A. V. Arutyunov
PFU, Moscow, Russia

B. Marinković
Faculty of Mining and Geology, University of Belgrade, Serbia

2-regularity condition in discrete calculus of variations problems

Abstract

This paper is devoted to nonlinear discrete calculus of variations problems with equality and inequality type of constraints on trajectories and on endpoints. We derive new nontrivial first and second-order necessary optimality conditions.

Key words. discrete calculus of variations, mathematical programming, 2-regularity.

1. Introduction

We shall consider the following discrete calculus of variations problem:

$$\text{minimize}\, f_0(x_0, x_N) + \sum_{i=0}^{N-1} f(x_{i+1}, \Delta x_i); \qquad (1)$$

$$\varphi(x_i) = 0, \quad i = \overline{0, N}, \quad \psi(x_i) \leq 0 \quad i = \overline{0, N}, \qquad (2)$$

$$K_1(x_0, x_N) = 0, \quad K_2(x_0, x_N) \leq 0, \qquad (3)$$

where

$$f_0(x_0, x_N) : R^n \times R^n \to R, \quad f(x, u) : R^n \times R^n \to R,$$
$$\varphi(x) : R^n \to R^{m_1}, \quad \psi(x) : R^n \to R^{m_2},$$
$$K_1(x_0, x_N) : R^n \times R^n \to R^{k_1}, \quad K_2(x_0, x_N) : R^n \times R^n \to R^{k_2}$$

are twice continuously differentiable functions and where the forward difference operator is denoted by Δ, i.e., $\Delta x_i = x_{i+1} - x_i$. We assume that $m_1 \leq n$ and $k_1 \leq 2n$.

Vector $\xi = (x_0, x_1, \ldots, x_N)$ is called a trajectory. If the conditions (2) and (3) are satisfied then we say that the trajectory ξ is feasible. The discrete calculus of variations problem is to minimize the function

$$J(\xi) = K_0(x_0, x_N) + \sum_{i=0}^{N-1} f(x_{i+1}, \Delta x_i), \tag{4}$$

on the set of feasible trajectories. A feasible trajectory $\hat{\xi} = (\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_N)$ is called a local minimum for the problem (1)-(3) if for some $\epsilon > 0$, the trajectory $\hat{\xi}$ minimizes $J(\xi)$ over all feasible trajectories ξ satisfying

$$||x_i - \hat{x}_i||_n < \epsilon, \quad i = \overline{0, N},$$

where $|| \cdot ||_n$ is any norm in R^n.

Optimality conditions for the initial problem without inequality type of endpoint constraints and without constraints on trajectories have been a research focus of many authors; see [5] and references therein. In [3] discrete 2-regularity condition was used for sensitivity analysis in discrete optimal control problems. We also refer to [4] where discrete optimal control problems with equality type of constraints are considered, and there were obtained first and second-order necessary optimality conditions based on the discrete 2-regularity condition .

The obtained results in this paper are based on the general theory developed in [1].

2. First-order optimality conditions

Suppose that $\hat{\xi}$ is the optimal trajectory for the initial problem.

For convenience, we discard all constraints corresponding to indices j such that

$$K_{2,j}(\hat{x}_0, \hat{x}_N) < 0,$$

and we assume that

$$K_2(\hat{x}_0, \hat{x}_N) = 0.$$

Put

$$\frac{\partial \varphi}{\partial x}(\hat{x}_i) = V_i, \quad \frac{\partial \psi}{\partial x}(\hat{x}_i) = W_i, \quad \frac{\partial \psi_j}{\partial x}(\hat{x}_i) = W_{ji}$$

$$\frac{\partial K_1}{\partial x_0}(\hat{x}_0, \hat{x}_N) = K_{10}, \quad \frac{\partial K_1}{\partial x_N}(\hat{x}_0, \hat{x}_N) = K_{1N},$$

$$\frac{\partial K_2}{\partial x_0}(\hat{x}_0, \hat{x}_N) = K_{20}, \quad \frac{\partial K_2}{\partial x_N}(\hat{x}_0, \hat{x}_N) = K_{2N},$$

