§42 Hauptachsentransformation

Zusammenfassung

Zur Formulierung des Satzes über die Hauptachsentransformation benötigen wir den Begriff des selbstadjungierten (oder symmetrischen) Operators auf einem endlichdimensionalen euklidischen Vektorraum. Der Beweis basiert dann wesentlich darauf, dass ein komplexes Polynom immer eine Nullstelle hat und daher eine selbstadjungierte Matrix immer einen reellen Eigenwert.

In diesem Paragrafen ist $(V, < \, , \, >)$ stets ein n–dimensionaler euklidischer Vektorraum, $n > 0$.

A. Dualität und Adjungierte

Lineare Abbildungen zwischen Vektorräumen werden auch *lineare Operatoren* oder kurz *Operatoren* genannt.

Für jeden Vektor $w \in V$ ist

$$j_w : V \to \mathbb{R},\ v \mapsto <w, v>,\ v \in V,$$

d.h. $j_w := <w, \ >$, eine Linearform, also ein Element von $V^\star = \mathrm{Hom}(V, \mathbb{R})$. Das Skalarprodukt liefert daher eine Abbildung

$$j : V \to V^\star,\ w \mapsto j_w,\ w \in V.$$

(42.1) Satz (Riesz): Die Abbildung $j : V \to V^\star$ ist ein Isomorphismus von Vektorräumen.

Beweis: Offensichtlich ist j linear. Denn es gilt

$$j(sv + w) = <sv + w, \ > = s<v, \ > + <w, \ > = sj(v) + j(w))$$

für $v, w \in V$ und $s \in \mathbb{R}$.

Im endlichdimensionalen Fall ist $\dim V = \dim V^\star$ (vgl. 19.6 oder 25.1). Es genügt – nach dem Äquivalenzsatz 15.10 für lineare Abbildungen zwischen gleichdimensionalen Vektorräumen – zu zeigen, dass j injektiv ist. Sei dazu $w \in V$ mit $j(w) = 0$ gegeben. Dann gilt $j_w(w) = 0$, also $<w, w> = 0$. Weil das Skalarprodukt positiv definit ist, folgt $w = 0$, also ist j injektiv.

Der Isomorphismus $j : V \to V^\star$ liefert zu jeder linearen Abbildung $f : V \to V$ eine Abbildung $f^\star : V \to V$ durch

$$<f(v), w> = <v, f^\star(w)>, \text{ für alle } \quad v, w \in V.$$

$f^\star(w)$ ist der Vektor, für den $j_{f^\star(w)}$ die Linearform $v \mapsto <f(v), w>$ ist, also diese Linearform darstellt. Um Existenz und Eindeutigkeit von $f^\star$ genauer zu beschreiben, vergleichen wir $f^\star$ mit der linearen Abbildung

$$\mathrm{Ad}(f) : V^\star \longrightarrow V^\star, \quad \mu \mapsto \mu \circ f,$$

die gelegentlich auch die Adjungierte von f genannt wird.

(42.2) Definition–Satz: Sei $f \in \mathrm{Hom}(V,V)$.

1° $f^\star$ ist wohldefiniert und linear. $f^\star$ heißt die *Adjungierte* zu f (bzw. der zu f *adjungierte Operator*). Es gilt: $j \circ f^\star = \mathrm{Ad}(f) \circ j$.

2° Die Abbildung $f \mapsto f^\star$, $f \in \mathrm{Hom}(V,V)$ ist ein Isomorphismus

$$^\star : \mathrm{Hom}(V,V) \longrightarrow \mathrm{Hom}(V,V)$$

von Vektorräumen. Es gilt $^\star\,^\star = \mathrm{id}$, i.e. $f^{\star\star} = f$ und $(g \circ f)^\star = f^\star \circ g^\star$.

3° f heißt *selbstadjungiert*, wenn $f = f^\star$ gilt.

Beweis: 1°: Für $\tilde{f} := j^{-1} \circ \mathrm{Ad}(f) \circ j : V \to V$,

$$\begin{array}{ccc} V & \xrightarrow{\tilde{f}} & V \\ {\scriptstyle j}\big\downarrow & & \big\downarrow{\scriptstyle j} \\ V^\star & \xrightarrow[\mathrm{Ad}(f)]{} & V^\star \end{array}$$

also $j \circ \tilde{f} = \mathrm{Ad}(f) \circ j$, gilt wegen $j_{\tilde{f}(w)} = \mathrm{Ad}(f)(j_w) = j_w \circ f$ für alle $w \in V$:

$$<\tilde{f}(w), v> = <w, f(v)> \quad \text{für alle} \quad v \in V.$$

Also ist $\tilde{f}$ die gesuchte Abbildung $f^\star$, d.h. $f^\star = \tilde{f}$ ist wohldefiniert und als Komposition von drei linearen Abbildung auch linear.

