Acta Sci. Math. (Szeged)
57 (1993), 569–586

On pseudoconvex functions

SÁNDOR KOMLÓSI*

Dedicated to Professor Béla Szőkefalvi-Nagy on his 80th birthday

1. Introduction

The importance of the convexity concept for functions in several branches of mathematics is rather well known. Since convex functions admits very nice extremal properties, therefore this type of functions have a particular importance in Optimization Theory (see [14]). From the standpoint of this theory the following properties are the most relevant: if $f(x)$ is convex, then

(i) the lower level sets $\{x : f(x) \le c\}$ are convex for any $c \in R$,

(ii) if in addition $f(x)$ is differentiable, then any stationary point of $f(x)$ ($\nabla f(x) = 0$) is a global minimum point of the given function.

The introduction of several kinds of generalizations of the convexity concept has been characteristic of the second half of our century. Numerous generalizations of convex functions have been derived which proved to be useful for extending optimality conditions, previously restricted to convex programs, to larger classes of optimization problems. For a review of generalized convexity and its applications see [1,4].

Characterizations of convex functions using first order (gradient) or second order (Hessian) information belong to the classical chapters of Function Theory. In this paper our attention will be focused only to pseudoconvexity introduced by Mangasarian for differentiable functions [13]. The most remarkable properties of pseudoconvex functions are properties (i) and (ii).

Received July 2, 1992.
* This research was supported by the National Science Foundation of Hungary (OTKA 1313/91).

Throughout this work, we shall be concerned with the real, single-valued, differentiable function $f(x)$ defined on the nonempty open set D in the n-dimensional Euclidean space $\mathbb{R}^n$. Let $\nabla f(y)$ denote the gradient vector of $f(x)$ at $y \in D$ and let "$^\top$" denote the transpose. First we recall some definitions.

Definition 1.1. $f(x)$ is called *pseudoconvex* on the convex set $C \subset D$ if condition (PCX) is satisfied:

$$\text{(PCX)} \qquad x, y \in C, f(x) < f(y) \text{ implies } \nabla f(y)^\top (x - y) < 0.$$

$f(x)$ is called *strictly pseudoconvex* on the convex set $C \subset D$ if condition (SPCX) is satisfied:

$$\text{(SPCX)} \qquad x, y \in C, x \neq y, f(x) \leq f(y) \text{ implies } \nabla f(y)^\top (x - y) < 0.$$

The aim of this paper is to provide first and second order characterizations for pseudoconvex functions using the so called quasi-Hessian approach developed by the author in [10] and has been used in analysing quadratic [11] and pseudolinear functions [12].

2. Local characterizations of pseudoconvexity

The method of our investigation is of analytical character based on the local analysis of the given function. To this end we will utilize local concepts of pseudoconvexity.

Definition 2.1. Assume that $f(x)$ is defined and differentiable around $a \in \mathbb{R}^n$. The function $f(x)$ is called *locally pseudoconvex at a* if there exists a neighborhood G of a such that condition (LPCX) holds:

$$\text{(LPCX)} \qquad \begin{array}{l} x \in G, \;\; f(x) < f(a) \text{ implies } \nabla f(a)^\top (x - a) < 0 \text{ and} \\ x \in G, \;\; f(x) = f(a) \text{ implies } \nabla f(a)^\top (x - a) \leq 0. \end{array}$$

The function $f(x)$ is called *locally strictly pseudoconvex* at a if there exists a neighborhood G of a such that condition (LSPCX) holds:

$$\text{(LSPCX)} \qquad x \in G, \;\; x \neq a, \;\; f(x) \leq f(a) \text{ implies } \nabla f(a)^\top (x - a) < 0.$$

The following theorem provides the link between the global and local concepts of (strict) pseudoconvexity.

Proposition 2.2. *Let the differentiable function $f(x)$ be defined on the open convex set $C \subset \mathbb{R}^n$. Then $f(x)$ is (strictly) pseudoconvex on C if and only if it is locally (strictly) pseudoconvex at each point of C.*

Proof. It is obvious that the global property implies the local one. Assume that $f(x)$ is locally pseudoconvex at each point of C. Let $x, y \in C$ be arbitrary and assume that $f(x) < f(y)$. We have to prove that $\nabla f(x)^\top (x - y) < 0$. Let us consider the following functions:

$$z(t) = (1-t)x + ty, \quad t \in [0,1] \qquad \text{and} \qquad \beta(t) = f(z(t)), \quad t \in [0,1].$$

