsto (f_1, f_2, f_3, \dots).$$

Darin liegt der $\mathbb{K}$–Untervektorraum $V \le \mathbb{K}^{\mathbb{N}}$ der P–rekursiven Folgen:
$V = \ker P(s) = \{\, f : \mathbb{N} \to \mathbb{K} \mid \forall k \ge n \;:\; f_k + p_1 f_{k-1} + \dots + p_n f_{k-n} = 0 \,\}$
Dieser Unterraum ist s–invariant, $s(V) \subseteq V$, mit Dimension $\dim_{\mathbb{K}} V = n$.

Genau dann ist der (so eingeschränkte) Verschiebeoperator $s : V \to V$ über $\mathbb{K}$ diagonalisierbar, wenn das Polynom P über $\mathbb{K}$ einfach zerfällt,

$$P(X) = (X - \lambda_1) \cdots (X - \lambda_n) \quad \text{mit } \lambda_i \ne \lambda_j \text{ für } i \ne j.$$

Zu jedem Eigenwert λ haben wir die Eigenfolge $e_\lambda : \mathbb{N} \to \mathbb{K} : n \mapsto \lambda^n$.
Wir erhalten so die Eigenbasis $(e_{\lambda_1}, \dots, e_{\lambda_n})$ von V bezüglich s.

Lineare Rekursion und ihre Eigenfolgen

Einfache Zerfällung ist immer hinreichend für Diagonalisierbarkeit (M2N).
Satz M2R besagt speziell für den Verschiebeoperator, dass einfache Zerfällung auch notwendig ist und gibt explizit eine Eigenbasis an.

Aufgabe: Beweisen Sie diesen allgemeinen Satz nach dem Vorbild des Fibonacci–Beispiels. Bestimmen Sie (1) die Dimension von V, (2) eine darstellende Matrix zu $s : V \to V$, (3) alle Eigenwerte und (4) Eigenräume, (5) eine Eigenbasis und (6) den Basiswechsel.

Lösung: (1) Wir betrachten die Projektion

$$q \;:\; \mathbb{K}^{\mathbb{N}} \ge V \to \mathbb{K}^n \;:\; f \mapsto (f_0, \dots, f_{n-1}).$$

Zu je n beliebig vorgegebenen Startwerten $f_0, \dots, f_{n-1} \in \mathbb{K}$ existiert genau eine P–rekursive Folge $f = (f_0, \dots, f_{n-1}, f_n, \dots)$.

Somit ist $q : V \xrightarrow{\sim} \mathbb{K}^n$ ein Isomorphismus, insbesondere $\dim_{\mathbb{K}} V = n$.
Diese einfache Vorgehensweise ist ebenso elegant wie effizient.

🙂 Durch diesen Isomorphismus q bestimmen wir die Dimension und erhalten nach Wunsch auch Basen, etwa $(q^{-1}(e_1), \dots, q^{-1}(e_n))$.

Lineare Rekursion und ihre Eigenfolgen

(2) Den Verschiebeoperator $s : V \to V$ können wir wie folgt darstellen:

$$\begin{tikzcd} V \arrow[r, "s"] \arrow[d, "q"', "\cong"] & V \arrow[d, "q"', "\cong"] \\ \mathbb{K}^n \arrow[r, "A"] & \mathbb{K}^n \end{tikzcd}$$

$$\begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{bmatrix} \mapsto \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ -p_n & -p_{n-1} & -p_{n-2} & \dots & -p_1 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{bmatrix}$$

Das charakteristische Polynom ist $\chi_A = P$, dank $A = C(P)^{\intercal}$ und M2Q.
Wir sehen $\dim \operatorname{Eig}(s, \lambda) \le 1$ und können direkt Satz M1I anwenden:

(3) Genau dann ist $s : V \to V$ diagonalisierbar, wenn P einfach zerfällt, $P(X) = (X - \lambda_1) \cdots (X - \lambda_n)$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und $\lambda_i \ne \lambda_j$ für $i \ne j$.

Lineare Rekursion und ihre Eigenfolgen

(4) Es gilt $\operatorname{Eig}(s, \lambda) = \langle\, e_\lambda \,\rangle^!_{\mathbb{K}}$ mit der Eigenfolge $e_\lambda : \mathbb{N} \to \mathbb{K} : n \mapsto \lambda^n$.

(5) Wir erhalten so die Eigenbasis $(e_{\lambda_1}, \dots, e_{\lambda_n})$ von V bezüglich s.

(6) Hierdurch wird der Verschiebeoperator $s : V \to V$ diagonalisiert:

$$T^{-1} A T = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \quad \text{mit } T = \begin{bmatrix} \lambda_1^0 & \lambda_2^0 & \dots & \lambda_n^0 \\ \lambda_1^1 & \lambda_2^1 & \dots & \lambda_n^1 \\ \vdots & & \vdots & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{bmatrix}$$

🙂 Die Basiswechselmatrix $T = \operatorname{VDM}(\lambda_1, \lambda_2, \dots, \lambda_n)^{\intercal}$ ist die berühmte Vandermonde–Matrix (siehe Satz B3A), hier transponiert. Alles ist gut.

Übung: Machen Sie die Probe $AT = TD$.

🙂 Die inverse Matrix T^{-1} ist mühsam, wie so oft. Für die Probe ist sie gar nicht nötig: Es genügt sicherzustellen, dass T invertierbar ist (B3A), und statt $T^{-1}AT = D$ die äquivalente Gleichung $AT = TD$ zu prüfen.

Trigonalisierung eines Endomorphismus

Definition M3A: Trigonalisierung eines Endomorphismus

Sei K ein Ring, wir denken insbesondere an Körper wie $\mathbb{C}, \mathbb{R}, \mathbb{Q}, \mathbb{F}_p, \dots$.

(1) Sei $f : V \to V$ linear über K. Eine **trigonalisierende Basis** zu f ist eine Basis $\mathcal{B}$ von V, für die die darstellende Matrix von f trigonal ist:

$$\mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Existiert eine solche Basis $\mathcal{B}$ von V, so nennen wir f **trigonalisierbar**.

(2) Sei $A \in K^{n \times n}$. Ein **trigonalisierender Basiswechsel** zu A über K ist eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$, so dass $T^{-1}AT$ trigonal ist:

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Existiert eine solche Matrix T, so nennen wir A **trigonalisierbar**.

Trigonalisierung und Zerfällung

Wie können wir feststellen, ob f bzw. A trigonalisierbar ist?

Satz M3B: Trigonalisierung ⇔ Zerfällung

Sei $f : V \to V$ linear über dem Körper K mit $\dim_K V = n < \infty$. Genau dann ist f über K trigonalisierbar, wenn χ_f über K zerfällt:

$$\chi_f(X) = (X - \lambda_1) \cdots (X - \lambda_n) \quad \text{mit } \lambda_1, \dots, \lambda_n \in K$$

Dasselbe gilt für die Trigonalisierung einer Matrix $A \in K^{n \times n}$ über K.

Beispiel: Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{C}$. Dann ist jeder $\mathbb{C}$–Endomorphismus $f : V \to V$ trigonalisierbar.

Beweis: „⇒“: Zu f sei $\mathcal{B}$ eine trigonalisierende Basis von V:

$$B = \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Demnach gilt $\chi_f = \chi_B = \det(XI - B) = (X - \lambda_1) \cdots (X - \lambda_n)$.

Trigonalisierung und Zerfällung

„⇐“: Wir führen Induktion über n. Für $n = 1$ ist die Aussage trivial. Sei nun $n \ge 2$, und die Behauptung für $n - 1$ sei bereits bewiesen.

Zum Eigenwert $\lambda_1 \in K$ existiert ein Eigenvektor $v_1 \in \mathrm{Eig}(f, \lambda_1) \le V$. Diesen ergänzen wir zu einer Basis $\mathcal{A} = (v_1, u_2, \dots, u_n)$ von V. (J2B)

$$\begin{matrix} A = \\ \mathrm{M}^{\mathcal{A}}_{\mathcal{A}}(f) = \end{matrix} \left[\begin{array}{c|ccc} \lambda_1 & * & * & * \\ \hline 0 & & & \\ 0 & & A' & \\ 0 & & & \end{array}\right] \implies \begin{matrix} B = \\ \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \end{matrix} \left[\begin{array}{c|ccc} \lambda_1 & * & * & * \\ \hline 0 & \lambda_2 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_n \end{array}\right]$$

Wir zerlegen $V = U \oplus V'$ in $U = \langle\, v_1 \,\rangle^{!}_{K}$ und $V' = \langle\, u_2, \dots, u_n \,\rangle^{!}_{K}$. Einschränkung und Projektion ergibt $f' = p_{V'} \circ f \circ \iota_{V'} : V' \to V'$. In der Basis $\mathcal{A}' = (u_2, \dots, u_n)$ gilt $A' = \mathrm{M}^{\mathcal{A}'}_{\mathcal{A}'}(f')$ wie oben gezeigt. Wir haben $\chi_f = (X - \lambda_1) \cdot \chi_{f'}$, also zerfällt auch das Polynom $\chi_{f'}$.

Dank $\dim V' = n - 1$ greift nun unsere Induktionsvoraussetzung: Zu f' existiert eine trigonalisierende Basis $\mathcal{B}' = (v_2, \dots, v_n)$ von V'. Die Basis $\mathcal{B} = (v_1, v_2, \dots, v_n)$ von V trigonalisiert $f : V \to V$. QED

Trigonalisierung und Zerfällung

☺ Um Induktion über die Dimension führen zu können, müssen wir die Dimension reduzieren. Der entscheidende Konstruktionsschritt ist hier die Projektion auf den strikt kleineren Teilraum $V' < V$:

$$\begin{tikzcd} V \arrow[r, "f"] & V \arrow[d, two heads, "p_{V'}"] \\ V' \arrow[u, hook, "\iota_{V'}"] \arrow[r, dashed, "f'"] & V' \end{tikzcd}$$

Konkret bedeutet das: Zu jedem $v' = \sum_{k=2}^{n} \alpha_k u_k \in V'$ betrachten wir $f(v') = \beta_1 v_1 + \sum_{k=2}^{n} \beta_k u_k$ und setzen $f'(v') := \sum_{k=2}^{n} \beta_k u_k \in V'$. Wir ignorieren also ganz dreist-naiv den Störterm $\beta_1 v_1 \in U$.

☺ Dass $f' : V' \to V'$ linear ist, sehen wir an $f' = p_{V'} \circ f \circ \iota_{V'}$.

☺ In der Zerlegung $V = U \oplus V'$ als direkte Summe ist die Wahl des Komplements V' willkürlich. Mit dem Quotienten $W = V/U$ gelingt dies natürlich und elegant, aber auch abstrakter. Dies führen wir nun aus.

Trigonalisierung: der Quotiententrick

Lemma M3C: Endomorphismus einer kurzen exakten Sequenz

Sei $f : V \to V$ linear über K und $U \le V$ ein invarianter Unterraum.

(1) Wir haben die Einschränkung $g = f_U^U$ und auf dem Quotienten $W = V/U$ die lineare Abbildung $h : W \to W : x + U \mapsto f(x) + U$.

$$\begin{tikzcd}
U \arrow[r, hook, "i"] \arrow[d, "g"] \arrow[dr, "f \circ i"] & V \arrow[r, two heads, "q"] \arrow[d, "f"] \arrow[dr, "q \circ f"] & W \arrow[d, "h"] \\
U \arrow[r, hook, "i"] & V \arrow[r, two heads, "q"] & W
\end{tikzcd}$$

(2) Aus je zwei Basen $\mathcal{A}$ von U und $\mathcal{C}$ von W erhalten wir eine Basis $\mathcal{B}$ von V gemäß Satz J2M. Für die darstellenden Matrizen gilt dann:

$$\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} \mathrm{M}_{\mathcal{A}}^{\mathcal{A}}(g) & * \\ 0 & \mathrm{M}_{\mathcal{C}}^{\mathcal{C}}(h) \end{bmatrix}$$

Für die charakteristischen Polynome folgt somit $\chi_f = \chi_g \cdot \chi_h$.

Trigonalisierung: der Quotiententrick

Aufgabe: Beweisen Sie Lemma M3C und damit erneut Satz M3B.

Beweis von Lemma M3C: (1) Wir erhalten $g : U \to U$ und $h : W \to W$ durch lineare Faktorisierung über eine Injektion I2F bzw. Surjektion I2E.
(2) Die Matrixdarstellung folgt aus (1) und der Konstruktion der Basis $\mathcal{B}$.

Beweis von Satz M3B: Wir zeigen hier nur die Rückrichtung „$\Leftarrow$“:
Wir führen Induktion über $n = \dim V$. Für $n = 1$ ist die Aussage trivial.
Sei nun $n \ge 2$, und die Behauptung für $n - 1$ sei bereits bewiesen.

Zum Eigenwert $\lambda_1 \in K$ existiert ein Eigenvektor $v_1 \in \operatorname{Eig}(f, \lambda_1) \le V$.
Der Unterraum $U = \langle v_1 \rangle_K^! < V$ ist f–invariant. Lemma M3C beschert uns die lineare Abbildung $h : W \to W$ auf dem Quotienten $W = V/U$.
Dabei gilt $\chi_f = (X - \lambda_1) \cdot \chi_h$, also zerfällt auch das Polynom χ_h.

