Teil 6. Bilinearformen.

6A. Kongruenz quadratischer Matrizen.

Sei K ein Körper, sei $A \in M(n \times n, K)$ eine quadratische Matrix. Wie wir zu Beginn von Teil 3 gesehen haben, liefert A durch

$$\langle x, y\rangle_A = {}^t x A y$$

eine Bilinearform auf dem n-dimensionalen K-Vektorraum K^n. Ist $P \in \mathrm{GL}(n, K)$, und bezeichnen wir die Spalten der Matrix P mit $v_1, \ldots, v_n$, so gilt

$$\langle v_i, v_j\rangle_A = \langle e_i, e_j\rangle_{{}^t PAP}$$

(wegen $Pe_i = v_i$ handelt es sich hier um die Gleichung: ${}^t v_i A v_j = {}^t(Pe_i)A(Pe_j) = {}^t e_i {}^t PAPe_j$). Die Bilinearform $\langle -, -\rangle$ wird also bezüglich der Basis $v_1, \ldots, v_n$ durch die Matrix ${}^t PAP$ beschrieben. Wir sehen auf diese Weise, wie sich die Matrizendarstellung einer Bilinearform bei einem beliebigen Basiswechsel ändert.

Die quadratischen Matrizen A, B heißen *kongruent*, wenn es eine invertierbare Matrix P mit $B = {}^t PAP$ gibt. *Eine Bilinearform wird also bezüglich verschiedener Basen eines Vektorraums durch kongruente Matrizen beschrieben.*

Beachte: Im Fall reeller Matrizen arbeitet man häufig mit Paaren von Matrizen, die gleichzeitig ähnlich und kongruent sind: denn man wählt oft zu einer gegebenen Matrix A eine orthogonale Matriz P, und bildet $B = {}^t PAP$; wegen $P^{-1} = {}^t P$ sind die Matrizen A, B sowohl kongruent als auch ähnlich. Die beiden Beziehungen der Kongruenz und der Ähnlichkeit sind aber völlig verschiedenartig: Ähnliche Matrizen sind im allgemeinen nicht kongruent (Beispiel: Die Matrizen $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ und $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ sind ähnlich, aber nicht kongruent – jede zu einer symmetrischen Matrix kongruente Matrix ist wieder symmetrisch), und kongruente Matrizen sind üblicherweise nicht ähnlich. Beispiel: Sei K ein Körper mit $\mathrm{char}(K) \neq 2, 3$, die Matrizen $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ und $\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$ sind kongruent (da $\mathrm{char}(K) \neq 2$), aber nicht ähnlich (da $\mathrm{char}(K) \neq 3$).

6B. Symmetrische Bilinearformen.

In diesem Abschnitt wird vorausgesetzt, daß K ein Körper der Charakteristik $\mathrm{char}(K) \neq 2$ *ist.* Sei $\langle -, -\rangle$ eine symmetrische Bilinearform auf dem K-Vektorraum V. Zur Erinnerung: ein Vektor $v \in V$ wird *isotrop* genannt, falls $\langle v, v\rangle = 0$ gilt, andernfalls *anisotrop*. Ein Unterraum U, der nur aus isotropen Vektoren besteht, heißt *total-isotrop*. Ist $U \subset V$, so setzt man

$$\begin{aligned} U^\perp &= \{v \in V \mid \langle u, v\rangle = 0 \quad \text{für alle} \quad u \in U\} \\ &= \{v \in V \mid \langle v, u\rangle = 0 \quad \text{für alle} \quad u \in U\}. \end{aligned}$$

(Hier ist einiges zu wiederholen, was wir im Fall $K = \mathbb{R}$ schon wissen... . Dies wird im Abschnitt 6C noch einmal diskutiert.)

