KYBERNETIKA — VOLUME 23 (1987), NUMBER 1

# ON NUMERICAL SOLUTION OF OPTIMAL CONTROL PROBLEMS WITH NONSMOOTH OBJECTIVES: APPLICATION TO ECONOMIC MODELS

JIŘÍ V. OUTRATA

The possibility to use bundle methods for the numerical solution of optimal control problems with nonsmooth locally Lipschitz objectives is investigated. Conditions are specified, under which elements of the generalized gradients of Clarke can be easily computed via an appropriate adjoint equation. The approach is applied to some economic models.

## 1. INTRODUCTION

Optimal control problems with nonsmooth objectives appear frequently in many practical situations, see e.g. [2] or [9]. For such problems the necessary optimality conditions in form of an appropriate maximum principle have been derived by Clarke [1], allowing even the system dynamics to be nonsmooth. Sufficient optimality conditions for a certain class of such problems can be found in [4]. Existing numerical methods enable us to solve problems of this kind numerically; however, we face a lot of technical difficulties in this context. In this paper we study one of these obstacles in connection with so called bundle methods, cf. [13]. It contains a generalization of the results from [9], where also the same problem has been attacked.

We consider the following general model:

$$(\mathscr{P}) \qquad \begin{aligned} &J(x, u) \to \inf \\ \text{subj. to} \quad &A(x, u) = 0 \\ &u \in \omega, \end{aligned}$$

where $x \in X$ and $u \in U$ are the state and control variable, respectively.

We assume that

(i) the optimality criterion $J[X \times U \to \mathbb{R}]$ is locally Lipschitz over $X \times U$;

(ii) the operator $A[X \times U \to X]$ is continuously Fréchet differentiable over $X \times U$, the system equation $A(x, u) = 0$ defines a unique implicit function $x = \mu(u)$ and the operator $\mu[U \to X]$ is continuously Fréchet differentiable over $U$.

The problem $(\mathscr{P})$ may be rewritten into the mathematical programming form

$$(\mathscr{M}\mathscr{P}) \qquad \begin{aligned} &\Phi(u) = J(\mu(u), u) \to \inf \\ &\text{subj. to} \\ &u \in \omega\,. \end{aligned}$$

Assuming that the structure of the set of feasible controls $\omega$ is sufficiently simple (e.g. merely upper and lower bounds), a bundle method of the type derived in [6] may be used for the numerical solution of $(\mathscr{M}\mathscr{P})$ provided we are able to compute for any $u \in \omega$ at least one element of $\partial\Phi(u)$ – the generalized gradient of Clarke of $\Phi$ at $u$. However, this operation may be generally very complicated and it is necessary that it does not spend too much computational time – otherwise it is useless. In the papers [8] and [9] two different ways are proposed how to tackle this question. In this paper we continue in the direction of [9] and substantionally generalize the appropriate results (Sect. 3). Section 4 illustrates the proposed approach on two economic models. The first one was taken from [3] and concerns the employment and wage policies of a monopolistic firm, the second is a generalized production-inventory model of the type discussed in [4]. As we have no meaningful economic data, we could not perform any numerical experiments; however, the approach has been successfully tested on another practical problems cf. [7].

In both economic models in Section 4 we face also inequality state-space constraints of the form

$$-q(x) \in D \subset Z\,, \tag{1.1}$$

where $Z$ is a Banach space (termed usually the *constraint space*), $D$ is a closed convex cone with the vertex at the origin and $q$ is a locally Lipschitz map $[X \to Z]$. We will augment these constraints into the objective by means of exact penalties of the form

$$P(x) = r\,\text{dist}_Z(-q(x), D)\,, \tag{1.2}$$

where $r > 0$ is so called *penalty parameter*. The augmented problem is for a sufficiently large $r$ equivalent to the original one under some very mild assumptions, cf. [5] for the complete explanation. Henceforth, the augmented cost will be denoted by $\Theta$, i.e.

$$\Theta(u) = \Phi(u) + P \circ \mu(u)\,. \tag{1.3}$$

We refer to [2] for the definition and the basic properties of the generalized gradient of Clarke as well as for the concept of regularity. A brief review about generalized gradients may also be found e.g. in [4]. Some basic calculus rules, the definition of partial generalized gradients and some auxiliary results needed in Section 3 are collected for convenience in the next section.

We employ the standard notation in nondifferentiable optimization $(f', f^0, \partial f)$

as used in the papers by Clarke. $\nabla^s F(x)$, $\nabla F(x)$ is the strict, Fréchet derivative of the function $F$ at $x$, respectively, $D^*$ is the positive polar cone to a cone $D$, $(x)^D$ denotes the projection of $x$ onto $D$, $(\lambda)^+ = \max\{0, \lambda\}$ for $\lambda \in \mathbb{R}$ and $|\cdot|_n$ is (any) norm in $\mathbb{R}^n$. For $X$ being a normed space $B_X^\varepsilon(a) = \{x \in X \mid \|x - a\| \leqq \varepsilon\}$; if $X = \mathbb{R}^n$ we write simply $B_n^\varepsilon(a)$. For $x \in X$, $\Omega \subset X$ $\operatorname{dist}_X(x, \Omega) = \inf_{a \in \Omega} \|a - x\|$.

