Klaus Tammer

Notwendige und hinreichende Bedingungen für (strenge) Konvexität, Pseudokonvexität und (strenge) Quasikonvexität einer quadratischen Funktion bezüglich einer konvexen Menge

Aplikace matematiky, Vol. 21 (1976), No. 2, 97–110

Persistent URL: `http://dml.cz/dmlcz/103628`

SVAZEK 21 (1976) APLIKACE MATEMATIKY ČÍSLO 2

NOTWENDIGE UND HINREICHENDE BEDINGUNGEN FÜR (STRENGE) KONVEXITÄT, PSEUDOKONVEXITÄT UND (STRENGE) QUASIKONVEXITÄT EINER QUADRATISCHEN FUNKTION BEZÜGLICH EINER KONVEXEN MENGE

KLAUS TAMMER

(Eingegangen 11. Februar 1974)

1. EINLEITUNG

In der Theorie der nichtlinearen Optimierung spielt bekanntlich der Begriff der konvexen Funktion eine entscheidende Rolle. Das liegt natürlich an einigen sehr wichtigen Eigenschaften, die mit diesem Begriff verbunden sind. Man vergleiche etwa die Zusammenstellung bei Elster in [5].

In den letzten Jahren wurde dieser Begriff in verschiedenen Varianten verallgemeinert mit dem Ziel, neben den bereits bekannten konvexen Funktionen weitere Klassen von Funktionen zu erhalten, für welche wenigstens einige wichtige dieser Eigenschaften konvexer Funktionen auch noch zutreffen. Die für die Anwendung unseres Erachtens nach wichtigsten verallgemeinerten Konvexitätsbegriffe sind die der Pseudokonvexität sowie der Quasikonvexität und strengen Quasikonvexität. Man vergleiche hierzu ebenfalls etwa [5] bzw. [17].

Die Überprüfung dieser Begriffe für konkrete Funktionen mittels der entsprechenden Definition ist aber oft nicht einfach. Deshalb ist es wichtig, hierfür notwendige, hinreichende und möglichst auch notwendige und hinreichende Bedingungen zur Verfügung zu haben.

Einige solcher Bedingungen sind bereits bekannt, insbesondere, was Konvexitätskriterien betrifft. Die grösste Anwendung fand wohl bisher ein Kriterium für die Konvexität bzw. strenge Konvexität einer zweimal stetig differenzierbaren Funktion über einer absolut offenen konvexen Teilmenge des E_n. Man vergleiche etwa die recht allgemeine Formulierung bei Stoer und Witzgall [20].

Für die Quasikonvexität und Pseudokonvexität von zweimal stetig differenzierbaren Funktionen (teilweise unter weiteren Voraussetzungen) über absolut offenen bzw. n-dimensionalen konvexen Mengen wurden bisher von Arrow und Enthoven [1], Ferland [6], [8], Gerencsér [9] und Mangasarian [12] notwendige (aber i. allg.

nicht hinreichende) sowie hinreichende (aber i. allg. nicht notwendige) Bedingungen entwickelt.

Das Ziel dieser Arbeit soll es sein, für die Klasse der quadratischen Funktionen der Gestalt

$$(1) \qquad Q(x) = x'Cx + p'x$$

(C symmetrische (n,n) – Matrix, $p, x \in E_n$)

notwendige und hinreichende Bedingungen für die strenge Konvexität, Konvexität, Pseudokonvexität, strenge Quasikonvexität und Quasikonvexität bezüglich einer beliebigen konvexen Menge $G \subseteq E_n$ zu entwickeln, wobei wir eine einheitliche Darstellungsweise angestrebt haben.

Im ersten Teil werden wir den Fall dim $G = n$ behandeln. Die Frage der (strengen) Konvexität von $Q(x)$ über einer n-dimensionalen konvexen Menge G ist ja bereits seit langem geklärt und das entsprechende Konvexitätskriterium ist in jedem Lehrbuch der nichtlinearen Optimierung zu finden. Für die Pseudokonvexität bzw. Quasikonvexität von $Q(x)$ sind auch bereits verschiedene Bedingungen entwickelt worden. Die Untersuchungen beschränkten sich bisher aber nur auf den Fall dim $G = n$, wobei ausserdem in den meisten Arbeiten nur gewisse Spezialfälle für $Q(x)$ bzw. G betrachtet wurden.

Martos [13], [14] sowie Cottle und Ferland [3], [4] entwickelten notwendige und hinreichende Bedingungen für die Quasikonvexität von $Q(x)$ über $G = E_n^+$ sowie hinreichende Bedingungen für die Pseudokonvexität von $Q(x)$ über $G = E_n^+ \setminus \{0\}$.

Keri [10] legt dagegen eine beliebige n-dimensionale Bezugsmenge $G \subseteq E_n$ zugrunde. Seine notwendigen und hinreichenden Bedingungen für die Quasi- bzw. Pseudokonvexität von $Q(x)$ betreffen aber nur den Fall $p = 0$.

Notwendige und hinreichende Bedingungen für die Quasi- bzw. Pseudokonvexität einer beliebigen quadratischen Funktion bezüglich einer n-dimensionalen konvexen Teilmenge des E_n werden (neben anderen Resultaten) von Schaible [18] (vergleiche auch [19]) entwickelt. Er geht davon aus, dasss sich jede quadratische Funktion mit Hilfe einer geeingneten regulären linearen Transformation eindeutig auf eine gewisse Normalform zurückführen lässt. Da bei einer solchen Transformation eventuell vorhandene Konvexitätseigenschaften erhalten bleiben, kann er sich auf die Charakterisierung quasi- bzw. pseudokonvexer quadratischer Funktionen beschränken, die in einer solchen Normalform vorliegen. Dabei erhält er äquivalente Ergebnisse wie Ferland [7], der notwendige und hinreichende Bedingungen für die Quasi- bzw. Pseudokonvexität von solchen quadratischen Funktionen entwickelte, deren Matrix C in Diagonalgestalt vorliegt.