Lemma 1. *Die Bilinearform $\langle -,- \rangle$ sei nicht die Nullform. Dann gibt es anisotrope Vektoren.*

Beweis: Da $\langle -,- \rangle$ nicht die Nullform ist, gibt es Vektoren v, w mit $\langle v,w \rangle \neq 0$. Ist $\langle v,v \rangle = 0, \langle w,w \rangle = 0$, so ist $\langle v+w, v+w \rangle = 2\langle v,w \rangle \neq 0$ (hier wird die Voraussetzung $\mathrm{char}(K) \neq 2$ gebraucht!). Wir sehen also: *Ist $\langle v,w \rangle \neq 0$, so ist mindestens einer der drei Vektoren $v, w, v+w$ anisotrop.*

Lemma 2. *Ist w ein anisotroper Vektor in V, so ist $V = Kw \oplus \{w\}^\perp$*

Beweis: Es ist $Kw \cap \{w\}^\perp = 0$, denn ist $\lambda \in K$ mit $\langle w, \lambda w \rangle = 0$, so ist $\lambda = 0$. Um zu sehen, daß $V = Kw + \{w\}^\perp$ gilt, betrachte einen beliebigen Vektor $v \in V$. Man rechnet sofort nach, daß $v' = v - \frac{\langle v,w \rangle}{\langle w,w \rangle} w$ zu $\{w\}^\perp$ gehört, und natürlich ist $v = \frac{\langle v,w \rangle}{\langle w,w \rangle} w + v'$, also in $Kw + \{w\}^\perp$.

Eine Basis $v_1, \dots, v_n$ von V heißt *Orthogonalbasis* (bezüglich der symmetrischen Bilinearform $\langle -,- \rangle$), falls $\langle v_i, v_j \rangle = 0$ für alle $i \neq j$ gilt. *Die darstellende Matrix einer symmetrischen Bilinearform bezüglich einer Orthogonalbasis ist eine Diagonalmatrix.*

Satz. *Sei* $\mathrm{char}(K) \neq 2$. *Ein endlich-dimensionaler K-Vektorraum mit einer symmetrischen Bilinearform $\langle -,- \rangle$ besitzt eine Orthogonalbasis $v_1, \dots, v_n$.*

Zusatz. *Ist $K = \mathbb{R}$, so kann man zusätzlich erreichen $\langle v_i, v_i \rangle \in \{1, -1, 0\}$. Ist $K = \mathbb{C}$, so kann man sogar $\langle v_i, v_i \rangle \in \{1, 0\}$ erreichen.*

Beweis: Induktion nach der Dimension von V. Ist $\langle -,- \rangle$ die Nullform, so ist nichts zu zeigen. Sei also $\langle -,- \rangle$ nicht die Nullform. Lemma 1 liefert einen anisotropen Vektor v_1, Lemma 2 besagt, daß $V = Kv_1 \oplus (Kv_1)^\perp$ gilt. Die Einschränkung von $\langle -,- \rangle$ auf $(Kv_1)^\perp$ ist natürlich wieder eine symmetrische Bilinearform. Nach Induktion besitzt $(Kv_1)^\perp$ eine Orthogonalbasis $v_2, \dots, v_n$. Offensichtlich ist $v_1, \dots, v_n$ eine Orthogonalbasis von V.

Sei nun $K = \mathbb{R}$ und $c_i = \langle v_i, v_i \rangle$. Ist $c_i \neq 0$, so setze $v_i' = \frac{1}{\sqrt{|c_i|}} v_i$. Dann ist $\langle v_i', v_i' \rangle \in \{1, -1\}$. Ist $c_i = 0$, so sei $v_i' = v_i$. Die Vektoren $v_1', \dots, v_n'$ bilden dann eine Orthogonalbasis mit $\langle v_i', v_i' \rangle \in \{1, -1, 0\}$.

Ist $K = \mathbb{C}$, und $c_i = \langle v_i, v_i \rangle \neq 0$, so wähle d_i mit $d_i^2 = c_i$ und setze $v_i' = \frac{1}{d_i} v_i$.

Noch ein Zusatz: Sei $\mathrm{char}(K) \neq 2$. Sei V ein n-dimensionaler K-Vektorraum mit einer symmetrischen Bilinearform $\langle -,- \rangle$ und einer Orthogonalbasis $v_1, \dots, v_n$. *Die Anzahl r der Vektoren v_i mit $\langle v_i, v_i \rangle = 0$ hängt nur von der Bilinearform, nicht von der Auswahl der Orthogonalbasis ab.* Zwei Beschreibungen von r: Es ist r die Dimension des Unterraum $V^\perp$ (man nennt $V^\perp$ das *Radikal* der Bilinearform), und es ist $n - r$ der Rang einer darstellenden Matrix. Man nennt eine symmetrische Bilinearform *nicht-ausgeartet,* wenn $r = 0$ gilt.

