KYBERNETIKA — VOLUME 12 (1976), NUMBER 1

The Second-Order Methods in Discrete Optimal Control Problems

JIŘÍ V. OUTRATA

In this paper a simple and effective method of the evaluation of the Hessian matrices in discrete optimal control problems is suggested. Thus, various second-order optimization techniques can be applied to solve this problem. Especially the switching from the first order to the second method near to the extremum appeares to be very efficient. The presented paper contains an algorithm of this type for solving of the unconstrained minimization problems using the Newton-Raphson formulae.

1. INTRODUCTION

For solving of discrete optimal control problems preferably the first-order methods are used, for which there are several reasons. First, in linear-quadratic case, there is no need to use higher order methods, because already first-order ones will find the optimum with a sufficient accuracy and some of them even in a finite number of steps [1]; [2]; [3]. Second, the evaluation of the Hessian matrices made the computation very time-consuming. Moreover, several additional assumptions must be satisfied, which further restrict the applicability of higher oder methods.

On the other hand, the second-order methods are of better rate of convergence. (See [1].) It leads us to the mixed first-second order methods, where only additional differentiability assumptions are required. For the evaluation of the Hessian matrices considerable a simple method is proposed, which at no time requires us to handle matrices of order higher than the order of the solved discrete system. An algorithm, based on this method is presented and its convergence is proved by means of abstract algorithm theory.

The following notation is employed: E^n is the Euclidean n-space, $\|\cdot\|$ is the norm, $\langle \cdot, \cdot \rangle$ is the scalar product, x^j is the j-th coordinate of a vector x, A^{T} is the transpose of a matrix A.

2. UNCONSTRAINED DISCRETE OPTIMAL CONTROL PROBLEM (UDOCP)

Problem. Given a discrete dynamical system described by the equation

(1) $$x_{i+1} = f_i(x_i, u_i), \quad x_i \in E^n, \quad u_i \in E^m, \quad i = 0, 1, \ldots, k-1,$$

where x_i is the state of the system and u_i is the control applied to the system at time i, $f_i[E^n \times E^m \to E^n]$ and k is a given integer. Let

(2) $$x_0 = \hat{x}_0$$

be the given initial state of the system. Find a control sequence $\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_{k-1}$ and a corresponding trajectory $\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_k$, determined by (1), which minimize the cost functional

(3) $$J = \Phi(x_k),$$

where $\Phi[E^n \to E^1]$.

This problem can be easily transcripted into the form of an unconstrained optimization problem

(4) $$f^0(z) \to \min$$

where $f^0[E^\nu \to E^1]$. We set usually

(5) $$z = (u_0^1, u_0^2, \ldots, u_0^m, u_1^1, \ldots, u_1^m, \ldots, u_{k-1}^1, \ldots, u_{k-1}^m)$$

i.e.

(6) $$\nu = km.$$

Hence

(7) $$f^0(z) = \Phi(x_k(z)).$$

Now, following assumptions must be satisfied:

AS 1: $\Phi(x_k)$ is two times continuously differentiable function and $f_i(x_i, u_i)$, $i = 0, 1, \ldots, k-1$, are continuously differentiable functions.

AS 2: The function $f^0(z) = \Phi(x_k(z))$ is bounded from below.

AS 3: We can find a $z_0 \in E^\nu$ such that the set

(8) $$C(z_0) = \{z \mid f^0(z) - f^0(z_0) \leqq 0\}$$

is compact.

AS 4: For all $z \in C(z_0)$

$$\frac{\partial^2 \Phi(x_k(z))}{\partial x_k^2} \neq 0 .$$

For the increment of the cost functional we can write

$$\delta\Phi(x_k(z)) \cong \frac{\partial \Phi(x_k(z))}{\partial x_k} \delta x_k + \frac{1}{2} \left\langle \delta x_k, \frac{\partial^2 \Phi(x_k(z))}{\partial x_k^2} \delta x_k \right\rangle . \tag{9}$$