Bei Schaible [18] findet man aber auch ein für beliebige quadratische Funktionen formuliertes Quasi- bzw. Pseudokonvexitätskriterium. Allerdings erfordert das die

Kenntnis aller Eigenwerte und zugehöriger Eigenvektoren der Matrix C. Das läuft aber ebenso wie die Transformation auf Normalform insbesondere auf eine vollständige Hauptachsentransformation der Matrix C hinaus. Da dies aber bei grossem n im allgemeinen numerisch schwer zu realisieren ist, wollen wir hierzu äquivalente Bedingungen angeben, welche sich leichter überprüfen lassen. Möglichkeiten für diese Überprüfung werden in [24] aufgezeigt.

Für die entsprechenden Spezialfälle stimmen diese von uns entwickelten Bedingungen natürlich mit den Resultaten der oben zitierten Autoren überein.

Interessant dürfte es sein, dass einige wesentliche Ergebnisse von Swarup [21], [22], [23] aus den hier bewiesenen Resultaten direkt folgen.

Unter Ausnutzung einiger weiterer Hilfssätze werden wir dann im zweiten Teil alle im ersten Teil angegebenen notwendigen und hinreichenden Bedingungen auf konvexe Teilmengen G des E_n mit beliebiger Dimension r verallgemeinern.

Bezüglich der Konvexität von $Q(x)$ über einer niederdimensionalen konvexen Menge können wir uns dabei auf Ergebnisse von Orden [16] und Cottle [2] stützen.

Für die Anwendung dürfte die Erkenntnis von grosser Bedeutung sein, dass die für n-dimensionale Mengen notwendigen und hinreichenden Kriterien für niederdimensionale Mengen zwar natürlich noch hinreichend, jedoch i. allg. nicht mehr notwendig sind.

2. EINIGE BEGRIFFE UND HILFSSÄTZE

Wir wollen zunächst die von uns verwendeten verallgemeinerten Konvexitätsbegriffe angeben.

Es sei $f(x)$ eine reellwertige Funktion. Um die Darstellung einheitlich zu halten, wollen wir annehmen, dass $f(x)$ auf einer im E_n (d.h. absolut) offenen Menge definiert sei, welche die konvexe Menge G umfasse.

Für die Definition der Pseudokonvexität benötigen wir noch die Differenzierbarkeit von $f(x)$ über G.

Definition 1. *Die Funktion $f(x)$ heisst konvex über G*[1]*) wenn für alle $x_1, x_2 \in G (x_1 \neq x_2)$ und $t \in (0,1)$ die Ungleichung $f(tx_1 + (1-t)\,x_2) \leqq tf(x_1) + (1-t)\,.\,f(x_2)$ gilt.*
Sie heisst streng konvex über G, wenn die entsprechende Ungleichung stets echt erfüllt ist.

Definition 2. *Es sei $f(x)$ auf G differenzierbar. Dann heisst $f(x)$ über G pseudokonvex, wenn für alle $x_1, x_2 \in G$ mit $f(x_2) < f(x_1)$ die Ungleichung $(x_2 - x_1)'$ grad $f(x_1) < 0$ gilt.*

---

[1]) Ist keine Bezugsmenge G explizit angegeben, so ist immer $G = E_n$ gemeint und wir sprechen von einer konvexen Funktion schlechthin.

Definition 3. *Die Funktion $f(x)$ heisst streng quasikonvex über G, wenn für alle $x_1, x_2 \in G$ mit $f(x_2) < f(x_1)$ und $t \in (0,1)$ die Ungleichung $f(tx_1 + (1 - t) x_2) < < f(x_1)$ gilt.*

Definition 4. *Die Funktion $f(x)$ heisst quasikonvex über G, wenn für alle $x_1, x_2 \in G$ mit $f(x_2) < f(x_1)$ und $t \in (0,1)$ die Ungleichung $f(tx_1 + (1 - t) x_2) \leqq f(x_1)$ gilt.*

Neben den angegebenen werden insbesondere für die Quasikonvexität in der Literatur auch verschiedene andere Definitionen benutzt. Wie in [17] bewiesen, sind diese aber mindestens im Falle stetig differenzierbarer Funktionen äquivalent.

Die Beziehungen zwischen den verschiedenen Konvexitätstypen sind folgendermassen (vgl. etwa [9], [11], [17]).

Lemma 1.

1° *Ist $f(x)$ streng konvex über G, so ist $f(x)$ konvex über G.*

2° *Ist $f(x)$ konvex und differenzierbar über G, so ist $f(x)$ pseudokonvex über G.*

3° *Ist $f(x)$ pseudokonvex über G, so ist $f(x)$ streng quasikonvex über G.*

4° *Ist $f(x)$ streng quasikonvex über G, so ist $f(x)$ quasikonvex über G.*

Im allgemeinen gilt nach [11] in keiner der Aussagen aus Lemma 1 die Umkehrung. Für eine spezielle Funktionenklasse, zu der auch die quadratischen Funktionen gehören, kann aber die Umkehrung von Aussage 4° gezeigt werden.

Lemma 2. *Ist $f(x)$ quasikonvex über G und analytisch, so ist $f(x)$ streng quasikonvex über G.*

Beweis. Ist $f(x)$ quasikonvex aber nicht streng quasikonvex über G, dann existieren zwei Punkte $x_1, x_2 \in G$ mit $f(x_2) < f(x_1)$ und ein $t_0 \in (0,1)$ mit $f(t_0 x_1 + (1 - t_0) x_2) = = f(x_1)$ Aus der Quasikonvexität erhält man hieraus sofort $f(tx_1 + (1 - t) x_2) = = f(x_1)$ für alle $t \in [t_0, 1]$. Da aber $f(x)$ analytisch ist, so erhält man wegen $t_0 < 1$ hieraus sofort $f(tx_1 + (1 - t) x_2) = f(x_1)$ für alle $t \in [0,1]$. Das steht aber im Wiederspruch zur Annahme $f(x_2) < f(x_1)$.

