GLOBAL INVEXITY

B.D. CRAVEN

Abstract. 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.

1. Introduction

The *invex* property [6], [1], [4] of a vector function has more often been postulated, than shown to hold, for some classes of constrained optimization problem. Conditions for *local invexity* (thus, invexity in a local domain) were obtained in [2], 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. Therefore, another type of approximation is required.

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

$$(\forall x \in \mathbb{R}^n) \quad F(x) - F(p) \geq F'(p)\,\eta(x,p)\,. \tag{1.1}$$

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 $\omega(x-p)$. Since a term in $F'(p)^{-1}(0)$ can be subtracted from $\eta(x,p)$, it can be assumed (see [2]) that $\eta(x,p) = x - p + o(\|x-p\|)$.

It is observed that a function may be invex at one point, but not at other neighbouring points. An example is: $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = -x^2$. Invexity at $p = 0$ would require that $(\forall x) - x^2 \geq -0 \cdot \omega(x)$, which is a contradiction. However, invexity at $p > 0$ requires that (setting $z = x - p$) $-z^2 \geq (-2p)\,\omega(z)$, 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 $x \in \mathbb{R}^n$, 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 is:

$$F_1'(p) + \sum_{j \geq 1} \rho_j F_j'(p) = 0, \quad \rho_j \geq 0, \; \rho_j F_j(p) = 0 \quad (j \geq 1)\,. \tag{1.2}$$

Date: August 23, 2000.
1991 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18.
Key words and phrases. invex, global, duality, saddlepoint.

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 multiplier $\rho_j > 0$, 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

$$F_1'(x) + \sum_{j\geq 1} \rho_j F_j'(x) = 0, \quad \rho_j \geq 0. \tag{1.3}$$

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

$$\begin{aligned} F(x) &= F(p) + F'(p)(x-p) + F^{\#}(x-p,p); \\ \omega(x) &= \omega(p) + \omega'(p)(x-p) + \omega^{\#}(x-p,p); \end{aligned} \tag{2.1}$$

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

$$(\forall x) \quad F^{\#}(x-p,p) \geq F'(p)\,\omega^{\#}(x-p,p). \tag{2.2}$$

Hanson and Mond [7] also considered a variant of *invex* (called *type I invex*) defined (in present notation), for $F = (f,g)$, by

$$(\forall x \in \mathbb{R}^n) \quad f(x) - f(p) \geq f'(p)\,\omega(x-p,p); \quad g(x) \geq g'(p)\,\omega(x). \tag{2.3}$$

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

$$(\forall x) \quad \Phi(p) + \Phi^{\#}(x-p,p) \geq F'(p)\,\omega^{\#}(x-p,p). \tag{2.4}$$

This property implies (see Hanson and Mond [7]) 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 $\lambda \geq 0$,

$$f(x) - f(p) \geq f'(p)\,\omega(x-p,p) = -\lambda^T g'(p)\,\omega(x-p,p) \geq 0. \tag{2.5}$$

3. Characterization

The following characterizations can be given for the invex properties (2.2) and (2.4).

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

$$\left[0 \neq \alpha \in \mathbb{R}^m_+,\ \alpha^T F'(p) = 0\right] \Rightarrow \alpha^T F^{\#}(x-p,p) \geq 0. \tag{3.1}$$

Proof. For a fixed $x \in \mathbb{R}^n$, set $\theta := F^{\#}(x-p;p)$ and $\zeta := \omega^{\#}(x-p;p)$, and set $M := F'(p)$. Then, substituting $\zeta = \frac{z}{t}$, and then $u := (z,t)$,

$$\begin{aligned}
(2.2) \iff & (\exists\zeta)\, M\zeta \le \theta \\
\iff & (\exists z \in \mathbb{R}^n,\ t \in \mathbb{R})\, Mz \le t\theta,\ t > 0 \\
\iff & \left(\exists u \in \mathbb{R}^{n+1}\right)\ Ku \le 0,\ Nu < 0 \ \text{ where } \ K := [M, -\theta],\ N := [0,-1] \\
\iff & \text{NOT}\left(\exists\alpha \in \mathbb{R}^m_+,\ \exists\beta \in \mathbb{R}\right)\ \alpha^T K + \beta N = 0,\ \alpha \ge 0,\ 0 \ne \beta \ge 0 \\
& \text{by Motzkin's alternative theorem} \\
\iff & \text{NOT}\left(\exists\alpha \in \mathbb{R}^m_+\right)\ \alpha^T M = 0,\ 0 \ne \alpha^T\theta \le 0 \\
\iff & \left[0 \ne \alpha \in \mathbb{R}^m_+,\ \alpha^T M = 0\right] \Rightarrow \alpha^T\theta \ge 0.
\end{aligned}$$