$$\frac{\partial^2 \varphi}{\partial x^2}(\hat{x}_i) = V_i^2, \quad \frac{\partial^2 \psi}{\partial x^2}(\hat{x}_i) = W_i^2, \quad \frac{\partial^2 \psi_j}{\partial x^2}(\hat{x}_i) = W_{ji}^2$$

$$\frac{\partial^2 K_1}{\partial x_0^2}(\hat{x}_0, \hat{x}_N) = K_{10}^2, \ \frac{\partial^2 K_1}{\partial x_N^2}(\hat{x}_0, \hat{x}_N) = K_{1N}^2, \ \frac{\partial^2 K_1}{\partial x_0 \partial x_N}(\hat{x}_0, \hat{x}_N) = K_{10N}^2$$

and

$$\frac{\partial^2 K_2}{\partial x_0^2}(\hat{x}_0, \hat{x}_N) = K_{20}^2, \ \frac{\partial^2 K_2}{\partial x_N^2}(\hat{x}_0, \hat{x}_N) = K_{2N}^2, \ \frac{\partial^2 K_2}{\partial x_0 \partial x_N}(\hat{x}_0, \hat{x}_N) = K_{20N}^2$$

For a fixed vector $\overline{h} = (\overline{h}_0, \overline{h}_1, \dots, \overline{h}_N)^T \in R^{n(N+1)}$ let us define the linear operator

$$N(\overline{h}) : R^{n(N+1)} \to R^{m_1(N+1)} \times R^{k_1}$$

by

$$N(\overline{h})h = (\tilde{y}_0, \dots, \tilde{y}_N, \tilde{z})^T$$

where

$$\tilde{y}_i = V_i^2[\overline{h}_i, h_i], \quad i = \overline{0, N},$$

$$\tilde{z} = K_{10}^2[\overline{h}_0, h_0] + K_{1N}^2[\overline{h}_N, h_N] + K_{10N}^2[\overline{h}_0, h_N] + K_{10N}^2[h_0, \overline{h}_N].$$

Let $Q = \{(\tilde{y}, \tilde{z}) \,|\, \tilde{v} \in R^{m_1(N+1)}, \ \tilde{w} \in R^{k_1}\}$ be the solution set of the following system of the linear equations:

$$V_0^T \tilde{y}_0 + K_{10}^T \tilde{z} = 0, \tag{5}$$

$$V_i^T \tilde{y}_i = 0, \quad i = \overline{1, N-1}, \tag{6}$$

$$V_N^T \tilde{y}_N + K_{1N}^T \tilde{z} = 0. \tag{7}$$

and denote by P the orthogonal projector onto the linear subspace $Q \subset R^{m_1(N+1)} \times R^{k_1}$.

Put

$$I = \{j : \psi_j(\hat{x}_i) = 0\}, \ i = \overline{0, N}.$$

Denote by A, B and C the the following block matrix

$$A = \begin{bmatrix} V_0 & 0 & \dots & 0 \\ 0 & V_1 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & V_N \\ K_{10} & 0 & \dots & K_{1N} \end{bmatrix}, \quad B = \begin{bmatrix} W_0 & 0 & \dots & 0 \\ 0 & W_1 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & W_N \\ K_{20} & 0 & \dots & K_{2N} \end{bmatrix},$$

and

$$C = \begin{bmatrix} \frac{\partial f_0}{\partial x_0}(\hat{x}_0, \hat{x}_N) - \frac{\partial f}{\partial u}(\hat{x}_1, \hat{x}_1 - \hat{x}_0) \\ \frac{\partial f}{\partial x}(\hat{x}_1, \hat{x}_1 - \hat{x}_0) + \frac{\partial f}{\partial u}(\hat{x}_1, \hat{x}_1 - \hat{x}_0) - \frac{\partial f}{\partial x}(\hat{x}_2, \hat{x}_2 - \hat{x}_1) \\ \vdots \\ \frac{\partial f_0}{\partial x_N}(\hat{x}_0, \hat{x}_N) + \frac{\partial f}{\partial x}(\hat{x}_N, \hat{x}_N + \hat{x}_{N-1}) + \frac{\partial f}{\partial u}(\hat{x}_N, \hat{x}_N - \hat{x}_{N-1}) \end{bmatrix}^T$$

where by 0 we denote the zero matrix of the corresponding size.

