Prof. Dr. H. Brenner | Osnabrück WS 2009/2010

Mathematik I

Vorlesung 18

Vielfachheiten und diagonalisierbare Abbildungen

Satz 18.1. *Es sei K ein Körper und es sei V ein endlichdimensionaler K-Vektorraum. Es sei*

$$\varphi : V \longrightarrow V$$

eine lineare Abbildung. Dann ist φ genau dann diagonalisierbar, wenn das charakteristische Polynom χ_φ in Linearfaktoren zerfällt und wenn für jede Nullstelle λ mit der algebraischen Vielfachheit μ_λ die Gleichheit

$$\mu_\lambda = \dim(\operatorname{Eig}_\lambda(\varphi))$$

gilt.

Beweis. Wenn φ diagonalisierbar ist, so kann man sofort annehmen, dass φ bzgl. einer Basis aus Eigenvektoren durch eine Diagonalmatrix beschrieben wird. Die Diagonaleinträge dieser Matrix sind die Eigenwerte, und diese wiederholen sich gemäß ihrer geometrischen Vielfachheit. Das charakteristische Polynom lässt sich auch direkt aus dieser Diagonalmatrix ablesen, jeder Diagonaleintrag λ trägt als Linearfaktor $X - \lambda$ bei. Für die Umkehrung seien $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte und

$$\mu_i = \dim(\operatorname{Eig}_{\lambda_i}(\varphi)) = \mu(\lambda_i)$$

seien die geometrischen=algebraischen Vielfachheiten. Da nach Voraussetzung das charakteristische Polynom in Linearfaktoren zerfällt, muss die Summe dieser Zahlen gleich n sein. Es seien

$$v_{ij},\, j = 1, \ldots, \mu_i\,,$$

Basen der Eigenräume $\operatorname{Eig}_{\lambda_i}(\varphi)$ für $i = 1, \ldots, k$. Dies sind insgesamt n Vektoren. Sei

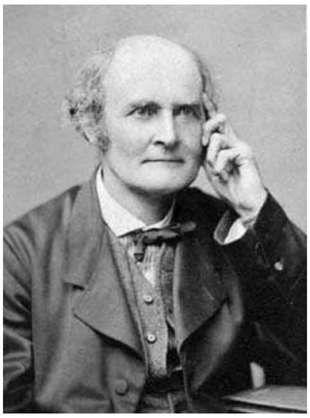
$$\sum_{ij} b_{ij} v_{ij} = 0$$

eine Darstellung der 0. Mit

$$w_i = \sum_{j=1}^{\mu_i} b_{ij} v_{ij} \in \operatorname{Eig}_{\lambda_i}(\varphi)$$

ergibt sich $\sum_{i=1}^k w_i = 0$, wobei die w_i aus den verschiedenen Eigenräumen sind. Nach Lemma 16.11 sind diese linear unabhängig, also müssen alle $w_i = 0$ sein. Damit müssen auch alle $b_{ij} = 0$ sein und die gewählten Basisvektoren der Eigenräume sind linear unabhängig. Daher bilden sie eine Basis. $\square$

Der Satz von Cayley-Hamilton



Arthur Cayley (1821-1895)

William Hamilton (1805-1865)

Einer der Höhepunkte dieses Kurses ist der Satz von Cayley-Hamilton. Um ihn formulieren zu können müssen wir uns zunächst klar machen, dass man in Polynome auch quadratische Matrizen einsetzen kann. 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

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

$$= \begin{pmatrix} 40 & 16 \\ 12 & 36 \end{pmatrix}.$$

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

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

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

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

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

Satz 18.2. *(Der Satz von Cayley-Hamilton) Es sei K ein Körper und sei M eine $n \times n$-Matrix über K. Es sei*

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

das charakteristische Polynom zu M. Dann gilt

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

Das heißt, dass das charakteristische Polynom die Matrix annulliert.