Dank $\dim W = \dim V - \dim U < n$ greift die Induktionsvoraussetzung:
Zu $h : W \to W$ existiert eine trigonalisierende Basis $\mathcal{C} = (w_2, \dots, w_n)$.
Wir wählen Urbilder $v_2, \dots, v_n \in V$ mit $q(v_i) = w_i$. Dank Lemma M3C wird f durch die Basis $\mathcal{B} = (v_1, v_2, \dots, v_n)$ trigonalisiert. $\boxed{\text{QED}}$

Trigonalisierung: Fahne von Unterräumen

Definition M3D: Fahne von Unterräumen

(1) Sei V ein K–Vektorraum der Dimension n. Eine **Fahne** ist eine Kette

$$\{0\} = V_0 < V_1 < \dots < V_\ell = V$$

von Unterräumen. Dabei gilt $0 = \dim V_0 < \dim V_1 < \dots < \dim V_\ell = n$.
Eine **vollständige Fahne** erfüllt $\ell = n$ und $\dim V_i = i$ für $i = 0, 1, \dots, n$.

Die ersten Teilräume $V_0 < V_1 < V_2$ mit $\dim V_i = i$ entsprechen Punkt, Gerade, Ebene; daher der geometrisch anschauliche Name „Fahne“.

Beispiel: Ist $(v_1, \dots, v_n)$ eine Basis von V über K, so definieren die aufgespannten Unterräume $V_i = \langle v_1, \dots, v_i \rangle_K$ für $i = 0, 1, \dots, n$ eine vollständige Fahne $\{0\} = V_0 < V_1 < \dots < V_n = V$.

Definition M3D: invariante Fahne von Unterräumen

(2) Sei $f : V \to V$ eine K–lineare Abbildung.
Ein Unterraum $U \le V$ heißt ***f*–invariant**, wenn $f(U) \subseteq U$ gilt.
Eine Fahne $(V_i)_i$ heißt ***f*–invariant**, wenn $f(V_i) \subseteq V_i$ für alle i gilt.

Trigonalisierung: Fahne von Unterräumen

Lemma M3E: Trigonalisierung ⇔ invariante Fahne

Genau dann ist $f : V \to V$ trigonalisierbar, wenn eine vollständige, f–invariante Fahne $\{0\} = V_0 < V_1 < \dots < V_n = V$ existiert.

$$\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

Aufgabe: Beweisen Sie dieses Lemma.

Lösung: „$\Rightarrow$“: Angenommen, zu f existiert eine trigonalisierende Basis $(v_1, \dots, v_n)$ von V. Die Unterräume $V_i = \langle v_1, \dots, v_i \rangle_K$ für $i = 0, 1, \dots, n$ sind eine vollständige Fahne und zudem f–invariant, denn $f(V_i) \subseteq V_i$.

„$\Leftarrow$“: Durch schrittweise Ergänzung erhalten wir eine Basis $(v_1, \dots, v_i)$ von V_i. Die Basis $\mathcal{B} = (v_1, v_2, \dots, v_n)$ von V trigonalisiert $f : V \to V$.

Spur und Determinante aus Eigenwerten

Satz M3F: Spur und Determinante

Sei $A \in K^{n\times n}$ eine Matrix mit zerfallendem charakteristischen Polynom

$$\chi_A(X) = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n).$$

(1) Die Spur von A ist die Summe der Eigenwerte:

$$\boxed{\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n}$$

(2) Die Determinante ist das Produkt der Eigenwerte:

$$\boxed{\det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n}$$

☺ Die Spur $\operatorname{tr}(A)$ der Matrix A ist besonders leicht zu berechnen. Wenn wir also $n-1$ Eigenwerte kennen, so ist der letzte gratis!

☺ Über $\mathbb{C}$ zerfällt jedes Polynom, die Voraussetzung ist also erfüllt. Dasselbe gilt über jedem algebraisch abgeschlossenen Körper.

Aufgabe: Beweisen Sie diesen Satz! (Zwei Wege sind möglich.)

Spur und Determinante aus Eigenwerten

Erster Beweis: Wir multiplizieren zur Monomform aus:

$$\begin{aligned}\chi_A(X) &= (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n)\\ &= X^n - a_{n-1}X^{n-1} \pm \cdots + (-1)^n a_0\end{aligned}$$

Dank Satz M2C kennen wir die extremen Koeffizienten:

$$\begin{aligned}\operatorname{tr}(A) &\overset{\text{M2C}}{=} a_{n-1} &&= \lambda_1 + \lambda_2 + \cdots + \lambda_n,\\ \det(A) &\overset{\text{M2C}}{=} a_0 &&= \lambda_1 \cdot \lambda_2 \cdots \lambda_n.\end{aligned}$$

Zweiter Beweis: Dank M3B können wir A trigonalisieren zu

$$B = T^{-1}AT = \begin{bmatrix}\lambda_1 & * & *\\ 0 & \ddots & *\\ 0 & 0 & \lambda_n\end{bmatrix}$$

Dank Satz M2F gilt dabei $\chi_A = \chi_B$, also insbesondere

$$\begin{aligned}\operatorname{tr}(A) &\overset{\text{M2F}}{=} \operatorname{tr}(B) &&= \lambda_1 + \lambda_2 + \cdots + \lambda_n,\\ \det(A) &\overset{\text{M2F}}{=} \det(B) &&= \lambda_1 \cdot \lambda_2 \cdots \lambda_n.\end{aligned}$$

☺ Der zweite Beweis ist noch schöner, dank stärkerer Werkzeuge.

Elementarsymmetrische Polynome

Der obige Satz M3F erklärt die Spur $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ und die Determinante $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$ aus den Eigenwerten. Allgemein:

Definition M3G: elementarsymmetrische Polynome

Im Polynomring $\mathbb{Z}[T, X_1, \dots, X_n]$ betrachten wir das Produkt

$$(T+X_1)(T+X_2)\cdots(T+X_n) = T^n + \sigma_1 T^{n-1} + \sigma_2 T^{n-2} + \cdots + \sigma_n.$$

Der hier auftretende Koeffizient $\sigma_k \in \mathbb{Z}[X_1, \dots, X_n]$ für $k = 1, 2, \dots, n$ ist das **elementarsymmetrische Polynom** vom Grad k in $X_1, X_2, \dots, X_n$:

$$\begin{aligned}\sigma_1(X_1, X_2, \dots, X_n) &= X_1 + X_2 + \cdots + X_n\\ \dots\\ \sigma_k(X_1, X_2, \dots, X_n) &= \textstyle\sum_{i_1<i_2<\dots<i_k} X_{i_1}X_{i_2}\cdots X_{i_k}\\ \dots\\ \sigma_n(X_1, X_2, \dots, X_n) &= X_1 X_2 \cdots X_n\end{aligned}$$

Das Polynom σ_k ist die Summe von $\binom{n}{k}$ Monomen, jedes vom Grad k.

Elementarsymmetrische Polynome

Satz M3H: Wurzelsatz von Vieta

Vorgelegt sei ein normiertes Polynom $P \in K[X]_n^1$, einerseits als Summe von Monomen und andererseits als Produkt von Linearfaktoren:

$$P = X^n + a_1 X^{n-1} + \cdots + a_n = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n)$$

Dann gilt $a_k = (-1)^k \sigma_k(\lambda_1, \lambda_2, \dots, \lambda_n)$, also ausgeschrieben

$$a_k = (-1)^k \textstyle\sum_{i_1<i_2<\dots<i_k} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}.$$

Dies gilt insb. für das charakteristische Polynom und die Eigenwerte!

Aus den Nullstellen $\lambda_1, \dots, \lambda_n \in K$ von P können wir also ganz leicht die Koeffizienten $a_1, \dots, a_n \in K$ bestimmen: Ausmultiplizieren genügt.

Umgekehrt bestimmen die Koeffizienten $a_1, \dots, a_n \in K$ von P eindeutig die Nullstellen $\lambda_1, \dots, \lambda_n$ bis auf die Reihenfolge.

Die Faktorisierung ist wesentlich schwieriger als das Ausmultiplizieren!

Die Spur lesen: Der letzte Eigenwert ist gratis!

Aufgabe: Gegeben ist die Matrix $A \in \mathbb{R}^{4\times 4}$ mit vier Vektoren:

$$A = \begin{bmatrix} -18 & -50 & 10 & 10 \\ 4 & 11 & -1 & -3 \\ -12 & -32 & 10 & 4 \\ 2 & 2 & 2 & -2 \end{bmatrix}, \; b = \begin{bmatrix} 0 \\ 1 \\ 4 \\ 0 \end{bmatrix}, \; c = \begin{bmatrix} -3 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \; d = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \; e = \begin{bmatrix} -1 \\ 0 \\ -1 \\ -1 \end{bmatrix}$$

(1) Welche davon sind Eigenvektoren von A? Zu welchen Eigenwerten?
(2) Bestimmen Sie alle Eigenwerte, (3) Eigenräume, (4) eine Eigenbasis.

Lösung: (1) Wir setzen die Definition $Av = \lambda v$ ein und rechnen:

$$Ab = A \begin{bmatrix} 0 \\ 1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} -10 \\ 7 \\ 8 \\ 10 \end{bmatrix} \notin \mathbb{R}b, \qquad Ac = A \begin{bmatrix} -3 \\ 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -6 \\ 2 \\ 0 \\ -2 \end{bmatrix} = 2c,$$

$$Ad = A \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0d, \qquad Ae = A \begin{bmatrix} -1 \\ 0 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ -2 \\ -2 \end{bmatrix} = 2e$$

Die Spur lesen: Der letzte Eigenwert ist gratis!

☺ Teil (1) ist eine leichte Fingerübung: Um zu prüfen, ob $v \in K^n$ ein Eigenvektor der Matrix $A \in K^{n\times n}$ ist, genügt es, die Definition $Av = \lambda v$ einzusetzen… und im günstigen Fall den Eigenwert λ abzulesen.

⚠ Es lohnt sich, die wertvollen Informationen aus Teil (1) zu nutzen! Die hier angebotenen Daten sind nicht zufällig, sondern mit Bedacht vorbereitet; in einer Klausur sind sie eine gezielte Hilfestellung.
Auch außerhalb von Klausuren kann diese Situation entstehen.

☹ Alternativ, aber ungeschickt, kann man die hilfreichen Daten aus Teil (1) ignorieren und für (2–4) stur das Standardverfahren einsetzen: Das charakteristische Polynom entwickeln und seine Nullstellen finden, Eigenräume berechnen und eine Eigenbasis wählen. Das ist mühsam.

☺ Sie arbeiten umso effizienter, je genauer Sie verstehen, was Sie tun!
… insbesondere, was Sie schon haben und was Sie noch suchen.
Sie können dann geschickt vom Standardverfahren abweichen,
je nach konkretem Bedarf und möglichen Abkürzungen.

Die Spur lesen: Der letzte Eigenwert ist gratis!

$$A = \begin{bmatrix} -18 & -50 & 10 & 10 \\ 4 & 11 & -1 & -3 \\ -12 & -32 & 10 & 4 \\ 2 & 2 & 2 & -2 \end{bmatrix}, \; b = \begin{bmatrix} 0 \\ 1 \\ 4 \\ 0 \end{bmatrix}, \; c = \begin{bmatrix} -3 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \; d = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \; e = \begin{bmatrix} -1 \\ 0 \\ -1 \\ -1 \end{bmatrix}$$

(2) Dank (1) wissen wir $\operatorname{Eig}(A,0) \ge \langle\, d \,\rangle_{\mathbb{R}}^{!}$ und $\operatorname{Eig}(A,2) \ge \langle\, c, e \,\rangle_{\mathbb{R}}^{!}$.
Damit kennen wir drei der vier Eigenwerte: $\lambda_1 = 0$, $\lambda_2 = 2$, $\lambda_3 = 2$.
Dank $\operatorname{tr}(A) = 1$ bekommen wir den vierten gratis: $\lambda_4 = -3$.
Ohne weitere Mühe schließen wir $\chi_A = X(X-2)^2(X+3)$.

(3) Wir berechnen den Eigenraum $\operatorname{Eig}(A,-3)$ wie üblich mit Gauß:

$$A + 3I = \begin{bmatrix} -15 & -50 & 10 & 10 \\ 4 & 14 & -1 & -3 \\ -12 & -32 & 13 & 4 \\ 2 & 2 & 2 & 1 \end{bmatrix} \xrightarrow[\text{RZSF}]{\text{Gauß}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Demnach gilt $\operatorname{Eig}(A,-3) = \langle\, f \,\rangle_{\mathbb{R}}^{!}$ mit $f = (-2,1,0,2)^\intercal$.

(4) Damit haben wir zu A unsere Eigenbasis (d,c,e,f) gefunden.

Die Spur lesen: Der letzte Eigenwert ist gratis!

☺ Dank der Informationen aus (1) sind unsere Rechnungen (2–4) in dieser Anwendung wesentlich effizienter als das Standardverfahren!

Vielleicht erscheint Ihnen die obige Aufgabenstellung etwas künstlich, das will ich gar nicht bestreiten. Ganz unrealistisch ist das Vorgehen jedoch auch nicht, denn in Anwendungen verfügen Sie manchmal über Teilinformationen, und diese sollten Sie geschickt nutzen.

Das Standardverfahren M2I zur Diagonalisierung hat natürlich trotzdem seine Berechtigung: Wenn Sie keine weiteren Daten oder Ideen haben, dann bietet es Ihnen eine allgemeine und verlässliche Methode, um das Diagonalisierungsproblem zu lösen.

Variantenreiche Beispiele wie das vorige zeigen Ihnen jedoch ebenso, dass Sie sich nicht stur auf vorgefertigte Rezepte verlassen sollten. Manchmal sind geschickte Abkürzungen oder raffinierte Tricks möglich. Diese sollten Sie erkennen, um sich unnötige Arbeit zu ersparen.

☺ Denken hilft!