Zum Beweis vergleiche auch [15].

Es sei noch darauf hingewiesen, dass eine Funktion, welche über einer konvexen Menge eine der angegebenen Konvexitätseigenschaften besitzt, diese auch über jeder konvexen Teilmenge dieser Menge besitzt.

In den meisten Lehrbüchern zur nichtlinearen Optimierung wird nur ausgegangen von solchen Funktionen, die über dem gesamten E_n konvex (bzw. pseudokonvex oder quasikonvex) sind. Spätestens mit den Ergebnissen von Mangasarian [11] dürften aber auch Funktionen von Interesse sein, welche nur über gewissen Teilmengen des E_n eine der Konvexitätseigenschaften besitzen. Der folgende Hilfssatz zeigt, dass das man im Falle quadratischer Funktionen durch Beschränkung auf solche, welche über dem gesamten E_n quasikonvex sind, nur wieder die konvexen quadratischen

Funktionen erhält und somit das Ziel der Erweiterung der Klasse der konvexen Funktionen nicht erreichen kann. Diese Aussage wurde bereits in [14] angegeben und ist sehr leicht beweisbar.

Lemma 3. *Ist $Q(x)$ quasikonvex über dem E_n, so ist $Q(x)$ konvex.*

Für die weiteren Untersuchungen benötigen wir noch einige weitere Hilfssätze.

Lemma 4. *Es sei G eine konvexe Teilmenge des E_n der Dimension r, N eine (n,r)-Matrix, deren Spalten eine Basis der affinen Hülle von G (aff G) bilden, und $x_0 \in$ aff G. Dann ist die Menge $G^* = \{y \in E_r / x_0 + Ny \in G\}$ konvex und es gilt dim $G^* = r$.*

Der Beweis der Konvexität von G^* dürfte klar sein. Die Dimensionsaussage folgt sofort daraus, dass im E_n jede konvexe Menge stets relativ innere Punkte besitzt.

Lemma 5. *Es seien die Voraussetzungen von Lemma 4 erfüllt. Die Funktion $f(x)$ sei auf G definiert und es sei $f^*(y) = f(x_0 + Ny)$ ($f^*(y)$ ist offenbar auf G^* definiert). Dann ist $f(x)$ genau dann über G streng konvex (konvex, pseudokonvex, streng quasikonvex, quasikonvex), wenn $f^*(y)$ über G^* streng konvex (konvex, pseudokonvex, streng quasikonvex, quasikonvex) ist.*

Beweis. Der Beweis folgt aus der Tatsache, dass zwischen aff G und dem E_r und speziell zwischen den Elementen der Mengen G und G^* eine eineindeutige Abbildung angegeben werden kann, die gewisse Eigenschaften erfüllt. Zu jedem $y \in G^*$ existiert nämlich genau ein $x \in G$ und, weil die Spalten von N eine Basis von aff G bilden, zu jedem $x \in G$ genau ein $y \in G^*$ derart, dass die Beziehung $x = x_0 + Ny$ besteht. Bei dieser Abbildung gilt offenbar stets $f(x) = f^*(y)$, wenn $x = x_0 + Ny$ ist. Weiter gilt noch folgende Eigenschaft: Sind $y_1, y_2 \in G^*$ Bilder der Punkte $x_1, x_2 \in G$ und ist $t \in (0,1)$, so stellt der Punkt $ty_1 + (1 - t)\, y_2$ das Bild des Punktes $tx_1 + (1 - t)\, x_2$ dar. Schliesslich folgt natürlich aus der Differenzierbarkeit von $f(x)$ über G die Differenzierbarkeit von $f^*(y)$ über G^*. Mit Hilfe dieser Aussagen folgt der Beweis sehr leicht. Man vergleiche auch [18]. Es sei noch darauf hingewiesen, dass die Aussage dieses Satzes natürlich unabhängig davon ist, welche spezielle Basis N und welches x_0 man auswählt.

Eine besondere Rolle bei unseren Untersuchungen wird die Menge

$$L = \{x / 2Cx = -p\}$$

spielen. Ist sie nicht leer, so stellt sie offensichtlich einen affinen Unterraum dar. Je nachdem, ob C positiv semidefinit, negativ semidefinit oder indefinit ist, sind die Elemente von L gerade die Punkte, in denen $Q(x)$ das freie Minimum oder das freie Maximum annimmt, oder die freie Sattelpunkte von $Q(x)$.

Die Aussagen des folgenden Hilfssatzes sind aus der linearen Algebra wohlbekannt, sodass sich ein Beweis hier erübrigt.

Lemma 6. Es seien $\lambda_1, \ldots, \lambda_n$ *die Eigenwerte der Matrix* C *(wobei etwa die ersten* k *die von Null verschiedenen sind) und* $h_1, \ldots, h_n$ *die zugehörigen Eigenvektoren, die bereits so gewählt seien, dass sie eine Orthonormalbasis des* E_n *bilden.*

Die Funktion $Q(x)$ *aus* (1) *lässt sich dann mit Hilfe der Matrix* $N = (h_1, \ldots, h_n)$ *und eines geeigneten* $x_0 \in E_n$ *durch eine reguläre lineare Transformation* $x = x_0 + Ny$ *überführen auf die Form*

$$Q^*(y) = Q(x_0 + Ny) = \sum_{i=1}^{k} \lambda_i y_i^2 + \sum_{i=k+1}^{n} a_i y_i + s\,.$$

Es existiert genau dann eine reguläre Transformation $x = \phi(y)$ *derart, dass* $Q^*(y) = Q(\phi(y)) = \sum_{i=1}^{n} \alpha_i y_i^2 + s$ *gilt, wenn* $L \neq \emptyset$ *ist. Ist dies der Fall, so können wir* $\phi(y) = x_0 + Ny$ *mit* $x_0 \in L$ *ansetzen und erhalten diese Gestalt mit* $\alpha_i = \lambda_i$ *für* $i = 1, \ldots, n$.