Let $H_1(\hat{\xi})$ be the cone of all vectors $h = (h_0, h_1, \dots, h_N)^T$ such that $Ah = 0$, $W_{ji}h_i \le 0$, $j \in I$, $i = \overline{0, N}$ and $K_{20}h_0 + K_{2N}h_N \le 0$ holds.

Let $H_2(\hat{\xi})$ be the set of all vectors $h \in H_1(\hat{\xi})$ for which there exists $u \in R^{n(N+1)}$ such that

$$Au = -N(h)h, \ W_{ji}u_i + W_{ji}^2[h_i, h_i] \le 0, \ \forall j \in I, \ i = \overline{0, N},$$

$$K_{20}h_0 + K_{2N}h_N + K_{20}^2[h_0, h_0] + K_{2N}^2[h_N, h_N] \le 0$$

holds.

Denote by $H_2^1(\hat{\xi})$ the set of all vectors $h \in H_2(\hat{\xi})$ such that for every $w \in H_1(\hat{\xi})$ holds

$$PN(h)w = Q$$

and there exists vectors $\eta^1 \in R^{n(N+1)}$ and $\eta^2 \in H_1(\hat{\xi})$ such that the following conditions

$$A\eta^1 = -N(h)\eta^2, \quad W_{ji}\eta_i^1 + W_{ji}^2[h_i, \eta_i^2] < 0, \quad \forall j \in I, \quad i = \overline{0, N},$$

$$K_{20}\eta_0^1 + K_{2N}\eta_N^1 + K_{20}^2[h_0, \eta_0^2] + K_{2N}^2[h_N, \eta_N^2] < 0$$

hold.

Define the functions

$$H(\xi,p^1,p^2,q^1,q^2,h_i) : R^{n(N+1)} \times R^{m_1} \times R^{m_1} \times R^{m_2} \times R^{m_2} \times R^n \to R, i = \overline{0,N},$$

and

$$l(x_0,x_N,\lambda^1,\lambda^2,\mu^1,\mu^2,h_0,h_N) : R^n \times R^n \times R^{k_1} \times R^{k_1} \times R^{k_1} \times R^{k_1} \times R^n \times R^n \to R$$

by

$$H(\xi,p^1,p^2,q^1,q^2,h_i) = \langle p^1,\varphi(x_i)\rangle + \left\langle p^2,\frac{\partial\varphi}{\partial x}(x_i)h_i\right\rangle + \langle q^1,\psi(x_i)\rangle$$
$$+\left\langle q^2,\frac{\partial\psi}{\partial x}(x_i)h_i\right\rangle + f_0(x_0,x_N) + \sum_{i=0}^{N-1} f(x_{i+1},\Delta x_i), \quad i = \overline{0,N},$$

and

$$l(x_0,x_N,\lambda^1,\lambda^2,\mu^1,\mu^2,h_0,h_N) = \langle\lambda^1,K_1(x_0,x_N)\rangle + \langle\mu^1,K_2(x_0,x_N)\rangle$$
$$+\left\langle\lambda^2,\frac{\partial K_1}{\partial(x_0,x_N)}(x_0,x_N)(h_0,h_N)\right\rangle + \left\langle\mu^2,\frac{\partial K_2}{\partial(x_0,x_N)}(x_0,x_N)(h_0,h_N)\right\rangle.$$

Theorem 2.1. *Let $\hat{\xi}$ be the optimal solution for the problem (1)-(3). Then there exists Lagrange multiplier $\lambda = (p^1,p^2,q^1,q^2,\lambda^1,\lambda^2,\mu^1,\mu^2)$, $p^1,p^2 \in R^{m_1(N+1)}$, $q^1,q^2 \in R^{m_2(N+1)}$, λ^1, $\lambda^2 \in R^{k_1}$, μ^1, $\mu^2 \in R^{k_2}$, such that for every $h \in H_2^1(\hat{\xi})$, such that $Ch \le 0$ holds, the following conditions are satisfied:*

$$\frac{\partial H}{\partial x_0}(\hat{\xi},p_0^1,p_0^2,q_0^1,q_0^2,h_0) = -\frac{\partial l}{\partial x_0}(x_0,x_N,\lambda^1,\lambda^2,\mu^1,\mu^2,h_0,h_N), \quad (8)$$