It is well-known from literature, that the term

$$\frac{\partial \Phi(x_k(z))}{\partial x_k} \delta x_k$$

can be expressed in the following way

$$\frac{\partial \Phi(x_k(z))}{\partial x_k} \delta x_k = - \sum_{i=0}^{k-1} \frac{\partial H_i(x_i, u_i, p_{i+1})}{\partial u_i} \delta u_i , \tag{10}$$

where $H_i(x_i, u_i, p_{i+1}) = \langle p_{i+1}, f_i(x_i, u_i) \rangle$, $i = 0, 1, \ldots, k - 1$, and vectors $p_i \in E^n$, $i = 1, 2, \ldots, k$, are obtained as the solution of the adjoint equation

$$p_i = \frac{\partial f_i(x_i, u_i)^{\mathrm{T}}}{\partial x_i} p_{i+1} , \quad i = 0, 1, \ldots, k - 1 , \tag{11}$$

with the terminal condition

$$p_k = - \frac{\partial \Phi(x_k(z))}{\partial x_k} . \tag{12}$$

We shall express the second-order term in relation (9) in the similar way in the next section.

3. THE EVALUATION OF THE HESSIAN MATRIX

As the Hessian matrix we denote the $[v \times v]$ matrix $H(z)$, satisfying the relation

$$\left\langle \delta x_k, \frac{\partial^2 \Phi(x_k(z))}{\partial x_k^2} \delta x_k \right\rangle = \langle \delta z, H(z) \delta z \rangle , \tag{13}$$

where

$$\delta z = (\delta u_0^1, \delta u_0^2, \ldots, \delta u_0^m, \delta u_1^1, \ldots, \delta u_1^m, \ldots, \delta u_{k-1}^1, \ldots, \delta u_{k-1}^m) . \tag{14}$$

Clearly

$$H(z) = \begin{bmatrix} \frac{\partial^2 \Phi(x_k(z))}{\partial u_0^1 \, \partial u_0^1} & \frac{\partial^2 \Phi(x_k(z))}{\partial u_0^1 \, \partial u_0^2} & \dots & \frac{\partial^2 \Phi(x_k(z))}{\partial u_0^1 \, \partial u_{k-1}^m} \\ \frac{\partial^2 \Phi(x_k(z))}{\partial u_0^2 \, \partial u_0^1} & \frac{\partial^2 \Phi(x_k(z))}{\partial u_0^2 \, \partial u_0^2} & \dots & \frac{\partial^2 \Phi(x_k(z))}{\partial u_0^2 \, \partial u_{k-1}^m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 \Phi(x_k(z))}{\partial u_{k-1}^m \, \partial u_0^1} & \dots & \dots & \frac{\partial^2 \Phi(x_k(z))}{\partial u_{k-1}^m \, \partial u_{k-1}^m} \end{bmatrix}. \tag{15}$$

Involving matrix adjoint difference equation

$$P_i = \frac{\partial f_i(x_i, u_i)^{\mathrm{T}}}{\partial x_i} P_{i+1} \frac{\partial f_i(x_i, u_i)}{\partial x_i}, \quad i = 0, 1, \dots, k-1, \tag{16}$$

where P_i are $[n \times n]$ matrices with the terminal condition

$$P_k = - \frac{\partial^2 \Phi(x_k(z))}{\partial x_k^2}, \tag{17}$$

we can write

$$\begin{aligned} \left\langle \delta x_k, \frac{\partial^2 \Phi(x_k(z))}{\partial x_k^2} \delta x_k \right\rangle &= \left\langle \frac{\partial x_k}{\partial z} \delta z, \frac{\partial^2 \Phi(x_k)}{\partial x_k^2} \frac{\partial x_k}{\partial z} \delta z \right\rangle = \\ &= - \left\langle \frac{\partial f_{k-1}(x_{k-1}, u_{k-1})}{\partial x_{k-1}} \delta x_{k-1} + \right. \\ &+ \frac{\partial f_{k-1}(x_{k-1}, u_{k-1})}{\partial u_{k-1}} \delta u_{k-1}, P_k \left(\frac{\partial f_{k-1}(x_{k-1}, u_{k-1})}{\partial x_{k-1}} \delta x_{k-1} + \right. \\ + \frac{\partial f_{k-1}(x_{k-1}, u_{k-1})}{\partial u_{k-1}} \delta u_{k-1} \left. \right) \Bigg\rangle &= - \sum_{i=0}^{k-1} \left(\left\langle \frac{\partial f_i(x_i, u_i)}{\partial u_i} \delta u_i, P_{i+1} \frac{\partial f_i(x_i, u_i)}{\partial u_i} \delta u_i \right\rangle + \right. \\ &+ 2 \left\langle \frac{\partial f_i(x_i, u_i)}{\partial x_i} \delta x_i, P_{i+1} \frac{\partial f_i(x_i, u_i)}{\partial u_i} \delta u_i \right\rangle \Bigg). \end{aligned} \tag{18}$$

