2

Invex Functions (The Smooth Case)

2.1 Introduction

Usually, generalized convex functions have been introduced in order to weaken as much as possible the convexity requirements for results related to optimization theory (in particular, optimality conditions and duality results), to optimal control problems, to variational inequalities, etc. For instance, this is the motivation for employing pseudo-convex and quasi-convex functions in [142, 143]; [228] use convexlike functions to give a very general condition for minimax problems on compact sets. Some approaches to generate new classes of generalized convex functions have been to select a property of convex functions which is to be retained and then forming the wider class of functions having this property: both pseudo-convexity and quasi-convexity can be assigned to this perspective. Other generalizations have been obtained through altering the expressions in the definition of convexity, such as the arcwise convex functions in [8] and [9], the (h, ϕ)-convex function in [17], the (α, λ)-convex functions in [27], the semilocally generalized convex functions in [113], etc.

The reasons for Hanson's conception of invex functions [83] may have stemmed from any of these motivating forces, although in that paper Hanson dealt only with the relationships of invex functions to the Kuhn–Tucker conditions and Wolfe duality. More precisely, Hanson [83] noted that the usual convexity (or pseudo-convexity or quasi-convexity) requirements, appearing in the sufficient Kuhn–Tucker conditions for a mathematical programming problems, can be further weakened. Indeed, in the related proofs of the said conditions, there is no explicit dependence of the linear term $(x - y)$, appearing in the definition of differentiable convex, pseudo-convex and quasi-convex functions. This linear term was therefore substituted by an arbitrary vector-valued function, usually denoted by η and sometimes called "kernel function."

2.2 Invex Functions: Definitions and Properties

Definition 2.1. *Assume $X \subseteq R^n$ is an open set. The differentiable function $f : X \to R$ is invex if there exists a vector function $\eta : X \times X \to R^n$ such that*

$$f(x) - f(y) \geq \eta(x,y)^T \nabla f(y), \quad \forall x, y \in X. \tag{2.1}$$

It is obvious that the particular case of (differentiable) convex function is obtained from (2.1) by choosing $\eta(x,y) = x - y$. The term "invex" is due to Craven [43] and is an abbreviation of "invariant convex," since it is possible to create an invex function with the following method:
Let $g : R^n \to R$ be differentiable and convex and $\Phi : R^r \to R^n (r \geq n)$ be differentiable with $\nabla \Phi$ of rank n. Then $f = g \circ \Phi$ is invex, $\forall x, y \in R^r$, we have

$$f(x) - f(y) = g(\Phi(x)) - g(\Phi(y)) \geq (\Phi(x) - \Phi(y))^T \nabla g(\Phi(y)).$$

As $\nabla f(y) = \nabla \Phi(y) \nabla g(\Phi(y))$ and $\nabla \Phi(y)$ is of rank n, the equation $(\Phi(x) - \Phi(y))^T \nabla g(\Phi(y)) = \eta(x,y)^T \nabla f(y)$ has a solution $\eta(x,y) \in R^r$. Hence, $f(x) - f(y) \geq \eta(x,y)^T \nabla f(y), \forall x, y \in R^r$ for some $\eta : R^r \times R^r \to R^r$.

This characterization of invexity is closely related to (h, F)-convexity, a generalization of convexity based on the use of generalized means (see, e.g., [146,169]). The class of (h, F)-convex functions, with h, h^{-1} and F differentiable, from a subclass of invex functions. It was stated earlier that invexity was used by Hanson [83] to obtain sufficient optimality conditions (in terms of Kuhn–Tucker conditions) for a nonlinear programming problem. This is possible, an invex function shares with convex function the property that every stationary point is a global minimum point. Craven and Glover [45] and Ben-Israel and Mond [18] established the basic relationship between this property and the function η of Definition 2.1.

Theorem 2.2. *Let $f : X \to R$ be differentiable. Then f is invex if and only if every stationary point is a global minimizer.*

Proof. Necessity: Let f be invex and assume $\bar{x} \in X$ with $\nabla f(\bar{x}) = 0$. Then $f(x) - f(\bar{x}) \geq 0, \forall x \in X$, so $\bar{x}$ is a global minimizer of f over X.
Sufficiency: Assume that every stationary point is a global minimizer. If $\nabla f(y) = 0$, let $\eta(x,y) = 0$. If $\nabla f(y) \neq 0$, let

$$\eta(x,y) = \frac{[f(x) - f(y)] \nabla f(y)}{\nabla f(y)^T \nabla f(y)}.$$

Then f is invex with respect to η.

□

This is, of course, not the only possible choice of η. Indeed, if $\nabla f(y) = 0$, then $\eta(x,y)$ may be chosen arbitrarily, and if $\nabla f(y) \neq 0$, then

$$\eta(x,y) \in \left\{ \frac{[f(x) - f(y)] \nabla f(y)}{\nabla f(y)^T \nabla f()y} + v : v^T \nabla f(y) \leq 0 \right\},$$

a half-space in R^n.

This importance of functions with the stationary points as global minimizers had been recognized also by Zang et al. [255], who however, did not pursue any further analysis and applications.

Let us denote by $L_f(\alpha)$ the lower α-level set of a function $f : X \to R$, i.e., the set $L_f(\alpha) = \{x : x \in X, f(x) \leq \alpha\}, \forall \alpha \in R$. Zang et al. [255] characterized by means of the sets $L_f(\alpha)$ the functions whose stationary points are global minimizers, i.e., the class of invex functions.

Definition 2.3. *If $L_f(\alpha)$ is non-empty, then it is said to be strictly lower semi-continuous if, for every $x \in L_f(\alpha)$ and sequence $\{\alpha_i\}$, with $\alpha_i \to \alpha, L_f(\alpha_i)$ non-empty, there exist $k \in N$, a sequence $\{x^i\}$, with $x^i \to x$ and $\beta(x) \in R, \beta(x) > 0$, such that*

$$x^i \in L_f[\alpha_i - \beta(x) \left\| x^i - x \right\|], \quad i = k, k+1, \ldots.$$

The authors proved the following result.

Theorem 2.4. *A function $f : X \to R$, differentiable on the open set $X \subseteq R^n$, is invex if and only if $L_f(\alpha)$ is strictly lower semi-continuous, for every α such that $L_f(\alpha) \neq \Phi$.*

Proof. See Zang et al. [255].

□

Another characterization of invex functions stemming from Theorem 2.2, can be obtained through the conjugation operation. Let $f : X \to R$, $X \subseteq R^n$; given $\xi \in R^n$, we consider the collection of all affine functions $\xi^T x - \alpha$, with slope ξ, that minorize $f(x)$, i.e., $\xi^T - \alpha \leq f(x), \forall x \in X$. This collection, if non-empty, gives rise to the smallest α^* for which the above relation holds. If there is no affine function with slope ξ minorizing $f(x)$, we agree to set $\alpha^* = +\infty$. In any case $\alpha^* = f^*(\xi) = \sup_x \{\xi^T x - f(x)\}$ is precisely what is called the conjugate function of f (see [211]). By reiterating the operation $f \to f^*$ on X, we get the biconjugate of $f(x)$, defined by

$$f^{**} = \sup_{\xi} \{\xi^T x - f^*(\xi)\}.$$

It can be proved (see [91]) the following result.

Theorem 2.5. *Let $f : X \to R$ be differentiable on the open set $X \subseteq R^n$. Then $x^0 \in X$ is a (global) minimum point of f on X if and only if: (i) $\nabla f(x^0)$, and (ii) $f^{**}(x^0) = f(x^0)$. In such a case f^{**} is differentiable at x^0 and $\nabla f^{**}(x^0) = 0$.*

Proof. See Hiriart-Urruty [91].

□

Thus Theorem 2.5 gives another characterization of an invex function: it is a C differentiable function whose value at stationary points equals the value of its biconjugate.

Hanson and Rueda [89] sufficient conditions for invexity of a function are established through the use of linear programming. We shall revert to this question when we shall treat the applications of invexity to nonlinear programming problems. From Theorem 2.2 we get immediately that if f has no stationary points, then f is invex. Furthermore, Theorem 2.2 will be useful to state some relationships between invex functions and other classes of generalized convex functions. Some nice properties of convex functions are however lost in the invex case. In fact, unlike convex (or pseudo-convex) case, the restriction of an invex function on a not open set does not maintain the local/global property. Let us consider the following example.

Example 2.6. Let $f(x,y) = y(x^2-1)^2$, considered on the closed set $S = \left\{(x,y) \in R^2 : x \geq -\frac{1}{2}, y \geq 1\right\}$. Every stationary point of f on S is a global minimum point of f on S and therefore f is invex on S. The point $(-\frac{1}{2}, 1)$ is a local minimum point of f on S, with

$$f(-\frac{1}{2}, 1) = \frac{9}{16} > f(1,y) = f(-1,y) = 0.$$

The points $(1,y)$, $(-1,y), y \geq 1$, are the global minimizers for f on S.

If f is invex on an open set $X \subseteq R^n$, contrary to what asserted in Pini [201], it is not true that the set $A = \{x \in X, \nabla f(x) = 0\}$ is a convex set (as for convex functions). Let us consider the following example.

Example 2.7. Let $f(x,y) = y(x^2-1)^2$, defined on the open set $S = \left\{(x,y) \in R^2 : x \in R, y > 0\right\}$. The set of all its stationary points coincides with the set of all its minimum points (i.e., f on S). This set is given by $\{(1,y) : y > 0\} \bigcup \{(-1,y) : y > 0\}$, which is not a convex set in R^2.

As a consequence, for an invex function the set of all minimum points (the set of all stationary points if f is defined on an open set) is not necessarily a convex set. Ben-Israel and Mond [18] observed that there is an analogue of Theorem 2.2 for pseudo-convex functions.

Theorem 2.8. *A differentiable function on the open set $X \subseteq R^n$ is pseudo-convex on X if and only if*

$$(x-y)^T \nabla f(y) = 0 \Rightarrow f(y) \leq f(y + t(x-y)), \quad \forall t > 0. \tag{2.2}$$

Proof. Necessity: Obvious from the definition of pseudo-convexity. Here (2.2) holds for all real t.
Sufficiency: Suppose f is not pseudo-convex; that is, there exists (x,y) such that $(x-y)^T \nabla f(y) \geq 0$ and $f(x) < f(y)$. If $(x-y)^T \nabla f(y) = 0$, then (2.2) is contradicted. If $(x-y)^T \nabla f(y) > 0$, then there exists v which maximizes f on the line segment from y to x. Thus $\nabla f(v) = 0$ and therefore $(x-y)^T \nabla f(v) = 0$ and

$$f(v) \geq f(y) > f(x) = f(v + 1(x-v)),$$

contradicting (2.2).