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

$$\mathrm{Adj}\,(XE_n - M)$$

liegt ebenfalls in $\mathrm{Mat}_n(K)$. 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, so dass in den Einträgen der adjungierten Matrix die Variable maximal in der $(n-1)$-ten Potenz vorkommt. Wir schreiben

$$\mathrm{Adj}\,(XE_n - M) = X^{n-1}A_{n-1} + X^{n-2}A_{n-2} + \ldots + XA_1 + A_0$$

mit Matrizen

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

d.h. man schreibt die einzelnen Einträge als Polynom und fasst dann zu X^i die Koeffizienten zu einer Matrix zusammen. Aufgrund von Satz 15.10 gilt

$$\begin{aligned} \chi_M E_n &= (XE_n - M)\circ \mathrm{Adj}\,(XE_n - M) \\ &= (XE_n - M)\circ (X^{n-1}A_{n-1} + X^{n-2}A_{n-2} \\ &\quad + \ldots + XA_1 + A_0) \\ &= X^nA_{n-1} + X^{n-1}(A_{n-2} - M\circ A_{n-1}) \\ &\quad + X^{n-2}(A_{n-3} - M\circ A_{n-2}) \\ &\quad + \ldots + X^1(A_0 - M\circ A_1) - M\circ A_0. \end{aligned}$$

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

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

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

$$\begin{array}{ccc} E_n & = & A_{n-1} \\ c_{n-1}E_n & = & A_{n-2} - M \circ A_{n-1} \\ c_{n-2}E_n & = & A_{n-3} - M \circ A_{n-2} \\ \vdots & \vdots & \vdots \\ c_1 E_n & = & A_0 - M \circ A_1 \\ c_0 E_n & = & -M \circ A_0 \,. \end{array}$$

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

$$\begin{array}{ccc} M^n & = & M^n \circ A_{n-1} \\ c_{n-1}M^{n-1} & = & M^{n-1} \circ A_{n-2} - M^n \circ A_{n-1} \\ c_{n-2}M^{n-2} & = & M^{n-2} \circ A_{n-3} - M^{n-1} \circ A_{n-2} \\ \vdots & \vdots & \vdots \\ c_1 M^1 & = & M A_0 - M^2 \circ A_1 \\ c_0 E_n & = & -M \circ A_0 \,. \end{array}$$

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 Teilsummand $M^{i+1} \circ A_i$ einmal positiv und einmal negativ vorkommt. Also ist $\chi_M(M) = 0$. □

Euklidische Vektorräume

Im Anschauungsraum kann man nicht nur Vektoren addieren und skalieren, sondern ein Vektor hat auch eine Länge, und das Verhältnis von zwei Vektoren zueinander wird durch den Winkel zwischen ihnen ausgedrückt. Länge und Winkel werden beide durch den Begriff des *Skalarprodukts* präzisiert. Dafür muss ein reeller Vektorraum[1] vorliegen.

Definition 18.3. Sei V ein reeller Vektorraum. Ein *Skalarprodukt* auf V ist eine Abbildung

$$V \times V \longrightarrow \mathbb{R},\ (v, w) \longmapsto \langle v, w \rangle \,,$$

mit folgenden Eigenschaften:

(1) Es ist

$$\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle$$

für alle $\lambda_1, \lambda_2 \in \mathbb{R}$, $x_1, x_2, y \in V$ und ebenso in der zweiten Komponente.

[1] Auch für komplexe Vektorräume gibt es Skalarprodukte, was wir aber nicht behandeln werden.

(2) Es ist
$$\langle v, w\rangle = \langle w, v\rangle$$
für alle $v, w \in V$.

(3) Es ist $\langle v, v\rangle \geq 0$ für alle $v \in V$ und $\langle v, v\rangle = 0$ genau dann, wenn $v = 0$ ist.

Die dabei auftretenden Eigenschaften heißen *Bilinearität* (das ist nur eine andere Bezeichnung für multilinear, wenn vorne zwei Vektorräume stehen), *Symmetrie* und *positive Definitheit*.

Beispiel 18.4. Auf dem $\mathbb{R}^n$ ist die Abbildung
$$\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},\, (v, w) = ((v_1, \ldots, v_n), (w_1, \ldots, w_n)) \longmapsto \sum_{i=1}^{n} v_i w_i\,,$$
ein Skalarprodukt, das man das *Standardskalarprodukt* nennt. Eine einfache Rechnung zeigt, dass dies in der Tat ein Skalarprodukt ist.

Definition 18.5. Ein reeller, endlichdimensionaler Vektorraum, der mit einem Skalarprodukt versehen ist, heißt *euklidischer Vektorraum*.

Zu einem euklidischen Vektorraum V ist jeder Untervektorraum $U \subseteq V$ selbst wieder ein euklidischer Vektorraum, da man das Skalarprodukt auf U einschränken kann und dabei die definierenden Eigenschaften erhalten bleiben.

