GENERALIZED GRADIENT METHOD FOR DYNAMIC LINEAR PROGRAMMING

A. Propoi

December 1976 WP-76-32

2361
Laxenburg
Austria

International Institute for Applied Systems Analysis

Preface

There are two fields of application for nondifferentiable optimization (NDO): the first are some practical problems, arising in applied projects; the second are the methods which themselves have a large field of application (e.g. structurized LP problems).

This working paper describes shortly the general scheme of the NDO method's application to dynamic LP problems. The paper is supposed and is believed to be a ground for joint work on developing the method for DLP, revealing all the possibilities of NDO [(1-4)] for structurized LP problems.

GENERALIZED GRADIENT METHOD FOR DYNAMIC LINEAR PROGRAMMING

A. Propoi

Abstract

A general scheme of application of nondifferentiable optimization methods (NDO) to dynamic linear programming (DLP) problems is considered.

1. *Statement of the Problem*

Let us consider the DLP problem in the following form [5]:

Problem 1P. To find a control

$$u^* = \{u^*(0),\ldots,u^*(T-1)\}$$

and a trajectory

$$x^* = \{x^*(0),\ldots,x^*(T)\} \quad ,$$

satisfying the state equations:

$$x(t+1) = A(t)x(t) + B(t)u(t) + s(t) \tag{1}$$

$$x(0) = x^0 \qquad (t=0,\ldots,T-1) \tag{2}$$

and the constraints

$$G(t)x(t) + D(t)u(t) \leq f(t) \tag{3}$$

$$R(t)u(t) \leq q(t) \tag{4}$$

$$u(t) \geq 0 \tag{5}$$

which maximizes the performance index

$$J_{1P} = (a(T),x(T)) + \sum_{t=0}^{T-1} [(a(t),x(t))+(b(t),u(t))] \quad . \quad (6)$$

Here $u(t) \in E^r$; $x(t) \in E^n$. Matrices $A(t)(n \times n)$, $B(t)(n \times r)$, $G(t)(m \times n)$, $D(t)(m \times r)$, $R(t)(\ell \times r)$ and vectors x^0, $s(t)(n \times 1)$, $f(t)(m \times 1)$, $q(t)(\ell \times 1)$ are supposed to be fixed.

Choosing a control u for some initial state x(0), we obtained from (1),(2) the corresponding trajectory:

$$x^u = x(x(0),u) \quad . \qquad (7)$$

Problem 1P is associated with the dual [6]:

<u>Problem 1D</u>. Find dual controls

$$\lambda^* = \{\lambda^*(T-1),\ldots, *(0))\} \quad , \qquad \sigma^* = \{\sigma^*(T-1),\ldots,\sigma^*(0)\}$$

and a dual trajectory

$$p^* = \{p^*(T),\ldots,p^*(0)\}$$

subject to costate equations:

$$p(t) = A^T(t)p(t+1) - G^T(t)\lambda(t) + a(t) \qquad (8)$$

$$p(T) = a(T) \qquad (9)$$

and constraints

$$B^T(t)p(t+1) - D^T(t)\lambda(t) - R^T(t)\sigma(t) \leq -b(t) \qquad (10)$$

$$\lambda(t) \geq 0 \quad ; \qquad \sigma(t) \geq 0$$

which minimizes

$$J_{1D}(\lambda,\sigma) = (p(0),x^0) + \sum_{t=0}^{T-1} [(p(t+1),s(t))+(\lambda(t),f(t))+(\sigma(t),q(t))] \quad . \qquad (11)$$

Here $p(t) \in E^n$, $\lambda(t) \in E^m$, $\sigma(t) \in E^\ell$, are the Lagrange multipliers for constraints (1)-(4) respectively.

Again, choosing a dual control λ for some boundary condition (10), we obtained from (8),(9) the corresponding dual trajectory

$$p^{\lambda} = p(p(T),\lambda) \quad . \tag{12}$$

2. *Method*

Let us introduce the Lagrange function for Problem 1P as follows:

$$L(u,\lambda) = J_{1P}(u) + \sum_{t=0}^{T-1} (\lambda(t),f(t)-G(t)x^{u}(t)-D(t)u(t)) \quad , \tag{13}$$

where in (13) $x^{u}(t)\,(t=0,\ldots,T-1)$ are supposed to satisfy state equations (1) for some u and x(0), (see (7)), so the Lagrange multipliers p(t) for these constraints are not necessary.

Propositions.

(i) Problem 1P has a solution.

(ii) The sets

$$U(t) = \{u(t)\,|\,R(t)u(t)\leq q(t),\ u(t)\geq 0\} \tag{14}$$

are bounded for all $t = 0,1,\ldots,T-1$.

