BULL. AUSTRAL. MATH. SOC. 90C30
VOL. 24 (1981), 357-366.

INVEX FUNCTIONS AND CONSTRAINED LOCAL MINIMA

B.D. CRAVEN

If a certain weakening of convexity holds for the objective and all constraint functions in a nonconvex constrained minimization problem, Hanson showed that the Kuhn-Tucker necessary conditions are sufficient for a minimum. This property is now generalized to a property, called K-invex, of a vector function in relation to a convex cone K. Necessary conditions and sufficient conditions are obtained for a function f to be K-invex. This leads to a new second order sufficient condition for a constrained minimum.

1. Introduction

A real function $f : \mathbb{R}^n \to \mathbb{R}$ will be called *invex*, with respect to η, if for the function $\eta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$,

(1) $$f(x) - f(u) \geq f'(u)\eta(x, u)$$

holds for each x and u in the domain of f. Here $f'(u)$ is the Fréchet derivative of f at u. Hanson [5] (see also [6], [7]) introduced this concept, and showed that, if all the functions f_i in the (nonconvex) constrained minimization problem,

(2) Minimize $f_0(x)$ subject to $f_i(x) \leq 0 \quad (i = 1, 2, \ldots, m)$,

are invex, with respect to the same η, then the Kuhn-Tucker conditions

Received 27 July 1981.

necessary for a global minimum of (2) are also sufficient. In fact, Hanson's proof [5] does not require (1) at all points x and u, since u may be fixed at the point where the Kuhn-Tucker conditions hold. Define, therefore, a real function f to be *invex in a neighbourhood at* u if f satisfies (1) for a given u, and for all x such that $\|x-u\|$ is sufficiently small. With this definition, the Kuhn-Tucker necessary conditions become also sufficient for a local minimum; the proof is the same as Hanson's. Craven [2] has shown that f has the invex property when $f = h \circ \phi$, with h convex, ϕ differentiable, and ϕ' having full rank. Thus some invex functions, at least, may be obtained from convex functions by a suitable transformation of the domain space. Such transformations destroy convexity, but not the invex property; the term *invex*, from invariant convex, was introduced in [2] to express this fact. (In (1), f is convex if $\eta(x, u) = x - u$.)

The requirement that all the functions f_i in (2) are invex with respect to the same function η may be expressed by forming a vector f, whose components are f_i $(i = 0, 1, 2, \ldots, m)$, and then requiring that

$$(3) \qquad f(x) - f(u) - f'(u)\eta(x, u) \in \mathbb{R}_+^{m+1} ,$$

where $\mathbb{R}_+^{m+1}$ denotes the nonnegative orthant in $\mathbb{R}^{m+1}$. More generally, let $K \subset \mathbb{R}^{m+1}$ be a convex cone. The vector function $f : \mathbb{R}^n \to \mathbb{R}^{m+1}$ will be called *K-invex*, with respect to η, if

$$(4) \qquad f(x) - f(u) - f'(u)\eta(x, u) \in K$$

for all x and u. If u is fixed, and (4) holds whenever $\|x-u\|$ is sufficiently small, then f will be called *K-invex*, with respect to η, *in a neighbourhood at* u. It is noted that, if f is K-invex in a neighbourhood at u, and if $v \in K^*$ (the dual cone of K, thus $v(K) \subset \mathbb{R}_+ \equiv [0, \infty)$), then $v^T f$ is invex in a neighbourhood at u, with respect to the same η.