Scherungsalgorithmus zur Diagonalisierung einer symmetrischen Bilinearform. Der vorangehende Beweis liefert einen effektiven Algorithmus, dabei sind zwei mögliche Schritte zu unterscheiden: Wahl eines anisotropen Vektors und

Orthogonalisieren. Das Orthogonalisieren erfolgt durch 'Scherungen', wie folgt: Wir gehen davon aus, daß wir eine Matrix der folgenden Form gegeben haben:

$$A = \begin{bmatrix} a_1 & & 0 & b_1 \\ & \ddots & & \vdots \\ 0 & & a_m & b_m \\ b_1 & \dots & b_m & c \end{bmatrix}$$

mit $a_i \neq 0$ für alle $1 \leq i \leq m$ (dies sei die darstellende Matrix für unsere Bilinearform auf einem $(m+1)$-dimensionalen Unterraum). Sei $\beta_i = -\frac{b_i}{a_i}$, für $1 \leq i \leq m$. Betrachte die Matrix

$$S = \begin{bmatrix} 1 & & 0 & \beta_1 \\ & \ddots & & \vdots \\ & & 1 & \beta_m \\ 0 & & & 1 \end{bmatrix}$$

(eine derartige Transformation wird geometrisch als 'Scherung' bezeichnet). Es ist

$$^tSAS = \begin{bmatrix} a_1 & & & \\ & \ddots & & \\ & & a_m & \\ & & & a_{m+1} \end{bmatrix}$$

mit einem a_{m+1}, das sich aus der Rechnung ergibt und natürlich auch Null sein kann. Entsprechend erhält man für eine Matrix A der folgenden Form, wobei D eine invertierbare Diagonalmatrix ist, eine invertierbare Matrix S wie angegeben:

$$A = \begin{bmatrix} D & B \\ ^tB & C \end{bmatrix}, \quad S = \begin{bmatrix} I_n & B' \\ 0 & I_{n-m} \end{bmatrix}, \quad \text{mit} \quad ^tSAS = \begin{bmatrix} D & 0 \\ 0 & C' \end{bmatrix}.$$

Man geht also induktiv vor: Ist $C' \neq 0$, so muss man gegebenfalls einen Basiswechsel wie in Lemma 1 durchführen: Vertauschung zweier Basisvektoren, oder Addition eines Vektors w zu v (auch letzteres ist wieder eine Scherung.) Insgesamt sieht man also:

Satz (Neue Formulierung). *Sei K ein Körper mit* char $K \neq 2$. *Zu jeder symmetrischen Matrix $A \in M(n \times n, K)$ gibt es eine invertierbare Matrix S, sodass ^tSAS eine Diagonalmatrix ist. Dabei kann S als Produkt von Permutationsmatrizen und von Matrizen der Form $I_n + tE_{ij}$ mit $i < j$ und $t \in K$ gewählt werden.*

6C. Der Dualraum eines Vektorraums mit einer nicht-ausgearteten symmetrischen Bilinearform.

Sei K ein Körper und V ein K-Vektorraum mit einer nicht-ausgearteten symmetrischen Bilinearform $\langle -,- \rangle$. Dann erhält man für jedes $v \in V$ durch $\langle v,- \rangle$ eine Linearform auf V (die Abbildung $\langle v,- \rangle$ ordnet dem Vektor $w \in V$ den Wert $\langle v,w \rangle \in K$ zu; daß dies eine K-lineare Abbildung ist, folgt aus den Axiomen einer Bilinearform). *Die Zuordnung $v \mapsto \langle v,- \rangle$ ist eine injektive Abbildung $\eta \colon V \to V^*$.* (Beweis: Da η eine lineare Abbildung ist, ist nur zu zeigen, daß $\eta(v) = 0$ nur für $v = 0$ gilt. Dies folgt aber gerade aus der Voraussetzung, daß die Bilinearform $\langle -,- \rangle$ nicht-ausgeartet ist.)

