Characterizing invex and related properties

B. D. Craven

University of Melbourne, Auatralia

Presented at the 7th. International Symposium on Generalized Convexity/Monotonicity at Hanoi, Vietnam, August 27-31, 2002

Abstract A characterization of *invex* , given by Glover and Craven, is extended to functions in abstract spaces. *Pseudoinvex* for a vector function coincides with *invex* in a restricted set of directions. The *V-invex* property of Jeyakumar and Mond is also characterized. Some differentiability properties of the invex *scale function* are also obtained.

1. Introduction

A differentiable vector function F is *invex* at a point p if

$$F(.) - F(p) \geq F'(p)\eta(.,p)$$

for some *scale function* η . As is well known, with this property necessary Lagragian optimization conditions are also sufficient, and various duality results hold. It is important to find when the *invex* property holds. This paper extends the characterizations of *invex* given by Craven and Glover (1985) and Craven (2002) to also characterize *V-invex* (Jeyakumar andr Mond, 1992), and to show that a related *pseudoinvex* property of a vector function concides with *invex* in a restricted set of directions. Differentiability properties of the *scale function* can also be characterized.

2. Characterizing invex

The differentiable vector function $F : \mathbf{R}^n \to \mathbf{R}^m$ is (globally) *invex* at $p \in \mathbf{R}^n$ if, for some differentiable *scale function* $\eta : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$,

$$(\forall x \in \mathbf{R}^n \ ) \ F(x) - F(p) \geq F'(p)\eta(x,p) \ . \tag{1}$$

If F and ω are twice-differentiable, then they have Taylor expansions

$$F(x) = F(p) + F'(p)(x-p) + F^{\#}(x-p;p); \quad \omega(x-p) \equiv \eta(x,p) = x - p + \omega^{\#} \text{ (x-p)}.$$

Substituting in the definition of invexity, invexity at p is equivalent to

$$(\forall x \in \mathbf{R}^n \ ) \ F^{\#} \text{ (x-p;p) } \geq F'(p)\omega^{\#} \text{ (x-p;p)} \ . \tag{2}$$

This is applied to an optimization problem:

$$\text{MIN } F_1(x) \text{ subject to } F_j(x) \leq 0 \ \ (j = 2, 3, \ldots \text{ , m}) \ . \tag{3}$$

Here, F is called *active-invex* (*a-invex*) at p if $F_1(p) = 0$ (by replacing $F_1(.)$ by $F_1(.) - F(p)$), and F^A (.), obtained from $F(.)$ by omitting those components F_j for which F_j (p) $\neq 0$, is invex at p. If α is a vector of Lagrange multipliers, with $\alpha_1 = 1$, then the Lagrangian:

$$L(x) := \alpha^T F(x) = \alpha^T (F(x) \text{ -F(p) } - F'(p)(x-p)) \equiv \alpha^T F^{\#} \text{ (x-p;p)} \tag{4}$$

if $L'(p) = \alpha^T F'(p) = 0$.

Consider, more generally, an optimization problem:

$$\text{MIN } F_1(x) \text{ subject to } G(x) := (F_2(x), \ldots, F_m \text{ (x)}) \ \in \ \text{-S}, \tag{5}$$

where S is a closed convex cone. Let $E \subset \mathbf{R}^n$. By definition, $F(.)$ is invex at p on E with respect to the convex cone $U := \mathbf{R}_+ \times S$ if:

$$(\exists \eta : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n)(\forall x \in \text{ E}) \; F(x) - F(p) - F'(p)\eta(x,p) \in U, \tag{6}$$

or equivalently if:

$$(\forall x \in \mathbf{E} \;) \; F^{\#} \text{ (x-p;p) } - F'(p)\omega^{\#} \text{ (x-p;p) } \in \text{ U}. \tag{7}$$

(Note that this definition restricts the set of points x for which the invex property is considered. If E is not stated, then $E = \mathbf{R}^n$ is assumed.)

