Normalformen von Matrizen

Klaus Bongartz

8. Mai 2006

1 Vorbemerkungen

Bekanntlich sind nicht alle Endomorphismen oder Matrizen diagonalisierbar. Aus theoretischen und praktischen Gründen möchte man aber dennoch eine Liste von 'Normalformen' haben, d.h. eine Liste möglichst einfacher Matrizen, in der jede Matrix bis auf Ähnlichkeit vorkommt. Ferner hätte man gern ein konstruktives Verfahren, um zu einer vorgegebenen Matrix die zugehörige 'Normalform' zu bestimmen und so insbesondere zu entscheiden, ob zwei vorgelegte Matrizen ähnlich sind oder nicht.

Es gibt nun zwei Typen von Normalformen, die rationale Normalform - abgekürzt als RNF - und die Jordansche Normalform, kurz JNF. Für theoretische Zwecke ist meist die RNF vorzuziehen: Sie ist beispielsweise über beliebigen Grundkörpern definiert und lässt sich explizit in endlich vielen Schritten nur durch rationale Rechenoperationen (d.h. Addition, Subtraktion, Multiplikation und Division) aus den Koeffizienten der Ausgangsmatrix berechnen. Hingegen ist die JNF, die die Kenntnis der Eigenwerte voraussetzt, besser geeignet für die meisten Anwendungen, etwa für die Berechnung hoher Potenzen der Ausgangsmatrix oder für das explizite Lösen linearer Differentialgleichungssysteme mit konstanten Koeffizienten. Wir werden in diesem Kapitel zunächst die RNF und danach die JNF völlig unabhängig voneinander behandeln.

Oft werden die Normalformen hergeleitet aus den entsprechenden Struktursätzen für Moduln über Hauptidealringen. Dazu müssen die Studenten kritisch analysieren, was von ihren gerade erst erlernten Kenntnissen über Vektorräume für Moduln gültig bleibt. Außerdem wird aus Zeitgründen die Spezialisierung der allgemeinen Struktursätze auf die Existenz und Eindeutigkeit der Normalformen häufig so knapp behandelt, dass man danach bestenfalls weiß, wie die Normalformen aussehen, aber nicht, wie man zu einer gegebenen Matrix die zugehörigen Normalformen und einen geeigneten Basiswechsel findet. Dieser algorithmische Aspekt steht in der nachfolgenden Behandlung stets im Vordergrund. Ferner verlassen wir nie die Sprache der Vektorräume, obwohl natürlich implizit durch Einsetzen von Matrizen oder Endomorphismen in Polynome auch Moduln über dem Polynomring betrachtet werden. Für ein durchdringendes Verständnis ist dieser Standpunkt natürlich unerlässlich, und jeder Student wird ihn später in der Algebravorlesung kennenlernen.

2 Das lokale Minimalpolynom

Lemma 1 *Sei A aus $k^{n\times n}$ und v aus k^n ein beliebiger Spaltenvektor. Dann gilt:*

a)Es gibt ein eindeutig bestimmtes normiertes Polynom $\mu_{A,v}$ minimalen Grades mit $\mu_{A,v}(A)v = 0$. Jedes Polynom H mit $H(A)v = 0$ wird von $\mu_{A,v}$ geteilt.

b)Sei i der Grad von $\mu_{A,v}$. Der von den Potenzen $A^j v$, $j = 0, 1, 2, \ldots i-1$ erzeugte Unterraum $M(v)$ ist der kleinste A-stabile Unterraum, der v enthält.

c) Die Vektoren $A^j v$, $j = 0, 1, 2, \ldots i-1$ bilden eine geordnete Basis, bezüglich der der von A induzierte Endomorphismus von $M(v)$ gerade durch die Begleitmatrix $B(\mu_{A,v})$ dargestellt wird.

Beweis: Man berechnet solange iterierte Bilder $v, Av, A^2v, \ldots, A^i v$ von v unter A, bis diese linear abhängig werden. Man erhält dann also eine Gleichung

$$A^i v + a_{i-1}A^{i-1}v + \ldots a_0 v = 0.$$

Das Polynom $P = X^i + a_{i-1}X^{i-1} + \ldots a_0$ ist das gesuchte Polynom. Gilt nämlich $H(A)v = 0$, so liefert Division mit Rest im Polynomring die Gleichung $H = QP + S$ für ein Polynom S echt kleineren Grades als P. Ist dabei $S = s_m X^m + s_{m-1}X^{m-1} + \ldots s_0$ vom Nullpolynom verschieden, so erhält man

$$0 = H(A)v = Q(A)P(A)v + S(A)v,$$

also $S(A)v = 0$. Dann sind aber schon die Vektoren $v, Av, A^2v, \ldots, A^m v$ mit $m \leq i-1$ linear abhängig im Widerspruch zur Definition von P. Somit ist $S = 0$ und P ein Teiler von H.

Die beiden anderen Aussagen sieht man leicht ein.

q.e.d.

Wir nennen das Polynom $\mu_{A,v}$ aus dem Lemma das lokale Minimalpolynom von A an der Stelle v. Natürlich ist dies ein Teiler des 'globalen' Minimalpolynoms μ_A. Klar ist auch, dass das Minimalpolynom das kleinste gemeinsame Vielfache der lokalen Minimalpolynome an den kanonischen Basisvektoren e_j für $j = 1, \ldots, n$ ist. Wie der nächste Satz zeigt, gibt es aber erstaunlicherweise sogar immer einen Vektor x, dessen lokales Minimalpolynom das globale Minimalpolynom ist. Darüber hinaus kann man einen solchen Vektor immer in endlich vielen Schritten mit Hilfe des Euklidischen und des Gauß-Algorithmus konstruieren.

Satz 1 *Sei eine Matrix A in $k^{n\times n}$ gegeben.*

a) Man kann zu zwei Vektoren v und w einen Vektor u konstruieren, so dass $\mu_{A,u}$ das kleinste gemeinsame Vielfache V von $P = \mu_{A,v}$ und $Q = \mu_{A,w}$ ist.

b) Man kann einen Vektor x konstruieren, so dass $\mu_{A,x}$ das Minimalpolynom von A ist.

Beweis: a) Sei G der mit dem Euklidischen Algorithmus berechnete größte gemeinsame Teiler von P und Q, dessen Grad g sei. Dann ist also $P = G\tilde{P}$

und $Q = G\tilde{Q}$ mit teilerfremden normierten Polynomen $\tilde{P}$ und $\tilde{Q}$. Sei H der wieder mit dem Euklidischen Algorithmus berechnete größte gemeinsame Teiler von G und $\tilde{Q}^g$. Es ist dann $G = HK$, wobei in H nur Primfaktoren von $\tilde{Q}$ auftreten, in K hingegen keine. Ist nämlich $\tilde{Q} = L_1^{n_1} L_2^{n_2} \dots L_t^{n_t}$ und $G = L_1^{m_1} L_2^{m_2} \dots L_t^{m_t} L_{t+1}^{m_{t+1}} L_{t+2}^{m_{t+2}} \dots L_s^{m_s}$ mit paarweise verschiedenen irreduziblen Polynomen L_i, $n_i \geq 1$ und $m_i \geq 0$, so ist $m_i \leq gn_i$ für $i \leq t$, also $H = L_1^{m_1} L_2^{m_2} \dots L_t^{m_t}$ und $K = L_{t+1}^{m_{t+1}} L_{t+2}^{m_{t+2}} \dots L_s^{m_s}$. Für das kleinste gemeinsame Vielfache $V = G\tilde{Q}\tilde{P}$ von P und Q hat man also eine Zerlegung in zwei teilerfremde Faktoren $H\tilde{Q}$ und $K\tilde{P}$ gefunden.

Der gesuchte Vektor ist nun $u = H(A)v + K(A)w$. Denn $V(A)u = 0$ ist klar. Sei L ein weiteres Polynom mit $L(A)u = 0$, d.h. $L(A)H(A)v = -L(A)K(A)w$. Anwenden von $K(A)\tilde{P}(A)$ auf beiden Seiten liefert

$$0 = L(A)P(A)v = -L(A)K^2(A)\tilde{P}(A)w.$$

Also teilt $Q = HK\tilde{Q}$ das Polynom $\tilde{P}K^2L$. Wegen der Teilerfremdheit von $K\tilde{P}$ und $H\tilde{Q}$ wird sogar L von $H\tilde{Q}$ geteilt. Symmetrisch sieht man, dass L auch von $K\tilde{P}$ geteilt wird, also wieder wegen der Teilerfremdheit sogar von $H\tilde{Q}K\tilde{P}$, also von V. Es folgt $\mu_{A,u} = V$ wie gewünscht.

b) Sei zuerst $v \neq 0$ ein beliebiger Vektor. Man berechnet dazu $P = \mu_{A,v}$ wie in Lemma 1. Ist dann $Kern\, P(A) = k^n$, so ist $P = \mu_A$ und v der gesuchte Vektor.