Insbesondere sehen wir: *Ist V endlich-dimensional, so ist die Zuordnung $v \mapsto \langle v,- \rangle$ ein Isomorphismus $V \to V^*$*; ein endlich-dimensionaler euklid'scher Vektorraum V wird durch sein inneres Produkt mit seinem Dualraum identifiziert! Hier sei noch einmal explizit formuliert, was die Surjektivität der Abbildung η bedeutet: *Zu jeder linearen Abbildung $\alpha \colon V \to K$ gibt es ein $v \in V$ mit $\alpha(-) = \langle v,- \rangle$*; dabei ist v eindeutig bestimmt (weil η injektiv ist). (**Warnung:** Der Isomorphismus $V \to V^*$ hängt von der Bilinearform ab!)

Wie im Spezialfall eines euklid'schen Vektorraums setzen wir für $M \subseteq V$

$$M^\perp = \{ v \in V \mid \langle v,u \rangle \quad \text{für alle} \quad u \in M \},$$

und nennen $M^\perp$ das *orthogonale Komplement* von M in V.

Sei V endlich-dimensional. Unter dem Isomorphismus $\eta \colon V \to V^*$ wird für jede Teilmenge $M \subseteq V$ das orthogonale Komplement $M^\perp$ auf M° abgebildet, entsprechend ist für jede Teilmenge $\Phi \subseteq V^*$ der Annullator Φ° nichts anderes als $(\eta^{-1}(\Phi))^\perp$; es gilt also

$$M^\perp = \eta^{-1}(M^\circ), \quad \text{und} \quad \Phi^\circ = (\eta^{-1}(\Phi))^\perp.$$

Beweis: Sei $v \in V$. Genau dann ist $v \in M^\perp$, wenn $\langle v,u \rangle = 0$ für alle $u \in M$ gilt, also genau dann wenn $\eta(v)(u) = 0$ für alle $u \in M$ gilt, also genau dann, wenn $\eta(v) \in M^\circ$, und dies bedeutet gerade, daß $v \in \eta^{-1}(M^\circ)$ gilt.

Zweite Gleichheit: Sei wieder $v \in V$. Sei $\phi \in V^*$. Es ist $\phi(v) = \langle \eta^{-1}(\phi), v \rangle$. Es ist $v \in \Phi^\circ$ genau dann, wenn $\phi(v) = 0$ für alle $\phi \in \Phi$ gilt, wenn also $\langle \eta^{-1}(\phi), v \rangle = 0$ für alle $\phi \in \Phi$ gilt, also wenn v zu $(\eta^{-1}(\Phi))^\perp$ gehört.

Mit Hilfe dieser Entsprechung können wir alle in Teil 2 formulierten Aussagen über Annullatoren übertragen. (Natürlich können die entsprechenden Aussagen auch direkt bewiesen werden. Da ziemlich alle Beweise im Teil 2 ganz einfach waren, wäre es ebenso einfach, hier die direkten Beweise zu notieren. Der Verweis auf Teil 2 dient also **nicht** der Beweis-Vereinfachung, sondern soll die Parallelität der Aussagen betonen. Auf diese Weise soll sichtbar werden, daß die Bildung $(-)^\perp$ des orthogonalen Komplements der Annullator-Bildung der Dualitätstheorie entspricht.)

Sei also V ein endlich-dimensionaler Vektorraum mit einer nicht-ausgearteten symmetrischen Bilinearform $\langle -,- \rangle$. Seien $M, M_1, M_2 \subseteq V$ Untermengen, seien U, U_1, U_2 Unterräume von V.

(0) *Es ist* $M^\perp = (\text{span}\{M\})^\perp$

(1) $M^\perp$ *ist ein Unterraum von* V.

(2) *Ist* $M_1 \subseteq M_2$, *so ist* $M_1^\perp \supseteq M_2^\perp$.

(3) *Es ist* $M \subseteq M^{\perp\perp}$.

(x) *Es ist* $\dim U + \dim U^\perp = \dim V$.

(x′) *Es ist* $U^{\perp\perp} = U$.

(4) *Es ist* $(M_1 \cup M_2)^\perp = M_1^\perp \cap M_2^\perp$.

(4′) *Es ist* $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.

(5) *Es ist* $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.

Warnung. Im Fall eines euklid'schen Vektorraums V gilt die folgende Verschärfung von (x): Es ist $U \oplus U^\perp = V$. Im allgemeinen gilt dies nicht! Ist etwa U total-isotrop, so ist $U \subseteq U^\perp$. Beispiel: Sei $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ und $U = Ke_1$. Die Matrix A ist symmetrisch, und invertierbar, also ist die Bilinearform $\langle -,- \rangle_A$ symmetrisch und nicht ausgeartet. Hier ist $U = U^\perp$. (Ist $K = \mathbb{R}$, so handelt es sich gerade um eine symmetrische Bilinearform mit Signatur $(1,1,0)$.)

