Journal of Mathematics Research; Vol. 10, No. 3; June 2018
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

A Classroom Note on Twice Continuously Differentiable Strictly Convex and Strongly Quasiconvex Functions

Giorgio Giorgi[1]

[1] Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, Italy.

Correspondence: Giorgio Giorgi, Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, Italy.

Received: February 22, 2018 Accepted: March 8, 2018 Online Published: March 27, 2018

doi:10.5539/jmr.v10n3p42 URL: https://doi.org/10.5539/jmr.v10n3p42

Abstract

We provide some remarks and clarifications for twice continuously differentiable strictly convex and strongly quasiconvex functions. Characterizations of these classes and their relationships with other classes of generalized convex functions are also examined.

Keywords: convexity, strict convexity, strong quasiconvexity, quasiconvexity, pseudoconvexity

Mathematics Subject Classification (2010): 26A51, 90C25.

1. Introduction

It is well-known that if $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a twice continuously differentiable function on the open convex set $X \subseteq \mathbb{R}^n$, then the following results hold:

a) The function f is convex on X if and only if its Hessian matrix $Hf(x)$ is positive semidefinite at every point $x \in X$.

b) If $Hf(x)$ is positive definite at every point $x \in X$, then f is strictly convex on X.

Condition *b*) is therefore only sufficient for the strict convexity of a twice continuously differentiable function on an open convex set. Consider, e. g., the function $f(x_1, x_2) = (x_1)^4 + (x_2)^4$, which is obviously strictly convex on the whole $\mathbb{R}^2$, but for which $Hf(0, 0)$ is the zero matrix. We have to note that the above results *a*) and *b*) can be better precised in the following results. See, e. g., Bertsekas (2009), Hiriart-Urruty and Lemaréchal (1993).

Theorem 1. Let $X \subseteq \mathbb{R}^n$ be a nonempty convex set and let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be twice continuously differentiable over an open set that contains X.

i) If $Hf(x)$ is positive semidefinite for all $x \in X$, then f is convex on X.

ii) If $Hf(x)$ is positive definite for all $x \in X$, then f is strictly convex on X.

iii) If X is *open* and f is convex on X, then $Hf(x)$ is positive semidefinite for all $x \in X$.

Remark 1. Theorem 4.28 in Güler (2010) is not entirely correct, on the grounds of Theorem 1. Indeed, if f is convex over a convex set that is not open, $Hf(x)$ may not be positive semidefinite at any point of X : take for example $X = \{(x_1, 0) : x_1 \in \mathbb{R}\}$ and $f(x_1, x_2) = (x_1)^2 - (x_2)^2$. However, it can be shown that the conclusions of Theorem 1 also holds if X has a nonempty interior instead of being open (i. e. X is a *solid* convex set).

Remark 2. Theorem 1 holds also under the assumption that f is twice Fréchet differentiable. A further weakening by means of twice Gâteaux differentiability is made by Borwein and Vanderwerff (2010).

This paper is organized as follows.

In Section 2 we shall make some comments on the above classical results *a*) and *b*), in order to get a simple characterization of twice continuously differentiable strictly convex functions of several variables on an open convex set $X \subseteq \mathbb{R}^n$ and of twice differentiable strictly convex functions of one variable on an open interval $I \subseteq \mathbb{R}$.

In Section 3 we shall make some comments on twice continuously differentiable strongly quasiconvex functions on an open convex set $X \subseteq \mathbb{R}^n$.

2. Twice Continuously Differentiable Strictly Convex Functions

For the reader's convenience we recall the basic characterizations of strictly convex functions.

Definition 1. Let $f : X \longrightarrow \mathbb{R}$ be defined on the convex set $X \subseteq \mathbb{R}^n$. Then f is strictly convex on X if

$$\forall x^1, x^2 \in X,\ x^1 \neq x^2,\ \forall \lambda \in (0,1) : f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda) f(x^2).$$

Theorem 2. Let $f : X \longrightarrow \mathbb{R}$ be defined on the convex set $X \subseteq \mathbb{R}^n$. Then f is strictly convex on X if and only if:

(i) For each $x \in X$, for each $y \in \mathbb{R}^n$, $y \neq 0$, the function

$$\varphi_{x,y}(t) = f(x + ty)$$

is strictly convex on the interval

$$T_{x,y} = \{t \in \mathbb{R} : x + ty \in X\}.$$

Equivalently, if and only if:

(ii) For each $x^1, x^2 \in X$, $x^1 \neq x^2$, the function

$$\psi_{x^1,x^2}(\lambda) = f(\lambda x^1 + (1-\lambda)x^2)$$

is strictly convex on $(0, 1)$.