In this paper, conditions are obtained necessary, or sufficient, for f to be K-invex with respect to some η. This involves an investigation of appropriate functions η for (4). To motivate the generalization to cones, consider problem (2) generalized to

(5) Minimize $f_0(x)$ subject to $-g(x) \in S$,

where $g(x) = \left(f_1(x), f_2(x), \ldots, f_m(x)\right)^T$, and $S \subset \mathbb{R}^m$ is a convex cone. The Kuhn-Tucker conditions necessary (assuming a constraint qualification) for a minimum of (5) at $x = u$ are that a Lagrange multiplier $\theta \in S^*$ exists, for which

$$(6) \qquad f_0'(u) + \theta^T g'(u) = 0 \; ; \quad \theta^T g(u) = 0 \; .$$

Set $\lambda = (1, \theta)$, and $K = \mathbb{R}_+ \times S^*$. (If $S = \mathbb{R}_+^m$ then $K = \mathbb{R}_+^{m+1}$.) Then (6) may be rewritten as

$$(7) \qquad \lambda^T f'(u) = 0 \; ; \quad \lambda^T f(u) = f_0(u) \; ; \quad \lambda \in K^* \; .$$

The following converse Kuhn-Tucker theorem then holds.

THEOREM 1. *Let* u *be feasible for problem* (5); *let the Kuhn-Tucker conditions* (7) *hold at* u , *with* $\lambda_0 = 1$; *let* f *be* K*-invex, with respect to some* η , *in a neighbourhood at* u . *Then* u *is a local minimum of* (5).

Proof. Let x be any feasible point for (5), with $\|x-u\|$ sufficiently small. Then

$$\begin{aligned} f_0(x) - f_0(u) &\geq \lambda^T f(x) - \lambda^T f(u) \quad \text{since} \quad \theta^T g(x) \leq 0 \quad \text{and} \quad \lambda^T f(u) = f_0(u) \\ &\geq \lambda^T f'(u)\eta(x, u) \quad \text{by the invex hypothesis} \\ &= 0 \quad \text{by the Kuhn-Tucker conditions.} \end{aligned}$$

So u is a local minimum for (5). □

The result generalizes [7], Theorem 2, which applies to polyhedral cones S only.

If the problem (5) is not convex, then the hypotheses (and proof) of Theorem 1 lead to a local (but not necessarily global) minimum. A local minimum also follows if f is U-invex, where U is a convex cone containing K , and $\lambda \in U^*$. Here the vector function f is less restricted than in Theorem 1, and the Lagrange multiplier λ is more restricted.

The problem (2) is *equivalent* to the transformed problem,

(8) Minimize $\phi_0 \circ f_0(x)$ subject to $\phi_i \circ f_i(x) \le 0$ $(i = 1, 2, \ldots, m)$,

in the sense that both problems (2) and (8) have the same feasible set and the same minimum (local or global), provided that $\phi_0 : \mathbb{R} \to \mathbb{R}$ is strictly increasing and, for $i = 1, 2, \ldots, m$, $\phi_i(\mathbb{R}_+) \subset \mathbb{R}_+$ and $\phi_i(-\mathbb{R}_+) \subset -\mathbb{R}_+$. (The ϕ_i , for $i = 1, 2, \ldots, m$, may be monotone, but need not be so.) Let F denote the vector whose components are $\phi_0 \circ f_0, \phi_1 \circ f_1, \ldots, \phi_m \circ f_m$. The converse Kuhn-Tucker property will hold for the original problem (2) if it holds for the transformed problem (8), thus if F is $\mathbb{R}_+^{m+1}$-invex, with respect to some η . Although (8) is not generally a convex problem, a local minimum for (8) at u follows, as in Theorem 1; and this implies a local (and hence global) minimum at u for the convex problem (2).

2. Conditions necessary or sufficient for an invex function

Assume now that the vector functions f and η are twice continuously differentiable. For fixed u , the Taylor expansion of $\eta(x, u)$ in terms of $v = x - u$ gives, up to quadratic terms

(9) $$\eta(x, u) = \eta_0 + Av + \tfrac{1}{2}v^T Q_\bullet v + o\left(\|v\|^2\right) \quad (v = x - u) ,$$

where A is an $n \times n$ matrix of first partial derivatives, and $v^T Q_\bullet v$ is a coordinate-free notation (see [3]) for the vector whose kth component is

(10) $$\sum_{i,j=1}^{n} v_i Q_{k,ij} v_j \text{ , where } Q_{k,ij} = \left.\frac{\partial^2 \eta_k(x,u)}{\partial x_i \partial x_j}\right|_{x=u} .$$

Of course, $\eta_0 \equiv \eta(u, u)$, A , and $Q_\bullet$ depend on u . If q is a row vector with n components, let $qQ_\bullet$ denote the matrix Q_k whose elements are $\sum_{k=1}^{n} q_k Q_{k,ij}$. Thus $q\left(v^T Q_\bullet v\right) = v^T (qQ_\bullet)_v$, an ordinary quadratic form.