Anderenfalls konstruieren wir einen neuen Vektor v' mit $P' = \mu_{A,v'}$, derart dass $Kern\, P'(A)$ echt größer als $Kern\, P(A)$ ist. Dazu wählen wir einen beliebigen Vektor w außerhalb von $Kern\, P(A)$, bestimmen $Q = \mu_{A,w}$ und konstruieren u wie in Teil a). Der gesuchte Vektor ist dann $v' = u$ mit $P' = V$. Da P und Q Teiler von P' sind, ist nämlich $Kern\, P(A)$ echt in $Kern\, P'(A)$ enthalten. Ist nun $Kern\, P'(A) = k^n$, so ist analog zum ersten Fall v' der gesuchte Vektor. Anderenfalls wiederholen wir das soeben durchgeführte Verfahren, das aus Dimensionsgründen nach spätestens $n-1$-maliger Anwendung mit dem Fall $Kern\, P'(A) = k^n$ endet.

q.e.d.

In der 'Praxis' endet das in b) beschriebene Verfahren meist sehr schnell. Die Vektoren v mit $\mu_{A,v} \neq \mu_A$ liegen nämlich in der Vereinigung der endlich vielen echten Teilräume $Kern\, P(A)$, wobei P ein normierter echter Teiler von μ_A ist. Nun ist über einem unendlichen Körper ein Vektorraum niemals die Vereinigung endlich vieler echter Unterräume (Beweis als Übung) und die Chance, bei zufälliger Wahl eines Vektors v direkt beim Start des Verfahrens einen gesuchten 'guten' Vektor zu finden, liegt bei nahezu hundert Prozent, weil die guten Vektoren etwa für $\mathbf{R}$ eine offene dichte Teilmenge von $\mathbf{R^n}$ bilden.

Entgegen der in manchen Lehrbüchern vertretenen Meinung ist also die Berechnung des Minimalpolynoms auch ohne Kenntnis der Zerlegung des charakteristischen Polynoms in Linearfaktoren recht einfach. Es ist sogar nach Lemma 1 im Gegenteil so, dass bei beliebiger Wahl von $v \neq 0$ die Matrix A ähnlich ist zur Matrix

$$\begin{bmatrix} B(\mu_{A,v}) & A_2 \\ 0 & A_4 \end{bmatrix},$$

was einen induktiven Weg weist zur Berechnung des charakteristischen Polynoms von A und seiner Nullstellen. Dabei sei allerdings nicht verschwiegen, dass für die 'meisten' Matrizen Minimalpolynom und charakteristisches Polynom sowieso übereinstimmen, da etwa über den reellen Zahlen diese Matrizen wieder eine offene dichte Teilmenge bilden.

Es wird nun höchste Zeit, dass wir all diese Gedankengänge an einem numerischen Beispiel illustrieren.

Sei A die folgende ganzzahlige 4×4-Matrix

$$\begin{bmatrix} 5 & -1 & -1 & -13 \\ 2 & 0 & -1 & -7 \\ -4 & 2 & 3 & 14 \\ 2 & -1 & -1 & -6 \end{bmatrix}.$$

Wir wollen einen Vektor x mit $\mu_{A,x} = \mu_A$ konstruieren. Wir versuchen es mit dem ersten kanonischen Basisvektor e_1 und erhalten:

$$e_1, Ae_1 = \begin{bmatrix} 5 \\ 2 \\ -4 \\ 2 \end{bmatrix}, A^2 e_1 = e_1.$$

Also ist $X^2 - 1 = \mu_{A,e_1}$. Der zweite kanonische Basisvektor e_2 liegt nicht im Kern von $A^2 - 1$. Wir bestimmen μ_{A,e_2} und finden

$$e_2, Ae_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \\ -1 \end{bmatrix}, A^2 e_2 = \begin{bmatrix} 6 \\ 3 \\ -4 \\ 2 \end{bmatrix}, A^3 e_2 = \begin{bmatrix} 5 \\ 2 \\ -2 \\ 1 \end{bmatrix},$$

woraus $\mu_{A,e_2} = X^3 - X^2 - X + 1 = (X^2 - 1)(X - 1)$ folgt. Also ist μ_{A,e_2} das kleinste gemeinsame Vielfache von μ_{A,e_1} und μ_{A,e_2}, und wir können einfach e_2 als den nach Lemma 2 a) zu konstruierenden Vektor u nehmen. Man überprüft, dass auch $\mu_{A,e_3} = X^3 - X^2 - X + 1 = (X^2 - 1)(X - 1)$ gilt. Somit ist $P_1 = X^3 - X^2 - X + 1$ das Minimalpolynom, weil der Kern von $P_1(A)$ die ersten drei kanonischen Basisvektoren enthält sowie Vektoren , deren vierte Komponente nicht verschwindet.

3 Die rationale Normalform

Definition 1 *Eine rationale Normalform R vom Format n über k ist eine Blockdiagonalmatrix*

$$\begin{bmatrix} B(P_1) & 0 & ... & 0 \\ 0 & B(P_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & B(P_r) \end{bmatrix}$$

mit lauter Begleitmatrizen $B(P_1), B(P_2), \dots, B(P_r)$ auf der Hauptdiagonalen, wobei alle Polynome Koeffizienten in k haben, und jeweils P_{i+1} ein Teiler von P_i ist für $1 \leq i \leq r-1$. Wir schreiben dafür $R = R(P_1, P_2, \dots, P_r)$.

Da eine Begleitmatrix $B(Q)$ stets Q als Minimalpolynom und als charakteristisches Polynom hat, ist also P_1 das Minimalpolynom von $R(P_1, P_2, \dots, P_r)$ und das Produkt $P_1 P_2 \dots P_r$ das charakteristische Polynom.

Wir formulieren nun das Hauptergebnis dieses Kapitels.

Satz 2 *Jede Matrix A aus $k^{n \times n}$ ist zu genau einer rationalen Normalform $R = R(P_1, P_2, \dots, P_r)$ ähnlich über k. Dabei lassen sich sowohl R als auch eine konjugierende Matrix T durch endlich viele rationale Rechenoperationen aus den Koeffizienten von A bestimmen.*

Wir geben vor dem Beweis einige wichtige Folgerungen an.

Folgerung 1 *Genau dann sind zwei Matrizen ähnlich über k, wenn sie gleiche RNF haben. Dies lässt sich durch endlich viele rationale Rechenoperationen entscheiden.*

Eine RNF über k bleibt auch in jedem Erweiterungskörper K eine RNF. Dies impliziert die folgende auf den ersten Blick erstaunliche Tatsache.

Folgerung 2 *Sei k in einem anderen Körper K enthalten. Genau dann sind zwei Matrizen A, B aus $k^{n \times n}$ ähnlich über K, wenn sie schon über k ähnlich sind.*

Schließlich erhalten wir einen unabhängigen Beweis für den Satz von Cayley-Hamilton.

Folgerung 3 *Für A in $k^{n \times n}$ ist stets das Minimalpolynom μ_A ein Teiler des charakteristischen Polynoms χ_A und χ_A ein Teiler von $(\mu_A)^n$.*

Charakteristisches Polynom und Minimalpolynom ändern sich nicht beim Übergang zu ähnlichen Matrizen. Wir können daher A als RNF $R(P_1, P_2, \dots, P_r)$ annehmen, und dann gilt die Behauptung nach obiger Bemerkung.

Wir beweisen als erstes die Eindeutigkeit der rationalen Normalform.

Lemma 2 *Sind zwei rationale Normalformen ähnlich, so sind sie schon gleich.*

Beweis: Seien also $R = R(P_1, P_2, \ldots, P_r)$ und $R' = R(Q_1, Q_2, \ldots, Q_s)$ rationale Normalformen über k, die ähnlich sind über k. Es gibt dann eine invertierbare Matrix T mit Koeffizienten in k, so dass $T^{-1}RT = R'$ gilt. Für jedes Polynom $P(X)$ in $k[X]$ sind dann auch $P(R)$ und $P(R')$ konjugiert unter T.

