Yugoslav Journal of Operations Research
16 (2006), Number 2, 153-160

FIRST AND SECOND ORDER NECESSARY OPTIMALITY CONDITIONS FOR DISCRETE OPTIMAL CONTROL PROBLEMS

Boban MARINKOVIĆ

Faculty of Mining and Geology, University of Belgrade
Belgrade, Serbia
mboban@verat.net

Received: October 2004 / Accepted: June 2005

Abstract: Discrete optimal control problems with varying endpoints are considered. First and second order necessary optimality conditions are obtained without normality assumptions.

Keywords: Discrete optimal control, mathematical programming, abnormal extremal.

1. INTRODUCTION

Consider discrete optimal control problem with varying endpoints.

$$\text{minimize} \sum_{i=1}^{N-1} f_i(x_i, u_i)\ ; \tag{1}$$

$$x_{i+1} = \varphi_i(x_i, u_i)\ ,\ i = \overline{0, N-1}\ , \tag{2}$$

$$K(x_0, x_N) = 0\ , \tag{3}$$

where $f_i(x,u): R^n \times R^r \to R$ is twice continuously differentiable function, $\varphi_i(x,u)$: $R^n \times R^r \to R^n$ and $K(x_0, x_N): R^n \times R^n \to R^k$ are twice continuously differentiable mappings. Here $x_i \in R^n$ is state variable, $u_i \in R^r$ is a control parameter, N is given number of steps. Vector $\varepsilon = (x_0, x_1, ..., x_n)$ is called a trajectory, $\omega = (u_0, u_1, ..., u_{N-1})$ is called a control, x_0 is a starting point and x_N is an end point of corresponding trajectory.

Let x_0 be a starting point and let be a control. Then the pair (x_0,ω) defines the corresponding directory $\varepsilon=(x_0,x_1,...,x_n)$. If the condition (3) is satisfied then we say that the pair (x_0,ω) is feasible.

The discrete optimization problem is to minimize the function

$$J(x_0,\omega)=\sum_{i=0}^{N-1} f_i(x_i,u_i)$$

on the set of feasible pairs.

The aim of this paper is to obtain first and second order necessary optimality conditions for the problem (1)-(3) without normality assumptions.

2. FIRST ORDER NECESSARY OPTIMALITY CONDITIONS

Let $(\hat{x}_0,\hat{\omega})$ be a feasible pair and let $\hat{\varepsilon}=(x_0,x_1,...,x_N)$ be a corresponding trajectory. Suppose that the pair $(\hat{x}_0,\hat{\omega})$ is optimal.

Put

$$\frac{\partial\varphi}{\partial x}=C_k,\ \frac{\partial\varphi}{\partial u}=D_k$$

and

$$\frac{\partial^2\varphi^k}{\partial x^2}=C_k^2,\frac{\partial^2\varphi^k}{\partial u^2}=D_k^2,\frac{\partial^2\varphi^k}{\partial x\partial u}=M_k^2\ .$$

Let (h,v), $h=(h_0,h_1,...,h_N)\in R^{n(N+1)}$, $v=(v_0,v_1,...,v_{N-1})\in R^{rN}$, be the vector for which there exists a vector $(h^{'},v^{'})\in R^{n(N+1)}\times R^{rN}$ such the following conditions are satisfied:

$$h_k=\prod_{s=0}^{k-1}C_sh_0+\sum_{j=0}^{k-2}\prod_{j<s\le k-1}C_sD_jv_j+D_{k-1}v_{k-1},\ k=\overline{1,N}\ , \tag{4}$$

$$\frac{\partial K}{\partial(x_0,x_N)}(\hat{x}_0,\hat{x}_N)(h_0,h_N)=0\,, \tag{5}$$

$$-C_k^2[h_k]^2-D_k^2[v_k]^2-2M_k^2[h_k,v_k]=h_{k+1}^{'}-C_kh_k^{'}-D_kv_k^{'},\ k=0,...,N-1. \tag{6}$$

Here [-,-] stands for the arguments of a bilinear form (or, more generally, bilinear mapping).

Denote by K set of all vectors (*h*,*v*) for which the preceding conditions are satisfied.

