Kapitel 7

Bilinearformen und quadratische Formen

7.1 Bilinearformen

7.1.1 Linearformen

Definition 7.1 Es sei V ein linearer Raum über dem Körper $\mathbb{K}$. Als *lineares Funktional* oder *Linearform* bezeichnet man die Elemente von $\mathrm{L}(V,\mathbb{K})$, das heißt die linearen Abbildungen von V, deren Werte skalar sind. Der Raum $\mathrm{L}(V,\mathbb{K})$ heißt auch *Dualraum* von V und wird mit V^* bezeichnet.

Beispiel 7.1 (a) $V=\mathbb{R}^n$. Es sei $B=\{b_1,\dots,b_n\}$ eine Basis von V, dann ist für jedes $i=1,\dots,n$, die Abbildung

$$V\ni x_1b_1+x_2b_2+\cdots+x_nb_n\mapsto x_i\in\mathbb{K}$$

ein lineares Funktional. Es heißt $i$tes *Koordinatenfunktional*, und man bezeichnet es mit b_i^*. Folglich gilt für alle $i,j=1,\dots,n$, dass $b_i^*(b_j)=\delta_{ij}$.
Es gilt $(\mathbb{R}^n)^*\cong\mathbb{R}^n$, denn jedes lineare Funktional F auf $\mathbb{R}^n$ ist durch einen Zeilenvektor $(a_1,\dots,a_n)\in\mathbb{R}^n$ gegeben über

$$\mathbb{R}^n\ni x=(x_1,\dots,x_n)\mapsto\sum_{i=1}^n a_ix_i.$$

In diesem Falle gilt $F=\sum_{i=1}^n a_i\mathrm{e}_i^*$. Umgekehrt gibt es für jedes lineare Funktional $F\in(\mathbb{R}^n)^*$ einen solchen Vektor $a=(a_1,\dots,a_n)$ so dass F die obige Form hat, $F=\sum_{i=1}^n F(\mathrm{e}_i)\mathrm{e}_i^*$.
(b) Es sei $V=\mathrm{C}[0,1]$, $g\in V$ fest. Setzt man für $f\in V$

$$T_g(f)=\int_0^1 f(x)g(x)\,\mathrm{d}x,$$

so ist dadurch eine Linearform auf V definiert.
(c) Sei $V=\mathrm{C}(\mathbb{R})$ und $a\in\mathbb{R}$ fest. Dann definiert

$$F_a(f)=f(a),\quad f\in V$$

ein lineares Funktional auf V.
(d) Es sei $(V,\langle\cdot\,,\cdot\rangle)$ ein unitärer Raum und $a\in V$ fixiert. Dann definiert $F(v)=\langle v\,,a\rangle$ eine Linearform auf V.

7.1.2 Bilinearformen – Einfachste Eigenschaften

Definition 7.2 Es sei V ein Vektorraum. Eine Abbildung $B\colon V\times V\to\mathbb{K}$ heißt *Bilinearform* auf V, falls für alle $u,v,w\in V$ und $\lambda,\mu\in\mathbb{K}$ gilt:

$$\begin{aligned}B(\lambda v+\mu w,u)&=\lambda B(v,u)+\mu B(w,u), &&\text{Linearität im ersten Argument}\\ B(u,\lambda v+\mu w)&=\lambda B(u,v)+\mu B(u,w) &&\text{Linearität im zweiten Argument.}\end{aligned}$$

Eine Bilinearform B heißt *symmetrisch*, wenn für alle $u,v\in V$ gilt $B(u,v)=B(v,u)$; sie heißt *schiefsymmetrisch*, wenn für alle $u,v\in V$ gilt $B(u,v)=-B(v,u)$.

Beispiel 7.2 (a) Ein euklidisches Skalarprodukt ist also eine symmetrische Bilinearform, die zusätzlich positiv definit ist.
(b) Es sei $A\in\mathbb{R}^{n\times n}$. Dann definiert $B(x,y)=\langle x\,,\,Ay\rangle=x^\top\cdot A\cdot y$, $x,y\in\mathbb{R}^n$, eine Bilinearform auf $\mathbb{R}^n$. Wir werden sehen, dass es für jede Bilinearform auf dem $\mathbb{R}^n$ eine Matrix $A\in\mathbb{R}^{n\times n}$ gibt, so dass $B(x,y)=\langle x\,,\,Ay\rangle$. Ist $A^\top=A$, so ist B symmetrisch, ist $A^\top=-A$, so ist B schiefsymmetrisch. Die Determinantenfunktion auf $\mathbb{R}^2$ ist ein Beispiel für eine schiefsymmetrische Bilinearform.
(c) Sind $f,g\in V^*$, so definiert $B(x,y)=f(x)g(y)$, $x,y\in V$ eine Bilinearform.
(d) Jede Bilinearform B ist die Summe aus einer symmetrischen Bilinearform S und einer schiefsymmetrischen Bilinearform T. Setzt man nämlich für alle $x,y\in V$,