Definition 18.6. Sei V ein Vektorraum über $\mathbb{R}$ mit einem Skalarprodukt $\langle -, -\rangle$. Man nennt zwei Vektoren $v, w \in V$ *orthogonal* zueinander (oder *senkrecht*), wenn
$$\langle v, w\rangle = 0$$
ist.

Definition 18.7. Sei V ein euklidischer Vektorraum und $U \subseteq V$ ein Untervektorraum. Dann heißt
$$U^{\perp} = \{v \in V \mid \langle v, u\rangle = 0 \text{ für alle } u \in U\}$$
das *orthogonale Komplement* von U.

Definition 18.8. Sei V ein euklidischer Vektorraum. Eine Basis $v_1, \ldots, v_n$ von V heißt *Orthonormalbasis*, wenn gilt
$$\langle v_i, v_i\rangle = 1 \text{ für alle } i \text{ und } \langle v_i, v_j\rangle = 0 \text{ für } i \neq j$$

Mit Hilfe des *Schmidtschen Orthonormalisierungsverfahren* kann man leicht zeigen, dass es in jedem euklidischen Vektorraum Orthonormalbasen gibt, siehe Aufgabe 18.21.

Norm und Abstand

Mit einem Skalarprodukt kann man die Länge eines Vektors und damit auch den Abstand zwischen zwei Vektoren erklären.

Definition 18.9. Sei V ein Vektorraum über $\mathbb{R}$ mit einem Skalarprodukt $\langle -, - \rangle$. Dann nennt man zu einem Vektor $v \in V$ die reelle Zahl

$$|| v || = \sqrt{\langle v, v \rangle}$$

die *Norm* von v.

Die Elemente in einer Orthonormalbasis haben alle die Norm 1 und sie stehen senkrecht aufeinander.

Satz 18.10. *(Ungleichung von Cauchy-Schwarz) Sei V ein Vektorraum über $\mathbb{R}$ mit einem Skalarprodukt $\langle -, - \rangle$ und der zugehörigen Norm $|| - ||$. Dann gilt die* Cauchy-Schwarzsche Ungleichung*, nämlich*

$$| \langle v, w \rangle | \leq || v || \cdot || w ||$$

für alle $v, w \in V$.

Beweis. Bei $w = 0$ ist die Aussage richtig. Sei also $w \neq 0$ und damit auch $|| w || \neq 0$. Damit hat man die Abschätzungen

$$\begin{aligned} 0 &\leq \langle v - \frac{\langle v, w \rangle}{|| w ||^2} w, v - \frac{\langle v, w \rangle}{|| w ||^2} w \rangle \\ &= \langle v, v \rangle - \frac{\langle v, w \rangle}{|| w ||^2} \langle w, v \rangle - \frac{\langle v, w \rangle}{|| w ||^2} \langle v, w \rangle + \frac{\langle v, w \rangle \langle v, w \rangle}{|| w ||^4} \langle w, w \rangle \\ &= \langle v, v \rangle - \frac{\langle v, w \rangle^2}{|| w ||^2}. \end{aligned}$$

Multiplikation mit $|| w ||^2$ und Wurzelziehen ergibt das Resultat. □

Bemerkung 18.11. Für zwei von null verschiedene Vektoren v und w in einem euklidischen Vektorraum V folgt aus der Ungleichung von Cauchy-Schwarz, dass

$$-1 \leq \frac{\langle v, w \rangle}{|| v || \cdot || w ||} \leq 1$$

ist. Damit kann man mit Hilfe der trigonometrischen Funktion *Kosinus* bzw. der Umkehrfunktion den Winkel zwischen den beiden Vektoren definieren, nämlich durch

$$\angle(v, w) = \arccos \frac{\langle v, w \rangle}{|| v || \cdot || w ||} .$$

Die trigonometrischen Funktionen werden wir bald einführen.