$$\frac{\partial H}{\partial x_i}(\xi,p_i^1,p_i^2,q_i^1,q_i^2,h_i) = 0, \quad i = \overline{1,N-1}, \quad (9)$$

$$\frac{\partial H}{\partial x_N}(\xi,p_N^1,p_N^2,q_N^1,q_N^2,h_N) = -\frac{\partial l}{\partial x_N}(x_0,x_N,\lambda^1,\lambda^2,\mu^1,\mu^2,h_0,h_N), \quad (10)$$

$$A^T(p^2,\lambda^2)^T + B^T(q^2,\mu^2)^T = 0, \quad (11)$$

$$q^1 \ge 0,\ \langle q_i^1,\psi(\hat{x}_i)\rangle = 0,\ i = \overline{0,N}, \quad q^2 \ge 0,\ \langle q_i^2,\psi(\hat{x}_i)\rangle = 0, \quad i = \overline{0,N}. \quad (12)$$

Proof. Define the functions $f(\xi)$, $F(\xi)$ and $g(\xi)$ by

$$f(\xi) = f_0(x_0, x_N) + \sum_{i=0}^{N-1} f(x_{i+1}, \Delta x_i),$$

$$F(\xi) = (\varphi(x_0), \ldots, \varphi(x_N), K_1(x_0, x_N))^T,$$

$$g(\xi) = (\psi(x_0), \ldots, \psi(x_N), K_2(x_0, x_N))^T.$$

We shall reformulate the initial problem into a mathematical programming problem (see [2, 3, 4]):

$$\text{minimize}\, f(\xi); \tag{13}$$

$$F(\xi) = 0, \quad g(\xi) \leq 0. \tag{14}$$

Obviously, $\hat{\xi}$ is a local minimum for the preceding problem.

Put

$$\tilde{I} = \{i : \, g_i(\hat{\xi}) = 0\}.$$

Let us consider the following sets:

1. $\tilde{H}_1(\hat{\xi}) = \{h \in R^{n(N+1)} \,|\, F'(\hat{\xi})h = 0,\ \langle g_i'(\hat{\xi}), h\rangle \leq 0,\ \forall i \in \tilde{I}\}$.

2. $\tilde{H}_2(\hat{\xi})$ is the set of all $h \in \tilde{H}_1(\hat{\xi})$ for which there exists $v \in R^{n(N+1)}$ such that

$$F'(\hat{\xi})v + F''(\hat{\xi})[h,h] = 0,\ \langle g_i'(\hat{\xi}), v\rangle + g_i''(\hat{\xi})[h,h] \leq 0, \forall i \in \tilde{I}$$

holds.

3. $\tilde{H}_2^1(\hat{\xi})$ is the set of all $h \in \tilde{H}_2(\hat{\xi})$ such that

$$\operatorname{im} F'(\hat{\xi}) + F''(\hat{\xi})[h, H_1(\hat{\xi})] = R^{m_1(N+1)+k_1} \tag{15}$$

and for which there exists $\eta^1 \in R^{n(N+1)}$ and $\eta^2 \in H_1(\hat{\xi})$ such that

$$F'(\hat{\xi})\eta^1 + F''(\hat{\xi})[h, \eta^2] = 0,\ \langle g_i'(\hat{\xi}), \eta^1\rangle + g_i''(\hat{\xi})[h, \eta^2] < 0, \forall i \in \tilde{I} \tag{16}$$

holds.

4. $\tilde{C}_2(\hat{\xi}) = \{h \in \tilde{H}_2(\hat{\xi}) \,|\, \langle f'(\hat{\xi}), h\rangle \leq 0\}$.

5. $\tilde{C}_2^1(\hat{\xi}) = \tilde{C}_2(\hat{\xi}) \cap \tilde{H}_2^1(\hat{\xi})$.

Let us introduce the generalized Lagrangian function by

$$L_2(\xi, \tilde{\lambda}, h) = f(\xi) + \langle \tilde{p}^1, F(\xi) \rangle + \left\langle \tilde{p}^2, F'(\xi)h \right\rangle + \langle \tilde{q}^1, g(\xi) \rangle + \left\langle \tilde{q}^2, g'(\xi)h \right\rangle,$$

where $\tilde{\lambda} = (\tilde{p}^1, \tilde{p}^2, \tilde{q}^1, \tilde{q}^2)$ and where $h \in \tilde{C}_2^1$ is a parameter.