$$S(x,y)=\frac{1}{2}\big(B(x,y)+B(y,x)\big),\qquad T(x,y)=\frac{1}{2}\big(B(x,y)-B(y,x)\big),$$

so ist $B=S+T$, wobei S symmetrisch und T schiefsymmetrisch ist. Der Raum der symmetrischen Bilinearformen (bzw. symmetrischen $n\times n$-Matrizen) hat die Dimension $\frac{1}{2}n(n+1)$; der Raum der schiefsymmetrischen Bilinearformen über V (der schiefsymmetrischen $n\times n$-Matrizen) hat die Dimension $\frac{1}{2}n(n-1)$.

7.1.3 Die Matrix einer Bilinearform

In Bemerkung 5.1 haben wir zu einem Skalarprodukt die *Gramsche* Matrix $G=(\langle e_i\,,\,e_j\rangle)_{\substack{i=1\dots n\\ j=1\dots n}}$ definiert, wobei $E=\{e_1,\dots,e_n\}$ eine Basis von V sei. Analog definiert man die *Matrix einer Bilinearform B*, $A=(B(e_i,e_j))_{\substack{i=1\dots n\\ j=1\dots n}}$; Bezeichnung: $A=M_E(B)$. Für $v=\sum_{i=1}^n x_ie_i$ und $w=\sum_{j=1}^n y_je_j$ ist dann

$$B(v,w)=B\left(\sum_{i=1}^n x_ie_i,\sum_{j=1}^n y_je_j\right)=\sum_{i,j=1}^n x_iy_j\,B(e_i,e_j)=\sum_{i,j}x_ia_{ij}y_j=\langle x\,,\,Ay\rangle=x^\top\cdot A\cdot y.$$

Wegen der Symmetrie von A gilt auch $B(v,w)=\langle x\,,\,Ay\rangle=\langle Ax\,,\,y\rangle=y^\top\cdot A\cdot x$.
Es sei V ein reeller Vektorraum mit den Basen $E=\{e_1,\dots,e_n\}$ und $F=\{f_1,\dots,f_n\}$. Die Basistransformation sei gegeben durch eine Matrix $S\in\mathbb{R}^{n\times n}$, $S=(s_{ij})$, in üblicher Weise:

$$f_j=\sum_{i=1}^n s_{ij}e_i,\quad j=1,\dots,n; \tag{7.1}$$

siehe auch Abschnitt 3.3.5. Die Bilinearform B werde bezüglich E beschrieben durch die Matrix $A=(a_{ij})$ und bezüglich F durch die Matrix $A'=(a'_{ij})$. Dann gilt:

$$\begin{aligned} a'_{kl} &= B(f_k, f_l) = B\left(\sum_{i=1}^{n} s_{ik}\, e_i, \sum_{j=1}^{n} s_{jl}\, e_j\right) \\ &= \sum_{i,j=1}^{n} s_{ik} s_{jl} B(e_i, e_j) = \sum_{i,j=1}^{n} s_{ik} s_{jl} a_{ij} = \sum_{i,j=1}^{n} (S^\top)_{ki} a_{ij} s_{jl} = (S^\top A S)_{kl}; \end{aligned}$$

also gilt die folgende *Transformationsformel für Bilinearformen*

$$\boxed{A' = S^\top A S.}$$

Man beachte, dass sich diese Formel von der Transformationsformel für Endomorphismen unterscheidet. Die Matrizen A und A' haben i. a. unterschiedliche charakteristische Polynome und damit unterschiedliche Eigenwerte.

Bemerkung 7.1 Eine Bilinearform B heißt *nicht-entartet* (oder nicht-ausgeartet), falls gilt

$$B(x, y) = 0 \quad \forall\, y \in V \Longrightarrow x = 0.$$

Andernfalls heißt B entartet. Für jede Basis E gilt: B ist genau dann nicht-entartet, wenn die zugehörige Matrix $A = M_E(B)$ regulär (invertierbar) ist.

7.2 Quadratische Formen und Bilinearformen

7.2.1 Definition und Beispiel

Definition 7.3 Es sei $B\colon V \times V \to \mathbb{R}$ eine Bilinearform auf V. Dann heißt $Q_B(x) = B(x, x)$, $x \in V$, die zugehörige *quadratische Form*. Es sei $A = M_E(B) \in \mathbb{R}^{n\times n}$ die Matrix zu B bezüglich einer Basis E, dann ist $Q_B(x) = \langle x\,,\, Ax\rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ die zugehörige quadratische Form. Sei umgekehrt $Q\colon V \to \mathbb{R}$ eine quadratische Form, dann definiert

$$C(x, y) = \frac{1}{4}\big(Q(x + y) - Q(x - y)\big)$$

eine *symmetrische* Bilinearform C.