Let $X \subseteq \mathbb{R}^n$ be open and convex and let f be differentiable on X. Then f is strictly convex on X if and only if:

(iii) For each $x^1, x^2 \in X$, $x^1 \neq x^2$,

$$f(x^1) - f(x^2) > \nabla f(x^2)(x^1 - x^2).$$

Equivalently, if and only if:

(iv) For each $x^1, x^2 \in X$, $x^1 \neq x^2$,

$$f(x^1) - f(x^2) < \nabla f(x^1)(x^1 - x^2).$$

Equivalently, if and only if:

(v) For each $x^1, x^2 \in X$, $x^1 \neq x^2$,

$$\left[\nabla f(x^1) - \nabla f(x^2)\right](x^1 - x^2) > 0.$$

For what concerns convex functions, their characterizations corresponding to (i) and (ii) of Theorem 2 are, respectively: f is convex on the convex set $X \subseteq \mathbb{R}^n$ if and only if:

(i)′ For each $x \in X$, for each $y \in \mathbb{R}^n$ the function

$$\varphi_{x,y}(t) = f(x + ty)$$

is convex on the interval

$$T_{x,y} = \{t \in \mathbb{R} : x + ty \in X\}.$$

Equivalently, if and only if:

(ii)′ For each $x^1, x^2 \in X$, the function

$$\psi_{x^1,x^2}(\lambda) = f(\lambda x^1 + (1-\lambda)x^2)$$

is convex on $[0, 1]$.

Characterizations (i) and (ii) of Theorem 2 and the above characterizations (i)′ and (ii)′ show that the concept of convex and strictly convex functions is genuinely unidimensional: f is convex (strictly convex) on a convex set $X \subseteq \mathbb{R}^n$ if and only if the restriction of f to each line segment contained in X is a convex (strictly convex) function. It is therefore useful to recall the main properties concerning convex and strictly convex functions of one single variable.

Theorem 3. Let $\varphi : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, with I open interval.

1. If φ is differentiable on I, the function φ is convex on I if and only if its derivative φ' is increasing on I.

2. If φ is differentiable on I, the function φ is strictly convex on I if and only if its derivative φ' is strictly increasing on I.

It is well-known that:

a) A differentiable function $\varphi : I \longrightarrow \mathbb{R}$ is increasing (respectively: decreasing) on the open interval $I \subseteq \mathbb{R}$ if and only if $\varphi'(x) \geqq 0$ (respectively: $\varphi'(x) \leqq 0$), $\forall x \in I$.

b) A differentiable function $\varphi : I \longrightarrow \mathbb{R}$ is strictly increasing (respectively: strictly decreasing) on the open interval $I \subseteq \mathbb{R}$ if $\varphi'(x) > 0$ (respectively: $\varphi'(x) < 0$), $\forall x \in I$.

The following result is less known; for the reader's convenience we give a proof.

Theorem 4. Let $f : I \longrightarrow \mathbb{R}$ be differentiable on the open interval $I \subseteq \mathbb{R}$. Then f is strictly increasing (resp.: strictly decreasing) on I *if and only if* $f'(x) \geqq 0$ (resp.:$f'(x) \leqq 0$), $\forall x \in I$, and there exists no subinterval of I where $f'(x) = 0$ in all points of the said subinterval. In other words, it must hold $f'(x) \geqq 0$ (resp.: $f'(x) \leqq 0$), $\forall x \in I$ and the set

$$\begin{array}{rl} & A = \{x : f'(x) > 0\} \\ \text{(resp.:} & A = \{x : f'(x) < 0\}\,) \end{array}$$

is dense in I.

Proof.

i) The condition is necessary. It is well-known that it must hold $f'(x) \geqq 0$ (resp.:$f'(x) \leqq 0$), $\forall x \in I$. If there would exist an interval $J \subseteq I$ such that, for each $x \in J$, we have $f'(x) = 0$, the restricftion of f to J would be constant on J, which is absurd.

ii) The condition is sufficient. Let us consider any pair of distinct points of I, say x' and x'' with $x' < x''$. Thanks to the *mean value theorem* (*or Lagrange mean value theorem*), applied to the restriction of f to $[x', x'']$, we have

$$f(x'') - f(x') = f'(\xi)(x'' - x'),$$

where ξ is a suitable point of (x', x''). Therefore, taking the assumptions into account:

$$\begin{array}{rl} & x' < x'' \Longrightarrow f(x') \leqq f(x'') \\ \text{(resp.:} & x' < x'' \Longrightarrow f(x') \geqq f(x'')), \end{array}$$

i. e. f is increasing (resp.: decreasing) on I. From this, obviously we have also:

$$\begin{array}{rl} & x' < x < x'' \Longrightarrow f(x') \leqq f(x) \leqq f(x'') \\ \text{(resp.:} & x' < x < x'' \Longrightarrow f(x') \geqq f(x) \geqq f(x'')). \end{array}$$

Hence, if $f(x') = f(x'')$, the restriction of f to $[x', x'']$ would be constant on $[x', x'']$ and this would imply that every $x \in [x', x'']$, it holds $f'(x) = 0$, against the assumption that no subinterval of I exists such that the derivative of f is identically zero on the same subinterval. It results therefore

$$\begin{array}{rl} & x'x'' \in I,\ \ x' < x'' \Longrightarrow f(x') < f(x'') \\ \text{(resp.:} & x'x'' \in I,\ \ x' < x'' \Longrightarrow f(x') > f(x'')), \end{array}$$

i. e. the thesis. □

We can therefore formulate the following result.