Ferner ist bei einer Blockdiagonalmatrix M mit Blöcken M_i auf der Hauptdiagonalen $P(M)$ wieder Blockdiagonalmatrix mit den $P(M_i)$ auf der Diagonalen. Der Rang von einer Blockdiagonalmatrix ist die Summe der Ränge der Diagonalblöcke.

Mit Hilfe dieser einfachen Bemerkungen zeigen wir nun sukzessive die Beziehungen $P_i = Q_i$ und $r = s$. Der Induktionsanfang ist klar, weil P_1 das Minimalpolynom von R und Q_1 das von R' ist, und weil ähnliche Matrizen gleiches Minimalpolynom haben.

Im Induktionsschritt sei für $j \leq i$ bereits $P_j = Q_j$ gezeigt. Falls $i = r$ oder $i = s$, so gilt auch $r = s$, da die Summe der Grade aller Polynome jeweils n ergibt. Sonst betrachten wir $P_{i+1}(R)$ und $P_{i+1}(R')$, die als ähnliche Matrizen gleichen Rang haben. Nun sind bei $P_{i+1}(R)$ aber höchstens die ersten i Diagonalblöcke von Null verschieden, und die ersten i Blöcke von $P_{i+1}(R')$ stimmen per Induktionsannahme mit diesen überein. Aus Ranggründen verschwinden also auch alle anderen Blöcke von $P_{i+1}(R')$. Insbesondere ist $P_{i+1}(B(Q_{i+1})) = 0$, also Q_{i+1} ein Teiler von P_{i+1}. Aus Symmetriegründen folgt der Induktionsschluß.

q.e.d.

Der Nachweis der Existenz und die Konstruktion sind schwieriger, wobei wir allerdings die Hauptarbeit bereits im letzten Abschnitt geleistet haben. Sei also eine Matrix $A_0 = A$ vorgelegt. Wir geben nacheinander für $i = 1, 2, 3$ drei invertierbare Matrizen T_i an, so dass $A_i = T_i^{-1} A_{i-1} T_i$ immer speziellere Gestalt hat und A_3 die gesuchte RNF zu A ist.

Als erstes berechnen wir mit Hilfe des letzten Lemmas das Minimalpolynom P_1 von A und einen Vektor x mit $P_1 = \mu_{A,x}$. Sei im folgenden p_i der Grad der teilweise noch zu bestimmenden P_i. Wir definieren eine erste Transformationsmatrix T_1 durch Angabe der zugehörigen Spaltenvektoren, indem wir die linear unabhängigen Spalten $x, Ax, \ldots, A^{p_1 - 1}x$ zu einer Basis von k^n ergänzen. Die konjugierte Matrix $A_1 = T_1^{-1} A_0 T_1$ ist dann die Darstellungsmatrix der durch A_0 definierten linearen Abbildung bezüglich der Spalten von T_1, hat also per Konstruktion die obere Blocktrigonalgestalt

$$\begin{bmatrix} B(P_1) & B_2 \\ 0 & B_4 \end{bmatrix}.$$

Beim Einsetzen dieser Matrix in P_1 erhält man immer noch die Nullmatrix, weshalb das Minimalpolynom P_2 von B_4 ein Teiler von P_1 ist. Per Induktion nach n haben wir bereits zu B_4 eine invertierbare Matrix T_2' mit $n - p_1$ Zeilen gefunden, derart dass $T_2'^{-1} B_4 T_2'$ eine rationale Normalform ist der Form

$$\begin{bmatrix} B(P_2) & 0 & ... & 0 \\ 0 & B(P_3) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & B(P_r) \end{bmatrix}.$$

Setzt man mit der Einheitsmatrix E_{p_1} in $k^{p_1 \times p_1}$ nun $T_2 = \begin{bmatrix} E_{p_1} & 0 \\ 0 & T_2' \end{bmatrix}$, so hat $A_2 = T_2^{-1} A_1 T_2$ die Gestalt

$$\begin{bmatrix} B(P_1) & C_2 & ... & C_r \\ 0 & B(P_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & B(P_r) \end{bmatrix}.$$

Wir werden anschließend noch durch einen geeigneten Koordinatenwechsel die Matrizen C_i in der ersten Zeile zum Verschwinden bringen, kennen aber jetzt schon die RNF zu A.

Für jeden Index $i \geq 2$ sei v_i der kanonische Basisvektor mit dem Index $p_1 + p_2 + \ldots + p_{i-1} + 1$ und es sei $v_1 = e_1$. Dann ist $A_2^j v_1 = e_{j+1}$ für $j < p_1$. Ferner ist $P_i(A_2)v_i$ gerade die Spalte von $P_i(A_2)$ mit Index $p_1 + p_2 + \ldots + p_{i-1} + 1$. Nur die ersten p_1 Koeffizienten davon sind eventuell nicht Null. Also ist $P_i(A_2)v_i = H_i(A_2)v_1$ für ein eindeutig bestimmtes Polynom H_i vom Grad höchstens $p_1 - 1$. Da P_2 und damit alle P_i ein Teiler des Minimalpolynoms P_1 sind, gibt es Polynome R_i mit $P_1 = P_i R_i$. Daraus ergibt sich

$$0 = P_1(A_2)v_i = (R_i P_i)(A_2)v_i = (R_i H_i)(A_2)v_1.$$

Deshalb ist $P_1 = \mu_{A_2,v_1}$ ein Teiler von $R_i H_i$, also $H_i = P_i S_i$. Setze jetzt $w_i = v_i - S_i(A_2)v_1$ für $i = 2, 3, \ldots, r$. Nach Konstruktion ist dann jeweils $P_i(A_2)w_i = 0$. Ferner ist die Matrix T_3 mit Spalten

$$v_1, A_2 v_1, \ldots, A_2^{p_1-1} v_1, w_2, A_2 w_2, \ldots, A_2^{p_2-1} w_2, \ldots, w_r, A_2 w_r, \ldots, A_2^{p_r-1} w_r$$

eine obere Dreiecksmatrix mit lauter Einsen auf der Diagonalen, also invertierbar. $T_3^{-1} A_2 T_3$ ist dann die gesuchte rationale Normalform

$$\begin{bmatrix} B(P_1) & 0 & ... & 0 \\ 0 & B(P_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & B(P_r) \end{bmatrix}.$$

q.e.d.

Wir wollen den Algorithmus konkret auf unser Beispiel aus dem letzten Abschnitt anwenden, wo wir bereits e_2 als einen Kandidaten für einen guten Vektor x erkannt hatten. Als erste Transformationsmatrix T_1 auf dem Weg zur RNF können wir also jede Matrix mit Spalten e_2, Ae_2, A^2e_2, y nehmen, für die y nicht im Erzeugnis der drei ersten Spalten liegt. Wir wählen für y der Einfachheit halber e_3. Um die letzte Spalte von $A_1 = T_1^{-1}AT_1$ zu erhalten, müssen wir Ae_3 als Linearkombination der Spalten von T_1 schreiben und finden mit dem Gauß-Algorithmus

$$Ae_3 = (-1)e_2 + (1)Ae_2 + (0)A^2e_2 + (1)e_3,$$

also

$$A_1 = \begin{bmatrix} 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Der zweite Schritt im Existenzbeweis für die RNF entfällt, da B_4 bereits die Begleitmatrix $B(X-1)$ ist. Es ist also $r = 2$, $P_2 = X - 1$ und $A_2 = A_1$.

Der dritte Schritt ist nahezu trivial. Es ist $v_2 = e_4$ und $P_2(A_2)v_2 = -e_1 + e_2 = H_2(A_2)v_1$ mit $H_2 = X - 1$. Wie im Beweis des dritten Schrittes behauptet ist nun $H_2 = P_2S_2$, wobei hier $S_2 = 1$ ist. Nach dem allgemeinen Rezept ist dann die Transformationsmatrix T_3 die Matrix mit den Spalten e_1, e_2, e_3 und $e_4 - e_1$.