Wir schreiben auch Q_A für Q_B, wenn $A = M_E(B)$ die Matrix von B ist. Man beachte, dass für $\lambda \in \mathbb{R}$ gilt $Q(\lambda x) = \lambda^2 Q(x)$, $x \in V$. Außerdem ist $Q(v + w) = Q(v) + B(v, w) + B(w, v) + Q(w)$, $v, w \in V$.

Beispiel 7.3 Auf $\mathbb{R}^2$ mit der Standardbasis ist durch $B(x, y) = x_1 y_1 + x_1 y_2 + 3x_2 y_1$ eine Bilinearform gegeben. Die zugehörige Matrix von B ist dann $A = \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix}$. Die zugehöriger quadratischer Form lautet

$$Q(x_1, x_2) = x_1^2 + 4x_1 x_2.$$

Die zugehörige symmetrische Bilinearform ist dann

$$\begin{aligned} C(x,y) &= \frac{1}{4}\left((x_1+y_1)^2+4(x_1+y_1)(x_2+y_2)-(x_1-y_1)^2-4(x_1-y_1)(x_2-y_2)\right) = x_1y_1+2x_1y_2+2x_2y_1 \\ &= \frac{1}{2}\left(B(x,y)+B(y,x)\right). \end{aligned}$$

Die C entsprechende Matrix ist $A' = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$. Wenn man mit einer quadratischen Form startet, kann man sich also immer auf symmetrische Bilinearformen beschränken.

7.2.2 Definitheit

Für lokale Extrema von Funktionen mehrerer Veränderlicher ist von entscheidender Bedeutung, welche Werte eine gewisse quadratische Formen annehmen kann. Das Verhalten der *Hesse*-Matrix von f an der Stelle x_0, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right)_{i,j}$ ist hier entscheidend.

Definition 7.4 Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische quadratische Matrix und $Q(x) = \sum_{i,j=1}^n a_{ij}x_ix_j$, $x=(x_1,\dots,x_n)\in\mathbb{R}^n$, die zugehörige quadratische Form. Q heißt

(a) *positiv definit*, falls für alle $x\in V$ mit $x\neq 0$ gilt $Q(x)>0$.
(b) *negativ definit*, falls für alle $x\in V$ mit $x\neq 0$ gilt, $Q(x)<0$.
(c) *indefinit*, falls es Vektoren $x_1, x_2\in V$ gibt mit $Q(x_1)>0>Q(x_2)$.
(d) *positiv semidefinit*, falls für alle $x\in V$, $Q(x)\geq 0$.
(e) *negativ semidefinit*, falls für alle $x\in V$, $Q(x)\leq 0$.

Man sagt, dass die symmetrische Matrix A bzw. die symmetrische Bilinearform B positiv definit, negativ definit usw. ist, falls Q_A positiv definit, negativ definit usw. ist.
Im komplexen Fall betrachtet man in der Regel *Sesquilinearformen* B, die im ersten Argument linear und im zweiten Argument antilinear sind. B heißt hermitesch, falls zusätzlich gilt $B(x,y)=\overline{B(y,x)}$. Diese Formen sind durch hermitesche Matrizen $A^*=A$ gegeben über $B(x,y)=\sum_{i,j=1}^n a_{ij}x_i\overline{y_j}$. Man assoziiert dann zu B die Form $Q(x)=\sum_{i,j=1}^n a_{ij}x_i\overline{x}_j$.

Lemma 7.1 *Es sei A ein quadratische symmetrische, reelle Matrix. Die quadratische Form Q_A ist positiv definit genau dann, wenn alle Eigenwerte positiv sind, negativ definit, genau dann, wenn alle Eigenwerte von A negativ sind und indefinit, falls es einen positiven und einen negativen Eigenwert gibt. Q_A ist positiv semidefinit, falls alle Eigenwerte von A nichtnegativ sind und negativ semidefinit, falls alle Eigenwerte von A kleiner gleich Null sind.*

Beweis. Nach dem Satz über die Hauptachsentransformation besitzt $\mathbb{R}^n$ eine Orthonormalbasis aus Eigenvektoren $\{e_1,\dots,e_n\}$ zu den reellen Eigenwerten $\lambda_1,\dots,\lambda_n$. Für $x=\sum_{i=1}^n x_ie_i$, $x\neq 0$, folgt dann

$$Q_A(x) = \langle x\,,Ax\rangle = \left\langle \sum_{i=1}^n x_ie_i\,,\sum_{j=1}^n \lambda_jx_je_j \right\rangle = \sum_{i,j=1}^n x_ix_j\lambda_i\delta_{ij} = \sum_{i=1}^n \lambda_i x_i^2.$$