Let

$$U = U(0)\times\cdots\times U(T-1)$$

$$\psi(\lambda) = \max_{u\in U} L(u,\lambda) \tag{15}$$

$$U^{\lambda} = \{u\,|\,L(u,\lambda) = \psi(\lambda), u\in U\} \tag{16}$$

Lemma 1. With the propostions stated above, the following statements are true:

(i) $\psi(\lambda)$ is a concave, piece-wise linear, continuous function defined for all $\lambda \geq 0$;

(ii) the function $\psi(\lambda)$ is bounded on the set $\{\lambda\geq 0\}$ and achieves its minima, i.e. λ^* exists, such that

$$\psi(\lambda^*) = \min_{\lambda \geq 0} \psi(\lambda) = \min_{\lambda \geq 0} \max_{u \varepsilon U} L(u,\lambda) \tag{17}$$

Proof follows from the general results stated in [1].

Lemma 2. Any solution λ^* of (17) corresponds to some solution $\{\lambda^*, \sigma^*, p^*\}$ of Problem 1D.

Proof: The problem of (15) for some fixed $\lambda = \lambda^0$ is the DLP problem with state equations (1),(2), constraints (4),(5) and performance index (13). As the sets (14) are bounded for all t, it has a finite solution $u^{\lambda^0} \varepsilon U^{\lambda^0}$. Therefore the dual to problem (15), which coincides with the Problem 1D, if $\lambda = \lambda^0$, $p = p^0 = (\lambda^0)$, has also a solution $\{\sigma^0 = \sigma(\lambda^0),\ p^0 = p(\lambda^0)\}$ and

$$J_{1D}(\sigma^0, \lambda^0) \leq J_{1D}(\sigma, \lambda^0)$$

for all feasible σ.

In particular, if $\lambda^0 = \lambda^*$ - a solution of (17), then

$$J_{1D}(\sigma^*, \lambda^*) \leq J_{1D}(\sigma, \lambda^*), \quad \sigma^* = \sigma(\lambda^*) \quad . \tag{18}$$

Choosing $\sigma = \sigma(\lambda^*, p^*)$ in (18) from a solution of the problem (10)-(11) with $\lambda = \lambda^*$, $p = p^* = p(\lambda^*)$, we obtain, obviously,

$$J_{1D}(\sigma^*, \lambda^*) = J_{1D}(\sigma(\lambda^*, p^*), \lambda^*) \quad .$$

Thus, the solution of dual Problem 1D is equivalent to the solution of the problem (17). For finding a solution of (17), we shall use the generalized gradient method [1]:

$$\lambda^{\nu+1} = P^+\{\lambda^\nu - \alpha_\nu \frac{\partial\psi(\lambda^\nu)}{\|\partial\psi(\lambda^\nu)\|}\} \quad (\nu = 0, 1, \ldots,) \tag{19}$$

Here P^+ is a projection operator on positive orthant, that is

$$P^+(x) = \begin{matrix} x_i, & \text{if } x_i \geq 0 \\ 0, & \text{if } x_i < 0 \end{matrix} \quad .$$

$||x||$ is the eucledean norm of vector **x**, $\partial\psi(\lambda)$ is generalized gradient of function $\psi(\lambda)$.

Lemma 3. The function $\psi(\lambda)$ is differentiable at any point $\lambda_0 \varepsilon E^{mT}$ in any direction $\gamma \varepsilon E^{mT}$ and

$$\partial_\gamma \psi(\lambda_0) = \max_{u \varepsilon U^{\lambda_0}} (\partial\psi(\lambda_0), \gamma) \tag{20a}$$

where

$$\partial\psi(\lambda) = \{f(t)-G(t)x(t)-D(t)u(t) \ \ (t=0,\ldots,T-1)\} \quad u = \{u(t)\} \ \varepsilon \ U^\lambda \ . \tag{20b}$$

From definitions of $L(u,\lambda)$ and u^λ in (13) and (16) one can obtain that the set U^λ is determined by the solutions of T LP problems:

$$(p(t+1)B(t)-\lambda(t)D(t)+b(t),u(t)) \rightarrow \max$$

$$R(t)u(t) \leq q(t) \quad ; \qquad u(t) \geq 0 \tag{21}$$

$$(t=0,\ldots,T-1)$$

where $p(t+1)$ and $\lambda(t)$ are linked by costated equations (8),(9).

According to Lemma 2, minimization of the function $\psi(\lambda)$ over λ implies the solution of the problem, dual to 1. This minimization can be carried out by different ways using different NDO approaches.