Theorem 5. Let $f : I \longrightarrow \mathbb{R}$ be twice differentiable on the open interval $I \subseteq \mathbb{R}$. Then f is strictly convex on I if and only if

$$\begin{cases} f''(x) \geqq 0, & \forall x \in I \\ A = \{x : f''(x) > 0\} & \text{is dense in } I. \end{cases}$$

Example 1. The function

$$f(x) = x^2 - \sin^2 x$$

is strictly convex on $\mathbb{R}$. We have indeed

$$f''(x) = 2(1 - \cos 2x) \geqq 0, \quad \forall x \in \mathbb{R},$$

but the infinitely many points where $f''(x) = 0$ are all isolated points.

The previous results provide a simple way to obtain the characterization of convex functions of several variables, which are twice continuously differentiable on some open convex set $X \subseteq \mathbb{R}^n$. First we restate a part of Theorem 1 in its "classical" version.

Theorem 1′. Let $f : X \longrightarrow \mathbb{R}^n$ be twice continuously differentiable on a nonempty open convex set $X \subseteq \mathbb{R}^n$. Then f is convex on X if and only if its Hessian matrix $Hf(x)$ is positive semidefinite at each point $x \in X$.

Proof. Let $x^0 \in X$, $v \in \mathbb{R}^n$, $v \neq 0$ and consider the restriction

$$\varphi(t) = f(x^0 + tv)$$

such that $x^0 + tv \in X$. It is sufficient to note that

$$\varphi''(t) = v^\top Hf(x^0 + tv)v.$$

□

We recall that a positive (negative) semidefinite quadratic form $Q(x) = x^\top Ax$ (A symmetric) is positive (negative) definite if and only if A is a nonsingular matrix (see, e. g. Hestenes (1966), Theorem 6.3). In other words, for positive semidefinite matrices we may have $x^\top Ax = 0$ even if $x \neq 0$, but in this case we neverthless have $Ax = 0$, and being $x \neq 0$, it must hold $|A| = 0$. Indeed, if $y \in \mathbb{R}^n$ and $t \in \mathbb{R}$ are arbitrary, then

$$(x + ty)^\top A(x + ty) = x^\top Ax + 2ty^\top Ax + t^2 y^\top Ay \geqq 0.$$

If $x^\top Ax = 0$, $x \neq 0$, this implies $y^\top Ax = 0$. Now let $y = Ax$, so $(Ax)^\top Ax = 0$, whence finally $Ax = 0$. Therefore, being $x \neq 0$, it must hold $|A| = 0$.

Contrary to the unidimensional case, the condition:

- " $Hf(x)$ is positive semidefinite for every $x \in X$ and $|Hf(x)|$ is not identically zero on any segment belonging to X"

is only sufficient for the strict convexity of the twice continuously differentiable function f on the open convex set $X \subseteq \mathbb{R}^n$ (see Fenchel (1953), Ortega and Rheinboldt (1970)).

Example 2.

i) See also Bernstein and Toupin (1962). Suppose $n = 2$, $X = \{(x_1, x_2) : x_2 < 0\}$ and

$$f(x) = (x_1)^2(1 + e^{x_2}).$$

The determinant of its Hessian matrix is

$$|Hf(x)| = 2(x_1)^2 e^{x_2}(1 - e^{x_2}),$$

which is positive throughout X except on the line $x_1 = 0$ where it vanishes. It is seen that f is convex on X but not strictly convex.

ii) Consider $f(x, y) = x^4 + y^4$, strictly convex on $\mathbb{R}^2$. Its Hessian matrix is

$$Hf(x) = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}$$

and $|Hf(x)| = 0$ on the two axes.

Necessary and sufficient conditions for the strict convexity of twice continuously differentiable functions have been established by Bernstein and Toupin (1962) and by Diewert and others (1981). Following these last authors, we "translate" the conditions of Bernstein and Toupin with a more convenient notation and statement.

Theorem 6. Let $f : X \longrightarrow \mathbb{R}$ be twice continuously differentiable on an open convex set $X \subseteq \mathbb{R}^n$. Then f is strictly convex on X if and only if:

$$\left\{x^0 \in X,\ v \in \mathbb{R}^n,\ v \neq 0,\ \bar{t} > 0,\ x^0 + \bar{t}v \in X\right\} \qquad \Longrightarrow$$

i) $v^\top Hf(x^0)v \geqq 0$ and
ii) the set $\left\{t \in \mathbb{R} : v^\top Hf(x^0 + tv)v > 0\right\}$ is dense in $[0, \bar{t}]$.

Remark 3. Diewert and others (1981) prove that another (equivalent) characterization of twice continuously differentiable strictly convex functions (on an open convex set $X \subseteq \mathbb{R}^n$) is:

$$x^0 \in X,\ v \in \mathbb{R}^n,\ v \neq 0 \Longrightarrow i)\ v^\top Hf(x^0)v > 0 \ \text{ or}$$

$$ii) \quad v^\top Hf(x^0)v = 0 \text{ and } h(t) \equiv f(x^0 + tv) + t\nabla f(x^0)v \text{ attains a strict local minimum at } t = 0.$$

The recent paper of Stein (2012) is concerned with the more general case of twice differentiable strictly convex functions defined on a convex set $X \subseteq \mathbb{R}^n$, not necessarily open.