Sind nun alle Eigenwerte von A positiv, so ist offenbar auch $Q_A(x) > 0$ für alle $x \neq 0$. Ist umgekehrt $Q_A(x) > 0$ für alle von Null verschiedenen Vektoren, so ist insbesondere $Q_A(e_i) = \lambda_i > 0$. Also sind alle Eigenwerte von A positiv.
Die anderen Aussagen (A negativ definit, indefinit usw.) folgen analog. ■

7.2.3 Die Normalform einer symmetrischen Bilinearform

Satz 7.2 *Es sei V ein reeller Vektorraum mit* $\dim V = n$ *und $B\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf V.*
Dann gibt es eine Basis $E = \{e_1, \dots, e_n\}$ von V, so dass die Matrixdarstellung $A = M_E(B)$ folgende Gestalt hat

$$A = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & \\ 0 & 0 & 0 \end{pmatrix},$$

wobei die angegebenen Blöcke auch die Zeilenzahl 0 *haben dürfen. Mit anderen Worten, es gibt Zahlen $p, q \geq 0$, $p + q \leq n$, so dass bezüglich E die Bilinearform B die Gestalt*

$$B((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1 y_1 + x_2 y_2 + \dots + x_p y_p - x_{p+1} y_{p+1} - \dots - x_{p+q} y_{p+q}$$

hat. Die Zahl $p + q$ heißt Rang *von B und $p - q$ heißt* Signatur *von B.*

Beweis. Wir statten V mit einem Skalarprodukt aus, wählen eine beliebige Basis E_1 und wenden die Hauptachsentransformation auf die symmetrische Matrix $A' = M_{E_1}(B)$ an. Nach dem Satz über die Hauptachsentransformation existiert eine Orthonormalbasis E_2 von V aus Eigenvektoren von A' zu den ausschließlich reellen Eigenwerten von A'. Wir ordnen nun die Eigenvektoren $\{f_1, \dots, f_n\}$ so in E_2 an, dass zunächst alle positiven Eigenwerte $\{\lambda_1, \dots, \lambda_p\}$, dann alle negativen Eigenwerte $\{\lambda_{p+1}, \dots, \lambda_{p+q}\}$ und dann alle Eigenwerte 0 kommen. In der Basis E_2 hat also B die Gestalt

$$B(x, y) = \lambda_1 x_1 y_1 + \lambda_2 x_2 y_2 + \dots + \lambda_p x_p y_p - \left|\lambda_{p+1}\right| x_{p+1} y_{p+1} - \dots - \left|\lambda_{p+q}\right| x_{p+q} y_{p+q}.$$

Insbesondere ist also $B(f_k, f_k) = \lambda_k$ für alle k. Im letzten Schritt reskalieren wir die Basisvektoren von E_2 und erhalten die gewünschte Basis E:

$$e_k = \frac{1}{\sqrt{|\lambda_k|}} f_k, \quad k = 1, \dots, p + q.$$

In der Tat ist nun $B(e_k, e_k) = B\left(\frac{f_k}{\sqrt{|\lambda_k|}}, \frac{f_k}{\sqrt{|\lambda_k|}}\right) = \frac{1}{|\lambda_k|} \lambda_k = \operatorname{sign} \lambda_k$. Dann hat B in der Basis E die gewünschte Gestalt. ■

Für symmetrische Matrizen A bedeutet das, dass eine invertierbare Matrix $S \in \mathrm{GL}(n, \mathbb{R})$ existiert mit

$$A = S^\top \operatorname{diag}(I_p, -I_q, 0) S. \tag{7.2}$$

Man beachte, dass S in der Regel nicht mehr orthogonal ist und auch nicht eindeutig bestimmt. Im nächsten Abschnitt werden wir aber sehen, dass p und q nicht von der Wahl der Basis abhängen.

7.2.4 Der Sylvestersche Trägheitssatz

Satz 7.3 (Sylvesterscher Trägheitssatz) *Es sei B eine symmetrische Bilinearform auf dem reellen Vektorraum V, $\dim V = n$.*
Dann existieren Unterräume V_0, V_+ und V_- von V mit

$$V = V_0 \oplus V_+ \oplus V_-,$$

wobei $V_0 = \{x \in V \mid B(x, y) = 0 \ \forall y \in V\}$, $B(x,x) > 0$ für alle $x \in V_+ \setminus \{0\}$ und $B(x,x) < 0$ für alle $x \in V_- \setminus \{0\}$. Die Zahlen

$$p = \dim V_+ \quad \textit{und} \quad q = \dim V_-$$

sind Invarianten von B und hängen nicht von der speziellen Wahl der Unterräume V_+ und V_- ab. Die Differenz $p - q$ heißt Signatur *und $p + q$ heißt* Rang *der symmetrischen Bilinearform B.*