Abbildungen. Wieder sei V ein endlich-dimensionaler K-Vektorraum mit einer nicht-ausgearteten symmetrischen Bilinearform $\langle -,- \rangle$. Sei $f\colon V \to V$ eine Abbildung. Für jedes $v \in V$ ist $\langle v, f(-)\rangle\colon V \to K$ eine lineare Abbildung, also $\langle f^{\text{ad}}(v), - \rangle$ für ein (eindeutig bestimmtes) Element $f^{\text{ad}}(v) \in V$. Man erhält auf diese Weise eine Abbildung $f^{\text{ad}}\colon V \to V$ mit

$$\langle f^{\text{ad}}(v), w \rangle = \langle v, f(w) \rangle \quad \text{für alle} \quad v, w \in V,$$

man nennt f^{ad} die zu f *adjungierte* Abbildung, dies ist eine lineare Abbildung $f^{\text{ad}}\colon V \to V$. Man nennt f *selbstadjungiert*, falls $f = f^{\text{ad}}$ gilt. (Erinnerung: im Fall eines euklid'schen Raums haben wir schon selbstadjungierte Abbildungen untersucht).

Lemma. *Sei* V *ein endlich-dimensionaler* K*-Vektorraum mit einer nicht-ausgearteten symmetrischen Bilinearform* $\langle -,- \rangle$. *Sei* $f\colon V \to V$ *eine lineare Abbildung. Dann ist*

$$\operatorname{Im} f^{\text{ad}} = (\operatorname{Ker} f)^\perp \quad \text{und} \quad \operatorname{Ker} f^{\text{ad}} = (\operatorname{Im} f)^\perp.$$

Beweis: Wir zeigen $(\operatorname{Im} f^{\text{ad}})^\perp = \operatorname{Ker} f$. Wenden wir $(-)^\perp$ darauf an, so erhalten wir wegen (g) die erste Behauptung. Wegen $(f^{\text{ad}})^{\text{ad}} = f$ folgt daraus auch die zweite Behauptung. Um $\operatorname{Ker} f \subseteq (\operatorname{Im} f^{\text{ad}})^\perp$ zu zeigen, betrachten wir $w \in \operatorname{Ker} f$ und $v \in V$ und berechnen $\langle f^{\text{ad}}(v), w\rangle = \langle v, f(w)\rangle = 0$. Umgekehrt sieht man, daß auch die Inklusion $(\operatorname{Im} f^{\text{ad}})^\perp \subseteq \operatorname{Ker} f$ gilt: Sei nämlich $w \in V$ mit $\langle f^{\text{ad}}(v), w\rangle = 0$ für alle $v \in V$. Dann ist $\langle v, f(w)\rangle = 0$ für alle $v \in V$. Da die Bilinearform $\langle -,- \rangle$ nicht-ausgeartet ist, folgt $f(w) = 0$, also $w \in \operatorname{Ker}(f)$.

Das Transponieren von Matrizen. Ist A eine $(n \times n)$-Matrix mit Koeffizienten in K, so können wir die Abbildung $f_A \colon K^n \to K^n$ betrachten. Auf K^n ist durch $\langle x, y\rangle = {}^t xy$ für $x, y \in K^n$ eine nicht-entartete symmetrische Bilinearform gegeben. Offensichtlich ist ${f_A}^{\mathrm{ad}} = f_{{}^tA}$. Wir erhalten also:

$$\operatorname{Im} f_{{}^tA} = (\operatorname{Ker} f_A)^{\perp} \quad \text{und} \quad \operatorname{Ker} f_{{}^tA} = (\operatorname{Im} f_A)^{\perp}.$$

