J. Austral. Math. Soc. Ser. B 28 (1986), 1–9

WHAT IS INVEXITY?

A. BEN-ISRAEL[1] AND B. MOND[2]

(Received 8 February 1985; revised 28 May 1985)

Abstract

Recently it was shown that many results in Mathematical Programming involving convex functions actually hold for a wider class of functions, called *invex*. Here a simple characterization of invexity is given for both constrained and unconstrained problems. The relationship between invexity and other generalizations of convexity is illustrated. Finally, it is shown that invexity can be substituted for convexity in the saddle point problem and in the Slater constraint qualification.

1. Introduction

In [13], Zang, Choo and Avriel studied functions whose stationary points are global minima and applied their results to Mathematical Programming.

A few years later, Hanson [5] considered differentiable functions from R^n into R for which there exists a vector function $\eta(x, u) \in R^n$ such that

$$f(x) - f(u) \geqslant [\eta(x, u)]'\nabla f(u), \tag{1}$$

where ∇ denotes the gradient. Clearly, differentiable convex functions satisfy (1), since in that case one can take $\eta(x, u) = x - u$. Hanson [5] showed that if, instead of the usual convexity conditions, the objective function and each of the constraints of a nonlinear programming problem all satisfy (1) for the same $\eta(x, u)$, weak duality and sufficiency of the Kuhn-Tucker conditions still hold.

Hanson's paper inspired a great deal of additional work. For example, in [1], Craven called functions satisfying (1) *invex* and established duality theorems for

[1] Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, U.S.A.
[2] Department of Mathematics, La Trobe University, Bundoora, Vic. 3083, Australia.

fractional programs with such functions. In [12], Mond and Hanson extended the concept of invexity to polyhedral cones while in [2] Craven defined it for more general cones and gave second-order conditions for invexity. In [3], Craven and Glover showed that the class of invex functions is equivalent to the class of functions whose stationary points are global minima.

A number of other forms of invexity have also recently appeared. In [11], Martin defines *Kuhn-Tucker invexity* and *weak duality invexity*, in [6] Hanson and Mond speak of Type I and Type II functions, while in [7] Jeyakumar defines strong and weak invex functions.

Our purpose here is first to give a simple characterization of invexity for both constrained and unconstrained problems and, secondly, to present some new results for invex functions. These include a sufficient condition for functions to be invex and proofs that invexity can be substituted for convexity in the saddle point problem and in the Slater constraint qualification.

2. Unconstrained optimization

Some of the results in this section were first given by Craven and Glover [3], where invexity for quasi-differentiable functions was defined, and by Jeyakumar [7], where strong and weak invexity were discussed. Some of their results are repeated here for completeness, and in order to present them in a simpler setting. The next theorem, for which we provide a simple proof, was first stated in [3].

THEOREM 1. *f is invex if and only if every stationary point is a global minimum.*

PROOF. Clearly, if f is invex, then $\nabla f(u) = 0$ implies $f(x) \geqslant f(u)$. Assume now that

$$\nabla f(u) = 0 \Rightarrow f(x) \geqslant f(u).$$

If $\nabla f(u) = 0$, take $\eta(x, u) = 0$. If $\nabla f(u) \neq 0$, take

$$\eta(x,u) = \frac{[f(x) - f(u)]}{[\nabla f(u)]'\nabla f(u)} \nabla f(u).$$

COROLLARY. *If f has no stationary points, then f is invex.*

Hanson [5] noted that there are simple extensions of (1) to

$$[\eta(x,u)]'\nabla f(u) \geqslant 0 \Rightarrow f(x) - f(u) \geqslant 0 \tag{2}$$

and

$$f(x) - f(u) \leqslant 0 \Rightarrow [\eta(x,u)]'\nabla f(u) \leqslant 0. \tag{3}$$

https://doi.org/10.1017/S0334270000005142 Published online by Cambridge University Press

Functions satisfying (2) and (3) were subsequently [1] called *pseudo-invex* and *quasi-invex* respectively.

It is clear, by taking $\eta(x, u) = (x - u)$, that convex, pseudo-convex and quasi-convex functions are, respectively, invex, pseudo-invex and quasi-invex. It is also obvious from the definitions that invex functions are both pseudo-invex and quasi-invex. It follows easily, from Theorem 1, that both pseudo-convex and pseudo-invex functions are invex. Thus (unlike pseudo-convex and convex) there is no distinction between pseudo-invex and invex functions.