Einzig der Nullraum V_0 ist eindeutig bestimmt. Die Räume V_+ und V_- kann man variieren indem man einzelne Basisvektoren durch Vektoren von V_0 abändert.
Beweis. (a) Existenz der Zerlegung. Die Existenz folgt sofort aus dem Satz 7.2, indem man $V_+ = \operatorname{span}\{e_1, \ldots, e_p\}$, $V_- = \operatorname{span}\{e_{p+1}, \ldots, e_{p+q}\}$ und $V' := \operatorname{span}\{e_{p+q+1}, \ldots, e_n\}$ setzt. Es ist klar, dass dann V_+ und V_- die Bedingungen der Aufgabe erfüllen, dass $V = V_+ \oplus V_- \oplus V'$ und dass $V' \subset V_0$. Wir müssen noch zeigen, dass $V' \supset V_0$ gilt. Dazu sei $x = \sum_{i=1}^n x_i e_i \in V_0$. Also gilt $B(x, e_i) = x_i = 0$ für alle $i = 1, \ldots, p$ und $B(x, e_i) = -x_i = 0$ für alle $i = p+1, \ldots, p+q$. Somit hat x nur noch von Null verschiedene Koeffizienten x_i, $i > p + q$, also $x \in V'$; also $V' = V_0$.
(b) Eindeutigkeit von p und q. Es sei $V = V_0 \oplus W_+ \oplus W_-$ eine zweite Zerlegung dieser Art. Für $v \in W_+ \cap (V_0 \oplus V_-) \setminus \{0\}$ gilt dann einerseits

$$B(v, v) > 0, \quad \text{weil } v \in W_+$$

und andererseits gilt

$$B(v, v) \le 0, \quad \text{weil } v = v_0 + v_- \in V_0 \oplus V_- \quad B(v_0 + v_-, v_0 + v_-) = 0 + 0 + 0 + B(v_-, v_-) \le 0.$$

Dies ist aber ein Widerspruch. Folglich ist $W_+ \cap (V_0 \oplus V_-) = \{0\}$. Somit gilt nach dem Dimensionssatz 2.9 wegen $W_+ + ((V_0 \oplus V_-) \subset V$

$$\begin{aligned} \dim V &\ge \dim(W_+ + (V_0 \oplus V_-)) = \dim W_+ + \dim V_0 \oplus V_- = \dim W_+ + \dim V_0 + \dim V_- \\ &= \dim W_+ + (n - p - q) + q \quad \Longrightarrow \quad p = \dim V_+ \ge \dim W_+. \end{aligned}$$

Aus Symmetriegründen (Vertauschen der Rollen von W_+, W_- und V_+, V_-) muss aber auch die umgekehrte Ungleichung gelten; somit gilt Gleichheit $\dim W_+ = \dim V_+$ und damit aus Dimensionsgründen auch $\dim V_- = \dim W_-$, was zu zeigen war. ∎

Folgerung 7.4 *Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix, $S \in \mathrm{GL}(n,\mathbb{R})$ und $A' = S^\top AS$. Dann haben A und A' die gleiche Anzahl von positiven und und die gleiche Anzahl von negativen Eigenwerten. Insbesondere ist A positiv definit (negativ definit, indefinit, usw.) genau dann, wenn A' positiv definit (negativ definit, indefinit, usw.) ist.*

Beweis. A und A' beschreiben ein und dieselbe Bilinearform nur in einer anderen Basis (weil S invertierbar ist). Daher ist $p = \dim V_+$ und $q = \dim V_-$ die Anzahl der positiven bzw. negativen Eigenwerte von A – und auch von A'. ∎

7.2.5 Sylvesterkriterium für positiv definite Matrizen

Bemerkung 7.2 (a) Jede positiv (negativ) definite symmetrische Bilinearform ist nichtentartet. Das folgt sofort aus dem Trägheitssatz da alle Eigenwerte von A positiv (negativ) sind und somit A invertierbar.

(b) Die quadratische reelle, symmetrische Matrix $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ ist genau dann positiv definit, wenn

$$\det A = ad - b^2 > 0 \quad \text{und} \quad a > 0.$$

Zum Beweis betrachten wir die Gleichung, $x = (x_1, x_2) \in \mathbb{R}^2$,

$$a\,Q_A(x) = a(ax_1^2 + 2bx_1x_2 + dx_2^2) = (ax_1 + bx_2)^2 + (ad - b^2)x_2^2.$$

sind die obigen beiden Bedingungen erfüllt, so gilt offenbar $Q_A(x) > 0$ für alle $x \neq 0$. Ist umgekehrt $Q_A(x) > 0$, so insbesondere $Q((1,0)) = a > 0$ und somit auch $aQ((-b,a)) = (ad - b^2)a^2 > 0$, also $\det A > 0$.