For problem (5), F is called *a-invex* at p if $F_1(p) = 0$ and F is invex at p with respect to the convex cone $U^A := \{\gamma(u + F(p)) : u \in U, \gamma \geq 0\}$. This reduces to the previous case when $S = \mathbf{R}_+^{m-1}$. The dual cone of U is denoted by U^*; then the dual cone of U^A is $U^{A*} = \{u^* \in U^* : u^{*T}F(p) = 0\}$.

The characterization of *invex* depends on the following consequence (see Craven and Glover, 1985; Craven 2002) of Motzkin's alternative theorem. It is stated here in abstract spaces, so that it may also be applied to optimal control problems.

Theorem 1 (Characterization) Let X and Y be normed spaces (or *lctvs*); let $M : X \to Y$ be a continuous linear mapping; let $V \subset Y$ be a closed convex cone; let $\theta \in$ Y; let the convex cone $K^*(V^*)$ be weak * closed, where K^* is the adjoint of $K := [M, -\theta]$, and V^* is the dual cone of V. Then:

$$(\exists \zeta : X \to X) - M\zeta + \theta \in V \Leftrightarrow$$

$$\{[0 \neq \alpha \in V^*, \alpha M = 0] \Rightarrow \; [0 \; \neq \alpha \in V^*, < \alpha, \theta > \geq 0]\} \text{ (8)}$$

Proof For a fixed $x \in \mathbf{R}^n$, set $\theta := F^{\#}$ (x-p;p) and $\zeta := \omega^{\#}$ (x-p;p), and set $M := F'(p)$. Then,

$$(\exists \zeta) - M\zeta + \theta \in V$$

$$\Leftrightarrow (\exists z \in X) - Mz + t\theta \in V, t \in \text{ int } \mathbf{R}_+$$

$$\text{substituting } \zeta = z/t, u = (z,t)$$

$$\Leftrightarrow (\exists u \in X \times \mathbf{R} \;) \text{ -Ku } \in V, \text{ -Nu } \in \text{ int } \mathbf{R}_+ \text{ where } N := [0, -1]$$

$$\Leftrightarrow \mathbf{NOT}(\exists \alpha \in V^*, \exists \beta \in \mathbf{R} \;) \; \alpha^T K + \beta N = 0, \alpha \geq \; 0, \; 0 \neq \beta \geq 0$$

$$\text{by Motzkin's alternative theorem, since } K^*(V^*) \text{ is weak * closed}$$

$$\Leftrightarrow \; \mathbf{NOT} \; (\exists \alpha \in V^*) \; \alpha M = 0, 0 \neq < \alpha, \theta > \leq 0$$

$$\Leftrightarrow \; [0 \; \neq \alpha \in V^*, \alpha M = 0] \Rightarrow < \alpha, \theta > \geq 0 \; . \qquad //$$

Theorem 2 In the differentiable optimization problem (5), F is invex [resp. active-invex] at a point p (where $F_1(p) = 0$) satisfying $\alpha \in U, \alpha^T F'(p) = 0$ [resp. $\alpha \in U, \alpha^T F'(p) = 0, \alpha^T F(p) = 0\}$ if and only if, for each $x, \alpha^T F^{\#}$ (x-p;p) ≥ 0, provided that the cone

$$[F'(p), -F^{\#}(x-p;p)]^*(U^*) \text{ [resp. } [F'(p), -F^{\#}(x-p;p)]^*(U^{A*}]$$

is closed. This condition holds automatically for problem (3).

Proof Apply Theorem 1 with $\theta = F^{\#}$ (x-p;p), for each fixed $x, M = F'(p)$,

and $V = U$ [resp. $V = U^A$]. For problem (3) the cone is polyhedral, hence closed. //

Discussion

Often α does not depend on x, in particular if α is a unique vector of Lagrange multipliers.

Theorem 2 also applies to infinite-dimensional problems, such as optimal control in continuous time, provided that the mentioned cone is assumed to be weak * closed.