2° sieht man durch Einsetzen:

$<v, (f+g)^\star(w)> = <(f+g)(v), w> = <f(v), w> + <g(v), w> = <v, f^\star(w) + g^\star(w)>$, also $(f+g)^\star = f^\star + g^\star$.

$<v, (sf)^\star(w)> = <sf(v), w> = s<f(v), w> = <v, s(f^\star(w))>$, also $(sf)^\star = s(f^\star)$ für $s \in \mathbb{R}$.

$<v, f^{\star\star}(w)> = <f^\star(v), w> = <w, f^\star(v)> = <f(w, v>$, also $f^{\star\star} = f$.

$<v, (g \circ f)^\star(w)> = <(g \circ f)(v), w> = <f(v), g^\star(w)> = <v, f^\star \circ g^\star(w)>$, also $(g \circ f)^\star = f^\star \circ g^\star$.

B. Selbstadjungierte Operatoren

(42.3) Lemma: Für ein lineare Abbildung $f \in \mathrm{Hom}(V,V)$ sei $A = A_f$ die darstellende Matrix bezüglich einer ONB. Dann ist f genau dann selbstadjungiert, wenn A symmetrisch ist: $A = A^\top$.

Beweis: Sei $B = (b_1, b_2, \ldots, b_n)$ die ONB (nach Voraussetzung ist V ja endlichdimensional). Die darstellende Matrix hat die Koeffizienten A_i^ν, die durch $f(b_i) = A_i^\nu b_\nu$ gegeben sind. Es gilt $<f(b_i), b_j> = <A_i^\nu b_\nu, b_j> = A_i^\nu <b_\nu, b_j> = A^j{}_i$ und analog $<b_i, f(b_j)> = <b_i, A_j^\mu b_\mu> = A_j^i$. Also ist $<f(b_i), b_j> = <b_i, f(b_j)>$ gleichbedeutend mit $A_i^j = A_j^i$.

(42.4) Satz: Sei $f \in \mathrm{Hom}(V,V)$ selbstadjungiert. Dann sind die Eigenvektoren von f zu verschiedenen Eigenwerten orthogonal zueinander.

Beweis: Sei $f(v) = \lambda v, f(w) = \mu w$ für Vektoren $v, w \in V \backslash \{0\}$. Aus $<f(v), w> = <v, f(w)>$ folgt $<\lambda v, w> = <v, \mu w>$, also $(\lambda - \mu)<v, w>$. Daraus folgt die Behauptung.

(42.5) Satz: Sei $f \in \mathrm{Hom}(V, V)$ selbstadjungiert. Für jeden Eigenvektor v von f ist $v^\perp = \{v\}^\perp$ f–invariant, d.h. $f(v^\perp) \subset v^\perp$.
Beweis: Für alle $w \in v^\perp$ ist $<v, w> = 0$, also gilt $<v, f(w)> = <f(v), w> = \lambda<v, w> = 0$, daher $f(w) \in v^\perp$.

C. Hauptachsentransformation

(42.6) Satz: Ein selbstadjungierter Operator hat immer einen Eigenwert.
Beweis: Für einen Operator $f \in \mathrm{Hom}(V, V)$ hat das charakteristische Polynom χ_f nach dem Fundamentalsatz der Algebra (vgl. 38.8) immer eine komplexe Nullstelle $\lambda = \xi + i\eta \in \mathbb{C}$ mit $\xi, \eta \in \mathbb{R}$. Es genügt offenbar zu zeigen, dass λ reell ist, d.h. dass $\eta = 0$ gilt. Sei A die darstellende Matrix bezglich einer Basis von V, und $z \in \mathbb{C}^n$ ein Spaltenvektor $z \neq 0$ mit $Az = \lambda z$. Dann hat z eine eindeutige Zerlegung $z = x + iy$ in reelle Spaltenvektoren $x, y \in \mathbb{R}^n$. Es gilt

$$Ax + iAy = Az = \lambda z = (\xi + i\eta)(x + iy) = (\xi x - \eta y) + i(\xi y + \eta x)$$

und daher

$$Ax = \xi x - \eta y \quad \text{und} \quad Ay = \xi y + \eta x.$$

Wenn jetzt f selbstadjungiert ist, so gilt insbesondere $<Ax, y> = <x, Ay>$, also $<\xi x - \eta y, y> = <x, \xi y + \eta x>$. Daher ist $\xi<x, y> - \eta<y, y> = \xi<x, y> + \eta<x, x>$ und $\eta(<x, x> + <y, y>) = 0$. Es folgt $\eta = 0$, weil $z \neq 0$ gilt.
Es ist dann $\lambda = \xi$ und $Ax = \lambda x$ und $Ay = \lambda y$, also ist x oder y Eigenvektor zu λ.