Define the functions

$$H_A^i(x,u,p,q,\lambda_0,h,v)=\langle p,\varphi_i(x,u)\rangle+\langle q,\frac{\partial\varphi_i}{\partial x}(x,u)h+\frac{\partial\varphi_i}{\partial u}(x,u)v\rangle-$$

$$-\lambda_0 f_i(x,u),\quad i=0,N-1\,,$$

$$l_A(x_0,x_N,\lambda_1,\lambda_2,h_0,h_N)=\langle\lambda_1,K(\hat{x}_0,\hat{x}_N)\rangle+\langle\lambda_2,\frac{\partial K}{\partial(x_0,x_N)}(\hat{x}_0,\hat{x}_N)(h_0,h_N)\rangle,$$

where

$$\lambda_0\in R,\ \ \lambda_1,\lambda_2\in R^k,\ \ p,q\in R^n.$$

Theorem 1: *Let* $(\hat{x}_0,\hat{\omega})$ *be the optimal solution for the problem (1)-(3). Then there exists Lagrange multiplier* $\lambda_A=(\lambda_0,\lambda_1,\lambda_2,p,q)$, $\lambda_0\in R$, $\lambda_1,\lambda_2\in R^k$, $p\in R^{n(N+1)}$, $q\in R^{nN}$, $\lambda_0+|(q,\lambda_2)|\neq 0$ *such that for every* $(h,v)\in K$ *the following conditions are satisfied:*

(i)

$$p_0=\frac{\partial l_A}{\partial x_0}(\hat{x}_0,\hat{x}_N,\lambda_1,\lambda_2,h_0,h_N), \tag{7}$$

$$p_i=\frac{\partial H_A^i}{\partial x}(\hat{x}_i,\hat{u}_i,p_{i+1},q_{i+1},\lambda_0,h_i,v_i),\quad i=0,N-1, \tag{8}$$

$$p_N=-\frac{\partial l_A}{\partial x_N}(\hat{x}_0,\hat{x}_N,\lambda_1,\lambda_2,h_0,h_N), \tag{9}$$

$$\frac{\partial H_A^i}{\partial u}(\hat{x}_i,\hat{u}_i,p_{i+1},q_{i+1},\lambda_0,h_i,v_i),\quad i=0,N-1. \tag{10}$$

(ii) There exists vector $(h',v')\in R^{n(N+1)}\times R^{rN}$ such that

$$p_k=h'_k-C_{k-1}h'_{k-1}-D_{k-1}v'_{k-1},\ \ k=1,N, \tag{11}$$

$$\lambda_1=\frac{\partial K_0}{\partial x_0}(\hat{x}_0,\hat{x}_N)h'_0, \tag{12}$$

(iii)

$$-C_0^*q_1+\frac{\partial K}{\partial x_0}(\hat{x}_0,\hat{x}_N)^*\lambda_2=0, \tag{13}$$

$$q_k-C_k^*q_{k+1}=0,\ \ k=1,...N-1, \tag{14}$$

$$q_N-\frac{\partial K}{\partial x_N}(\hat{x}_0,\hat{x}_N)^*\lambda_2=0, \tag{15}$$

$$-D_{k-1}^*q_k=0,\ \ k=1,...,N. \tag{16}$$

Proof: We shall formulate the problem (1)-(3) as a mathematical programming problem, and we shall apply results from [2].

Define the functions

$$f(\varepsilon,\omega): R^{n(N+1)} \times R^{rN} \to R,\ F_i(\varepsilon,\omega): R^{n(N+1)} \times R^{rN} \to R^n,\ i=\overline{1,N},$$

and

$$F(\varepsilon,\omega): R^{n(N+1)} \times R^{rN} \to R^{nN},$$

by

$$f(\varepsilon,\omega) = \sum_{i=0}^{N-1} f_i(x_i,u_i),$$

$$F_{i+1}(\varepsilon,\omega) = x_{i+1} - \varphi_i(x_i,u_i),\quad i=0,N-1,$$

$$F(\varepsilon,\omega) = (F_1(\varepsilon,\omega),\dots,F_N(\varepsilon,\omega)).$$

Consider the following mathematical programming problem:

$$\text{Minimize } f(\varepsilon,\omega); \tag{17}$$

$$F(\varepsilon,\omega) = 0, \tag{18}$$

$$K(x_0,x_N) = 0. \tag{19}$$

The point $(\hat{\varepsilon},\hat{w})$ is the local minimum for preceding problem.
Consider the operator

$$\mathcal{F}(\varepsilon,\omega): R^{n(N+1)} \times R^{rN} \to R^{nN} \times R^k$$

given by

$$\mathcal{F}(\varepsilon,\omega) = (F(\varepsilon,\omega), K(x_0,x_N)).$$

Define the Lagrange – Avakov function

$$L_A(\varepsilon,\omega,\tilde{\lambda}_A,h,v): R^{n(N+1)} \times R^{rN} \times R \times R^{nN+k} \times \tilde{K} \to R$$