∎

Remark 1. *Glover* [5] *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 (2.4) at* p *if and only if* $\alpha^T F^{\#}(x-p,p) \ge 0$ *whenever* p *is a KKT point, with Lagrange multiplier* $\alpha = (1,\lambda)$, *for minimization of* $f(\cdot)$ *subject to* $g(\cdot) \le 0$.

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

$$\begin{aligned}
(2.4) \iff & (\exists\zeta)\, M\zeta \le c + \theta \iff (\exists\zeta) - M\zeta + \theta \in \mathbb{R}^m_+ - c \\
\iff & (\exists z \in \mathbb{R}^n,\ t \in \mathbb{R}) - Mz + t\theta \in T := \text{cone}\left(\mathbb{R}^m_+ - c\right),\ t > 0 \\
\iff & \left(\exists u \in \mathbb{R}^{n+1}\right) - Ku \in T,\ Nu < 0 \ \text{ where } \ K := [M, -\theta],\ N := [0,-1] \\
\iff & \text{NOT}\left(\exists\alpha \in \mathbb{R}^m_+,\ \exists\beta \in \mathbb{R}\right)\ \alpha^T K + \beta N = 0,\ \alpha \in T^*,\ 0 \ne \beta \ge 0 \\
& \text{by Motzkin's alternative theorem, where the dual cone } T^* \\
& \text{is characterised by } \alpha \in T^* \iff \left[\alpha \in \mathbb{R}^m_+,\ \alpha^T c = 0\right] \\
\iff & \text{NOT}\left(\exists\alpha \in \mathbb{R}^m_+,\ \alpha^T c = 0\right)\ \alpha^T M = 0,\ -\alpha^T\theta - \beta = 0,\ 0 \ne \beta \le 0 \\
\iff & \left[\alpha \in \mathbb{R}^m_+,\ \alpha^T c = 0,\ \alpha^T M = 0\right] \Rightarrow \alpha^T\theta \ge 0.
\end{aligned}$$

∎

4. The Wolfe Dual

The Wolfe dual for the minimization problem:

$$MIN\ f(x) \ \text{ subject to } g(x) \ge 0, \tag{4.1}$$

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

$$MAX\ f(u) + v^T g(u) \ \text{ subject to } v \ge 0,\ f'(u) + v^T g'(u) = 0. \tag{4.2}$$

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

Theorem 3. *For each feasible point* (u,v) *for (4.2) , assume that* $g'(u)$ *has full rank,* $f'(u) \ne 0$, *and*

$$(\forall z) \quad f(z) + v^T g(z) \ge f(u) + v^T g(u). \tag{4.3}$$

Then (f,g) is invex at (u,v); and weak duality holds for (4.1) and (4.2).

Remark 2. *Condition (4.3) 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 (4.2).*

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 (4.1), then L satisfies the saddlepoint property:

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

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

For fixed v, denote $\Psi(z) := f(z) + v^T g(z)$. Since $\Psi'(u) = 0$, (4.3) is equivalent to the second-order condition

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

Under hypotheses of Theorem 3, the Wolfe dual (4.2) is equivalent to

$$\text{(4.4)} \qquad MAX_{u,v}\ MIN_z f(z) + v^T g(z) \text{ subject to } v \geq 0,\ f'(u) + v^T g'(u) = 0.$$

Since, given invexity, $f(\cdot) + v^T g(\cdot)$ is minimized when $f'(u) + v^T g'(u) = 0$, problem (4.4) is equivalent to the Lagrangian dual to (P), namely:

$$\text{(4.5)} \qquad MAX_{v \geq 0}\ MIN_z f(z) + v^T g(z).$$

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

$$\text{(4.6)}\ \begin{aligned} \begin{bmatrix} 1 & v^T \end{bmatrix} F^{\#}(z-u;u) &= \begin{bmatrix} 1 & v^T \end{bmatrix} (F(z) - F(u) - F'(u)(z-u)) \\ &= \begin{bmatrix} 1 & v^T \end{bmatrix} (F(z) - F(u)) \\ &= \left[f(z) + v^T g(z)\right] - \left[f(u) + v^T g(u)\right]. \end{aligned}$$