Remark 4. Ginsberg (1973) defines the class of *strongly convex functions,* as those twice continuously differentiable functions on an open convex set $X \subseteq \mathbb{R}^n$ for which all leading principal minors (see Section 3, after Theorem 7) of their Hessian matrix $Hf(x)$ are positive, for each $x \in X$. Obviously, this one is a sufficient condition for f to be a strictly convex function on X (Theorem 1), however the above definition is misleading, as in the current literature on Convex Analysis (see, e. g., Diewert and others (1981), Avriel and others (1981), Rockafellar (1976), Vial (1982)) twice continuously differentiable strongly convex functions (on an open convex set $X \subseteq \mathbb{R}^n$) are characterized by the property:

- There exists $\alpha > 0$ such that

$$x \in X \Longrightarrow Hf(x) - \alpha I$$

is positive semidefinite (I is the identity matrix).

Remark 5. For the special case of *quadratic functions,* i. e. of the functions

$$\varphi(x) = x^\top Ax + c^\top x,$$

where A is a symmetric matrix of order n, some of the previous results can be stated as follows.

a) Martos (1975) proved for $\varphi(x)$ what already pointed out in Remark 1: $\varphi(x)$ is convex on any solid convex set $X \subseteq \mathbb{R}^n$ if and only if A is positive semidefinite.

b) The quadratic function $\varphi(x)$ is strictly convex on $\mathbb{R}^n$ if and only if A is positive definite (the same result holds with reference to a solid convex set $X \subseteq \mathbb{R}^n$). Indeed, if φ is strictly convex and $h \in \mathbb{R}^n$, $h \neq 0$, the first-order characterization of strictly convex functions gives

$$\varphi(x) + \nabla\varphi(x)h < \varphi(x + h) = \varphi(x) + \nabla\varphi(x)h + \frac{1}{2}h^\top Ah.$$

We have $h^\top Ah > 0$ for all $h \in \mathbb{R}^n$, $h \neq 0$, that is A is positive definite. Conversely, if A is positive definite and $h \neq 0$, then

$$\varphi(x + h) = \varphi(x) + \nabla\varphi(x)h + \frac{1}{2}h^\top Ah > \varphi(x) + \nabla\varphi(x)h,$$

which implies that φ is strictly convex.

c) Martos (1975) has proved that $\varphi(x)$ is quasiconvex (see Section 3) on $\mathbb{R}^n$ if and only if it is convex on $\mathbb{R}^n$. This shows that there is no reason to study quadratic functions that are quasiconvex, without being convex, on the whole $\mathbb{R}^n$.

3. Twice Continuously Differentiable Strongly Quasiconvex Functions

First we recall some basic definitions.

Definition 2. Let f be defined on a convex set $X \subseteq \mathbb{R}^n$; then f is said to be *quasiconvex* on X if

$$f(\lambda x^1 + (1 - \lambda)x^2) \leqq \max\left\{f(x^1), f(x^2)\right\}$$

for every $x^1, x^2 \in X$ and for every $\lambda \in [0, 1]$ or, equivalently,

$$f(x^1) \geqq f(x^2) \Longrightarrow f(x^1) \geqq f(\lambda x^1 + (1 - \lambda)x^2)$$

for every $x^1, x^2 \in X$ and for every $\lambda \in [0, 1]$.

Definition 3. A function f be defined on a convex set $X \subseteq \mathbb{R}^n$ is said to be *strictly quasiconvex* on X if

$$f(\lambda x^1 + (1 - \lambda)x^2) < \max\left\{f(x^1), f(x^2)\right\}$$

for every $x^1, x^2 \in X$, $x^1 \neq x^2$, and for every $\lambda \in (0,1)$ or, equivalently,

$$f(x^1) \geqq f(x^2) \Longrightarrow f(x^1) > f(\lambda x^1 + (1-\lambda)x^2)$$

for every $x^1, x^2 \in X$, $x^1 \neq x^2$, and for every $\lambda \in (0,1)$.

Definition 4. A function f be defined on a convex set $X \subseteq \mathbb{R}^n$ is said to be *semistrictly quasiconvex* on X if

$$f(\lambda x^1 + (1-\lambda)x^2) < \max\left\{f(x^1), f(x^2)\right\}$$

for every $x^1, x^2 \in X$, with $f(x^1) \neq f(x^2)$, and for every $\lambda \in (0,1)$ or, equivalently,

$$f(x^1) > f(x^2) \Longrightarrow f(x^1) > f(\lambda x^1 + (1-\lambda)x^2)$$

for every $x^1, x^2 \in X$, and for every $\lambda \in (0,1)$.

Under lower semicontinuity of f we have the following inclusion diagram.

$$\begin{array}{ccc}
\boxed{\text{strictly convex}} & \Longrightarrow & \boxed{\text{strictly quasiconvex}} \\
\Downarrow & & \Downarrow \\
\boxed{\text{convex}} & \Longrightarrow & \boxed{\text{semistrictly quasiconvex}} \\
 & & \Downarrow \\
 & & \boxed{\text{quasiconvex}}
\end{array}$$

In their pioneering paper on quasiconcave functions and quasiconcave programming, Arrow and Enthoven (1961) give the following necessary conditions for a twice continuously differentiable function to be quasiconvex on the open convex set $X \subseteq \mathbb{R}^n$.

