ACTA MATHEMATICA VIETNAMICA
Volume 26, Number 3, 2001, pp. 249-256

ON NONCONVEX OPTIMIZATION

B. D. CRAVEN

Dedicated to Pham Huu Sach on the occasion of his sixtieth birthday

Abstract. Constrained minimization problems, in particular optimal control problems, are considered, where convexity requirements may not be satisfied. Various criteria for existence, or uniqueness, of a minimum are discussed. Criteria for necessary (Karush-Kuhn-Tucker or Pontryagin) conditions to be also sufficient for a minimum are discussed as well. Some of these depend on generalized-convexity properties such as invexity or V-invexity.

1. Introduction

If a constrained optimization problem in continuous variables reaches a minimum, and if certain regularity holds, then necessary Lagrangian conditions hold. Under additional assumptions of convexity, or generalized convexity, the necessary conditions are also sufficient for a minimum. Consider, however, a constrained optimization problem in infinite-dimensional spaces − typically an optimal control problem − for which existence of an optimum is hard to prove, in the absence of compactness properties, and perhaps also of coercivity properties. Moreover, an optimal control problem includes an equality constraint, in the form of a dynamic differential equation; if this equality constraint is nonlinear, then convexity cannot hold, and a generalized convexity property such as invexity is often very hard to verify.

A number of criteria will be given, for existence, or uniqueness, of an optimum, and for when a generalized convexity property − *invex* or *V-invex* − holds.

2. Indirect existence of a minimum

Assuming, subject to verification, that a local, or global, minimum is reached. Then necessary Lagrangian conditions may be deduced. If these conditions can be solved, and if they are also sufficient for a minimum, then the attainment of a minimum has been indirectly established. This situation often applies to optimal control problems, as in Craven (1978), for which the Pontryagin principle, together with the adjoint differential equation, are equivalent necessary conditions (assuming certain regularity) to the Karush-Kuhn-Tucker (KKT) necessary conditions for the corresponding mathematical programming problem. When is

Received April 3, 2001; in revised form November 14, 2001.

a solution guaranteed to these necessary conditions? This happens under generalized convex assumptions (see Section 6), and also under uniqueness conditions (see Section 3).

3. Solution uniqueness

A mathematical program may often have several *KKT points* at which the necessary optimality conditions hold. Under what conditions is a KKT point unique? If a minimum is known to be reached (see Section 4), and if also there is exactly one KKT point, then that point must be the global minimum. There is then no need to assume generalized convexity hypotheses. This situation happens with a substantial class of optimal control problems.

4. Proof of the solution existence by approximations

Consider the constrained minimization problem:

$$\text{MIN } F(z) \text{ subject to } G(z) \leq 0, \quad K(z) = 0, \tag{1}$$

in which F, G and K are continuous functions, and, in general, z lies in an infinite-dimensional space Z. Consider, in particular, the optimal control problem, in which z is replaced by a pair $(x(.), u(.))$ of state and control functions:

$$\text{MIN}_{x(.),u(.)} F(x,u) := \int_0^T f(x(t), u(t), t)dt + \Phi(x(T)) \tag{2}$$

subject to

$$x(0) = x_0, K(x,u)(t) := -\dot{x}(t) + m(x(t), u(t), t) = 0, \quad (0 \leq t \leq T) \tag{3}$$

$$u(t) \in \Delta(t) \quad (0 \leq t \leq T). \tag{4}$$

(Here (4) represents $G(z) \leq 0$). Consider a modification of (1), by adjoining a constraint $z \in Z_n$, where Z_n is a finite-dimensional subspace of Z, such that the sequence $\{Z_n\}$ expands to Z as $n \to \infty$. The corresponding modification for the control problem is to adjoin a constraint $u \in U_n$, where U_n is a finite-dimensional subspace, which may be obtained (e.g. Teo, Goh and Womg, 1991) by subdividing $[0, T]$ into finitely many subintervals, and restricting the control u to a step-function, constant on each of the subintervals,

Assume that the modified problem reaches a unique minimum at $z = \bar{z}_n$, for $n = 1, 2, \ldots$. Assume that a subsequence of $\{\bar{z}_n\}$ converges, say to a limit $\bar{z}$; by abuse of notation, denote this convergent subsequence again by $\{\bar{z}_n\}$. Denote by Γ the feasible set $\{z : G(z) \leq 0,\ K(z) = 0\}$. Now

$$(\forall z \in \Gamma \cap Z_n) F(z) \geq F(\bar{z}_n) \geq F(\bar{z}) - \epsilon_n \tag{5}$$

for some sequence $\{\varepsilon_n\} \to 0$ as $n \to \infty$. It follows that

$$(\forall z \in Z) F(z) \geq F(\bar{z}); \tag{6}$$

thus $\bar{z}$ is a global minimum for the problem (1).