Das lokale Minimalpolynom

Sei $f : V \to V$ linear über dem Körper K und $\dim_K(V) = n < \infty$.
(1) Jeder Vektor $v \in V$ erzeugt seinen **f–zyklischen Unterraum**

$$Z := \langle\, f^k(v) \mid k \in \mathbb{N} \,\rangle_K \le V.$$

Wegen $\dim_K(Z) \le n$ sind $f^0(v), f^1(v), \dots, f^n(v)$ linear abhängig (J2K).
(2) Sei $m \in \mathbb{N}$ minimal, sodass $f^0(v), \dots, f^m(v)$ linear abhängig sind.
Somit ist die Familie $f^0(v), \dots, f^{m-1}(v)$ noch linear unabhängig,
und wir erhalten $a_0 f^0(v) + \ldots + a_{m-1} f^{m-1}(v) + f^m(v) = 0$
mit eindeutigen Koeffizienten $a_0, \dots, a_{m-1} \in K$.

Definition M3I: das lokale Minimalpolynom

Wir nennen $\mu = \mu_f^v := a_0 + a_1 X + \dots + a_{m-1} X^{m-1} + X^m \in K[X]$ das **lokale Minimalpolynom** von $f : V \to V$ bezüglich des Vektors $v \in V$.

Das Polynom $P = \mu_f^v$ ist normiert und annulliert v gemäß $P(f)(v) = 0$.
Unter allen Polynomen in $K[X]$ mit diesen beiden Eigenschaften hat es den kleinsten Grad, daher der Name **Minimalpolynom** von f bzgl. v.

Das lokale Minimalpolynom

Allgemein gilt $f^0 = \mathrm{id}_V$ und $f^1 = f$ sowie rekursiv $f^{k+1} = f \circ f^k$.
Dies sind die Potenzen von $f : V \to V$ im Endomorphismenring

$$(\mathrm{End}_K(V), +, 0, \circ, \mathrm{id}_V).$$

Für $v \in V$ gilt $f^0(v) = v$ und $f^1(v) = f(v)$ sowie $f^{k+1}(v) = f(f^k(v))$.
Die Familie $f^0(v), f^1(v), f^2(v), \dots$ in V entsteht demnach aus dem Startvektor $v \in V$ durch wiederholte Anwendung von $f : V \to V$.

🙂 Die obige Definition M3I erklärt zugleich einen Algorithmus:

Algo M3I: Berechnung des lokalen Minimalpolynoms

Eingabe: ein Endomorphismus $f : V \to V$ und ein Vektor $v \in V$
Ausgabe: das lokale Minimalpolynom $\mu_f^v \in K[X]$

```
Wähle eine Basis von $V$ und schreibe $f^0(v), f^1(v), \dots, f^m(v)$ als
    Spalten in eine Matrix $M$ bis der Rang stagniert bei $\operatorname{rang} M = m$.
Bestimme den Kern $\ker M = \langle (a_0, a_1, \dots, a_{m-1}, 1)^\intercal \rangle$.
return $\mu_f^v = a_0 + a_1 X + \dots + a_{m-1} X^{m-1} + X^m$
```

Der Satz von Cayley–Hamilton

Lemma M3J: Cayley–Hamilton

Das lokale Minimalpolynom μ_f^v teilt das charakteristische Polynom χ_f.

Aufgabe: Beweisen Sie dieses Lemma nach folgender Anleitung:
(1) Berechnen Sie zu $g = f|_Z^Z$ das charakteristische Polynom χ_g.
(2) Vergleichen Sie dies mit dem charakteristischen Polynom χ_f.

Lösung: (1) Wir nutzen weiterhin die obigen Bezeichnungen (M3I).
Wir haben $Z = \langle\, f^k(v) \mid k < m \,\rangle_K^! \le V$ mit der Basis $\mathcal{B} = (f^k(v))_{k=0}^{m-1}$.
Nach Konstruktion gilt $f(Z) \subseteq Z$, das heißt, der Raum Z ist f–invariant.
Die Einschränkung $g = f|_Z^Z$ wird dargestellt durch die **Begleitmatrix**

$$B := \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(g) = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -a_{m-1} \end{bmatrix} =: \mathrm{C}(\mu_f^v).$$

Das charakteristische Polynom ist $\chi_g = \det(XI - B) = \mu_f^v$ dank M2Q.

Der Satz von Cayley–Hamilton

(2) Wir ergänzen die Basis $\mathcal{B}$ von Z zu einer Basis $\mathcal{A}$ von V.
Bezüglich dieser Basis wird $f : V \to V$ dargestellt durch die Matrix

$$A := \mathrm{M}_{\mathcal{A}}^{\mathcal{A}}(f) = \begin{bmatrix} B & * \\ 0 & C \end{bmatrix}.$$

Für diese Block-Dreiecksmatrix gilt (dank L2V)

$$\chi_f = \det(XI - A) = \det(XI - B)\det(XI - C) = \chi_g \cdot Q.$$

In Worten ausformuliert ist das genau die Behauptung: Das lokale Minimalpolynom $\mu_f^v = \chi_g$ teilt das charakteristische Polynom χ_f. $\boxed{\text{QED}}$

Satz M3K: Cayley–Hamilton

Sei $f \in \mathrm{End}_K(V)$ und $\dim_K(V) < \infty$. Dann gilt $\chi_f(f) = 0$ in $\mathrm{End}_K(V)$:
Das charakteristische Polynom von f annulliert den Endomorphismus f.

Beweis: Sei $v \in V$. Dank obigem Lemma gilt $\chi_f = Q \cdot \mu_f^v$ mit $Q \in K[X]$,
also $\chi_f(f)(v) = Q(f)\mu_f^v(f)(v) = 0$. Wir schließen $\chi_f(f) = 0$. $\boxed{\text{QED}}$

Cayley–Hamilton für Matrizen: falscher Beweis

Der Satz M3K von Cayley–Hamilton für Endomorphismen ist äquivalent zu folgender Formulierung für quadratische Matrizen:

Satz M3K: Cayley–Hamilton für Matrizen

Für jede Matrix $A \in K^{n\times n}$ und $\chi_A \in K[X]^1_n$ gilt $\chi_A(A) = 0$ in $K^{n\times n}$: Das charakteristische Polynom von A annulliert die Matrix A.

⚠ Verlockend simpel, aber unsinnig, ist folgendes Scheinargument: Das charakteristische Polynom ist definiert als die Determinante

$$\chi_A(X) := \det(XI - A) \in K[X].$$

Wenn wir formal X durch A ersetzen, so erhalten wir scheinbar

$$\chi_A(A) \overset{?}{=} \det(AI - A) = \det(A - A) = \det(0) = 0.$$

Aufgabe: Bitte schützen Sie sich gegen solcherart (Selbst)Betrug! Warum ist das kein Beweis von Cayley–Hamilton, sondern Unsinn?

Cayley–Hamilton für Matrizen: falscher Beweis

Lösung: Links ist $\chi_A(A) \in K^{n\times n}$ eine Matrix: Diese entsteht, indem wir in das Polynom $\chi_A(X) = X^n - a_{n-1}X^{n-1} \pm \dots + (-1)^n a_0$ die Matrix einsetzen, ausgeschrieben $\chi_A(A) = A^n - a_{n-1}A^{n-1} \pm \dots + (-1)^n a_0 A^0$. Rechts hingegen ist $\det(AI - A) = \det(0) = 0 \in K$ ein Skalar.

Das obige Scheinargument behauptet, naiv oder frech, nach obskurer Rechnung „Matrix = Skalar“ und ist deshalb offensichtlich unsinnig.

⚠ Schon die rein syntaktische Typen-Prüfung schlägt hier fehl! Auf der linken und der rechten Seite stehen verschiedene Objekte. Die frech-naiv behauptete Gleichung „$\overset{?}{=}$“ gilt ganz sicher nicht.

☺ Erstaunlicherweise lässt sich dieses unsinnige Scheinargument reparieren zu einem Beweis (nach S. Lang: *Algebra.* 2002, §XIV.3):

Aufgabe: (1) Was erhalten Sie tatsächlich, wenn Sie in der Matrix $C(X) = (XI - A)^\intercal$ die Variable X durch die Matrix A ersetzen?
(2) Multiplizieren Sie die Matrix $C(A)$ mit dem Vektor $(e_1, \dots, e_n)^\intercal$.
(3) Multiplikation mit der Adjunkten $B(A) = \operatorname{adj} C(A)$ zeigt $\chi_A(A) = 0$.

Cayley–Hamilton für Matrizen: reparierter Beweis

Lösung: (1) Wir betrachten die Matrix $C(X) = (XI - A)^\intercal$, also

$$C(X) = \begin{bmatrix} X - a_{11} & \dots & -a_{n1} \\ \vdots & & \vdots \\ -a_{1n} & \dots & X - a_{nn} \end{bmatrix} \in K[X]^{n\times n}.$$

Wir ersetzen die Variable X korrekt durch die Matrix A und erhalten

$$C(A) = \begin{bmatrix} A - a_{11}I & \dots & -a_{n1}I \\ \vdots & & \vdots \\ -a_{1n}I & \dots & A - a_{nn}I \end{bmatrix} \in K[A]^{n\times n}.$$

(2) Angewendet auf die Standardbasis $e_1, \dots, e_n \in K^n$ gilt:

$$C(A) \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} Ae_1 - a_{11}e_1 - \dots - a_{n1}e_n \\ \vdots \\ -a_{1n}e_1 - \dots + Ae_n - a_{nn}e_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Wir arbeiten hier über dem kommutativen Ring $K[A] \le K^{n\times n}$. Der Ring $K[A]$ operiert von links auf dem Raum $K^n = K^{n\times 1}$.

Cayley–Hamilton für Matrizen: reparierter Beweis

(3) Aus der Gleichung (2) folgern wir nun $\det C(A) = 0$. Für die adjunkte Matrix $B(X) = \operatorname{adj} C(X)$ gilt dank L2S:

$$B(X) \cdot C(X) = \det C(X) \cdot I = \begin{bmatrix} \chi_A(X) & & 0 \\ & \ddots & \\ 0 & & \chi_A(X) \end{bmatrix} \in K[X]^{n\times n}$$

Wir ersetzen wieder X durch A und erhalten:

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = C(A) \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = B(A) \cdot C(A) \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} \chi_A(A) & & 0 \\ & \ddots & \\ 0 & & \chi_A(A) \end{bmatrix} \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

Somit gilt $\chi_A(A)\, e_i = 0$ in K^n für alle i, also $\chi_A(A) = 0$ in $K^{n\times n}$. $\boxed{\text{QED}}$

☺ Diese Rechnung folgt der naiven Eingebung „ersetze X durch A“. Die korrekte Ausführung erfordert allerdings extrem genaue Notation.

☺ Determinante und Adjunkte haben wir über jedem kommutativen Ring, insbesondere also über unserem Matrixunterring $K[A] \le K^{n\times n}$.

Das (globale) Minimalpolynom

Sei $f : V \to V$ linear über dem Körper K und $\dim_K(V) = n < \infty$.

(1) Der Endomorphismenring $\operatorname{End}_K(V) \cong K^{n\times n}$ ist ein K–Vektorraum der Dimension $d = n^2$ (K1I). Demnach sind in $\operatorname{End}_K(V)$ die Potenzen $f^0, f^1, f^2, \dots, f^d$ linear abhängig über K (J2K).

(2) Sei $m \in \mathbb{N}$ minimal, sodass $f^0, f_1, \dots, f^m$ linear abhängig sind. Somit ist die Familie $f^0, f_1, \dots, f^{m-1}$ noch linear unabhängig, und wir erhalten $a_0 f^0 + a_1 f^1 + \ldots + a_{m-1} f^{m-1} + f^m = 0$ mit eindeutigen Koeffizienten $a_0, a_1, \dots, a_{m-1} \in K$.

Definition M3L: das (globale) Minimalpolynom

Wir nennen $\mu = \mu_f := a_0 + a_1 X + \dots + a_{m-1} X^{m-1} + X^m \in K[X]$ das **(globale) Minimalpolynom** des Endomorphismus $f : V \to V$.

Das Polynom $P = \mu_f$ ist normiert und annulliert f gemäß $P(f) = 0$. Unter allen Polynomen in $K[X]$ mit diesen beiden Eigenschaften hat es den kleinsten Grad, daher der Name **Minimalpolynom**.

☺ Diese Definition erklärt zugleich einen Algorithmus!

Das (globale) Minimalpolynom

Algo M3L: Berechnung des (globalen) Minimalpolynoms

Eingabe: ein Endomorphismus $f : V \to V$, wobei $\dim V = n < \infty$
Ausgabe: das (globale) Minimalpolynom $\mu_f \in K[X]$

```
Wähle eine Basis $V \cong K^n$ und somit $\operatorname{End}_K(V) \cong K^{n^2}$.
Schreibe $f^0, f^1(v), \dots, f^m(v) \in \operatorname{End}_K(V)$ so als Spalten
    in eine Matrix $M$ bis der Rang stagniert bei $\operatorname{rang} M = m$.
Bestimme den Kern $\ker M = \langle (a_0, a_1, \dots, a_{m-1}, 1)^\intercal \rangle$.
return $\mu_f = a_0 + a_1 X + \dots + a_{m-1} X^{m-1} + X^m$
```

Bemerkung: Der Gauß–Algorithmus zeigt: Das Minimalpolynom ändert sich nicht, wenn wir es in einem Erweiterungskörper $\bar{K} \geq K$ berechnen.

☹ Die Spalten haben hier die Länge n^2; das ist spürbar aufwändiger als die Dimension $n = \dim_K V$ beim lokalen Minimalpolynom (M3I).