Die gesuchte Transformationsmatrix ist T_1T_3, also

$$\begin{bmatrix} 0 & -1 & 6 & 0 \\ 1 & 0 & 3 & -1 \\ 0 & 2 & -4 & 1 \\ 0 & -1 & 2 & 0 \end{bmatrix}.$$

4 Die Hauptraumzerlegung

In diesem Abschnitt arbeiten wir zu Abwechslung einmal wieder mit Endomorphismen und empfehlen dem Leser sehr, jeweils die entsprechenden Aussagen für Matrizen selbst zu formulieren.

Lemma 3 *Sei $f : V \to V$ ein Endomorphismus, $Q = Q_1Q_2$ ein Polynom in $k[X]$ mit teilerfremden Polynomen Q_1 und Q_2. Dann sind die Unterräume $Kern\, Q_i(f)$ beide f-stabil und man hat die Zerlegung*

$$Kern\, Q(f) = Kern\, Q_1(f) \oplus Kern\, Q_2(f).$$

Beweis: Konstruiere mit dem Euklidischen Algorithmus Polynome R_1, R_2 in $k[X]$ mit

$$1 = R_1Q_1 + R_2Q_2.$$

Einsetzen von f für die Variable liefert in $EndV$ die Beziehung

$$id = R_1(f)Q_1(f) + R_2(f)Q_2(f).$$

Da jedes Q_i ein Teiler von Q ist, liegen die trivialerweise f-stabilen Unterräume $Kern\, Q_i(f)$ in $Kern\, Q(f)$. Ist v im Durchschnitt der beiden Kerne, so folgt

$$v = id(v) = (R_1(f)Q_1(f) + R_2(f)Q_2(f))v = 0.$$

Sei nun v aus $Kern\, Q(f)$. Dann ist also

$$v = id(v) = R_1(f)Q_1(f)v + R_2(f)Q_2(f)v,$$

wobei der erste Term von $Q_2(f)$, der zweite von $Q_1(f)$ auf Null abgebildet wird.
q.e.d.

Es sei nochmals darauf hingewiesen, dass die Zerlegung eines Polynoms in irreduzible Faktoren oder auch nur in zwei teilerfremde Faktoren ein schwieriges, nicht-lineares Problem ist, dessen Lösung man kennen muß, um das Lemma anwenden zu können. Natürlich kann man dieses schöne Zerlegungslemma per Induktion sofort auf mehr als zwei Faktoren verallgemeinern und findet:

Folgerung 4 *Ist $Q = Q_1Q_2 \dots Q_r$ mit paarweise teilerfremden Polynomen Q_i, so gilt*

$$Kern\, Q(f) = \bigoplus_{i=1}^{r} Kern\, Q_i(f).$$

Besonders wichtig ist der Fall, wo $Q = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$ das Minimalpolynom von f ist, das über k in verschiedene Linearfaktoren zerfällt. Falls dabei alle $m_i = 1$ sind, so besagt die Folgerung gerade, dass V (der Kern von $Q(f)$ die direkte Summe der Eigenräume $E(\lambda_i)$ (der Kerne der $Q_i(f)$) ist, d.h. dass f diagonalisierbar ist. Wir haben also die nichttriviale Implikation bewiesen in folgendem früher ohne Beweis angegebenen Satz.

Folgerung 5 *Genau dann ist ein Endomorphismus über k diagonalisierbar, wenn sein Minimalpolynom in $k[X]$ in lauter verschiedene Linearfaktoren zerfällt.*

Wir definieren nun für jeden Eigenwert λ von f den Hauptraum (oder verallgemeinerten Eigenraum) $H(\lambda)$ als die Menge aller Vektoren v, so dass eine natürliche Zahl m existiert mit $(f - \lambda)^m v = 0$. Man verifiziert sofort, dass ein Hauptraum stets ein f-stabiler Unterraum ist, der den Eigenraum $E(\lambda)$ umfasst.

Satz 3 *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen k-Vektorraumes V. Das charakteristische Polynom χ_f zerfalle über k in verschiedene Linearfaktoren $\chi_f = \prod_{i=1}^{r}(X - \lambda_i)^{n_i}$. Weiter sei $Q = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$ ein Polynom mit $Q(f) = 0$. Dann gilt:*

a) Für jeden Index ist der Hauptraum zum Eigenwert λ_i gerade der Kern von $(f - \lambda_i)^{m_i}$. Seine Dimension ist n_i.

b) V ist die direkte Summe der Haupträume. Auf jedem Hauptraum $H(\lambda_i)$ induziert f einen Endomorphismus f_i, für den $f_i - \lambda_i$ nilpotent ist

Beweis: Da V der Kern von $Q(f)$ ist, besagt obige Folgerung gerade, dass V die direkte Summe der $V_i = Kern\,(f-\lambda_i)^{m_i}$ ist. Klar ist auch, dass jedes V_i im entsprechenden Hauptraum liegt.

Wäre nun etwa $H(\lambda_1)$ echt größer als V_1, so gäbe es ein v und ein $l > m_1$ mit $v \in Kern\,(f-\lambda_1)^l \setminus V_1$. Wieder nach der Folgerung wäre dann V die direkte Summe von $Kern\,(f-\lambda_1)^l$ und den V_i für $i \geq 2$, was aus Dimensionsgründen nicht geht.

Die V_i's sind also wirklich die Haupträume. Nun ist das Minimalpolynom von f_i jeweils ein Teiler von $(X-\lambda_i)^{m_i}$, hat also nur $X-\lambda_i$ als irreduziblen Faktor. Das charakteristische Polynom von f_i ist also $(X-\lambda_i)^{r_i}$, wobei r_i die Dimension von V_i ist. Wegen der f-Stabilität der V_i's ist aber χ_f das Produkt der χ_{f_i}'s. Die Dimensionsaussage über die Haupträume folgt somit aus der Eindeutigkeit der Primfaktorzerlegung in $k[X]$.

q.e.d.

Als Spezialfall dieses Satzes finden wir unser altes Ergebnis, dass f genau dann diagonalisierbar ist, wenn das charakteristische Polynom in Linearfaktoren zerfällt und für alle Eigenwerte die geometrische Vielfachheit mit der algebraischen übereinstimmt.

Wir wollen die Hauptraumzerlegung für $f = f_A$ aus unserem Standardbeispiel durchführen. Wie bereits gesehen ist das Minimalpolynom $\mu_A = X^3 - X^2 - X + 1 = (X+1)(X-1)^2$. Wir bestimmen mit dem Gauß-Algorithmus Basen von $K_1 = Kern\,(A+1)$ und $K_2 = Kern\,(A-1)^2$ und erhalten so die Hauptraumzerlegung. Nehmen wir diese Basen als Spalten einer Transformationsmatrix S_1, so erhalten wir mit $S_1^{-1}AS_1$ die Darstellungsmatrix von f_A bezüglich dieser der Hauptraumzerlegung angepassten Basis von V. Eine mögliche Wahl ist

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

Dann ergibt sich $S_1^{-1}AS_1$ als

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & -2 & -1 & -2 \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

5 Die Jordansche Normalform für nilpotente Endomorphismen

Wieder sei k ein beliebiger Körper.

Definition 2 *Ein Jordankästchen $J(m,\lambda)$ vom Format $m > 0$ zum Eigenwert*

λ aus k ist eine $m \times m$-Matrix folgender Form

$$\begin{bmatrix} \lambda & 0 & 0 & ... & 0 \\ 1 & \lambda & 0 & ... & 0 \\ 0 & 1 & \lambda & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & ... & 1 & \lambda \end{bmatrix}.$$

Eine Jordansche Normalform J über k vom Format n mit r Kästchen ist eine Blockdiagonalmatrix der Gestalt

$$\begin{bmatrix} J(m_1, \lambda_1) & 0 & ... & 0 \\ 0 & J(m_2, \lambda_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & J(m_r, \lambda_r) \end{bmatrix},$$

wobei alle λ_i aus k sind. Es bezeichne $v((J, m, \lambda))$ die Vielfachheit, mit der $J(m, \lambda)$ auf der Hauptdiagonalen von J als Kästchen auftritt.

Matrizen in Jordanscher Normalform kann man also sofort erkennen, während man bei einer Blockdiagonalmatrix mit lauter Begleitmatrizen auf der Hauptdiagonalen zuerst noch die Teilbarkeitsbedingungen überprüfen muss, um sie als RNF zu identifizieren. Dies ist ein Vorteil der JNF.