□

We note that the class of functions differentiable on an open set X and all invex with respect to the same $\eta(x,y)$, is closed under addition on any domain contained in X, unlike the classes of quasi-convex and pseudo-convex functions which do not retain this property of convex functions. However, the class of functions invex on an open set X, but not necessarily with respect to the same $\eta(x,y)$, need not be closed under addition. For instance (see, [178,224]), consider $f_1 : R \to R$ and $f_2 : R \to R$ defined by $f_1(x) = 1 - e^{-(x+5)^2}$. Both f_1 and f_2 are invex, but $f_1 + f_2$ has a stationary point at $\bar{x} = 0$ which is not a global minimizer. In fact, for a given $\eta(x,y)$, the set of functions invex with respect to $\eta(x,y)$, form a convex cone; that is, the set is closed under addition and positive scalar multiplication. Therefore, we can state the following result.

Theorem 2.9. *Let $f_1, f_2, \ldots, f_m : X \to R$ all invex on the open set $X \subseteq R^n$, with respect to the same function $\eta(x,y) : X \times X \to R^n$. Then:*

1. *For each $\alpha \in R, \alpha > 0$, the function $\alpha f_i, i = 1, \ldots, m$, is invex with respect to the same η.*
2. *The linear combination of $f_1, f_2, \ldots, f_m$, with nonnegative coefficients is invex with respect to the same η.*

Following Smart [224] and Mond and Smart [179], a natural question is now the following: given two (or more) invex functions, how do we know if they are invex with respect to a common η. It is convenient to first prove a result characterizing functions for which no common η exists.

Lemma 2.10. *Let $f : X \to R, g : X \to R$ be invex. There does not exist a common η, with respect to which f and g are both invex if and only if there exists $x, y \in X, \lambda > 0$ such that $\nabla f(y) = -\lambda \nabla g(y)$ and $f(x) - f(y) + \lambda(g(x) - g(y)) < 0$.*

Proof. (a) Sufficiency: Assume there exist $x, y \in X, \lambda > 0$ such that $\nabla f(y) = -\lambda \nabla g(y)$ and $f(x) - f(y) + \lambda(g(x) - g(y)) < 0$. We wish to show that the system

$$f(x) - f(y) \geq \eta(x,y)^T \nabla f(y)$$
$$g(x) - g(y) \geq \eta(x,y)^T \nabla g(y)$$

has no solution $\eta(x,y) \in R^n$. Assume such an $\eta(x,y)$ exists. Now, as $\lambda > 0$, $g(x) - g(y) \geq \eta(x,y)^T \nabla g(y) \Rightarrow \lambda[g(x) - g(y)] \geq \lambda\eta(x,y)^T \nabla g(y)$. Therefore,

$$\begin{aligned} f(x) - f(y) + \lambda(g(x) - g(y)) &\geq \eta(x,y)^T \nabla f(y) + \lambda\eta(x,y)^T \nabla g(y) \\ &= \eta(x,y)^T [\nabla f(y) + \lambda \nabla g(y)] \\ &= 0, \end{aligned}$$

which contradicts $f(x) - f(y) + \lambda(g(x) - g(y)) < 0$. Hence, no common function $\eta(x,y)$ exists.

(b) Necessity: Assume no common function $\eta(x, y)$ exists. Then there exists $x, y \in X$ such that the system

$$f(x) - f(y) \geq \eta(x, y)^T \nabla f(y)$$

$$g(x) - g(y) \geq \eta(x, y)^T \nabla g(y)$$

has no solution $\eta(x, y)^T \in R^n$.

Rewrite the system as $A\eta(x, y) \leq C$, where $A = \begin{pmatrix} \nabla f(y)^T \\ \nabla g(y)^T \end{pmatrix}, C = \begin{pmatrix} f(x) - f(y) \\ g(x) - g(y) \end{pmatrix}$.

By Gale's Theorem of the alternative for linear inequalities (see, e.g., [143]), there exists $y \in R^2, y = (y_1, y_2)^T$, such that $A^T y = 0, C^T y = -1, y \geq 0$, that is,

$$\nabla f(y) y_1 + \nabla g(y) y_2 = 0,$$

$$[f(x) - f(y) y_1] + [g(x) - g(y) y_2] = -1,$$

$$y_1 \geq 0, \quad y_2 \geq 0.$$

Now, if $y_1 = 0$, then $\nabla g(y) y_2 = 0, [g(x) - g(y)] y_2 = -1, y_2 \geq 0$, which implies that $\nabla g(y) = 0$ and $g(x) - g(y) < 0$, which contradicts the invexity of g. Hence, $y_1 \geq 0$. Similarly, $y_2 > 0$. Thus,

$$\nabla f(y) = -\frac{y_2}{y_1} \nabla g(y) = -\lambda \nabla g(y), \quad \text{where} \quad \lambda = \frac{y_2}{y_1} > 0$$

and

$$f(x) - f(y) + \frac{y_2}{y_1}[g(x) - g(y)] = -1,$$

that is,

$$f(x) - f(y) + \lambda[g(x) - g(y)] < 0.$$

□

The negation of the Lemma 2.10 yields the next result.

Theorem 2.11. *Let $f : X \to R, g : X \to R$ be invex. A common η, with respect to which both f and g are invex, exists if and only if $\forall x, y \in X$ either*

1. $\nabla f(y) \neq \lambda \nabla g(y)$ for any $\lambda > 0$ or
2. $\nabla f(y) = -\lambda \nabla g(y)$ for some $\lambda > 0$ and

$$f(x) - f(y) \geq -\lambda[g(x) - g(y)].$$

Using Theorem 2.11, it is possible to give a more useful characterization of invex functions with respect to a common η.

Theorem 2.12. *Let $f : X \to R, g : X \to R$ be invex. A common η, with respect to which both f and g are invex, exists if and only if $f + \lambda g$ is invex for all $\lambda > 0$.*

Proof. (a) Necessity: this follows since the set of functions invex with respect to η is a convex cone.
(b) Sufficiency: assume $f+\lambda g$ is invex, for all $\lambda > 0$. Then, whenever $\nabla f(y) = -\lambda \nabla g(y)$ for some $\lambda > 0$, we have

$$f(x) + \lambda g(x) \geq f(y) + \lambda g(y), \quad \forall x \in X,$$

by invexity of $f + \lambda g$. That is,

$$\nabla f(y) = -\lambda \nabla g(y) \Rightarrow f(x) - f(y) \geq -\lambda[g(x) - g(y)], \quad \forall x \in X.$$

By Theorem 2.11, a common η exists.

□

Theorem 2.12 generalizes to any finite number of functions, and is useful for the requirements of invexity in sufficiency and duality results in optimization.

Corollary 2.13. *Let $f : X \to R, g_1, g_2, \ldots, g_m : X \to$ be invex. A common η with respect to which $f, g_1, g_2, \ldots, g_m$ are invex, exists if and only if $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_m g_m$ is invex for all $\lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_m > 0$.*

Proof. By induction; the case $m = 1$ is proved in Theorem 2.12. Assume the statement is true for some $k \in N$. Now $f, g_1, g_2, \ldots, g_{k+1}$ have a common η if and only if $f, g_1, g_2, \ldots, g_k$ have a common η with respect to which g_{k+1} is also invex. Now $f, g_1, g_2, \ldots, g_k$ have a common η if and only if $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_k g_k$ is invex for all $\lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_k > 0$. Therefore, $f, g_1, g_2, \ldots, g_{k+1}$ have a common η if and only if $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_k g_k$ is invex with respect to same η independent of $\lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_k > 0$, and g_{k+1} is invex with respect to the same η. But $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_k g_k$ and g_{k+1} have a common η if and only if $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_{k+1} g_{k+1}$ is invex for all $\lambda_{k+1} > 0$. Therefore, $f, g_1, g_2, \ldots, g_{k+1}$ have a common η if and only if $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_{k+1} g_{k+1}$ is invex for all $\lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_{k+1} > 0$.

□

Since it is assumed in Corollary 2.13 that f is invex, the necessary and sufficient condition could also be expressed as: $f + \lambda_1 g_1 + \lambda_2 g_2 + \cdots + \lambda_m g_m$ is invex for all $\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_m \geq 0$.

Like convex functions, invex functions with respect to a certain η are transformed into invex functions with respect to the same η, by a suitable class of monotone functions.

Theorem 2.14. *Let $\psi : R \to R$ be a monotone increasing differentiable convex function. If f is invex on X with respect to η, then the composite function $\psi \circ f$ is invex with respect to the same η.*

Proof. By the fact that $\psi(x + h) \geq \psi(x) + \psi'(x)h, \forall x, h \in R$, we get

$$\begin{aligned}
\psi(f(x)) &\geq \psi(f(y)) + \nabla f(y)\eta(x, y) \\
&\geq \psi(f(y)) + \psi'(f(y))\nabla f(y)\eta(x, y) \\
&= \psi(f(y)) + \nabla(f \circ \psi)(y)\eta(x, y).
\end{aligned}$$

□

Further generalizations of invexity are possible; indeed Hanson [83] introduced also the following classes of generalized convex functions.

