Daniel Plaumann
Technische Universität Dortmund
Sommersemester 2017

9. ENDLICH ERZEUGTE MODULN UND GANZHEIT

Arbeitsblatt: Der Satz von Cayley-Hamilton und Anwendungen

Lesen Sie den Text sorgfältig und lösen Sie möglichst viele der Übungsaufgaben. Diskutieren Sie die Lösungen.

Sei im folgenden wie immer R ein kommutativer Ring mit Eins.

Der klassische Satz von Cayley-Hamilton ist aus der linearen Algebra bekannt: Ist A eine $n \times n$-Matrix über einem Körper K und $f = \det(tI_n - A) \in K[t]$ ihr charakteristisches Polynom, dann gilt $f(A) = 0$. Unser Ziel ist eine analoge Aussage für endlich erzeugte Moduln.

Übung 9.1. Die kleine Rechnung $\det(AI_n - A) = \det(0) = 0$ ist *kein* Beweis des Satzes von Cayley-Hamilton, sondern ein Missverständnis der Aussage. Warum?

Übung 9.2. Überprüfen Sie den Satz von Cayley-Hamilton an einer 2×2-Matrix Ihrer Wahl (vielleicht nicht gerade einer Diagonalmatrix…).

Als erstes brauchen wir den Matrizen- und Determinantenkalkül über Ringen: Sei M ein R-Modul und φ ein **Endomorphismus** von M, also ein Homomorphismus $M \to M$ von R-Moduln. Die Endomorphismen bilden einen Ring $\mathrm{End}(M)$ mit der Addition $(\varphi + \psi)(x) = \varphi(x) + \psi(x)$ für $x \in M$ und der Multiplikation $(\varphi\psi)(x) = \varphi(\psi(x))$ gegeben durch Komposition. Dieser Ring ist fast nie kommutativ.

Ist M endlich erzeugt mit n Erzeugern $x_1, \ldots, x_n$ und ist $\varphi \in \mathrm{End}(M)$, dann können wir die Bilder $\varphi(x_1), \ldots, \varphi(x_n)$ wieder durch $x_1, \ldots, x_n$ darstellen, also

$$\varphi(x_i) = \sum_{j=1}^{n} a_{ij} x_j$$

mit $a_{ij} \in R$ schreiben, und bekommen wie gewohnt eine darstellende Matrix $A \in \mathrm{Mat}_n(R)$ von φ bezüglich der Erzeuger $x_1, \ldots, x_n$. (Die brauchen keine Basis zu bilden). Wir schreiben I_n für die Einheitsmatrix der Größe n.

Die **Determinante** einer Matrix $A \in \mathrm{Mat}_n(R)$ definieren wir wie gewohnt durch die Leibniz-Formel

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Vieles gilt wie gewohnt, ohne dass wir das im einzelnen nachprüfen: Zum Beispiel ist die Determinante multiplikativ:

$$\det(AB) = \det(A)\det(B).$$

Vertauscht man zwei Zeilen, dann ändert sich das Vorzeichen, und wenn zwei Zeilen gleich sind, ist die Determinante 0. Außerdem ist die Determinante linear in jeder einzelnen Zeile und jeder einzelnen Spalte (multilinear), woraus sich der Entwicklungssatz ergibt: Zu $A \in \mathrm{Mat}_n(R)$ und (i,j) mit $1 \leqslant i,j \leqslant n$ bilden wir den $(n-1)$-**Minor**

$$\det(A[i,j]),$$

wobei $A[i,j]$ aus A durch Streichung der i-ten Zeile und der j-ten Spalte entsteht.

Proposition 9.1 (Entwicklungssatz von Laplace)**.** *Für jedes $k \in \{1,\dots,n\}$ gilt*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{i,k} \det(A[i,k])$$

(Entwicklung nach der k-ten Spalte).

Beweisskizze. Unter Beachtung der Vorzeichen kann man auf den Fall $k = 1$ reduzieren. Dann beweist man die Aussage, indem man die Matrix A in n Summanden mit jeweils nur einem Eintrag ungleich 0 in der ersten Spalte zerlegt. Die Behauptung ergibt sich dann aus der Leibniz-Formel. ∎

Alle $(n-1)$-Minoren bilden zusammen die **adjunkte Matrix**

$$A^{\mathrm{adj}} = \left((-1)^{i+j} \det(A[i,j])_{i,j=1,\dots,n}\right)^T.$$

Diese Matrix oder ihre Transponierte heißen oft auch *Cofaktormatrix* oder *Cramer-Matrix*. Der Sinn dieser horrenden Matrix steckt in der folgenden Aussage.

