ZUSATZ ZU KAPITEL 5.7
HERMITESCHE SESQUILINEARFORMEN

THOMAS WILLWACHER

Für hermitesche Sesquilinearformen auf einem komplexen Vektorraum bleiben die Resultate des vorherigen Abschnitts (5.7) für reelle symmetrische Bilinearformen mit kleineren Anpassungen gültig. Insbesondere hat man wieder eine Version der Hauptachsentransformation, und des Trägheitssatzes von Sylvester, die wir wie folgt zusammenfassen wollen.

Theorem. *Sei $A \in M(n \times n, \mathbb{C})$ eine hermitesche Matrix, also $\bar{A}^T = A$. Sei s die entsprechende hermitesche Sesquilinearform auf dem $\mathbb{C}^n$, also $s(x,y) = x^T A \bar{y}$. Dann gilt:*

(1) *Es gibt eine Orthonormalbasis $\mathcal{B}$ des $\mathbb{C}^n$ (bzgl. des kanonischen Skalarproduktes), so dass*

$$M_{\mathcal{B}}(s) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} =: D$$

diagonal ist, mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ den Eigenwerten von A.

Äquivalent, in Matrizenform: Es gibt eine unitäre Matrix $U \in U(n)$ so dass

$$U^T A \bar{U} = D.$$

(2) *Es gibt eine Basis $\mathcal{B}'$ des $\mathbb{C}^n$ und Zahlen k, l, so dass*

$$M_{\mathcal{B}'}(s) = \begin{pmatrix} E_k & 0 & 0 \\ 0 & -E_l & 0 \\ 0 & 0 & 0 \end{pmatrix} =: D'.$$

Äquivalent, in Matrizenform: Es gibt eine invertierbare Matrix $S \in GL(n, \mathbb{C})$ so dass

$$S^T A \bar{S} = D'.$$

(3) *Die Zahlen k, l aus (2) sind die Anzahl der positiven bzw. negativen Eigenwerte von A, mit Vielfachheit gezählt, und unabhängig von der Wahl von $\mathcal{B}'$.*

Alternativ, in Matrizenform: Für jede invertierbare Matrix $S \in GL(n, \mathbb{C})$ so dass $S^T A \bar{S}$ diagonal ist, befinden sich genau k positive und genau l negative Einträge auf der Diagonalen von $S^T A \bar{S}$.

Beweis. Die Beweise sind jeweils identisch wie die im reellen Fall.

(1) Wir verwenden für $\mathcal{B} = (v_1, \dots, v_n)$ eine ONB aus Eigenvektoren von $A^T = \bar{A}$, die nach dem Spektralsatz existiert. Dann ist

$$s(v_i, v_j) = (A^T v_i)^T \bar{v}_j = \lambda_i v_i^T \bar{v}_j = \lambda_i \delta_{ij}.$$

Man beachte, dass die Eigenwerte von $\bar{A}$ genau gleich den Eigenwerten von A sind, da alle Eigenwerte reell sind, also $\bar{\lambda}_i = \lambda_i$.

(2) Wir nehmen o.B.d.A. an, die Basis von (1) ist schon so geordnet, dass die Eigenvektoren zu positiven Eigenwerten vor denen zu negativen und diese vor den Eigenvektoren zum Eigenwert 0 stehen. Dann setzten wir, wie im reellen Fall, $\mathcal{B}' = (v_1', \dots, v_n')$ mit

$$v_i' = \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} v_i & \text{falls } \lambda_i \neq 0 \\ v_i & \text{sonst} \end{cases}.$$

Dann gilt immer noch $s(v_i', v_j') = 0$ für $i \neq j$. Ausserdem gilt für i so dass $\lambda_i \neq 0$ jetzt auch

$$s(v_i', v_i') = \left(\frac{1}{\sqrt{|\lambda_i|}} \right)^2 \lambda_i = \frac{\lambda_i}{|\lambda_i|} \in \{\pm 1\}.$$

