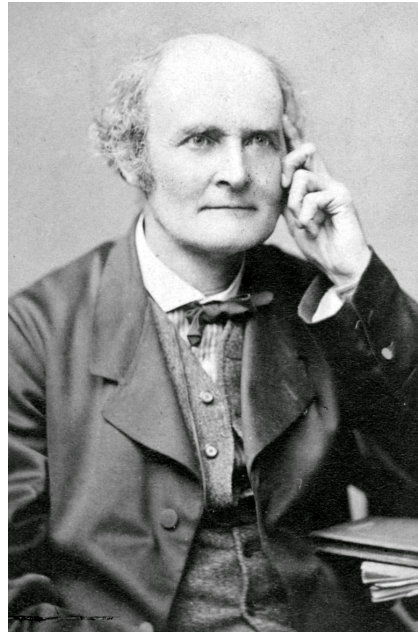


Lineare Algebra

Vorlesung 24

Der Satz von Cayley-Hamilton



Arthur Cayley (1821-1895)



William Hamilton (1805-1865)

Einer der Höhepunkte der linearen Algebra ist der Satz von Cayley-Hamilton. Um ihn formulieren zu können erinnern wir daran, dass man in Polynome quadratische Matrizen einsetzen kann, siehe die 20. Vorlesung. Dabei ersetzt man an jeder Stelle die Variable X durch die Matrix M und muss die Potenzen M^i als das i -te Matrixprodukt von M mit sich selbst verstehen und die Addition als die (komponentenweise) Addition von Matrizen interpretieren. Ein Skalar a wird dabei als das a -fache der Einheitsmatrix interpretiert. Für das Polynom

$$P = 3X^2 - 5X + 2$$

und die Matrix

$$M = \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}$$

ist also

$$\begin{aligned} P(M) &= 3 \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}^2 - 5 \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} + 2 \\ &= \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 16 & 12 \\ 9 & 13 \end{pmatrix} + \begin{pmatrix} -5 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 40 & 16 \\ 12 & 36 \end{pmatrix}. \end{aligned}$$

Zu einer fixierten Matrix $M \in \text{Mat}_n(K)$ gibt es also eine *Einsetzungsabbildung*

$$K[X] \longrightarrow \text{Mat}_n(K), P \longmapsto P(M).$$

Dies ist - ebenso wie die Einsetzungsabbildung zu $a \in K$ - ein Ringhomomorphismus, d.h. es gelten die Beziehungen (vergleiche Lemma 20.3)

$$(P+Q)(M) = P(M) + Q(M), (P \cdot Q)(M) = P(M) \circ Q(M) \text{ und } 1(M) = E_n.$$

Der Satz von Cayley-Hamilton beantwortet nun die Frage, was passiert, wenn man eine Matrix in ihr charakteristisches Polynom einsetzt.

SATZ 24.1. *Es sei K ein Körper und sei M eine $n \times n$ -Matrix über K . Es sei*

$$\chi_M = X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0$$

das charakteristische Polynom zu M . Dann gilt

$$\chi_M(M) = M^n + c_{n-1}M^{n-1} + \cdots + c_1M + c_0 = 0.$$

Das heißt, dass die Matrix das charakteristische Polynom annulliert.

Beweis. Wir fassen die Matrix $XE_n - M$ als eine Matrix auf, deren Einträge im Körper $K(X)$ liegen. Die adjungierte Matrix

$$(XE_n - M)^{\text{adj}}$$

liegt ebenfalls in $\text{Mat}_n(K(X))$. Die einzelnen Einträge der adjungierten Matrix sind nach Definition Determinanten von $(n-1) \times (n-1)$ -Untermatrizen

von $XE_n - M$. In den Einträgen dieser Matrix kommt die Variable X maximal in der ersten Potenz vor, sodass in den Einträgen der adjungierten Matrix die Variable maximal in der $(n-1)$ -ten Potenz vorkommt. Wir schreiben

$$(XE_n - M)^{\text{adj}} = X^{n-1}A_{n-1} + X^{n-2}A_{n-2} + \cdots + XA_1 + A_0$$

mit Matrizen

$$A_i \in \text{Mat}_n(K),$$

d.h. man schreibt die einzelnen Einträge als Polynom und fasst dann zu X^i die Koeffizienten zu einer Matrix zusammen. Aufgrund von Satz 17.9 gilt

$$\begin{aligned} \chi_M E_n &= (XE_n - M) \circ (XE_n - M)^{\text{adj}} \\ &= (XE_n - M) \circ (X^{n-1}A_{n-1} + X^{n-2}A_{n-2} \\ &\quad + \cdots + XA_1 + A_0) \\ &= X^n A_{n-1} + X^{n-1}(A_{n-2} - M \circ A_{n-1}) \\ &\quad + X^{n-2}(A_{n-3} - M \circ A_{n-2}) \\ &\quad + \cdots + X^1(A_0 - M \circ A_1) - M \circ A_0. \end{aligned}$$