Für den Polynomgrad haben wir zunächst die grobe Schranke $m \leq n^2$.

☺ Der folgende Satz M3O zeigt $\mu_f \mid \chi_f$ und reduziert dies zu $m \leq n$. Diese Schranke ist eine enorme Verbesserung und scharf (M3M, M3N).

Zahlenbeispiele zum Minimalpolynom

Aufgabe: Berechnen Sie das Minimalpolynom zu

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \text{und } B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{und } C = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}.$$

Lösung: Wir finden direkt nach Definition:

$$A^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ A^1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \Longrightarrow \ \mu_A = X - 2$$

$$B^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ B^1 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \ B^2 = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix} \qquad \Longrightarrow \ \mu_B = X^2 - 4X + 4$$

$$C^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ C^1 = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}, \ C^2 = \begin{bmatrix} 7 & 4 \\ 12 & 7 \end{bmatrix} \qquad \Longrightarrow \ \mu_C = X^2 - 4X + 1$$

Für 2×2–Matrizen wie A, B, C sind die Rechnungen allzu einfach, dennoch können wir hieran schön die Definition M3L anwenden und den zugehörigen Algorithmus M3L als allgemeines Verfahren illustrieren.

Zahlenbeispiele zum Minimalpolynom

Um das Minimalpolynom von $C \in K^{n\times n}$ zu bestimmen, suchen wir die erste nicht-triviale Relation der Potenzen $C^0, C^1, C^2, \dots \in K^{n\times n}$.

Hierzu füllen wir die Koeffizienten der ersten $m+1$ Matrizen $C^0, \dots, C^m$ spaltenweise in die Matrix M bis der Rang stagniert bei $\operatorname{rang} M = m$:

$$M = \begin{bmatrix} 1 & 2 & 7 \\ 0 & 3 & 12 \\ 0 & 1 & 4 \\ 1 & 2 & 7 \end{bmatrix} \quad \Longrightarrow \quad \ker M = K \begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad \mu_C = X^2 - 4X + 1$$

Allgemein nutzen wir Gauß $M \to SM$ zur RZFS. Gilt $\operatorname{rang} M = m+1$, so erweitere zu $M' = (M, v) \to (SM, Sv)$, reduziere $M' \to S'M'$, usw.

Für $C \in K^{n\times n}$ erwarten wir $M \in K^{n^2 \times m}$ mit $m \leq n^2$, besser nur $m \leq n$ dank M3O. Für große n ist das aufwändig, aber es ist immer möglich.

Die folgenden Beispiele zeigen willkommene Vereinfachungen für Jordan–Blöcke M3M, Diagonalmatrizen M3N und Begleitmatrizen M3P.

Bemerkung: Das Minimalpolynom ist invariant unter Ähnlichkeit (M3Q).

Strukturbeispiel: ein Jordan–Block

Beispiel M3M: ein Jordan–Block

In Dimension $n \in \mathbb{N}$ betrachten wir den nilpotenten Jordan–Block

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in K^{n \times n}.$$

In diesem Falle gilt $\mu_B = X^n$, und dies stimmt mit $\chi_B = X^n$ überein.

Beweis: (a) Das Polynom $P = X^n$ annulliert B, denn es gilt $B^n = 0$.
(b) Für jedes Polynom $Q = a_0 + \dots + a_{n-1}X^{n-1}$ kleineren Grades gilt:

$$Q(B) = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ 0 & a_0 & \ddots & \vdots \\ 0 & 0 & \ddots & a_1 \\ 0 & 0 & 0 & a_0 \end{bmatrix}$$

Demnach ist $Q(B) = 0$ und $\deg Q < n$ nur für $Q = 0$ möglich.
Somit ist $P = X^n$ tatsächlich das Minimalpolynom von B. QED

Strukturbeispiel: ein diagonalisierbarer Endomorphismus

Beispiel M3N: ein diagonalisierbarer Endomorphismus

Sei $f : V \to V$ diagonalisierbar mit charakteristischem Polynom

$$\chi_f = (X - \lambda_1)^{r_1} \cdots (X - \lambda_k)^{r_k}$$

wobei $\lambda_i \neq \lambda_j$ für $i \neq j$. Für das Minimalpolynom gilt dann

$$\mu_f = (X - \lambda_1) \cdots (X - \lambda_k).$$

Beweis: (a) Für $P = (X - \lambda_1) \cdots (X - \lambda_k)$ zeigen wir $P(f) = 0$, also $P(f)(v) = 0$ für alle $v \in V$. Wir haben $V = \bigoplus_{i=1}^{k} \operatorname{Eig}(f, \lambda_i)$.
Es genügt also, Eigenvektoren $v_i \in \operatorname{Eig}(f, \lambda_i)$ zu betrachten:

$$P(f)(v_i) = P(\lambda_i)\, v_i = 0$$

(b) Jedes Polynom $Q \neq 0$ vom Grad $d < k$ hat höchstens d Nullstellen dank Satz G3K. Demnach gilt $Q(\lambda_i) \neq 0$ für ein $i \in \{1, \dots, k\}$, also

$$Q(f)(v_i) = Q(\lambda_i)\, v_i \neq 0.$$

Somit ist P tatsächlich das Minimalpolynom von f. QED

Das Minimalpolynom teilt das charakteristische Polynom.

🙂 Minimalität bezüglich Teilbarkeit und Grad stimmen überein:

Satz M3O: μ_f^v teilt μ_f teilt χ_f, kurz $\mu_f^v \mid \mu_f \mid \chi_f$

Sei $f : V \to V$ eine K–lineare Abbildung und $\dim_K V = n < \infty$.
Sei $v \in V$ ein Vektor und $P \in K[X]$ ein Polynom über K.

(1) Genau dann gilt $P(f) = 0$, wenn P ein Vielfaches von μ_f ist.

$$P(f) = 0 \quad \Longleftrightarrow \quad \exists Q \in K[X] \,:\, P = Q \cdot \mu_f$$

(2) Das Minimalpolynom μ_f teilt das charakteristische Polynom χ_f.
Für den Polynomgrad gilt insbesondere $m = \deg \mu_f \le \deg \chi_f = n$.

(3) Genau dann gilt $P(f)(v) = 0$, wenn P ein Vielfaches von μ_f^v ist.

$$P(f)(v) = 0 \quad \Longleftrightarrow \quad \exists Q \in K[X] \,:\, P = Q \cdot \mu_f^v$$

(4) Das globale Minimalpolynom μ_f ist das kleinste gemeinsame Vielfache im Ring $K[X]$ aller lokalen Minimalpolynome μ_f^v für $v \in V$.
Zur Berechnung genügt ein $K[f]$–Erzeugendensystem $(v_i)_{i \in I}$ von V.

Das Minimalpolynom teilt das charakteristische Polynom.

Beweis: (1) Sei $P(f) = 0$. Euklidische Division (G3H) von P durch μ_f ergibt $P = \mu_f \cdot Q + R$ mit $Q, R \in K[X]$ und $\deg R < \deg \mu_f = m$.
Daraus folgt $0 = P(f) = \mu_f(f)Q(f) + R(f) = R(f)$, da $\mu_f(f) = 0$.
Da $\mu_f \in K[X]^1_m$ minimalen Grad hat, bleibt nur $R = 0$.

(2) Dank Cayley–Hamilton (M3K) gilt $\chi_f(f) = 0$. Dank (1) folgt $\mu_f \mid \chi_f$.

(3) Sei $P(f)(v) = 0$. Wie in (1) folgt $\mu_f^v \mid P$ durch euklidische Division.

(4a) Aus $\mu_f(f) = 0$ folgt $\mu_f(f)(v_i) = 0$ für $i \in I$, dank (3) also $\mu_f^{v_i} \mid \mu_f$.
Das bedeutet, μ_f ist ein gemeinsames Vielfaches von $(\mu_f^v)_{v \in V}$.

(4b) Sei $P \in K[X]$ mit $\mu_f^{v_i} \mid P$ für alle $i \in I$. Dank (3) gilt $P(f)(v_i) = 0$, somit $P(f)(v) = 0$ für alle $v \in V$, also $P(f) = 0$. Dank (1) folgt $\mu_f \mid P$.

Somit ist μ_f ein kleinstes gemeinsames Vielfaches von $(\mu_f^{v_i})_{i \in I}$. QED

🙂 Das Minimalpolynom μ_f ist zudem normiert, wir sprechen daher von *dem* (normierten) kgV der lokalen Minimalpolynome $(\mu_f^{v_i})_{i \in I}$.

Strukturbeispiel: eine Begleitmatrix

Beispiel M3P: eine Begleitmatrix

(1) Sei $P = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0 \in K[X]^1_n$ ein normiertes Polynom und $C = C(P)$ seine Begleitmatrix (M2Q). Dann gilt $\mu_C = \chi_C$.

(2) Sei $f : V \to V$ linear über K mit $\dim_K V = n < \infty$. Ist V zudem f–zyklisch, also $V = \langle\, f^k(v) \mid k < n \,\rangle^!_K$ für ein $v \in V$, so folgt $\mu_f = \chi_f$.

Aufgabe: Beweisen Sie dies nach dem Vorbild von Lemma M3J.

Lösung: (2) Bezüglich der Basis $\mathcal{B} = (f^k(v))_{k=0}^{n-1}$ von V haben wir

$$B := \mathrm{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{bmatrix} =: \mathrm{C}(P).$$

(a) Das charakteristische Polynom ist $\chi_f = \det(XI - B) = P$ dank M2Q.
(b) Das lokale Minimalpolynom von f bezüglich v ist ebenfalls $\mu_f^v = P$.
(c) Dank M3O gilt $\mu_f^v \mid \mu_f \mid \chi_f$, mit (a) und (b) folgt $\mu_f^v = \mu_f = \chi_f = P$.

Das Minimalpolynom ist invariant unter Ähnlichkeit.

Lemma M3Q: Das Minimalpolynom ist invariant unter Ähnlichkeit.

(1) Für jedes Polynom $P \in K[X]$ sowie $A \in K^{n\times n}$ und $T \in \mathrm{GL}_n K$ gilt:

$$\boxed{P\big(T^{-1}AT\big) = T^{-1}\,P(A)\,T}$$

(2) Aus $P(A) = 0$ folgt $P(B) = 0$ für alle ähnlichen Matrizen $B \sim A$.

(3) Das Minimalpolynom ist invariant unter Ähnlichkeit, also $\mu_A = \mu_B$.

Beweis: (1) Wir setzen die konjugierte Matrix $B = T^{-1}AT$ in unser Polynom $P(X) = \sum_{i=0}^n p_iX^i$ ein und erhalten:

$$\begin{aligned} P(T^{-1}AT) &= \textstyle\sum_{i=0}^n p_i(T^{-1}AT)^i = \sum_{i=0}^n p_i(T^{-1}A^iT) \\ &= \textstyle\sum_{i=0}^n T^{-1}(p_iA^i)T = T^{-1}(\sum_{i=0}^n p_iA^i)T = T^{-1}P(A)T \end{aligned}$$

Daraus folgt (2) und daraus wiederum (3) nach Definition M3L. $\boxed{\text{QED}}$

☺ Der Algorithmus M3L ist demnach unabhängig von der Basiswahl. Die Definition M3L ist ohnehin schon basisunabhängig formuliert.

Die Nullstellen des Minimalpolynoms

Satz M3R: Nullstellen des Minimalpolynoms

Sei $f : V \to V$ eine K–lineare Abbildung und $\dim_K V = n < \infty$.

(1) Falls χ_f über K in Linearfaktoren zerfällt, so gilt

$$\boxed{\begin{array}{ll} \chi_f = (X-\lambda_1)^{r_1}\cdots(X-\lambda_k)^{r_k} & \text{mit } \lambda_i \ne \lambda_j \text{ für } i \ne j \text{ und} \\ \mu_f = (X-\lambda_1)^{s_1}\cdots(X-\lambda_k)^{s_k} & \text{mit } 1 \le s_i \le r_i \text{ für alle } i. \end{array}}$$

(2) Das Minimalpolynom μ_f teilt das charakteristische Polynom χ_f, und umgekehrt teilt χ_f eine hinreichend hohe Potenz $(\mu_f)^a$ mit $a \in \mathbb{N}$.
Beide Polynome haben dieselben Nullstellen: die Eigenwerte von f.

(3) Allgemein gilt demnach

$$\begin{aligned} \chi_f &= (X-\lambda_1)^{r_1}\cdots(X-\lambda_k)^{r_k}\,Q, \\ \mu_f &= (X-\lambda_1)^{s_1}\cdots(X-\lambda_k)^{s_k}\,P, \end{aligned}$$

wobei $P, Q \in K[X]$ keine Nullstellen in K haben und $P \mid Q \mid P^a$ erfüllen.

Die Nullstellen des Minimalpolynoms

Beweis: (1) Dank M3O gilt $\mu_f \mid \chi_f$. Daraus folgt $s_i \le r_i$ für alle i.
Zu jedem Eigenwert λ_i existiert ein Eigenvektor $v_i \in \mathrm{Eig}(f, \lambda_i) \smallsetminus \{0\}$.
Aus $v_i \ne 0$ und $f(v_i) = \lambda_iv_i$ erhalten wir das lokale Minimalpolynom $\mu_f^{v_i} = X - \lambda_i$. Dank M3O gilt $\mu_f^{v_i} \mid \mu_f$, also tatsächlich $s_i \ge 1$.