First we prove that $\beta(t)$ attains its maximum on the closed interval $[0,1]$ only at 0. Assume for contradiction that the positive number t_0, $0 < t_0 < 1$, is the last maximum point of $\beta(t)$ in $[0,1]$. Setting $z_0 = (1-t_0)x + t_0 y$ and using the fact that

$$\beta'(t_0) = \nabla f(z_0)^\top (x - y) = 0$$

we get that for every $t \in [0,1]$

$$\nabla f(z_0)^\top (z(t) - z_0) = 0. \tag{2.1}$$

By assumption $f(x)$ is locally pseudoconvex at z_0 thus there exists a neighborhood G of z_0 such that condition (LPCX) is fulfilled. For this G one can find a positive number δ such that $t_0 < \delta < 1$ and $z(\delta) \in G$. From the definition of t_0 it is obvious that $f(z(\delta)) < f(z_0)$ and by (LPCX) it follows that

$$\nabla f(z_0)^\top (z(\delta) - z_0) < 0$$

contradicting to (2.1). So we have proved that $\beta(t) < \beta(0)$ for every $t \in [0,1]$. Taking into account that $f(x)$ is locally pseudoconvex at y there exists a positive number s such that $0 < t < s$ and $f(z(t)) < f(y)$ implies

$$\nabla f(y)^\top (z(t) - y) = t \nabla f(y)^\top (x - y) < 0. \tag{2.2}$$

Since $f(x(t)) = \beta(t) < \beta(0) = f(y)$ therefore the right hand side inequality in (2.2) gives the desired statement, $\nabla f(y)^\top (x - y) < 0$. Strict pseudoconvexity can be proved by similar reasoning.

■

Proposition 2.3. *Let $f(x)$ be a real function defined and differentiable around $a \in \mathbb{R}^n$. Let us consider the following 'level-surface' condition:*

$$\text{(LSC)} \qquad \begin{array}{c} \textit{there exists a neighborhood } G \textit{ of } a \textit{ such that:} \\ x \in G,\ \ f(x) = f(a) \ \textit{implies}\ \nabla f(a)^\top (x-a) \le 0. \end{array}$$

If $\nabla f(a) \neq 0$, then conditions (LPCX) *and* (LSC) *are equivalent.*

Proof. The implication (LPCX) $\Rightarrow$ (LSC) is trivial. Assume that condition (LSC) holds. We have to prove, that

$$x \in G,\ \ f(x) < f(a) \text{ implies } \nabla f(a)^\top (x-a) < 0. \tag{2.3}$$

First we show that (LSC) implies the following property:

$$z \in G,\ \ f(z) < f(a) \text{ implies } \nabla f(a)^\top (z-a) \le 0. \tag{2.4}$$

Assume for contradiction that there exists $z \in G$ such that $f(z) < f(a)$ and $\nabla f(a)^\top (z-a) > 0$. Since $z-a$ is a strict ascent direction of $f(x)$ at a, therefore there exists a real number $0 < t < 1$ such that $f(u) = f(a)$ for $u = tz + (1-t)a$. By (LSC) it follows that

$$\nabla f(a)^\top (u-a) = t\nabla f(a)^\top (z-a) \le 0, \tag{2.5}$$

which is in contradiction to the 'indirect' assumption $\nabla f(a)^\top (z-a) > 0$.

Now we prove (2.3). Assume that $x \in G$ and $f(x) < f(a)$. Since $\nabla f(a) \neq 0$, therefore there exists a direction $d \in \mathbb{R}^n$ such that $\nabla f(a)^\top d > 0$. Let δ be a sufficiently small positive number. Since $f(x)$ is continuous and $z = x + rd$ is sufficiently close to a, therefore $f(z) < f(a)$. By (2.4) it follows that

$$\nabla f(a)^\top (z-a) = \nabla f(a)^\top (x-a) + r\nabla f(a)^\top d \le 0$$

and thus $\nabla f(a)^\top (x-a) \le -r\nabla f(a)^\top d < 0$. ■

Remark 2.4. The second part of the proof is due to T. Rapcsák (cf.[15, Lemma 2.6]).

Applying the same reasoning the strict version of Proposition 2.3 can be proved.

Proposition 2.5. *Let $f(x)$ be a real function defined and differentiable around $a \in \mathbb{R}^n$. Let us consider the following strict 'level-surface' condition:*

$$\text{(SLSC)} \qquad \begin{array}{c} \textit{there exists a neighborhood } G \textit{ of } a \textit{ such that} \\ x \in G,\;\; x \neq a,\;\; f(x) = f(a) \textit{ implies } \nabla f(a)^\top (x-a) < 0. \end{array}$$

If $\nabla f(a) \neq 0$, then conditions (LSPCX) *and* (SLSC) *are equivalent.*

Combining the previous results let us formulate the following theorem motivating our further investigations on local (strict) pseudoconvexity.

Theorem 2.6. *Let $\nabla f(a) \neq 0$. Then $f(x)$ is locally (strictly) pseudoconvex at a if and only if there exists a neighborhood G of a such that the 'level surface' condition* (LSC) ((SLSC)) *holds.*

In the following chapter we shall use this characterization in order to characterize local pseudoconvexity.

3. First order characterization of local pseudoconvexity

In our further investigation the local analysis of the level surface $f(x) = f(a)$ will play a central role, whereas our main tool will be the implicit function theorem. To be able to apply this theorem we will assume throughout this section that $f(x)$ is continuously differentiable. Further on we assume, that $\nabla f(a) \neq 0$. We shall use the implicit function theorem in the following way.

