GLOBAL INVEXITY AND DUALITY IN MATHEMATICAL PROGRAMMING

Asia-Pacific Journal of Operational Research 19 (2002), 169-175

B D CRAVEN
Department of Mathematics and Statistics,
University of Melbourne,
Victoria 3010, Australia

Global invexity is characterized by a condition which is independent of the scale function describing the invexity. Consequently, weak duality holds for the Wolfe, or Mond-Weir, dual problem when a sufficient invexity hypothesis is replaced by a suitable inequality condition. This holds exactly when the Wolfe dual is equivalent to the Lagrangian dual. Results are given for differentiable, and for locally Lipschitz, functions.

Keywords Global invex, Wolfe dual, Mond-Weir dual, Lipschitz

1. Introduction

The *invex* property (see Hanson (1980), Craven (1981a), and Craven (1995) of a vector function has more often been postulated, than shown to hold, for some class of constrained optimization problem. Conditions for *local invexity* (thus, invexity in a local domain) were obtained in Craven (1981b), by quadratic approximation of the function. Such conditions extend only until some parabola has a turning point, and so cannot provide a condition for global invexity. Moreover, the class of quadratic functions that are invex but not convex is quite restricted. So another sort of approximation is required.

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* $\eta : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$,

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

The gradient $F'(p)$ is a $m \times n$ matrix, with $m < n$. Since $\eta(p,p)$ is normally zero, the scale function (at given p) may be written as $\omega(x-p,p)$. Expand $\omega(x-p,p) = C(x-p) + \mathbf{o}(\|x-p\|)$, where C is linear, and set $x=p+\alpha v$ for fixed v and $\alpha \in \mathbf{R}$. Then, from (1):

$$(\forall \alpha \in \mathbf{R}) \quad F'(p)\alpha v + \mathbf{o}(|\alpha|) \geq F'(p)(C\alpha v + \mathbf{o}(|\alpha|)) \,.$$

Hence $F'(p)C(x-p) = 0$; thus the term $(C-I)(x-p)$ may be subtracted from $\omega(x-p,p)$. So it can be assumed (see Craven (1981b)) that $\eta(x-p) = x-p + \mathbf{o}(\|x-p\|)$.

It is observed that a function may be invex at one point, but not at other neighbouring points. An example is $f : \mathbf{R} \to \mathbf{R}$ given by $f(x) = -x^2$. Invexity at $p=0$ would require that $(\forall x)\ -x^2 \geq -0.\omega(x)$, which is contradicted. However, invexity at $p>0$ requires that (setting $z=x-p$) $-2pz-z^2 \geq (-2p)(z+\omega(z))$, thus that $\omega(z) \geq (2p)^{-1}z^2$; and similarly for $p<0$. Thus f is invex at each point $p\neq 0$, but the scale function becomes unbounded as $p\to 0$.

Invexity is often used for minimization problems, where $f(x) := F_1(x)$ is minimized over $x \in \mathbf{R}^n$, subject to constraints $g_j(x) := F_j(x) \leq 0$. If a minimum is reached at $x=p$, and a constraint qualification holds, then Karush-Kuhn-Tucker (KKT) conditions hold, that

$$F_1'(p) + \sum_{j>1} \rho_j F_j'(p) = 0,\ \ \rho_j \geq 0,\ \rho_j F_j(p) = 0\ (j>1). \qquad (2)$$

Thus $F'(p)$ does *not* have full rank. If F is invex, then KKT is sufficient for a minimum. This suggests studying invexity under the restriction that $F'(p)$ has rank $m-1$. It is useful to assume a further restriction, that each multipler $\rho_j > 0$, thus excluding cases where both $\rho_j = 0$ and $F_j(p) = 0$ for some j.

Invexity is also applied to dual problems, containing constraints of the form

$$F_1'(x) + \sum_{j>1} \rho_j F_j'(x) = 0,\ \rho_j \geq 0. \qquad (3)$$

It is relevant also here to consider invexity under a rank restriction on $F'(x)$.

2. Invexity is a second-order property

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

$$F(x) = F(p) + F'(p)(x-p) + F^{\#}(x-p;p)\ ; \qquad (4)$$

$$\omega(x) = \omega(p) + \omega'(p)(x-p) + \omega^{\#}(x-p;p)\ ;$$

in which the higher-order terms satisfy $F^{\#}(x-p;p) = \mathbf{o}(\|x-p\|)$and $\omega^{\#}(x-p;p) = \mathbf{o}(\|x-p\|)$. Substituting into the definition of invexity, invexity at p is equivalent to :

$$(\forall x) \qquad F^{\#}(x-p;p) \geq F'(p)\omega^{\#}(x-p;p)\ . \qquad (5)$$

Hanson and Mond (1987) also considered a variant of *invex* (called *type I invex*), defined (in present notation), for $F = (f,g)$, by

$$(\forall x\in\mathbf{R}^n)\ f(x) - f(p) \geq f'(p)\omega(x-p;p); \qquad -g(p) \geq g'(p)\omega(x). \qquad (6)$$