Aus (1) folgt (2), falls das charakteristische Polynom χ_f über K zerfällt.
Falls χ_f über K nicht zerfällt, so behelfen wir uns mit folgendem Trick:

Allgemein betrachten wir ohne Einschränkung $f : K^n \to K^n : x \mapsto Ax$.
Es existiert eine Körpererweiterung $\bar K \ge K$, sodass χ_f über $\bar K$ zerfällt.
Wir betrachten nun die lineare Abbildung $\bar f : \bar K^n \to \bar K^n : x \mapsto Ax$ über $\bar K$.
Für diese gilt Aussage (1) und somit Teilbarkeit (2) $\mu_{\bar f} \mid \chi_{\bar f} \mid \mu_{\bar f}^a$ in $\bar K[X]$.
Es gilt $\chi_f = \chi_{\bar f}$ und $\mu_f = \mu_{\bar f}$ in $K[X] \le \bar K[X]$, also $\mu_f \mid \chi_f \mid \mu_f^a$ in $K[X]$.
(Für χ ist das klar, für μ siehe die Bemerkung nach Algorithmus M3L.)
Somit gilt gegenseitige Teilbarkeit (2) tatsächlich über jedem Körper K.

Aus der Teilbarkeit (2) folgt sofort die Aussage (3). $\boxed{\text{QED}}$

Die Nullstellen des Minimalpolynoms

Aufgabe: Beweisen Sie Satz M3R(2) ohne Erweiterungskörper:

$$\boxed{\text{Es gilt } \mu_f \mid \chi_f \mid (\mu_f)^a \text{ für ein } a \in \mathbb{N}.}$$

Anleitung: Führen Sie Induktion über die Dimension $n = \dim_K V$ und folgen Sie dem Vorbild von Lemma M3J zu Cayley–Hamilton.

🙂 Warum sollten wir einen Satz wie M3R zweimal beweisen?
Im vorliegenden Fall ist der erste Beweis elegant-schön-trickreich, beruht jedoch auf einem recht starken Hilfsmittel: der Existenz eines Erweiterungskörpers $\bar{K} \geq K$, über dem χ_f in Linearfaktoren zerfällt.

🙂 Der nachfolgende Beweis hingegen kommt ohne diesen Trick aus, das ist zwar etwas länger, dafür aber elementar – und eine gute Übung!

Lösung: Für $n = 0$ gilt $\chi_f = \mu_f = 1$. Für $n \geq 1$ existiert $v \in V \smallsetminus \{0\}$.
Wir betrachten den f–zyklischen Unterraum $Z := \langle\, f^k(v) \mid k \in \mathbb{N} \,\rangle_K$.
Für $m := \dim_K Z$ gilt $1 \leq m \leq n$ und $Z = \langle\, f^k(v) \mid k < m \,\rangle_K^!$,
das heißt, die Familie $\mathcal{B} = (f^k(v))_{k=0}^{m-1}$ ist eine Basis von Z.

Die Nullstellen des Minimalpolynoms

Die Einschränkung $g = f|_Z^Z$ wird dargestellt durch die **Begleitmatrix**

$$B := \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(g) = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -a_{m-1} \end{bmatrix}.$$

Hier wissen wir bereits $\mu_B = \chi_B$ aus Beispiel M3P, also $\mu_B \mid \chi_B \mid \mu_B^1$.
Wir ergänzen die Basis $\mathcal{B}$ von Z zu einer Basis $\mathcal{A}$ von V und erhalten

$$A := \mathrm{M}_{\mathcal{A}}^{\mathcal{A}}(f) = \begin{bmatrix} B & * \\ 0 & C \end{bmatrix}.$$

Nach Induktionsvoraussetzung gilt hier $\mu_C \mid \chi_C \mid \mu_C^\gamma$ für ein $\gamma \in \mathbb{N}$.
Das Minimalpolynom μ_A annulliert B und C, also $\mu_B \mid \mu_A$ und $\mu_C \mid \mu_A$.
Für das charakteristische Polynom folgt $\chi_A = \chi_B \cdot \chi_C \mid \mu_B \cdot \mu_C^\gamma \mid \mu_A^{\gamma+1}$.
Die grundlegende Teilbarkeit $\mu_A \mid \chi_A$ verdanken wir Satz M3O.
Gleiches gilt für den Endomorphismus f, also $\mu_f \mid \chi_f \mid \mu_f^{\gamma+1}$.

Minimalpolynom in un/endlicher Dimension

🙂 Die folgenden Aussagen erfordern keine endliche Dimension.
In günstigen Fällen kann dies unsere Rechnungen vereinfachen.

Satz M3S: Minimalpolynom in un/endlicher Dimension

Sei $f : V \to V$ eine lineare Abbildung über dem Körper K.
Angenommen wir finden ein Polynom $P \in K[X]_n^1$ mit $P(f) = 0$.

(1) Das Minimalpolynom $\mu_f \in K[X]_m^1$ existiert und erfüllt $\mu_f \mid P$.

(2) Weiterhin ist μ_f das kgV aller lokalen Minimalpolynome μ_f^v.

(3) Jeder Eigenwert von f ist eine Nullstelle des Polynoms P.

(4) Das Spektrum ist $\sigma(f; K) = \{\, \lambda \in K \mid \mu_f(\lambda) = 0 \,\}$.

(1) Im Falle $\dim_K V = n < \infty$ erfüllt das charakteristische Polynom $P = \chi_f \in K[X]_n^1$ diese Rolle, denn $\chi_f(f) = 0$ dank Cayley–Hamilton.

(3) Nicht alle Nullstellen von P müssen tatsächlich Eigenwerte sein, doch die Kenntnis von P schränkt die möglichen Kandidaten stark ein.

(4) Beim Minimalpolynom hingegen ist jede Nullstelle ein Eigenwert.

Minimalpolynom in un/endlicher Dimension

Beweis: (1) Existenz und Eindeutigkeit folgen wie zuvor in M3L.

(2) Satz M3O(3,4) gilt weiterhin, unabhängig von der Dimension.

(1) Angenommen $v \in V \smallsetminus \{0\}$ und $\lambda \in K$ erfüllen $f(v) = \lambda v$.
Aus $P(f) = 0$ folgt $0 = P(f)(v) = P(\lambda)\, v$, also $P(\lambda) = 0$.

(2) Die Inklusion „$\subseteq$“ folgt aus (1), „$\supseteq$“ folgt aus M3O. QED

Aufgabe: Sei $A \in K^{n\times n}$ und $\{\chi_A, \mu_A\} \ni P = a_0 + a_1 X + \dots + a_m X^m$.
Genau dann ist A in $K^{n\times n}$ invertierbar, wenn a_0 in K invertierbar ist; in diesem Falle gilt $A^{-1} = B := -a_0^{-1}(a_1 + a_2 A + \dots + a_m A^{m-1})$.

Lösung: Wir haben folgende Äquivalenzen:

$$A \in \mathrm{GL}_n K \overset{\text{B2D}}{\iff} \ker A = \{0\} \overset{\text{M1B}}{\iff} 0 \notin \sigma(f) \overset{\text{M3R}}{\iff} a_0 \neq 0$$

Im Falle $a_0 \neq 0$ gilt $AB = I - a_0^{-1} P(A) = I$ dank $P(A) = 0$ (M3K, M3L).

🙂 Wir können so die Inverse A^{-1} als ein Polynom in A ausdrücken.
Hierzu genügt jedes annullierende Polynom P mit $P(A) = 0$ und $a_0 \neq 0$.

Anwendungsbeispiel: Nilpotenz und Minimalpolynom

Sei V ein Vektorraum. Ein Endomorphismus $f : V \to V$ heißt **nilpotent**, falls $f^k = 0$ für einen (hinreichend großen) Exponenten $k \in \mathbb{N}$ gilt. Der kleinste Exponent $m \in \mathbb{N}$ mit $f^m = 0$ heißt **Nilpotenzindex** von f.

Beispiel M3T: ein nilpotenter Endomorphismus

Für $f \in \operatorname{End}_K(V)$ und $\dim_K(V) = n < \infty$ gilt:

$$
\begin{aligned}
f \text{ ist nilpotent} \quad &\iff \quad \mu_f = X^m \quad \iff \quad \chi_f = X^n \\
&\qquad\qquad\qquad\quad\;\; \implies \quad \operatorname{tr}(f) = \det(f) = 0
\end{aligned}
$$

Die letzte Bedingung ist notwendig, aber für $n \ge 3$ nicht hinreichend. Speziell in Dimension $n = 2$ gilt $\chi_f = X^2 - \operatorname{tr}(f)X + \det(f)$, die Bedingung $\operatorname{tr}(f) = \det(f) = 0$ ist hier auch hinreichend.

Beweis: „$\Rightarrow$“: Ist f nilpotent, so gilt $f^k = 0$ für einen Exponenten $k \in \mathbb{N}$. Dank M3O gilt $\mu_f \mid X^k$, also $\mu_f = X^m$ mit dem Nilpotenzindex $m \le k$. Dank M3R gilt $\chi_f \mid (\mu_f)^a$ für ein $a \in \mathbb{N}$, und somit folgt $\chi_f = X^n$.

„$\Leftarrow$“: Die Umkehrung ist klar, dank Cayley–Hamilton (M3K). $\boxed{\text{QED}}$

Anwendungsbeispiel: Nilpotenz und Minimalpolynom

Aufgabe: Sei $f : V \to V$ ein nilpotenter Endomorphismus. Unter welcher Bedingung ist f diagonalisierbar?

Lösung: Nilpotenz bedeutet $\mu_f = X^m$ mit $m \in \mathbb{N}$ dank M3T. Diagonalisierbarkeit ist dann äquivalent zu $m \le 1$ nach M3N Das bedeutet $m = 0$: $V = \{0\}$, oder $m = 1$: $V \ne \{0\}$ und $f = 0$. Jeder nilpotente Endomorphismus $f \ne 0$ ist nicht diagonalisierbar.

Aufgabe: Sei $f \in \operatorname{End}_K(V)$ nilpotent vom Index m. Zeigen Sie (ohne M3T), dass $f^0, \dots, f^{m-1}$ über K linear unabhängig sind.

Lösung: Wir wissen $f^m = 0$ und $f^{m-1} \ne 0$. Gegeben sei $0 \le r < m$ und Koeffizienten $a_r, \dots, a_{m-1} \in K$ mit $a_r f^r + \dots + a_{m-1} f^{m-1} = 0$. Komposition mit f^{m-r-1} ergibt $a_r f^{m-1} = 0$, also $a_r = 0$. So fortfahrend schließen wir $a_r = \dots = a_{m-1} = 0$.

Anwendungsbeispiel: Idempotenz und Eigenraumzerlegung

Beispiel M3U: Idempotenz und Eigenraumzerlegung

Sei $f : V \to V$ eine K–lineare Abbildung mit $f^2 = f$.

Das Minimalpolynom μ_f teilt demnach $X^2 - X = X(X - 1)$:

0. Im Falle $\mu_f = 1$ ist $V = 0$ der Nullraum.
1. Im Falle $\mu_f = X$ ist $f = 0$ die Nullabbildung.
2. Im Falle $\mu_f = X - 1$ ist $f = \operatorname{id}_V$ die Identität.
3. Im Falle $\mu_f = X(X - 1)$ haben wir die Eigenraumzerlegung I2L:

$$
\begin{aligned}
&V = \operatorname{Eig}(f, 0) \oplus \operatorname{Eig}(f, 1) \;\text{ mit} \\
&\operatorname{Eig}(f, 0) = \ker(f - 0) = \operatorname{im}(f - 1) \ne \{0\}, \\
&\operatorname{Eig}(f, 1) = \ker(f - 1) = \operatorname{im}(f - 0) \ne \{0\}.
\end{aligned}
$$

Anwendungsbeispiel: Idempotenz und Eigenraumzerlegung

Idempotenz $f^2 = f$ bedeutet $P(f) = 0$ für $P = X^2 - X = X(X - 1)$. Aus dieser Faktorisierung folgt die Kernzerlegung

$$V = \ker(f^2 - f) = \ker(f) \oplus \ker(f - 1).$$

🙂 Die Eigenraumzerlegung $V = \operatorname{Eig}(f, 0) \oplus \operatorname{Eig}(f, 1)$ haben wir in I2L explizit ausgerechnet durch die Projektoren $\varphi_1 = f$ und $\varphi_2 = \operatorname{id}_V - f$. Diese Beobachtung werden wir im folgenden Satz M3V optimieren zu einem möglichst allgemeinen Zerlegungssatz.

🙂 Beispiel M3U gilt unabhängig von der Dimension $\dim_K(V)$, egal ob endlich oder unendlich. Das ist bemerkenswert: Das charakteristische Polynom steht uns bei $\dim_K(V) = \infty$ zwar nicht mehr zur Verfügung, aber ein Minimalpolynom ist in glücklichen Fällen dennoch möglich.

Teilerfremde Faktorisierung und Kernzerlegung

Satz M3v: teilerfremde Faktorisierung und Kernzerlegung

Sei $P = P_1 \cdots P_n$ mit $P_1, \dots, P_n \in K[X]$ und $\operatorname{ggT}(P_i, P_j) = 1$ für $i \neq j$.

(1) Für jeden K–Endomorphismus $f : V \to V$ mit $P(f) = 0$ gilt

$$V = \ker P(f) = \ker P_1(f) \oplus \cdots \oplus \ker P_n(f).$$

(2) Jeder Summand dieser Zerlegung ist f–invariant und erfüllt zudem

(3) $\ker P_i(f) = \operatorname{im} P_i^*(f)$ mit dem Cofaktor $P_i^* = P_1 \cdots P_{i-1} P_{i+1} \cdots P_n$.

(4) Bézout liefert $Q_i, Q_i^* \in K[X]$ mit $Q_i P_i + Q_i^* P_i^* = \operatorname{ggT}(P_i, P_i^*) = 1$,

(5) und $\varphi_i := Q_i^*(f) P_i^*(f) : V \to V$ projiziert auf den $i$ten Summanden.