Theorem 7. Let $f : X \longrightarrow \mathbb{R}$ be twice continuously differentiable on the open convex set $X \subseteq \mathbb{R}^n$; if f is quasiconvex on X, then

$$\Delta_r(x) \leqq 0, \quad \forall x \in X, \quad \forall r = 2, 3, ..., n, \tag{1}$$

where

$$\Delta_r(x) = \begin{vmatrix} 0 & \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_r} \\ \frac{\partial f}{\partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_r} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f}{\partial x_r} & \frac{\partial^2 f}{\partial x_r \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_r \partial x_r} \end{vmatrix}.$$

See also Avriel (1972), Kemp and Kimura (1978) and, for characterizations of twice continuously differentiable quasiconvex functions, Crouzeix (1980), Crouzeix and Ferland (1982), Diewert and others (1981). Needless to say, condition (1) is trivially satisfied for $r = 1$. In the same paper Arrow and Enthoven show that the following condition

$$\Delta_r(x) < 0, \quad \forall x \in X, \quad \forall r = 1, 2, 3, ..., n \tag{2}$$

is sufficient for the quasiconvexity of f on X. Indeed, this condition is even sufficient for strict quasiconvexity and more: see, e. g., Ginsberg (1973), Diewert and others (1981).

Relations (1) and (2) require a brief review on the main properties concerning quadratic forms subject to a system of homogeneous linear constraints. Given a (real) symmetric matrix A, of order n, and its associated quadratic form

$$Q(x) = x^\top Ax, \quad x \in \mathbb{R}^n, \tag{3}$$

we are interested in the sign of (3), but when $x \in S$, S being the set of non trivial solutions of the homogeneous system of linear equations

$$Bx = 0,$$

where B is a (real) (m,n) matrix, with $m < n$. This problem has been treated by several authors; see, e. g., Chabrillac and Crouzeix (1984), Debreu (1952), Farebrother (1977), Giorgi (2003, 2017), Murata (1977), Samuelson (1947). The following result may be considered a generalization of the *Sylvester criterion* for unconstrained quadratic forms, to the problem under examination. We recall that, given a square matrix A, of order n, its *leading principal minor (or successive principal minor or North-West principal minor) of order* k, $k = 1, ..., n$, is the determinant of the (k, k) matrix consisting of the first k rows and columns of A. Its *principal minor of order* k is the k-th order leading principal minor of $P^\top AP$, where P is some permutation matrix of order n. There are $n!/k!(n-k)!$ possible k-th order principal minors, whereas obviously there are in all n leading principal minors.

Theorem 8. Let us suppose that $rank(B) = m$ and, without loss of generality, that the first m columns of B are linearly independent. Then:

(i) $Q(x) = x^\top Ax > 0$ for all $x \neq 0$ such that $Bx = 0$ if and only if the leading principal minors of order $2m + p$, $p = 1, 2, ..., n - m$, of the bordered matrix

$$H = \begin{bmatrix} 0 & B \\ B^\top & A \end{bmatrix}$$

have the sign of $(-1)^m$. By denoting with Δ_k the k-th leading principal minor of H, we must therefore have

$$(-1)^m \Delta_k > 0, \quad k = 2m + 1, ..., m + n.$$

(ii) $Q(x) = x^\top Ax < 0$ for all $x \neq 0$ such that $Bx = 0$ if and only if the leading principal minors of H, of order $2m + p$, $p = 1, 2, ..., n - m$, have the sign of $(-1)^{m+p}$.

Remark 6. We note that in the previous theorem statement (ii) derives from statement (i) by observing that A is negative definite under constraints if and only if $-A$ is positive definite under the same constraints.

Remark 7. In some papers and books the following bordered matrix is considered:

$$\bar{H} = \begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix}.$$

In this case obviously the previous conditions (i) and (ii) have to be suitably modified, by making reference to the leading principal minors of $\bar{H}$.

Remark 8. In the case of only one constraint, of the type $bx = 0$, with $b \in \mathbb{R}^n$ and with $b_1 \neq 0$, the previous conditions (i) and (ii) of Theorem 8 become, respectively,

a) $Q(x)$ is positive definite on the set of the nontrivial solutions of $bx = 0$ if and only if

$$\Delta_3 = \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{vmatrix} < 0, ..., |H| < 0.$$

b) $Q(x)$ is negative definite on the set of the nontrivial solutions of $bx = 0$ if and only if

$$\Delta_3 = \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{vmatrix} > 0, \qquad \Delta_4 = \begin{vmatrix} 0 & b_1 & b_2 & b_3 \\ b_1 & a_{11} & a_{12} & a_{13} \\ b_2 & a_{21} & a_{22} & a_{23} \\ b_3 & a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \text{ etc.}$$