Wir können auch die Matrix links nach den Potenzen von X aufteilen, dann ist

$$\chi_M E_n = X^n E_n + X^{n-1}c_{n-1}E_n + X^{n-2}c_{n-2}E_n + \cdots + X^1c_1E_n + c_0E_n.$$

Da diese zwei Polynome übereinstimmen, müssen jeweils ihre Koeffizienten übereinstimmen. D.h. wir haben ein System von Gleichungen

$$\begin{aligned} E_n &= A_{n-1} \\ c_{n-1}E_n &= A_{n-2} - M \circ A_{n-1} \\ c_{n-2}E_n &= A_{n-3} - M \circ A_{n-2} \\ &\vdots \\ c_1E_n &= A_0 - M \circ A_1 \\ c_0E_n &= -M \circ A_0. \end{aligned}$$

Wir multiplizieren diese Gleichungen von links von oben nach unten mit $M^n, M^{n-1}, M^{n-2}, \dots, M^1, E_n$ und erhalten das Gleichungssystem

$$\begin{aligned} M^n &= M^n \circ A_{n-1} \\ c_{n-1}M^{n-1} &= M^{n-1} \circ A_{n-2} - M^n \circ A_{n-1} \\ c_{n-2}M^{n-2} &= M^{n-2} \circ A_{n-3} - M^{n-1} \circ A_{n-2} \\ &\vdots \\ c_1M^1 &= MA_0 - M^2 \circ A_1 \\ c_0E_n &= -M \circ A_0. \end{aligned}$$

Wenn wir die linke Spalte dieses Gleichungssystem aufsummieren, so erhalten wir gerade $\chi_M(M)$. Wenn wir die rechte Seite aufsummieren, so erhalten wir 0, da jeder Teilschritt $M^{i+1} \circ A_i$ einmal positiv und einmal negativ vorkommt. Also ist $\chi_M(M) = 0$.

□

SATZ 24.2. *Es sei V ein endlichdimensionaler Vektorraum über einem Körper K und es sei*

$$f: V \longrightarrow V$$

eine lineare Abbildung. Dann gilt für das charakteristische Polynom die Beziehung

$$\chi_f(f) = 0.$$

Beweis. Dies folgt direkt aus Satz 24.1. □

Minimalpolynom und charakteristisches Polynom

KOROLLAR 24.3. *Es sei V ein endlichdimensionaler Vektorraum über einem Körper K und es sei*

$$f: V \longrightarrow V$$

eine lineare Abbildung. Dann ist das charakteristische Polynom χ_f ein Vielfaches des Minimalpolynoms μ_f zu f .

Beweis. Dies folgt direkt aus Satz 24.2 und Korollar 20.12. □

Insbesondere ist der Grad des Minimalpolynoms zu $\varphi: V \rightarrow V$ durch die Dimension des Vektorraums V beschränkt. Minimalpolynom und charakteristisches Polynom stimmen in verschiedener Hinsicht überein, beispielsweise besitzen sie die gleichen Nullstellen.

LEMMA 24.4. *Es sei V ein endlichdimensionaler Vektorraum über einem Körper K und es sei*

$$f: V \longrightarrow V$$

eine lineare Abbildung. Es sei $v \in V$ ein Eigenvektor von f zum Eigenwert λ und es sei $P \in K[X]$ ein Polynom. Dann ist

$$(P(f))(v) = P(\lambda)v.$$

Insbesondere ist v ein Eigenvektor von $P(f)$ zum Eigenwert $P(\lambda)$. Der Vektor $v \neq 0$ gehört genau dann zum Kern von $P(f)$, wenn λ eine Nullstelle von P ist.

Beweis. Es ist

$$(f^k)(v) = \lambda^k v.$$

Daher folgt alles daraus, dass die Zuordnung $P \mapsto P(f)$ mit der Addition und der Skalarmultiplikation verträglich ist. □

LEMMA 24.5. *Es sei V ein endlichdimensionaler Vektorraum über einem Körper K und es sei*

$$f: V \longrightarrow V$$

eine lineare Abbildung. Dann besitzen das charakteristische Polynom χ_f und das Minimalpolynom μ_f die gleichen Nullstellen.