Similarly, f has an expansion

(11) $$f(x) - f(u) = Bv + \tfrac{1}{2}v^T M_\bullet v + o\left(\|v\|^2\right) ,$$

https://doi.org/10.1017/S0004972700004895 Published online by Cambridge University Press

where $B = f'(u)$ is a matrix of first derivatives $\partial f_k(x)/\partial x_i|_{x=u}$, and

(12) $$M_{k,ij} = \left.\frac{\partial^2 f_k(x)}{\partial x_i \partial x_j}\right|_{x=u}$$

Let S be a convex cone in $\mathbb{R}^n$. The quadratic expression $v^T M_\bullet v$ will be called *S-positive semidefinite* if $v^T M_\bullet v \in S$ for every $v \in \mathbb{R}^n$. The expression $v^T M_\bullet v$ will be called *S-positive definite* if $v^T M_\bullet v \in \operatorname{int} S$ for every nonzero $v \in \mathbb{R}^n$. Here int S denotes the interior of S , supposed nonempty. Now $v^T Q_k v$ can be expressed, by rotation of axes, in the form $\sum_{i=1}^{n} \rho_{ki}\alpha_{ki}$, where the ρ_{ki} $(i = 1, 2, \ldots, n)$ are the eigenvalues of Q_k , and the α_{ki} , depending on v , are nonnegative. For each k , denote by ρ_k the vector whose components are $\rho_{k1}, \rho_{k2}, \ldots, \rho_{kn}$. If, for some ordering of the eigenvalues of each Q_k , every vector ρ_k lies in S (respectively in int S), then it follows that $v^T Q_\bullet v$ is S-positive semidefinite (respectively S-positive definite). This sufficient condition for S-positive semidefiniteness would be also necessary if the Q_k are simultaneously diagonalizable, but that is not usually the case. It is convenient to say that $Q_\bullet$ is S-positive (semi-) definite when $v^T Q_\bullet v$ is.

If S is a polyhedral cone, then the dual cone S^* has a finite set, G , of generators (considered as row vectors). Since a vector $a \in S$ if and only if $qs \geq 0$ for each $q \in G$, it follows that Q is S-positive (semi-) definite if and only if, for each $q \in G$, qQ is positive (semi-) definite in the usual sense.

Let $r = m + 1$. If B is an $r \times n$ matrix, define $v^T(BQ_\bullet)v$ for $v \in \mathbb{R}^n$ as the vector whose kth component is

https://doi.org/10.1017/S0004972700004895 Published online by Cambridge University Press

(13) $$\sum_{i,j=1}^{n} v_i C_{k,ij} v_j \quad \text{where} \quad C_{k,ij} = \sum_{t=1}^{r} B_{kt} Q_{t,ij} .$$

Let I denote the $n \times n$ identity matrix.

Observe that the K-invex property (4) is unaffected by subtracting from η any term in the nullspace of $f'(u) \equiv B$.

THEOREM 2. *Let $f : \mathbb{R}^n \to \mathbb{R}^r$ be twice continuously differentiable: let $K \subset \mathbb{R}^r$ be a closed convex cone, satisfying $K \cap (-K) = \{0\}$. If f is K-invex in a neighbourhood at u , with respect to a twice continuously differentiable function η , for which $\eta(u, u) = 0$, then, after subtraction of a term in the nullspace of B , η has the form*

(14) $$\eta(u+v, u) = v + \tfrac{1}{2}v^T Q_\bullet v + o\left(\|v\|^2\right) ,$$

where $M_\bullet - BQ_\bullet$ is K-positive semidefinite. Conversely, if η has the form (14), *and if $M_\bullet - BQ_\bullet$ is K-positive definite, then f is K-invex in a neighbourhood at u , with respect to this η .*

