Vietnam Journal of Mathematics **32**:1 (2004) 87–94

Vietnam Journal
of
MATHEMATICS

Is Invexity Weaker Than Convexity?

Hoang Xuan Phu

Institute of Mathematics, 18 Hoang Quoc Viet Road, 10307 Hanoi, Vietnam

Received June 26, 2003
Revised October 14, 2003

Abstract. Invexity is useful only if all the needed functions are invex with respect to the same directional function η. If one of them does not satisfy this condition, then it is hard to assert something, even if it is convex. Assertions made in such an incomplete situation might be false or the invexity assumption might be redundant. Since convex functions are rarely accepted as invex partners with respect to a given function η, it is not safe to say that invexity is weaker than convexity. Some wrong theorems of D. V. Luu and P. T. Kien (2000) are analyzed in this paper to illustrate this opinion.

1. Introduction

It is well known that the Kuhn-Tucker conditions for a constrained minimization problem become also sufficient for a (global) minimum if the functions are assumed to be convex. It remains true if the functions satisfy certain generalized convex properties. In 1981, Hanson [5] introduced a class of differentiable functions satisfying

$$\phi(x) - \phi(x') \geq \langle \nabla\phi(x'), \eta(x, x') \rangle \quad \text{for all} \quad x,\, x' \in C, \tag{1}$$

for some arbitrary given function η defined on $C \times C$, to prove the following sufficient condition:

Let f_0, f_1,...,f_m be differentiable functions on $C \subset \mathbb{R}^n$ satisfying (1) *for some η defined on $C \times C$. If there exist $x^* \in C$ and $y^* \in \mathbb{R}^m$ satisfying the Kuhn-Tucker conditions*

$$\begin{aligned} &\nabla f_0(x^*) + \nabla\big(y^* \bar{f}(x^*)\big) = 0, \\ &y^* \geq 0, \quad \bar{f}(x^*) \leq 0, \quad y^* \bar{f}(x^*) = 0, \end{aligned} \tag{2}$$

where $\bar{f} = (f_1, ..., f_m)^T$, then

$$f_0(x^*) = \min\{f_0(x) : x \in C,\ \bar{f}(x) \le 0\}.$$

Functions satisfying (1) are called *invex* by Craven [1]. Although its proof is not longer than the theorem, Hanson's result is interesting enough to initiate a lively research direction concerned with invexity and its generalizations. Many papers have been published, but not all of them are eligible.

Since every convex function is invex with respect to $\eta(x, x') = x - x'$, invexity is always considered as a generalization of convexity. One even says that invexity is much weaker than convexity. This is only partially true. In fact, it is senseless to discuss invexity without mentioning the function η. After fixing the function η, the class of invex functions with respect to this η is not as large as one used to believe. Especially, convex functions are often not invex with respect to η if $\eta(x, x') \ne x - x'$. From this point of view, invexity is not weaker than convexity, and convexity is not stronger than invexity. If this simple fact is always well understood, then some errors can be avoided.

A more detailed explanation on the fact just mentioned is given in Sec. 2. For illustration, some errors made in [7] are analyzed in Secs. 3 and 4.

2. Convexity as a Rarely Accepted Member of Invexity

Let

$$\varphi(x) = x_1 - x_2^2, \quad \psi(x) = -x_1, \quad x = (x_1, x_2) \in \mathbb{R}^2. \tag{3}$$

These functions are continuously differentiable on $\mathbb{R}^2$ and $\nabla\varphi(x) = (1, -2x_2)$ and $\nabla\psi(x) = (-1, 0)$. For

$$\eta(x, y) = x - y + r(x, y), \quad r(x, y) = \big(-(x_2 - y_2)^2, 0\big), \tag{4}$$

for all $x = (x_1, x_2) \in \mathbb{R}^2$ and $y = (y_1, y_2) \in \mathbb{R}^2$, we have

$$\begin{aligned}\langle \nabla\varphi(y), \eta(x, y)\rangle &= x_1 - y_1 - (x_2 - y_2)^2 - 2y_2(x_2 - y_2)\\ &= x_1 - y_1 - x_2^2 + y_2^2\\ &= \varphi(x) - \varphi(y),\end{aligned}$$

which implies that both φ and $-\varphi$ are invex on $\mathbb{R}^2$ with respect to η.