(c) Ist $A = (a_{ij})$ positiv definit, so gilt
(1) $a_{ii} > 0$ für alle $i = 1,\dots,n$.
(2) $a_{ii}a_{jj} - a_{ij}^2 > 0$ für alle $1 \le i < j \le n$.

Ist A negativ definit, so gilt
(1) $a_{ii} < 0$ für alle $i = 1,\dots,n$.
(2) $a_{ii}a_{jj} - a_{ij}^2 > 0$ für alle $1 \le i < j \le n$.

Die erste Behauptung folgt aus $Q_A(\mathrm{e}_i) > 0$ bzw. aus $Q_A(x_1\mathrm{e}_i + x_2\mathrm{e}_j) > 0$.

(d) Ist A' eine $k \times k$ Untermatrix von A in der linken oberen Ecke von A, Also

$$A = \begin{pmatrix} A' & * \\ * & * \end{pmatrix},$$

so ist A' ebenfalls positiv definit. Man nennt $\det A'$ den *k ten Hauptminor* von A.
Ist nämlich A' eine $r \times r$-Matrix, so wählt man $x = (x_1,\dots,x_r,0,\dots,0)$ und erhält $Q_A(x) = x^\top Ax = x'^\top A'x' > 0$, wobei $x' = (x_1,\dots,x_r)$.

Wir verallgemeinern die Eigenschaft (b) für beliebige Dimension.

Satz 7.5 (Sylvestersches Definitheitskriterium) *Es sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, $A = A^\top$, eine symmetrische reelle Matrix. Für $k = 1,\dots,n$ bezeichen $A_k = (a_{ij})_{\substack{i=1\dots k\\ j=1\dots k}}$ die „abgeschnittenen“ $k \times k$-Matrizen. Außerdem setzen wir $d_k = \det A_k$.*
Dann gilt:

(a) Q_A *ist genau dann positiv definit, wenn* $d_k > 0$ *für alle* $k = 1, \dots, n$.
(b) Q_A *ist genau dann negativ definit, wenn* $(-1)^k d_k > 0$ *für alle* $k = 1, \dots, n$

Man nennt die die Determinanten d_k die *Hauptminore* von A.
Beweis. Wir beweisen (a). (1) Die *Notwendigkeit* der Bedingung ist klar: Ist A positiv definit, so ist nach Bemerkung 7.2 (d) für alle k die Matrix A_k positiv definit. Damit sind alle ihre Eigenwerte $\lambda_1^{(k)}, \dots, \lambda_k^{(k)}$ positiv. Ferner gilt $\det A_k = \lambda_1^{(k)} \cdots \lambda_k^{(k)} > 0$.
(2) *Hinlänglichkeit.* Seien alle $d_k > 0$. Insbesondere sind alle A_k invertierbar. Man schreibt nun $A = A_n$ in Kästchenform wie folgt auf:

$$A = \begin{pmatrix} A_{n-1} & v \\ v^\top & a \end{pmatrix},$$

mit der symmetrischen $(n-1)\times(n-1)$-Matrix A_{n-1}, einem Spaltenvektor $v \in \mathbb{R}^{n-1}$ und $a \in \mathbb{R}$. Man rechnet nach, dass nun

$$A = S^\top \begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix} S, \quad S = \begin{pmatrix} I_{n-1} & A_{n-1}^{-1} v \\ 0 & 1 \end{pmatrix}, \quad b = a - v^\top A_{n-1}^{-1} v.$$

Nach Folgerung 7.4 ist A genau dann positiv definit, wenn die mit S transformierte Matrix $\begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix}$ positiv definit ist. Dies ist wiederum äquivalent zur positiven Definitheit von A_{n-1} und $b > 0$. Nach Induktionsvoraussetzung ist aber $d_1, d_2, \dots, d_{n-1} > 0$ und damit A_{n-1} positiv definit. Außerdem folgt aus der Zerlegung

$$\det A = \det S \det A_{n-1} b \det S.$$

Nun ist aber S aber eine obere Dreiecksmatrix mit lauter Einsen auf der Diagonale; also ist $\det S = 1$. Folglich gilt $\det A = b \det A_{n-1}$. Mit $\det A > 0$ und $\det A_{n-1} > 0$ folgt $b > 0$ und die positive Definitheit von A ist gezeigt.
(3) Sei die Bedingung in (b) erfüllt. Wir betrachten die symmetrische Matrix $-A$. Alle Hauptminore von $-A$ sind dann positiv und damit ist $-A$ nach (a) positiv definit und folglich ist A negativ definit. Die Notwendigkeit folgt analog. ∎

Beispiel 7.4 Die Matrix

$$A = \begin{pmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{pmatrix}$$

ist positiv definit, da ihre Hauptminore

$$\det A_1 = 2, \quad \det A_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3, \quad \det A = \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} = 1$$

alle positiv sind.