Proof. Let f be K-invex with respect to η in a neighbourhood at u . Substituting the expansion (9) into the expansion (11), and setting $\eta_0 = 0$, the inequality

(15) $$\left[Bv+\tfrac{1}{2}v^T M_\bullet v+o\left(\|v\|^2\right)\right] - B\left[Av+\tfrac{1}{2}v^T Q_\bullet v+o\left(\|v\|^2\right)\right] \in K$$

must hold, whenever $\|v\|$ is sufficiently small. Considering the terms linear in v , $Bv - BAv + o(\|v\|) \in K$ for each $v \in \mathbb{R}^n$. Hence, for each $q \in K^*$, and each v , $q\left(B(I-A)v+o(\|v\|)\right) \geq 0$, hence $qB(I-A)v \geq 0$. Hence $B(I-A)v \in K \cap (-K) = \{0\}$. Therefore $B(I-A) = 0$. But the definition (4) of K-invex allows any term in the nullspace of B to be added to η . Hence f is also K-invex with respect to η , now modified by replacing A by I . The quadratic terms then require that, for each v ,

(16) $$v^T(M_\bullet - BQ_\bullet)v + o\left(\|v\|^2\right) \in K .$$

Hence, for each $q \in K^*$ and each $\alpha > 0$, replacing v by αv ,

$$q\left[v^T(M_\bullet - BQ_\bullet)v\right] + o\left(\alpha^2\right)/\alpha^2 \geq 0 .$$

Hence $q\left[v^T(M_\bullet - BQ_\bullet)v\right] \geq 0$ for each $q \in K^*$, hence

(17) $$v^T(M_\bullet - BQ_\bullet)v \in K \text{ for each } v \in \mathbb{R}^n .$$

Thus $M_\bullet - BQ_\bullet$ is K-positive semidefinite.

Conversely, assume that $M_\bullet - BQ_\bullet$ is K-positive definite. A reversal of the above argument shows that (15), with $A = I$, is satisfied up to quadratic terms, with K replaced by int K . Here the quadratic terms dominate any higher order terms, so that (15) itself holds, whenever $\|v\|$ is sufficiently small. Thus (4) follows, and f is K-invex, in a neighbourhood at u , with η given by (14). □

If a nonzero term $\eta_0 \equiv (0, 0)$ is included in (9), then the K-invex property for f requires that $-B\eta_0 \in K$, on setting $v = 0$. Suppose that, for the constrained minimization problem (5), the Kuhn-Tucker conditions (6) hold at the point u . Then $\lambda^T B = 0$, for some nonzero $\lambda \in K^*$. Suppose that $\lambda \in \text{int } K^*$ (for problem (2), this means that each Lagrange multiplier $\lambda_i > 0$). From this, if $0 \neq -B\eta_0 \in K$, there follows (see [1], page 31) $\lambda^T B\eta_0 < 0$, contradicting $\lambda^T B = 0$. So the assumption that $B\eta_0 = 0$ is a relevant one, when the K-invex property is to be applied to Kuhn-Tucker conditions and Theorem 1. In Theorem 2, $\eta_0 = 0$ was assumed, since a vector in the nullspace of B may be subtracted from η .

In Theorem 2, the sufficient conditions for f to be K-invex involve first and second derivatives of f . Combining this with the sufficient Kuhn-Tucker theorem (Theorem 1), it has been shown that the Kuhn-Tucker conditions are sufficient for a local minimum of a nonconvex problem, if the first and second derivatives of f at the Kuhn-Tucker point u are suitably restricted. The Lagrangian $f_0(x) + \lambda^T g(x)$ for the problem (5) has $\lambda^T M$ as its matrix of second derivatives. The second order sufficiency conditions, given by Fiacco and McCormick [4], page 30, require that (in the present notation) each component of $v^T(\lambda^T M_\bullet)v > 0$ for each nonzero v in a certain cone. This is related to, but not the same as, the hypothesis (from Theorem 2) that $M_{\bullet} - BQ_{\bullet}$ is K-positive definite, for some choice of $Q_{\bullet}$. However, the construction of a suitable Q , given f , is a nontrivial matter, since the eigenvalues of each $M_k - (BQ_{\bullet})_k$ matrix are involved.