Lemma 9.2. *Es gilt*

$$A \cdot A^{\mathrm{adj}} = A^{\mathrm{adj}} \cdot A = \det(A) \cdot I_n.$$

Beweisskizze. Für die Diagonaleinträge von $A^{\mathrm{adj}} \cdot A$ ist das genau der Entwicklungssatz. Die Nicht-Diagonaleinträge an der Stelle (k,l) mit $k \neq l$ entsprechen gerade der Entwicklung nach der l-ten Spalte der Matrix, die aus A entsteht, indem man die l-te Spalte durch die k-te ersetzt. Diese Matrix hat zwei gleiche Spalten und deshalb die Determinante 0. ∎

Übung 9.3. Folgern Sie: Genau dann ist eine Matrix $A \in \mathrm{Mat}_n(R)$ invertierbar, wenn $\det(A) \in R^*$ eine Einheit ist. (Relevante Beispiele: $R = \mathbb{Z}$, $R = K[x]$.)

All diese Formeln (und die daraus resultierende sogenannte Cramersche Regel für die Lösungen eines inhomogenen Gleichungssystems) sind praktisch nicht besonders nützlich, weil es so mühsam ist, Determinanten auszurechnen. In Beweisen können sie aber sehr hilfreich sein:

Satz 9.3 (Cayley-Hamilton)**.** *Es sei M ein R-Modul, der von n Elementen erzeugt ist und $\varphi\colon M \to M$ ein Homomorphismus. Dann gibt es ein normiertes Polynom f vom Grad n in $R[t]$ mit $f(\varphi) = 0$.*
Zusatz: Ist $I \subset R$ ein Ideal mit $\varphi(M) \subset IM$, dann kann f von der Form $f = t^n + \sum_{j=0}^{n-1} a_j t^j$ mit $a_{n-j} \in I^j$ für $j = 1, \dots, n$ gewählt werden.

Dabei muss man wieder richtig verstehen, was die Einsetzung $f(\varphi)$ bedeutet: Weil $\mathrm{End}(M)$ einen Ring und einen R-Modul bildet, kann man φ potenzieren, mit Elementen aus R skalieren und solche Terme addieren. Formal können wir das Einsetzungslemma für Polynomringe verwenden (siehe Algebra I, Prop. 2.2.4): Die Abbildung $R \to \mathrm{End}(M)$ gegeben durch $a \mapsto a \cdot \mathrm{id}_M$ ist ein Ringhomomorphismus. Dieser setzt fort zu einem Ringhomomorphismus $R[t] \to \mathrm{End}(M)$ mit $t \mapsto \varphi$.

Dadurch wird M zu einem $R[t]$-Modul: Für $x \in M$ und $f \in R[t]$ definieren wir

$$f(t) \cdot x = (f(\varphi))(x).$$

Übung 9.4. Ist $\varphi = \mathrm{id}_M$, dann gilt $f(\varphi)(x) = f(1) \cdot x$ für alle $f \in R[t]$.

Beweis von Satz 9.3. Es seien $x_1, \dots, x_n$ Erzeugende von M und sei φ durch die Matrix $A = (a_{ij})$, also durch

$$\varphi(x_i) = \sum_{j=1}^{n} a_{ij} x_j$$

gegeben. Sei $x = (x_1, \dots, x_n)^T$ (Spaltenvektor). Wegen $tx = \varphi(x)$ gilt dann

$$(t \cdot I_n - A) \cdot x = 0.$$

Multiplikation der linken Seite mit $(tI_n - A)^{\mathrm{adj}}$ ergibt

$$\det(tI_n - A) \cdot I_n \cdot x \;=\; 0,$$

also $\det(tI_n - A)x_i = 0$ für $i = 1, \dots, n$. Weil M von $x_1, \dots, x_n$ erzeugt ist, folgt daraus

$$\det(tI_n - A) \cdot M = 0.$$

Es folgt, dass das Polynom $f(t) = \det(tI_n - A) \in R[t]$ die gewünschte Eigenschaft hat.

Für den Zusatz bemerken wir, dass $\varphi(M) \subset IM$ gerade $a_{ij} \in I$ für alle i, j bedeutet. Aus der Laplace-Entwicklung der Determinanten sieht man dann die Behauptung über die Koeffizienten von f. ∎

Mit Cayley-Hamilton beweisen wir nun eine Reihe von Folgerungen, die sich für Vektorräume auch leicht anders beweisen lassen. Für Moduln sind die Beweise dagegen ziemlich trickreich.

Korollar 9.4. *Sei M ein endlich-erzeugter R-Modul.*

(1) Jeder surjektive Homomorphismus $\alpha\colon M \to M$ ist ein Isomorphismus.

(2) Ist $M \cong R^n$ ein freier R-Modul, dann ist jede Menge von n Erzeugenden von M linear unabhängig, also eine Basis.

(3) Es gilt $R^m \cong R^n$ nur für $m = n$.

Übung 9.5. Begründen Sie diese Aussagen für Vektorräume.