Remark 9. We have to note that the assumption that the rank of B ($rank(B) = m$) is given by the first m columns of B is essential to have necessary and suffcient conditions. In absence of this assumption, conditions (i) and (ii) of Theorem 8 are only sufficient conditions and they imply $rank(B) = m$. This is the case of the usual sufficient second order optimality conditions imposed on the Lagrangian function of optimization problems with equality constraints. Consider, e. g., the following simple example. It is obvious that the quadratic form

$$Q(x) = (x_1)^2 + (x_2)^2 - (x_3)^2$$

is positive definite on the constraint $x_3 = 0$ (hence here vector b is $b = [0, 0, 1]$). It results $\Delta_4 = |H| = -1$, however

$$\Delta_3 = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0.$$

Sound and correct proofs of necessary and sufficient conditions for the positive and negative semidefiniteness of $Q(x)$ subject to $Bx = 0$ are given, e. g., by Chabrillac and Crouzeix (1984) and by Debreu (1952).

Similarly to what holds for unconstrained quadratic forms, the relation between positive (negative) definiteness of a constrained quadratic form and positive (negative) semidefiniteness of the same constrained quadratic form, is expressed by the following result.

Theorem 9. Let the previous assumptions on A and B be verified. Let $x^\top Ax \geqq 0$ (resp. $x^\top Ax \leqq 0$) for all $x \neq 0$ such that $Bx = 0$. Then $x^\top Ax > 0$ (resp. $x^\top Ax < 0$) for all $x \neq 0$ such that $Bx = 0$ if and only if

$$|H| = \begin{vmatrix} 0 & B \\ B^\top & A \end{vmatrix} \neq 0. \tag{4}$$

Proof.

i) (Sufficiency). Suppose $x^\top Ax = 0$ for some $x \neq 0$. Then, since such x attains an optimum of the quadratic form under the linear constraints which are linearly independent, there exists a set of multipliers $\lambda = [\lambda_1, \lambda_2, ..., \lambda_m]^\top$ such that

$$\begin{cases} 2Ax + B^\top\lambda = 0 \\ Bx = 0. \end{cases}$$

These equations can be rewritten as

$$\begin{bmatrix} 0 & B \\ B^\top & A \end{bmatrix} \begin{bmatrix} \frac{1}{2}\lambda \\ x \end{bmatrix} = 0, \tag{5}$$

which implies, being $x \neq 0$, that the bordered matrix of (5) is singular. This condradicts assumption (4).

ii) (Necessity). Suppose that $|H| = 0$. Then, there exist vectors $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^m$ such that

$$(x^\top, \lambda^\top) \neq 0$$

and

$$\begin{bmatrix} 0 & B \\ B^\top & A \end{bmatrix} \begin{bmatrix} \lambda \\ x \end{bmatrix} = 0,$$

that is

$$Bx = 0 \tag{6}$$

$$B^\top\lambda + Ax = 0. \tag{7}$$

If $x = 0$, then $B^\top\lambda = 0$, which implies $\lambda = 0$, since the linear constraints are linearly independent. Therefore $x \neq 0$. From (6) and (7) we have

$$x^\top Ax = -x^\top B^\top\lambda = 0.$$

his contradicts the assumption that A is positive definite or negative definite under the constraints $Bx = 0$. □

Several authors have introduced the concept of *strongly quasiconvex* and *strongly quasiconcave* functions, called also *strongly pseudoconvex* and *strongly pseudoconcave* functions. See, e. g., Avriel and others (1981), Diewert and others (1981), Barten and Böhm (1982), Ginsberg (1973), Newman (1969), Leroux (1984).

Definition 5. A real differentiable function f defined on an open convex set $X \subseteq \mathbb{R}^n$ is called *strongly quasiconvex* (or *strongly pseudoconvex*) on X if f is strictly quasiconvex on X and in addiction:

$$x^0 \in X,\ v \neq 0,\ \nabla f(x^0)v = 0 \Longrightarrow \text{ there exist } \varepsilon > 0 \text{ and } \alpha > 0 \text{ such that}$$

$$(x^0 + \varepsilon v) \in X \text{ and } f(x^0 + tv) \geqq f(x^0) + \alpha t^2 \text{ for } t \in [0, \varepsilon]\,.$$

Obviously strong quasiconvexity implies strict quasiconvexity. Diewert and others (1981) prove the following result.

Theorem 10. Let f be twice continuously differentiable on the open convex set $X \subseteq \mathbb{R}^n$. Then f is strongly quasiconvex on X if and only if

$$v^\top Hf(x)v > 0 \text{ for all } x \in X \text{ and all } v \in \mathbb{R}^n,\ v \neq 0,\ \text{ such that } \nabla f(x)v = 0. \tag{8}$$

On the grounds of the previous theorem, we have that if f is twice continuously differentiable on the open convex set $X \subseteq \mathbb{R}^n$ and if $\nabla f(x) = 0$, for $x \in X$, then the thesis of Theorem 10 requires that $Hf(x)$ has to be positive definite on X; if