Beweis. Dass die Nullstellen des Minimalpolynoms auch Nullstellen des charakteristischen Polynoms sind, folgt direkt aus Cayley-Hamilton. Umgekehrt sei $\lambda \in K$ eine Nullstelle des charakteristischen Polynoms und sei $v \in V$ ein Eigenvektor von f zum Eigenwert λ , den es nach Satz 23.2 gibt. Das Minimalpolynom schreiben wir als

$$\mu_f = (X - \lambda_1)^{m_1} \cdots (X - \lambda_k)^{m_k} Q,$$

wobei Q nullstellenfrei sei. Dann ist

$$\begin{aligned} 0 &= \mu_f(f) \\ &= ((X - \lambda_1)^{m_1} \cdots (X - \lambda_k)^{m_k} Q)(f) \\ &= (f - \lambda_1 \text{Id}_V)^{m_1} \cdots (f - \lambda_k \text{Id}_V)^{m_k} Q(f). \end{aligned}$$

Wir wenden dies auf v an. Nach Lemma 24.4 bilden die Faktoren den Vektor v auf $(\lambda - \lambda_i)^{m_i} v$ bzw. auf $Q(\lambda)v$ ab. Insgesamt wird somit v auf

$$(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k} Q(\lambda)v$$

abgebildet. Da die Gesamtabbildung die Nullabbildung und $Q(\lambda) \neq 0$ ist, muss ein $\lambda_i = \lambda$ sein.

□

Weitere Beispiele

Den folgenden Begriff werden wir im Moment ausschließlich für invertierbare Matrizen anwenden.

DEFINITION 24.6. Es sei G eine Gruppe und $g \in G$ ein Element. Dann nennt man die kleinste positive Zahl n mit $g^n = e_G$ die *Ordnung* von g . Man schreibt hierfür $\text{ord}(g)$. Wenn alle positiven Potenzen von g vom neutralen Element verschieden sind, so setzt man $\text{ord}(g) = \infty$.

Wir betrachten lineare Abbildungen

$$\varphi: V \longrightarrow V$$

mit der Eigenschaft, dass eine Potenz davon die Identität ist, sagen wir

$$\varphi^k = \text{Id}_V,$$

dass φ also endliche Ordnung besitzt. Typische Beispiele sind Drehungen um einen Winkel der Form $\frac{360}{k}$ oder Permutationsmatrizen. Das Polynom $X^k - 1$ annulliert dann diesen Endomorphismus und ist daher ein Vielfaches des Minimalpolynoms.

DEFINITION 24.7. Es sei K ein Körper und $n \in \mathbb{N}$. Dann heißen die Nullstellen des Polynoms

$$X^n - 1$$

in K die n -ten *Einheitswurzeln* in K .

LEMMA 24.8. *Es sei $n \in \mathbb{N}_+$. Die Nullstellen des Polynoms $X^n - 1$ über \mathbb{C} sind*

$$e^{2\pi ik/n} = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, k = 0, 1, \dots, n-1.$$

In $\mathbb{C}[X]$ gilt die Faktorisierung

$$X^n - 1 = (X - 1)(X - e^{2\pi i/n}) \dots (X - e^{2\pi i(n-1)/n}).$$

Beweis. Der Beweis verwendet einige Grundtatsachen über die komplexe Exponentialfunktion. Es ist

$$(e^{2\pi ik/n})^n = e^{2\pi ik} = (e^{2\pi i})^k = 1^k = 1.$$

Die angegebenen komplexen Zahlen sind also wirklich Nullstellen des Polynoms $X^n - 1$. Diese Nullstellen sind alle untereinander verschieden, da aus

$$e^{2\pi ik/n} = e^{2\pi i\ell/n}$$

mit $0 \leq k \leq \ell \leq n-1$ sofort, durch Betrachten des Quotienten, $e^{2\pi i(\ell-k)/n} = 1$ folgt, und daraus

$$\ell - k = 0.$$

Es gibt also n explizit angegebene Nullstellen und daher müssen dies alle Nullstellen des Polynoms sein. Die explizite Beschreibung in Koordinaten folgt aus der eulerschen Formel. □

DEFINITION 24.9. Zu einer Permutation π auf $\{1, \dots, n\}$ nennt man die $n \times n$ -Matrix

$$M_\pi = (a_{ij}),$$

für die

$$a_{\pi(j),j} = 1$$

ist und sonst alle Einträge 0 sind, eine *Permutationsmatrix*.

Wir wollen das charakteristische Polynom zu einer Permutationsmatrix bestimmen. Dabei verwenden wir, dass eine Permutation ein Produkt von Zykeln ist. Zu einem Zykel der Form $1 \mapsto 2 \mapsto 3 \mapsto \dots \mapsto k \mapsto 1$ gehört die Permutationsmatrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Jeder Zykel kann (durch Umnummerierung) auf diese Gestalt gebracht werden.