Bemerkung. Es ist natürlich $\operatorname{Im} f_{{}^tA}$ der von den Spalten von ${}^tA$ erzeugte Unterraum von K^n, und $\operatorname{Ker} f_A = \text{Lös}(A, 0)$ ist die Menge der Lösungen des Gleichungssystems $AX = 0$. Die Orthogonalitätsbeziehung zwischen $\operatorname{Im} f_{{}^tA}$ und $\operatorname{Ker} f_A$ bezüglich der Bilinearform $\langle -, -\rangle$ ist nur Umformulierung von $AX = 0$, denn bei der Matrizenmultiplikation werden Zeilenvektoren mit Spaltenvektoren multipliziert. Entsprechend ist $\operatorname{Ker} f_{{}^tA}$ die Menge der Lösungen des Gleichungssystems $({}^tA)X = 0$, während $\operatorname{Im} f_A$ der von den Spalten von A erzeugte Unterraum von K^n ist. Die hier notierten Gleichungen $\operatorname{Im} f_{{}^tA} = (\operatorname{Ker} f_A)^{\perp}$ und $\operatorname{Ker} f_{{}^tA} = (\operatorname{Im} f_A)^{\perp}$ sind also nichts Überraschendes, diese neuen Formulierungen betonen allerdings die Gleichwertigkeit einerseits des Lösungsraums eines homogenen linearen Gleichungssystems und andererseits des Spaltenraums einer Matrix.

Der Spezialfall eines euklid'schen Vektorraums. Sei nun V ein endlichdimensionaler euklid'scher Raum. Hier gilt, wie wir wissen, die folgende Verschärfung von (f): *Ist U ein Unterraum von V, so ist $V = U \oplus U^{\perp}$.*

Ist demnach $f \colon V \to V$ ein selbstadjungierter Endomorphismus von V, so ist $V = \operatorname{Im} f \oplus \operatorname{Ker} f$ und diese beiden Unterräume $\operatorname{Im} f$ und $\operatorname{Ker} f$ sind zueinander orthogonal. Umformulierung für Matrizen: Ist A eine reelle symmetrische Matrix, so ist $\mathbb{R}^n = \operatorname{Im} f_A \oplus \text{Lös}(A, 0)$ und diese Unterräume $\operatorname{Im} f_A$ und $\text{Lös}(A, 0)$ sind zueinander orthogonal.

6D. Positiv definite (reelle) Matrizen.

Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Wir betrachten die Bilinearform $\langle x, y\rangle_A = {}^t xAy$. Es sei daran erinnert, daß jede reelle symmetrische Matrix diagonalisierbar ist, also genügend reelle Eigenwerte hat! *Genau dann ist $\langle -, -\rangle_A$ positiv definit, wenn alle Eigenwerte von A positiv sind.* (Derartige Matrizen nennt man daher auch *positiv definite* Matrizen.)

Ist eine beliebige symmetrische Bilinearform auf $\mathbb{R}^n$ gegeben, so können wir sie mit der kanonischen Bilinearform $\langle -, -\rangle$ vergleichen. Sei also A eine symmetrische reelle $(n \times n)$-Matrix und setze $\langle x, y\rangle_A = {}^t xAy$. Ist $\langle -, -\rangle_A$ positiv definit, so bezeichnen wir die zugehörige Norm mit $||x||_A = \sqrt{\langle x, x\rangle_A}$.

Satz. *Ist die symmetrische Bilinearform $\langle x, y\rangle_A = {}^t xAy$ auf $\mathbb{R}^n$ positiv definit, so gibt es eine reelle Zahlen $\gamma, \gamma' > 0$ mit*

$$\gamma ||x|| \le ||x||_A \le \gamma' ||x||.$$

Beweis: Es gibt eine orthogonale Matrix P, so daß $D = {}^tPAP$ eine Diagonalmatrix ist, mit Diagonalkoeffizienten d_i. Alle diese Koeffizienten d_i sind positiv. Sei

d das Minimum der d_i und $\gamma = \sqrt{d}$. Es ist also $d_i \geq \gamma^2 > 0$. Sei $y = P^{-1}x$, also $x = Py$. Es ist

$$\langle x, x\rangle_A = {}^t x A x = {}^t y {}^t P A P y = {}^t y D y = \sum d_i y_i \geq \gamma^2 \sum y_i^2 = \gamma^2 ||y||^2.$$

Da P orthogonal ist, ist $||y|| = ||x||$. Dies zeigt $||x||_A \geq \gamma ||x||$. Entsprechend sieht man $||x||_A \leq \gamma' ||x||$ für $\gamma' > 0$ mit $(\gamma')^2$ das Maximum der d_i.