Paradebeispiel: Zerfällt $P = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k)$ über K mit $\lambda_i \neq \lambda_j$ für $i \neq j$, so erhalten wir aus $P(f) = 0$ die Eigenraumzerlegung

$$V = \operatorname{Eig}(f, \lambda_1) \oplus \operatorname{Eig}(f, \lambda_2) \oplus \cdots \oplus \operatorname{Eig}(f, \lambda_k)$$

mit den expliziten Lagrange–Projektoren $\varphi_i = \prod_{j \neq i} (f - \lambda_j)/(\lambda_i - \lambda_j)$.

Teilerfremde Faktorisierung und Kernzerlegung

Bemerkung: Ohne die Voraussetzung $P(f) = 0$ gilt entsprechend nur

$$V \geq \ker P(f) = \ker P_1(f) \oplus \cdots \oplus \ker P_n(f).$$

Der Unterraum $U = \ker P(f)$ ist nämlich f–invariant, also $f(U) \subseteq U$. Somit können wir $f : V \to V$ einschränken zu $g : U \to U$ mit $P(g) = 0$. Der Satz zerlegt dann $\ker P(g) = \ker P(f) = U$ als direkte Summe der Unterräume $\ker P_i(f) = \ker P_i(g) \leq U$ wie oben angegeben.

Das ist eine bemerkenswerte Zerlegung, ebenso einfach wie elegant: Aus jeder Produktzerlegung $P = P_1 \cdots P_n$ in teilerfremde Faktoren $P_1, \dots, P_n \in K[X]$ erhalten wir die Summenzerlegung des Kerns!

Zudem erhalten wir jeden Projektor φ_i von $U = \ker P(g)$ auf $\ker P_i(g)$ als Polynom in g, durch eine explizite Formel wie im Satz angegeben. Im Beispiel ist dies besonders schön und einfach. Rechnen Sie es nach!

Der Beweis ist nicht minder elegant: Es handelt sich um eine genial effiziente Anwendung des Satzes von Beźout für Polynome in $K[X]$. Alle Daten liegen explizit vor, wir müssen es nur noch nachrechnen.

Kernzerlegung: teile und herrsche… mit Bézout!

Beweis: (3a) Es gilt $0 = P(f) = P_i(f) P_i^*(f)$, also $\operatorname{im} P_i^*(f) \subseteq \ker P_i(f)$.

(3b) Wir nutzen (4) $Q_i P_i + Q_i^* P_i^* = 1$. Angewendet auf f und $v \in V$ gilt

$$v = Q_i(f) P_i(f)(v) + P_i^*(f) Q_i^*(f)(v).$$

Für $v \in \ker P_i(f)$ folgt $v = P_i^*(f) Q_i^*(f)(v)$, also $\ker P_i(f) \subseteq \operatorname{im} P_i^*(f)$.

(2) Für $v \in \ker P_i(f)$ gilt $f(v) \in \ker P_i(f)$, denn

$$P_i(f) \circ f(v) = f \circ P_i(f)(v) = 0.$$

(1a) Wir haben $Q_i P_i + Q_i^* P_i^* = 1$. Für jeden Vektor $v \in V$ gilt

$$v = P_i(f) Q_i(f)(v) + P_i^*(f) Q_i^*(f)(v).$$

Somit erhalten wir die Summe

$$V = \operatorname{im} P_i(f) + \operatorname{im} P_i^*(f) \overset{(3)}{=} \ker P_i(f) + \ker P_i^*(f).$$

Hierzu nutzen wir (3) für die Zerlegung $P = P_i P_i^*$.

Kernzerlegung: teile und herrsche… mit Bézout!

(1b) Für $v \in \ker P_i(f) \cap \ker P_i^*(f)$ gilt

$$v = Q_i(f) P_i(f)(v) + Q_i^*(f) P_i^*(f)(v) = 0.$$

Wir erhalten so die Zerlegung als direkte Summe

$$V = \ker P_i(f) \oplus \ker P_i^*(f).$$

Der Projektor $\varphi_i := Q_i^*(f) P_i^*(f) : V \to V$ schickt $v_i + v_i^*$ auf v_i.

(1c) Der Unterraum $V^* = \ker P_i^*(f) \leq V$ ist f–invariant dank (2). Wir können daher f einschränken zu $f^* : V^* \to V^*$ mit $P_i^*(f^*) = 0$. Per Induktion über n gilt $V^* = \bigoplus_{j \neq i} \ker P_j(f)$, also insgesamt

$$V = \ker P_1(f) \oplus \ker P_2(f) \oplus \cdots \oplus \ker P_n(f).$$

Der Projektor $\varphi_i := Q_i^*(f) P_i^*(f)$ schickt $v_1 + \cdots + v_n$ auf v_i. $\boxed{\text{QED}}$

Insbesondere gilt $\varphi_i^2 = \varphi$ mit $\operatorname{im} \varphi_i = \ker P_i(f)$ und $\ker \varphi_i = \ker P_i^*(f)$ sowie $\varphi_1 + \cdots + \varphi_n = \operatorname{id}_V$ und $\varphi_i \circ \varphi_j = 0$ für $i \neq j$ siehe Satz I2K.

Diagonalisierung: notwendige und hinreichende Kriterien

Satz M3W: Kriterien zur Diagonalisierbarkeit

Sei $f : V \to V$ eine K–lineare Abbildung und $\dim_K V = n < \infty$.

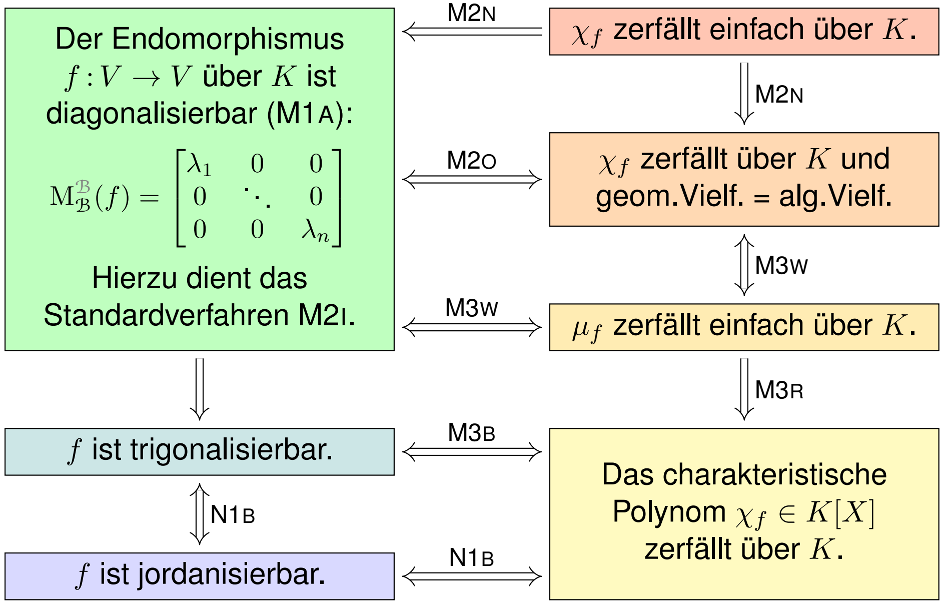
Dann sind die folgenden Bedingungen äquivalent:

1. Der Endomorphismus $f : V \to V$ ist diagonalisierbar.
2. Es gilt die Eigenraumzerlegung $V = \bigoplus_{\lambda\in\sigma(f)} \operatorname{Eig}(f,\lambda)$.
3. Es gilt die Dimensionsformel $n = \sum_{\lambda\in\sigma(f)} \dim_K \operatorname{Eig}(A,\lambda)$.
4. Das charakteristische Polynom χ_f zerfällt über K in Linearfaktoren und für jeden Eigenwert $\lambda \in \sigma(f)$ gilt $\dim_K \operatorname{Eig}(f,\lambda) = \operatorname{ord}(\chi_f,\lambda)$.
5. Das Minimalpolynom μ_f zerfällt einfach über K, ausgeschrieben $\mu_f = \prod_{\lambda\in\sigma(f)}(X-\lambda)$.

Beweis: Die Äquivalenz „(1) ⇔ (2) ⇔ (3)“ ist die Aussage von Satz M1I. Die Äquivalenz zu (4) ist Satz M2O. Neu ist nur die Äquivalenz zu (5):

Die Implikation „(1) ⇒ (5)“ haben wir in Beispiel M3N nachgerechnet. Die Umkehrung „(5) ⇒ (2)“ verdanken wir Satz M3V. QED

Diagonalisierung: notwendige und hinreichende Kriterien



Der Endomorphismus $f : V \to V$ über K ist diagonalisierbar (M1A):

$$\mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \lambda_n\end{bmatrix}$$

Hierzu dient das Standardverfahren M2I.

χ_f zerfällt einfach über K.

χ_f zerfällt über K und geom.Vielf. = alg.Vielf.

μ_f zerfällt einfach über K.

f ist trigonalisierbar.

f ist jordanisierbar.

Das charakteristische Polynom $\chi_f \in K[X]$ zerfällt über K.

Zusammenfassung zur Diagonalisierung

Grundlegende Begriffe und Techniken zur Diagonalisierung:

- Diagonalisierung (M1A), Eigenvektor, Eigenwert, Eigenraum, geometrische Vielfachheit, Eigenbasis, Spektrum (M1B)
- Jordan–Blöcke sind nicht-diagonalisierbar. (M1D)
- lineare Unabhängigkeit von Eigenvektoren (M1F)
- Eigenraumzerlegung und Diagonalisierung (M1I)
- Eigenwerte (M2A) und charakteristisches Polynom (M2C)
- Ähnlichkeit von Matrizen (M2D) und Invarianz (M2F)
- Standardverfahren zur Diagonalisierung (M2I)
- Einfache Zerfällung impliziert Diagonalisierbarkeit. (M2N)
- geometrische und algebraische Vielfachheit (M2O)
- Begleitmatrix (M2Q) und lineare Rekursion (M2R)

Zusammenfassung zur Diagonalisierung

Diagonalisierung hat zwei Bedingungen: (a) Das char. Polynom zerfällt und (b) die geometrische Vielfachheit erreicht die algebraische (M2O).

Letzteres ist nicht immer gegeben, statt Diagonalisierung begnügen wir uns dann mit einer Trigonalisierung; das ist wenig, aber immerhin etwas:

- Trigonalisierung (M3A) und Zerfällung (M3B)
- Spur und Determinante aus Eigenwerten (M3F)

Schließlich bündeln Minimalpolynome nützliche Informationen:

- lokales Minimalpolynom (M3I) und Cayley–Hamilton (M3K)
- das (globale) Minimalpolynom eines Endomorphismus (M3L)
- Teilbarkeit (M3O) und Nullstellen (M3R) der beiden Polynome
- teilerfremde Faktorisierung und Kernzerlegung (M3V)
- äquivalente Kriterien zur Diagonalisierbarkeit (M3W)

Anwendungsbeispiele und Übungen

Aufgabe: Zu untersuchen ist die reelle Matrix

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

(1) Bestimmen und zerlegen Sie die Polynome χ_A und μ_A in $\mathbb{R}[X]$.
(2) Warum ist A diagonalisierbar? Bestimmen Sie eine Eigenbasis zu A.
(3) Beschreiben Sie die Wirkung von A auf $\mathbb{R}^3$ geometrisch / in Worten.

Lösung: (1a) Für das charakteristische Polynom finden wir

$$\chi_A(X) := \det(XI - A) \quad = \quad X^3 - 2X^2 + X \quad = \quad X(X-1)^2.$$

(1b) Das Minimalpolynom von A ist die erste nicht-triviale Relation der Potenzen A^0, A^1, A^2, A^3. Wir berechnen $A^2 = A$ und schließen

$$\mu_A \quad = \quad X^2 - X \quad = \quad X(X-1).$$

(2a) Allein aus χ_A können wir auf Trigonalisierbarkeit schließen (M3B). Da zudem μ_A einfach zerfällt, ist die Matrix A diagonalisierbar (M3W).

Anwendungsbeispiele und Übungen

(2b) Wir bestimmen die zugehörigen Eigenräume:

$$A - 0I = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix} \xrightarrow[\text{RZSF}]{\text{Gauß}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Daraus lesen wir $\operatorname{Eig}(A,0) = \langle\, v_1 = (-1,1,1)^\intercal \,\rangle =: U$ ab.

$$A - 1I = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix} \xrightarrow[\text{RZSF}]{\text{Gauß}} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Daraus folgt $\operatorname{Eig}(A,1) = \langle\, v_2 = (-1,1,0)^\intercal,\ v_3 = (1,0,1)^\intercal \,\rangle := V$
Wie in (2a) vorhergesagt, gilt $\mathbb{R}^3 = \operatorname{Eig}(A,0) \oplus \operatorname{Eig}(A,1)$.

Mit $\mathcal{B} = (v_1, v_2, v_3)$ haben wir eine Eigenbasis von $\mathbb{R}^3$ zu A.
Für $T = (v_1, v_2, v_3) \in \mathrm{GL}_3\,\mathbb{R}$ gilt $T^{-1}AT = \operatorname{diag}(0,1,1)$.

(3) Die Abbildung $p : \mathbb{R}^3 \to \mathbb{R}^3 : x \mapsto Ax$ ist die Projektion auf die Ebene V parallel zur Geraden U: Es gilt $p^2 = p$ mit $\ker(p) = U$ und $\operatorname{im}(p) = V$.