Let the vectors $d, b_1, \ldots, b_{n-1}$ form an orthonormal basis in $\mathbb{R}^n$. We call this basis admissible for $f(x)$ at a if

$$\nabla f(a)^\top d \neq 0. \tag{3.1}$$

Let us form the $n \times (n-1)$ type matrix B from the basis vectors $\{b_1, b_2, \ldots, b_{n-1}\}$ in the following way:

$$B = [b_1\, b_2 \ldots b_{n-1}].$$

Let $x \in \mathbb{R}^n$ be arbitrary and let $u := B^\top x \in \mathbb{R}^{n-1}$ and $v := d^\top x \in \mathbb{R}$. It is clear that v and u are the coordinates of x in the basis $\{d, B\}$ which means that

$$x = vd + Bu. \tag{3.2}$$

Further on the vector x will be identified with the pair (v, u).

Consider now the equation $f(v, u) = f(a)$, where $a = (v_0, u_0)$ and the basis $\{d, B\}$ is admissible. By the implicit function theorem there exist a neighborhood G of a and a neighborhood N of u_0 and a unique function $h(u)$ defined and continuously differentiable on N such that

$$f(h(u), u) = f(a) \text{ and } (h(u), u) \in G \text{ for all } u \in N,$$

and

$$\text{for all } x \in G \text{ we have } \nabla f(x)^\top d \neq 0,$$

moreover

$$x \in G,\ \ f(x) = f(a) \text{ implies that } x = (h(u), u)$$

for some $u \in N$. Computing the gradient $\nabla h(u)$ from the implicit relation $f(h(u), u) = f(a)$, one gets

$$\nabla h(u)^\top = \frac{\nabla f(x)^\top B}{\nabla f(x)^\top d}, \tag{3.3}$$

where $u \in N$ and $x = (h(u), u)$.

The following modification of the implicit function $h(u)$ will play a central role in the sequel.

Definition 3.1. The function

$$H(u) := (-\nabla f(a)^\top d) h(u), \qquad u \in N, \tag{3.4}$$

is called the *corrected implicit function of $f(x)$ at a with respect to the admissible basis $\{d, B\}$.*

The following result is of technical importance and can be proved very easily applying (3.3) and (3.4).

Lemma 3.2. *Let the notations of the above implicit function theorem be applied. Then for each $x \in G$ and $u \in N$ such that $x = (h(u), u)$ one has*

$$H(u_0) + \nabla H(u_0)^\top (u - u_0) - H(u) = \nabla f(a)^\top (x - a). \tag{3.5}$$

In order to state the basic theorem of this section the following concept of local convexity will be of use.

Definition 3.3. Assume that $H(u)$ is defined and differentiable around $u_0 \in \mathbb{R}^m$. The function $H(u)$ is called *locally convex at* u_0, if there exists a neighborhood N of u_0 such that for every $u \in N$ one has

$$H(u) \geq H(u_0) + \nabla H(u_0)^\top (u - u_0). \tag{3.6}$$

If for $u \in N$, $u \neq u_0$ condition (3.6) holds with strict inequality, then $H(u)$ is called *locally strictly convex* at u_0.

Theorem 3.4. *Let the continuously differentiable function $f(x)$ satisfy condition $\nabla f(a) \neq 0$. Then $f(x)$ is locally (strictly) pseudoconvex at $a = (v_0, u_0)$ if and only if the corrected implicit function $H(u)$ is locally (strictly) convex at u_0.*

Proof. Necessity: Assume that $f(x)$ is locally pseudoconvex at a, where $\nabla f(a) \neq 0$. Let the notations of the above implicit function theorem be applied. Without loss of generality we may assume that the neighborhood G of a fulfils condition (LPCX), as well. Let $u \in N$ be arbitrary and set $x = (h(u), u)$. Since $f(x) = f(a)$, therefore, by combining (LPCX) and Lemma 3.2. we have that

$$H(u_0) + \nabla H(u_0)^\top (u - u_0) - H(u) = \nabla f(a)^\top (x - a) \leq 0,$$

which is equivalent to (3.6).

Sufficiency: Assume that condition (3.6) holds on N. Since for each $x \in G$ with $f(x) = f(a)$ one can find an $u \in N$ such that $x = (h(u), u)$, therefore applying Lemma 3.2 it follows that property (LSC) holds on G. By Theorem 2.6 it follows that condition (LPCX) holds on G, as well.

The statement on local strict pseudoconvexity of $f(x)$ at a and local strict convexity of $H(u)$ at u_0 can be proved by the same arguments. ∎

There is a very natural question concerning the above theorem as to how one can test the local (strict) convexity of the corrected implicit function $H(u)$ at u_0. There is a well known classical test for twice continuously differentiable functions involving the positive (semi)definiteness of the Hessian matrix of the implicit function $H(u)$ at u_0. The next section will be devoted to the study of this matrix.

4. The quasi-Hessian matrices

From now on let us assume that $f(x)$ is twice continuously differentiable. Then the corrected implicit function $H(u)$ is twice continuously differentiable, as well.

Definition 4.1. Let $f(x)$ be twice continuously differentiable around $a \in \mathbb{R}^n$ and suppose that the basis $\{d, B\}$ is admissible for $f(x)$ at $a = (u_0, v_0)$. The Hessian matrix $\nabla^2 H(u_0)$ is called the *quasi-Hessian of $f(x)$ at a associated with the admissible basis* $\{d, B\}$ and is denoted by $Q_B f(a, d)$.