Ist A eine $(n\times m)$-Matrix, so nennt man die Determinanten der Matrizen $A_t = (a_{ij})_{1\leq i,j \leq t}$ die *Hauptminoren* der Matrix A (man betrachtet also die quadratischen $(t\times t)$-Matrizen, die aus A durch Streichen der Zeilen und Spalten mit Index echt größer als t entstehen).

Minoren-Kriterium für Positivität. *Sei A eine reelle $(n\times n)$-Matrix. Genau dann ist die Bilinearform $\langle x, y\rangle_A$ positiv definit, wenn die Hauptminoren positiv sind.*

Beweis: Ist $\langle -, -\rangle_A$ positiv definit, so gibt es eine invertierbare Matrix P, so daß ${}^t PAP$ eine Diagonalmatrix ist; seien $d_1, \dots, d_n$ die Diagonalkoeffizienten, dann ist $d_i > 0$ für alle i. Es ist $(\det P)^2 \det A = \det {}^t PAP = d_1 \cdots d_n > 0$, also auch $\det A > 0$. Bei der Bildung des Hauptminors $\det A_t$ betrachtet man die Einschränkung von $\langle -, -\rangle_A$ auf den Unterraum $\operatorname{span}\{e_1, \dots, e_t\}$. Da auch diese Einschränkung positiv definit ist, folgt $\det A_t > 0$.

Sei umgekehrt $\det A_t > 0$ für $1 \leq t \leq n$. Wir können annehmen (Induktion), daß A_{n-1} positiv definit ist. Also gibt es eine Matrix $P \in \mathrm{GL}((n-1), \mathbb{R})$, so daß $D = {}^t P A_{n-1} P$ eine Diagonalmatrix ist (mit positiven Diagonalkoeffizienten). Für unsere Matrix $A = \begin{bmatrix} A_{n-1} & a \\ {}^t a & a' \end{bmatrix}$ mit $a \in M((n-1)\times 1, \mathbb{R})$ und $a' \in \mathbb{R}$ liefert dies:

$$B = \begin{bmatrix} {}^t P & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_{n-1} & a \\ {}^t a & a' \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} D & b \\ {}^t b & b' \end{bmatrix},$$

dabei ist $b \in M((n-1)\times 1, \mathbb{R})$ und $b' \in \mathbb{R}$. Die rechte Matrix kann durch eine Scherung diagonalisiert werden, wir erhalten eine Diagonalmatrix $\begin{bmatrix} D & \\ & d \end{bmatrix}$. Es ist $\det A > 0$ positiv, also auch $\det B > 0$, demnach auch $\det D \cdot d = \det \begin{bmatrix} D & \\ & d \end{bmatrix} > 0$. Wegen $\det D > 0$ sehen wir, daß d positiv ist. Da alle Diagonalkoeffizienten von $\begin{bmatrix} D & \\ & d \end{bmatrix}$ positiv sind, ist $\langle -, -\rangle$ positiv definit.

Nachtrag über Polynome: Berechnung der Koeffizienten bei Vorgabe der Nullstellen. Sei R ein kommutativer Ring, seien $\gamma_1, \ldots, \gamma_n \in R$. Sei

$$f(T) = \prod_{i=1}^{n}(T - \gamma_i) = \sum_{i=0}^{n} a_i T^{n-i},$$

dann lassen sich die Koeffizienten $a_i \in R$ von f wie folgt berechnen:

$$a_s = (-1)^s \sum_{1 \le i_1 < i_2 < \cdots < i_s \le n} \gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_s},$$

also

$$\begin{aligned} a_0 &= +1 \\ a_1 &= -\sum\nolimits_i \gamma_i \\ a_2 &= +\sum\nolimits_{i<j} \gamma_i\gamma_j \\ a_3 &= -\sum\nolimits_{i<j<k} \gamma_i\gamma_j\gamma_k \\ &\cdots \\ a_n &= (-1)^n \gamma_1\gamma_2\cdots\gamma_n. \end{aligned}$$

Man nennt die Funktion

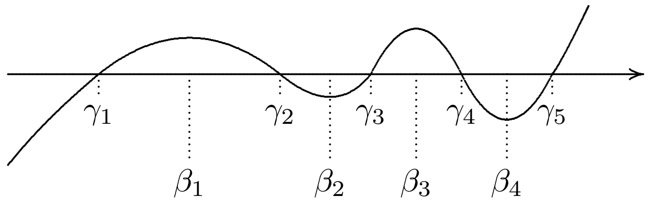
$$\sigma_s(T_1, \ldots, T_n) = \sum_{1 \le i_1 < i_2 < \cdots < i_s \le n} T_{i_1}T_{i_2}\cdots T_{i_s}$$

(die Anzahl der Summanden ist $\binom{n}{s}$) die s*-te elementarsymmetrische Funktion* in den Variablen $T_1, \ldots, T_n$, dabei ist $0 \le s \le n$.