Definition 2.15. *The differentiable function $f : X \to R$ is pseudo-invex if there exists $\eta : X \times X \to R^n$ such that for all $x, y \in X$,*

$$\eta(x,y)^T \nabla f(y) \geq 0 \Rightarrow f(x) - f(y) \geq 0;$$

f is quasi-invex if there exists $\eta : X \times X \to R^n$ such that $\forall x, y \in X$,

$$f(x) - f(y) \leq 0 \Rightarrow \eta(x,y)^T \nabla f(y) \leq 0.$$

We point out that if we do not specify the function η in the definition of quasi-invexity, it turns out that every function f is quasi-invex: it is sufficient to take η identically equal to zero. Definitions 2.1 and 2.15 can be further weakened if we consider, as in Kaul and Kaur [114], pointwise characterization at a point $x^0 \in X$. In this respect we say that a differentiable function $f : X \to R$ is invex at $x^0 \in X$, if there exists $\eta(x, x^0)$ such that $\forall x \in X$,

$$f(x) - f(x^0) \geq \eta(x, x^0)^T \nabla f(x^0).$$

f is pseudo-invex at $x^0 \in X$, if there exists $\eta(x, x^0)$ such that $\forall x \in X$,

$$\eta(x, x^0)^T \nabla f(x^0) \geq 0 \Rightarrow f(x) - f(x^0) \geq 0.$$

f is quasi-invex at $x^0 \in X$, if there exists $\eta(x, x^0)$ such that $\forall x \in X$,

$$f(x) - f(x^0) \leq 0 \Rightarrow \eta(x, x^0)^T \nabla f(x^0) \leq 0.$$

Craven [43] introduced further relaxations: the local invexity at a point and the invexity with respect to a cone.

Definition 2.16. *The differentiable function $f : X \to R, X \subseteq R^n, X$ open, is said to be locally invex at $x^0 \in X$, if there exist a function $\eta(x, x^0)$ and a positive scalar δ such that*

$$f(x) - f(x^0) \geq \eta(x, x^0)^T \nabla f(x^0), \quad \forall x \in X, \left\| x - x^0 \right\| < \delta.$$

Definition 2.17. *Let $f : X \to R^k$ be a differentiable vector-valued function; f is invex with respect to the cone K in R^k if*

$$f(x) - f(y) - \nabla f(y) \eta(x,y) \in K.$$

If K is polyhedral convex cone and $q^j, j = 1, \ldots, l$, denote the generating vectors of the dual cone K^* such that

$$K = \left\{ x \in R^k : q^j x \geq 0, j = 1, \ldots, l \right\},$$

the Definition 2.17 is nothing but the invexity with respect to η, Craven [43] has given a characterization of local invexity with respect to a cone. Assume

$f : R^n \to R^k$ and $\eta : R^n \times R^n \to R^n$ are functions of class C^2. Given, y, we write the Taylor expression of $\eta(\cdot, y)$ and $f(\cdot)$ up to quadratic terms as follows

$$\eta(x,y) = \eta(y,y) + A(x-y) + \frac{1}{2}(x-y)^T Q_0(x-y) + O(\|x-y\|^2)$$

$$f(x) = f(y) + B(x-y) + \frac{1}{2}(x-y)^T M_0(x-y) + O(\|x-y\|^2),$$

where A, B, Q and M_0 have the obvious significance. Then the following holds.

Theorem 2.18. *Let $f : R^n \to R^k$ be a function of class C^2; denote by K a closed convex cone in R^k such that $K \bigcap (-K) = 0$. If f is locally invex at y, with respect to η and with respect to the cone K and $\eta : R^n \times R^n \to R^n$ is a function of class C^2, for which $\eta(x,x) = 0$, then, after substitution of a term in the null space of B, η has the form*

$$\eta(x,y) = x - y + \frac{1}{2}(x-y)^T Q_0(x-y) + O(\|x-y\|^2), \tag{2.3}$$

where $M_0 - BQ_0$ is K-semidefinite. Conversely, if η has the form (2.3), and if $M_0 - BQ_0$ is K-positive definite, then f is locally invex at y, with respect to η and K.

Proof. See Craven [43].

□

Note that if f is a function defined on R and the cone K is the interval $[0, +\infty]$, the positive semidefiniteness of $M_0 - BQ_0$ is nothing but the condition

$$f''(y) - f'(y)\eta(y,y) \geq 0.$$

The conditions of Theorem 2.18 are however, from a computational point of view, difficult to apply. In Sect. 3, we shall see other sufficient conditions for invexity in nonlinear programming, through the use of linear programming. Further generalizations of invex functions can be obtained through notions similar to the ones utilized by Vial [239] to define strong and weak convex functions. On these lines Jeyakumar [100, 103] defined the following class of generalized invex functions.

Definition 2.19. *A differentiable function $f : X \to R, X \subseteq R^n$, is called ρ-invex with respect to the vector-valued function η and θ, if there exists some real number ρ such that, for every $x, y \in X$*

$$f(x) - f(y) \geq \eta(x,y)^T \nabla f(y) + \rho(\|\theta(x,y)\|^2).$$

If $\rho > 0$, then f is called strongly ρ-invex. If $\rho = 0$, we obviously get the usual definition of invexity and if $\rho < 0$, then f is called weakly ρ-invex.

It is clear that strongly ρ-invexity $\Rightarrow$ invexity $\Rightarrow$ weakly ρ-invexity. Rueda [213] points out that, under the assumption that $\|\nabla f(x)\| \neq 0$, Definition 2.19 is equivalent to invexity. Indeed, define

$$\eta_1(x,y) = \eta_x(x,y) + \rho(\|\theta(x,y)\|^2)\frac{\nabla f(x)}{[\nabla f(x)]^T \nabla f(x)}.$$

Thus, f is invex with respect to η_1.

Definition 2.20. *A differentiable function $f : X \to R$, is called ρ-pseudo-invex with respect to the vector-valued functions η and θ, if there exists some real number ρ such that, for every $x, y \in X$*

$$\eta(x,y)^T \nabla f(y) \geq -\rho(\|\theta(x,y)\|^2) \Rightarrow f(x) \geq f(y).$$

Definition 2.21. *A differentiable function $f : X \to R$, is called ρ-quasi-invex with respect to the vector-valued functions η and θ, if there exists some real number ρ such that, for every $x, y \in X$*

$$f(x) \leq f(y) \Rightarrow \eta(x,y)^T \nabla f(y) \leq -\rho(\|\theta(x,y)\|^2).$$

Pointwise definitions follow easily. The above definitions can be used to obtain general optimality and duality results for a nonlinear programming problem.

2.3 Restricted Invexity and Pointwise Invexity

The results characterizing invex functions as the class of functions for which stationary points are global minimizers, may be viewed as a special case of a more general theorem, due to Smart [224]; see also Mond and Smart [179], Molho and Schaible [166] and Chandra et al. [33].

For given $x, y \in R^n$, let $m(x,y)$ be a point in R^n and $\Lambda(x,y)$ a cone of R^n with vertex at $0 \in \Lambda$. Let $\Lambda^*(x,y)$ be the (positive) polar cone of $\Lambda(x,y)$, i.e.,

$$\Lambda^*(x,y) = \left\{ v \in R^n : v^T t \geq 0, \forall t \in \Lambda(x,y) \right\}.$$

Theorem 2.22. *Let $f : X \subseteq R^n \to R$ be differentiable. A necessary and sufficient condition for f to be invex with respect to $\eta : X \times X \to R^n$, subject to the restriction $\eta(x,y) \in m(x,y) + \Lambda(x,y), \forall x, y \in X$, is the following:*

$$\nabla f(y) \in \Lambda^*(x,y) \Rightarrow f(x) - f(y) - m(x,y)^T \nabla f(y) \geq 0.$$

Proof. Necessity: Assume f is invex with respect to $\eta(x,y) \in m(x,y) + \Lambda(x,y)$. Then $f(x) - f(y) \geq \eta(x,y)^T \nabla f(y) = (m(x,y) + t(x,y)^T)\nabla f(y)$, for some $t(x,y) \in \Lambda(x,y)$. Thus

$$\nabla f(y) \in \Lambda^*(x,y) \Rightarrow f(x) - f(y) \geq m(x,y)^T \nabla f(y), \quad \forall x \in X.$$

Sufficiency: assume that

$$\nabla f(y) \in \Lambda^*(x,y) \Rightarrow f(x) - f(y) - m(x,y)^T \nabla f(y) \geq 0.$$

Case (a). $\nabla f(y) \in \Lambda^*(x,y)$. Then take $\eta(x,y) = m(x,y)$. Since $0 \in \Lambda(x,y)$, we have $\eta(x,y) \in m(x,y) + \Lambda(x,y)$.
Case (b). $\nabla f(y) \notin \Lambda^*(x,y)$. Then there exists $t_1(x,y) \in \Lambda(x,y)$ such that $t_1(x,y)^T \nabla f(y) < 0$. If $f(x) - f(y) - m(x,y)^T \nabla f(y) \geq 0$, take $\eta(x,y) = m(x,y)$. On the other hand, if $f(x) - f(y) - m(x,y)^T \nabla f(y) < 0$, take $\eta(x,y) = m(x,y) + t_2(x,y)$, where

$$t_2(x,y) = \frac{f(x) - f(y) - m(x,y)^T \nabla f(y)}{t_1(x,y) \nabla f(y)} t_1(x,y).$$

Then

$$\begin{aligned} f(x) - f(y) - \eta(x,y)^T \nabla f(y) &= f(x) - f(y) - m(x,y)^T \nabla f(y) \\ &\quad - t_2(x,y)^T \nabla f(y) \\ &= f(x) - f(y) - m(x,y)^T \nabla f(y) \\ &\quad - \left(\frac{f(x) - f(y) - m(x,y)^T \nabla f(y)}{t_1(x,y)^T \nabla f(y)} \right)^T \\ &\quad \times t_1(x,y)^T \nabla f(y) \\ &= 0. \end{aligned}$$

Since

$$\frac{f(x) - f(y) - m(x,y)^T \nabla f(y)}{t_1(x,y)^T \nabla f(y)} > 0$$

and $\Lambda(x,y)$ is a cone, we have $t_2(x,y) \in \Lambda(x,y)$. Hence f is invex with respect to η, subject to the restriction $\eta(x,y) \in m(x,y) + \Lambda(x,y)$.