3. Examples

Consider the problem,

(18) $$\operatorname*{Minimize}_{x=(x_1,x_2)\in\mathbb{R}^2} \quad f_0(x) = \tfrac{1}{3}x_1^3 - x_2^2 \text{ subject to } f_1(x) = \tfrac{1}{2}x_1^2 + x_2^2 - 1 \le 0 .$$

This problem has a local minimum at $(0, 1)$, with Lagrange multiplier 1 . The matrices M_k are then

$$M_0 = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}, \text{ and } M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} ; \quad B = \begin{bmatrix} 0 & -2 \\ 0 & 2 \end{bmatrix} .$$

When do symmetric matrices Q_0 and Q_1 exist for which

$$\begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} - 0Q_0 - (-2)Q_1 \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} - 0Q_0 - 2Q_1$$

are both positive semidefinite, or both positive definite? Setting

$$Q_1 = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}, \text{ the two matrices are } \begin{bmatrix} 2\alpha & 2\beta \\ 2\beta & -2+2\gamma \end{bmatrix} \text{ and } \begin{bmatrix} 1-2\alpha & -2\beta \\ -2\beta & 2-2\gamma \end{bmatrix} .$$

Using the Routh-Hurwicz criterion, $0 \le \alpha \le 1$, $\gamma = 1$, $\alpha(-1+\gamma) - \beta^2 \ge 0$, and $(1-\alpha)(2-2\gamma) - 4\beta^2 \ge 0$, are required. Positive semidefinite matrices

$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}$$

are obtained, with $\alpha = \frac{1}{2}$, $\beta = 0$, $\gamma = 1$, but positive definite matrices are not possible. Thus the necessary conditions of Theorem 2, but not the sufficient condition, holds in this instance.

https://doi.org/10.1017/S0004972700004895 Published online by Cambridge University Press

For the same problem (18), the point $(-1, 2^{\frac{1}{2}})$ is a saddle point, with Lagrange multiplier 1 . Here

$$M_0 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, \quad M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2^{\frac{1}{2}} \\ -1 & 2^{\frac{1}{2}} \end{bmatrix}.$$

The matrices to consider are

$$\begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} + Q_0 - 2^{\frac{1}{2}}Q_1 \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} - Q_0 + 2^{\frac{1}{2}}Q_1 ,$$

and these cannot both be positive definite (or semidefinite), whatever the choice of the matrix $Q_0 - 2^{\frac{1}{2}}Q_1$. So the sufficient condition of Theorem 2 does *not* hold here.

References

[1] B.D. Craven, *Mathematical programming and control theory* (Chapman and Hall, London; John Wiley & Sons, New York; 1978).

[2] B.D. Craven, "Duality for generalized convex fractional programs", *Generalized concavity in optimization and economics* (Academic Press, New York, London, to appear).

[3] B.D. Craven and B. Mond, "Sufficient Fritz John optimality conditions for nondifferentiable convex programming", *J. Austral. Math. Soc. Ser. B* **19** (1975/76), 462-468.

[4] Anthony V. Fiacco, Garth P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques* (John Wiley and Sons, New York, London, Sydney, 1968).

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

[6] M.A. Hanson and B. Mond, "Further generalizations of convexity in mathematical programming" (Pure Mathematics Research Paper No. 80-6, Department of Mathematics, La Trobe University, Melbourne, 1980). See also: *J. Inform. Optim. Sci.* (to appear).

[7] B. Mond and M.A. Hanson, "On duality with generalized convexity" (Pure Mathematical Research Paper No. 80-4, Department of Mathematics, La Trobe University, Melbourne, 1980).

Department of Mathematics,
University of Melbourne,
Parkville,
Victoria 3052,
Australia.