If one uses the generalized gradient method technique [1], then the algorithm of finding the optimal value λ^* can be described as follows.

(1) Choose arbitrary dual control $\lambda^\nu = \{\lambda^\nu(t) \geq 0\}$ $(\nu=0,1,\ldots,)$

(2) From dual equations (7),(8) compute corresponding dual trajectory $p^\nu = p(a(T),\lambda^\nu)$.

(3) Using $\{\lambda^\nu(t), p^\nu(t+1)\}$, solve T LP problems (21). If the solution of these problems for some t is not unique, choose arbitrary $u^\nu(t) \ \varepsilon \ U^\lambda(t)$, where $U^\lambda(t)$ is the set of solutions of LP problem (21) for that t.

(4) From primal equations (1),(2) compute trajectory x^ν for $u = u^\nu$.

(5) Compute vector $\partial\psi(\lambda^\nu)$ from (20).

(6) Compute new value $\lambda^{\nu+1}$ from (19) etc.

We shall call this procedure A1.

Theorem 1. Let

$$\alpha_\nu \to 0, \qquad \nu \to \infty; \qquad \sum_{\nu=0}^{\infty} \alpha_\nu = 0 \quad .$$

Then

$$\psi(\lambda^\nu) \to \psi(\lambda^*)$$
$$\lambda^\nu \to \lambda^*$$

The proof of the Theorem follows from [1].

3. *Discussion*

From the proof of Lemma 2, it follows that the solution $\{\lambda^*,\sigma^*,p^*\}$ of dual Problem 1D can be determined from procedure A2.

(1) Compute from A1 the dual control λ^*, optimal for (17).

(2) Compute from the dual state equations (8),(9) the optimal dual trajectory $p^* = p(a(T),\lambda^*)$.

(3) Solve T LP problems:

$$(\sigma(t),q(t)) \to \min$$
$$R^T(t)\sigma(t) \geq b(t) + B^T(t)p^*(t+1) - D^T(t)\lambda^*(t) \qquad (22)$$
$$\sigma(t) \geq 0 \quad .$$

It follows from Lemma 2, that $\{\sigma^*,\lambda^*,p^*\}$, where $\sigma^* = \{\sigma^*(t)\}$ is optimal for problem (22), and is a solution of dual Problem 1D.

4. *Definition of Primal Solution*

If the optimal solution of (17) λ^* was determined, then $u^{\lambda^*} \varepsilon U^{\lambda^*}$ may be a non-optimal control of primal Problem 1. This is because of the unstable properties of the saddle-point set of Lagrange function (13) for LP problems.

To obtain the optimal control u^* of the primal Problem 1, one can use, for example, the complementary slackness conditions. That is, λ^* is supposed to be known. Then if $\lambda_i^*(t) > 0$, then

$$[G(t)x(t)+D(t)u(t)]_i = f_i(t) \tag{23}$$

if

$$[B^T(t)p(t+1)-D^T(t)\lambda(t)-R^T(t)\sigma(t)]_j <- b_j(t) \quad ,$$

then

$$u_j^*(t) = 0 \quad . \tag{24}$$

So, to find u^*, it is necessary to solve DLP Problem 1P with constraints (23),(24).

5. *Conclusion*

The algorithm considered above has two "control" parameters:

1. Choice of vectors $u^\lambda \varepsilon U^\lambda$ (which corresponds to the choice of direction in (19)),
2. Choice of the step size α^ν.

The choice of these parameters determines the properties of concrete realization of the algorithm and there is enormous room for application of different NDO approaches and ideas.

The investigation of different approaches to this problem may be of special interest, because here we deal with linear programming problems, so the choice of u^λ and the determination of u^* can be connected with "finite" procedures of symplex-method.

References

[1] Shor, N.Z., Generalized Gradient Methods for Minimization of Nondifferentiable Functions and Its Applications to Mathematical Programming, *Economika and Math. Methody*, XII, 2(1976), (in Russian).

[2] Lemarechal, C., An Extension of 'Davidon' Methods to Nondifferentiable Problems, *Nondifferentiable Optimization*, Mathematical Programming Study 3, M.L. Balinski and P. Wolfe, eds., (1975).

[3] Mifflin, R., An Algorithm of Nonsmooth Optimization, *Abstracts of IX International Symposium on Mathematical Programming*, Budapest, (1976).

[4] Nurminskii, E.A., The Convergence Conditions of Nonlinear Programming Algorithms, *Kibernetica*, no. 6, (1972), (in Russian).

[5] Propoi, A.I., *Problems of DLP*, RM-76-78, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

[6] Propoi, A.I., *The Duality Systems of DLP*, RM- -76, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976, (forthcoming).