For the control problem, an instance is given in Craven (1999) where $\bar{z}_n$ is independent of n, so convergence is trivial. More generally, consider a problem where *bang-bang control* is optimal for each modified problem; thus $u(.)$ is a step-function, which jumps between two extreme values at *switching times* $\tau_n^i \in [0,T]$ $(n = 1,2,\ldots;\ i = 1,2,\ldots,r)$.

Proposition 1. *For a control problem on a bounded time interval, where bang-bang control is optimal for each restriction of the control to a finite-dimensional subspace U_n of step-functions, assume also that the number of switching times is bounded as $n \to \infty$. Then the control problem reaches an optimum.*

Proof. It is sufficient to consider switching times $\tau_n^i \in [0,T]$ for $i = 1,2,\ldots,r$, with r not increasing with $n \to \infty$. By choosing appropriate subsequences, each of the bounded sequences $\{\tau_n^i\}_{n=1,2,\ldots}$ may be assumed to converge, say to a limit $\bar{\tau}^i$. Then the corresponding control functions u_n will converge, to a limit $\bar{u}$, in the $L^1[0,T]$ norm (which is appropriate to the Pontryagin theory). Hence the global minimum is reached. □

5. Solution uniqueness, for a class of optimal control models

Consider a class of optimal control models, defined by (2) and (3), but omitting (4), this assuming that the constraints on the control $u(.)$ are never active. Such control models are extensively used for models of economic growth (see e.g. Islam and Craven 2001). Assuming that a minimum is reached, and assuming the regularity hypotheses usual for the Pontryagin theory (see e.g. Craven 1995), necessary conditions for the minimum at $(\bar{x}(t), \bar{u}(t))$ are the dynamic equation (3), the adjoint differential equation:

$$-\dot{\lambda}(t) = f_x(\bar{x}(t), \bar{u}(t), t) + \lambda(t) m_x(\bar{x}(t), \bar{u}(t), t), \lambda(T) = \Phi'(x(T)), \tag{7}$$

together with Pontryagin's principle, which here requires that

$$f(\bar{x}(t), ., t) + \lambda(t) m(\bar{x}(t), ., t) \to \text{ MIN at } \bar{u}(t). \tag{8}$$

Here, f_x and m_x denote partial derivatives with respect to x. For (8), with no active constraints on $u(.)$, it is necessary that the gradient:

$$f_i(\bar{x}(t), \bar{u}(t), t) + \lambda(t) m_o(\bar{x}(t), \bar{u}(t), t) = 0. \tag{9}$$

Let $w(t) := (\bar{x}(t), \lambda(t))$. Combine the differential equations in (3) and (7) in the form:

$$\dot{w}(t) = Q(w(t), u(t), t), w(0) = w_0, \tag{10}$$

where the initial condition w_0 includes a parameter, to be adjusted to satisfy the terminal boundary condition in (7). The following Proposition extends a result of Islam and Craven (2001).

Proposition 2. *Assume that*

(a) *$Q(., u(t), t)$ satisfies a Lipschitz condition, uniformly in $u(t)$ in a neighbourhood of the optimum and $t \in [0,T]$, and*

(b) *Equation* (9) *can be solved uniquely for* $\bar{u}(t) = S(w(t), t)$, *where* $S(., t)$ *satisfies a Lipschitz condition uniformly in* t.

Then, if the optimal control problem, (1), (2), (3), *but omitting the control constraint* (4), *attains a minimum, then this minimum is unique.*

Proof. From (b),

$$\dot{w}(t) = Q(w(t), S(w(t), t), t), w(0) = w_0, \tag{11}$$

From (a), the function on the right of (11) satisfies a Lipschitz condition. From a well-known corollary of the contraction mapping theorem, (11) is solvable uniquely for $w(.)$, given w_0.