This notation of the quasi-Hessian matrix $Q_B f(a, d)$ is going to stress the dependence of $H(u)$ on the basis $\{d, B\}$. It is obvious that the quasi-Hessian matrix is not unique, since any admissible basis $\{d, B\}$ determines a quasi-Hessian matrix. However they are all similar. In order to prove this statement we show how the quasi-Hessian $Q_B f(a, d)$ is related to the gradient $\nabla f(a)$ and Hessian $\nabla^2 f(a)$ of $f(x)$ at a. Differentiating twice the implicit relation $f(h(u), u) = f(a)$ one gets that

$$Q_B f(a, d) = B^\top [H_1 - H_2 + H_3] B, \tag{4.1}$$

where

$$\begin{aligned} H_1 &= \nabla^2 f(a), \\ H_2 &= \frac{1}{(\nabla f(a)^\top d)} (\nabla f(a) d^\top \nabla^2 f(a) + \nabla^2 f(a) d \nabla f(a)^\top), \\ H_3 &= \frac{d^\top \nabla^2 f(a) d}{(\nabla f(a)^\top d)^2} \nabla f(a) \nabla (a)^\top. \end{aligned}$$

First we point to the differential geometric connection of the quasi-Hessian matrix. Let us consider the level surface, defined by the equation $f(x) = f(a)$, 'near' a. For any admissible basis $\{d, B\}$ the implicit function $h(u)$ provides a possible parameterization of the elementary surface under consideration. If we compute the matrix $M_B f(a, d)$ of the second fundamental form of this surface at a with respect to the above parameterization and normal vector $\nabla f(a)/\|\nabla f(a)\|$ we obtain the following (cf. [16]):

$$M_B f(a, d) = (-1/\|\nabla f(a)\|) Q_B f(a, d). \tag{4.2}$$

It is well known from the theory of elementary surfaces that the matrices $M_B f(a, d)$ and $M_{B'} f(a, d')$ corresponding to different admissible basises $\{d, B\}$ and $\{d', B'\}$ are similar. Taking into account (4.2) we have arrived to the following assertion.

Theorem 4.2. *Let $f(x)$ be twice continuously differentiable around a and suppose that $\nabla f(a) \neq 0$. Let $\{d, B\}$ and $\{d', B'\}$ be admissible basises for $f(x)$ at a. Then the quasi-Hessian matrices $Q_B f(a, d)$ and $Q_{B'} f(a, d')$ are similar.*

Despite of the rather complicated structure of the quasi-Hessian matrix, the flexibility in choosing the basis $\{d, B\}$ makes it possible to calculate with it. The following results shows that for special admissible basis formula (4.1) becomes significantly simpler. The proof is left for the Readers.

Proposition 4.3. *Let the basis $\{d, B\}$ be admissible for $f(x)$ at a and suppose that $d = \nabla f(a)/\|\nabla f(a)\|$. Then*

$$Q_B f(a, d) = B^\top \nabla^2 f(a) B.$$

Proposition 4.4. *Let the basis $\{d, B\}$ be admissible for $f(x)$ at a and suppose that*

$$d^\top \nabla^2 f(a) b_i = 0 \text{ for all } b_i \in B. \tag{4.3}$$

Then

$$Q_B f(a, d) = B^\top \left[\nabla^2 f(a) + \frac{d^\top \nabla^2 f(a) d}{(\nabla f(a)^\top d)^2} \nabla f(a)^\top \nabla f(a) \right] B. \tag{4.4}$$

If the vector d happens to be an eigenvector of the Hessian $\nabla^2 f(a)$ then the hypothesis of the above proposition holds, consequently it seems to be worth investigating basises containing eigenvectors of the Hessian $\nabla^2 f(a)$.

Let $s_1, s_2, \ldots, s_n \in \mathbb{R}^n$ be orthonormal eigenvectors of the symmetric matrix $\nabla^2 f(a)$ and let $h_1, h_2, \ldots, h_n \in \mathbb{R}$ be the corresponding eigenvalues. Since $\nabla f(a) \neq 0$, therefore there exists at least one index i such that $\nabla f(a)^\top s_i \neq 0$. Put

$$B_i = [s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n].$$

We focus our attention in the sequel to the quasi-Hessian associated with the admissible basis $\{s_i, B_i\}$ and which will be denoted from now on by a simpler symbol $Q_i f(a)$.

Let us introduce the following matrices:

$$S = [s_1, s_2, \ldots, s_n] \text{ and } E_k = [e_1, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n],$$

where e_j stands for the j-th unit vector of $\mathbb{R}^n$. Formula (4.4) gives the following in this case:

$$Q_k f(a) = E_k^\top S^\top \left[\nabla^2 f(a) + \frac{h_k}{(\nabla f(a)^\top s_k)^2} \nabla f(a) \nabla f(a)^\top\right] S E_k. \tag{4.5}$$

5. Conditions ensuring the positive semidefiniteness of the quasi-Hessian

This section is devoted to analysing connections between the positive semidefiniteness of the quasi-Hessian matrices of $f(x)$ at a and certain properties of the gradient $\nabla f(a)$ and Hessian $\nabla^2 f(a)$ of $f(x)$ at a.

Let us begin with listing some conditions and show them to be equivalent.