Let us consider the problems

(P_1) minimize $\varphi(x)$, subject to $\psi(x) \le 0$,

and

(P_2) minimize $\psi(x)$, subject to $\varphi(x) \le 0$.

For $x^* = (0, 0) \in \mathbb{R}^2$, we have $\varphi(x^*) = \psi(x^*) = 0$ and

$$\nabla\varphi(x^*) + \nabla\psi(x^*) = (1, 0) + (-1, 0) = (0, 0).$$

Thus, for $y^* = 1$, the Kuhn-Tucker conditions (2) are satisfied for both problems.

Since $\varphi(x^*) = 0$ and $\varphi(x) = -x_2^2 < 0$ for all $x \in \mathbb{R}^2$ satisfying $\psi(x) = -x_1 = 0$ and $x \ne x^*$, x^* is not a local minimizer of Problem (P_1).

Since $\psi(x^*) = 0$ and $\psi(x) = -x_1 < 0$ for all $x \in \mathbb{R}^2$ satisfying $\varphi(x) = x_1 - x_2^2 = 0$ and $x \ne x^*$, x^* is not a local minimizer of Problem (P_2).

We have seen that the invexity of φ and the convexity of ψ are not sufficient for a Kuhn-Tucker point of (P_1) or (P_2) to be a local minimizer. This situation may occur even if ψ is a strictly convex function. For instance, if

$$\psi(x) = -x_1 + \frac{1}{2}(x_1^2 + x_2^2)$$

then

$$\psi(x^*) = 0, \quad \nabla\varphi(x^*) + \nabla\psi(x^*) = (1,0) + (-1,0) = (0,0)$$

for $x^* = (0,0)$, and for all $x \in \mathbb{R}^2$ satisfying $x_1 = x_2^2$ and $0 < x_1 < 1$, we have $-x_1 + x_1^2 < 0$ and

$$\psi(x) = -x_1 + \frac{1}{2}(x_1^2 + x_2^2) = \frac{1}{2}(-x_1 + x_1^2) < 0,$$

i.e., the Kuhn-Tucker point $x^* = (0,0)$ is not a local minimizer of Problem (P_2). It is easy to show that this x^* is also a Kuhn-Tucker point of (P_1), which is not a local minimizer.

Note that φ stated in (3) is invex with respect to several η which are different from the one given in (4). For instance, for

$$\eta(x,y) = -\left(|x_1| + x_2^2 + |y_1|\right)\nabla\varphi(y)$$

and for all $x = (x_1, x_2) \in \mathbb{R}^2$, $y = (y_1, y_2) \in \mathbb{R}^2$, we have

$$\begin{aligned}\varphi(x) - \varphi(y) &= x_1 - x_2^2 - y_1 + y_2^2 \\ &\geq -\left(|x_1| + x_2^2 + |y_1|\right) \\ &\geq -\left(|x_1| + x_2^2 + |y_1|\right)\left(1 + 4y_2^2\right) \\ &= \langle \nabla\varphi(y), \eta(y,x)\rangle,\end{aligned}$$

which implies that φ is invex on $\mathbb{R}^2$ with respect to this η. But there exists no function η such that both functions φ and ψ are invex with respect to η, otherwise, due to Hanson's theorem, $x^* = (0,0)$ should be a minimizer of (P_1) and of (P_2).

The above examples show:

(a) It is too short to say that invexity is weaker than convexity. On the contrary, for a given invex function which is not convex, its invexity property is so strong that a lot of convex functions cannot be accepted as its invex partners with respect to the same function η.

(b) Invexity is useful only if *all the needed functions* are invex with respect to the same function η. If some of them do not satisfy this condition, then it is hard to assert something, even if the outsiders are convex or strictly convex. If only some functions are invex with respect to a common η and the others are not, then one must be careful because the desired assertion may remain true without this invexity, or it is possibly false.