□

Let us apply the above results to some special cases:

(a) For convexity, take $m(x,y) = x - y$ and $\Lambda(x,y) = \{0\}$. The necessary and sufficient condition is $f(x) - f(y) \geq (x-y)^T \nabla f(y), \forall x, y \in X$.
(b) For arbitrary invexity, take $m(x,y)$ arbitrary, $\Lambda(x,y) = R^n$, so the necessary and sufficient condition is

$$y \in X, \nabla f(y) = 0 \Rightarrow f(x) - f(y) \geq 0, \quad \forall x \in X.$$

(c) For invexity with $\eta(x,y) \geq x - y$, take $m(x,y) = x - y, \Lambda(x,y) = R^n_+$. The condition is

$$y \in X, \nabla f(y) \geq 0 \Rightarrow f(x) - f(y) \geq (x-y)^T \nabla f(y), \quad \forall x \in X.$$

(d) For invexity with $\eta(x, y) + y \geq 0$, take $m(x, y) = -y, \Lambda(x, y) = R^n_+$. The condition is

$$y \in X, \nabla f(y) \geq 0 \Rightarrow f(x) - f(y) \geq -y^T \nabla f(y), \quad \forall x \in X.$$

We remark that if there does not exist a $y \in X$ such that $\nabla f(y) \geq 0$, then it is immediate that f is invex with respect to the desired η in both cases (c) and (d). Consider the following (non-convex) example: $f : R^2 \to R$, $f(x_1, x_2) = -x_2(x_1^2 + 1) + g(x_1)$, where $g : R \to R$ is any differentiable function. As $\nabla f(x_1, x_2) = (-2x_1x_2 + \nabla g(x_1); -(x_1^2 + 1)^T)$, there is no $(y_1, y_2) \in R^2$ such $\nabla f(y_1, y_2) \geq 0$, so f is invex with respect to some η_1 with $\eta_1(x, y) \geq x - y$, and also with respect to some η_2 with $\eta_2(x, y) + y \geq 0$.

A further special case of Theorem 2.22 concerns quadratic functions; we postpone the analysis of this case to Chap. 8, due to its importance in mathematical programming. We have already given the definitions, due to Kaul and Kaur [114], of invexity at a point x^0. We now make some other considerations on this case, under the assumption of twice differentiability of the functions. Let us therefore consider invex functions that are twice continuous differentiable. If $\nabla f(x^0) = 0$ for some $x^0 \in X$, a necessary condition for (global) invexity is that the Hessian matrix $\nabla^2 f(x^0)$ of f at x^0 is positive semidefinite. Indeed, if $\nabla f(x^0) = 0$ and f is invex, then x^0 is a point of global minimum. Therefore, $\nabla^2 f(x^0)$ is positive semidefinite.

2.4 Invexity and Other Generalizations of Convexity

In this section, we examine the main relationships between invexity definitions and other forms of generalized convexity. Obviously, for any assertion on a generalized convexity concept there is a generalized concavity counterpart. For invexity, the "incavity" is defined in a natural way by replacing $\geq$ with $\leq$.

First of all we note that:

(I) A differentiable convex function is also invex (take $\eta(x, y) = x - y$) but the converse is not true. Take, for example, the function $f(x) = \log x, x \in R$, which has no stationary points and is therefore invex. Obviously $f(x) = \log(x), x \in R$, is not convex (it is strictly concave) on its domain.

(II) A differentiable pseudo-convex function is also pseudo-invex, but not conversely. This property will be best precised in Theorem 2.25.

(III) A differentiable quasi-convex function is also quasi-invex, but not conversely (recall that every differentiable function is trivially quasi-invex).

For the reader's convenience we recall the basic definitions and properties of quasi-convex and pseudo-convex functions.

Definition 2.23 (Mangasarian [143], Avriel et al. [10]). *The function $f : X \to R$ is said to be quasi-convex on the convex set $X \subseteq R^n$ if for each $x, y \in X$ such that $f(x) - f(y) \leq 0$ and for each $\lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) \leq f(y)$.*

It is well known that f is quasi-convex on C if and only if the lower level sets $L_f(\alpha)$ are convex sets in R^n for each $\alpha \in R$. In case f is differentiable on the open convex set X, then f is quasi-convex on X if and only if $x, y \in X, f(x) - f(y) \leq 0 \Rightarrow (x - y)^T \nabla f(y) \leq 0$; or equivalently, $x, y \in X, (x - y)^T \nabla f(y) > 0 \Rightarrow f(x) - f(y) > 0$.

Definition 2.24. *The function $f : X \to R$, differentiable on the open set $X \subseteq R^n$, is pseudo-convex on X if*

$$x, y \in X, (x - y)^T \nabla f(y) \geq 0 \Rightarrow f(x) - f(y) \geq 0;$$

or equivalently,

$$x, y \in X, f(x) - f(y) < 0 \Rightarrow (x - y)^T \nabla f(y) < 0.$$

Furthermore, we say that f is strictly pseudo-convex on X if

$$x, y \in X, f(x) - f(y) \leq 0 \Rightarrow (x - y)^T \nabla f(y) < 0,$$

and we say that f is strongly pseudo-convex on X if f is pseudo-convex and satisfies the following conditions: For every $x^0 \in X$ and for every $v \in R^n, \|v\| = 1$, such that $v^T \nabla f(x^0) = 0$, there exist positive ϵ and α such that

$$f(x^0 + tv) \geq f(x^0) + \frac{1}{2}\alpha t^2,$$

for every $t \in R, 0 \leq t \leq \epsilon$.

(IV) Every invex function is also pseudo-invex for the same function η, but not conversely (see [114]). We have already remarked that a (differentiable) function without stationary points is invex, thanks to Theorem 2.2. Moreover, it results that the class of invex and pseudo-invex functions are coincident. This is not in contrast with property (IV), which is established with respect to the same η. We may note that some authors (see, e.g., Hanson and Mond [87], Kim [118] still consider pseudo-invexity as a generalization of invexity. We can therefore assert the following property:

(V) Every pseudo-convex function is invex; every pseudo-invex function is quasi-invex, but not conversely. For what concerns property (II) or its equivalent statement expressed by the first part of property (V), we have the following results, due to Pini [201].

Theorem 2.25. *The class of pseudo-convex functions on $X \subseteq R^n$ is strictly included in the class of invex functions if $n > 1$; if $n = 1$ the two classes coincide.*

Instead of following the proof of Pini [201], it is more useful to prove the following lemma [178, 224]).

Lemma 2.26. *Let $f : X \to R$, where X is an interval (open, half-open or closed) in R. If f is invex on X then it is also quasi-convex on X.*

Proof. We show that for every $\alpha \in R$, the lower level sets $L_f(\alpha)$ are convex. Assume to contrary that there exists $\alpha \in R$ such that $L_f(\alpha)$ is not convex. Then $L_f(\alpha)$ is the union of more than one disjoint intervals in X. Consider any two such intervals, I_1 and I_2, which are consecutive. Without loss of generality, $x_1 \in I_1$ and $x_2 \in I_2 \Rightarrow x_1 < x_2$. By continuity of f, I_1 must be closed on the right and I_2 must be closed on the left.

That is, there exists $\bar{x}_1 \in I_1$ such that $x_1 \leq \bar{x}_1, \forall x_1 \in I_1$ and $f(\bar{x}_1) = \alpha$; and there exists $\bar{x}_2 \in I_2$ such that $x_2 \geq \bar{x}_2, \forall x_2 \in I_2$ and $f(\bar{x}_2) = \alpha$. By assumption, $f(\alpha) > \alpha, \forall x \in (\bar{x}_1, \bar{x}_2)$. Since f is differentiable, then by the Mean Value Theorem, there exists $\bar{x} \in (\bar{x}_1, \bar{x}_2)$ such that $\nabla f(\bar{x}) = 0$. As $f(\bar{x}) > \alpha$, then $\bar{x}$ is not a global minimizer, which contradicts f being invex. □

The converse of Lemma 2.26 does not hold: take, e.g., the function $f : X \to R$, $f(x) = x^3$, which is quasi-convex (quasi-concave) on R, but not invex, since $\bar{x} = 0$ is a stationary point which is not global minimizer. Moreover, Lemma 2.26 does not hold when $X \subseteq R^n$ with $n > 1$. Consider the following example: $f : R^2 \to R, f(x_1, x_2) = 1 + x_1^2 - e^{-x_2^2}$. The function f has one stationary point, namely $x^* = (0, 0)$, and x^* is a global minimizer of f, so f is invex. However, f is not quasi-convex; take, e.g., $x = (1.12, 2.32940995)$ and $y = (1.31, 1.64704975)$. Now, $f(x) \leq f(y)$, but $(x - y)^T \nabla f(y) > 0$.

Another example is given by Ben-Israel and Mond [18]: The function $f : R^2 \to R, f(x_1, x_2) = x_1^3 + x_1 - 10x_2^3 - x_2$ is invex, since there are no stationary points. Taking $y = (0, 0), x_1 = 2, x_2 = 1$, gives $f(x) < f(y) < 0$ but $(x - y)^T \nabla f(y) > 0$, so f is not quasi-convex.

Another result useful to detect the relationships between the different classes of functions here considered is the following Theorem, due to Crouzeix and Ferland [50] and Giorgi [69]. See also Smart [224] and Mond and Smart [178].

Theorem 2.27. *Let f be differentiable quasi-convex function on the open convex set $X \subseteq R^n$. Then f is pseudo-convex on X if and only if f has a global minimum point at $x \in X$, whenever $\nabla f(x) = 0$.*

Theorem 2.27 asserts, in other words, that, under the assumption of quasi-convexity, invexity and pseudo-convexity coincides. So for an invex function not to be pseudo-convex, it must also not be quasi-convex. Taking this result into account, together with Lemma 2.26 and the related remarks, the proof of Theorem 2.25 is immediate.