Thus the necessary Lagrangian conditions for a minimum of the control problem possess a unique solution for $(\bar{x}(t), \bar{u}(t), \lambda(t))$. Since a minimum is attained, this must be it. □

Remark. This result is not applicable to a control model which is linear in the control $u(t)$, since such a model would have no solution without active constraints on the control, and moreover would typically have a bang-bang regime, for which the Pontryagin principle would not allow $u(t)$ to be a Lipschitz function of the state $x(t)$ and costate $\lambda(t)$.

6. Invexity by function transformation

A vector function P is *invex* on a domain E, with respect to a *scale function* η and a convex order cone Q, if

$$(\forall z, p \in \text{ E}) \; P(z) - P(p) - P'(p)\eta(z, p) \in Q. \tag{12}$$

If L has r components relating to ≤ 0 inequalities, then s components relating to equalities, then (12) requires

$$P_i(z) - P_i(p) - P_i'(p)\eta(z, p) \geq 0 \quad (i = 1, 2, \ldots, r), \tag{13}$$

$$P_i(z) - P_i(p) - P_i'(p)\eta(z, p) = 0 \quad (i = r + 1, \ldots, r + s). \tag{14}$$

It is well known (see e.g. Craven 1995) that a (differentiable, invertible) transformation $z = \varphi(w)$ of the domain maps an invex function to an invex function, though with a different scale function. If, for some such φ, the composition $P_\circ\varphi$ happens to be convex, then this shows that P is invex.

If $L(z) := (F(z), G(z))$, and $\bar{z}$ is a KKT point for minimizing $F(z)$ subject to $G(z) \leq 0$ with Lagrange multiplier λ, with $F'(\bar{z}) \neq 0$ and $G'(\bar{z})$ having full rank, then (Craven, 2000, Theorem 3) $P(.)$ is invex at $\bar{z}$ exactly when the Lagrangian

$$L(.) := F(.) + \lambda G(.) = [1 \;\; \lambda] L(.)$$

is minimized at $\bar{z}$.

Proposition 3. *Consider the optimal control model* (1), (2), (3), *thus with the control constraint* (4) *assumed inactive, and the particular form of the dynamic* *equation* (3)*:*

$$\dot{x}(t) = b(t)\theta(x(t)) - u(t), \tag{15}$$

where $\theta(.)$ *is an increasing function, and the vector functions* $x(t)$ *and* $u(t)$ *areassumed to have the same dimension. If an invertible differentiable transformation* $X(t) = N(x(t))$, *with Jacobian matrix* $N'(x(t))$, *satisfies*

$$N'(x)\theta(x) = \mathbf{e} \tag{16}$$

(where $\mathbf{e}$ *is a column of ones), and if*

$$\Psi(X(t), U(t)) := N'(N^{-1}(X(t))f(N^{-1}(X(t)), U(t)/N'(N^{-1}(X(t))) \tag{17}$$

is positive definite over a domain E, *then the control problem is invex over* E.

Proof. The differential equation transforms by $X(t) = N(x(t))$ to :

$$\dot{X}(t) = N'(N^{-1}(X(t))[b(t)\theta(N^{-1}(X(t)) - \text{ u(t)}] \; = b(t)\mathbf{e} - U(t), \tag{18}$$

where

$$U(t) = N'(N^{-1}(X(t))u(t) \tag{19}$$

is a new control function, provided that N can be chosen so that $N'(x)$ satisfies (16).

Substituting into the objective function (2) gives the integrand as $\Psi(X(t), U(t))$. Since (17) is linear in $X(t)$ and $U(t)$, the transformed problem is convex if Ψ is convex. □

Remarks. In particular, consider $x(t)$ with two components p and q, and $\rho := \theta(x(t))$ with components $\varphi(t)$ and $\psi(t)$. Denote one row of $M := N'(x)$ by $v :=$ [a b]. Then $a\varphi + b\psi = 1$ when (for some s):

$$v = \alpha\rho + s\sigma, \text{ where } \sigma := [-\psi, \varphi]\perp\rho \text{ and } \alpha = 1/\|\rho\|. \tag{20}$$

In order that v be a gradient, it is required that the partial derivatives a $_q = b_p$. This happens if s satisfies the first-order linear partial differential equation:

$$(\varphi s)_p + (\psi s)_q = (\alpha\varphi)_q - (\alpha\psi)_p. \tag{21}$$

Hence (16) holds if N is constructed from two independent solutions of (19).