Lemma 7.

a) *Für alle* $x \in L$ *ist* $Q(x) = \text{const} = Q_0$.

b) *Ist* h_1 *ein Eigenvektor zum Eigenwert* $\lambda_1 \neq 0$ *von* C, *so ist* $h_1' x = \text{const} = t_0$ *für alle* $x \in L$.

c) *Ist* h_1 *ein normierter Eigenvektor zum Eigenwert* λ_1 *von* C *und* $x_0 \in L$, *so lässt sich jedes* $x \in E_n$ *eindeutig darstellen in der Form* $x = x_0 + th_1 + z$ *mit* $t \in E_1$ *und* $z'h_1 = 0$. *Bei dieser Darstellung lässt sich dann* $Q(x)$ *darstellen in der Form* $Q(x) = Q_0 + \lambda_1 t^2 + z'Cz$.

d) *Ist* λ_1 *der einzige negative Eigenwert von* C, *so gilt* $z'Cz \geqq 0$ *für alle* z *mit* $z'h_1 = 0$.

Beweis.

a) Es seien $x_1, x_2 \in L$. Dann gilt $Q(x_1) = x_1'Cx_1 + p'x_1 = -1/2p'x_1$ und analog $Q(x_2) = -1/2p'x_2$. Es bleibt zu zeigen, dass $p'x_1 = p'x_2$ gilt. Aus der Voraussetzung ergibt sich weiter $p = -2Cx_1 = -2Cx_2$. Daraus folgt $p'x_1 = -2x_2'Cx_1 = x_2'p = p'x_2$.

b) Nach Voraussetzung gilt $Ch_1 = \lambda_1 h_1$ mit $\lambda_1 \neq 0$. Damit ergibt sich für jedes $x \in L$
$h_1'x = 1/\lambda_1\, h_1'Cx = -1/2\lambda_1\, h_1'p$, was offensichtlich von x unabhängig ist.

c) Die eindeutige Darstellbarkeit in der angegebenen Weise dürfte allgemein bekannt sein. Unter Ausnutzung der Beziehungen $z'h_1 = 0$, $Ch_1 = \lambda_1 h_1$, $h_1'h_1 = 1$, $p = 2Cx_0$, $h_1'x_0 = 1/2\lambda_1\, h_1'p$ und $Q(x_0) = Q_0$ erhalten wir die Darstellung von $Q(x)$.

d) Ist $z'h_1 = 0$, so existiert eine eindeutige Darstellung der Form $z = \sum_{i=2}^{n} \alpha_i h_i$, da die Vektoren h_i $(i = 2, \ldots, n)$ eine Basis des orthogonalen Komplements zu h_1

bilden. Setzt man diese Darstellung ein und benutzt die Beziehungen $Ch_i = \lambda_i h_i$ und $h_i' h_j = \delta_{ij}$, so erhält man $z'Cz = \sum_{i=2}^{n} \alpha_i^2 \lambda_i$.
Das ist aber nach Voraussetzung nicht negativ.

3. NOTWENDIGE UND HINREICHENDE BEDINGUNGEN FÜR n-DIMENSIONALE MENGEN

Wir betrachten in diesem Abschnitt den Fall $\dim G = n$. Allgemein bekannt ist das folgende Konvexitätskriterium.

Satz 1. *Die Funktion $Q(x)$ ist genau dann konvex (streng konvex) über G, wenn alle Eigenwerte der Matrix C nicht negativ (positiv) sind.*

Bemerkung 1. Die Bedingung aus Satz 1 ist äquivalent damit, dass die Matrix C positiv semidefinit (positiv definit) ist.

Dieses Kriterium zeigt, dass für den Fall $\dim G = n$ die Konvexität von $Q(x)$ allein von der Matrix C und nicht von p oder gar G abhängt. Die Konvexität von $Q(x)$ über einer n-dimensionalen Menge G hat auch stets die Konvexität von $Q(x)$ schlechthin zur Folge. Wie wir sehen werden, ist das alles im Falle der Pseudo- bzw. Quasikonvexität nicht so.

Da nach Lemma 1 die konvexen quadratischen Funktionen natürlich auch pseudokonvex und quasikonvex über jeder konvexen Menge $G \subseteq E_n$ sind, wir diese Funktionenklasse aber schon in Satz 1 vollständig beschrieben haben, wollen wir uns im folgenden nur für nichtkonvexe quadratische Funktionen und Bedingungen für deren Pseudo- bzw. Quasikonvexität interessieren. Es seien etwa $\lambda_1 \leqq \lambda_2 \leqq \ldots \leqq \lambda_n$ die Eigenwerte der Matrix C der Grösse nach geordnet und $h_1, \ldots, h_n$ die zugehörigen Eigenvektoren, die wieder bereits so gewählt seien, dass sie eine Orthonormalbasis bilden. Nach obigen Überlegungen und Satz 1 werden wir nun annehmen, dass $\lambda_1 < 0$ ist.

Wir verwenden wieder die Bezeichnungen Q_0, t_0, h_1, L aus Lemma 7 und führen ausserdem noch folgende weitere ein.

Es sei $K(s) = \{x / Q(x) \leqq s\}$ für $s \in E_1$ und, falls $L \neq \emptyset$ ist

$$K_1(s) = \{x \in K(s) / x'h_1 \geqq t_0 + \sqrt{((s - Q_0)/\lambda_1)}\} ,$$

$$K_2(s) = \{x \in K(s) / x'h_1 \leqq t_0 - \sqrt{((s - Q_0)/\lambda_1)}\} .$$

Mit diesen Grössen werden wir folgenden Hilfssatz formulieren.