The last expression implies that

$$\frac{\partial^2 \Phi(x_k(z))}{\partial u_i \, \partial u_i} = - \frac{\partial f_i(x_i, u_i)^{\mathrm{T}}}{\partial u_i} P_{i+1} \frac{\partial f_i(x_i, u_i)}{\partial u_i} \quad \text{for} \quad i = 0, 1, \dots, k-1. \tag{19}$$

In order to express matrix terms $\partial^2 \Phi(x_k(z)) / \partial u_i \, \partial u_j$, $i \neq j$, in the similar way, we involve the matrix adjoint difference equation given in (21). We set first

$$P_i = S_i^i \quad \text{for} \quad i = 1, 2, \dots, k. \tag{20}$$

 Now for every matrix S_i^i, $i = 2, 3, \ldots, k$, as the terminal condition we solve the equation

$$S_{j-1}^i = \frac{\partial f_{j-1}(x_{j-1}, u_{j-1})^{\mathrm{T}}}{\partial x_{j-1}} S_j^i, \quad j = 2, 3, \ldots, i. \tag{21}$$

Clearly for $j < i$

$$\frac{\partial^2 \Phi(x_k(z))}{\partial u_j \, \partial u_i} = -\frac{\partial f_j(x_j, u_j)^{\mathrm{T}}}{\partial u_j} S_{j+1}^{i+1} \frac{\partial f_i(x_i, u_i)}{\partial u_i}. \tag{22}$$

Thus

$$H(z) = -\begin{bmatrix} \dfrac{\partial f_0^{\mathrm{T}}}{\partial u_0} S_1^1 \dfrac{\partial f_0}{\partial u_0} & \dfrac{\partial f_0^{\mathrm{T}}}{\partial u_0} S_1^2 \dfrac{\partial f_1}{\partial u_1} & \ldots & \dfrac{\partial f_0^{\mathrm{T}}}{\partial u_0} S_1^k \dfrac{\partial f_{k-1}}{\partial u_{k-1}} \\ \left[\dfrac{\partial f_0^{\mathrm{T}}}{\partial u_0} S_1^2 \dfrac{\partial f_1}{\partial u_1}\right]^{\mathrm{T}} & \dfrac{\partial f_1^{\mathrm{T}}}{\partial u_1} S_2^2 \dfrac{\partial f_1}{\partial u_1} & \ldots & \dfrac{\partial f_1^{\mathrm{T}}}{\partial u_1} S_2^k \dfrac{\partial f_{k-1}}{\partial u_{k-1}} \\ \ldots & \ldots & \ldots & \ldots \\ \left[\dfrac{\partial f_0^{\mathrm{T}}}{\partial u_0} S_1^k \dfrac{\partial f_{k-1}}{\partial u_{k-1}}\right]^{\mathrm{T}} & \left[\dfrac{\partial f_1^{\mathrm{T}}}{\partial u_1} S_2^k \dfrac{\partial f_{k-1}}{\partial u_{k-1}}\right]^{\mathrm{T}} & \ldots & \dfrac{\partial f_{k-1}^{\mathrm{T}}}{\partial u_{k-1}} S_k^k \dfrac{\partial f_{k-1}}{\partial u_{k-1}} \end{bmatrix}. \tag{23}$$