Consider the following small examples:

Example 1 : The point (0,0) is a Karush-Kuhn-Tucker (KKT) point for:

$$\text{MIN } F_1(x,y) = x^2 + y \text{ subject to}$$
$$F_2(x,y) = -y + \alpha y^2 \leq 0, \quad F_3(x,y) = -1 - x + \beta x^2 \leq 0,$$

with Lagrange multipliers 1 and 0. The function $F = (F_1, F_2, F_3)$ is invex at (0,0) if functions $\rho(x,y)$ and $\sigma(x,y)$ exist for which:

$$x^2 \geq 0\rho + 1\sigma, \alpha y^2 \geq 0 - 1\sigma, \beta x^2 \geq -1\rho + 0\sigma \ ;$$

which hold for $\alpha > 0$ and β of either sign. The Lagrangian at (0,0) is:

$$L(x,y) = x^2 + 1(\alpha y^2) + 0(\beta x^2),$$

so that $L(x,y) \geq L(0,0)$ provided that $\alpha \geq 0$. If $\alpha < 0$ then $L(x,y)$ is not minimized at (0,0), and invexity fails, since $(-\alpha)y^2 \leq \sigma \leq x^2$ does not generally hold.

Example 2 The point (0,0I is a KKT point for:

$$\text{MIN } F_1(x,y) = x + x^2 + y \text{ subject to}$$
$$F_2(x,y) = -y + \alpha y^2 \leq 0, \ F_3(x,y) = -1 - x + \beta x^2 \leq 0, \ F_4(x,y) = -x \leq 0,$$

with Lagrange multipliers 1,0,1. However $\alpha^T F'(0,0) = 0$, with $\alpha = (1, \lambda, \mu, v) \geq$ (0,0,0,0), only requires that $1 + 0\lambda - \mu - \nu = 0$ and $1 - \lambda + 0\mu + 0\nu = 0$, hence $\lambda = 1, \mu + \nu = 1$, so the multiplier μ for the inactive constraint can be any value in [0,1]. The Lagrangian at (0,0) is:

$$L(x,y) = x^2 + 1(\alpha y^2) + \mu(\beta x^2) + \nu(0) \geq L(0,0)$$

for each $\mu \in [0,1]$ provided that $\alpha > 0$ and $1 + \mu\beta \geq 0$, thus when $\beta \geq -1$. And F is invex at (0,0) when $x^2 \geq \rho + \sigma, \alpha y^2 \geq -\sigma, \beta x^2 \geq -\rho, 0 \geq -\rho$, which hold when $\alpha \geq 0$ and $(1+\beta)x^2 + \alpha y^2 \geq 0$, thus when $\beta \geq -1$.

However, $F^A = (F_1, F_2, F_4)$ is invex when F is invex with respect to U^A , which requires $\mu = 0$, so here β is unrestricted.

3. Vector pseudoinvex

A differentiable vector function $F : \mathbf{R}^n \to \mathbf{R}^m$ is *vector pseudoinvex* (*vpi*) at p with respect to the convex cone U (Craven, 2001) if:

$$F(x) - F(p) \in -\text{ int } U \Rightarrow (\exists \eta : \mathbf{R}^n \times \mathbf{R^n} \to \mathbf{R^n}) \ \ F'(p)\eta(x,p) \in -\text{ int } U \ . \tag{9}$$

Theorem 3 Let F be differentiable; let $U = \mathbf{R}_+ \times S$ be the convex cone in problem (5); for fixed p, denote $E := \{x : \ F(x) - F(p) \in -\text{ int } U\}$; assume that the cone $[F'(p), -F(x) + F(p)]$ is closed, for each $x \in$ E. Then F is vector pseudoinvex at p with respect to U if and only if F is invex at p on E with respect to U.