What does *"all the needed functions"* mean? If invexity is used to show that a Kuhn-Tucker point is a minimizer, then the objective function and all the constraint functions must be invex respect to the same η. If invexity is used to show $\lambda > 0$, where λ is the multiplier in the Kuhn-Tucker conditions corresponding to the objective function, then only all constraint functions must be invex respect to the same η.

In the next sections, we discuss some errors which are easily recognized with the help of the preceding remarks.

3. An Error Related to Optimization with Pseudo-Invex Objective Function

Let X be a Banach space, Q_1,..., Q_m, Q_{m+1} be subsets of X, and f be a real-valued function defined on X. Consider the problem

$$(P_3) \qquad \text{minimize} \quad f(x), \quad \text{subject to} \quad x \in Q,$$

where $Q = \bigcap_{i=1}^{m+1} Q_i$. Let $x^* \in Q$ and let

- K_0 be the cone of decreasing directions v of f at x^* defined by: There are a neighborhood U of v and numbers $\alpha < 0$ and $\epsilon_0 > 0$ such that $f(x^* + \epsilon u) \leq f(x^*) + \epsilon\alpha$ for every $\epsilon \in (0, \epsilon_0)$ and $u \in U$;
- $K_1, ..., K_m$ be the cones of admissible directions v of the inequality-type constraints $Q_1, ..., Q_m$ at x^* defined by: There are a neighborhood U of v and a number $\epsilon_0 > 0$ such that $x^* + \epsilon u \in Q_i$ for every $\epsilon \in (0, \epsilon_0)$ and $u \in U$, where $i = 1, ..., m$, respectively;
- K_{m+1} be the cone of tangent directions of the equality-type constraint Q_{m+1} at x^*.

Denote by K_i^* the dual cone of K_i:

$$K_i^* = \{\xi \in X^* : \langle \xi, x \rangle \geq 0, \ \ \forall x \in K_i\}.$$

Let us consider the following sufficient condition:

Assertion 1. (Luu and Kien [7], Theorem 3.1)
Assume that

(a) *The function f is locally Lipschitz at x^* and it has directional derivative at x^* in any directions; f is pseudo-invex at x^* on Q with the scale function*

$$\eta(x, x^*) = x - x^* + r(x, x^*),$$

where

$$\|r(x, x^*)\| / \|x - x^*\| \to 0 \quad \textit{whenever} \quad \|x - x^*\| \to 0;$$

(b) *Q_1,..., Q_m, and Q_{m+1} are convex sets such that there exists*

$$\hat{x} \in \bigcap_{i=1}^{m} (\operatorname{int} Q_i) \cap Q_{m+1},$$

where $\operatorname{int} Q_i$ *denotes the interior of Q_i;*

(c) *There exist $\xi \in K_i^*$ $(i = 0, 1, ..., m+1)$, not all zero, such that*

$$\xi_0 + \xi_1 + ... + \xi_{m+1} = 0.$$

Then x^* is a local minimum of f over Q.

Recall that a function $f : X \to \mathbb{R}$ having directional derivatives at $y \in Q \subset X$ is said to be *pseudo-invex at* y on Q with respect to $\eta : X \times X \to X$ if, for all $x \in Q$,

$$f'(y;\eta(x,y)) \geq 0 \quad \text{implies} \quad f(x) - f(y) \geq 0.$$

Obviously, this kind of pseudo-invexity follows from the invexity defined for Fréchet differentiable functions as follows: f is called *invex at* y with respect to η if

$$f(x) - f(y) \geq f'(y)\eta(x,y) \quad \text{for all } \; x \in X.$$

and f is named *invex* with respect to η if it is invex at every $y \in X$ (see [1]).

In Assertion 1, only the objective function f is assumed to be pseudo-invex, while the other constraint functions defining Q_1,..., Q_m, and Q_{m+1} are convex, i.e., not necessarily pseudo-invex or invex with respect to the same η. Thus, according to our remarks in Sec. 2, the assumption of pseudo-invexity could be redundant, or the assertion could be wrong. In fact, to show that the assertion is wrong, we now present a simple example with an invex objective function.