Beweis. (1) Fassen M durch α als $R[s]$-Modul auf, also $sx = \alpha(x)$ für $x \in M$. Sei $I = \langle s \rangle \subset R[s]$. Dass α surjektiv ist sagt gerade, dass $IM = M$ gilt. Wir können deshalb Cayley-Hamilton auf den $R[s]$-Modul M und den Homomorphismus $\varphi = \mathrm{id}_M$ anwenden. Es gibt dann ein Polynom $f \in R[s][t]$ mit $f(\mathrm{id}_M)M = 0$, also $f(1)M = 0$. Außerdem hat f Koeffizienten in I, das heißt f ist von der Form $f = t^n + \sum_{i=0}^{n-1} f_i(s)st^i$. Setze

$$g(s) = -(f_0(s) + \cdots + f_{n-1}(s))s$$

dann ist also $f(1) = 1 - g(s)s$. Es gilt

$$(1 - g(s)s)M = 0$$

und damit $g(\alpha)\alpha = \mathrm{id}_M$. Also ist $g(\alpha)$ das Inverse von α und damit α ein Isomorphismus.

(2) Nach Voraussetzung gibt es einen Isomorphismus $\gamma\colon M \xrightarrow{\sim} R^n$. Sind $x_1, \ldots, x_n \in M$ Erzeuger, dann erhalten wir eine Surjektion $\beta\colon R^n \to M$ gegeben durch $e_i \mapsto x_i$. Also ist $\gamma \circ \beta\colon M \to M$ surjektiv und damit nach (1) ein Isomorphismus. Also ist auch $\beta = (\beta \circ \gamma) \circ \gamma^{-1}$ ein Isomorphismus. Das zeigt, dass $x_1, \ldots, x_n$ eine Basis bilden.

(3) Es sei $\alpha\colon R^m \xrightarrow{\sim} R^n$ ein Isomorphismus, ohne Einschränkung mit $m \leqslant n$, und sei $e_1, \ldots, e_m$ eine Basis von R^m. Setze $y_{m+1} = \cdots y_n = 0$. Weil α ein Isomorphismus ist, sind $\alpha(e_1), \ldots, \alpha(e_m)$ Erzeugende von R^n und damit auch $\alpha(e_1), \ldots, \alpha(e_m), y_{m+1}, \ldots, y_n$. Nach (2) ist dieses System linear unabhängig. Weil 0 aber niemals linear unabhängig ist, folgt daraus $m = n$. ∎

Korollar 9.5. *Ist M ein endlich erzeugter R-Modul und I ein Ideal in R mit $IM = M$, dann gibt es $r \in I$ mit $rx = x$ für alle $x \in M$, also $(1 - r)M = 0$.*

Beweis. Sei $\varphi = \mathrm{id}_M$ und wende Cayley-Hamilton an: Wir erhalten wieder ein Polynom $f \in R[t]$ mit $f(1)M = 0$. Ist $f = t^n + \sum_{i=0}^{n-1} a_i t^i$ so bedeutet das also

$$(1 + a_0 + \cdots + a_{n-1})M = 0$$

mit $a_{n-j} \in I^j$. Setze also $r = -(a_0 + \cdots + a_{n-1})$. ∎

Korollar 9.6 (Lemma von Nakayama)**.** *Sei R ein lokaler Ring mit maximalem Ideal Q und M ein endlich erzeugter R-Modul.*

(1) Falls $QM = M$, so ist $M = 0$.

(2) Genau dann sind $x_1, \ldots, x_n \in M$ Erzeuger von M, wenn ihre Restklassen im Faktormodul M/QM Erzeuger sind.

Beweis. (1) Nach dem vorangehenden Korollar gibt es $r \in Q$ mit $(1-r)M = 0$. Aber da r im einzigen maximalen Ideal von R liegt, ist $1-r$ nach Lemma 7.4 eine Einheit. Also folgt $M = 1M = (1-r)(1-r)^{-1}M = 0$.

(2) Sei $N = M/\langle x_1, \ldots, x_n\rangle$. Wir müssen $N = 0$ zeigen. Nach Voraussetzung sind die Restklassen von $x_1, \ldots, x_n$ in M/QM Erzeuger dieses Moduls. Das bedeutet gerade $M = \langle x_1, \ldots, x_n\rangle + QM$. Unter Verwendung der Isomorphiesätze aus Aufgabe 23 folgt

$$QN = QM/(\langle x_1, \ldots, x_n\rangle \cap QM) \cong (QM + \langle x_1, \ldots, x_n\rangle)/\langle x_1, \ldots, x_n\rangle = M/\langle x_1, \ldots, x_n\rangle = N$$

und nach (1) damit $N = 0$. ∎

Übung 9.6. Überzeugen Sie sich, dass das Lemma von Nakayama allgemeiner für endlich erzeugte Moduln über einem beliebigen kommutativen Ring R und für jedes Ideal I, das im Durchschnitt aller maximalen Ideale von R enthalten ist (Jacobson-Radikal), gültig ist.