Lemma 18.12. *Sei V ein Vektorraum über $\mathbb{R}$ mit einem Skalarprodukt $\langle -, - \rangle$. Dann gelten für die zugehörige Norm folgende Eigneschaften.*

(1) $|| v || \geq 0$,

(2) $||\, v\, ||= 0$ *genau dann, wenn* $v = 0$ *ist.*
(3) *Für* $\lambda \in \mathbb{R}$ *und* $v \in V$ *gilt*

$$||\, \lambda v\, ||=|\, \lambda\, | \cdot ||\, v\, || \ .$$

(4) *Für* $v, w \in V$ *gilt*

$$||\, v + w\, ||\leq||\, v\, || + ||\, w\, || \ .$$

Beweis. Die ersten beiden Eigenschaften folgen direkt aus der Definition des Skalarprodukts. Die Multiplikativität folgt aus

$$||\, \lambda v\, ||^2 = \ \langle \lambda v, \lambda v\rangle \ = \ \lambda \langle v, \lambda v\rangle \ = \ \lambda^2 \langle v, v\rangle \ = \ \lambda^2 \ ||\, v\, ||^2 \ .$$

Zum Beweis der Dreiecksungleichung schreiben wir

$$||\, v + w\, ||^2 = \ \langle v{+}w, v{+}w\rangle \ = ||\, v\, ||^2 + ||\, w\, ||^2 + 2\langle v, w\rangle \ \leq ||\, v\, ||^2 + ||\, w\, ||^2 + 2 \mid \langle v, w\rangle \mid$$

Aufgrund von Satz 18.10 ist dies $\leq (||\ v\ || \ + \ ||\ w\ ||)^2$. Diese Abschätzung überträgt sich auf die Quadratwurzeln. □

Lemma 18.13. *Sei* V *ein Vektorraum über* $\mathbb{R}$ *mit einem Skalarprodukt* $\langle -, -\rangle$ $\langle -, -\rangle$ *und der zugehörigen Norm* $||-||$. *Dann gilt die Beziehung*

$$\langle v, w\rangle = \frac{1}{2}(||\, v + w\, ||^2 - ||\, v\, ||^2 - ||\, w\, ||^2)\, .$$

Beweis. Siehe Aufgabe 18.9. □

Definition 18.14. Sei V ein Vektorraum über $\mathbb{R}$ mit einem Skalarprodukt $\langle -, -\rangle$. Zu zwei Vektoren $v, w \in V$ nennt man

$$d(v, w) :=||\, v - w\, ||$$

den *Abstand* zwischen v und w.

Lemma 18.15. *Sei* V *ein Vektorraum über* $\mathbb{R}$ *mit einem Skalarprodukt* $\langle -, -\rangle$. *Dann besitzt der zugehörige Abstand die folgenden Eigenschaften (dabei sind* $u, v, w \in V$*).*

(1) *Es ist* $d(v, w) \geq 0$.
(2) *Es ist* $d(v, w) = 0$ *genau dann, wenn* $v = w$.
(3) *Es ist* $d(v, w) = d(w, v)$.
(4) *Es ist*

$$d(u, w) \leq d(u, v) + d(v, w)\, .$$

Beweis. Siehe Aufgabe 18.11. □

Isometrien

Definition 18.16. Es seien V und W zwei euklidische Vektorräume und sei
$$\varphi : V \longrightarrow W$$
eine lineare Abbildung. Dann heißt φ eine *Isometrie*, wenn für alle $v, w \in V$ gilt:
$$\langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle \, .$$

Lemma 18.17. *Es seien V und W zwei euklidische Vektorräume und sei*
$$\varphi : V \longrightarrow W$$
eine lineare Abbildung. Dann sind folgende Aussagen äquivalent

(1) *φ ist eine Isometrie.*
(2) *Für alle $u, v \in V$ ist $d(\varphi(u), \varphi(v)) = d(u, v)$.*
(3) *Für alle $v \in V$ ist $\| \varphi(v) \| = \| v \|$.*

Beweis. Die Richtungen (1) $\Rightarrow$ (2) und (2) $\Rightarrow$ (3) sind Einschränkungen und (3) $\Rightarrow$ (1) folgt aus Lemma 18.13. □

Satz 18.18. *Sei V ein euklidischer Vektorraum und sei*
$$\varphi : V \longrightarrow V$$
eine lineare Isometrie. Dann besitzt jeder Eigenwert von φ den Betrag 1.

Beweis. Es sei $\varphi(v) = \lambda v$ mit $v \neq 0$, d.h. v ist ein Eigenvektor zum Eigenwert λ. Wegen der Isometrieeigenschaft gilt
$$\| v \| = \| \varphi(v) \| = \| \lambda v \| = | \lambda | \cdot \| v \| \, .$$
Wegen $\| v \| \neq 0$ folgt daraus $| \lambda | = 1$, also $\lambda = \pm 1$. □

Im Allgemeinen muss es keine Eigenwerte geben (bei ungerader Dimension allerdings schon).

Abbildungsverzeichnis