by

$$L_A(\varepsilon,\omega,\tilde{\lambda}_A,h,v) = \lambda_0 f(\varepsilon,\omega) + \langle \tilde{p}, \mathcal{F}(\varepsilon,\omega)\rangle + \left\langle \tilde{q}, \frac{\partial \mathcal{F}}{\partial(\varepsilon,\omega)}(\varepsilon,\omega)(h,v)\right\rangle,$$

where $\tilde{\lambda}_A = (\lambda_0,\tilde{p},\tilde{q})$, $\lambda_0 \in R$, $\tilde{p},\tilde{q} \in R^{nN+k}$ and $\tilde{K}$ is the set of all $(h,v) \in R^{n(N+1)} \times R^{rN}$ such that the following conditions are satisfied:

1)

$$\frac{\partial \mathcal{F}}{\partial(\varepsilon,\omega)}(\tilde{\varepsilon},\hat{\omega})(h,v) = 0,$$

2)

$$\frac{\partial^2 \mathcal{F}}{\partial(\varepsilon,\omega)^2}(\tilde{\varepsilon},\hat{\omega})[(h,v),(h,v)] \in im \frac{\partial \mathcal{F}}{\partial(\varepsilon,\omega)}(\tilde{\varepsilon},\hat{\omega}).$$

Note that the vector (h,v) is the parameter in the Lagrange – Avakov function.

Put $A = \frac{\partial F}{\partial(\varepsilon,\omega)}(\tilde{\varepsilon},\tilde{\omega})$.

From [2], theorem 10.1, we have that $\exists \tilde{\lambda}_A,\ \lambda_0 \geq 0,\ \lambda_0 + |\tilde{q}| \neq 0$ such that for every $(h,v) \in \tilde{K}$,

$$\frac{\partial L_A}{\partial(\varepsilon,\omega)}(\tilde{\varepsilon},\tilde{\omega},\tilde{\lambda}_A,h,v) = 0,\ \ \tilde{p} \in im\,A,\ \ \tilde{q} \in (im\,A)^{\perp}. \tag{20}$$

From the fact that

$$L_A(\varepsilon,\omega,\tilde{\lambda}_A,h,v) = \lambda_0 \sum_{i=0}^{N-1} f_i(x_i,u_i) + \sum_{i=0}^{N-1} \langle p_{i+1}, x_{i+1} - \varphi(x_i,u_i)\rangle + \langle \lambda_1, K(\tilde{x}_0,\tilde{x}_N)\rangle +$$

$$+\sum_{i=0}^{N-1} \langle q_{i+1}, h_{i+1} - C_i h_i - D_i v_i\rangle + \left\langle \lambda_2, \frac{\partial K}{\partial(x_0,x_N)}(\tilde{x}_0,\tilde{x}_N)(h_0,h_N)\right\rangle,$$

where $\tilde{p} = (p_1,...,p_N,\lambda_1)$ and $\tilde{q} = (q_1,...,q_N,\lambda_2)$, we have

$$\frac{\partial \tilde{L}_A}{\partial x_0} = \frac{\partial l_A}{\partial x_0}(\tilde{x}_0,\tilde{x}_N,\lambda_1,\lambda_2,h_0,h_N) - \frac{\partial H_A^0}{\partial x}(\tilde{x}_0,\tilde{u}_0,p,q,\lambda_0,h,v) = 0, \tag{21}$$

$$\frac{\partial \tilde{L}_A}{\partial x_i} = p_i - \frac{\partial H_A^i}{\partial x}(\tilde{x}_i,\tilde{u}_i,p_{i+1},q_{i+1},\lambda_0,h_i,v_i) = 0\ \ i = \overline{1,N-1}, \tag{22}$$

$$\frac{\partial \tilde{L}_A}{\partial x_N} = p_N + \frac{\partial l_A}{\partial x_N}(\tilde{x}_0,\tilde{x}_N,\lambda_1,\lambda_2,h_0,h_N) = 0, \tag{23}$$

$$\frac{\partial \hat{L}_A}{\partial u_i} = \frac{\partial H_A^i}{\partial u}(\hat{x}_i,\hat{u}_i,p_{i+1},q_{i+1},\lambda_0,h_i,v_i) = 0\ \ i = 0,N-1. \tag{24}$$

Put

$$p_0 = \frac{\partial H_A^0}{\partial x}(\hat{x}_0,\hat{u}_0,p,q,\lambda_0,h,v). \tag{25}$$

From (25) and (21) we have

$$p_0 = \frac{\partial l_A}{\partial x_0}(\hat{x}_0,\hat{x}_N,\lambda_1,\lambda_2,h_0,h_N).$$

Obviously that from (22), (23) and (24) we obtain that (8), (9) and (10) hold.