Theorem 1 has the following analogue for pseudo-convex functions.

PROPOSITION. *A differentiable function f is pseudo-convex if and only if for all x, u*

$$(x - u)'\nabla f(u) = 0 \Rightarrow f(u) \leqslant f(u + t(x - u)), \quad \textit{for all } t > 0. \tag{4}$$

PROOF. *Only if*. Obvious from the definition of pseudo-convexity. Here (4) holds for all real t.

If. Suppose f is not pseudo-convex; that is, there exists x, u such that

$$(x - u)'\nabla f(u) \geqslant 0 \quad \text{and} \quad f(x) < f(u).$$

If $(x - u)'\nabla f(u) = 0$, then (4) is contradicted. If $(x - u)'\nabla f(u) > 0$, then there exists v which maximizes f on the line segment from u to x. Thus $\nabla f(v) = 0$ and, therefore

$$(x - v)'\nabla f(v) = 0 \quad \text{and} \quad f(v) \geqslant f(u) > f(x) = f(v + 1(x - v)),$$

contradicting (4).

Although pseudo-convex and pseudo-invex functions are both invex, this is not the case with quasi-convex and quasi-invex functions.

EXAMPLE 1. $f(x) = x^3$.

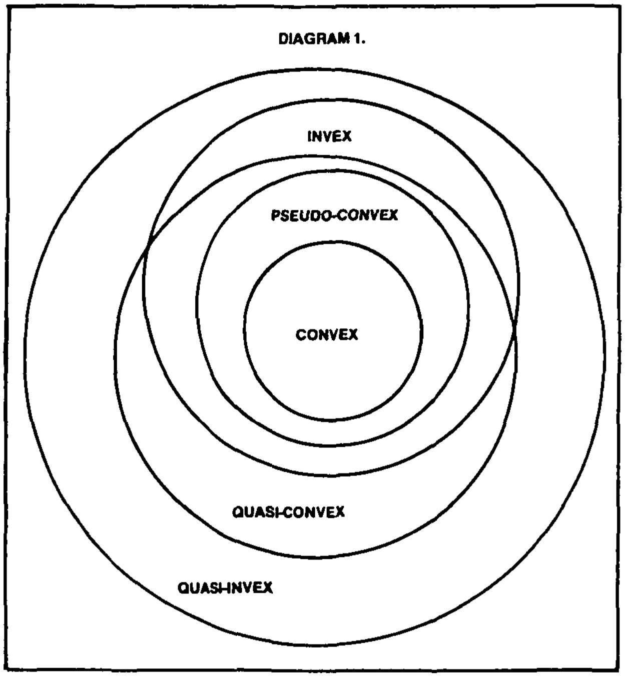
f is not invex, since the stationary point $x = 0$ is not a global minimum. As is known, x^3 is quasi-convex and hence, also quasi-invex.

Although an invex function is quasi-invex, it need not be quasi-convex.

EXAMPLE 2. $f(x) = x_1^3 + x_1 - 10x_2^3 - x_2$.

Since there are no stationary points, f is invex. Taking $u = (0,0)$, $x_1 = 2$, $x_2 = 1$, gives $f(x) - f(x) < 0$ but $(x - u)'\nabla f(u) > 0$ so that f is not quasi-convex.

The relationship between the different forms of convexity and invexity is illustrated in Diagram 1.



REMARK. In [8], Kaul and Kaur call functions satisfying (1), (2) and (3), η-convex, η-pseudoconvex and η-quasiconvex, respectively. Their result that "Every η-convex function is η-pseudoconvex for the same η but the converse is not true" does not contradict the result stated here and in [3] that all pseudo-invex functions are invex. An η-pseudoconvex function may not be η-convex *for the same* η but will be η-convex for *some* η. Indeed, it is easy to see that the counterexample from [8] of a function that is η-pseudoconvex but not η-convex for the same η is, indeed, η-convex for the η described in the proof of our Theorem 1.

We now give a sufficient condition for invexity.

THEOREM 2. *If f is differentiable and there exists an n-dimensional vector function $\eta(x, u)$ such that*

$$f(u + \lambda\eta(x,u)) \leqslant \lambda f(x) + (1 - \lambda)f(u), \qquad 0 \leqslant \lambda \leqslant 1, \tag{5}$$

then f is invex.