Lemma 8. *Es gelten folgende Voraussetzungen:*

1° *Die Matrix C besitze genau einen negativen Eigenwert λ_1.*
2° *Es sei $L \neq \emptyset$.*
3° *Es sei $s \leqq Q_0$.*

Dann gilt:

a) $K(s) = K_1(s) \cup K_2(s)$.

b) $K_1(s) \cap K(s) = \varnothing$ *für* $s < Q_0$ *und* $K_1(Q_0) \cap K_2(Q_0) = L$.

c) *Es ist* $x \in K_1(s)$ *genau dann, wenn* $2x_0 - x \in K_2(s)$ *ist für ein beliebiges* $x_0 \in L$.

d) $K_1(s)$ *und* $K_2(s)$ *sind konvex, abgeschlossen und unbeschränkt und es gilt* int $K_i(s) = \{x \in K_i(s)/Q(x) < s\} \neq \varnothing$ *für* $i = 1, 2$. *Die Mengen* $K_1(Q_0)$ *und* $K_2(Q_0)$ *sind zusätzlich noch Kegel.*

Beweis. Die meisten Aussagen folgen unmittelbar aus [7] (z.T. auch aus [10]) durch Anwendung von Lemma 6 und Lemma 7. Die Mengen $K_1(Q_0)$ und $K_2(Q_0)$ entsprechen offensichtlich den in [7] Satz 3 ff. charakterisierten Mengen $\overline{T}^+$ und und $\overline{T}^-$ bei der Abbildung aus Lemma 6. Somit gelten die Eigenschaften a) und d) für $K_1(Q_0)$ und $K_2(Q_0)$, da sie bei einer regulären linearen Transformation erhalten bleiben. Für $s < Q_0$ folgt die Aussage a) sofort aus Lemma 7 c) und d). Die Aussage b) ist für $s < Q_0$ offensichtlich und folgt für $s = Q_0$ durch Ausrechnen ebenso wie die Aussage c).

Die Konvexität von $K_1(s)$ und $K_2(s)$ ergibt sich als Folgerung aus der Quasikonvexität von $Q(x)$ unter den Voraussetzungen der folgenden beiden Sätze. Sie liesse sich aber auch unabhängig davon zeigen. Die übrigen Aussagen von d) dürften klar sein.

Die Kriterien für die Pseudokonvexität bzw. Quasikonvexität von $Q(x)$ über G aus [7], [10] usw. lassen sich nun folgendermassen auf den allgemeinen Fall übertragen.

Satz 2. *Die Matrix C sei nicht positiv semidefinit, besitze also mindestens einen negativen Eigenwert* λ_1.

Dann ist $Q(x)$ *genau dann quasikonvex über* G, *wenn folgende Bedingungen erfüllt sind:*

1° *Die Matrix C besitzt genau einen negativen Eigenwert.*
2° *Es ist* $L \neq \varnothing$.
3° *Es ist* $G \subseteq K_1(Q_0)$ oder $G \subseteq K_2(Q_0)$.

Der Beweis ergibt sich mit den gleichen Überlegungen wie bei Lemma 8 aus Theorem 32 und Satz 24 in [7] unter Anwendung von Lemma 5 auf die in Lemma 6 angegebene Transformation.

Bemerkung 2. Bedingung 3° aus Satz 2 kann für $n \geqq 2$ ersetzt werden durch die Bedingung $Q(x) \leqq Q_0$ für alle $x \in G$. Ausserdem gilt noch $K_1(s) = \{x \in K(s)/x'h_1 \geqq t_0\}$ und $K_2(s) = \{x \in K(s)/x'h_1 \leqq t_0\}$.

Satz 3. *Die Matrix C sei nicht positiv semidefinit, besitze also mindestens einen negativen Eigenwert* λ_1. *Dann ist* $Q(x)$ *genau dann pseudokonvex über* G, *wenn folgende Bedingungen erfüllt sind:*

1° *Die Matrix C besitzt genau einen negativen Eigenwert.*
2° *Es ist* $L \neq \emptyset$.
3° *Es ist* $G \subseteq K_1(Q_0) \setminus L$ oder $G \subseteq K_2(Q_0) \setminus L$.

Der Beweis folgt analog wie bei Satz 2 diesmal aus Theorem 26, Theorem 33 und Satz 24 aus [7]. Er ergibt sich aber auch mit Satz 2 und folgender Bemerkung.

Bemerkung 3. Es sei $Q(x)$ über G quasikonvex aber nicht konvex. Dann ist $Q(x)$ über G genau dann pseudokonvex, wenn für alle $x \in G$ die Bedingung grad $Q(x) \neq 0$ erfüllt ist.

Beweis. Ist $Q(x)$ quasikonvex über G und grad $Q(x) \neq 0$, so ist nach [8] Theorem 12 die Funktion $Q(x)$ pseudokonvex über G.

Existiert dagegen ein Punkt $x_0 \in G$ mit grad $Q(x_0) = 0$, so ist zunächst offensichtlich $x_0 \in L$. Da aber $Q(x)$ quasikonvex aber nicht konvex über G ist und $\dim G = n$ gilt, erhält man mit Hilfe von Satz 2 und Lemma 8d) die Existenz eines $x_1 \in G$ mit $Q(x_1) < Q(x_0) = Q_0$. Für dieses Punktepaar ist aber offensichtlich Definition 2 nicht erfüllt, weshalb dann $Q(x)$ über G nicht pseudokonvex wäre.

Die Sätze 2 und 3 zeigen, dass die Quasikonvexität bzw. Pseudokonvexität von $Q(x)$ über G nicht allein von der Matrix C abhängt, sondern auch von p und der Lage von G in Bezug auf den Funktionsverlauf von $Q(x)$ beeinflusst wird.

Wegen Lemma 2 liefert Satz 3 gleichzeitig ein notwendiges und hinreichendes Kriterium für die strenge Quasikonvexität von $Q(x)$ über G.

In Erweiterung von Lemma 8 können wir noch mit Hilfe von Satz 2 gewisse quadratische Hyperflächen im E_n charakterisieren, welche die n-dimensionale Verallgemeinerung der Hyperbel bzw. des zweischaligen Hyperboloiden darstellen. Wir sehen gleichzeitig einen gewissen geometrischen Hintergrund für die Sätze 2 und 3.