Counter example 1. Consider a concrete example of (P_3), where $m = 1$,

$$f(x) = x_1 - x_2^2, \quad x = (x_1, x_2),$$
$$Q_1 = \{(x_1,x_2) \in \mathbb{R}^2 : x_1 \geq -1\},$$
$$Q_2 = \{(x_1,x_2) \in \mathbb{R}^2 : x_1 = 0\},$$

and $x^* = (0,0) \in Q = Q_1 \cap Q_2 = Q_2$. As shown in Sec. 2, $f(x) = x_1 - x_2^2$ is invex on $\mathbb{R}^2$ with respect to

$$\eta(x,y) = x - y + r(x,y), \quad r(x,y) = \big(-(x_2-y_2)^2, 0\big).$$

Since

$$0 \leq \frac{\|r(x,y)\|}{\|x-y\|} = \frac{(x_2-y_2)^2}{\big((x_1-y_1)^2 + (x_2-y_2)^2\big)^{\frac{1}{2}}} \leq \frac{(x_2-y_2)^2}{|x_2-y_2|} = |x_2-y_2|$$

for $x_2 \neq y_2$ and since $\|r(x,y)\| = 0$ for $x_2 = y_2$, we have

$$\|r(x,y)\|/\|x-y\| \to 0 \quad \text{as} \quad \|x-y\| \to 0.$$

Obviously, Q_1 and Q_2 are convex and $x^* \in \operatorname{int} Q_1 \cap Q_2$. Since $K_0 = \{(v_1,v_2) \in \mathbb{R}^2 : v_1 < 0\}$, $K_1 = \mathbb{R}^2$, and $K_2 = \{(v_1,v_2) \in \mathbb{R}^2 : v_1 = 0\}$, for $\xi_0 = (-1,0) \in K_0^*$, $\xi_1 = (0,0) \in K_1^*$, and $\xi_2 = (1,0) \in K_2^*$, it holds $\xi_0 + \xi_1 + \xi_2 = 0$. Hence, all the assumptions of Assertion 1 are satisfied. Therefore, it ensures that $x^* = (0,0)$ is a local minimizer. But, on the admissible set $Q = \{(x_1,x_2) \in \mathbb{R}^2 : x_1 = 0\}$, the objective function is $f(x) = -x_2^2$, which has no local minimizer. This example shows that Assertion 1, i.e., Theorem 3.1 in [7], is false.

4. Error Related to Optimization with Invex Constraint Functions

Let f be a real-valued function defined on a Banach space X, and g a mapping from X into a Banach space Y. Let K be a closed convex cone in Y. Consider the problem

$$(P_4) \qquad \begin{cases} \text{minimize } \ f(x), \\ \text{subject to } g(x) \in -K. \end{cases}$$

For $x^* \in M = \{x \in X : g(x) \in -K\}$, denote

$$K_{g(x^*)} = K + \{\lambda g(x^*) : \lambda \in \mathbb{R}\},$$
$$L_M(x^*) = \{v \in X : g'(x^*)v \in -K_{g(x^*)}\}.$$

Suppose that f and g are Fréchet differentiable of first and second-order at x^*.

For

$$L(x, y^*, \lambda) = \lambda f(x) + \langle y^*, g(x)\rangle \quad (x \in X,\ \lambda \in R,\ y^* \in Y^*),$$

assume that

$$\exists y^* \in Y^* :\ L_x(x^*, y^*, \lambda) = 0,\ \langle y^*, g(x^*)\rangle = 0, \tag{5}$$

and

$$\begin{aligned} &\exists \delta > 0\ \exists \beta > 0 : L_{xx}(x^*, y^*, 1)(v, v) \geq \delta \|v\|^2 \\ &\text{for all } v \in L_M(x^*) \cap \{v \in X : \langle y^*, g'(x^*)v\rangle \leq \beta \|v\|\}. \end{aligned} \tag{6}$$

Assertion 2. (Luu and Kien [7], Theorem 4.4)
Let $x^* \in M$. Assume that (5) and (6) hold. Suppose, in addition, that the mapping g is K-invex at x^* on M with a scale mapping η satisfying

$$\eta(x, x^*) = x - x^* + r(x, x^*) \quad \textit{for all } \ x \in M,$$

where $\|r(x, x^*)\|/\|x - x^*\| \to 0$ as $\|x - x^*\| \to 0$. Then there exist number $\alpha > 0$ and $\rho > 0$ such that

$$f(x) \geq f(x^*) + \alpha\|x - x^*\| \quad \textit{for all } \ x \in M \cap B(x^*; \rho), \tag{7}$$

where $B(x^*; \rho) = \{x \in X : \|x - x^*\| < \rho\}$.