$\nabla f(x) \neq 0$, for $x \in X$, then $Hf(x)$ has to be positive definite on the subspace orthogonal to the gradient vector $\nabla f(x)$. In this last case Theorem 8 can be usefully applied. First we recall that $f : X \longrightarrow \mathbb{R}$, f differentiable on the open convex set $X \subseteq \mathbb{R}^n$, is *pseudoconvex* on X if:

$$x^1, x^2 \in X, \;\; f(x^2) < f(x^1) \Longrightarrow \nabla f(x^1)(x^2 - x^1) < 0.$$

Pseudoconvex functions are semistrictly quasiconvex and therefore also quasiconvex, but not strictly quasiconvex and not strongly quasiconvex. Under differentiability assumption we have the following relationships.

$$\begin{tikzcd}
\boxed{\text{strongly quasiconvex}} \arrow[r, Rightarrow] \arrow[dr, Rightarrow] & \boxed{\text{pseudoconvex}} \arrow[r, Rightarrow] & \boxed{\text{semistrictly quasiconvex}} \arrow[r, Rightarrow] & \boxed{\text{quasiconvex}} \\
 & \boxed{\text{strictly quasiconvex}} \arrow[ur, Rightarrow] & &
\end{tikzcd}$$

It turns out that (8) is a sufficient condition for a twice continuously differentiable function $f : X \longrightarrow \mathbb{R}$ to be pseudoconvex and also strictly quasiconvex on the open convex set $X \subseteq \mathbb{R}^n$.

Then we recall a classical result on pseudoconvexity and quasiconvexity of twice continuously differentiable functions; see, e. g., Crouzeix (1980), Crouzeix and Ferland (1982), Diewert and others (1981), Ferland (1972), Giorgi (2013), Otani (1983), Simon and Blume (1994).

Theorem 11. Let $f : X \longrightarrow \mathbb{R}$ be twice continuously differentiable on the open convex set $X \subseteq \mathbb{R}^n$ and let $\nabla f(x) \neq 0$ for all $x \in X$. Then f is quasiconvex on X and also pseudoconvex on X if and only if

$$v^\top Hf(x)v \geqq 0 \text{ for all } x \in X \text{ and all } v \in \mathbb{R}^n, \; v \neq 0, \text{ such that } \nabla f(x)v = 0. \tag{9}$$

In other words, under the assumptions of Theorem 11, the Hessian matrix of f has to be positive semidefinite on the subspace orthogonal to $\nabla f(x)$, for $x \in X$.

Remark 10. Crouzeix and Ferland (1982) prove that, if $\nabla f(x) \neq 0$ for all $x \in X$, then any one of the following conditions is equivalent to condition (9):

I) Either $Hf(x)$ is positive semidefinite for every $x \in X$, or $Hf(x)$ has one simple negative eigenvalue, for every $x \in X$, and there exists a vector $b \in \mathbb{R}^n$ such that $Hf(x)b = (\nabla f(x))^\top$ and $\nabla f(x)b \leqq 0$.

II) If we denote by $M(x)$ the following bordered matrix (of order $(n + 1)$):

$$M(x) = \begin{bmatrix} 0 & \nabla f(x) \\ (\nabla f(x))^\top & Hf(x) \end{bmatrix},$$

then $M(x)$ has one simple negative eigenvalue, for every $x \in X$.

III) All the principal minors of $M(x)$ (and not only its leading principal minors) are less than or equal to zero, for every $x \in X$.

Taking the previous theorems into account, the following results are at hand.

Theorem 12. Let $f : X \longrightarrow \mathbb{R}$ be twice continuously differentiable on the open convex set $X \subseteq \mathbb{R}^n$ and let $\frac{\partial f}{\partial x_1} \neq 0$ for all $x \in X$. Let us denote again by $M(x)$ the bordered matrix of Remark 10. Then, f is strongly quasiconvex on X if and only if the leading principal minors of $M(x)$ of order $3, 4, ..., n + 1$, are all negative for all $x \in X$.

Theorem 13. Let $f : X \longrightarrow \mathbb{R}$ be twice continuously differentiable on the open convex set $X \subseteq \mathbb{R}^n$ and let f be quasiconvex on X, with $\nabla f(x) \neq 0$ for all $x \in X$. Then f is strongly quasiconvex on X if and only if $|M(x)| \neq 0$, for all $x \in X$.

References

ARROW, K. J., & ENTHOVEN, A. C. (1961). Quasiconcave programming. *Econometrica, 29*, 779-800. https://doi.org/10.2307/1911819

AVRIEL, M. (1972). r-convex functions. *Math. Programming, 2*, 309-323. https://doi.org/10.1007/BF01584551

AVRIEL, M., DIEWERT, W. E., SCHAIBLE, S., & ZIEMBA, W. T. (1981). Introduction to concave and generalized concave functions. in S. Schaible and W. T. Ziemba (Eds.). *Generalized Concavity in Optimization and Economics, Academic Press*, New York, 21-50.

BARTEN, A. P., & BÖHM, V. (1982). Consumer theory. in K. J. Arrow and M. D. Intriligator (Eds.). *Handbook of Mathematical Economics, 2*, North Holland, Amsterdam, 381-429.

BERNSTEIN, B., & TOUPIN, R. A. (1962). Some properties of the Hessian matrix of a strictly convex function, *J. Reine und Angew. Math., 210*, 65-72.