An instance where $x(t)$ and $u(t)$ have each one component was discussed in Islam and Craven (2001). There, $\theta(x(t))$ had a power-law form $x(t)^\beta$ with $\beta < 1$, so that $N(.)$ was also a power law; and the integrand was $u(t)^\sigma$ for $\sigma = 0.1$. Invexity was verified.

7. V-invexity for optimal control

For the differentiable optimization problem (1) in finite dimensions, denote

$$F_0(z) := F(z), (F_1(z), \dots, F_r(z)) := G(z), (F_{r+1}(z), \dots, F_{r+s}(z)) := K(z).$$

The problem (1) is called *V-invex* at a feasible point $\bar{z}$ (Jeyakumar and Mond 1992) if, for some (vector) *scale function* $\eta(.,.)$ and some strictly positive (scalar) weight functions $w_j(.)$:

$$(\forall z)(\forall j) F_j(z) - F_j(\bar{z}) \geq w_j(z) F_j'(\bar{z}) \eta(z, \bar{z}) \quad (j = 0, 1, \dots, r) \tag{22}$$

$$(\forall z)(\forall j) F_j(z) - F_j(\bar{z}) = w_j(z) F_j'(\bar{z}) \eta(z, \bar{z}) \quad (j = r+1, \dots, r+s). \tag{23}$$

Assume that η and w_j are continuously differentiable, and $\eta(\bar{z}, \bar{z}) = 0$. Setting $\lambda_0 = 1$ and $\lambda_j (j \geq 1)$ as the Lagrange multipliers at a KKT point $\bar{z}$, define a modified Lagrangian as:

$$\hat{L}(z) := \sum_j \lambda_j (w_j(z))^{-1} F_j(z). \tag{24}$$

As in Jeyakumar and Mond (1992),

$$\begin{aligned} (\forall z) F_0(z) - F_0(\bar{z}) &\geq \hat{L}(z) - \hat{L}(\bar{z}) \\ &\geq \sum_j \lambda_j (w_j(z))^{-1} w_j(z) F_j'(\bar{z}) \eta(z, \bar{z}) = 0, \end{aligned} \tag{25}$$

from KKT; so $\bar{z}$ is a minimum point.

Note that with *invexity* or *V-invexity*, the requirements (13), (14) or (22), (23) need only be assumed for all *feasible* points z, since points z that do not satisfy the constraints do not enter the proof that $\bar{z}$ is a minimum point. Moreover, inactive constraints may be omitted; for them, $\lambda_j = 0$. The invexity and V-invexity properties, restricted to feasible points z and only active constraints, will be called *restricted invexity* and *restricted V-invexity*.

Since η is differentiable, $\eta(z, \bar{z}) = C(z - \bar{z}) + \mathbf{o}(\|z - \bar{x}\|)$, for some linear mapping C. Substitution into (22) shows that

$$(\forall v := z - \bar{z})(w_j(\bar{z})\mathbf{1} - \ \mathrm{C})\mathrm{F}\ '_j(\bar{z}) v \geq \ 0,$$

where $\mathbf{1}$ is the identity mapping. Hence

$$w_j(\bar{z})\mathbf{1} - C \in F_j'(\bar{z})^{-1}(0). \tag{26}$$

For a fixed $\bar{z}, w_j(\bar{z}) = 1$ can be assumed, since any other value can be absorbed into C.

Proposition 4. *Restricted V-invexity at a point $\bar{z}$ is equivalent to restricted invexity at $\bar{z}$.*

Proof. Since $w_j(z) > 0$, the given constraint $F_j(z) \leq 0$ and the modified constraint $(w_j(z))^{-1} F_j(z) \leq 0$ defines the same points z. Denote by E the feasible set. Assume that $w_j(\bar{z}) = 1$, and denote $c_j(z) := 1/w_j(z)$. If $F_j(\bar{z}) = 0$, then

$$c_j(.)F_j(.) \text{ is } \textit{restricted invex} \text{ at } \bar{z}$$

$$\Leftrightarrow (\forall z \in E)(w_j(z))^{-1}F_j(z) - (w_j(\bar{z}))^{-1}F_j(\bar{z}) \geq (c_j(.)F_j(.))'(\bar{z})\eta(z,\bar{z})$$