Es sei nochmals daran erinnert, dass ein Endomorphismus oder eine Matrix nilpotent heißt, wenn eine Potenz davon die Null ist. Das nächste Lemma ist ein einfacher Spezialfall des Satzes über die RNF, da bei nilpotenten Matrizen die RNF auch eine JNF ist. Wir wollen aber einen unabhängigen Beweis haben, der zudem sehr anschaulich ist.

Lemma 4 *Sei V ein Vektorraum der Dimension n und f ein nilpotenter Endomorphismus. Dann kann man ein Koordinatensystem ϕ konstruieren, so dass die zugehörige Darstellungsmatrix $M_\phi(f)$ eine JNF mit 0 als einzigem Eigenwert ist.*

Beweis: Sei $f^{r+1} = 0 \neq f^r$. In $K := Kern\, f$ betrachten wir die Kette von Unterräumen

$$K \cap f^r V \subseteq K \cap f^{r-1} V \subseteq \ldots K \cap fV \subseteq K.$$

Wir konstruieren eine Basis $x_{r,1}, x_{r,2}, \ldots, x_{r,n_r}$ von $K \cap f^r V$, ergänzen diese durch $x_{r-1,1}, x_{r-1,2}, \ldots, x_{r-1,n_{r-1}}$ zu einer Basis von $K \cap f^{r-1}V$, und so weiter. Dabei ist also per Konstruktion n_i die Differenz der Dimensionen von $K \cap f^i V$ und $K \cap f^{i+1} V$, so dass ohne weiteres der Fall $n_i = 0$ auftreten kann. Es ist aber

$$\{x_{i,j} | 0 \leq i \leq r, 1 \leq j \leq n_i\}$$

eine Basis von K. Per Definition gibt es Vektoren $y_{i,j}$ mit $f^i y_{i,j} = x_{i,j}$. Wir werden gleich zeigen, dass die Menge

$$\mathcal{B} = \{f^l y_{i,j} | 0 \leq i \leq r, 1 \leq j \leq n_i, 0 \leq l \leq i\}$$

eine Basis von V ist. Zur Illustration haben wir die Elemente der Menge $\mathcal{B}$ für $r = 3$ in das nachstehende treppenförmige Schema angeordnet:

$$\begin{array}{cccccccccccc}
y_{3,1} & \cdot\cdot & y_{3,n_3} & & & & & & & & & \\
fy_{3,1} & \cdot\cdot & fy_{3,n_3} & y_{2,1} & \cdot\cdot & y_{2,n_2} & & & & & & \\
f^2y_{3,1} & \cdot\cdot & f^2y_{3,n_3} & fy_{2,1} & \cdot\cdot & fy_{2,n2} & y_{1,1} & \cdot\cdot & y_{1,n_1} & & & \\
f^3y_{3,1} & \cdot\cdot & f^3y_{3,n_3} & f^2y_{2,1} & \cdot\cdot & f^2y_{2,n_2} & fy_{1,1} & \cdot\cdot & fy_{1,n_1} & y_{0,1} & \cdot\cdot & y_{0n_0}
\end{array}$$

Dabei bildet f jeden Vektor auf den darunterstehenden Vektor ab oder auf 0, falls kein Vektor darunter steht. Ordnet man diese Basis bei festgehaltenen Indizes i, j so, dass l von 0 bis i läuft, d.h. bewegt man sich im Schema spaltenweise von oben nach unten, so ist die zugehörige Darstellungsmatrix von der gewünschten Form. Außerdem benötigen alle bisherigen Konstruktionen nur Matrizenmultiplikationen und den Gauß-Algorithmus, sind also wirklich berechenbar.

Es bleibt nur zu zeigen, dass $\mathcal{B}$ eine Basis ist. Das ist nicht schwer, aber wie der gesamte Beweis von den Notationen her aufwendig. Man orientiere sich wieder an dem Schema oben.

Sei also $\sum \lambda_{l,i,j} f^l y_{i,j} = 0$. Wir wenden f^r auf diese Gleichung an und erhalten $0 = \sum \lambda_{l,i,j} f^{r+l} y_{i,j} = \sum \lambda_{0,r,j} x_{r,j}$, d.h. $\lambda_{0,r,j} = 0$ für alle j. Danach wenden wir f^{r-1} an und finden $0 = \sum \lambda_{l,i,j} f^{l+r-1} y_{i,j} = \sum \lambda_{1,r,j} x_{r,j} + \sum \lambda_{0,r-1,j} x_{r-1,j} = 0$, d.h. alle $\lambda_{1,r,j}$ und alle $\lambda_{0,r-1,j}$ sind 0. So fortfahrend sieht man, dass alle $\lambda_{l,i,j}$ verschwinden.

Um einzusehen, dass $\mathcal{B}$ Erzeugendensystem ist, reicht es, induktiv nachzuweisen, dass $Kern f^i$ im Erzeugnis $\langle \mathcal{B} \rangle$ von $\mathcal{B}$ liegt. Für $i = 1$ ist das per Konstruktion richtig. Sei also v aus $Kern f^i$, d.h. $f^{i-1}v$ aus $K \cap f^{i-1}V$. Per Konstruktion gibt es dann einen Vektor w in $\langle \mathcal{B} \rangle$ mit $f^{i-1}v = f^{i-1}w$, d.h. $v - w$ liegt in $Kern f^{i-1}$ also in $\langle \mathcal{B} \rangle$ per Induktionsvoraussetzung.

q.e.d.

Im Beweis haben wir r so gewählt, dass $n_r \neq 0$ gilt. Nun ist die Dimension von V gerade die Kardinalität von B, woraus die Beziehung

$$\sum_{i=0}^{r} (i+1) n_i = n$$

folgt. Also ist $r < n$.

Um sicher zu stellen, dass der Leser die vorangegangene Argumentation so weit verstanden hat, dass er in der Lage ist, in kleinen Dimensionen die JNF samt einer Transformationsmatrix zu berechnen, wollen wir nun für $n = 4$ und den Fall, wo der Endomorphismus durch eine nilpotente Matrix B gegeben ist, alle Möglichkeiten diskutieren. Dabei konstruieren wir jeweils auch die Spalten v_1, v_2, v_3, v_4 einer Transformationsmatrix S, für die $S^{-1}BS$ eine JNF ist. Die

Korrektheit unserer Behauptungen folgt immer aus dem Beweis des Lemmas. Wir geben im folgenden nur die von Null verschiedenen Indizes an.

Fall 1: $B^3 \neq 0$

Dann ist $n_3 = 1$. Es gibt mindestens einen Index i mit $B^3_{.i} = B^3 e_i \neq 0$. Die gesuchte Basis ist $\{v_j = B^{j-1}e_i | 1 \leq j \leq 4\}$.

Fall 2: $B^2 \neq 0 = B^3$

Dann ist $n_2 = 1 = n_0$. Sei i ein Index mit $B^2_{.i} \neq 0$. Setze $v_j = B^{j-1}e_i, 1 \leq j \leq 3$, und wähle für v_4 einen von v_3 linear unabhängigen Vektor in $Kern\, B$.

Fall 3: $B \neq 0 = B^2$

Dann gibt es zwei Lösungen, nämlich $n_1 = 2$ oder $n_1 = 1, n_0 = 2$.

Unterfall 3.1: $n_1 = 2$

Dann ist 2 der Rang von B. Es gibt also zwei linear unabhängige Spalten $B_{.i}, B_{.j}$. Setze $v_1 = e_i, v_2 = Be_i, v_3 = e_j, v_4 = Be_j$.

Unterfall 3.2: $n_1 = 1, n_0 = 2$

Sei dann $B_{.i} \neq 0$. Setze $v_1 = e_i, v_2 = Be_i$ und ergänze v_2 durch geeignete Vektoren v_3, v_4 zu einer Basis von $Kern\, B$.