Proof For a given $x \in E$, denote $c := F(x)-$ F(p) and $d := F'(p)\eta(x,p)$. If (9) holds, and $c \in -$ int U, then $d + N \in -$ int U for some open ball N. For some $\epsilon > 0, -\epsilon c \in$ N. Hence $d - \epsilon c \in -$ int U. Conversely, if $c \in -$ int U and $d - \epsilon c \in -$ int U, then $d \in -$ int $U-$ int $U = -$ int U. Hence (9), for a given x, is equivalent to $d - \epsilon c \in -$ int U, for some $\epsilon > 0$. Hence, for $k = 1/\epsilon > 0$, (9) is equivalent to:

$$c - \text{ kd } \in \text{ int } U \subset \text{ U}. \tag{10}$$

Applying Theorem 1 with $M = F'(p), \theta = k^{-1}(F(x) - F(p)) \in -$ int U, and $V = U$ shows that η exists, satisfying (10), if and only if:

$$[\alpha \in U^*, \alpha^T F'(p) = 0] \Rightarrow \alpha^T k^{-1}(F(x) - F(p)) \geq 0 \ ;$$

or equivalently if and only if:

$$[\alpha \in U^*, \alpha^T F'(p) = 0] \Rightarrow \alpha^T F^{\#}(x-p;p) \geq 0 \ . \tag{11}$$

From Theorem 2, (11) for each $x \in E$ holds if and only if F is invex at p on E with respect to U. //

Remark Thus *pseudoinvex* at a point reduces to *invex* at the point in a restricted set of directions, This result explains the scarcity in the literature of examples of functions that are pseudoinvex but not invex. However, such a function could be constructed by changing an invex function F at some points x for which $F(x) - F(p) \notin -$ int U.

4. Description of V-invex

In (Jeyakumar and Mond, 1992), a vector function F is called *V-invex* if:

$$F_i(x) - F_i(p) \geq \gamma_i(x,p)F'(p)\eta(x,p) \tag{12}$$

holds for each x and each component F_i , with some positive scalar coefficients, here denoted γ_i (x, p). They showed that property (12) can replace invex, in proving sufficient KKT conditions. Here, p is fixed, and a characterization is obtained for the property (12), using Theorem 1. For a given x, denote $r_i := \gamma_i(x,p)^{-1}$; let $v = x - p$. Then (12) may be written:

$$\theta_i := (r_i - 1)F_i'(p)v + r_i F_i^{\#}\ \text{(v;p)}\ \geq F_i'(p)\omega^{\#}(v;p). \tag{13}$$

Applying Theorem 1 with $\theta = (\theta_1, \ldots, \theta_m\ )$, $M = F'(p), V = U$ gives the equivalent statement:

$$[0 \neq \alpha \in U^*, \alpha^T F'(p) = 0] \Rightarrow \sum \alpha_i(r_i - 1)F_i'(p)v + \sum \alpha_i r_i F_i^{\#}(v,p) \geq 0. \tag{14}$$

Theorem 4 Let F be differentiable; let $U = \mathbf{R}_+ \times S$ be the convex cone in problem (5); for each x, assume that the cone $[F'(p), -\theta]$, with θ from (13), is closed, Then F is V-invex at p, with respect to the cone U, if and only if:

$$[0 \neq \alpha \in U^*,\ \alpha^T F'(p) = 0] \Rightarrow \sum_i \beta_i(F_i'(p)(x-p) + F_{i\#}\ \text{(x-p;p)}\ \geq\ 0, \tag{15}$$

for some coefficients $\beta_i \equiv \beta_i(x) \geq 0$, with

$$\beta_i > 0 \Leftrightarrow \alpha_i > 0. \tag{16}$$

Proof From (14), with $\beta_i := \alpha_i r_i$. //

5. Properties of the scale function

Apply now the definition (7) of invex at p, but with cone U^A , in the form:

$$F^{\#}(.;p) - \mathbf{M}\omega^{\#}(.;p) \in \mathbf{U}^A. \tag{17}$$

considering E as a compact subset of $\mathbf{R}^n$, with $F^{\#}(.:p) : C(E) \to \mathbf{R}^m$, $\omega^{\#}$ (.;p) : $C(E) \times C(E) \to C(E)$, and $\mathbf{U}^A := \{bu \in C(E) : (\forall t \in E)\ \mathbf{u}\ (t) \in U^A\}$, and define $\mathbf{M}$: $C(E, \mathbf{R}^m\ ) \to C(E, \mathbf{R}^m\ )$ by $(\mathbf{Mu})(t) := M\mathbf{u}(t)$, with $M := F'(p)$. Define $\mathbf{K}$ by $(\mathbf{Ku})(x) := K(x)\mathbf{u}(x)$, with $K(x) := [M, -\theta(x)]$. Since Theorem 1 is formulated in abstract spaces, it can be applied to (17). Assume now that F is C^2, and express (17) as:

$$(\exists \sigma(.) \in C^1 E))\theta(.) - \mathbf{M}\sigma(.) \in \mathbf{U}^A. \tag{18}$$

with $\theta(.) := F^{\#}(.;p)/\|x-p\|^2$, and $\sigma(.) := \omega^{\#}(.;p)/\|x-p\|^2$.

The elements of the dual cone $\mathbf{U}^{A*}$are represented by signed vector measures μ . Then:

$$\mu \in \mathbf{U}^{A*} \Leftrightarrow (\forall \mathbf{z} \in \mathbf{U}^A)\ 0 \leq\ <\mu, \mathbf{z}>\ = \int_E \mu(x)\mathbf{z}(x)dx, \tag{19}$$

For each interval I, let φ be a smooth approximation to the indicator function of I. Substituting $\mathbf{z} = \mathbf{c}\varphi, 0 \leq \int_E \mu(x)\mathbf{c}\varphi(x)dx$, for each $\mathbf{c} \in U^A$, hence $\int_E \mu(x)\varphi(x)dx \in U^{A*}$. Taking a limit of suitable φ,

$$\mu(E \cap I) \in U^{A^*}. \tag{20}$$

If $\mathbf{w} \in C(E, \mathbf{R}^m\ )$, then

$$< \mathbf{K}^*\mu, \mathbf{w} >= \int_E < \mu(x), K(x)\mathbf{w}(x) > dx \tag{21}$$

may be approximated by the integral of a step-function, taking constant values on intervals I. With (21), this approximation has the form $\int_E < \mathbf{q}$ (x),$\mathbf{w}$ $(x) > dx$, where $\mathbf{q}$ (x) $\in K(x)^T(U^{A*})$. If the cone $K(x)^T(U^{A*}$ is closed, then (as a limiting case) $\{< \mathbf{K}^*\mu$,$\mathbf{w} >: \mu \in \mathbf{U}^{A*}\}$ is closed, for each $\mathbf{w}$. Thus $\mathbf{K}^*(\mathbf{U}^{A*}$ is weak * closed,

Theorem 5 (Property of scale function) Assume that:
(a) the function F is C^2 and a-invex at each point $x \in E$,
(b) the convex cone $K(x)^T(U^{A^*})$ is closed, for each $x \in E$,
(c) $[\alpha \in U, \alpha^T F'(p) = 0, \alpha^T F(p) = 0]$ defines a unique $\alpha = \hat{\alpha}$.
Then there exists a scale function ω such that:

$$\omega\#(x-p;p) = \psi(x)\|x-p\|^2,$$

with $\psi(.)$ continuous.
Proof Since F is a-invex at each point x, and the cone $K(x)^T(U^{A*})$is closed, Theorem 2 shows that:

$$(\forall x \in E)[0 \neq \alpha(x) \in U^A, \alpha(x)^T F'(p) = 0, \alpha^T F(p) = 0] \Rightarrow \alpha(x)^T F^{\#}(x-p;p) \geq 0 \tag{22}$$

A multiplier μ satisfying $[0 \neq \mu \in \mathbf{U}^{A*}, < \mu, \mathbf{M} > = 0]$ is a signed vector measure for which $\mu(I) = \hat{\alpha}(I)$ for each interval I, and this μ is unique since α is assumed unique. Hence (22) implies that:

$$[0 \neq \mu \in \mathbf{U}^{A^*}, < \mu, \mathbf{M} = 0] \Rightarrow < \alpha, \theta(.) \geq 0. \tag{23}$$