(3) Durch den Beweis von (2) ist klar, dass für unser $\mathcal{B}'$ die Zahlen k und l gerade die Anzahl der positiven und negativen Eigenwerte sind. Wir müssen noch zeigen, dass, falls für eine andere Basis $\mathcal{C}$ die Matrix $M_{\mathcal{C}}(s)$ diagonal ist mit $\tilde{k}$ positiven und $\tilde{l}$ negativen Einträgen, so ist $k = \tilde{k}$ und $l = \tilde{l}$. Setzt man nun $V_+ = \mathrm{span}(v_1', \dots, v_k')$, $V_- = \mathrm{span}(v_{k+1}', \dots, v_{k+l}')$, $V_0 = \ker A^T$, und definiert analog $\tilde{V}_\pm$ für $\mathcal{C}$, so haben wir zwei Zerlegungen

$$V = V_+ \oplus V_- \oplus V_0 = \tilde{V}_+ \oplus \tilde{V}_- \oplus V_0.$$

Diese haben jeweils die Eigenschaften, dass $s(v,v) > 0$ ist für alle $v \neq 0$ im ersten Untervektorraum, und $s(v,v) < 0$ für alle $v \neq 0$ im zweiten Untervektorraum. (Und natürlich $s(v,v) = 0$ für alle $v \in V_0$.) Es ist $\dim V_+ = k$, $\dim V_- = l$, $\dim \tilde{V}_+ = \tilde{k}$, $\dim \tilde{V}_- = \tilde{l}$. Wie im reellen Fall können wir uns aus Symmetriegründen darauf beschränken zu zeigen, dass $k \geq \tilde{k}$. Denn wäre $k < \tilde{k}$, so gäbe es (mit dem gleichen Argument wie im reellen Fall) ein

$$0 \neq v \in (V_- \oplus V_0) \cap \tilde{V}_+,$$

und für dieses wäre dann

$$s(v,v) \leq 0,$$

ein Widerspruch.

□

Ebenso verallgemeinern die besprochenen Kriterien für die Definitheit von reellen symmetrischen Matrizen bzw. Bilinearformen auf hermitesche Matrizen bzw. Sesquilinearformen. Dies fassen wir wie folgt zusammen.

Theorem. *Sei $A \in M(n \times n, \mathbb{C})$ eine hermitesche Matrix. Dann sind äquivalent:*

(1) *A ist positiv definit, also $x^T A \bar{x} > 0$ für alle $0 \neq x \in \mathbb{C}^n$.*
(2) *Alle Eigenwerte von A sind positiv.*
(3) *Es gibt eine Matrix $S \in GL(n, \mathbb{C})$ so dass*

$$S^T A \bar{S}$$

diagonal ist mit positiven Diagonaleinträgen.
(4) *Es gibt eine Matrix $T \in GL(n, \mathbb{C})$ so dass*

$$A = \bar{T}^T T.$$

(5) *Sei $P_A(t) = (-t)^n + \alpha_{n-1} t^{n-1} + \dots + \alpha_1 t + \alpha_0 \in \mathbb{R}[t]$ das charakteristische Polynom von A. Dann gilt*

$$(-1)^j \alpha_j > 0 \quad \text{für } j = 0, \dots, n-1.$$

(6) *Sei $A_k \in M(k \times k, \mathbb{C})$ die $k \times k$ Untermatrix von A "oben links". Dann gilt*

$$\det A_k > 0 \quad \text{für } k = 1, \dots, n.$$

Der Beweis ist wieder gleich wie im reellen Fall.

Beweis. (1) ⇔ (2) ⇔ (3) folgt direkt aus dem Theorem oben.

(3) ⇔ (4): Zunächst ist $S^T A \bar{S} = D$ äquivalent zu $A = (S^{-1})^T D \bar{S}^{-1}$. Also gilt schonmal (4) ⇒ (3). (Setze $D = E_n$, $S^{-1} = \bar{T}$.) Sei nun D diagonal mit positiven Diagonaleinträgen, und W diagonal mit Diagonaleinträgen den Wurzeln von den Diagonaleinträgen von D, so dass $D = W^2$. Dann ist

$$A = (S^{-1})^T D \overline{S^{-1}} = (W S^{-1})^T \overline{(W S^{-1})} = \bar{T}^T T,$$

mit $T := \overline{(W S^{-1})}$.

(5) ⇔ (2) folgt mit der Vorzeichenregel genau wie im reellen Fall. Man beachte hier, dass das charakteristische Polynom reelle Koeffizienten hat, da alle Eigenwerte von A reell sind und A diagonalisierbar ist. (Also ist $P_A = \pm(t - \lambda_1) \cdots (t - \lambda_n)$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.)