Suppose that

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

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

$$\begin{bmatrix} 1 & v^T \end{bmatrix} F'(u) = 0 \Rightarrow \begin{bmatrix} 1 & v^T \end{bmatrix} F^{\#}(z-u;u) \geq 0$$

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 (4.3) holds; and invexity is sufficient for weak duality. ∎

Remark 3. *The Mond-Weir dual for (4.1) is the problem:*

$$\text{(4.7)} \qquad MAX\ f(u) \textit{ subject to } v \geq 0,\ f'(u) + v^T g'(u) = 0,\ v^T g(u) \geq 0.$$

A proof similar to that of Theorem 3 shows that the hypotheses of Theorem 3 are also sufficient for duality of (4.1) and (4.7).

5. Locally Lipschitz Functions

Suppose now that F is locally Lipschitz at p, but not necessarily differentiable there. If $\xi = \partial F(p)$ (The Clarke generalised Jacobian of F at p), denote $F^{\#}(x-p;p;\xi) := F(x) - F(p) - \xi(x-p)$. Now F has been defined [3] to be *generalised invex* at p if $\eta(\cdot,\cdot)$ exists so that

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

This is equivalent to the following analog of the condition (2.2):

$$(\forall \xi \in \partial F(p)) \quad F^{\#}(x-p;p;\xi) \geq \xi \omega^{\#}(x-p;p). \tag{5.1}$$

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

The Wolfe dual for problem (P) now becomes:

$$MIN\ f(u) + v^T g(u) \text{ subject to } v \geq 0,\ 0 \in \partial\left(f + v^T g\right)(u). \tag{5.2}$$

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

Theorem 4. *F is generalised invex at p if and only if*

$$(\forall \xi \in \partial F(p)) \quad \left[0 \neq \alpha \in \mathbb{R}^m_+,\ \alpha^T \xi = 0\right] \Rightarrow \alpha^T F^{\#}(x-p;p;\xi) \geq 0. \tag{5.3}$$

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 (5.2), assume that each element of $\partial g(u)$ has full rank, $0 \notin \partial f(u)$, and*

$$(\forall z) \quad f(z) + v^T g(z) \geq f(u) + v^T g(u). \tag{5.4}$$

Then $F := (f,g)$ is invex at (u,v); and weak duality holds for (4.1) and (5.2).

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

$$\begin{aligned} \begin{bmatrix} 1 & v^T \end{bmatrix} F^{\#}(z-u;u;\xi) &= \begin{bmatrix} 1 & v^T \end{bmatrix} \left(F(z) - F(u) - F'(u)(z-u)\right) \\ &= \begin{bmatrix} 1 & v^T \end{bmatrix} \left(F(z) - F(u)\right) \\ &= \left[f(z) + v^T g(z)\right] - \left[f(u) + v^T g(u)\right]. \end{aligned}$$

Suppose that

$$\alpha^T = \begin{bmatrix} \gamma & v^T \end{bmatrix} \geq \begin{bmatrix} 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \gamma & v^T \end{bmatrix} \xi = 0.$$

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

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

holds if and only if

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

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

References

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

[2] B.D. CRAVEN, Invex functions and constrained local minima, *Bull. Astral. Math. Soc.,* **25** (1981), 37-46.

[3] B.D. CRAVEN, Nondifferential optimization by smooth approximations, *Optimization,* **17** (1986), 1-17.

[4] B.D. CRAVEN, *Control and Optimization,* Chapman & Hall, London (1995).

[5] B.D. CRAVEN and B.M. GLOVER, Invex functions and duality, *J. Austral. Math. Soc. Series A,* **39** (1985), 1-20.

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

[7] M.A. HANSON and B. MOND, Necessary and sufficient conditions on constrained optimization, *Mathematical Programming,* **37** (1987), 51-58.

Department of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria 3052, Australia.

E-mail address: `B.Craven@ms.unimelb.edu.au`