Für positiv definite Matrizen können wir Folgendes zusammenfassend feststellen.

Satz 7.6 (Äquivalenzsatz für positiv definite Matrizen) *Für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$, $A^\top = A$, sind äquivalent*

i. *A ist positiv definit.*
ii. *Es gibt eine invertierbare quadratische Matrix $S \in \mathrm{GL}(n,\mathbb{R})$ mit $A = S^\top S$.*
iii. *Alle Hauptminore sind positiv.*
iv. *Alle Eigenwerte von A sind positiv.*
v. *Es gibt eine orthogonale Matrix $U \in \mathrm{O}(n)$ und eine Diagonalmatrix A' mit positiven Diagonalelementen, so dass $A = U^\top A' U = U^{-1} A' U$.*
vi. *S ist positiv semidefinit und $\det S \neq 0$.*
vii. *S ist invertierbar und S^{-1} ist positiv semidefinit.*

7.2.6 Beispiel zur Hauptachsentransformation

Geben sei die quadratische Gleichung

$$f(x_1,x_2) = x_1^2 + 6x_1x_2 + x_2^2 + 8x_1 + 8x_2 + 1 = 0, \quad x_1,x_2 \in \mathbb{R},$$

welche eine Kurve $C = \{(x_1,x_2) \in \mathbb{R}^2 \mid f(x_1,x_2) = 0\}$ in der Ebene $\mathbb{R}^2$ definiert. Ist C die leere Menge, ein Geradenpaar, eine Ellipse, eine Hyperbel oder eine Parabel?
Wir schreiben $f(x)$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, in Matrixform,

$$f(x) = x^\top A x + b^\top x + c \quad A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 8 \end{pmatrix}, \quad c = 1.$$

und führen mit A die Hauptachsentransformation durch. Wegen $\det A = 1 - 9 = -8$ und $\operatorname{tr} A = 2$ ist $\chi_A(\lambda) = \lambda^2 - 2\lambda - 8 = (\lambda-4)(\lambda+2)$ und damit sind $\lambda_1 = 4$ und $\lambda_2 = -2$ die Eigenwerte von A. Die Eigenräume sind:

$$V_{\lambda_1} = \mathrm{Lös}(A-4,0) = \{x \in \mathbb{R}^2 \mid -3x_1 + 3x_2 = 0\} = \mathbb{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$V_{\lambda_2} = \mathrm{Lös}(A+2,0) = \{x \in \mathbb{R}^2 \mid 3x_1 + 3x_2 = 0\} = \mathbb{R}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Die Orthonormalbasis aus Eigenvektoren lautet also $E = \{(1,1)/\sqrt{2}, (1,-1)/\sqrt{2}\}$. Die Transformationsmatrix (vergleiche Abschnitt 3.3.6) $S = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ führt auf die neuen Koordinaten

$$y = S^{-1}x, \quad x = Sy = \begin{pmatrix} \frac{1}{\sqrt{2}}(y_1+y_2) \\ \frac{1}{\sqrt{2}}(y_1-y_2) \end{pmatrix}, \quad x_1 = \frac{1}{\sqrt{2}}(y_1+y_2), \quad x_2 = \frac{1}{\sqrt{2}}(y_1-y_2).$$

Setzt man x_1 und x_2 in die Gleichung der Kurve C ein, so erhält man

$$0 = 4y_1^2 - 2y_2^2 + 8\sqrt{2}y_1 + 1 = \left(y_1 + 4\sqrt{2}\right)^2 - 2y_2^2 - 31 = 4z_1^2 - 2z_2^2 - 31.$$

wobei $z_1 = y_1 + 4\sqrt{2}$, $z_2 = y_2$ eine weitere Koordinatentransformation (Verschiebung in Nullpunktslage) ist. Weil A jeweils einen positiven und einen negativen Eigenwert hat, handelt es sich bei C um eine Hyperbel.

7.3 Geometrie in euklidischen Räumen

Im Folgenden sei V ein euklidischer Raum mit Skalarprodukt.