Proof (of Theorem 2.27 Giorgi [69]). The necessary part of the theorem follows from the definition of pseudo-convex functions. As for sufficiency, let

$x^0 \in X, \nabla f(x^0) = 0 \Rightarrow x^0$ is a global minimum point of $f(x)$ on X, i.e., $(x - x^0)^T \nabla f(x^0) = 0 \Rightarrow f(x) \geq f(x^0), \forall x \in X$. It is obvious that $f(x)$ is then pseudo-convex at x^0 with respect to X. Let us now prove that: $f(x)$ quasi-convex on $X; x^0 \in X, \nabla f(x^0) \neq 0$ implies $f(x)$ pseudo-convex at x^0, i.e., $(x - x^0)^T \nabla f(x^0) \geq f(x) \geq f(x^0), \forall x \in X$. Let us consider a point $x^1 \in X$, such that

$$(x^1 - x^0)^T \nabla f(x^0) \geq 0 \tag{2.4}$$

but for which it is

$$f(x^1) < f(x^0). \tag{2.5}$$

Thus x^1 belongs to the nonvoid set

$$X_0 = \left\{x : x \in X, f(x) \leq f(x^0)\right\}$$

whose elements, thanks to the quasi-convexity of $f(x)$, verify the relation

$$x \in X_0 \Rightarrow (x - x^0)^T \nabla f(x^0) \leq 0. \tag{2.6}$$

Let us now consider the sets, both non-void,

$$W = \left\{x : x \in X, (x - x^0)^T \nabla f(x^0) \geq 0\right\}, \quad \text{and} \quad X_{00} = X_0 \cap W.$$

the following implication obviously holds:

$$x \in X_{00} \Rightarrow x \in H_0 = \left\{x : x \in X, (x - x^0)^T \nabla f(x^0) = 0\right\}.$$

It is therefore, evident that X_{00} is included in the hyperplane (recall that $\nabla f(x^0) \neq 0$) $H = \left\{x : x \in R^n, (x - x^0)^T \nabla f(x^0) = 0\right\}$, a hyperplane supporting X_0 covering to (2.6). Relation (2.4) and (2.5) point out that x^1 belongs to W and X_0 and hence to X_{00}, H_0 and H. Moreover, (2.5) says that x^1 lies in the interior of X_0; therefore x^1 at the same time belongs to the interior of a set and to a hyperplane supporting the same set, which is absurd. So relation (2.5) is false and (2.4) implies $f(x^1) \geq f(x^0)$.

□

We remark that the previous result states that a quasi-convex function $f(x)$ is thus pseudo-convex at every point $x \in X$ whenever $\nabla f(x) \neq 0$. Consequently we note that those sufficient conditions to test the quasi-convexity of a function in a convex set X where $\nabla f(x) \neq 0, \forall x \in X$, really locate the class of pseudo-convex functions. This is for example, the case of determinantal conditions for twice continuously differentiable functions, established by Arrow and Enthoven [5].

We can therefore add to the previous result, the following ones:

(VI) The classes of invex and pseudo-invex functions coincide.
(VII) The classes of quasi-convex and invex functions have only a partial overlapping.

We consider again pseudo-invex and quasi-invex functions. For what concerns pseudo-invex functions, we already know that if we do not impose further specifications on the choice of the kernel function η, this class coincides with the class of invex functions. However, if we consider the properties of these two classes of functions (invex and pseudo-invex) with respect to a specific function η, these properties are not the same. For example, unlike invex functions, the sum of pseudo-invex functions with respect to the same η is not pseudo-invex, with respect to that η. Consider, e.g., the following functions: $f(x) = \log x$ and $g(x) = -2x^2$ both defined on $X = \{x \in R : x > 0\}$. Both functions are pseudo-invex for $\eta(x,y) = x - y$. Indeed, $f(x) = \log x$ is strictly increasing function, being $f'(x) = \frac{1}{x} > 0, \forall x \in X$; therefore, $\eta(x,y)f'(x) \geq 0 \Leftrightarrow \eta(x,y) \geq 0$. Thus $\eta(x,y) = x - y \geq 0 \Leftrightarrow x \geq y \Rightarrow f(x) \geq f(y)$. So f is pseudo-invex with respect to $\eta(x,y) = x - y$.

The function g is strictly decreasing on X, as $g'(x) = -4x < 0, \forall x \in X$. We have $\eta(x,y)g'(y) \geq 0 \Leftrightarrow \eta(x,y) \leq 0; \eta(x,y) = x - y \leq 0 \Leftrightarrow x \leq y \Rightarrow g(x) \geq g(y)$, so g is pseudo-invex with respect to $\eta(x,y) = x - y$. The sum $z = f + g$ is $z = \log x - 2x^2, x > 0$. We have $z' = \frac{1}{x} - 4x = \frac{1-4x^2}{x}$. Thus $z' \geq 0 \Leftrightarrow 1 - 4x^2 \geq 0 \Rightarrow x \leq \frac{1}{2}$. Therefore $z(x)$ has a maximum point at $x = \frac{1}{2}$, so it is not pseudo-invex.

As for what concerns quasi-invex functions, we know that the class of pseudo-invex functions (i.e., invex functions) is strictly contained in the class of quasi-invex functions.

However, if we consider a pseudo-invex function f with respect to a certain function η, it is no longer true that f is also quasi-invex with respect to the same η. The converse also holds. Consider the following example.

Example 2.28. Let $f(x) = x^2 - 2x$ defined on R and

$$\eta(x,y) = \begin{cases} -1, & \forall (x,y) = (2,0) \\ 1, & \forall (x,y) = (x,1) \\ \frac{(x-y)(x+y-2)}{2(y-1)}, & \forall (x,y) \neq (x,1). \end{cases}$$

Let us verify that f is pseudo-invex with respect to $\eta(x,y)$; we have $f'(y) = 2y - 2$. If $(x-y) \neq (2,0)$ and $(x,y) \neq (x,1)$, then

$$\begin{aligned} \eta(x,y)f'(y) &= (x-y)(x+y-2) \\ &= x^2 - 2x - (y^2 - 2y) \geq 0. \end{aligned}$$

$$\Rightarrow x^2 - 2x \geq y^2 - 2y \Leftrightarrow f(x) \geq f(y).$$

If $(x,y) = (x,1)$, then $\eta(x,1)f'(1) = 0$ and $f(x) \geq f(1)$, being $x^2 - 2x \geq -1 \Leftrightarrow (x-1)^2 \geq 0, \forall x \in R$. If $(x,y) = (2,0)$, then $\eta(2,0)f'(0) = 2$ and

$f(2) = f(0) = 0$. So, f is not quasi-invex with respect to the same η; indeed if we choose $x = 2$ and $y = 0$, we have $f(x) \leq f(y)$, but $\eta(2,0)f'(0) = 2 > 0$.

To verify that a quasi-invex function f with respect to a certain η, may not be pseudo-invex with respect to the same η, consider the function $f(x) = \tan x, x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ which is quasi-convex and also therefore quasi-invex with respect to $\eta(x,y) = x - y$ but not pseudo-convex. Choose, e.g., $x = -\frac{\pi}{4}, y = 0$; we have $f(\frac{\pi}{4}) = 1 < f(0) = 0$, but $(x-y)f'(y) = 0$. So f is not pseudo-invex with respect to $\eta(x,y) = x - y$ (and it is not invex with respect to the same η).

We recall again that, if no specification is made on the choice of η, the class of quasi-invex functions coincides with the class of differentiable functions.

Similar to pseudo-invex functions, the sum of quasi-invex functions with respect to the same functions η, need not be quasi-invex with respect to that η. For example:

Example 2.29. Consider the functions $f(x) = \arctan(x)$ and $g(x) = -x^2$, both defined on $X = \{x \in R : x \geq 0\}$. Both functions are quasi-convex and therefore quasi-invex with respect to $\eta(x,y) = x - y$. The sum $z = \arctan(x) - x^2$ is not quasi-convex on X : Choose $x = 0$ and $y = 0.8$. We have $z(x) = 0 < z(y) = 0.03$. Therefore, we should have $z(x) \leq 0.03$ for every $x \in \left(0, \frac{8}{10}\right)$. But if we consider $x = 0.5$, we have $z(x) = 0.21 > 0.03$. So $z(x)$ is not quasi-invex with respect to $\eta(x,y) = x - y$.

We now give following results from Pini [201] which ensure that an invex function is pseudo-convex or quasi-convex.

Theorem 2.30. *Assume that $X \subseteq R^n$ is an open convex set and $f : X \to R$ is an invex function, with respect to η. If*

$$(x-y)^T \nabla f(y) \leq \eta(x,y)^T \nabla f(y), \quad \forall x, y \in X, \tag{2.7}$$

such that $f(x) < f(y)$, then f is pseudo-convex. If

$$(x-y)^T \nabla f(y) < \eta(x,y)^T \nabla f(y), \quad \forall x, y \in X \tag{2.8}$$

such that $f(x) \leq f(y)$, then f is strictly pseudo-convex.

Proof. If $x, y \in X$ and $f(x) < f(y)$, by the hypothesis of invexity and (2.7), we get

$$\begin{aligned}(x-y)^T \nabla f(y) &= [(x-y) - \eta(x,y)]^T \nabla f(y) + \eta(x,y) \nabla f(y) \\ &\leq [(x-y) - \eta(x,y)]^T \nabla f(y) + f(x) - f(y) \\ &< [(x-y) - \eta(x,y)]^T \nabla f(y) < 0.\end{aligned}$$

If $x, y \in X$ and $f(x) \leq f(y)$, then (2.8) implies that

$$\begin{aligned}(x-y)^T \nabla f(y) &= [(x-y) - \eta(x,y)]^T \nabla f(y) + \eta(x,y) \nabla f(y) \\ &\leq [(x-y) - \eta(x,y)]^T \nabla f(y) + f(x) - f(y) \\ &\leq [(x-y) - \eta(x,y)]^T \nabla f(y) < 0.\end{aligned}$$

□

Theorem 2.31. *Assume that for every $y \in R^n$ the function $x \to \eta(x, y)$ is differentiable at the point $x = y, \eta(x, y) = 0$ and $\eta_x(y, y) = 1$. If $f : X \to R$ is invex with respect to η and*

$$f(x) < f(y) \Rightarrow \eta(x, y)^T \nabla f(y) \leq (x - y)^T \nabla f(y)$$

and

$$v^T \nabla f(y) = 0 \Rightarrow (v^T \eta_{xx}(y, y)v) \nabla f(y) > 0,$$

then f is strongly pseudo-convex.

Proof. Choose $x^0 \in X$ and $v \in R^n$, with $\|v\| = 1$ such that $v^T \nabla f(x^0) = 0$. Since f is invex, we have

$$f(x^0 + tv) - f(x^0) \geq [\eta(x^0 + tv, x^0) - tv]^T \nabla f(x^0).$$

Since

$$\frac{d}{dt}[\eta(x^0 + tv, x^0) - tv]^T \nabla f(x^0)_{t=0} = 0,$$

it is sufficient to prove that

$$\frac{d^2}{dt^2}[\eta(x^0 + tv, x^0) - tv]^T \nabla f(x^0)_{t=0} > 0;$$

this is equivalent to

$$\nabla f(x^0)[v^T \eta_{xx}(x^0, x^0)v] > 0,$$

which is indeed true by assumption.