Recall that a Fréchet differentiable function $g : X \to Y$ is said to be K-invex at $y \in M$ on M with respect to $\eta : X \times X \to X$ if

$$g(x) - g(y) - g'(y)\eta(x, y) \in K \quad \text{for all } \ x \in M.$$

Function g is called K-invex with respect to η if this property holds for all $x \in X$ and $y \in X$.

Assertion 3. (Luu and Kien [7], Theorem 4.5)
Let $x^* \in M$. Assume that (5) and (6) hold. Suppose, furthermore, that the following condition is fulfilled:

$$0 \in \text{ int}\big(g'(x^*)X + K\big). \tag{8}$$

Then there exist number $\alpha > 0$ and $\rho > 0$ such that (7) holds.

Assertion 2 states a sufficient condition for a special kind of local minima. For this purpose, if invexity is really useful then it should be assumed both for the objective and the constraint functions. But, in Assertion 2, the K-invexity is only assumed for the constraint function g and not for the objective function f.

Due to the remarks in Sec. 2, the assumption on K-invexity could be redundant, or the assertion could be wrong. Indeed, we will show that both cases appear, i.e., the invexity is redundant and the assertion is wrong.

Counter example 2. Let $X = Y = \mathbb{R}$, $K = \mathbb{R}_+$, and $f(x) = g(x) = x$. For $x^* = 0$, $\lambda = 1$, and $y^* = -1$, (5) is fulfilled. Since $K_{g(x^*)} = K = \mathbb{R}_+$ and $M = -\mathbb{R}_+$, we have

$$L_M(x^*) = \{v \in \mathbb{R} : v \le 0\} = -\mathbb{R}_+$$

and, for $\beta = 1/2$,

$$\{v \in X : \langle y^*, g'(x^*)v\rangle \le \beta\,\|v\|\} = \{v \in \mathbb{R} : -v \le |v|/2\} = \mathbb{R}_+,$$

which yields

$$L_M(x^*) \cap \{v \in \mathbb{R} : \langle y^*, g'(x^*)v\rangle \le \beta\,\|v\|\} = \{0\}.$$

Hence, it follows from $L(x, y^*, 1) = x - x = 0$ that (6) is satisfied for an arbitrary $\delta > 0$ and $\beta = 1/2$. Since g is linear, and therefore K-invex (everywhere) with respect to $\eta(x, y) = x - y$, Assertion 2 implies that $x^* = 0$ is a local minimizer. Since $g'(x^*)X + K = \mathbb{R} + \mathbb{R}_+ = \mathbb{R}$, (8) is fulfilled. Therefore, Assertion 3 also yields the same. But the function $f(x) = x$ has no local minimizer in $M = -\mathbb{R}_+$. Hence, both Assertion 2 and Assertion 3, i.e., Theorems 4.4 and 4.5 in [7], are false.

In the above example, both objective and constraint functions f and g are linear. Nevertheless, it has no effect. The invexity assumption is really useless and redundant here.

Since there are properties which are only true in higher dimensional spaces, i.e., only if $\dim X \ge 2$, let us describe an example with $X = \mathbb{R}^2$. Moreover, unlike the previous example, in the next example the set $L_M(x^*) \cap \{v \in X : \langle y^*, g'(x^*)v\rangle \le \beta\,\|v\|\}$ has nonempty interior.