We shall now demonstrate this method in the linear-quadratic case. Considering the stationary linear system described by the equation

$$x_{i+1} = Ax_i + Bu_i, \quad i = 0, 1, \ldots, k-1, \tag{24}$$

where A resp. B is $[n \times n]$ resp. $[n \times m]$ constant matrix and quadratic cost functional

$$\Phi(x_k) = \langle x_k, Qx_k \rangle, \tag{25}$$

we conclude that

$$\begin{aligned} P_k &= S_k^k = -2Q, \\ P_i &= S_i^i = -2\underbrace{(A^{\mathrm{T}}A^{\mathrm{T}} \ldots A^{\mathrm{T}})}_{(k-i)\times} Q \underbrace{(AA \ldots A)}_{(k-i)\times}, \\ S_j^i &= -2\underbrace{(A^{\mathrm{T}}A^{\mathrm{T}} \ldots A^{\mathrm{T}})}_{(i-j)\times} S_i^i, \\ \frac{\partial f_i(x_i, u_i)}{\partial u_i} &= B. \end{aligned} \tag{26}$$

According to formulae (23) we see, that if for example $k = 2$, $m = 1$, $n = 2$, then

$$H(z) = 2\begin{bmatrix} B^{\mathrm{T}}A^{\mathrm{T}}QAB & B^{\mathrm{T}}A^{\mathrm{T}}QB \\ B^{\mathrm{T}}QAB & B^{\mathrm{T}}QB \end{bmatrix}. \tag{27}$$

First we transcribe the UDOCP into the form (4) by means of (5), (7). Then we apply the following algorithm:

Algorithm A.

0. Select a $z_0 \in E^\nu$ such that the set $C(z_0)$ is compact; select an $\alpha \in]0, 1[$ and a $\beta \in [0, 1]$.
1. Set $i = 0$.
2. Compute $\nabla f^0(z_i)$.
3. If $\nabla f^0(z_i) = 0$, stop; else compute $H(z_i)$.
4. If $H^{-1}(z_i)$ exists, compute $h(z_i)$ by solving

$$H(z_i)\, h(z_i) = -\nabla f^0(z_i) \tag{28}$$

and go to step 5; else, set

$$h(z_i) = -\nabla f^0(z_i) \tag{29}$$

and go to step 7.

5. Compute $\langle \nabla f^0(z_i), h(z_i) \rangle$.
6. If $\langle \nabla f^0(z_i), h(z_i) \rangle \geqq 0$ then set $h(z_i) = -\nabla f^0(z_i)$ and go to step 7.
7. Set $\lambda = 1$.
8. Compute

$$\Delta = f^0(z_i + \lambda\, h(z_i)) - f^0(z_i) - \lambda\alpha \langle \nabla f^0(z_i), h(z_i) \rangle\,. \tag{30}$$

9. If $\Delta \leqq 0$, set $\lambda_i = \lambda$ and go to step 10; else, set $\lambda = \lambda\beta$ and go to step 8.
10. Set $z_{i+1} = z_i + \lambda_i\, h(z_i)$, set $i = i + 1$ and go to step 2.

The presented algorithm is of the form of the following algorithm model, constructing points in the closed subset T of a Banach space B which have property P. (Further we shall call them "desirable"). Let $a[T \to T]$ be a search function and $c[T \to E^1]$ be a stop rule.

Model M.

0. Compute a $z_0 \in T$.
1. Set $i = 0$.
2. Compute $a(z_i)$.
3. Set $z_{i+1} = a(z_i)$.
4. If $c(z_{i+1}) \geqq c(z_i)$, stop; else, set $i = i + 1$ and go to step 2.

 The convergence properties of this model are specified in the following theorem, proved in [1].

Theorem 1. Suppose that

(i) $c(z)$ is continuous in all nondesirable points $z \in T$.

(ii) for every $z \in T$ which is not desirable, there exists an $\varepsilon(z) > 0$ and $a\ \delta(z) < 0$ such that

$$c(a(z')) - c(z') \leqq \delta(z) \quad \text{for all} \quad z' \in B(z, \varepsilon(z)),$$

where $B(z, \varepsilon(z)) = \{z' \in T \mid \|z' - z\|_B \leqq \varepsilon(z)\}$.

Then, either the sequence $\{z_i\}$ constructed by algorithm model M is finite and its next to last element is desirable, or else it is infinite and every accumulation point of $\{z_i\}$ is desirable.

Referring to the model M, we shall say that the point z is desirable if $\nabla f^0(z) = 0$, we set $c(z) = f^0(z)$ and we define $a(z)$ as follows:

$$a(z) = z + \lambda\, h(z)$$

where $\lambda > 0$ is determined in step 9 and $h(z)$ is determined either in step 4 or in step 6. Algorithm A then clearly satisfies the assumption (i) of the previous theorem. We shall prove that the assumption (ii) is satisfied as well.