PROOF. Rewrite (5) as

$$f(u + \lambda\eta(x,u)) - f(u) \leqslant \lambda[f(x) - f(u)],$$

https://doi.org/10.1017/S0334270000005142 Published online by Cambridge University Press

assume $\lambda > 0$ and divide by λ to obtain

$$\frac{1}{\lambda}[f(u+\lambda\eta(x,u)) - f(u)] \leqslant f(x) - f(u).$$

Taking the limit as $\lambda \to 0_+$ gives (1).

Note that unlike the situation for convexity (where $\eta(x,u) \equiv x - u$), (5) plus differentiability implies (1), but not conversely.

3. Constrained optimization

Consider the nonlinear programming problem

(P) Minimize $f(x)$ subject to $g(x) \leqslant 0$,

where f and g are differentiable functions from R^n into R and R^m. Since f and g_i, $i = 1,\ldots,m$ invex, for the same function η, implies that for any $\in R^m$ such that

$$y \geqslant 0, \tag{6}$$

the function

$$f(x) + y'g(x) \tag{7}$$

is invex, the role of invexity for problem (P) can be discerned by applying Theorem 1 to the unconstrained Lagrangian function (7). From Theorem 1, it follows that (for fixed y) (7) is invex if and only if

$$\nabla f(u) + \nabla y'g(u) = 0 \tag{8}$$

implies

$$f(u) + y'g(u) \leqslant f(x) + y'g(x) \tag{9}$$

for all x, $u \in R^n$. If x is feasible, $y'g(x) \leqslant 0$ and (9) yields

$$f(u) + y'g(u) \leqslant f(x). \tag{10}$$

Thus, (10) holds if x is feasible for (P), (u, y) satisfies (6) and (8), and (7) is invex. This establishes weak duality between (P) and its Wolfe dual: (D) maximize $f(u) + y'g(u)$ subject to (6) and (8), the result stated by Hanson [5].

Sufficiency of the Kuhn-Tucker conditions when (7) is invex also follows readily from (10). Thus, if u is feasible for (P) and $y_i = 0$ if $g_i(u) < 0$, then $y'g(u) = 0$ and (10) becomes

$$f(u) \leqslant f(x) \tag{11}$$

for all feasible x.

https://doi.org/10.1017/S0334270000005142 Published online by Cambridge University Press

It should be stressed that for sufficiency our proof requires that the Lagrangian (7) be invex, or, equivalently, that a stationary point of (7) be a global minimum of (7). Even if f and each g_i are individually invex, $f + y'g$ may not be invex. This is equivalent to f and g_i being invex for different η but $f + y'g$ is not invex for any η. An example of a problem where both f and g are invex but $f + y'g$ is not, is the following:

EXAMPLE 3. Minimize $\frac{1}{2}(x-1)^2$ subject to $x^3 + x \leqslant 0$.
If we take $x = 0$, $y = 1$, the Kuhn-Tucker conditions are all satisfied. Here both f, which is convex and g, which is pseudo-convex, are invex, but, for $y = 1$, the Lagrangian $f + y'g = x^3 + \frac{1}{2}x^2 + \frac{1}{2}$ is not; so that neither the sufficiency result proved here nor in [5] is applicable (although, of course, $x = 0$ is minimal).

Martin [11] and Hanson and Mond [6] observed that by weakening slightly the requirement that all g_i be invex, invexity not only remains sufficient for a Kuhn-Tucker point to be optimal for (P), but becomes necessary as well.

Recall that g_i invex means

$$g_i(x) - g_i(u) \geqslant \eta(x,u)'\nabla g_i(u). \tag{12}$$

If x is feasible and g_i is a tight constraint at u, we have

$$0 \geqslant \eta(x,u)'\nabla g_i(u). \tag{13}$$

This is the requirement on the tight constraints of g at u set by Martin [11], who gives necessary and sufficient conditions for every Kuhn-Tucker point to be a global minimizer.

Hanson and Mond [6] require instead

$$-g_i(u) \geqslant \eta(x,y)'\nabla g_i(u) \tag{14}$$

for all g_i. It is clear that, if (11) holds, there exists an $\eta(x,u)$ satisfying (1) and (13) or (14). Simply set $\eta(x,u) = 0$. Hanson and Mond [6] give conditions for the existence of a non-zero $\eta(x,u)$ satisfying (1) and (14).