(C1) For any admissible basis $\{d, B\}$ the quasi-Hessian $Q_B f(a, d)$ is positive semidefinite.

(C2) Either $\nabla^2 f(a)$ is positive semidefinite, or, assuming that h_k is a negative eigenvalue of it, the following conditions hold:

$$\nabla f(a)^\top s_k \neq 0, \tag{5.1a}$$

$$h_j \geq -h_k \frac{(\nabla f(a) s_j)^2}{(\nabla f(a) s_k)^2}, \qquad j \neq k, \tag{5.1b}$$

$$1 + \frac{h_k}{(\nabla f(a) s_k)^2} \left[\sum_{\substack{j \neq k \\ h_j \neq 0}} \frac{(\nabla f(a) s_j)^2}{h_j}\right] \geq 0. \tag{5.1c}$$

(C3) A Bowman-Gleason-type condition [3,6]: Either $\nabla^2 f(a)$ is positive semidefinite, or, assuming that h_k is a negative eigenvalue of it, the following conditions hold:

$$\nabla f(a)^\top s_k \neq 0, \tag{5.2a}$$

$$\begin{gathered} h_j \geq 0 \text{ for all } j \neq k \text{ and} \\ h_j = 0 \text{ implies } \nabla f(a)^\top s_j = 0, \end{gathered} \tag{5.2b}$$

$$\text{(5.2c)} \qquad \sum_{h_j \neq 0} \frac{(\nabla f(a) s_j)^2}{h_j} \leq 0.$$

(C4) The Crouzeix-Ferland condition [5]: Either $\nabla^2 f(a)$ is positive semidefinite, or it has one simple negative eigenvalue and there exists a vector $r \in \mathbb{R}^n$ such that

$$\text{(5.3)} \qquad \nabla^2 f(a) r = \nabla f(a) \text{ and } \nabla f(a)^\top r \leq 0.$$

(C5) The Katzner condition [9]:

$$\text{(5.4)} \qquad p \in \mathbb{R}^n, \;\; \nabla f(a)^\top p = 0 \text{ implies } p^\top \nabla^2 f(a) p \geq 0.$$

(C6) The Ferland condition [7]: The bordered matrix

$$\begin{bmatrix} 0 & \nabla f(a)^\top \\ \nabla f(a) & \nabla^2 f(a) \end{bmatrix}$$

has one simple negative eigenvalue.

Theorem 5.1. *Let $s_1, s_2, \ldots, s_n \in \mathbb{R}^n$ be orthonormal eigenvectors of the Hessian $\nabla^2 f(a)$, let $h_1, h_2, \ldots, h_n \in \mathbb{R}$ be the corresponding eigenvalues and suppose that $\nabla f(a) \neq 0$. Let $\{d, B\}$ be an admissible basis for $f(x)$ at a. Then conditions* (C1)-(C6) *are equivalent.*

Proof. (C1)⇒(C2): Assume that $Q_B f(a,d)$ is positive semidefinite. Let us consider an index i such that $\nabla f(a)^\top s_i \neq 0$. It follows that the quasi-Hessian $Q_i f(a)$ does exist. According to Theorem 4.2 the matrices $Q_B f(a,d)$ and $Q_i f(a)$ are similar, consequently the matrix $Q_i f(a)$ is positive semidefinite, too. By formula (4.5) it may be the case that the Hessian $\nabla^2 f(a)$ is positive semidefinite. Now let us suppose that it is not the case. Then the matrix $\nabla^2 f(a)$ has at least one negative eigenvalue. Let h_k denote one of them.

To prove (5.1a) assume for contradiction that $\nabla f(a)^\top s_k = 0$. Put $u = E_i^\top e_k \in \mathbb{R}^{n-1}$. Then formula (4.5) gives the following:

$$u^\top Q_i f(a) u = s_k^\top \nabla^2 f(a) s_k = h_k < 0.$$

This inequality contradicts the positive semidefiniteness of $Q_i f(a)$ and this contradiction proves the necessity of condition (5.1a).

To prove (5.1b) let us consider the quasi-Hessian $Q_k f(a)$ which is, by Theorem 4.2 positive semidefinite, as well. Let j be an arbitrary index different from k. Put $u = E_k^\top e_j \in \mathbb{R}^{n-1}$. Applying (4.5) we have

$$0 \le u^\top Q_k f(a) u = h_j + h_k \frac{\left(\nabla f(a)^\top s_j\right)^2}{\left(\nabla f(a)^\top s_k\right)^2},$$

which is equivalent to (5.1b).