Put $p = (p_0,p_1,...,p_N)$, $q = (q_1,q_2,...,q_N)$ and $\lambda_A = (\lambda_0,\lambda_1,\lambda_2,p,q)$. We proved that for considered λ_A hold (7), (8), (9) and (10).

From the fact that $\tilde{p} \in im A$ we have that there exists vector $(h', v') \in R^{n(N+1)} \times R^{rN}$ such that (11) and (12) hold.

Since $\tilde{q} \in (im A)^{\perp}$ we have that $A^* \cdot \tilde{q} = 0$ or, in other words, we obtain that (13)-(16) hold.

From $A(h, v) = 0$ we obtain the system of the equations

$$h_k - C_{k-1}h_{k-1} - D_{k-1}v_{k-1} = 0, \ k = 1, N \tag{26}$$

with the boundary conditions

$$\frac{\partial K}{\partial(x_0, x_N)}(\hat{x}_0, \hat{x}_N)(h_0, h_N) = 0 .$$

Solving the equations (26) we obtain that

$$h_k = \prod_{s=0}^{k=1} C_s h_0 + \sum_{j=0}^{k-2} \prod_{j<s\le k-1} C_s D_j v_j + D_{k-1}v_{k-1}, \ k = 1, N ,$$

holds. It's easy to see that from

$$\frac{\partial^2 F}{\partial(\varepsilon, \omega)^2}(\hat{\varepsilon}, \hat{\omega})[(h, v), (h, v)] \in im A$$

we obtain the equation (6). We conclude that $K = \tilde{K}$.

3. SECOND ORDER NECESSARY OPTIMALITY CONDITIONS

Suppose that the function $f_i(x, u)$ and mappings $\varphi(x, u)$ and $K(x_0, x_N)$ are the three times continuously differentiable.

Put

$$\frac{\partial^2 \hat{l}_A}{\partial x_0^2} = \frac{\partial^2 l_A}{\partial x_0^2}(\hat{x}_0, \hat{x}_N, \lambda_1, \frac{\lambda_2}{3}, h_0, h_N), \ \frac{\partial^2 \hat{H}_A^i}{\partial x^2} = \frac{\partial^2 H_A^i}{\partial x^2}(\hat{x}_i, \hat{u}_i, p_{i+1}, \frac{q_{i+1}}{3}, \lambda_0, h_i, v_i) .$$

For a given Lagrange multiplier $\lambda_A = (\lambda_0, \lambda_1, \lambda_2, p, q)$ define the bilinear form

$$\Omega[(h, v), (h, v)] = \frac{\partial^2 \hat{l}_A}{\partial x_0^2}[h_0, h_0] + \frac{\partial^2 \hat{l}_A}{\partial x_N^2}[h_N, h_N] + 2\frac{\partial^2 \hat{l}_A}{\partial x_0 \partial x_N}[h_0, h_0] -$$

$$-\sum_{i=0}^{N-1} \frac{\partial^2 \hat{H}_A^i}{\partial x^2}[h_i, h_i] - \sum_{i=0}^{N-1} \frac{\partial^2 \hat{H}_A^i}{\partial u^2}[v_i, v_i] - 2\sum_{i=0}^{N-1} \frac{\partial^2 \hat{H}_A^i}{\partial x \partial u}[h_i, v_i] ,$$

Where $\frac{\partial^2 \hat{H}_A^i}{\partial x_N^2}, \frac{\partial^2 \hat{H}_A^i}{\partial u^2}, \frac{\partial^2 \hat{H}_A^i}{\partial x_0 \partial x_N}$ *and* $\frac{\partial^2 \hat{H}_A^i}{\partial x \partial u}$ are introduced analogously as above.

Theorem 2. *Let* $(\hat{x}_0, \hat{\omega})$ *be the optimal solution for the problem (1)-(3). Then there exists Lagrange multiplier* $\lambda_A = (\lambda_0, \lambda_1, \lambda_2, p, q)$ *such that the assertions of the theorem 1 hold, and for every* $(h, v) \in K$ *holds:*

$$\Omega_\lambda[(h,v),(h,v)] \geq 0 . \tag{27}$$

Proof: Analogously as in the proof of theorem 1 we consider the mathematical programming problem (17)-(19). It's easy to see that for Lagrange – Avakov function hold:

$$\frac{\partial^2 L_A}{\partial x_i \partial x_j} \equiv 0, \forall (i,j) : i \neq j, (i,j) \neq (0,N); \frac{\partial^2 L_A}{\partial x_i \partial u_j} \equiv 0, \forall_i \neq j$$

From [2], theorem 10.2, and from preceding facts we obtain that the assertion of theorem 2 hold.