$$\Leftrightarrow (\forall z \in E)c_j(z)F_j(z) \geq c_j(\bar{z})F_j'(\bar{z})\eta(z,\bar{z}) \text{ (using } F_j(\bar{z}) = 0)$$

$$\Leftrightarrow (\forall z \in E)F_j(z) - F_j(\bar{z}) \geq w_j(z)F_j'(\bar{z})\eta(z,\bar{z})$$

$$\Leftrightarrow F_j(.) \text{ is } \textit{restricted V-invex} \text{ at } \bar{z}, \text{ with weight function } w_j(.)$$

For an equality constraint $F_j(z) = 0$, *restricted V-invexity* at $\bar{z}$ requires:

$$(\forall z \in E)0 - 0 = F_j(z) - F_j(\bar{z}) = w_j(z)F_j'(\bar{z})\eta(z,\bar{z})$$

$$\Leftrightarrow (\forall z \in E)0 = F_j'(\bar{z})\eta(z,\bar{z}),$$

which is the same requirement as for *restricted invexity* at $\bar{z}$. □

For the optimal control problem (2), (3), (4), *restricted V-invexity* at $\bar{z} = (\bar{x}, \bar{u})$ for the equality constraint (2) reduces to the form:

$$\begin{aligned} 0 = {} & m_x(\bar{x}(t),\bar{u}(t),t)\eta^{(x)}(x(t)-\bar{x}(t),u(t)-\bar{u}(t)) \\ & + m_u(\bar{x}(t),\bar{u}(t),t)\eta^{(u)}(x(t)-\bar{x}(t),u(t)-\bar{u}(t)), \end{aligned} \tag{27}$$

where $\eta^{(x)}$ and $\eta^{(u)}$ are two components of the scale function η, and subscripts $_x$ and $_u$ mean partial derivatives. The requirement for the integrand of the objective is:

$$\begin{aligned} & f(x(t),u(t),t) - f(\bar{x}(t),\bar{u}(t),t) \geq \\ & w^{(f)}(x(t),u(t))[f_x(\bar{x}(t),\bar{u}(t),t)\eta^{(x)}(x(t)-\bar{x}(t),u(t)-\bar{u}(t)) \\ & + f_u(\bar{x}(t),\bar{u}(t),t)\eta^{(u)}(x(t)-\bar{x}(t),u(t)-\bar{u}(t))]. \end{aligned} \tag{28}$$

Since $w^{(f)}(.) > 0$, (28) means that the left side of (28) must have the same sign as the expression in $[\ldots]$ on the right side. For a constraint $\gamma(u) \leq 0$, the V-invexity requirement is:

$$\gamma(u(t)) - \gamma(\bar{u}(t)) \geq w^{(\gamma)}(u(t))\gamma_u(\bar{x}(t),\bar{u}(t),t)\eta^{(u)}(x(t)-\bar{x}(t),u(t)-\bar{u}(t)) \tag{29}$$

for $w^{(\gamma)}(.) > 0$. Similarly, this means that the left side of (29) must have the same sign as $\gamma_u(.)\eta^{(u)}(.)$.

Acknowledgement

A referee is thanked for some detail corrections.

References

[1] B. D. Craven, *Mathematical Programming and Control Theory*, Chapman & Hall, London, 1978.
[2] B. D. Craven, *Control and Optimization*, Chapman & Hall, London, 1995.
[3] B. D. Craven, *Optimal control for an obstruction problem*, J. Optimization Theory. Appl. **100** (1999), 435-439.
[4] B. D. Craven, *Global Invexity*, RGMIA Research Report Collection, (4), Article 8, 2000.
[5] S. M. N. Islam and B. D. Craven, *Computation of nonlinear continuous optimal growth models*: Experiments with optimal control algorithms and computer programs, Economic Modelling (2001, to appear).
[6] V. Jeyakumar and B. Mond, *On generalised convex mathematical programming*, J. Austral. Math. Soc. Series B, **34** (1992), 43-53.
[7] K. L. Teo, C. J. Goh, and K. H. Wong, *A Unified Computational Approach to Optimal Control*, Longman Scientific and Technical, London, 1991.

Department of Mathematics & Statistics
University of Melbourne, Australia