Denote $\Phi(x) = (f(x) - f(p), g(x))$. Then this variant of invex is equivalent to:

$$(\forall x) \qquad \Phi(p) + \Phi^{\#}(x-p;p) \geq \Phi'(p)\omega^{\#}(x-p;p)\ . \qquad (7)$$

This property implies (see Hanson and Mond (1987)) that a KKT point is a minimum of $f(x)$, subject to $g(x) \leq 0$, since $0 \geq g(x) \geq g'(p)\omega(x-p;p)$,

hence for some multiplier $\lambda \geq 0$,

$$f(x) - f(p) \geq f'(p)\omega(x-p;p) = -\lambda^T g'(p)\omega(x-p;p) \geq 0. \qquad (8)$$

3. Characterization

The following characterizations can be given for the invex properties (5) and (7).

Theorem 1 *F is invex at p if and only if*

$$[0 \neq \alpha \in \mathbf{R}_+^m,\ \alpha^T F'(p) = 0] \Rightarrow \alpha^T F^{\#}(x-p;p) \geq 0\ . \qquad (9)$$

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, substituting $\zeta = z/t$, and then $u := (z,t)$,

$$\begin{aligned}
(5\ ) &\Leftrightarrow (\exists\zeta)\ M\zeta \leq \theta \\
&\Leftrightarrow (\exists z\in\mathbf{R}^n,\ t\in\mathbf{R})\ Mz \leq t\theta,\ t > 0 \\
&\Leftrightarrow (\exists u \in \mathbf{R}^{n+1})\quad Ku \leq 0,\ Nu < 0 \quad \text{where } K := [M,-\theta],\ N := [0,-1] \\
&\Leftrightarrow \textbf{NOT}\ (\exists\alpha \in \mathbf{R}_+^m,\ \exists\ \beta \in \mathbf{R})\ \ \alpha^T K + \beta N = 0,\ \alpha \geq 0,\ 0 \neq \beta \geq 0 \\
&\qquad \text{by Motzkin's alternative theorem} \\
&\Leftrightarrow \textbf{NOT}\ (\exists\alpha \in \mathbf{R}_+^m)\ \ \alpha^T M = 0,\ \ 0 \neq \alpha^T\theta \leq 0 \\
&\Leftrightarrow [0 \neq \alpha \in \mathbf{R}_+^m,\ \alpha^T M = 0] \Rightarrow \alpha^T\theta \geq 0\ . \qquad \square
\end{aligned}$$

Remark 1 Glover (1985) gave a similar result, but referring to $F(x) - F(0)$ instead of $F^{\#}(x-p;p)$. The point p in the theorem is not always a KKT point, because the complementarity slackness condition was not assumed. It is assumed in the next theorem.

Theorem 2 *$F := (f,g)$ has the modified invex property (7) at p if and only if $\alpha^T F^{\#}(x-p;p) \geq 0$ whenever p is a KKT point, with Lagrange multiplier $\alpha = (1,\lambda)$, for minimization of $f(.)$ subject to $g(.) \leq 0$.*

Proof For a fixed $x \in \mathbf{R}^n$, set $\theta := \Phi^{\#}(x-p;p)$, $\zeta := \omega^{\#}(x-p;p)$, $M = \Phi'(p)$, and $c := \Phi(p)$. Then, as in the proof of Theorem 1,

$$\begin{aligned}
(7\ ) &\Leftrightarrow (\exists\zeta)\ M\zeta \leq c + \theta \Leftrightarrow (\exists\zeta)\ \ -M\zeta + \theta \in \mathbf{R}_+^m - c \\
&\Leftrightarrow (\exists z\in\mathbf{R}^n,\ t\in\mathbf{R})\ \ -Mz + t\theta \in T := \text{cone}(\ \mathbf{R}_+^m - c),\ t > 0 \\
&\Leftrightarrow (\exists u \in \mathbf{R}^{n+1})\ -Ku \in T,\ Nu < 0,\ \text{where } K := [M,-\theta],\ N := [0,-1] \\
&\Leftrightarrow \textbf{NOT}\ (\exists\alpha \in \mathbf{R}_+^m,\ \exists\ \beta \in \mathbf{R})\ \ \alpha^T K + \beta N = 0,\ \alpha \in T^*,\ 0 \neq \beta \geq 0 \\
&\qquad \text{by Motzkin's alternative theorem, where the dual cone } T^* \\
&\qquad \text{of } T \text{ is characterized by } \alpha \in T^* \Leftrightarrow [\ \alpha \in \mathbf{R}_+^m,\ \alpha^T c = 0] \\
&\Leftrightarrow \textbf{NOT}\ \ (\exists\alpha \in \mathbf{R}_+^m,\ \alpha^T c = 0)\ \ \alpha^T M = 0,\ \ -\alpha^T\theta - \beta = 0,\ 0 \neq \beta \geq 0 \\
&\Leftrightarrow [\alpha \in \mathbf{R}_+^m,\ \alpha^T c = 0,\ \ \alpha^T M = 0] \Rightarrow \alpha^T\theta \geq 0\ . \qquad \square
\end{aligned}$$