□

Theorem 2.32. *Let $f : X \to R$ be invex on the open convex set $X \subseteq R^n$, with respect to the kernel function η. If $(x - y)^T \nabla f(y) > 0 \Rightarrow \eta(x, y)^T \nabla f(y) \geq (x - y)^T \nabla f(y)$, for every $x, y \in X$, then f is quasi-convex on X.*

Proof. We estimate the difference $f(x) - f(y)$ whenever $(x - y)^T \nabla f(y) > 0$. We readily get

$$\begin{aligned} f(x) - f(y) &\geq \eta(x, y)^T \nabla f(y) \\ &= [\eta(x, y) - (x - y)]^T \nabla f(y) + (x - y)^T \nabla f(y) \\ &> [\eta(x, y) - (x - y)]^T \nabla f(y) > 0. \end{aligned}$$

□

Recall now the following definitions (see [10]).

Definition 2.33. *Let f be a function defined on the convex set $X \subseteq R^n$. We say that f is semi-strictly quasi-convex on X if*

$$f(x) < f(y) \Rightarrow f(\lambda x + (1 - \lambda)y) < f(y), \quad \forall x, y \in X, x \neq y, \lambda \in (0, 1).$$

Following Pini [201], we can give a sufficient condition for semi-strictly quasi-convexity.

Theorem 2.34. *Suppose that $f : X \to R$ is invex with respect to η, and that for every $x^0 \in X$ and $v \in R^n, \|v\| = 1$, such that $v\nabla f(x^0) = 0$, one of the two following conditions hold:*

$$\eta(x^0 + tv, x^0)^T \nabla f(x^0) > 0, \quad t \in [-a, b] \tag{2.9}$$

or

$$\eta(x^0 + tv, x^0)^T \nabla f(x^0) \geq 0, \quad t \in (-a, b),$$

$$\eta(x^0 - av, x^0)^T \nabla f(x^0) > 0, \quad \eta(x^0 + bv, x^0)^T \nabla f(x^0) > 0, \tag{2.10}$$

for some suitable $a, b > 0$. Then f is semi-strictly quasi-convex.

Proof. By Theorem 3.34 of Avriel et al. [10], it is sufficient to show that if $v\nabla f(x^0) = 0$, then the function $F(t) = f(x^0 + tv)$ does not admit a one-sided semi-strict local minimum at $t = 0$. Since f is an invex function, we have that

$$f(x^0 + tv) - f(x^0) \geq \eta(x^0 + tv, x^0)^T \nabla f(x^0),$$

that is,

$$F(0) \leq F(t) - \eta(x^0 + tv, x^0)^T \nabla f(x^0).$$

From (2.9), (2.10) it follows that $F(0) < F(-a), F(0) < F(b)$ and $F(0) \leq F(t), \forall t \in (-a, b)$. The thesis follows from the definition of one-sided semi-strict local maximum.

□

2.5 Domain and Range Transformations: The Hanson–Mond Functions

We follow here the approach of Smart [224], Mond and Smart [178] and Rueda [213]. These authors analyze in particular the article of Horst [94] dealing with non-convex nonlinear programs which may be transformed into convex programs via domain and/or range transformations in order to employ algorithms developed for convex programs. Convex range transformable functions, or F-convex functions, were first introduced by De Finetti [54].

Definition 2.35. *Let $f : X \to R, X$ a convex set in R^n. f is said to be convex range transformable or F-convex, if there exists a continuous, strictly monotone increasing function $F : range(f) \to R$, such that $F \circ f$ is convex on X. That is:*

$$F\left[f(\lambda x + (1-\lambda)y)\right] \leq \lambda F\left[f(x)\right] + (1-\lambda) F\left[f(x)\right], \tag{2.11}$$

$\forall x, y \in X, \forall \lambda \in [0, 1]$.

p-Convex functions (or power convex functions) and r-convex functions (see, [6–8, 94, 127, 146, 169]) are included in the class of convex range transformable functions. Concerning this subject we recall that the rth-generalized mean of $f(x)$ and $f(y)$, with $f(x)$ and $f(y)$ real and positive, defined as follows:

$$\begin{aligned} M_r\left(f(x), f(y), \lambda\right) &= M_r(f, \lambda) \\ &= \left[\lambda(f(x))^r + (1-\lambda)(f(y))^r\right]^{\frac{\lambda}{r}}, \end{aligned} \tag{2.12}$$

if $r \neq 0, \lambda \in [0,1]$.

It is possible to generalize (2.12) to the following cases:

$$M_0(f, \lambda) = \lim_{r \to 0} M_r(f, \lambda) = [f(y)]^{\lambda} \cdot [f(x)]^{1-\lambda}$$

$$M_{+\infty}(f, \lambda) = \lim_{r \to +\infty} (f, \lambda) = \max\left[f(x), f(y)\right].$$

$$M_{-\infty}(f, \lambda) = \lim_{r \to -\infty} (f, \lambda) = \min\left[f(x), f(y)\right].$$

Definition 2.36. *The function $f(x) > 0$ defined on the convex set $X \subseteq R^n$ is p-convex on X if there exists $p \geq 1$ such that $F \circ f = f^p$ is convex on X, i.e.,*

$$f(\lambda x + (1-\lambda)y) \leq M_p(f, \lambda), \quad \forall x, y \in X, \;\; \forall \lambda \in [0,1].$$

The previous inequality gives the usual definition of convexity for $p = 1$. If $1 < p < +\infty$, p-convexity is a special case of pseudo-invexity, i.e., of invexity. Indeed, if f is p-convex, then f^p is convex by definition and therefore it is invex. Since pseudo-invexity is equivalent to invexity for real functions, then there exists η such that

$$\eta(x,y)^T (\nabla f(y))^p \geq 0 \Rightarrow (f(x))^p \geq (f(y))^p.$$

Hence $f(x) \geq f(y)$, which proves that f is pseudo-invex. Note that invex functions need not be p-convex.

Example 2.37. Let $f : \left(0, \frac{\pi}{2}\right) \to R$ be defined by $f(x) = \sin x$. Then f is invex but it is not p-convex as can be seen by taking $y = \frac{\pi}{4}, x = \frac{\pi}{3}, p = 2$ and $\lambda = \frac{1}{2}$. In order to get rid of the restriction $f(x) > 0$, Avriel [6], Martos [146] and others proposed the following definition.

Definition 2.38. *The function $f : X \to R$ is r-convex on the convex set $X \subseteq R^n$, if for all $r, \lambda, -\infty \leq r \leq +\infty, 0 \leq \lambda \leq 1$, satisfies*

$$f(\lambda x + (1-\lambda)y) \leq \log M_r(e^{(f(x))}, e^{(f(y))}, \lambda).$$

Avriel [6] has proved that $f(x)$ is r-convex, with $r \neq 0$, if and only if the function $e^{rf(x)}$ is convex for $r > 0$ and concave $r < 0$. For $r > 0$, this is just the definition of r-convexity given by Horst [94]:

f is said to be r-convex if there exists $r > 0$ such that $F \circ f = e^{rf}$ is convex. For $0 < r < \infty$, we shall show that r-convexity is a special case of pseudo-invexity. If f is r-convex it follows that e^{rf} is convex and therefore it is invex (differentiability is assumed).

From

$$e^{rf(x)} - e^{rf(y)} \geq \eta(x,y)^T e^{rf(y)\nabla f(y)} \nabla f(y)$$

it follows that

$$e^{r(f(x)-f(y))} - 1 \geq \eta(x,y)^T.$$

Assume $\eta(x,y)^T \nabla f(y) \geq 0$. From the inequality above $e^{r(f(x)-f(y))} \geq 1$, which implies $r(f(x) - f(y)) \geq 0$. Since $r > 0$, it follows that $f(x) \geq f(y)$, which proves that f is pseudo-invex. From the previous example it follows that invex functions need not be r-convex. More generally, convex range transformable functions are quasi-convex [54, 94]. If in addition, a differentiable function f is convex range transformable with respect to a differentiable F, then f is invex. This may be seen by noting that, $\forall x, y \in X$,

$$F \circ f(x) - F \circ f(y) \geq (x-y)^T \nabla (F \circ f)(y) = (x-y)^T \nabla F(f(y)) \nabla f(y),$$

by convexity of $F \circ f$ and the by chain rule.

If $\nabla f(y) = 0$, then $F \circ f(x) \geq F \circ f(y), \forall x \in X$. By monotonicity of F, this implies that $f(x) \geq f(y), \forall x \in X$, so f is invex. By Theorem 2.27, f must also be pseudo-convex. Thus the class of differentiable convex range transformable (F-convex)functions, with F differentiable, form a strict sub-class of the invex functions.

A more general classification is obtained by incorporating a domain transformation [94].

Definition 2.39. *Let $f : X \to R, X \subseteq R^n, X$ convex. f is said to be (h, F)-convex if there exists a continuous one-to-one mapping $h : X \to h(X) \subseteq R^n$, and a continuous strictly monotone increasing function F : range$(f) \to R$ such that $h(X)$ is a convex set and $F \circ f \circ h^{-1}$ is a convex function on $h(X)$, i.e., $\forall x, y \in X$, and $\lambda \in [0, 1]$, we have*

$$f\left[h^{-1}(\lambda h(x) + (1-\lambda)h(y))\right] \leq F^{-1}\left[\lambda F(f(x)) + (1-\lambda)F(f(y))\right].$$

Horst [94] has shown that (h, F)-convex functions need not be quasi-convex; the purpose of the domain transformation h is to obtain a quasi-convex function which is F-convex. Assuming that h and F are differentiable with ∇h of full rank, so that h^{-1} is differentiable, (h, F)-convexity implies invexity. This follows, since $\forall x, y \in X$,

$$\begin{aligned}
(F \circ f \circ h^{-1})(x) &- (F \circ f \circ h^{-1})(y) \\
&\geq \nabla (F \circ f \circ h^{-1})(y) \\
&= \nabla F(f \circ h^{-1})(y) \cdot \nabla f(h^{-1}(y)) \nabla h^{-1}(y).
\end{aligned}$$

If $\nabla f(x^*) = 0$, then as h is onto,there exists $y \in h(X)$ such that $h(x^*) = y$ and $h^{-1}(y) = x^*$. Therefore, $(F \circ f \circ h^{-1})(x) - (F \circ f)(x^*) \geq 0, \forall x \in h(X)$. As F is monotonic increasing, then $(f \circ h^{-1})(x) \geq f(x^*), \forall x \in h(X)$. Since h is onto, $f(z) \geq f(x^*), \forall z \in X$. Hence, every stationary point of f yields a global minimum on X, so f is invex on X.