Descartes'sche Vorzeichenregel: *Sei* $f(T) = \sum_{i=0}^{n} a_i T^{n-i}$ *ein normiertes reelles Polynom mit nur reellen Nullstellen. Dann sind die beiden folgenden Bedingungen äquivalent:*
(i) *Alle Nullstellen sind positiv.*
(ii) *Es ist* $a_i > 0$ *für* i *gerade, und* $a_i < 0$ *für* i *ungerade.*

Beweis. Sind alle Nullstellen positiv, so ist auch $\sigma_s(\gamma_1, \ldots, \gamma_n) > 0$, für $1 \le i \le n$. Wegen $a_s = (-1)^s \sigma_s(\gamma_1, \ldots, \gamma_n)$ ist die Bedingung (ii) erfüllt.

Sei nun umgekehrt (ii) erfüllt. Wir betrachten zuerst den Fall, daß die Nullstellen γ_i paarweise verschieden sind. Wir führen Induktion nach n. Ist $n = 1$, so ist $a_1 = -\gamma_1$, also gilt in diesem Fall (i). Sei nun n beliebig. Wir betrachten die Ableitung $f'(T)$ von $f(T)$. Seien $\gamma_1 < \gamma_2 < \cdots < \gamma_n$ die Nullstellen von $f(T)$. Für $1 \le i < n$ liegt zwischen γ_i und γ_{i+1} mindestens ein lokales Extremum, also eine Nullstelle β_i von $f'(T)$. Insgesamt finden wir auf diese Weise mindestens $n - 1$ paarweise verschiedene Nullstellen $\beta_1, \dots, \beta_{n-1}$ von $f'(T)$. Da $f'(T)$ den Grad $n-1$ hat, gibt es keine weiteren Nullstellen: $f'(T)$ hat also nur reelle Nullstellen.



Die Koeffizienten von $f'(T)$ sind $(n-i)a_i$, es ist also die zu (ii) analoge Regel für $f'(T)$ erfüllt. Nach Induktion wissen wir, daß die Nullstellen β_i von $f'(T)$ positiv sind, für $1 \leq i < n$, also ist auch $\gamma_{i+1} > 0$ für $1 \leq i < n$. Schließlich wissen wir, daß $(-1)^n a_n = \gamma_1 \dots \gamma_n$ positiv ist. Da schon gezeigt ist, daß die γ_i mit $i > 1$ positiv sind, sehen wir, daß auch $\gamma_1 > 0$ gilt.

Es bleibt zu zeigen, daß die Implikation (ii) $\Longrightarrow$ (i) auch dann richtig ist, wenn einige Nullstellen zusammenfallen. Sei $\epsilon = \frac{1}{2} \min |a_i|$. Die Funktionen σ_s sind stetig, also gibt es ein $\delta > 0$, so daß sich bei einem Abändern der Zahlen γ_i um weniger als δ die Koeffizienten a_i um weniger als ϵ ändern, daß sich insbesondere also das Signum nicht ändert. Wir wählen nun $0 \leq \delta_i < \delta$, so daß die Zahlen $\gamma_i' = \gamma_i - \delta_i$ paarweise verschieden sind und betrachten das Polynom $g(T) = \prod_i (T - \gamma_i')$. Da die Vorzeichenregel (ii) für das Polynom $g(T)$ gilt, schließen wir, daß die Zahlen $\gamma_i - \delta_i$ positiv sind, also ist $\gamma_i > 0$, für alle i.

Bemerkung. Allgemeiner kann man zeigen, daß für jedes reelle Polynom mit nur reellen Nullstellen die Anzahl der Vorzeichenwechsel in der Koeffizientenfolge $(a_0, a_1, a_2, \dots, a_n)$ gleich der Anzahl der positiven Nullstellen ist.