In theorem 3 from [1] was proved that there exists Lagrange multiplier $\tilde{\lambda}$ such that for every $h \in \tilde{C}_2^1(\hat{\xi})$ the following conditions hold:

$$\frac{\partial L_2}{\partial \xi}(\hat{\xi}, \tilde{\lambda}, h) = 0, \tag{17}$$

$$(F'(\hat{\xi}))^T \tilde{p}^2 + (g'(\hat{\xi}))^T \tilde{q}^2 = 0, \tag{18}$$

$$\tilde{q}^1 \geq 0,\ \langle \tilde{q}^1, g(\hat{\xi}) \rangle = 0, \quad \tilde{q}^2 \geq 0,\ \langle \tilde{q}^2, g(\hat{\xi}) \rangle = 0. \tag{19}$$

First, it is easy to see that $F'(\hat{\xi})$ is given by the matrix A, $g'(\hat{\xi})$ is given by the matrix B and $f'(\hat{\xi})$ is given by the matrix C.

>From the definition of the operator $N(h)$ and from the preceding facts, we have that $\tilde{H}_2(\hat{\xi}) = H_2(\hat{\xi})$.

Also, from the fact that

$$\left(\operatorname{im}(F'(\hat{\xi})\right)^{\perp} = \ker F'(\hat{\xi})^T,$$

and from the equations (5)-(7), we have that $\tilde{H}_2^1(\hat{\xi}) = H_2^1(\hat{\xi})$.

Finally, it is easy to see that $\tilde{C}_2^1(\hat{\xi})$ is equal to the set of all $h \in H_2^1(\hat{\xi})$, such that

$$Ch \leq 0$$

holds.

Let us clarify the relation (17). From the fact that

$$\begin{aligned} L_2(\hat{\xi}, \tilde{\lambda}, h) &= f_0(\hat{x}_0, \hat{x}_N) + \sum_{i=0}^{N-1} f(\hat{x}_{i+1}, \Delta \hat{x}_i) + \sum_{i=0}^{N} \langle p_i^1, \varphi(\hat{x}_i) \rangle \\ &+ \sum_{i=0}^{N} \left\langle p_i^2, \frac{\partial \varphi}{\partial x}(\hat{x}_i) h_i \right\rangle + \sum_{i=0}^{N} \langle q_i^1, \psi(x_i) \rangle + \sum_{i=0}^{N} \left\langle q_i^2, \frac{\partial \psi}{\partial x}(\hat{x}_i) h_i \right\rangle \\ &+ \langle \lambda^1, K_1(\hat{x}_0, \hat{x}_N) \rangle + \left\langle \lambda^2, \frac{\partial K_1}{\partial(x_0, x_N)}(\hat{x}_0, \hat{x}_N)(h_0, h_N) \right\rangle \\ &+ \langle \mu^1, K_2(\hat{x}_0, \hat{x}_N) \rangle + \left\langle \mu^2, \frac{\partial K_2}{\partial(x_0, x_N)}(\hat{x}_0, \hat{x}_N)(h_0, h_N) \right\rangle \end{aligned}$$

where $\tilde{p}^1 = (p^1, \lambda^1)$, $\tilde{p}^2 = (p^2, \lambda^2)$, $\tilde{q}^1 = (q^1, \mu^1)$, and $\tilde{q}^2 = (q^2, \mu^2)$, we have

$$\frac{\partial L_2}{\partial x_0}(\hat{\xi}, \tilde{\lambda}, h) = \frac{\partial \hat{l}}{\partial x_0} + \frac{\partial \hat{H}}{\partial x_0}, \tag{20}$$

$$\frac{\partial L_2}{\partial x_i}(\hat{\xi}, \tilde{\lambda}, h) = \frac{\partial \hat{H}}{\partial x_i}, \quad i = \overline{1, N-1}, \tag{21}$$

$$\frac{\partial \hat{L}_2}{\partial x_N}(\hat{\xi}, \tilde{\lambda}, h) = \frac{\partial \hat{l}}{\partial x_N} + \frac{\partial \hat{H}}{\partial x_N}, \tag{22}$$