Suppose, that an $\varepsilon' > 0$ can be found such that for all $z' \in B(z, \varepsilon') = \{z' \mid \|z' - z\|_{E^v} \leqq \varepsilon'\}$ the vector $h(z')$ is computed in the same way, i.e. either according to (28) or (29).

The assumption AS 1 implies that the product $\langle \nabla f^0(z'), h(z') \rangle$ is a continuous function of z' for all $z' \in B(z, \varepsilon')$. As if z is nondesirable, then

$$\langle \nabla f^0(z), h(z) \rangle = 2\gamma(z) < 0,$$

there must exist an $\varepsilon(z) \leqq \varepsilon'$, $\varepsilon(z) > 0$ such that

$$\langle \nabla f^0(z'), h(z') \rangle \leqq \tfrac{1}{2} \langle \nabla f^0(z), h(z) \rangle = \gamma(z) < 0$$

for all $z' \in B(z, \varepsilon(z))$. For λ' satisfying the relation (30) it is true that

$$f^0(z' + \lambda' h(z')) - f^0(z') \leqq \lambda' \alpha \langle \nabla f^0(z'), h(z') \rangle = \lambda' \alpha\, \gamma(z) < 0.$$

The $\varepsilon' > 0$ cannot be found if and only if the matrix $H(z)$ is not invertible. Then the vector $h(z)$ is computed according to (29) and for every $\varepsilon(z)$ the set $B(z, \varepsilon(z))$ can contain points z' such that $h(z')$ are computed according to (28). However, for every $\varepsilon(z)$ an integer $q > 0$ can be found such that

$$\langle \nabla f^0(z'), h(z') \rangle \leqq \frac{1}{q} \langle \nabla f^0(z), h(z) \rangle < 0$$

for all $z' \in B(z, \varepsilon(z))$ what completes the proof.

Therefore, following theorem is true.

Theorem 2. Suppose that $\{z_i\}$ is a sequence constructed by algorithm A; then either the sequence $\{z_i\}$ is finite, terminating at z_k, where $\nabla f^0(z_k) = 0$, or else it is infinite and every accumulation point z' of $\{z_i\}$ satisfies $\nabla f^0(z') = 0$.

Corollary. Consider the set $Z = \{z \mid \nabla f^0(z) = 0\}$ and suppose that for every $z' \neq z''$ in Z, $f^0(z') \neq f^0(z'')$. Then any infinite sequence $\{z_i\}_{i=0}^{\infty}$ constructed by algorithm A must converge to a point $\hat{z} \in Z$.

The proof is quite evident.

5. CONCLUSIONS

If the vector $h(z_i)$ in the iteration scheme of algorithm A can be computed according to (28) and the step size would not be restricted in step 9, we obtain the usual Newton-Raphson formulae

$$z_{i+1} = z_i + H^{-1}(z_i)\,\nabla f^0(z_i)\,.$$

However, the step size adjustment of step 9 does not influence the rate of convergence, which was proved to be quadratic [1].

If the solved control problem has a constrained set of admissible controls or state space variables, another numerical method must be used, as for instance the second-order feasible direction method. In this case, the evaluation of the Hessian matrices, proposed in the section 2, can be applied as well.

We hope that the second-order methods, applied to the optimal control problems will, at least in some cases, spare the computation time due to better rate of convergence. Moreover, if by means of some first-order method some stationary point of a solved UDOCP is obtained, we can use the corresponding Hessian matrix for testing, whether this point is really a local minimum.

(Received September 10, 1975.)

REFERENCES

[1] E. Polak: Computational methods in Optimization. Academic Press, N. Y. 1971.

[2] M. D. Canon, C. D. Cullum, E. Polak: Theory of Optimal Control and Mathematical Programming. Mc Graw Hill, N. Y. 1971.

[3] P. Wolfe: The Simplex Method for Quadratic Programming. Econometrica *27*, (1959), 3, 382—398.

Ing. Jiří V. Outrata; Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 180 76 Praha 8. Czechoslovakia.