Fall 3: $B = 0$

Als konkretes Beispiel betrachten wir die nilpotente Matrix $B = A_2 - E_3$. Wir folgen dem Algorithmus aus Lemma 5. Es ist $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ ein Basisvektor von $Kern\, B \cap Bild\, B$, der durch $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ zu einer Basis von $Kern\, B$ ergänzt wird. Weiter ist $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ ein Vektor, der von B auf $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ abgebildet wird. Mit

$$S'_2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

ist also

$$S'^{-1}_2 B S'_2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

6 Die Jordansche Normalform

Der Satz über die Jordansche Normalform lautet:

Satz 4 *Sei k ein Körper und A aus $k^{n \times n}$. Genau dann ist A über k zu einer Jordanschen Normalform J über k ähnlich, wenn das charakteristische Polynom von A in $k[X]$ in Linearfaktoren zerfällt. Zwei Jordansche Normalformen J und J' sind genau dann zueinander ähnlich über k, wenn $v(J, m, \lambda) = v(J', m, \lambda)$ für alle m und λ gilt.*

Die Jordansche Normalform hat also im Vergleich zur RNF drei Nachteile.

1. Sie existiert für manche Matrizen nicht, wenn der Körper nicht algebraisch abgeschlossen ist.

2. Sie ist im Falle ihrer Existenz nur eindeutig bis auf Permutation der Kästchen.

3. Sie ist nur dann exakt berechenbar, wenn die Nullstellen des charakteristischen Polynoms bekannt sind.

Den Beweis des Satzes beginnen wir wieder mit der Eindeutigkeit:

Lemma 5 *Sei J eine Jordansche Normalform. Dann gilt*

$$v(J,l,\lambda) = Rang\,(J-\lambda E_n)^{l-1} + Rang\,(J-\lambda E_n)^{l+1} - 2Rang\,(J-\lambda E_n)^l.$$

Beweis: $J(m,\lambda) - \lambda E_m$ ist eine nilpotente Matrix mit lauter Einsen auf der ersteren unteren Nebendiagonalen. Infolgedessen nimmt der Rang beim Potenzieren jeweils um 1 ab bis zur m-ten Potenz, die die Nullmatrix ist. Für beliebiges i ist daher $Rang\,(J(m,\lambda)-\mu E_m)^i)-Rang\,(J(m,\lambda)-\mu E_m)^{i+1}$ nur dann 1, wenn $\lambda = \mu$ und $i < m$ ist, und sonst 0. Nun hat $(J-\lambda E_n)^i$ lauter derartige Kästchen auf der Hauptdiagonalen, weshalb $(Rang\,(J-\lambda E_n)^i - Rang\,(J-\lambda E_n)^{i+1}$ gleich der Summe aller $v(J,m,\lambda)$ mit $m > i$ ist. Durch Subtraktion folgt die Formel aus dem Lemma. q.e.d.

Offenbar bleibt die rechte Seite der Gleichung im Lemma unverändert beim Übergang zu einer ähnlichen Matrix. Daher stimmen bei zwei ähnlichen JNF's die Vielfachheiten der Kästchen überein. Ferner kann man bei zerfallendem charakteristischem Polynom $\chi_A = \prod_{i=1}^{s}(X-\lambda_i)^{n_i}$ von A sofort die zugehörige JNF J ohne einen Basiswechsel bestimmen , weil $v(J,l,\lambda_i) = Rang\,(A-\lambda_i E_n)^{l-1} + Rang\,(A-\lambda_i E_n)^{l+1} - 2Rang\,(A-\lambda_i E_n)^l$ für alle Eigenwerte λ_i gilt.

Kommen wir nun zur Existenz einer JNF. Klar ist, dass das Zerfallen des charakteristischen Polynoms von A eine notwendige Bedingung für die Existenz einer JNF ist. Sei nun umgekehrt $\chi_A = \prod_{i=1}^{s}(X-\lambda_i)^{n_i}$ mit s verschiedenen Körperelementen λ_i gegeben. Nach dem Satz über die Hauptraumzerlegung ist dann k^n die direkte Summe der Haupträume. Man erhält also durch Wahl einer geeigneten Basis eine zu A ähnliche Matrix mit s Blöcken A_i auf der Hauptdiagonalen. Dabei ist $A_i - \lambda_i E_{n_i}$ jeweils nilpotent. Es gibt daher nach dem letzten Abschnitt invertierbare Matrizen S_i, so dass

$$S_i^{-1}(A_i - \lambda_i E_{n_i})S_i = J_i$$

eine nilpotente JNF ist, also

$$S_i^{-1}A_iS_i = \lambda_i E_{n_i} + J_i$$

eine JNF. Konjugiert man daher A mit der Blockdiagonalmatrix S, deren Diagonalblöcke die S_i's sind, so erhält man die gewünschte JNF zu A.

q.e.d.

Wir wollen den im Beweis vorgestellten Algorithmus auf unser Beispiel anwenden, was nach unseren Vorarbeiten sehr einfach ist. Das Konjugieren mit $S_2 = \begin{bmatrix} 1 & 0 \\ 0 & S_2' \end{bmatrix}$ macht nämlich aus $\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ die gesuchte JNF

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Die konjugierende Matrix ist

$$S = S_1 S_2 = \begin{bmatrix} 2 & 4 & 3 & 3 \\ 1 & 0 & 1 & 0 \\ -2 & 0 & -2 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

7 Ergänzungen und Ausblicke

Als erstes überlegen wir uns, wie man von einer RNF zur JNF gelangt und umgekehrt. Sei zunächst $B(P)$ eine Begleitmatrix zu $P = \prod_{i=1}^{s}(X - \lambda_i)^{n_i}$ mit paarweise verschiedenen λ_i. Dann hat die zugehörige JNF J genau s Kästchen $J(n_i, \lambda_i)$, da anderenfalls das Minimalpolynom von J nicht mehr P ist, sondern ein echter Teiler davon. Die JNF zu einer RNF bestimmt man dann durch mehrfaches Anwenden dieses Verfahrens auf die Hauptdiagonalblöcke.

Ist umgekehrt eine JNF J gegeben mit Eigenwerten $\lambda_1, \ldots, \lambda_s$, so bestimmt man für jeden Eigenwert die maximale auftretende Kästchengröße $m_{1,i}$, entfernt jeweils ein Kästchen dieser Größe und berechnet $P_1 = \prod_{i=1}^{s}(X-\lambda_i)^{m_{1,i}}$. Danach sucht man wieder unter den übriggebliebenen Kästchen für jeden Eigenwert nach einem Kästchen mit maximalem Format $m_{2,i}$, wobei jetzt bereits eine dieser Zahlen Null sein kann. Es ist dann $P_2 = \prod_{i=1}^{s}(X - \lambda_i)^{m_{2,i}}$. Man fährt auf diese Weise solange fort, bis kein Kästchen mehr übrig ist.

Nun treten an vielen Stellen der reinen Mathematik aber auch bei Anwendungen etwa in der Physik ähnliche, aber viel schwierigere Klassifikationsprobleme auf, bei denen mehrere Endomorphismen gleichzeitig auf 'Normalform' zu bringen sind. Die zugehörige mathematische Disziplin heißt Darstellungstheorie. Übersetzt in die elementare Sprache der linearen Algebra handelt es sich um das Problem, ein Tupel $f_1, f_2, \ldots f_n$ von Endomorphismen eines Vektorraumes V, die gewissen Bedingungen genügen, durch besonders einfache Matrizen darzustellen. In Übung 10 etwa wird für endlichdimensionales V der Fall $n = 2$ mit der Bedingung $f_1 f_2 = f_2 f_1$ ein wenig diskutiert. Bereits hier ist eine vollständige Klassifikation unmöglich. Dennoch gibt es eine Menge interessanter und nichttrivialer Ergebnisse in der Darstellungstheorie mit vielen Anwendungen. Einige dieser Ergebnisse wollen wir nun ganz kurz skizzieren.

Man sieht leicht, dass ein Jordankästchen nie zu einer Blockdiagonalmatrix mit zwei Diagonalblöcken ähnlich ist. Die Verallgemeinerung hiervon ist der Begriff der unzerlegbaren Darstellung. Wie bei der JNF gilt nun ganz allgemein, dass die Zerlegung einer Darstellung in unzerlegbare Darstellungen im wesentlichen eindeutig ist. Allerdings ist die Klassifikation der Unzerlegbaren nur selten möglich.