4. The Wolfe dual

The Wolfe dual for the minimization problem:

$$\text{MIN} \quad f(x) \ \text{ subject to } g(x) \geq 0, \tag{10}$$

(with $f : \mathbf{R}^n \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}^m$ differentiable, $m < n$) is the problem:

$$\text{MAX } f(u) + v^T g(u) \text{ subject to } v \geq 0,\ f'(u) + v^T g'(u) = 0. \tag{11}$$

Let $F := (f,g)$. It is well known that duality holds if a constraint qualification holds, and if F is invex at each point.

Theorem 3 *For each feasible point (u,v) for (11), assume that $g'(u)$ has full rank, $f'(u) \neq 0$, and*

$$(\forall z) \quad f(z) + v^T g(z) \geq f(u) + v^T g(u)\ . \tag{12}$$

Then (f,g) is invex at (u,v); and weak duality holds for (10) and (11).

Remark 2 Condition (12) also follows from invexity, so that Theorem 3 gives a necessary and sufficient condition. Points (z,v) with $z \neq u$ are not feasible for (11).

Define the Lagrangian $L(z,w) := f(z) + w^T g(z)$. Under the hypotheses of Theorem 3, if u satisfies Karush-Kuhn-Tucker necessary conditions for (10), then L satisfies the saddlepoint property:

$$(\forall z)\ (\forall w \geq 0) \quad L(u,w) \leq L(u,v) \leq L(z,v)\ .$$

(The left inequality follows from $w \geq 0$, $g(u) \leq 0$, and $v^T g(u) = 0$.)

For fixed v, denote $\Psi(z) := f(z) + v^T g(z)$. Since $\Psi'(u) = 0$, (12) is equivalent to the second-order condition

$$(\forall z)\ \Psi^{\#}(z-u;u) \geq 0.$$

Under hypotheses of Theorem 3, The Wolfe dual (11) is equivalent to

$$\text{MAX}_{u,v}\ \text{MIN}_z\ f(z) + v^T g(z) \text{ subject to } v \geq 0,\ f'(u) + v^T g'(u) = 0. \tag{13}$$

Since, given invexity, $f(.) + v^T g(.)$ is minimized when $f'(u) + v^T g'(u) = 0$, problem (13) is equivalent to the Lagrangian dual to (P), namely:

$$\text{MAX}_{v \geq 0}\ \text{MIN}_z\ f(z) + v^T g(z)\ . \tag{14}$$

Proof Let (u,v) be feasible for (11). Since $g'(u)$ has full rank, there is no other feasible point (u,v') with $v' \neq v$. Now

$$\begin{aligned}
[1 \ \ v^T]F^{\#}(z-u;u) &= [1 \ \ v^T](F(z) - F(u) - F'(u)(z-u)) \\
&= [1 \ \ v^T](F(z) - F(u)) \ - 0 \\
&\qquad\qquad \text{from a constraint in (11)} \\
&= [f(z) + v^T g(z)] - [f(u) + v^T g(u)]\ .
\end{aligned} \tag{15}$$

In order to apply Theorem 1, let α satisfy:

$$\alpha^T := \quad [\gamma \ v^T] \geq [0 \ \ 0] \text{ and } [\gamma \ v^T]\begin{bmatrix} f'(u) \\ g'(u) \end{bmatrix} = 0$$

If $\gamma = 0$ then $v^T g'(u) = 0$, hence $v = 0$ since $g'(u)$ has full rank; so $f'(u) = 0$, contrary to hypotheses. So $\gamma = 1$ may be assumed. Then

$$[1 \;\; v^T]F'(u) = 0 \Rightarrow [1 \;\; v^T]F^{\#}(z-u;u) \geq 0$$

if and only if

$$f'(u) + v^T g'(u) = 0 \;\Rightarrow\; f(z) + v^T g(z) \geq f(u) + v^T g(u) .$$

Then, by Theorem 1, (f,g) is invex at (u,v) if and only if (12) holds; and invexity is sufficient for weak duality. □

Remark 3 The Mond-Weir dual for (10) is the problem:

$$\text{MAX } f(u) \text{ subject to } v \geq 0,\; f'(u) + v^T g'(u) = 0,\; v^T g(u) \geq 0. \qquad (16)$$

A proof similar to that of Theorem 3 shows that the hypotheses of Theorem 3 are also sufficient for duality of (10) and (16).