Counter example 3. Let $X = \mathbb{R}^2$, $Y = \mathbb{R}$, $K = \mathbb{R}_+$, $f(x) = x_1^2 - (x_2 + 1)^2$, and $g(x) = -x_2$. For $x^* = (0, 0)$, $\lambda = 1$, and $y^* = -2$, (5) is fulfilled. Since $K_{g(x^*)} = K = \mathbb{R}_+$ and $M = \{x \in \mathbb{R}^2 : x_2 \ge 0\}$, we have

$$L_M(x^*) = \{v \in \mathbb{R}^2 : -v_2 \le 0\} = \{v \in \mathbb{R}^2 : v_2 \ge 0\}$$

and, for $\beta = 1$,

$$\begin{aligned}\{v \in X : \langle y^*, g'(x^*)v\rangle \le \beta\,\|v\|\} &= \{v \in \mathbb{R}^2 : (-2)(-v_2) \le \|v\|\}\\ &= \{v \in \mathbb{R}^2 : 2v_2 \le (v_1^2 + v_2^2)^{1/2}\},\end{aligned}$$

which yields

$$L_M(x^*) \cap \{v : \langle y^*, g'(x^*)v\rangle \le \beta\,\|v\|\} = \{v \in \mathbb{R}^2 : v_2 \ge 0,\ 3v_2^2 \le v_1^2\}.$$

Since $L(x, y^*, 1) = x_1^2 - x_2^2 - 1$, it follows from $3v_2^2 \le v_1^2$ that

$$L_{xx}(x^*, y^*, 1)(v, v) = 2(v_1^2 - v_2^2) = v_1^2 + (v_1^2 - 2v_2^2) \ge v_1^2 + v_2^2 = \|v\|^2,$$

i.e., (6) is satisfied for $\delta = \beta = 1$. Since g is linear, and therefore K-invex (everywhere) with respect to $\eta(x,y) = x - y$, Assertion 2 implies that $x^* = (0,0)$ is a local minimizer. Since

$$g'(x^*)X + K = \{-x_2 : (x_1, x_2) \in \mathbb{R}^2\} + \mathbb{R}_+ = \mathbb{R} + \mathbb{R}_+ = \mathbb{R},$$

(8) is fulfilled. Therefore, Assertion 3 also yields the same. But $f(x^*) = -1$ and $f(x) = -(x_2+1)^2 < -1$ for all $x \in M$ satisfying $x_1 = 0$ and $x_2 > 0$, i.e., $x^* = (0,0)$ cannot be a local minimizer. Thus, this example shows again that both Assertion 2 and Assertion 3, i.e., Theorems 4.4 and 4.5 in [7], are false, even when $X = \mathbb{R}^2$.

5. Concluding Remarks

The counter examples in Secs. 3 and 4 show that all the main results of Luu and Kieu [7] are wrong. While demonstrating the power of invexity and its generalizations, several other essential errors have been made in [2-4, 6] by misplacing the originally interesting concept of Hanson [5]. Obviously, such an inflation does not heighten the image of invexity. Some of these errors are analyzed in [8] and [9].

Acknowledgment. The author thanks Prof. Dr. Nguyen Dong Yen and the referee for valuable comments.

References

1. B. D. Craven, Invex functions and constrained local minima, *Bull. Austral. Math. Soc.* **24** (1981) 357–366.
2. N. X. Ha, Optimality conditions for a class of nonsmooth optimization problems (Vietnamese), *Journal of Science and Technique of Military Technical Academy*, No. **100** (2002) 12–19.
3. N. X. Ha and D. V. Luu, Invexity of supremum and infimum functions, *Bull. Austral. Math. Soc.* **65** (2002) 289–306.
4. N. X. Ha and D. V. Luu, Sufficient conditions for invexity, *Bull. Austral. Math. Soc.* **68** (2003) 113–125.
5. M. A. Hanson, On sufficiency of the Kuhn-Tucker conditions, *J. Math. Anal. Appl.* **80** (1981) 545–550.
6. D. V. Luu and N. X. Ha, An invariant property of invex functions and applications, *Acta Math. Vietnam.* **25** (2000) 181–193.
7. D. V. Luu and P. T. Kien, Sufficient optimality conditions under invexity hypotheses, *Vietnam J. Math.* **28** (2000) 227–236.
8. H. X. Phu, On a necessary optimality condition with invexity, *Acta Math. Vietnam.* (to appear).
9. H. X. Phu, On some badly-solved problems with invexity, *Acta Math. Vietnam.* (to appear).