For a given pair of points x, u the condition (1) is a linear inequality in $\eta = \eta(x,u)$. If $\nabla f(u) \neq 0$, (1) has a solution for all x, and the general solution is

$$\eta(x,u) = \frac{f(x) - f(u)}{\nabla f(u)'\nabla f(u)}\nabla f(u) + v, \tag{15}$$

$$v'\nabla f(u) \leqslant 0.$$

Thus the general solution (15) is a closed half-space. For an invex function f and a point u with $\nabla f(u) = 0$, any η will do in (1) for all x, and the general solution is the whole space.

Consider now a given finite collection of invex functions $\{f_1, \ldots, f_m\}$, and the question of the existence of a common η satisfying

$$f_i(x) - f_i(u) \geqslant \eta(x,u)'\nabla f_i(u), \; i = 1, \ldots, m \tag{16}$$

for all x, u. Let

$$J(u) = \{i: \nabla f_i(u) \neq 0\}.$$

The existence of a common η is equivalent to the nonemptiness of the following intersection of closed half spaces

$$\bigcap_{i \in J(u)} \left\{ \frac{f_i(x) - f_i(u)}{\nabla f_i(u)'\nabla f_i(u)} \nabla f_i(u) - (\nabla f_i(u))^* \right\} \neq \varnothing \tag{17}$$

for all x, u. A sufficient condition for (17), stated in terms of u alone, is that the set $\{\nabla f_i(u): i \in J(u)\}$ be nonnegatively independent, i.e.

$$\left. \begin{array}{l} \sum_{J(u)} \alpha_i \nabla f_i(u) = 0 \\ \alpha_i \geqslant 0, \; i \in J(u) \end{array} \right\} \Rightarrow \alpha_i = 0, \qquad i \in J(u). \tag{18}$$

If (18) holds, then for any x there is an $\eta(x,u)$ satisfying (16). The condition (18) is a well-known constraint qualification.

We turn now to the saddle-point problem. A point $(\bar{x}, \bar{y})$, $\bar{y} \geqslant 0$ is said to be a solution of the saddle-point problem if

$$f(\bar{x}) + y'g(\bar{x}) \leqslant f(\bar{x}) + \bar{y}'g(\bar{x}) \leqslant f(x) + \bar{y}'g(x) \tag{19}$$

for all $x \in R^n$, $y \in R^m$, $y \geqslant 0$. It is well known (see e.g. [10]) that if $(\bar{x}, \bar{y})$ is a solution of (19), then $\bar{x}$ is an optimal solution of (P). On the other hand, if $\bar{x}$ is an optimal solution of (P), one needs a constraint qualification and convexity to assure the existence of $\bar{y}$ such that $(\bar{x}, \bar{y})$ is a solution of (19). We now show that this convexity requirement can be weakened to invexity.

THEOREM 3. *If $\bar{x}$ is optimal for* (P), *a constraint qualification is satisfied, and* (7) *is invex for $y \geqslant 0$, then there exists $\bar{y} \geqslant 0$ such that $(\bar{x}, \bar{y})$ is a solution of the saddle-point problem.*

PROOF. Since $\bar{x}$ is optimal for (P) and a constraint qualification is satisfied, then by the Kuhn-Tucker necessary conditions, there exists $\bar{y} \in R^m$ such that (with $u = \bar{x}$) (6) and (8) are satisfied as well as $\bar{y}'g(\bar{x}) = 0$. Since (7) is invex, (8) implies (9) by Theorem 1, which is the right side of (19). The left side holds since $y \geqslant 0$, $g(\bar{x}) \leqslant 0$, $\bar{y}'g(\bar{x}) = 0$.

COROLLARY 2. *Assume that a constraint qualification is satisfied at* $\bar{x}$ *and that* (7) *is invex for all* $y \geqslant 0$. *Then* $\bar{x}$ *is optimal for* (P) *if and only if there exists* $\bar{y} \geqslant 0$ *such that* $(\bar{x}, \bar{y})$ *is a solution of the saddlepoint problem.*

Finally, we show that invexity can be substituted for convexity in the Slater constraint qualification.