(42.7) Satz (Hauptachsentransformation): Zu jedem selbstadjungierten Operator $f \in \mathrm{Hom}(V, V)$ gibt es eine Orthonormalbasis $B = (b_1, b_2, \ldots, b_n)$, die aus lauter Eigenvektoren besteht. f hat also bezüglich dieser Basis die Matrixdarstellung

$$\boxed{A_f = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)}$$

mit den Eigenwerten $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, und die Basisvektoren $(b_1, b_2, \ldots, b_n)$ sind die „Hauptachsen“.
Insbesondere hat f die folgende Darstellung

$$\boxed{f(v) = \sum_{i=1}^{n} \lambda_i<b_i, v>b_i}$$

(vgl. 42.8.2°).
Beweis: Induktion nach $n = \dim V$:
Im Falle $n = 1$ gibt es immer einen Eigenwert $\lambda \in \mathbb{R}$ und einen Eigenvektor $v \in V$. Mit $b := v/\|v\|$ ist (b) die gesuchte ONB.

Im Falle $n > 1$ gibt es nach 42.6 einen Eigenwert $\lambda_1 \in \mathbb{R}$ und dazu einen Eigenvektor $v \in V$. Setze $b_1 := v/\|v\|$ und $U = v^\perp = (b_1)^\perp$. Es gilt dann $U \oplus U^\perp = V$ und $U^\perp = \mathbb{R}b_1$. Also ist $U^\perp$ eindimensional und es folgt $\dim U = n-1$ nach der Dimensionsformel. Nach 42.5 ist $f(U) \subset U$, deshalb kann man die Restriktion $g := f|_U$ von f auf U als eine lineare Abbildung von U nach U definieren. Natürlich ist g selbstadjungiert, weil f selbstadjungiert ist. Wegen $\dim U = n-1$ hat U eine ONB $(b_2, \ldots, b_n)$ aus Eigenvektoren von g, die natürlich auch Eigenvektoren von f sind. Wegen $b_1 \in U^\perp$ ist jetzt $(b_1, b_2, \ldots, b_n)$ eine ONB aus Eigenvektoren von f, und der Satz ist bewiesen.

(42.8) Bemerkung:

1° Der Name „Hauptachsentransformation" kommt von folgender Anwendung in der Geometrie: Es sei eine Quadrik Q in $\mathbb{R}^n$ mit Standard-Skalarprodukt und Standard-Orthonormalbasis durch eine quadratische Form $q(x) = x^\top Ax$ zu einer symmetrischen Matrix A gegeben; als $Q := q^{-1}(0)$. Die Aussage des Satzes bedeutet, dass es eine Orthonormalbasis, also einen Koordinatenwechsel innerhalb der kartesischen Koordinatensysteme, so gibt, dass diese Quadrik Q in den neuen Koordinaten $\bar{x}$ die (*Normal-*) Form

$$\boxed{Q = \{\bar{x} : \sum_{i=1}^{i=n} \lambda_i (\bar{x}^i)^2 = 0\}}$$

hat.

Für allgemeine Quadriken hat man gegebenenfalls zum einen noch eine Nullpunktverschiebung und zum anderen noch lineare Terme zu beachten.

2° Unter den Voraussetzungen des Satzes 42.7 seien $\lambda_1, \lambda_2, \ldots, \lambda_k$, $k \leq n$, die paarweise verschiedenen Eigenwerte von f und es seien $E_j := \mathrm{Ker}(f - \lambda_j \mathrm{id})$ die zugehörigen Eigenräume. Sei $P_j : V \to V$ die nach 41.13 existierende und eindeutig bestimmte orthogonale Projektion auf E_j, also $P_j \circ P_j = P_j$, P_j ist selbstadjungiert und es gilt $E_j = \mathrm{Im} P_j$. Dann hat f die Darstellung

$$\boxed{f = \sum_{j=1}^{j=k} \lambda_j P_j.}$$

Man nennt diese Darstellung auch die „*Spektralzerlegung*" von f, denn das *Spektrum* von f ist in dieser (endlichdimensionalen) Situation die Menge $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$. Wir kommen auf Spektralzerlegungen in den Paragrafen 44 (der endlichdimensionale komplexe Fall) und 46 (der unendlichdimensionale Fall) zurück.

(42.9) Korollar: Zu einer symmetrischen $n \times n$–Matrix $A \in \mathbb{R}^{n \times n}$ gibt es stets eine orthonormale Matrix $R \in \mathrm{O}(n, \mathbb{R})$, so dass RAR^{-1} eine Diagonalmatrix ist.

Das ist genau die Aussage des Satzes 42.7 angewandt auf die durch die Matrix A gegebene selbstadjungierte lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$.