Bemerkung 4.

a) Sind die Voraussetzungen 1°, 2° und 3° aus Lemma 8 erfüllt und ist $n \geqq 2$, so besteht die quadratische Hyperfläche

$$H(s) = \{x / Q(x) = s\}$$

aus genau zwei „Schalen", welche für $s < Q_0$ keine gemeinsamen Punkte besitzen, während für $s = Q_0$ ihr Durchschnitt gerade L ist. Diese Hyperfläche bewirkt eine Einteilung[3]) des E_n in drei zusammenhängende, abgeschlossene und unbeschränkte Mengen $K_1(s)$, $K_2(s)$ und $K_3(s) = \{x / Q(x) \geqq s\}$ mit $\operatorname{int} K_i(s) = \{x \in K_i(s) / Q(x) < s\} \neq \emptyset$ $(i = 1, 2)$ und $\operatorname{int} K_3(s) = \{x \in K_3(s) / Q(x) > s\} \neq \emptyset$, von denen die ersten beiden sogar konvex sind und zueinander bezüglich L symmetrisch liegen, während $K_3(s)$ nicht konvex ist.

b) Ist C nicht positiv semidefinit und ist mindestens eine der Voraussetzungen aus

[3]) Das bedeutet $K_1(s) \cup K_2(s) \cup K_3(s) = E_n$, $\dim K_i(s) = n$ $(i = 1, 2, 3)$ und $\operatorname{int} K_i(s) \cap \operatorname{int} K_j(s) = \emptyset$ für $i \neq j$.

Lemma 8 nicht erfüllt, so ist die Menge $K(s)$ nicht konvex und lässt sich auch nicht als Vereinigung endlich vieler konvexer Mengen darstellen.

Beweis. Die Aussagen für $K_1(s)$ und $K_2(s)$ ergeben sich aus Lemma 8. Für $s = Q_0$ und den dort behandelten Spezialfall wurden sie auch in [7] bewiesen.

Die Tatsache, dass $K_3(s)$ zusammenhängend ist, ergibt sich daraus, dass für jedes $x \in K_3(s)$ und jedes $x_0 \in L$ deren Verbindungsstrecke stets in $K_3(s)$ liegt. Die übrigen Eigenschaften von $K_3(s)$ ergeben sich nach den gleichen Überlegungen wie in [7] für $K_1(s)$ und $K_2(s)$. Die Tatsache, dass $K_3(s)$ nicht konvex ist, sowie die Aussage b) ergeben sich aus Satz 2 und der etwa in [17] bewiesenen Bedingung, dass $Q(x)$ genau dann über einer konvexen Menge S quasikonvex ist, wenn alle Mengen der Form $\{x \in S / Q(x) \leqq s\}$ mit $s \in E_1$ konvex sind.

Swarup untersuchte in [21], [22], [23] spezielle quadratische Funktionen der Gestalt (auf ein Minimumproblem umgeschrieben)

$$Q(x) = -(c'x + c_0)(d'x + d_0) \tag{3}$$

bezüglich spezieller konvexer Mengen $G \subseteq E_n$ unter der Voraussetzung, dass beide Faktoren über G positiv sind. Er zeigte u.a., dass unter dieser Voraussetzung jedes lokale Minimum von (3) über G auch globales ist (eine wichtige Eigenschaft streng quasikonvexer Funktionen). Diese Tatsache ergibt sich aber auch aus Satz 2 und der folgenden Aussage.

Bemerkung 5. Ist $G \subseteq E_n$ konvex und $c'x + c_0 > 0$ und $d'x + d_0 > 0$ über G, so ist entweder $Q(x)$ konvex oder es sind die Bedingungen 1°, 2° und 3° aus Satz 2 erfüllt und $Q(x)$ aus (3) streng quasikonvex über G.

Beweis. Die Funktion $Q(x)$ aus (3) lässt sich zunächst schreiben in der Form $Q(x) = -\frac{1}{2}x'(cd' + dc')x - (c_0d' + d_0c')x - c_0d_0$, wobei die Matrix $(cd' + dc')$ symmetrisch ist. Es lässt sich dann zeigen, dass diese Matrix höchstens den Rang zwei haben kann und somit höchstens zwei von Null verschiedene Eigenwerte besitzt. Hat sie den Rang zwei, so lässt sich weiter zeigen, dass die quadratische Form $x'(cd' + dc')x$ sowohl positive als auch negative Werte annehmen kann, sodass diese Matrix in diesem Fall genau einen negativen und einen positiven Eigenwert besitzt. Ist der Rang kleiner als zwei, so ist Bedingung 1° natürlich trivialerweise erfüllt. Bedingungen 2° und 3° lassen sich für die verschiedenen Fälle für den Rang der zugehörigen Matrix leicht verifizieren, wobei u.a. herauskommt, dass $Q_0 \geqq 0$ ist. Hierbei wird aber die Voraussetzung benötigt, dass die beiden Faktoren in (3) über G positiv sind. Man vergleiche auch [18], [19].

4. VERALLGEMEINERUNG DER BEDINGUNGEN AUF BELIEBIG-DIMENSIONALE BEZUGSMENGEN

Im folgenden sei $\dim G = r$ mit $0 \leqq r \leqq n$ und es seien wie in Lemma 4 und Lemma 5 die Grössen N, x_0, G^* und $Q^*(y) = Q(x_0 + Ny)$ definiert. Mit Hilfe von

Lemma 5 lassen sich die Bedingungen der Sätze 1, 2 und 3 nun verallgemeinern.

Zunächst ist offenbar

$$Q^*(y) = x_0'Cx_0 + p'x_0 + y'(2N'Cx_0 + N'p) + y'N'CNy$$

wieder eine quadratische Funktion. Mit Hilfe von Lemma 4, Lemma 5 und Satz 1 erhalten wir sofort:

Satz 4. *Die Funktion $Q(x)$ ist genau dann konvex (streng konvex) über G, wenn alle Eigenwerte der Matrix $N'CN$ nicht negativ (positiv) sind.*

Bemerkung 6. Die Bedingung aus Satz 4 ist natürlich wieder äquivalent damit, dass die Matrix $N'CN$ positiv semidefinit (positiv definit) ist.