THEOREM 4. *Assume that* $\bar{x}$ *is an optimal solution of* (P). *Assume, also, that there exists a point* x^* *such that* $g(x^*) < 0$ *and that all* g_i *for which* $g_i(\bar{x}) = 0$ *are invex with respect to the same vector function* $\eta(x, u)$. *Then there exists* $\bar{y} \in R^m$ *such that* $(\bar{x}, \bar{y})$ *satisfy the Kuhn-Tucker necessary conditions.*

PROOF. Let B_0 denote the set of active constraints at $\bar{x}$, i.e., $B_0 \equiv \{i\colon g_i(\bar{x}) = 0\}$. If we can show that

$$z^t \nabla g_i(\bar{x}) \leqslant 0 \ (\forall i \in B_0) \Rightarrow z^t \nabla f(\bar{x}) \geqslant 0 \tag{20}$$

the result will follow as in [9] by applying Farkas' Lemma and setting $y_i = 0$ for $i \notin B_0$.

Assume (20) does not hold, i.e., there exists $z \in R^n$ such that

$$z^t \nabla g_i(\bar{x}) \leqslant 0 \ \forall i \in B_0 \quad \text{and} \quad z^t \nabla f(\bar{x}) < 0. \tag{21}$$

Since by the assumed Slater-type condition, $g_i(x^*) - g_i(\bar{x}) < 0$, $i \in B_0$, by invexity there exists $\eta(x^*, \bar{x})$ such that

$$[\eta(x^*, \bar{x})]^t \nabla g_i(\bar{x}) < 0, \qquad i \in B_0. \tag{22}$$

Therefore

$$[z + \rho\eta(x^*, \bar{x})]^t \nabla g_i(\bar{x}) < 0, \qquad i \in B_0, \tag{23}$$

for all $0 < \rho$. Hence for some positive τ small enough

$$g_i(\bar{x} + \tau[z + \rho\eta(x^*, \bar{x})]) \leqslant g_i(\bar{x}) = 0, \qquad i \in B_0.$$

Acknowledgment

This research was supported by NSF Grant EGS-8214081 while one of the authors (A. B-I.) was a visitor at La Trobe University.

References

[1] B. D. Craven, "Duality for generalized convex fractional programs," in *Generalized concavity in optimization and economics*, (eds. S. Schaible and W. T. Ziemba), (Academic Press, New York, 1981), 473–489.

[2] B. D. Craven, "Invex functions and constrained local minima," *Bull. Austral. Math. Soc.* 24 (1981), 357–366.

[3] B. D. Craven, and B. M. Glover, "Invex functions and duality," *J. Austral. Math. Soc. Ser. A*, 39 (1985), 1–20.

[4] L. Ya. Gulenko, (Russian) "The relaxation method for mathematical programming problems with Hanson's functions," *Dokl. Akad. Nauk BSSR* 27 (1983), 105–107.

[5] M. A. Hanson, "On sufficiency of the Kuhn-Tucker conditions," *J. Math. Anal. Appl.* 80 (1981), 545–550.

[6] M. A. Hanson and B. Mond, "Necessary and sufficient conditions in constrained optimization," FSU Statistics Report M683, Florida State University, Department of Statistics, Tallahassee, Florida, U.S.A., 1984.

[7] V. Jeyakumar, "Strong and weak invexity in mathematical programming," *Method. Oper. Res.* 55 (1985), 109–125.

[8] R. N. Kaul and S. Kaur, "Optimality criteria in nonlinear programming involving nonconvex functions," *J. Math. Anal. Applic.* 105 (1985), 104–112.

[9] H. W. Kuhn, and A. W. Tucker, "Nonlinear programming." in *Proceedings of the Second Symposium on Mathematical Statistics and Probability*, (ed. J. Neyman), (University of California Press, Berkeley, Calif., 1951), 481–492.

[10] O. L. Mangasarian, *Nonlinear programming*, (McGraw-Hill, New York, 1969).

[11] D. H. Martin, "The essence of invexity," *J. Optim. Theory Appl.* 47 (1985), 65–76.

[12] B. Mond, and M. A. Hanson, "On duality with generalized convexity," *Math. Operationsforsch. Statist. Ser. Optim.* 25 (1984), 962–967.

[13] I. Zang, E. U. Choo, and M. Avriel, "On functions whose stationary points are global," *J. Optim. Theory Appl.* 22 (1977), 195–208.