where

$$\frac{\partial \hat{l}}{\partial x_0} = \frac{\partial l}{\partial x_0}(\hat{x}_0, \hat{x}_N, \lambda^1, \lambda^2, \mu^1, \mu^2, h_0, h_N),$$

$$\frac{\partial \hat{l}}{\partial x_N} = \frac{\partial l}{\partial x_N}(\hat{x}_0, \hat{x}_N, \lambda^1, \lambda^2, \mu^1, \mu^2, h_0, h_N)$$

and

$$\frac{\partial \hat{H}}{\partial x_i} = \frac{\partial H}{\partial x_i}(\hat{\xi}, p_i^1, p_i^2, q_i^1, q_i^2, h_i).$$

Obviously that from (20), (21) and (22) we obtain that (8), (9) and (10) hold. From (18) and (19) we have that (11) and (12) hold. Theorem was proved.

3. Second-order optimality conditions

Suppose that the functions $f_0(x_0, x_N)$, $f(x, u)$, $\varphi(x)$, $\psi(x)$, $K_1(x_0, x_N)$ and $K_2(x_0, x_N)$ are three times continuously differentiable.

For a given Lagrange multiplier λ define the bilinear form

$$\Omega_\lambda[h, h] = \frac{\partial^2 \hat{l}}{\partial x_0^2}[h_0, h_0] + \frac{\partial^2 \hat{l}}{\partial x_N^2}[h_N, h_N] +$$
$$2\frac{\partial^2 \hat{l}}{\partial x_0 \partial x_N}[h_0, h_N] + 2\frac{\partial^2 f_0}{\partial x_0 \partial x_N}(\hat{x}_0, \hat{x}_N)[h_0, h_N] +$$
$$\sum_{i=0}^{N} \frac{\partial^2 \hat{H}}{\partial x_i^2}[h_i, h_i] + 2\sum_{i=0}^{N-1} \frac{\partial^2 \hat{H}}{\partial x_i \partial x_{i+1}}[h_i, h_{i+1}].$$

where $\frac{\partial^2 \hat{l}_2}{\partial x_0^2}$, $\frac{\partial^2 \hat{l}_2}{\partial x_N^2}$, $\frac{\partial^2 \hat{l}_2}{\partial x_0 \partial x_N}$, $\frac{\partial^2 \hat{H}}{\partial x_i^2}$ and $\frac{\partial^2 \hat{H}}{\partial x_i \partial x_{i+1}}$ are introduced analogously as in the previous section only with $p_i^2/3$, $q_i^2/3$, $\lambda^2/3$ and $\mu^2/3$ instead of p_i^2, q_i^2, λ^2 and μ^2.

Theorem 3.1. *Let $\hat{\xi}$ be the optimal solution for the problem (1)-(3). Then there exists Lagrange multiplier λ such that the assertions of theorem 2.1 hold, and for every $h \in H_1^2(\hat{\xi})$, such that $C(h) \leq 0$ holds, we have*

$$\Omega_\lambda[h, h] \geq 0. \tag{23}$$

Proof. Analogously as in the proof of theorem 2.1 we consider the mathematical programming problem (13)-(14). It is easy to see that Ω_λ is the second derivative with respect to ξ of the function $L_2(\hat{\xi}, \tilde{\lambda}, h)$. From [1], theorem 4 and from the preceding facts we obtain that the assertions of theorem 3.1 hold.

References

[1] *E.R. Avakov, A.V. Arutyunov, A.F. Izmailov,* Necessary conditions for an extremum in a mathematical programming problem, Proceedings of the Steklov institute of mathematics, vol. 256, 2-25 (2007).

[2] *A.V. Arutyunov, B. Marinković,* Necessary conditions for optimality in discrete optimal control problems, Vestnik MGU. Ser. 15. **1** 43-48 (2005) (in Russian).

[3] *B. Marinković,* Sensitivity analysis for discrete optimal control problems, Mathematical Methods of Operations Research, **63** (3) 513-524 (2006).

[4] *B. Marinković,* Optimality conditions in discrete optimal control problems, Optimization Methods & Software, (2007) (To appear).

[5] *R. Hilscher, V. Zeidan* , Nonnegativity and positivity of a quadratic functional in the discrete calculus of variations, J. Differ. Equations Appl. **11** (2) 857–875 (2005).