5. Locally Lipschitz functions

Suppose now that F is locally Lipschitz at p, but not necessarily differentiable there. If $\xi \in \partial F(p)$ (the Clarke generalized Jacobian of F at p), denote $F^{\#}(x-p;p;\xi) := F(x) - F(p) - \xi(x-p)$. Now F has been defined in Craven (1986) to be *generalized invex* at p if $\eta(.,.)$ exists so that

$$(\forall \xi \in \partial F(p))\; F(x) - F(p) \geq \xi\eta(x,p) .$$

This is equivalent to the following analog of the condition (5):

$$(\forall \xi \in \partial F(p) \;\; F^{\#}(x-p;p;\xi) \geq \xi\omega^{\#}(x-p;p) \;. \qquad (17)$$

Note that $\omega(x-p,p) := \eta(x,p)$ is still assumed differentiable.

The Wolfe dual for problem (P) now becomes:

$$\text{MIN } f(u) + vTg(u) \text{ subject to } v \geq 0,\; 0 \in \partial(f + v^T g)(u) . \qquad (18)$$

Theorems 1 and 3 now generalize as follows to the locally Lipschitz case.

Theorem 4 *F is generalized invex at p if and only if*

$$(\forall \xi \in \partial F(p) \quad [0 \neq \alpha \in \mathbf{R}_+^m,\; \alpha^T \xi = 0] \Rightarrow \alpha^T F^{\#}(x-p;p;\xi) \geq 0 . \qquad (19)$$

Proof As for Theorem 1, replacing M by ξ and $F^{\#}(x-p;p)$ by $F^{\#}(x-p;p;\xi)$.

Theorem 5 *For each feasible point (u,v) for (18), assume that each element of $\partial g(u)$ has full rank, $0 \notin \partial f(u)$, and*

$$(\forall z) \quad f(z) + v^T g(z) \geq f(u) + v^T g(u) . \qquad (20)$$

Then F := (f,g) is invex at (u,v); and weak duality holds for (10) and (18).

Proof If $\xi \in \partial F(u)$, then

$$\begin{aligned} [1 \;\; v^T]F^{\#}(z-u;u;\xi) &= [1 \;\; v^T](F(z) - F(u) - \xi(z-u)) \\ &= [1 \;\; v^T](F(z) - F(u)) - 0 \\ &\qquad \text{since } (u,v) \text{ is feasible for (18) and } \xi = 0 \\ &= [f(z) + v^T g(z)] - [f(u) + v^T g(u)] . \end{aligned}$$

Suppose that

$$\alpha^T := [\gamma\ v^T] \geq [0\ \ 0] \text{ and } [\gamma\ v^T]\xi = 0.$$

The case $\gamma = 0$ is excluded by the hypothesis that each element of $\partial g(u)$ has full rank. So $\gamma = 1$ may be assumed. Then

$$(\alpha \geq 0 \text{ and } \alpha^T\xi = 0) \Rightarrow \alpha^T F^{\#}(z-u;u;\xi) \geq 0$$

holds if and only if

$$0 \in \partial(f + v^T g)(u) \Rightarrow [f(z) + v^T g(z)] \geq [f(u) + v^T g(u)] .$$

From Theorem 4, the latter condition is equivalent to generalized invexity of F, and so implies weak duality. □

Acknowleedgement

The author thanks two referees for a number of corrections.

References

Craven, B. D. (1981a), Duality for generalized convex fractional programs, in Schaible, S. and Ziemba, W. T. (eds.), *Generalized Concavity in Optimization and Economics*, Academic Press, New York , 473-490.

Craven, B. D. (1981b), Invex functions and constrained local minima, Bull. Austral. Math. Soc. 25 , 37- 46.

Craven, B. D. (1986), Nondifferentiable optimization by smooth approximations, Optimization **17** , 1-17.

Craven, B. D. (1995), *Control and Optimization*, Chapman & Hall, London

Craven, B. D. and Glover, B. M. (1985), Invex functions and duality J. Austral. Math. Soc. Series A, 39 , 1-20.

Hanson, M. A. (1980), On the sufficiency of the Kuhn-Tucker conditions, J. Math. Anal. Appl. **80**, 545-550.

Hanson, M. A. and Mond, B. (1987), Necessary and sufficient conditions in constrained optimization, Mathematical Programming **37**, 51-58.