Since the cone $K(x)^T(U^{A*})$ is closed, the cone $\mathbf{K}^{\ *}(\mathbf{U}^{A*})$ is weak * closed, by the earlier discussion, Hence Theorem 1 shows that (18) holds if and only if (24) holds:

$$[0 \neq \mu \in \mathbf{U}^{A*}, < \mu , \mathbf{M} > = 0] \Rightarrow < \alpha, \theta(.) \geq 0. \tag{24}$$

Since $\alpha(x) = \hat{\alpha}$ uniquely, independent of $x \in E, \int_I \mu(x)dx = \hat{\alpha}|I|$ for each interval I, where μ is the signed vector measure representing α . Hence (22) is a consequence of (23). Hence (18) holds. Hence a continuous $\sigma(.)$ exists. //

Discussion Suppose now that $F^{\#}$ (.;p) and $\omega = \#(.;p)$ are defined on spaces $C^1(E)$ instead of $C(E)$. If $\mathbf{U}^A$ is redefined with elements $\mathbf{u} \in C^1(E)$, then the dual cone $\mathbf{U}^{A\ *}$ is represented by Schwartz distributions μ , which are the weak derivatives of signed vector measures ρ . If ψ is a smooth vector function, then $< \mu, \psi > = - < \rho, \psi' >$, leading to $\rho(E \cap I) \in -U^{A\ *}$ instead of (20). A similar construction from (21), using ρ , shows again that the cone is weak * closed, Hence the invex property (18) also holds with $\sigma(.) \in C^1(E)$.

The dependence of the scale function on the point p can also be analysed. The property:

$$(\forall v \in X)(\forall p \in P) F^{\#}(v;p) \geq F'(p)\omega^{\#}(v;p), \tag{25}$$

where X and P are suitable domains, may be expressed as:

$$\Phi(p)(v) \geq \mathbf{L}\Omega(p)(v), \tag{26}$$

in which $\Omega(p)(v) := \omega^{\#}(v;p)/\|v\|^2, \Phi(p)(v) := F^{\#}(v;p)/\|v\|^2$, and the linear mapping $\mathbf{L}$ is the Cartesian product of the mappings $F'(p)$, for $p \in P$. This construction may be illustrated by the following case, where p takes only two values p_1 and p_2 :

$$\begin{bmatrix} F^{\#}(v;p_1) \\ F^{\#}(v;p_2) \end{bmatrix} \geq \begin{bmatrix} F'(p_1) & 0 \\ 0 & F'(p_2) \end{bmatrix} \begin{bmatrix} \omega^{\#}(v;p_1) \\ \omega^{\#}(v;p_2) \end{bmatrix} \tag{27}$$

If F ic C^2, then $\mathbf{L}$ is a continuous mapping of $C^1(E) \times C^1 P)$ into itself. Theorem (1) may be used to characterize (25), provided that a certain convex cone is weak * closed. (This can be described, similarly to Theorem 5) . If F is assumed invex at each point p, then it follows that a scale function exists, with $\omega^{\#}(v,p) = \|v\|^2\rho(v,p)$, where $\rho(.,.)$ is a C^1 function.

However, it does *not* follow that F is *convexifiable* ; there need not exist any invertible transformation γ, such that $F \circ \gamma$ is convex at all points p.

6. References

B. D. Craven and B. M. Glover (1985), *Invex functions and duality* , Journal of the Australian Mathematical Society, Series A, **39** , 1-20.

V. Jeyakumar and B. Mond (1992), *On generalized convex mathematical programming,* Journal of the Australian Mathematical Society, Series *B*, **34** , 43-53,

B. D. Craven (2001), *Vector generalized invex* , Opsearch **38** (4), 345-361.

B. D. Craven (2002), *Global invexity and duality in mathematical programming* , Asia-Pacific Journal of Operational Research, to appear.