To prove (5.1c) let us consider the vector $p \in \mathbb{R}^n$ whose coordinates p_j, $j = 1, 2, \ldots, n$ are defined in the following manner:

$$p_j = \begin{cases} \dfrac{\nabla f(a)^\top s_j}{h_j} & \text{if } h_j \ne 0, \\ \text{arbitrary} & \text{otherwise.} \end{cases} \tag{5.5}$$

From (5.1b) one can easily deduce (5.2b) and taking into account this condition we have got the following result:

$$h_j p_j = \nabla f(a)^\top s_j, \qquad j = 1, 2, \ldots, n. \tag{5.6}$$

It follows that

$$S^\top \nabla^2 f(a) S p = S^\top \nabla f(a). \tag{5.7}$$

Put $u = E_k^\top p \in \mathbb{R}^{n-1}$ and $z = E_k u = E_k E_k^\top p \in \mathbb{R}^n$. It is not difficult to notice that the vectors p and z differ from each other in their k-th coordinate only and we have $z_k = 0$. Applying (4.5) we have

$$0 \le u^\top Q_k f(a) u = z^\top S^\top \nabla^2 f(a) S z + h_k \frac{\left(\nabla f(a)^\top S z\right)^2}{\left(\nabla f(a)^\top s_k\right)^2}.$$

This inequality along with (5.7) yields

$$z^\top S^\top \nabla^2 f(a) S z = \nabla f(a)^\top S z \ge -h_k \frac{\left(\nabla f(a)^\top S z\right)^2}{\left(\nabla f(a)^\top s_k\right)^2} \ge 0,$$

from which we obtain

$$1 + \frac{h_k}{\left(\nabla f(a)^\top s_k\right)^2} \nabla f(a)^\top S z \ge 0.$$

This last inequality is equivalent to (5.1c) which can be verified by using (5.5) and the definition of vector z.

(C2)$\Rightarrow$(C3): This implication is obvious. We have already mentioned that (5.1b) implies (5.2b). Moreover (5.1c) and (5.2c) are equivalent.

(C3)$\Rightarrow$(C1): Assume first that the Hessian $\nabla^2 f(a)$ is positive semidefinite. By hypothesis one can find an index i such that $Q_i f(a)$ exists. It follows from (4.5) that this matrix is positive semidefinite and thus, by Theorem 4.2, the quasi-Hessian $Q_B f(a,d)$ is positive semidefinite, too.

Now let us suppose that $\nabla^2 f(a)$ has at least one negative eigenvalue. Let h_k be one of them and suppose that conditions (5.2a,b,c) hold. Condition (5.2a) ensures the existence of the quasi-Hessian $Q_k f(a)$. To prove the positive semidefiniteness of $Q_B f(a,d)$, by Theorem 4.2, it is enough to prove the positive semidefiniteness of $Q_k f(a)$.

Let $u \in \mathbb{R}^{n-1}$ be arbitrary and let us introduce the following notation:

$$E_k u = (z_1, z_2, \ldots, z_n)^\top \in \mathbb{R}^n.$$

It is obvious that $z_k = 0$. By using (4.5) we obtain

$$u^\top Q_k f(a) u = \sum_{j\neq k} h_j z_j^2 + \frac{h_k}{(\nabla f(a)^\top s_k)^2}\left[\sum_{j\neq k}(\nabla f(a)^\top s_k) z_j\right]^2. \tag{5.8}$$

Taking into account (5.2b) and applying the Cauchy-Schwarz inequality, we have

$$\Big(\sum_{j\neq k}(\nabla f(a)^\top s_j) z_j\Big)^2 \leq \left[\sum_{\substack{j\neq k\\ h_j\neq 0}} \frac{(\nabla f(a)^\top s_j)^2}{h_j}\right]\Big(\sum_{j\neq k} h_j z_j^2\Big).$$

The above inequality along with (5.2c), which has been proved to be equivalent to (5.1c), and (5.8) yields the desired conclusion:

$$z^\top Q_k f(a) z \geq$$
$$\geq \Big(\sum_{j\neq k} h_j z_j^2\Big)\left[1 + \frac{h_k}{(\nabla f(a) s_k)^2}\sum_{\substack{j\neq k\\ h_j\neq 0}}\frac{(\nabla f(a) s_j)^2}{h_j}\right] \geq 0.$$

This inequality proves the positive semidefiniteness of the quasi-Hessian $Q_k f(a)$.

(C3)$\Rightarrow$(C4): Assume that condition (C3) is fulfilled. It is enough to examine the case when the Hessian $\nabla^2 f(a)$ is not positive semidefinite. Let us consider a vector p defined by (5.5) and therefore satisfying (5.7). Put $r = Sp$. It follows from (5.7) that

$$\nabla^2 f(a) r = \nabla f(a). \tag{5.9}$$

Keeping (5.2c) in mind we have

$$\nabla f(a)^\top r = \nabla f(a)^\top Sp = \sum_{h_j \neq 0} \frac{(\nabla f(a) s_j)^2}{h_j} \leq 0. \tag{5.10}$$

It is not difficult to see that in our case every solution r of equation (5.9) satisfies inequality (5.10).

(C4)⇒(C3): Assume that condition (C4) is fulfilled. It is enough to examine the case when $\nabla^2 f(a)$ is not positive semidefinite.

Let $r \in \mathbb{R}^n$ be arbitrary such that $\nabla^2 f(a) r = \nabla f(a)$ and $\nabla f(a)^\top r \leq 0$. Then for $p = S^\top r$ equations (5.6) and (5.7) hold. By hypothesis the eigenvalue h_k is negative and $h_j \geq 0$ for all $j \neq k$. From (5.6) we easily obtain (5.2b).

On the other hand

$$\sum_{h_j \neq 0} \frac{(\nabla f(a) s_j)^2}{h_j} = \nabla f(a)^\top Sp = \nabla f(a)^\top r \leq 0$$

and thus (5.2c) holds, too.