(6) ⇔ (1): Der Beweis vom reellen Fall (Fischer 5.7.7) ist hier mit kleinen Anpassungen anwendbar. Zunächst hat eine positiv definite Matrix positive Determinante, denn aus (4) folgt

$$\det A = \overline{\det T} \det T = |\det T|^2 > 0.$$

Daraus erhält man (1) ⇒ (6). Die andere Richtung erfordert keine Änderungen. □

Wir können auch die (bzw. eine) Diagonalmatrix D und $S \in GL(n, \mathbb{C})$ so dass

$$S^T A \bar{S} = D$$

mittels elementarer Zeilen- und Spaltenumformungen finden. Hier empfiehlt es sich äquivalent $T(= \bar{S})$ zu suchen mit

$$\bar{T}^T A T = D.$$

Dann kann man genau wie im reellen Fall vorgehen, ausser dass bei den Zeilenumformungen von A jeweils die Koeffizienten komplex konjugiert werden müssen. Ein Beispiel soll dies verdeutlichen.

Beispielaufgabe: Finde zu

$$A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

Matrizen D, T wie oben und prüfe, ob A positiv definit ist.

Wir formen A um, indem wir zunächst das $-\frac{i}{2}$-fache der zweiten Spalte zur ersten hinzuaddieren. Wir müssen nun danach noch das $\overline{(-\frac{i}{2})} = \frac{i}{2}$-fache (sic!) der zweiten Zeile zur ersten addieren. Dies ergibt

$$\begin{pmatrix} 1 & i \\ -i & 0 \end{pmatrix}$$

Wir ziehen nun das i-fache der ersten von der zweiten Spalte ab, und danach das $-i = \bar{i}$-fache der ersten Zeile von der zweiten, und erhalten

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Also ist A nicht positiv definit. Durch Ausführen nur der Spaltenumformungen erhalten wir ferner aus der Einheitsmatrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ -\frac{i}{2} & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -i \\ -\frac{i}{2} & \frac{1}{2} \end{pmatrix} = T$$

Man rechnet leicht nach, dass tatsächlich $\bar{T}^T A T = D$ ist.

Negative Definitheit und Semidefinitheit. Man nennt eine Bilinearform bzw. Sesquilinearform s auf dem $\mathbb{K}$-Vektorraum V positiv semidefinit, falls $s(v,v) \geq 0$ ist für alle $v \in V$. Man nennt s negativ definit, falls $s(v,v) < 0$ für alle $0 \neq v \in V$ und negative semidefinit falls $s(v,v) \leq 0$ für alle $v \in V$. Eine symmetrische bzw. hermitesche Matrix A ist positiv oder negativ (semi)definit, falls entsprechendes für die zugehörige Bilinearform $s(x,y) = x^T A \bar{y}$ auf $V = \mathbb{K}^n$ gilt.

Negative Definitheit von A ist äquivalent zur positiven Definitheit von $-A$, mit unseren Kriterien kann man also direkt auch negative Definitheit testen. Ebenso ist negative Semidefinitheit von A äquivalent zu positiver Semidefinitheit von $-A$. Zum Prüfen der Semidefinitheit gibt es sehr ähnlich Kriterien, teilweise mit kleinen Abwandlungen. Wir verweisen dafür auf die Aufgaben der Übungsserie.

Physikerkonvention. Als Schlussbemerkung sei noch angefügt, dass in der "Physikerkonvention" Sesquilinearformen s immer semilinear im ersten statt im zweiten Argument sind. Man muss dann die Formeln oben marginal abändern. Es ist dann

$$s(x,y) = \bar{x}^T A y,$$

mit $A = (s(e_i, e_j))$. Die Transformationsformel wird zu

$$A \to \bar{S}^T A S,$$

und man kann z.B. für $\mathcal{B}$ im ersten Theorem eine Eigenbasis von A nehmen. Ausserdem ist die Physikerkonvention insofern schöner als die Mathematikerkonvention, dass im Vergleich zum reellen Fall in den Aussagen einfach $(-)^T$ ersetzt werden muss durch $(-)^\dagger = \overline{(-)}^T$. Um Verwirrung zu vermeiden werden wir uns dennoch immer an die "Mathematikerkonvention" halten, wie oben verwendet.