7.3.1 Geraden und Strecken

Eine *Gerade* g in V ist durch einen Punkt $p_0 \in V$ und einen Richtungsvektor $r \in V$, $r \neq 0$, gegeben,

$$g = \{p_0 + t\,r \mid t \in \mathbb{R}\}.$$

Ist $p_0 = 0$, so ist g sogar ein eindimensionaler Teilraum. Variiert der Parameter t nur in einem endlichen Intervall $[a, b]$, so nennt man $s = \{p_0 + t\,r \mid t \in [a, b]\}$ eine *Strecke*. Die Strecke $\overline{p_0 p_1}$ von p_0 nach p_1, $p_0, p_1 \in V$ hat die Gestalt $\overline{p_0 p_1} = \{(1-\lambda)p_0 + \lambda p_1 \mid \lambda \in [0, 1]\}$.

Definition 7.5 Es sei $p_0 \in V$ und $U \subset V$ ein linearer Teilraum von V, dann bezeichnet man

$$M = p_0 + U = \{p_0 + u \mid u \in U\}$$

als *affinen Teilraum* von V. Wir bezeichnen $\dim M := \dim U$ als *Dimension* des affinen Teilraumes M.

Punkte und Geraden sind also genau die null- bzw. eindimensionalen affinen Teilräume von V. Die $n-1$-dimensionalen affinen Teilräume bezeichnet man als *Hyperebenen*.

Beispiel 7.5 (a) Für jedes lineare Gleichungssystem $Ax = b$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, ist $\text{Lös}(A, b)$ ein affiner Teilraum von $\mathbb{R}^n$ der Dimension $\operatorname{def} A$.
(b) Ist $v \in V$, $v \neq 0$, so ist

$$v^\perp = \{x \in V \mid \langle v\,,\,x\rangle = 0\}$$

eine Hyperebene durch den Ursprung. Jede Hyperebene im $\mathbb{R}^n$ ist gegeben durch eine einzige lineare Gleichung:

$$H = \{(x_1, \dots, x_n) \mid a_1 x_1 + \cdots + a_n x_n = c\},$$

wobei $a = (a_1, \dots, a_n) \neq 0$, $c \in \mathbb{R}$. Die Hyperebenen des $\mathbb{R}^3$ sind die Ebenen, die des $\mathbb{R}^2$ sind die Geraden.

7.3.2 Kurven und Flächen zweiter Ordnung — Quadriken

Die allgemeine Gleichung 2. Grades im $\mathbb{R}^n$ hat die Gestalt (mit a_{ij}, a_k, $c \in \mathbb{R}$)

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j + 2\sum_{i=1}^{n} a_k x_k + c = 0, \quad a_{ij} = a_{ji}.$$

Fasst man die Koeffizienten zu eine symmetrischen Matrix $B = (a_{ij}) \in \mathbb{R}^{(n+1)\times(n+1)}$ zusammen, $B = \begin{pmatrix} A & a \\ a^\top & c \end{pmatrix}$, wobei a_{ij} die Koeffizienten aus der Gleichung sind, so kann man die Gleichung schreiben als $P_B(x) = \langle x\,,\,Ax\rangle + 2\,\langle a\,,\,x\rangle + c = 0$.

Als *Quadrik* bezeichnet man jede Menge (sofern sie nicht leer ist)

$$M_B = \{x \in \mathbb{R}^n \mid P_B(x) = 0\}, \quad A \neq 0.$$

Die Gestalt einer Quadrik ist vollständig bestimmt durch Rang und Signatur von A und B. Im Falle $n = 2$ erhält man die folgenden Normalformen von Quadriken (Kurven 2. Ordnung):

i. Ellipse, $a_1^2 x_1^2 + a_2^2 x_2^2 = 1$, wobei $a_1 a_2 \neq 0$.

ii. Hyperbel oder zwei sich schneidende Geraden, $a_1^2 x_1^2 - a_2^2 x_2^2 = 1$ oder $= 0$ mit $a_1 a_2 \neq 0$.

iii. Parabel, $x_1^2 + b x_2 = 0$ mit $b \neq 0$.

iv. $x_1^2 = c$ mit $c > 0$ (paralleles Geradenpaar) oder mit $c = 0$ (Doppelgerade).

Literaturverzeichnis

[Ant98] H. Anton. *Lineare Algebra. Einführung, Grundlagen, Übungen.* Spectrum Lehrbuch. Spektrum Akademischer Verlag, Heidelberg, 1998.

[Jän04] K. Jänich. *Lineare Algebra.* Springer-Lehrbuch. Springer-Verlag, Berlin, 10 edition, 2004.

[Koe97] M. Koecher. *Lineare Algebra und analytische Geometrie.* Springer-Verlag, Berlin, 4 edition, 1997.

[Kow79] H.-J. Kowalsky. *Lineare Algebra.* De Gruyter Lehrbuch. Walter de Gruyter, Berlin, New York, 1979.

PD Dr. A. Schüler
Mathematisches Institut
Universität Leipzig
04009 Leipzig
`mailto:Axel.Schueler@math.uni-leipzig.de`