To prove (5.2a) assume for contradiction that $\nabla f(a)^\top s_k = 0$. Since in this case

$$\nabla f(a)^\top r = \sum_{h_j > 0} \frac{(\nabla f(a) s_j)^2}{h_j} > 0,$$

we have a contradiction to the hypothesis $\nabla f(a)^\top r \leq 0$.

(C1)⇒(C5): Assume that condition (C1) is fulfilled. Let $p \in \mathbb{R}^n$ be arbitrary and suppose that $\nabla f(a)^\top p = 0$. We may assume, by Theorem 4.2, without loss of generality that $d = \nabla f(a)/\|\nabla f(a)\|$. By Proposition 4.4 we have that

$$0 \leq u^\top Q_B f(a,d) u = p^\top \nabla^2 f(a) p,$$

where $u = B^\top p \in \mathbb{R}^{n-1}$.

(C5)⇒(C1): Assume that condition (C5) holds. Now we prove that the matrix $Q_B f(a,d)$ is positive semidefinite. As in the previous part of the present proof we may assume that $d = \nabla f(a)/\|\nabla f(a)\|$. Let $u \in \mathbb{R}^{n-1}$ be arbitrary. Then we have $\nabla f(a)^\top Bu = 0$. Taking into account condition (C5) and Proposition 4.4 we obtain

$$0 \leq u^\top B^\top \nabla^2 f(a) Bu = u^\top Q_B f(a,d) u.$$

(C5)⇔(C6): This equivalence was shown by Ferland (cf. [7], Theorem 2.9). ∎

Remark 5.2. For comparing condition (C6) with the other ones our quasi-Hessian approach is not suitable. The equivalence of (C5) and (C4) was shown by Crouzeix and Ferland in [5, Theorem 4.1].

6. Conditions ensuring the positive definiteness of the quasi-Hessian

Theorem 6.1. *Let $s_1, s_2, \ldots, s_n \in \mathbb{R}^n$ be orthonormal eigenvectors of the Hessian $\nabla^2 f(a)$, let $h_1, h_2, \ldots, h_n \in \mathbb{R}$ be the corresponding eigenvalues and suppose that $\nabla f(a) \neq 0$. Let $\{d, B\}$ be an admissible basis for $f(x)$ at a. Then conditions* (D1)-(D5) *are equivalent.*

(D1) *The quasi-Hessian $Q_B f(a, d)$ is positive definite.*

(D2) *Either $\nabla^2 f(a)$ is positive definite, or, assuming that h_k is a nonpositive eigenvalue of it, the following conditions hold:*

$$\nabla f(a)^\top s_k \neq 0; \quad h_j > -h_k \frac{(\nabla f(a)^\top s_j)^2}{(\nabla f(a)^\top s_k)^2}, \quad j = 1, \ldots, n \text{ and } j \neq k;$$

and in case of $h_k < 0$, it is additionally

$$\sum_{j=1}^{n} \frac{(\nabla f(a) s_j)^2}{h_j} < 0.$$

(D3) *Either $\nabla^2 f(a)$ is positive definite, or it has one simple nonpositive eigenvalue h_k such that $\nabla f(a)^\top s_k \neq 0$, moreover if $h_k < 0$, then for the unique solution $r \in \mathbb{R}^n$ of equation $\nabla^2 f(a) r = \nabla f(a)$ we have $\nabla f(a)^\top r < 0$.*

(D4) *$p \in \mathbb{R}^n$, $p \neq 0$, $\nabla f(a)^\top p = 0$ implies $p^\top \nabla^2 f(a) p > 0$.*

(D5) *The bordered matrix*

$$\begin{bmatrix} 0 & \nabla f(a)^\top \\ \nabla f(a) & \nabla^2 f(a) \end{bmatrix}$$

is nonsingular and has one simple negative eigenvalue.

(D6) *There exists a positive number r_0 such that for all $r \geq r_0$ the augmented Hessian*

$$H(r) = \nabla^2 f(a) + r \nabla f(a) \nabla f(a)^\top$$

is positive definite.

Proof. The equivalence of conditions (D1),(D2),(D3) and (D4) can be proved using the same technique developed in the proof of the previous theorem, however, the examination of conditions (D5) and (D6) requires quite other ideas.

The equivalence of (D4) and (D6) is the well known Finsler Lemma [8]. The equivalence of (D4) and (D5) was shown by I. Zang [17, Proposition 7]. ∎

Remark 6.2. The implication (D1)⇒(D4) was essentially proved by T. Rapcsák in [15]. It should be noted that the positive semidefiniteness of the augmented Hessian $H(a)$ is not equivalent to the positive semidefiniteness of the quasi-Hessian $Q_B f(a,d)$ (combine [17, Theorem 2] and Theorem 5.1 of this paper).

7. Characterization of pseudoconvexity via the quasi-Hessian

In this chapter we combine the results of the previous parts in order to characterize pseudoconvexity or strict pseudoconvexity. First we prove a preparatory lemma.