Anwendungsbeispiele und Übungen

Aufgabe: Sei $\mathbb{K} = \mathbb{R}, \mathbb{C}$ sowie $A \in \mathbb{K}^{n\times n}$ und $m \in \mathbb{N}_{\ge 1}$ mit $A^m = I$.
Ist A diagonalisierbar? Geben Sie einen Beweis oder ein Gegenbeispiel.

Lösung: (1) Wir betrachten zunächst den reellen Fall $\mathbb{K} = \mathbb{R}$.

(1a) Das einfachste Gegenbeispiel ist die Vierteldrehung der Ebene $\mathbb{R}^2$:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

Für diese Matrix gilt $A^4 = I$, doch $\chi_A = X^2 + 1$ zerfällt nicht über $\mathbb{R}$, also ist A nicht diagonalisierbar, nicht einmal trigonalisierbar (M3B).

(1b) Wir betrachten eine Drehung der Ebene $\mathbb{R}^2$ um den Winkel $\theta \in \mathbb{R}$:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

Speziell $A = R(\theta)$ mit $\theta = \pi k/n$ und $k \in \{1, \dots, n-1\}$ erfüllt $A^{2n} = I$.
Das char. Polynom $\chi_A = (X - \cos\theta)^2 + (\sin\theta)^2$ zerfällt nicht über $\mathbb{R}$.
Das einfachste Beispiel erhalten wir für $\theta = \pi/2$, wie in (1a) gezeigt.

Anwendungsbeispiele und Übungen

(2) Der komplexe Fall $\mathbb{K} = \mathbb{C}$ ist wesentlich sympathischer, wie so oft:

(2a) Das Polynom $P = X^m - 1$ annulliert A, denn $P(A) = A^m - I = 0$.
Die Gleichung $z^m = 1$ hat in $\mathbb{C}$ die Lösungen $z_k = \mathrm{e}^{2\pi\mathrm{i}k/m}$. Demnach gilt

$$X^m - 1 = \textstyle\prod_{k=0}^{m-1} (X - \mathrm{e}^{2\pi\mathrm{i}k/m}).$$

Die komplexen Zahlen $z_0, z_1, \dots, z_{m-1}$ sind die $m$ten Einheitswurzeln.
Über $\mathbb{R}$ zerfällt $X^m - 1$ nur für $m \le 2$, dank $X^2 - 1 = (X-1)(X+1)$.

(2b) Das Minimalpolynom μ_A teilt das Polynom P gemäß Satz M3O.
Also zerfällt auch μ_A einfach, und A ist diagonalisierbar dank M3W.

Beispiel: Die Drehmatrix $A = R(\theta) \in \mathbb{R}^{2\times 2}$ aus (1b) hat zwei echt komplexe Eigenwerte, $\sigma(A;\mathbb{R}) = \emptyset$ und $\sigma(A;\mathbb{C}) = \{\, \cos\theta \pm \mathrm{i}\sin\theta \,\}$.
Über $\mathbb{C}$ zerfällt χ_A einfach, und A ist diagonalisierbar dank M2N.
Über $\mathbb{R}$ jedoch ist A nicht einmal trigonalisierbar wegen M3B.

Bemerkung: Das Argument (2) gilt über jedem Körper $\mathbb{K}$, über dem $X^m - 1$ zerfällt, zum Beispiel $\mathbb{K} = \mathbb{F}_p$ mit $p = nm + 1$ prim. Dank G3N gilt $X^{p-1} - 1 = (X^m - 1)(X^{(n-1)m} + \dots + X^m + 1) = \prod_{a \in \mathbb{F}_p^*} (X - a)$.

Simultane Diagonalisierung von Endomorphismen

Definition M4A: simultane Diagonalisierung von Endomorphismen

(1) Seien $f_1, \dots, f_\ell : V \to V$ lineare Abbildungen über K. Hierzu ist eine **diagonalisierende Basis** $\mathcal{B} = (v_i)_{i \in I}$ eine Basis von V, sodass $f_k(v_i) = \lambda_i^{(k)} v_i$ und $\lambda_i^{(k)} \in K$ für jeden Index $i \in I$ und alle $k = 1, \dots, \ell$.

Wir nennen $\mathcal{B}$ eine **(simultane) Eigenbasis** der Familie $(f_1, \dots, f_\ell)$.

Existiert zur Familie $(f_1, \dots, f_\ell)$ eine simultane Eigenbasis $\mathcal{B}$ von V, so nennen wir $(f_1, \dots, f_\ell)$ **simultan diagonalisierbar**.

(2) Seien $A_1, \dots, A_\ell \in K^{n \times n}$ quadratische Matrizen derselben Größe über K. Hierzu ist ein **(simultan) diagonalisierender Basiswechsel** eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$, so dass die konjugierten Matrizen $T^{-1} A_1 T, \dots, T^{-1} A_\ell T \in K^{n \times n}$ allesamt diagonal sind.

Existiert eine solche Matrix T, so nennen wir die Familie $(A_1, \dots, A_r)$ in $K^{n \times n}$ **simultan diagonalisierbar** über K.

Im Falle $\ell = 1$ ist dies die vorige Definition M1A zur Diagonalisierung.

Simultane Diagonalisierung von Endomorphismen

Wann ist simultane Diagonalisierung möglich? Der folgende Satz gibt hierzu ein einfaches Kriterium, sowohl notwendig als auch hinreichend:

Satz M4B: simultane Diagonalisierung von Endomorphismen

Seien $f_1, \dots, f_\ell : V \to V$ lineare Abbildungen über K. Äquivalent sind:

1. Die Familie $(f_1, \dots, f_\ell)$ ist simultan diagonalisierbar.
2. Jeder Endomorphismus $f_1, \dots, f_\ell$ ist einzeln diagonalisierbar und je zwei kommutieren, also $f_j \circ f_k = f_k \circ f_j$ für alle $j, k \in \{1, \dots, \ell\}$.

Die im Beweis geführte Konstruktion „(2) ⇒ (1)“ erklärt zugleich einen Algorithmus zur simultanen Diagonalisierung.

Beweis: „(1) ⇒ (2)“: Sei $\mathcal{B} = (v_i)_{i \in I}$ eine simultan diagonalisierende Basis, also $f_k(v_i) = \lambda_i^{(k)} v_i$ für jeden Index $i \in I$ und alle $k = 1, \dots, \ell$. Dann gilt $f_j \circ f_k = f_k \circ f_j$ auf der Basis $\mathcal{B}$ und somit auf ganz V:

$$\begin{aligned} f_j(f_k(v_i)) &= f_j(\lambda_i^{(k)} v_i) = \lambda_i^{(k)} f_j(v_i) = \lambda_i^{(k)} \lambda_i^{(j)} v_i, \\ f_k(f_j(v_i)) &= f_j(\lambda_i^{(j)} v_i) = \lambda_i^{(j)} f_k(v_i) = \lambda_i^{(j)} \lambda_i^{(k)} v_i. \end{aligned}$$

Simultane Diagonalisierung von Endomorphismen

„(2) ⇒ (1)“: Wir führen Induktion über ℓ. Der Fall $\ell = 1$ ist trivial. Sei also $\ell \ge 2$. Der Endomorphismus $f_\ell : V \to V$ ist diagonalisierbar, also gilt $V = \bigoplus_{\lambda \in K} \mathrm{Eig}(f_\ell, \lambda)$. Die entscheidende Beobachtung ist: Jeder Eigenraum $U = \mathrm{Eig}(f_\ell, \lambda)$ ist f_k–invariant: Für $v \in \mathrm{Eig}(f_\ell, \lambda)$ gilt

$$f_\ell(f_k(v)) = f_k(f_\ell(v)) = f_k(\lambda v) = \lambda f_k(v),$$

also $f_k(v) \in \mathrm{Eig}(f_\ell, \lambda)$. Wir können also $f_k : V \to V$ einschränken zu $g_k = f_k|_U^U$. Die Endomorphismen $g_1, \dots, g_{\ell-1}$ kommutieren weiterhin und sind einzeln diagonalisierbar dank dem nachfolgenden Lemma.

Nach Induktionsvoraussetzung existiert eine diagonalisierende Basis $\mathcal{B}_\lambda = (v_i)_{i \in I_\lambda}$ von U zu der Familie $(g_1, \dots, g_{\ell-1})$. Zudem gilt $g_\ell = \lambda\, \mathrm{id}_U$. Zusammengesetzt zu $I = \bigsqcup_{\lambda \in K} I_\lambda$ erhalten wir die diagonalisierende Basis $\mathcal{B} = (v_i)_{i \in I}$ von V zu der Familie $(f_1, \dots, f_{\ell-1}, f_\ell)$. QED

☺ Dieser Beweis ist konstruktiv: Das Vorgehen erklärt zugleich einen Algorithmus zur simultanen Diagonalisierung.

Simultane Diagonalisierung von Endomorphismen

⚠ Die Basis $\mathcal{B}_\lambda$ des Eigenraums $\mathrm{Eig}(f_\ell, \lambda)$ bestimmen wir nicht allein mit f_ℓ, sondern auch mit den anderen Endomorphismen $f_1, \dots, f_{\ell-1}$.

⚠ Satz M4B sagt nicht allgemein: „Kommutierende Endomorphismus sind diagonalisierbar.“ Der Satz sagt korrekt genau: „Kommutierende diagonalisierbare Endomorphismus sind simultan diagonalisierbar.“

Beispiel: Alle Matrizen $\left[\begin{smallmatrix}1 & a\\ 0 & 1\end{smallmatrix}\right]$ mit $a \in K^*$ kommutieren untereinander, aber keine davon ist diagonalisierbar. (Übung: Rechnen Sie es nach.)

☺ Angenommen, die Basis $\mathcal{B}$ diagonalisiert $f_1, \dots, f_\ell$ in $\mathrm{End}_K(V)$. Dann gilt dies für alle Endomorphismen in dem hiervon erzeugten kommutativen Unterring $K[f_1, \dots, f_\ell] \le \mathrm{End}_K(V)$.

Beispiel: Ohne Kommutativität bleibt Diagonalisierbarkeit nicht erhalten unter Summen und Produkten. Zum Beispiel sind die Matrizen $A = \left[\begin{smallmatrix}1 & 0\\ 0 & 2\end{smallmatrix}\right]$ und $B = \left[\begin{smallmatrix}0 & 1\\ 0 & -1\end{smallmatrix}\right]$ einzeln diagonalisierbar, nicht jedoch $A + B = \left[\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right]$. (Übung: Finden Sie ähnliche Gegenbeispiele für Produkte.)

Diagonalisierbarkeit und Einschränkung

Lemma M4C: Diagonalisierbarkeit und Einschränkung

Sei $f : V \to V$ diagonalisierbar und $U \le V$ ein f–invarianter Unterraum. Dann ist auch die Einschränkung $g = f|_U^U : U \to U$ diagonalisierbar:

$$U = \bigoplus\nolimits_{\lambda\in K} \mathrm{Eig}(g,\lambda) \quad \text{mit} \quad \mathrm{Eig}(g,\lambda) = \mathrm{Eig}(f,\lambda) \cap U$$

Beweis in endlicher Dimension: Genau dann ist f diagonalisierbar, wenn das Minimalpolynom μ_f einfach zerfällt (M3W). Zudem gilt $\mu_g \mid \mu_f$, denn $\mu_f(g) = \mu_f(f)|_U^U = 0$ (M3O). Daher zerfällt μ_g ebenfalls einfach, und somit ist auch g diagonalisierbar (M3W). QED

Allgemein: Es gilt $V = \bigoplus_{\lambda\in K} \mathrm{Eig}(f,\lambda)$. Jeder Vektor $u \in U$ schreibt sich $u = v_0 + \dots + v_n$ mit $v_i \in \mathrm{Eig}(f,\lambda_i)$, wobei $\lambda_i \ne \lambda_j$ für $i \ne j$ gelte. Wir zeigen $v_0,\dots,v_n \in U$ und schließen daraus die obige Zerlegung.

Beweis per Induktion über n: Der Fall $n = 0$ ist klar. Sei also $n \ge 1$. Anwendung von $f - \lambda_0$ ergibt $u' = (\lambda_1 - \lambda_0)v_1 + \dots + (\lambda_n - \lambda_0)v_n \in U$. Per Induktion folgt daraus $v_1,\dots,v_n \in U$, also auch $v_0 \in U$. QED

Diagonalisierbarkeit und Einschränkung

Beweis mit dem Vandermonde–Trick: Da U invariant unter f ist, gilt $u_k := f^k(u) \in U$ für alle $k \in \mathbb{N}$. Wir haben $f^k(u) = \lambda_0^k v_0 + \dots + \lambda_n^k v_n$. Für $k = 0, 1, \dots, r$ erhalten wir so das folgende Gleichungssystem:

$$\begin{bmatrix} \lambda_0^0 & \lambda_1^0 & \dots & \lambda_n^0 \\ \lambda_0^1 & \lambda_1^1 & \dots & \lambda_n^1 \\ \vdots & & \vdots & \\ \lambda_0^n & \lambda_1^n & \dots & \lambda_n^n \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_n \end{bmatrix}$$

Die Vandermonde–Matrix ist invertierbar dank Satz B3A:

$$\begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \lambda_0^0 & \lambda_1^0 & \dots & \lambda_n^0 \\ \lambda_0^1 & \lambda_1^1 & \dots & \lambda_n^1 \\ \vdots & & \vdots & \\ \lambda_0^n & \lambda_1^n & \dots & \lambda_n^n \end{bmatrix}^{-1} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_n \end{bmatrix}$$

Somit ist jeder Vektor $v_0,\dots,v_n$ eine Linearkombination von $u_0,\dots,u_n$, insbesondere gilt demnach $v_0,\dots,v_n \in U$, wie behauptet. QED

Diagonalisierbarkeit und Einschränkung

Beweis mit Lagrange–Projektoren: Wir nutzen die Endomorphismen

$$\varphi_i = \prod\nolimits_{j\ne i} (f - \lambda_j)/(\lambda_i - \lambda_j) \; : \; U \to U.$$

Damit gilt $v_i = \varphi_i(u) \in U$, und genau das war zu zeigen. QED

🙂 Warum sollten wir ein Lemma wie M4C mehrfach beweisen?