BERTSEKAS, D. P. (2009). *Convex Optimization, Athena Scientific*. Belmont, Mass.

BORWEIN, J. M., & VANDERWERFF, J. D. (2010). *Convex Functions: Constructions*. Characterizations and Counterexamples, Cambridge Univ. Press, Cambridge. https://doi.org/10.1017/CBO9781139087322

CHABRILLAC, Y., & CROUZEIX, J.-P. (1984). Definiteness and semidefiniteness of quadratic forms revisited, *Linear Algebra and Its Applications, 63*, 283-292. https://doi.org/10.1016/0024-3795(84)90150-2

CROUZEIX, J.-P. (1980). A second order condition for quasiconvexity. *Math. Programming, 18*, 349-352. https://doi.org/10.1007/BF01588330

CROUZEIX, J.-P., & FERLAND, J. A. (1982). Criteria for quasi-convexity and pseudo-convexity: relationships and comparisons. *Math. Programming, 23*, 193-205. https://doi.org/10.1007/BF01583788

DEBREU, G. (1952). Definite and semidefinite quadratic forms. *Econometrica, 20*, 295-300. https://doi.org/10.2307/1907852

DIEWERT, W. E., AVRIEL, M., & ZANG, I. (1981). Nine kinds of quasiconcavity and concavity. *Journal of Economic Theory, 25*, 397-420. https://doi.org/10.1016/0022-0531(81)90039-9

FAREBROTHER, R. W. (1977). Necessary and sufficient conditions for a quadratic form to be positive whenever a set of homogeneous linear constraints is satisfied. *Linear Algebra and Its Applications, 16*, 39-42. https://doi.org/10.1016/0024-3795(77)90017-9

FENCHEL, W. (1953). Convex Cones, Sets, and Functions, mimeographed lecture notes from Princeton University, Department of Mathematics.

FERLAND, J. A. (1972). Mathematical programming problems with quasi-convex objective functions. *Math. Programming, 3*, 296-301. https://doi.org/10.1007/BF01585002

GINSBERG, W. (1973). Concavity and quasiconcavity in economics. *Journal of Economic Theory, 6*, 596-605. https://doi.org/10.1016/0022-0531(73)90080-X

GIORGI, G. (2003). *Note sulle forme quadratiche libere e condizionate e sul riconoscimento del segno tramite minori.* in G. Crespi and others (Eds.). Recent Advances in Optimization, DATANOVA, Milan, 71-86.

GIORGI, G. (2013). Optimality conditions under generalized convexity revisited. *Annals of the University of Bucharest (Mathematical Series). 4*, LXII, 1-12.

GIORGI, G. (2017). Various proofs of the Sylvester criterion for quadratic forms. submitted. https://doi.org/10.5539/jmr.v9n6p55

GÜLER, O. (2010). *Foundations of Optimization*. Springer, New York. https://doi.org/10.1007/978-0-387-68407-9

HESTENES, M. R. (1966). Calculus of Variations and Optimal Control Theory, J. Wiley, New York.

HIRIART-URRUTY, J.-B. & LEMARECHAL, C. (1993). Convex Analysis and Minimization Algorithms, Vols I and II, Springer Verlag, Berlin. https://doi.org/10.1007/978-3-662-02796-7

LEROUX, A. (1984). Other determinental (sic) conditions for concavity and quasi-concavity. *Journal of Mathematical Economics, 13*, 43-49. https://doi.org/10.1016/0304-4068(84)90023-5

KEMP, M. C., & KIMURA, Y. (1978). Introduction to Mathematical Economics, Springer Verlag, New York. https://doi.org/10.1007/978-1-4612-6278-7

MARTOS, B. (1975). Nonlinear Programming, North Holland, Amsterdam.

MURATA, Y. (1977). Mathematics for Stability and Optimization of Economic Systems, Academic Press, New York.

NEWMAN, P. (1969). Some properties of concave functions. *Journal of Economic Theory, 1*, 291-314. https://doi.org/10.1016/0022-0531(69)90035-0

ORTEGA, J. M., & RHEINBOLDT, W. C. (1970). Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York.

OTANI, K. (1983). A characterization of quasi-concave functions. *Journal of Economic Theory, 31*, 194-196. https://doi.org/10.1016/0022-0531(83)90030-3

ROCKAFELLAR, R. T. (1976). Saddle points of Hamiltonian systems in convex Lagrange problems having a nonzero discount rate. *Journal of Economic Theory, 12*, 71-113. https://doi.org/10.1016/0022-0531(76)90028-4

SAMUELSON, P. A. (1947). *Foundations of Economic Analysis*. Harvard Univ. Press, Cambridge, Mass.

SIMON, C., & BLUME, L. (1994). *Mathematics for Economists.* Norton, New York.

STEIN, O. (2012). Twice differentiable characterizations of convexity notions for functions on full dimensional convex sets. *Schedae Informaticae, 21*, 55-63.

VIAL, J.-P. (1982). Strong convexity of sets and functions. *Journal of Mathematical Economics, 9*, 187-205. https://doi.org/10.1016/0304-4068(82)90026-X

Copyrights