Lemma 7.1. *Let $f(x)$ be defined and twice continuously differentiable around $a \in \mathbb{R}^n$, where $\nabla f(a) \neq 0$. Let the basis $\{d, B\}$ be admissible for $f(x)$ at a. Then*

(i) *if there exists a neighborhood G of a such that the quasi-Hessian $Q_B f(x,d)$ is positive semidefinite for all $x \in G$ then $f(x)$ is locally pseudoconvex at a;*

(ii) *if $Q_B f(a,d)$ is positive definite then $f(x)$ is locally strictly pseudoconvex at a.*

Proof. (i) Assume that $Q_B f(x,d)$ is positive semidefinite for all $x \in G$. Without loss of generality we may assume that neighborhood G is just the one figuring in the implicit function theorem. From this it can be deduced that for any $u \in N$ the Hessian $\nabla^2 H(u)$ is positive semidefinite, too and thus $H(u)$ is convex on N. Applying Theorem 3.4 we obtain the thesis.
(ii) Since $Q_B f(a,d)$ is positive definite, therefore the Hessian $\nabla^2 H(u_0)$ is positive definite, too. It follows that the corrected implicit function $H(u)$ is strictly locally convex at u_0. Applying Theorem 3.4 we get the thesis. ∎

It is not difficult now to extend the previous result for global (strict) pseudoconvexity.

Theorem 7.2. *Let $f(x)$ be twice continuously differentiable on the open convex set $C \subset \mathbb{R}^n$. If the following conditions are satisfied, then $f(x)$ is (strictly) pseudoconvex on C:*

(i) *for all $z \in C$ $\nabla f(z) = 0$ implies that $f(x)$ attains a (strict) local minima at z,*

(ii) *for all* $a \in C$ *such that* $\nabla f(a) \neq 0$ *one can find a basis* $\{d, B\}$, *admissible for* $f(x)$ *at* a, *for which the quasi-Hessian* $Q_B f(a, d)$ *is positive semidefinite (positive definite).*

For pseudoconvexity the appropriate conditions of this theorem are necessary, too.

Proof. The above conditions ensure that $f(x)$ is locally (strictly) pseudoconvex at each point of the open convex set C. Applying Proposition 2.2 the proof is completed.

■

Remark 7.3. It is worth noting that the quasi-Hessian $Q_B f(a, d)$ characterizes (strict) pseudoconvexity as the Hessian $\nabla^2 f(x)$ does (strict) convexity.

References

[1] M. Avriel, W.E. Diewert, S. Schaible and I. Zang, *Generalized Concavity*, Plenum Press, New York, 1988.

[2] M. Avriel and S. Schaible, Second order characterization of pseudoconvex functions, *Mathematical Programming*, **14** (1978), 170–185.

[3] V. J. Bowman and T. C. Gleason, *A note on second order conditions for quasiconvexity, Working Paper Nr.44.*, Carnegie-Mellon University, Pittsburgh, Pa., 1975.

[4] A. Cambini, E. Castagnoli, L. Martein, P. Mazzoleni and S. Schaible (Eds.), *Generalized Convexity and Fractional Programming with Economic Applications*, Springer-Verlag, Berlin – Heidelberg – New York, 1990.

[5] J.-P. Crouzeix and J. Ferland, Criteria for quasiconvexity and pseudoconvexity: relationships and comparisons, *Mathematical Programming*, **23** (1982), 193–205.

[6] J. A. Ferland, Matrix criteria for pseudoconvex functions in the class C^2, *Linear Algebra and its Applications*, **21** (1978), 47–57.

[7] J. A. Ferland, Matrix-theoretic criteria for the quasiconvexity of twice continuously differentiable functions, *Linear Algebra and its Applications*, **38** (1981), 51–63.

[8] P. Finsler, Über das Vorkommen definiter und semidefiniter Formen in scharen quadratischen Formen, *Comment. Math. Helv.*, **9** (1937), 188–192.

[9] D. W. Katzner, *Static Demand Theory*, The MacMillan Company, New York, 1970.

[10] S. Komlósi, *Contribution to the theory and methods of quasiconvex programming*, Ph.D. Dissertation, unpublished manuscript, Baku, 1984.

[11] S. Komlósi, Generalized convexity of quadratic functions, *Izv. Vyssh. Uchebn. Zaved, Mat.*, **9** (1984), 38–43.

[12] S. Komlósi, First and second order characterizations of pseudolinear functions, *European Journal of Operational Research*, **66** (1993), to appear.

[13] O. L. Mangasarian, Pseudoconvex functions, *SIAM Journal on Control*, **3** (1965), 281–290.

[14] B. Martos, *Nonlinear programming: theory and methods*, North Holland, Amsterdam, 1975.

[15] T. Rapcsák, The SUMT method for solving nonconvex programming problems, *Alkalmazott Matematikai Lapok*, **2** (1976), 427–437 (in Hungarian).

[16] T. Rapcsák, On the second order sufficiency conditions, *Journal of Information and Optimization Sciences*, **4** (1983), 183–191.

[17] I. Zang, Concavifiability of C^2 functions: A unified exposition, *Generalized Concavity in Optimization and Economics,* Academic Press, New York, 1981.

S. Komlósi, Faculty of Economics, Janus Pannonius University, H-7601 Pécs, P O Box 219