Etwas bescheidener ist die Frage nach den einfachen Darstellungen, d.h. nach den Tupeln von Endomorphismen, für die kein echter Unterraum stabil unter allen Endomorphismen bleibt. Für $n = 1$ und über den komplexen Zahlen ist dann immer V eindimensional und die Klassifikation wird durch die Menge der 1×1-Matrizen gegeben. Der Tatsache, dass jede Matrix trigonalisierbar ist mit bis auf eventuell mögliche Vertauschungen eindeutig festliegenden Diagonaltermen, entspricht nun die Aussage, dass jede Darstellung sich auf obere Blockdreiecksform bringen lässt mit einfachen Darstellungen als Diagonalblöcken, die wieder bis auf Vertauschungen und Isomorphie eindeutig bestimmt sind. Allerdings können diese Blöcke nun beliebig großes Format haben. Für $n = 2$ wird diese Frage in den Übungen 11 und 12 illustriert.

Selbst die Klassifikation aller einfachen Darstellungen ist im allgemeinen unmöglich, aber in vielen wichtigen Fällen bekannt. So sind etwa die 'endlichdimensionalen einfachen Darstellungen der komplexen halbeinfachen Lie-Algebren' seit etwa 100 Jahren alle klassifiziert, was wichtige Anwendungen in der Theorie der Elementarteilchen hat und teilweise diese Theorie erst ermöglicht hat. Vielleicht haben diese Andeutungen das Interesse einiger Studenten geweckt und sie besuchen in späteren Semestern die entsprechenden Vorlesungen.

8 Übungen

Übung 1 *Berechne die rationale und eine Jordansche Normalform zu folgenden komplexen Matrizen. Gib dabei auch jeweils eine Transformationsmatrix an:*

$$\begin{bmatrix} 8 & 4 & 12 \\ 0 & 1 & 0 \\ -3 & -2 & -4 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -2 & 0 & 1 \\ 3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 3 & 1 \end{bmatrix}$$

und

$$\begin{bmatrix} -3 & -1 & 4 & -3 & -1 \\ 1 & 1 & -1 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 4 & 1 & -4 & 5 & 1 \\ -1 & 0 & 1 & -1 & 1 \end{bmatrix}.$$

Übung 2 *a) Zwei 3×3-Matrizen sind genau dann zueinander ähnlich, wenn sie gleiches Minimalpolynom und charakteristisches Polynom haben.*

b) Welche der drei folgenden reellen Matrizen sind einander ähnlich:

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

Übung 3 *Bestimme die JNF's zu den folgenden komplexen* 3×3*-Matrizen:*

$$\begin{bmatrix} 1 & \pi & 17 \\ 0 & 2 & e \\ 0 & 0 & 3 \end{bmatrix}, \begin{bmatrix} 1 & \pi & 17 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & \epsilon & 1 \\ 0 & 0 & \epsilon^2 \end{bmatrix}.$$

Bei der letzten Matrix sei ϵ *eine nicht-negative reelle Zahl.*

Übung 4 *In dieser Aufgabe arbeiten wir in* $\mathbf{C}^{n\times n}$ *versehen mit der üblichen Norm.*

a) Zeige, dass die Links- oder Rechtsmultiplikation mit einer festen Matrix stetige Abbildungen von $\mathbf{C}^{n\times n}$ *nach* $\mathbf{C}^{n\times n}$ *sind.*

b) Zeige, dass es für jede Matrix A *aus* $\mathbf{C}^{n\times n}$ *eine Folge* A_j *mit Grenzwert* A *gibt, wobei alle* A_j *diagonalisierbar sind mit lauter verschiedenen (von* j *abhängigen Eigenwerten). Die diagonalisierbaren Matrizen liegen also dicht in der Menge aller Matrizen.*

(Tip: Man betrachte zunächst den Fall einer oberen Dreiecksmatrix und verwende Teil a))

c) Sei nun $n = 2$*. Zeige, dass die Menge aller diagonalisierbaren Matrizen mit zwei verschiedenen Eigenwerten offen in* $\mathbf{C}^{n\times n}$ *ist. Betrachte dazu die Diskriminante des charakteristischen Polynoms. (Dies verallgemeinert sich auf beliebige* n*, da in der Algebra I Diskriminanten für Polynome beliebigen Grades eingeführt werden, und auf beliebige Körper durch definition der sog. Zariski-Topologie, was in einer Einführung in die algebraische Geometrie geschieht.)*

Definition 3 *Eine Partition einer natürlichen Zahl* n *ist eine monoton fallende Folge natürlicher Zahlen echt größer als 0, deren Summe* n *ergibt. Es bezeichne* $\mathcal{P}(n)$ *die Menge aller solcher Partitionen und* $p(n)$ *deren Kardinalität.*

So sind z.B. $(4,3)$ und $(3,2,2)$ Partitionen von 7. Es ist bisher keine geschlossene Formel für die Anzahl $p(n)$ der Partitionen von n bekannt.

Übung 5 *Seien nun* $L_1, L_2, \ldots L_r$ *über* k *irreduzible Polynome. Ferner seien* $n_1, n_2, \ldots n_r$ *natürliche Zahlen. Wieviele RNF's gibt es mit charakteristischem Polynom*

$$\prod_{i=1}^{r} L_i^{n_i}?$$

Gib numerische Beispiele mit kleinem r *und kleinen* n_i*'s.*

Übung 6 *Gib eine Bijektion an zwischen* $\mathcal{P}(n)$ *und der Menge der Ähnlichkeitsklassen nilpotenter* $n \times n$*-Matrizen.*

Übung 7 *Sei A aus $\mathbf{k}^{n\times n}$ und v aus $\mathbf{k}^n$.*

a) Der von den Potenzen A^iv mit $0 \leq i$ erzeugte Raum hat als Dimension höchstens den Grad des Minimalpolynoms.

b) Genau dann gibt es einen Vektor v derart, dass die A^iv mit $0 \leq i \leq n-1$ eine Basis von $\mathbf{k}^n$ sind, wenn das Minimalpolynom und das charakteristische Polynom übereinstimmen. Das ist auch äquivalent dazu, dass A zu einer Begleitmatrix ähnlich ist.

Übung 8 *Sei A aus $\mathbf{k}^{n\times n}$.*

a) Minimalpolynom und charakteristisches Polynom von A und A^T stimmen überein.

b) A ist ähnlich zu A^T. (Daher findet man in vielen Lehrbüchern für die RNF und die JNF Matrizen, die zu den von uns betrachteten Normalformen transponiert sind .)

c) Für $n \geq 2$ gibt es keine invertierbare Matrix S, so dass $S^{-1}AS = A^T$ für alle Matrizen A gleichzeitig gilt.

d) Gib jedoch eine Permutationsmatrix an, die beim Konjugieren aus oberen Dreiecksmatrizen untere macht und umgekehrt.

Die nächsten Übungen beziehen sich auf das Kapitel 'Ergänzungen und Ausblicke'.

Übung 9 *Eine Matrix M aus $\mathbf{k}^{n\times n}$ heißt einfach bzw. unzerlegbar über k, falls M über k nicht ähnlich ist zu einer Matrix der Form*

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \textit{bzw.} \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

mit A aus $\mathbf{k}^{r\times r}$ für ein r mit $0 < r < n$.

a) Zeige,dass M genau dann einfach ist über k, wenn M zu einer Begleitmatrix $B(P)$ mit einem irreduziblen Polynom P in $k[X]$ ähnlich ist.

b) Zeige,dass M genau dann unzerlegbar ist über k, wenn M zu einer Begleitmatrix $B(P^s)$ mit einem irreduziblen Polynom P in $k[X]$ ähnlich ist.

c) Zeige, dass Einfachheit und Unzerlegbarkeit nicht erhalten bleiben beim Übergang von k zu einem größeren Körper K.

Übung 10 *Seien A, B aus $\mathbf{k}^{n\times n}$ mit $AB = BA$.*

a) Falls A und B über k diagonalisierbar sind, so existiert eine Matrix S, so dass $S^{-1}AS$ und $S^{-1}BS$ beide Diagonalmatrizen sind. Man sagt dann, A und B seien gleichzeitig diagonalisierbar.

b) Formuliere und beweise die analoge Aussage für Trigonalisierbarkeit.

c) Zeige anhand eines Beispiels, dass die analoge Aussage für die Normalformen nicht richtig ist.