Ist C positiv semidefinit, so ist offenbar auch $N'CN$ positiv semidefinit. Die Umkehrung gilt i. allg. nicht. Besitzt C genau k negative Eigenwerte, so liegt die Anzahl der negativen Eigenwerte von $N'CN$ zwischen $\max(0, k - n + r)$ und k, wobei es Fälle gibt, in denen die untere bzw. die obere Grenze angenommen wird.

Die Aussage von Satz 4 erhält man auch durch Kombination der Ergebnisse aus [2] und [16].

Im folgenden werden wir wieder annehmen, dass die Matrix $N'CN$ mindestens einen negativen Eigenwert λ_1^* mit zugehörigem Eigenvektor h_1^* besitzt.

Auf die Funktion $Q^*(y)$ lassen sich zunächst voll Lemma 7 und Lemma 8 anwenden. Wir versehen die dort auftretenden Grössen mit einem Stern, d.h., wir verwenden die Bezeichnungen L^*, Q_0^*, t_0^*, h_1^*, λ_1^*, $K^*(s)$, $K_1^*(s)$, $K_2^*(s)$. Es ist also insbesondere

$$\begin{aligned}
L^* &= \{y/2N'C(x_0 + Ny) + N'p = 0\}\,,\\
Q^*(y) &= \text{const} = Q_0^* \quad \text{für alle} \quad y \in L^*\,,\\
h_1^{*\prime}y &= \text{const} = t_0^* \quad \text{für alle} \quad y \in L^*\,,\\
K^*(s) &= \{y/Q^*(y) \leqq s\}\,,\\
K_1^*(s) &= \{y \in K^*(s)/y'h_1^* \geqq t_0^*\}
\end{aligned}$$

und

$$K_2^*(s) = \{y \in K^*(s)/y'h_1^* \leqq t_0^*\}\,.$$

Mit diesen Grössen können dann Satz 2 und 3 auf $Q^*(y)$ bezogen formuliert werden und wir hätten dann nach Lemma 5 Bedingungen für die Quasi- bzw. Pseudokonvexität von $Q(x)$ über G. Die Bedingungen 2° und 3° aus diesen Sätzen können wir aber durch folgende Überlegungen etwas umformen, so dass die Variablen y_i und der Punkt x_0 nicht mehr auftauchen und die Verallgemeinerung von Satz 2 bzw. Satz 3 deutlicher wird.

Bemerkung 7.

1° Es ist $L^* \neq \varnothing$ genau dann, wenn $L^{**} = \{x \in \operatorname{aff} G/2N'Cx + N'p = 0\} \neq \varnothing$ ist, wobei $y \in L^*$ genau dann gilt, wenn $x = x_0 + Ny \in L^{**}$ gilt.

2° Es gilt $Q(x) = \text{const} = Q_0^*$ für alle $x \in L^{**}$.

3° Es gilt $\begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} x = \text{const} = t_0^{**} \left(= t_0^* + \begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} x_0 \right)$ für alle $x \in L^{**}$, wenn die $(n, n - r)$-Matrix M so gewählt wird, dass ihre Spalten zusammen mit den r Spalten von N eine Basis des E_n bilden.

4° Bezeichnen wir

$$K_1^{**}(s) = \left\{ x \in K(s) \middle/ \begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} x \geqq t_0^{**} \right\},$$

$$K_2^{**}(s) = \left\{ x \in K(s) \middle/ \begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} x \leqq t_0^{**} \right\},$$

so gilt für $s \leqq Q_0^*$ genau dann $y \in K_i^*(s)$, wenn $x = x_0 + Ny \in K_i^{**}(s)$ und $y \in \operatorname{int} K_i^*(s)$ genau dann, wenn $x = x_0 + Ny \in \operatorname{int} K_i^{**}(s)$ für $i = 1, 2$.

Beweis. Die Aussagen 1° und 2° folgen sehr einfach auf Grund der bereits in Lemma 5 benutzten Eigenschaften der durch $x = x_0 + Ny$ gegebenen eineindeutigen Abbildung zwischen aff G und dem E_r.

Die Aussagen 3° und 4° lassen sich durch Ausrechnen zeigen. Es gilt nämlich für $x = x_0 + Ny$ (d.h. $x \in \operatorname{aff} G$)

$$\begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} x = \begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} (x_0 + Ny) = \begin{pmatrix} h_1^* \\ 0 \end{pmatrix}' (N, M)^{-1} x_0 + h_1^{*\prime} y .$$

Damit dürfte der Beweis klar sein.

Damit erhalten wir unmittelbar die folgenden beiden Aussagen:

Satz 5. *Es sei $Q(x)$ nicht konvex über G. Dann ist $Q(x)$ genau dann quasikonvex über G, wenn folgende Bedingungen erfüllt sind:*

1° *Die Matrix $N'CN$ besitzt genau einen negativen Eigenwert.*

2° *Es ist $L^{**} \neq \emptyset$.*

3° *Es ist $G \subseteq K_1^{**}(Q_0^*)$ oder $G \subseteq K_2^{**}(Q_0^*)$.*

Satz 6. *Es sei $Q(x)$ nicht konvex über G. Dann ist $Q(x)$ genau dann pseudokonvex über G, wenn folgende Bedingungen erfüllt sind:*

1° *Die Matrix $N'CN$ besitzt genau einen negativen Eigenwert.*

2° *Es ist $L^{**} \neq \emptyset$.*

3° *Es ist $G \subseteq K_1^{**}(Q_0^*) \setminus L^{**}$ oder $G \subseteq K_2^{**}(Q_0^*) \setminus L^{**}$.*

Das folgende kleine Beispiel soll die dargestellte Teorie etwas veranschaulichen und insbesondere die Bedeutung der Ergebnisse des vierten Abschnitts deutlich machen.