Rueda [213] has shown that invex functions need not be (h, F)-convex. For further considerations on the relationships between invexity and (h, F)-convexity, see Smart [224] and Mond and Smart [178]. (h, F)-convex functions are actually a special case of the arcwise convex functions described in Avriel [8] and Avriel and Zang [9]. We can consider any continuous path from x to y instead of the straight line between x and y. Let $p_{x,y}(\lambda)$, where $p_{x,y}(0) = x$ and $p_{x,y}(1) = y$, represents a continuous path from x to y in R^n such that $f(p_{x,y}(\lambda)), 0 \leq \lambda \leq 1$, is defined. Let h be a continuous strictly increasing scalar function that implies $f(x)$ and $f(y)$ in its domain. Then f is said to be arcwise convex or (p, θ)-convex if

$$f(p_{x,y}(\lambda)) \leq h^{-1}\left[\lambda h(f(x)) + (1-\lambda)h(f(y))\right],$$

for all x, y in the domain of f, $0 \leq \lambda \leq 1$. For (h, F)-convexity these paths (or arcs) are h-mean value functions given by

$$p_{x,y}(\lambda) = h^{-1}\left[\lambda h(x) + (1-\lambda)h(y)\right].$$

Rueda [213] has shown that an arcwise convex function, with path and range transformation assumed to be differentiable, is pseudo-invex, and hence invex, but the converse does not hold.

We now briefly treat the so-called Hanson–Mond functions. Hanson and Mond [86] introduced a generalization of convexity based on sublinear functionals, intending to generalize both convex and invex functions. However, this class of functions is in fact the class of invex functions.

Definition 2.40. *The functional $F : D \to R, D \subseteq R^n$ is said to be sublinear if*

(i) $F(a + b) \leq F(a) + F(b), \forall a, b \in D$,
(ii) $F(\alpha x) \leq \alpha F(x), \forall x \in D, \forall \alpha \geq 0$ *such that* $x \in D, \alpha x \in D$.

Note that (ii) implies $F(0) = 0$.

Definition 2.41 (Hanson and Mond [86]). *Let $f : X \to R$ be differentiable; f is said to be a Hanson–Mond function if there exists a sublinear functional $F(x, y; \cdot) : X \times X \times R^n \to R$ such that $\forall x, y \in X$,*

$$f(x) - f(y) \geq F(x, y; \nabla f(y)).$$

These functions are also called F-convex functions (see, e.g., [20, 32, 77, 185, 204]).

Invex functions are Hanson–Mond functions, since if f is invex with respect to η, we can define F in Definition 2.41 by $F(x, y; a) = \eta(x, y)^T a$. But, note also that if f is a Hanson–Mond function and $\nabla f(y) = 0$, then since $F(x, y; a) = 0, \forall x \in X, y$ is a global minimizer of f, so f is invex. Therefore, the Hanson–Mond functions correspond to the invex functions. Craven and Glover [45] proved the equivalence between the two said classes. Caprari [26] proved this equivalence also with regards to the Lipschitzian case and also other type of equivalence involving pseudo-Hanson–Mond functions and quasi-Hanson–Mond functions. In spite of this, there is still a lot of papers dealing with (generalized) Hanson–Mond functions, with the conviction that these classes are true generalizations of the corresponding classes of invex functions.

2.6 On the Continuity of the Kernel Function

The continuity of the kernel of invex functions was studied by Smart [224, 225]. Here we follow his analysis. Usually, in the main applications of invexity (mathematical programming, variational and control problems, etc.) there are no restrictions on the analytical properties on the kernel function η, such as continuity or differentiability, etc. However, there are some type of problems where assumptions about the kernel η need to be made. Smart [225] describes two examples where continuity of η must be imposed.

In Parida et al. [195] a variational-like inequality problem is examined and applied to an invex mathematical program with the condition that η be continuous (in fact, continuity of η is included in the definition of invexity in [195]). The variational-like inequality problem considered is as follows:

Given a closed convex set K of R^n, and two continuous maps $F : K \to R^n$ and $\eta : K \times K \to R^n$, find $\bar{x} \in K$ such that

$$F(\bar{x})^T \eta(x, \bar{x}) \geq 0, \quad \forall x \in K.$$

For the applications of this problem to mathematical programming, they assume f is a continuously differentiable real-valued function on K, invex with respect to η and take $F = \nabla f$. Consider the program (PSK) Min $f(x)$, Subject to $x \in K$.

Parida et al. [195] show that if $\bar{x}$ solves the variational-like inequality problem, then $\bar{x}$ is an optimal solution of the program (PSK). The existence of a solution to the variational-like inequality problem depends on the continuity of η, allowing the Kakutani fixed-point theorem to be invoked.

Secondly, Ponstein [203] established six equivalent definitions of quasi-convexity, of which two apply to differentiable functions. The problem is to know whether the equivalence for these two can be extended to quasi-invexity. In fact, this equivalence is possible under a continuity property of the kernel. First, we recall Ponstein's results: assume $f : X \to R$ differentiable on the open convex set $X \subseteq R^n$. Then f is quasi-convex on X if either

$$f(x^2) \leq f(x^1) \Rightarrow (x^2 - x^1)^T \nabla f(x^1) \leq 0, \tag{2.13}$$

or equivalently

$$f(x^2) < f(x^1) \Rightarrow (x^2 - x^1)^T \nabla f(x^1) \leq 0. \tag{2.14}$$

We recall the definition of a quasi-invex function:

$$f(x^2) \leq f(x^1) \Rightarrow \eta(x^2, x^1)^T \nabla f(x^1) \leq 0. \tag{2.15}$$

Smart [225] gives a condition on η to guarantee that (2.15) is equivalent to

$$f(x^2) < f(x^1) \Rightarrow \eta(x^2, x^1)^T \nabla f(x^1) \leq 0. \tag{2.16}$$

Note that this result subsumes the results of Ponstein, taking $\eta(x^2, x^1) = x^2 - x^1$.

Theorem 2.42. *If the function f satisfies $\eta(x^2, \cdot)$ continuous at x^1 whenever $f(x^2) = f(x^1)$ and f is continuously differentiable, then conditions (2.15) and (2.16) are equivalent.*

Proof. Clearly, if (2.15) holds then (2.16) holds. Conversely, if (2.16) holds we need only establish that $f(x^2) = f(x^1) \rightarrow \eta(x^2, x^1)\nabla f(x^1) \leq 0$. Assume there exist $x^1, x^2 \in X$ (not necessarily distinct) such that $f(x^2) = f(x^1)$ and $\eta(X^2, x^1)\nabla f(x^1) > 0$. Then, by continuity of f, there exists $\bar{\lambda} > 0$, such that $\forall \lambda < \bar{\lambda}, \lambda \neq 0$, we have

$$f(x^1 + \lambda\eta(x^2, x^1)) > f(x^1) = f(x^2).$$

By (2.16), this gives

$$\eta(x^2, x^1 + \lambda\eta(x^2, x^1))^T \nabla f(x^1 + \lambda\eta(x^2, x^1)) \leq 0.$$

Taking limits as $\lambda \downarrow 0$, we obtain by continuity of $\eta(x^2, \cdot)$ and ∇f that $\eta(x^2, x^1)^T \nabla f(x^1) \leq 0$, a contradiction. Thus, if (2.16) holds then (2.15) holds. □

Now, given $f : X \rightarrow R(X \subseteq R^n)$, differentiable and invex, we know that f is invex with respect to $\eta : X \times X \rightarrow R^n$ if for every $x, y \in X$

$$\eta(x, y) = \left\{ \frac{(f(x) - f(y))\nabla f(y)}{\nabla f(y)^T \nabla f(y)} + v; v^T \nabla f(y) \leq 0 \right\},$$

where $\nabla f(y) \neq 0$. Under what conditions on f can a continuous η be chosen subject to the above constraint? For a given $f : X \rightarrow R$, one choice of η is given in the proof of Theorem 2.2:

$$\eta(x, y) = \begin{cases} \frac{(f(x)-f(y))\nabla f(y)}{\nabla f(y)^T \nabla f(y)}, & \text{if } \nabla f(y) \neq 0 \\ 0, & \text{if } \nabla f(y) = 0 \end{cases}$$

For $f : R \to R, f(x) = x^2$, this choice of η gives:

$$\eta(x, y) = \begin{cases} \dfrac{x^2 - y}{2y}, & y \neq 0 \\ 0, & y = 0. \end{cases}$$

Thus, for fixed $x \in R, x \neq 0, \lim_{y\to 0} \eta(x, y) = \lim_{y\to 0} \frac{x^2-y}{2y}$ which does not exist, so $\eta(x, y)$ is not continuous at 0 for any $x \in R - \{0\}$. An alternative choice of η is

$$\eta(x, y) = \begin{cases} 0, & f(y) \geq f(y) \\ \dfrac{(f(x) - f(y))\nabla f(y)}{\nabla f(y)^T \nabla f(y)}, & f(y) < f(y). \end{cases}$$

This η is formed by choosing v so that $v^T\nabla f(y) = f(y) - f(x)$ whenever $f(y) \leq f(x)$ with $\nabla f(y) \neq 0$, choosing $v = 0$ whenever $f(y) > f(x)$, and putting $\eta(x, y) = 0$ when $\nabla f(y) = 0$.

In the simple example above, we obtain

$$\eta(x, y) = \begin{cases} \frac{x^2-y^2}{2y}, & \text{if } |y| > |x| \\ 0, & \text{if } |y| \leq |x| \end{cases}$$

which is continuous in y for each $x \in R$ and furthermore, is continuous on R^2. The following theorem due to Smart [224, 225], gives a sufficient condition for the continuity of the most recent choice of η.