Übung 11 *Für zwei Matrizen X, Y aus $\mathbf{k}^{n\times n}$ ist der Kommutator $[X, Y]$ definiert als $XY - YX$. Seien A, B aus $\mathbf{k}^{n\times n}$. Setze $K = \bigcap_{0\leq i,j,\leq n-1} Kern[A^i, B^j]$.*

a) Zeige mit Hilfe des Satzes von Cayley-Hamilton, dass

$$K = \bigcap_{0 \leq i,j, \leq n} Kern[A^i, B^j].$$

b) Zeige, dass K stabil unter A und B ist, und dass beide auf K kommutieren.

c) Sei nun k algebraisch abgeschlossen. Zeige, dass A und B genau dann einen gemeinsamen Eigenvektor haben, wenn $K \neq 0$ ist.

Übung 12 *Sei A aus $\mathbf{k}^{n\times n}$ eine Diagonalmatrix mit n verschiedenen Diagonaltermen und B die Begleitmatrix zu $X^n - 1$. B ist also eine Permutationsmatrix.*

a) Dann sind die Potenzen $A^i B^j$ mit $0 \leq i, j \leq n-1$ eine Basis von $\mathbf{k}^{n\times n}$.

b) Es gibt keinen echten Unterraum U, der unter A und B stabil ist.

Die nächste Übung verallgemeinert die JNF auf beliebige Körper

Übung 13 *Sei k ein Körper.*

a) Sei $P = P_1^{n_1} \dots P_r^{n_r}$ ein Polynom zerlegt in paarweise verschiedene irreduzible Polynome P_i. Zeige, dass $B(P)$ ähnlich ist zu einer Blockdiagonalmatrix mit den $B(P_i^{n_i})$ auf der Diagonalen.

b)) Seien P, Q beliebige nichtkonstante normierte Polynome der Grade p und q. Sei M die $q \times p$-Matrix mit $M_{1p} = 1$ und $M_{ij} = 0$ sonst. Dann ist

$$\begin{bmatrix} B(P) & 0 \\ M & B(Q) \end{bmatrix}$$

ähnlich zu $B(PQ)$.

c) Sei P irreduzibel. Zeige, dass $B(P^r)$ ähnlich ist zum folgenden 'Jordankästchen' $J(r, P)$ vom Format r zum irreduziblen Polynom P

$$J(r,P) = \begin{bmatrix} B(P) & 0 & ... & 0 & 0 \\ M & B(P) & ... & 0 & 0 \\ 0 & M & B(P) & ... & 0 \\ ... & ... & ... & ... & \\ 0 & 0 & ... & M & B(P) \end{bmatrix}.$$

Dabei ist M die Matrix aus Teil b) für $p = q = Grad\, P$. Es stehen r Kopien von $B(P)$ auf der Hauptdiagonalen.

d) Erkläre, wieso die wirklichen Jordankästchen ein Spezialfall der Konstruktion aus Teil c) sind, und verallgemeinere den Satz über die JNF. Was besagt dieser Satz für $\mathbf{R}$?

e) Was ist für beliebige Polynome P, Q die rationale Normalform von

$$\begin{bmatrix} B(P) & 0 \\ 0 & B(Q) \end{bmatrix}?$$

Kannst Du sogar eine Transformationsmatrix angeben?

Die Gruppe der nächsten Aufgaben beschäftigt sich mit dem Rechenaufwand, der zur Bestimmung der RNF nötig ist.

Definition 4 *Unter der Komplexität eines Algorithmus verstehen wir im folgenden einfach die Anzahl der dabei benötigten Multiplikationen.*

Übung 14 *(Grundaufgabe ggT) Seien P, Q Polynome in $k[X]$ der Grade p, q. Gib einen Algorithmus an, um den g.g.T. (P, Q) von P und Q zu bestimmen sowie eine Darstellung*

$$(P, Q) = RP + SQ,$$

und schätze dessen Komplexität ab nach oben durch eine polynomiale Funktion in p und q.

Übung 15 *(Grundaufgabe Kern) Sei A aus $k^{m\times n}$. Gib einen Algorithmus an, um eine Basis des Kerns der zugehörigen Abbildung zu bestimmen, und schätze dessen Komplexität durch ein Polynom in m und n ab.*

Übung 16 *(Grundaufgabe Bild) Sei A aus $k^{m\times n}$. Gib einen Algorithmus an, um eine Basis des Bildes der zugehörigen Abbildung zu bestimmen, und schätze dessen Komplexität durch ein Polynom in m und n ab.*

Übung 17 *(Grundaufgabe Produkt) Schätze die Komplexität, das Produkt von zwei Matrizen zu berechnen, durch ein Polynom in den Formaten der beteiligten Matrizen ab.*

Übung 18 *(Grundaufgabe Durchschnitt) Seien U und V zwei Unterräume in k^n, von denen Basen bekannt sind. Gib einen Algorithmus an, um eine Basis des Durchschnitts zu bestimmen, und gib dessen Komplexität an.*

Übung 19 *(Grundaufgabe Basisergänzung) In k^n seien zwei linear unabhängige Mengen L und M von Vektoren gegeben, derart dass das Erzeugnis von L in dem Erzeugnis V von M liegt. Gib einen Algorithmus an, um L durch Hinzunahme geeigneter Vektoren aus M zu einer Basis von V zu ergänzen, und schätze dessen Komplexität ab.*

Übung 20 *Zeige, dass bei der Berechnung der RNF nach dem angegebenen Algorithmus nur die obigen Grundaufgaben auftreten. Schließe daraus, dass die Komplexität der Berechnung der RNF nur polynomial mit dem Format der Matrix wächst. Insbesondere gilt dies für die Berechnung von Minimalpolynom und charakteristischem Polynom.*

Eine zu Übung 20 analoge Aussage über die JNF scheitert daran, dass wir die Nullstellen des charakteristischen Polynoms nicht kennen und im allgemeinen nur näherungsweise bestimmen können. Dabei kann sich aber die JNF drastisch ändern, wie in den Übungen 3 und 4 gezeigt.

Definition 5 *Für Matrizen M aus $\mathbf{k}^{m\times m}$ und N aus $\mathbf{k}^{n\times n}$ ist der Kommutant $H(M, N)$ definiert als die Menge aller Matrizen X aus $\mathbf{k}^{n\times m}$ mit $XM = NX$. Dies ist ein Unterraum, dessn Dimension wir mit $[M, N]$ abkürzen.*

Später werden wir den Kommutant als eine Menge von Modulhomomorphismen deuten. Die nächste umfangreiche Aufgabe ist teilweise nicht leicht.

Übung 21 *Sei M aus $\mathbf{k}^{m\times m}$ und N aus $\mathbf{k}^{n\times n}$.*

a) Wie verhalten sich sich $H(M,N)$ und $[M,N]$ beim Übergang zu ähnlichen Matrizen? Zeige, dass $[M,N]=[N^T,M^T]$ und schließe daraus $[M,N]=[N,M]$.

b) Für $M=N$ ist $H(M,M)$ eine Unteralgebra von $\mathbf{k}^{m\times m}$, die die Unteralgebra $k[M]$ bestehend aus allen Linearkombinationen von Potenzen von M enthält.

c) Sei M eine Blockdiagonalmatrix

$$\begin{bmatrix} M_1 & 0 & ... & 0 \\ 0 & M_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & M_r \end{bmatrix}$$

und N eine Blockdiagonalmatrix

$$\begin{bmatrix} N_1 & 0 & ... & 0 \\ 0 & N_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & N_s \end{bmatrix}.$$

Man zerlege dann X aus $\mathbf{k}^{n\times m}$ entsprechend in rs Blöcke X_{ij}. Dann liegt X in $H(M,N)$ genau dann, wenn für alle Blöcke gilt $X_{ij}\epsilon H(M_j,N_i)$.

d) Für eine Begleitmatrix $M=B(P)$ liefert die Abbildung $X\mapsto Xe_1$ einen Isomorphismus zwischen $H(B(P),N)$ und $Kern\,P(N)$.

d) Es ist $[B(P),B(Q)]$ gleich dem Grad des größten gemeinsamen Teilers von P und Q.

e) Gib eine Formel für $[R,R']$ an, wenn R und R' rationale Normalformen sind.

f) Stets gilt

$$dim k[M]\leq m\leq [M,M]$$

Gleichheit gilt bei der ersten Abschätzung genau dann, wenn sie bei der zweiten Abschätzung gilt. Dies ist dazu äquivalent, dass M zu einer Begleitmatrix ähnlich ist.

g) Genau dann sind zwei Matrizen M,N aus $\mathbf{k}^{n\times n}$ ähnlich über k, wenn gilt

$$[M,M]=[N,N]=[M,N].$$