Rein logisch gesehen genügt selbstverständlich ein einziger Beweis. Wenn sich mehrere anbieten, wählen wir je nach Bedarf den kürzesten, schönsten, nützlichsten, etc. Im vorliegenden Fall fällt die Wahl schwer, denn jeder unserer Beweise illustriert ein anderes schönes Werkzeug: Minimalpolynom, Induktion, Vandermonde oder Projektoren.

Aus didaktischen Gründen und zur Übung dieser Techniken halte ich es daher für sinnvoll, mehrere Beweise zu präsentieren. So sehen Sie ihre jeweiligen Vorzüge und können sich den für Sie schönsten aussuchen. Das bietet zudem eine breite Grundlage für zukünftige Anwendungen und vielfältige Erfahrung für eigene Rechnungen und Beweise.

Ausblick: Anwendungen in Physik und Chemie

Auch in der Physik treten Eigenwerte und Eigenvektoren häufig auf. In der Quantenmechanik werden physikalische Größen durch lineare Operatoren beschrieben. Eigenvektoren sind besonders einfache Basiszustände, ihre Eigenwerte entsprechen möglichen Messwerten.

Auf diese Weise erklärte Heisenberg seine Unschärferelation durch Begriffe der Linearen Algebra: Ortsoperator und Impulsoperator kommutieren nicht und haben keine gemeinsame Eigenbasis, sie können nicht beide gleichzeitig genau gemessen werden.

Die „Atom-Orbitale“, die Sie aus dem Chemie-Unterricht kennen, sind die Eigenfunktionen der zugehörigen Schrödinger–Gleichung. Auch hier entsprechen Messungen linearen Operatoren, Messwerte sind Eigenwerte, und gleichzeitige Messung entspricht simultaner Diagonalisierung. Letzteres gelingt nur bei Kommutativität.

Eigenvektoren spielen in vielen weiteren Anwendungen eine wichtige Rolle, etwa in der Graphentheorie, speziell beim PageRank–Verfahren, mit dem Google die Relevanz von Internetseiten bewertet.

Anwendungsbeispiel zur simultanen Diagonalisierung

Aufgabe: (1) Welche der folgenden Matrizen sind $\mathbb{R}$–diagonalisierbar?

$$A = \begin{bmatrix} -10 & -8 & -2 \\ 11 & 9 & 2 \\ 22 & 16 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} -9 & -8 & -2 \\ 18 & 15 & 3 \\ -6 & -4 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} -10 & -8 & -2 \\ 14 & 11 & 2 \\ 10 & 8 & 5 \end{bmatrix}$$

(2) Welche Teilfamilien von (A, B, C) sind simultan diagonalisierbar?
(3) Finden Sie jeweils eine simultane Eigenbasis, soweit möglich.
(4) Finden Sie jeweils alle simultanen Eigenbasen, soweit möglich.

Lösung: (1a) Wir berechnen $\chi_A = (X-1)^2(X-2)$. Das allein lässt die Frage der Diagonalisierbarkeit noch offen. Zur Klärung berechnen wir entweder zuerst $\mu_A = (X-1)(X-2)$ oder gleich die Eigenräume:

$$\begin{aligned} \operatorname{Eig}(A,1) &= \big\langle\, (-8,11,0)^\intercal,\ (-2,0,11)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \\ \operatorname{Eig}(A,2) &= \big\langle\, (-1,1,2)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \end{aligned}$$

Machen Sie die Probe! Somit ist die Matrix A über $\mathbb{R}$ diagonalisierbar. (Bonus: Die Eigenräume liefern erneut χ_A und μ_A wie angegeben.)

Anwendungsbeispiel zur simultanen Diagonalisierung

(1b) Wir berechnen $\chi_B = (X-2)^2(X-3)$. Das allein lässt die Frage der Diagonalisierbarkeit vorerst noch offen. Zur Klärung berechnen wir entweder zuerst $\mu_B = (X-2)(X-3)$ oder gleich die Eigenräume:

$$\begin{aligned} \operatorname{Eig}(B,2) &= \big\langle\, (2,-3,1)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \\ \operatorname{Eig}(B,3) &= \big\langle\, (-2,3,0)^\intercal,\ (-1,0,6)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \end{aligned}$$

Machen Sie die Probe! Somit ist die Matrix B über $\mathbb{R}$ diagonalisierbar. (Bonus: Die Eigenräume liefern erneut χ_B und μ_B wie angegeben.)

(1c) Wir berechnen $\chi_C = (X-1)(X-2)(X-3)$. Somit ist die Matrix C über $\mathbb{R}$ diagonalisierbar. Explizit berechnen wir hierzu die Eigenräume:

$$\begin{aligned} \operatorname{Eig}(C,1) &= \big\langle\, (2,-3,1)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \\ \operatorname{Eig}(C,2) &= \big\langle\, (1,-2,2)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \\ \operatorname{Eig}(C,3) &= \big\langle\, (0,-1,4)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \end{aligned}$$

Machen Sie die Probe! Dies zeigt erneut $\chi_C = (X-1)(X-2)(X-3)$.

Anwendungsbeispiel zur simultanen Diagonalisierung

(2) Jede Matrix A, B, C ist *einzeln* diagonalisierbar. Zur *simultanen* Diagonalisierbarkeit prüfen wir die Kommutativität (M4B) und finden:

$$AB = BA \quad \text{und} \quad BC = CB \quad \text{aber} \quad AC \neq CA$$

Die Teilfamilien (A, B) und (B, C) sind also simultan diagonalisierbar, die Teilfamilien (A, C) und (A, B, C) hingegen sind es nicht!

(3a) Wir können auf (A, B) den Algorithmus M4B anwenden oder aus den berechneten Eigenräumen eine gemeinsame Eigenbasis wählen:

$$\begin{aligned} &\operatorname{Eig}(A,1) = \big\langle\, (-8,11,0)^\intercal,\ (-2,0,11)^\intercal \,\big\rangle_{\mathbb{R}}^{!}, \ \operatorname{Eig}(A,2) = \big\langle\, (-1,1,2)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \\ &\operatorname{Eig}(B,2) = \big\langle\, (2,-3,1)^\intercal \,\big\rangle_{\mathbb{R}}^{!}, \ \operatorname{Eig}(B,3) = \big\langle\, (-2,3,0)^\intercal,\ (-1,0,6)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \end{aligned}$$

$$S = \begin{bmatrix} 2 & 0 & -1 \\ -3 & -1 & 1 \\ 1 & 4 & 2 \end{bmatrix}, \quad S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad S^{-1}BS = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

(4a) Das ist die einzige gemeinsame Eigenbasis des Paares (A, B), wie immer bis auf Vielfache und Vertauschung der Basisvektoren.

Anwendungsbeispiel zur simultanen Diagonalisierung

(3b) Wir können auf (B, C) den Algorithmus M4B anwenden oder aus den berechneten Eigenräumen eine gemeinsame Eigenbasis wählen:

$$\begin{aligned} &\operatorname{Eig}(B,2) = \big\langle\, (2,-3,1)^\intercal \,\big\rangle_{\mathbb{R}}^{!}, \ \operatorname{Eig}(B,3) = \big\langle\, (-2,3,0)^\intercal,\ (-1,0,6)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \\ &\operatorname{Eig}(C,\{1,2,3\}) = \big\{\, \big\langle\, (2,-3,1)^\intercal \,\big\rangle_{\mathbb{R}}^{!},\ \big\langle\, (1,-2,2)^\intercal \,\big\rangle_{\mathbb{R}}^{!},\ \big\langle\, (0,-1,4)^\intercal \,\big\rangle_{\mathbb{R}}^{!} \,\big\} \end{aligned}$$

$$T = \begin{bmatrix} 2 & 1 & 0 \\ -3 & -2 & -1 \\ 1 & 2 & 4 \end{bmatrix}, \quad T^{-1}BT = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad T^{-1}CT = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

(4b) Das ist die einzige Eigenbasis von C und somit auch von (B, C), wie immer bis auf Vielfache und Vertauschung der Basisvektoren.

Bemerkung: In den Eigenbasen zu (A, B) bzw. (B, C) gilt jeweils

$$S^{-1}CS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{bzw.} \quad T^{-1}AT = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

Dies zeigt eindrücklich, dass die jeweils dritte Matrix nicht gleichzeitig diagonalisiert wird, auch nicht kann, wie wir aus (2) bereits wissen.

Anwendungsbeispiel zur simultanen Diagonalisierung

Aufgabe: Wir betrachten weiterhin die drei reellen Matrizen

$$A = \begin{bmatrix} -10 & -8 & -2 \\ 11 & 9 & 2 \\ 22 & 16 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} -9 & -8 & -2 \\ 18 & 15 & 3 \\ -6 & -4 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} -10 & -8 & -2 \\ 14 & 11 & 2 \\ 10 & 8 & 5 \end{bmatrix}.$$

(5) Sind $A,\ B,\ A+B,\ A \cdot B$ diagonalisierbar? einzeln? simultan?
(6) Sind $B,\ C,\ B+C,\ B \cdot C$ diagonalisierbar? einzeln? simultan?
(7) Sind $A,\ C,\ A+C,\ A \cdot C$ diagonalisierbar? einzeln? simultan?

Lösung: (5) Wir haben das Paar (A, B) bereits simultan diagonalisiert:

$$S = \begin{bmatrix} 2 & 0 & -1 \\ -3 & -1 & 1 \\ 1 & 4 & 2 \end{bmatrix} \implies \begin{cases} S^{-1}AS &= \operatorname{diag}(1,1,2) \\ S^{-1}BS &= \operatorname{diag}(2,3,3) \\ S^{-1}(A+B)S &= \operatorname{diag}(3,4,5) \\ S^{-1}(A \cdot B)S &= \operatorname{diag}(2,3,6) \end{cases}$$

Durch S wird somit auch $(\ A,\ B,\ A+B,\ A \cdot B\ )$ simultan diagonalisiert.

Anwendungsbeispiel zur simultanen Diagonalisierung

(6) Wir haben das Paar (B, C) bereits simultan diagonalisiert:

$$T = \begin{bmatrix} 2 & 1 & 0 \\ -3 & -2 & -1 \\ 1 & 2 & 4 \end{bmatrix} \implies \begin{cases} T^{-1}BT &= \operatorname{diag}(2,3,3) \\ T^{-1}CT &= \operatorname{diag}(1,2,3) \\ T^{-1}(B+C)T &= \operatorname{diag}(3,5,6) \\ T^{-1}(B \cdot C)T &= \operatorname{diag}(2,6,9) \end{cases}$$

Durch T wird somit auch $(\ B,\ C,\ B+C,\ B \cdot C\ )$ simultan diagonalisiert.

(7) Wegen $AC \ne CA$ lässt sich (A, C) nicht simultan diagonalisieren. Über die Diagonalisierbarkeit von $A + C$ und $A \cdot C$ lässt sich daraus allein leider noch nichts schließen; beide Fälle sind noch möglich.

Um dies zu klären, müssen wir die Matrizen genauer untersuchen.
Langer Weg: Wir können das Standardverfahren M2I anwenden.
Abkürzung: Wir nutzen geschickt unsere Ergebnisse (1–4).

Anwendungsbeispiel zur simultanen Diagonalisierung

Zur Vereinfachung nutzen wir unsere Diagonalisierung von A:

$$S = \begin{bmatrix} 2 & 0 & -1 \\ -3 & -1 & 1 \\ 1 & 4 & 2 \end{bmatrix}, \quad S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad S^{-1}CS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Daraus lesen wir Summe und Produkt ab:

$$S^{-1}(A+C)S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \quad S^{-1}(A \cdot C)S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

Das zeigt uns, $A \cdot C$ ist (zufällig) diagonalisierbar, $A + C$ jedoch nicht. Ausführlich gilt $\chi_{A+C} = \mu_{A+C} = (X-1)(X-4)^2$, also greift Satz M3W: Somit sind zwar A und C diagonalisierbar, $A + C$ jedoch nicht.

Ebenso gilt $\chi_{AC} = \mu_{AC} = (X-1)(X-3)(X-4)$, also greift Satz M2N: Somit sind A und C und AC diagonalisierbar, jedoch nicht simultan, denn es gilt $AC \ne CA$ sowie $A(AC) \ne (AC)A$ und $C(AC) \ne (AC)C$.

Anwendungsbeispiel zur simultanen Diagonalisierung

Dieses Zahlenbeispiel zeigt erneut sehr eindrücklich:

☺ Starke theoretische Grundlagen liefern praktische Werkzeuge.
Diese strukturieren und vereinfachen die Rechnungen spürbar.

☺ Gute Notation und Standardverfahren erleichtern das Vorgehen.
Oft hilft es, vorige Informationen geschickt wiederzuverwenden.

☺ Sie arbeiten umso effizienter, je genauer Sie verstehen, was Sie tun!
. . . insbesondere, was Sie schon haben und was Sie noch suchen.
Sie können dann geschickt vom Standardverfahren abweichen,
je nach konkretem Bedarf und möglichen Abkürzungen.