Theorem 2.43. *Let $f : X \to R$ be continuously differentiable and invex. The function $\eta : X \times X \to R^n$ with respect to which f is invex, defined by*

$$\eta(x, y) = \begin{cases} 0, & \text{if } f(x) \geq f(y) \\ \frac{(f(x)-f(y))\nabla f(y)}{\nabla f(y)^T \nabla f(y)}, & \text{if } f(x) < f(y) \end{cases}$$

is continuous if, given y such that $\nabla f(y) = 0$, then for any sequence $\{y^n\}, y^n \to y, \nabla f(y^n) \neq 0$, we have

$$\lim_{n\to\infty} \frac{|f(y) - f(y^n)|}{\|\nabla f(y^n)\|} = 0,$$

where $\|\cdot\|$ is the usual Euclidean norm.

Proof. Let $(x, y) \in X \times X$ and assume $\{x^n\}$ and $\{y^n\}$ are sequences such that $(x^n, y^n) \in X \times X$ $x^n \to x$ and $y^n \to y$. We want to show that $\lim_{n\to\infty} \eta(x^n, y^n) = \eta(x, y)$. Three separate cases must be considered:
(a) From the definition of η, we have

$$\eta(x, y) = \frac{(f(x) - f(y))\nabla f(y)}{\nabla f(y)^T \nabla f(y)}.$$

By continuity of f, there exists an $N \in \aleph$ such that $\forall n \in N$, $f(x^n) < f(y^n)$. Therefore, for $n \geq N$,

$$\eta(x^n, y^n) = \frac{(f(x^n) - f(y^n))\nabla f(y^n)}{\nabla f(y^n)^T \nabla f(y^n)}.$$

by continuity of ∇f, $\lim_{n\to\infty} \eta(x^n, y^n) = \eta(x, y)$.
(b) By hypothesis, $\eta(x, y) = 0$. Again, by continuity of f there exists an $N \in \aleph$ such that $\forall n \in N$, $f(x^n) > f(y^n)$, and thus $\eta(x^n, y^n) = 0$. Therefore, $\lim_{n\to\infty} \eta(x^n, y^n) = \eta(x, y)$.
(c1) by continuity of f and $\nabla f, \forall \epsilon > 0$ there exists an $N \in \aleph$ such that $\forall n \geq N$,

$$|f(x^n) - f(x)| < \frac{\epsilon}{2}, |f(y^n) - f(y)| < \frac{\epsilon}{2}$$

and $f(y^n) \neq 0$.
Now, for $n \geq N$, if $f(x^n) \geq f(y^n)$, then $\eta(x^n, y^n) = 0$ and if $f(x^n) < f(y^n)$, then

$$\eta(x^n, y^n) = \frac{(f(x^n) - f(y^n))\nabla f(y^n)}{\nabla f(y^n) \nabla f(y^n)}.$$

We also have $|f(x^n) - f(y^n)| < \epsilon$. Hence, for $f(x^n) < f(y^n)$,

$$\begin{aligned}
\|\eta(x^n, y^n\|) &= \left\| \frac{(f(x^n) - f(y^n))\nabla f(y^n)}{\nabla f(y^n)^T \nabla f(y^n)} \right\| \\
&= \frac{\|(f(x^n) - f(y^n))\nabla f(y^n)\|}{\|\nabla f(y^n)\|^2} \\
&= \frac{|(f(x^n) - f(y^n))| \cdot \|\nabla f(y^n)\|}{\|\nabla f(y^n)\|^2} \\
&< \frac{\epsilon}{\|\nabla f(y^n)\|}.
\end{aligned}$$

As this holds $\forall \epsilon > 0$ and ∇f continuous, then $\lim_{n\to\infty} \|\eta(x^n, y^n)\| = 0$, so that $\lim_{n\to\infty} \eta(x^n, y^n) = 0 = \eta(x, y)$.
(c2) If $f(x^n) \geq f(y^n)$, then $\eta(x^n, y^n) = 0$. If $f(x^n) < f(y^n)$, then

$$\eta(x^n, y^n) = \frac{(f(x^n) - f(y^n))\nabla f(y^n)}{\nabla f(y^n)^T \nabla f(y^n)}$$

and so

$$\|\eta(x^n, y^n)\| = \frac{|f(x^n) - f(y^n)}{\|\nabla f(y^n)\|}.$$

Note that $\nabla f(y) = 0$ and $f(x) = f(y)$ implies that x and y are global minimizers, so that when $f(x^n) < f(y^n)$, we have $f(y) = f(x) \leq f(x^n) < f(y^n)$. This gives

$$|f(y) - f(y^n)| \geq |f(x^n) - f(y^n)|$$

and hence

$$\|\eta(x^n, y^n)\| \le \frac{|f(x^n) - f(y^n)|}{\|\nabla f(y^n)\|}.$$

Now if there exists an $N \in \aleph$ such that $\forall n \ge N, f(x^n) \ge f(y^n)$, then we immediately have $\lim_{n\to\infty} \eta(x^n, y^n) = 0 = \eta(x, y)$. Otherwise, there exists a sub-sequence $\{y_i^n\}$ or $\{y^n\}$ such that $y_i^n \to y, f(x_i^n) < f(y_i^n)$, and $\nabla f(y_i^n) \neq 0$. By the hypothesis of the theorem

$$\lim_{n\to\infty} \|\eta(x_i^n, y_i^n)\| \le \lim_{n_i\to\infty} \frac{|f(y) - f(y_i^n)|}{\|\nabla f(y_i^n)\|} = 0.$$

Therefore, $\lim_{n\to\infty} \eta(x^n, y^n) = 0 = \eta(x, y)$.

□

The next result gives a simple second-order sufficient condition for the limit property of Theorem 2.43 to be satisfied.

Theorem 2.44. *Let $f : X \to R$ be invex and assume $\nabla f(y) = 0$. If f is twice continuously differentiable in some open neighborhood of y and $\nabla^2 f(y)$ is positive definite, then for any sequence $y^n, y^n \in X, y^n \to y, \nabla f(y^n) \neq 0$, we have*

$$\lim_{n\to\infty} \frac{|f(y) - f(y^n)|}{\|\nabla f(y^n)\|} = 0.$$

Proof. As f is twice continuously differentiable in some open neighborhood of y, and y is a global and therefore local minimizer with $\nabla^2 f(y)$ positive definite, then by continuity of $\nabla^2 f$, there exists some $\epsilon > 0$ such that $\forall x \in N(y, \epsilon)$ (the open ball of radius ϵ centered at y), f is twice continuously differentiable at x and $\nabla^2 f(x)$ is positive semi-definite. Now, consider $x \in N(y, \epsilon), x \neq y$, and define $g : [0, 1] \to R$ by $g(t) = f(y + t(x - y))$; g is twice differentiable, and its derivatives are given by

$$g' = (x - y)^T \nabla f(y + t(x - y)), \quad g''(t) = (x - y)^T \nabla^2 f(y + t(x - y))(x - y).$$

Let $t \in [0, 1]$. By the Mean Value Theorem, there exists $\xi \in [0, t]$ such that

$$g'(\xi) = \frac{g(t) - g(0)}{t},$$

that is, $g(t) - g(0) = tg'(\xi)$. But, as $\nabla f(x)$ is positive semi-definite on $\aleph(y, \epsilon)$, then $g'' \ge 0$ on $[0, 1]$. Hence g' is a non-decreasing function, so $g'(\xi) \le g'(t)$. Therefore, $g(t) - g(0) \le tg'(t)$. In particular, $g(1) - g(0) \le g'(1)$; that is, $f(x) - f(y) \le (x - y)^T \nabla f(x)$. Since the invexity of f implies that $f(x) \ge f(y)$, then by Cauchy–Schwarz inequality,

$$|f(x) - f(y)| \le \left|(x - y)^T \nabla f(x)\right| \le \|(x - y)\| \cdot \|\nabla f(x)\|.$$

Thus, if $\nabla f(x) \neq 0$, then

$$\frac{|f(x) - f(y)|}{\|\nabla f(x)\|} \le \|x - y\|.$$

Now, for any sequence $\{y^n\}, y^n \in X, y^n \to y, \nabla f(y^n) \neq$, there exists $N \in \aleph$ such that $\forall n \geq N$, we have $y^n \in N(y, \epsilon)$ and consequently

$$\frac{|f(y) - f(y^n)|}{\|\nabla f(y^n)\|} \leq \|y^n - y\| .$$

Therefore, by the squeeze principle,

$$\lim_{n \to \infty} \frac{|f(y) - f(y^n)|}{\|\nabla f(y^n)\|} = 0.$$

□

The limit property of Theorem 2.44 does not hold for all continuously differentiable invex functions. In the following example, due to Smart [224, 225], the property does not hold. Furthermore, for invex functions of one variable if there exists $\bar{x} \in X$ such that $\bar{x}$ is a strict minimum and $\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x})}{f'(x)} \neq 0$, then there is no continuous η with respect to which f is invex [225].

Example 2.45. Define $f : R \to R$ by

$$f(x) = \begin{cases} 0, & x = 0 \\ \dfrac{n^2+n+1}{n+1}x^2 + \dfrac{-2n^2-2n-1}{(n+1)^2}x + \dfrac{4n^2+5n+2}{4(n+1)^3}, & \\ \qquad \dfrac{1}{n+1} \leq x \leq \dfrac{2n+1}{2n(n+1)}, & n = 1, 2, \ldots \\ \dfrac{1-n^2}{n}x^2 + \dfrac{2n^2-1}{n^2}x + \dfrac{-4n^2+n+1}{4n^3}, \dfrac{2n+1}{2n(n+1)} \leq x \leq \dfrac{1}{n}, & n = 1, 2, \ldots \\ x - \dfrac{1}{2}, & x \geq 1 \\ f(-x), & x < 0. \end{cases}$$

It is very easy to check that f is continuously differentiable, with $f'(y) = 0$ if and only if $y = 0$, which is a global minimizer. Consider the sequence y^n with $y^n = \frac{1}{n}, n = 1, 2, \ldots$. We have

$$f(y^n) = \frac{n+1}{4n^3} \quad \text{and} \quad f'(y^n) = \frac{1}{n^2}.$$

Therefore,

$$\lim_{n \to \infty} \frac{|f(y) - f(y^n)|}{|f'(y^n)|} = \lim_{n \to \infty} \frac{n^2(n+1)}{4n^3} = \frac{1}{4}.$$

Therefore, for this example, there is no choice of η which is continuous.

Springer

http://www.springer.com/978-3-540-78561-3

Invexity and Optimization
Mishra, S.K.; Giorgi, G.
2008, X, 266 p., Hardcover
ISBN: 978-3-540-78561-3