4. CONCLUDING REMARKS

First we shall compare the number of variables and number of equations from theorem 1.

We have that the number of variables $\hat{\varepsilon},\ \hat{\omega},\ (h', v')\ and\ \lambda_A$ from the theorem 1 is equal to:

$$2(n(N+1) + rN) + 1 + 2(nN + k) .$$

From (2), (3), and (8)-(17) we obtain:

$$N_n + k + n(N+1) + rN + nN + k + n(N+1) + rN$$

equations.

Since $|\lambda_A| \neq 0$ then, without loss of generality, we may take that $|\lambda_A| = 1$. It follows that the number of variables is equal to the number of the equations and we have complete system of the equations associated with $\hat{\varepsilon},\ \hat{\omega},\ (h', v')\ and\ \lambda_A$.

The pair $(\hat{x}_0, \hat{\omega}_0)$ is said to be extreme for the problem (1)-(3) if feasible and if we have satisfied conditions from theorem 1.as in [3] we shall define the normal extreme for problem (1)-(3).

Definition 1: *The extremal is to be normal if*

$$im\, A = R^{nN+k}$$

and abnormal otherwise. Note that $A = \dfrac{\partial F}{\partial(\varepsilon, \omega)}(\hat{\varepsilon}, \hat{\omega})$.

First we suppose that the extreme $(\hat{x}_0, \hat{\omega}_0)$ is normal. Then we have that $(q, \lambda_2) = 0$ and we obtain classical first order optimality conditions which are known for a long time (see [4, 7, 8]). Also theorem 2 becomes a known second order optimality conditions (see [9]).

Suppose that the extreme $(\hat{x}_0, \hat{\omega}_0)$ is abnormal. We shall define two regular constraint mapping for the problem (1)-(3).

Denote by $K_{\overline{(h,v)}}$ the set

$$K_{\overline{(h,v)}} = \left\{ (h,v) : A(h,v) = 0, \frac{\partial^2 F}{\partial(\varepsilon,\omega)^2}(\hat{\varepsilon},\hat{\omega})[\overline{(h,v)},(h,v)] \in im\, A \right\}.$$

Definition 2: *The constraint mapping* $F(\varepsilon,\omega)$ *is sad to be 2- regular at the point* $(\hat{\varepsilon},\hat{\omega})$ *with respect to a direction* $(h,v) \in K$ *if*
$co\dim K_{\overline{(h,v)}} = nN + k$.

Suppose the extreme $(\hat{x}_0, \hat{\omega}_0)$ is abnormal and that for every $(h,v) \in K$ the mapping $F(\varepsilon,\omega)$ is not 2-regular at the point $(\hat{x},\hat{\omega})$ with respect to a direction (h,v). Then we have that assertions of the theorem 1 and theorem 2 are satisfied for every minimizing function f . It follows that in that case we have trivial theorem.

The most interesting case is when the mapping $F(\varepsilon,\omega)$ is 2-regular at the point $(\hat{\varepsilon},\hat{\omega})$ with respect to a direction $\overline{(h,v)} \in K$. Then we obtain nontrivial first and second order optimality conditions for abnormal extremes.

For details of the preceding facts we refer to [2].

REFERENCES

[1] Avakov, E.R., "Extremum conditions in smooth problems with equality type of constraints", *Comput. Math. and Math. Phys.*, 25 (5) (1985) 690-693 (in Russian).

[2] Arutyunov, A.V., *Optimality conditions: Abnormal and Degenerate Problems*, Moscow, 1997 (in Russian).

[3] Arutyunov, A.V., "Second order conditions in optimal control problems", *Doklady Math.*, 371 (1) (2000) 10-13 (in Russian).

[4] Boltyanskii, V.G., *Optimal Control of Discrete Systems*, Moscow, 1973 (in Russian).

[5] Vasilev, F.P., *Numerical Methods of Optimal Control Problems*, Moscow, 1988 (in Russian).

[6] Vasilev, F.P., *Optimization Methods*, Moscow, 2002 (in Russian).

[7] Ioffe, A.D., and Tihomirov, V.M., *Theory of Extremal Problems*, Moscow 1974 (in Russian).

[8] Propoii, A.I., *Elements of the Theory of Optimal Discrete Processes*, Moscow, 1973 (in Russian).

[9] Hilscher, R., and Zeidan, V., "Discrete optimal control: second order optimality conditions", *J. Differ. Equations Appl.*, Vol.8 (10) (2002) 875-896.