Es sei $Q(x) = 3x_1^2 - x_2^2 - x_3^2$ und

$$G = \{x \in E_3 / x_1 - x_2 \quad = 1 ,\\ -x_1 \quad + x_3 \quad = 1 ,\\ x_i \geqq 0 \qquad i = 1, 2, 3\} .$$

Diese Funktion ist nach Satz 2 über keiner dreidimensionalen konvexen Menge quasikonvex oder gar konvex. In diesem Fall kann man etwa $N = (1, 1, 1)'$ wählen. Dann ist $N'CN = 1 > 0$ und somit $Q(x)$ über G sogar streng konvex.

Wie wir hier sehen, dürften die Kriterien aus Abschnitt vier besonders wichtig sein bei der Anwendung von Verfahren der konvexen Optimierung auf quadratische Optimierungsprobleme mit zunächst nichtkonvexer Zielfunktion beim Vorliegen von Gleichungsrestriktionen. Möglichkeiten der Bestimmung einer geeigneten Matrix N wurden für diesen Fall etwa in [2] angegeben.

Literaturverzeichnis

[1] *Arrow, K. J., Enthoven, A. C.:* Quasi-concave programming. Econometrica 29 (1961), 779—800.

[2] *Cottle, R. W.:* On the convexity of quadratic forms over convex sets. Operations Research 15 (1967), 170—172.

[3] *Cottle, R. W., Ferland, J. A. :* Matrix theoretic criteria for the quasi-convexity and pseudo-convexity of quadratic functions. Linear Algebra and its Applications 5 (1972), 123—136.

[4] *Cottle, R. W., Ferland, J. A.:* On pseudo-convex functions of nonnegative variables. Mathematical Programming 1 (1971), 95—101.

[5] *Elster, K. H.:* Ergebnisse und Probleme der nichtlinearen Optimierung. Wiss. Z. d. Technischen Hochschule Ilmenau 15 (1969), 37—61.

[6] *Ferland, J. A.:* Quasi-convex and pseudo-convex functions on solid convex sets. Techn. Rept. No 71884, Operations Research House, Stanford University, Stanford, Calif. (April 1971).

[7] *Ferland, J. A.:* A maximal domains of quasi-convexity and pseudo-convexity for quadratic functions. Mathematical Programming 3 (1972), 178—192.

[8] *Ferland, J. A.:* Mathematical programming problems with quasi-convex objective functions, Mathematical Programming 3 (1972), 296—301.

[9] *Gerencsér, L.:* On a close relation between quasiconvex and convex functions and related investigations. Math. Operationsforsch. u. Statist. 4 (1973), 197—182.

[10] *Keri, K.:* An examination of nonnegativity and quasiconvexity conditions of quadratic forms on the nonnegative orthant. Studia Scientarum Mathematicarum Hungarica 7 (1972), 11—20.

[11] *Mangasarian, O. L.:* Pseudo-convex functions. SIAM Journal on Control 3 (1965), 281—290.

[12] *Mangasarian, O. L.:* Convexity, pseudo-convexity and quasiconvexity of composite functions, Cahiers du Centre d'Etudes de Recherche Opérationelle 12 (1970), 114—122.

[13] *Martos, B.:* Subdefinit matrices and quadratic forms. SIAM Journal on Applied Mathematics 17 (1969), 1215—1223.

[14] *Martos, B.:* Quadratic programming with a quasi-convex objective function. Operations Research 19 (1971), 87—97.

[15] *Martos, B.:* Quasi-convexity and quasi-monocity in nonlinear programming, Studia Scintarium Mathematicarum Hungarica 2 (1967), 265—273.
[16] *Orden, A.:* Minimization of indefinit quadratic functions with linear constraints; in: Recent advances in mathematical programming. Graves, R. L., Wolfe, P. (ed.) New York 1963.
[17] *Ponstein, J.:* Seven kinds of convexity. SIAM Review 9 (1967), 115—119.
[18] *Schaible, S.:* Beträge zur quasikonvexen Programmierung, Dissertation, Köln 1971
[19] *Schaible, S.:* Quasi-concave, strictly quasiconcave and pseudo-concave functions; in: Operations Research Verfahren, Henn, R., Künzi, H. P., Schubert, H. (ed.), Meiseuheim 1973.
[20] *Stoer, J., Witzgall, Ch.:* Convexity and Optimization in Finite Dimensions I. Springer-Verlag, Berlin—Heidelberg—New York 1970.
[21] *Swarup, K.:* Programming with indefinit quadratic functions with linear constraints. Cahiers du Centre d'Etudes de Recherche Opérationelle 8 (1966), 132—136.
[22] *Swarup, K.:* Indefinit quadratic programming with nonlinear constraints. ZAMM 47 (1967), 411—413.
[23] *Swarup, K.:* Indefinit quadratic programming with a quadratic constraint. Ekonom.-Mat. Obzor 4 (1968), 69—75.
[24] *Tammer, K.:* Über eine Klasse von verallgemeinerten quadratischen Optimierungsproblemen mit nichtkonvexer Zielfunktion, die auf quasi- bzw. pseudokonvexe Probleme zurückführbar sind. Aplikace matematiky, eingereicht 1974.

Souhrn

NUTNÉ A POSTAČUJÍCÍ PODMÍNKY PRO (SILNOU) KONVEXNOST, PSEUDOKONVEXNOST A (SILNOU) KVASIKONVEXNOST FUNKCE VZHLEDEM KE KONVEXNÍ MNOŽINĚ

KLAUS TAMMER

V příspěvku jsou jednotným způsobem uvedeny nutné a postačující podmínky pro konvexnost, pseudokonvexnost a kvasikonvexnost kvadratické funkce na libovolné konvexní podmnožině n-rozměrného euklidovského prostoru.

Anschrift des Verfassers: Dr. *Klaus Tammer*, Humboldt-Universität Berlin, Sektion Mathematik, 1086 Berlin, PSF 1297, DDR.