LEMMA 24.10. *Das charakteristische Polynom einer Permutationsmatrix M_ρ zu einem Zykel $\rho \in S_n$ der Ordnung k ist*

$$\chi_M = (X - 1)^{n-k}(X^k - 1).$$

Beweis. Wir können von einem Zykel der Form $1 \mapsto 2 \mapsto 3 \mapsto \dots \mapsto k \mapsto 1$ ausgehen. Die zugehörige Permutationsmatrix M_ρ ist bezüglich e_{k+1}, \dots, e_n die Einheitsmatrix und hat bezüglich der ersten k Standardvektoren die Gestalt

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Die Determinante zu $XE_n - M_\rho$ ist $(X - 1)^{n-k}$ multipliziert mit der Determinante von

$$\begin{pmatrix} X & 0 & \dots & 0 & -1 \\ -1 & X & 0 & \dots & 0 \\ 0 & -1 & X & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & X \end{pmatrix}.$$

Die Entwicklung nach der ersten Zeile liefert

$$X^k + (-1)^{k+1}(-1)(-1)^{k-1} = X^k - 1.$$

□

LEMMA 24.11. *Zu einer Permutationsmatrix M_ρ über \mathbb{C} zu einem Zykel $\rho \in S_n$ mit $\rho : i_1 \mapsto i_2 \mapsto \dots \mapsto i_k \mapsto i_1$ und einer k -ten Einheitswurzel ζ sind die Vektoren*

$$v_\zeta := \zeta^{k-1}e_{i_1} + \zeta^{k-2}e_{i_2} + \dots + \zeta e_{i_{k-1}} + e_{i_k}$$

Eigenvektoren von M_ρ zum Eigenwert ζ . Insbesondere ist eine Permutationsmatrix zu einem Zykel über \mathbb{C} diagonalisierbar.

Beweis. Es ist

$$\begin{aligned} M_\rho(v_\zeta) &= M_\rho(\zeta^{k-1}e_{i_1} + \zeta^{k-2}e_{i_2} + \dots + \zeta e_{i_{k-1}} + e_{i_k}) \\ &= \zeta^{k-1}M_\rho(e_{i_1}) + \zeta^{k-2}M_\rho(e_{i_2}) + \dots + \zeta M_\rho(e_{i_{k-1}}) + M_\rho(e_{i_k}) \\ &= \zeta^{k-1}e_{i_2} + \zeta^{k-2}e_{i_3} + \dots + \zeta e_{i_k} + e_{i_1} \\ &= e_{i_1} + \zeta^{k-1}e_{i_2} + \zeta^{k-2}e_{i_3} + \dots + \zeta e_{i_k} \\ &= \zeta(\zeta^{k-1}e_{i_1} + \zeta^{k-2}e_{i_2} + \dots + \zeta e_{i_{k-1}} + e_{i_k}) \\ &= \zeta v_\zeta. \end{aligned}$$

Da es k verschiedene k -te Einheitswurzeln in \mathbb{C} gibt, sind diese Vektoren nach Lemma 22.3 linear unabhängig und erzeugen einen k -dimensionalen Untervektorraum U von K^n , und zwar gilt

$$U = \langle e_{i_j}, j = 1, \dots, k \rangle.$$

Da die Vektoren $e_i, i \neq i_j$, Fixvektoren sind, bilden die v_ζ zusammen mit den $e_i, i \neq i_j$, eine Basis aus Eigenvektoren von M_ρ und daher ist M_ρ diagonalisierbar.

□

SATZ 24.12. *Eine Permutationsmatrix ist über \mathbb{C} diagonalisierbar.*

Beweis. Siehe Aufgabe 24.23.

□

Abbildungsverzeichnis

- Quelle = Arthur Cayley.jpg , Autor = Benutzer Zuirdj auf Commons,
Lizenz = PD 1
- Quelle = WilliamRowanHamilton.jpeg , Autor = Benutzer auf PD,
Lizenz = 2
- Erläuterung: Die in diesem Text verwendeten Bilder stammen aus
Commons (also von <http://commons.wikimedia.org>) und haben eine
Lizenz, die die Verwendung hier erlaubt. Die Bilder werden mit ihren
Dateinamen auf Commons angeführt zusammen mit ihrem Autor
bzw. Hochlader und der Lizenz. 9
- Lizenzklärung: Diese Seite wurde von Holger Brenner alias
Bocardodarapti auf der deutschsprachigen Wikiversity erstellt und
unter die Lizenz CC-by-sa 3.0 gestellt. 9