## 2. CALCULUS OF GENERALIZED GRADIENTS

Let $Y$ and $V$ be two Banach spaces. In what follows we will extensively exploit the following basic propositions of the calculus of generalized gradients:

**Proposition 2.1.** Let functions $f, f_1, \ldots, f_n$ $[V \to \mathbb{R}]$ be Lipschitz near $v_0 \in V$ and $\lambda$ be any scalar. Then,

$$\partial(\lambda f)(v_0) = \lambda\, \partial f(v_0)\,, \tag{2.1}$$

$$\partial\Big(\sum_{i=1}^{n} f_i\Big)(v_0) \subset \sum_{i=1}^{n} \partial f_i(v_0)\,. \tag{2.2}$$

Equality holds in (2.2) if either all but at most one of the functions $f_i$ are strictly. differentiable at $v_0$, or each $f_i$ is regular at $v_0$ (in which case the sum is also regular).

**Proposition 2.2.** Let $F$ be a map from $Y$ to $V$ which is strictly differentiable at $y_0$. Assume that the function $g[V \to \mathbb{R}]$ is Lipschitz near $F(y_0)$. Then, for $f = g \circ F$ one has

$$\partial f(y_0) \subset (\nabla^s F(y_0))^*\, \partial g(F(y_0))\,. \tag{2.3}$$

Equality holds in (2.3) if either $g$ is regular at $F(y_0)$ (in which case $f$ is regular at $y_0$) or $-g$ is regular at $F(y_0)$ or $F$ maps every neighbourhood of $y_0$ to a set which is dense in a neighbourhood of $F(y_0)$.

**Proposition 2.3.** Let function $f_1, f_2, \ldots, f_n$ $[V \to \mathbb{R}]$ be Lipschitz near $v_0$. Then for $f = \max_{i=1,2,\ldots,n} f_i$

$$\partial f(v_0) \subset \operatorname{co}\{\partial f_i(v_0) \mid i \in I(v_0)\}\,, \tag{2.4}$$

where $I(v) = \{i \in \{1, 2, \ldots, n\} \mid f(v) = f_i(v)\}$.
Equality holds in (2.4) if $f_i$ is regular at $v_0$ for each $i \in I(v_0)$ and then also $f$ is regular at $v_0$.

**Proposition 2.4.** Let $f[[0, 1] \times \mathbb{R}^m \to \mathbb{R}]$ and

$$G(v) = \int_0^1 f(t, v(t))\, \mathrm{d}t \tag{2.5}$$

be an integral functional defined over $L_\infty[0, 1, \mathbb{R}^m]$. We suppose that

(i) $G$ is (finitely) defined at a function $v_0 \in L_\infty[0, 1, \mathbb{R}^m]$;

(ii) there exists $\varepsilon > 0$ and a function $k \in L_1[0, 1]$ such that for a.e. $t \in [0, 1]$

and for all $v_1, v_2 \in v_0(t) + B_m^\varepsilon(0)$

$$|f(t, v_1) - f(t, v_2)| \leqq k(t)\,|v_1 - v_2|_m\,; \tag{2.6}$$

(iii) the mapping $t \mapsto f(t, v)$ is measurable for each $v \in \mathbb{R}^m$.

Then $G$ is Lipschitz in a neighbourhood of $v_0$ and one has

$$\partial G(v_0) \subset \int_0^1 \partial f_t(v_0(t))\,\mathrm{d}t\,, \tag{2.7}$$

where $\partial f_t(v_0(t))$ is the generalized gradient of the function $v \mapsto f(t, v)$ at $v_0(t)$. Moreover, if $f(t, \cdot)$ is regular at $v_0(t)$ for each $t \in [0, 1]$, then $G$ is regular at $v_0$ and equality holds in (2.7)

The proof of all preceding propositions may be found in [2].

If we compute points of generalized gradients of functions defined over the Carthesian product $V_1 \times V_2 \times \ldots \times V_n$ of Banach spaces $V_i$, $i = 1, 2, \ldots, n$, we may often utilize the concept of partial generalized gradients. So, let $f[V_1 \times V_2 \times \ldots \ldots V_n \to \mathbb{R}]$ be Lipschitz near $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n)$. For $i = 1, 2, \ldots, n$ we introduce the *partial generalized gradients* $\partial_{v_i} f(\bar{v}_1, \bar{v}_2, \ldots \bar{v}_n)$ as the generalized gradients of the functions $v_i \mapsto f(\bar{v}_1, \ldots, \bar{v}_{i-1}, v_i, \bar{v}_{i+1}, \ldots, \bar{v}_n)$ at points $\bar{v}_i$.

The computation of points belonging to partial generalized gradients is usually easier than those of $\partial f$; hence it is convenient for our purposes if for all $(v_1, v_2, \ldots \ldots, v_n) \in V_1 \times V_2 \times \ldots \times V_n$

$$\mathop{\times}_{i=1}^{n} \partial_{v_i} f(v_1, v_2, \ldots, v_n) \subset \partial f(v_1, v_2, \ldots, v_n)\,. \tag{2.8}$$

Besides the trivial situation $f(v_1, v_2, \ldots, v_n) = \varphi(v_1, v_2, \ldots, v_n) + \sum_{i=1}^{n} f_i(v_i)$ with $\varphi$ continuously Fréchet differentiable over $V_1 \times V_2 \times \ldots V_n$ in which inclusion (2.8) holds as an equality, an inclusion of the type (2.8) holds still e.g. in the following cases:

**Proposition 2.5.** Let $f[Y \times V \to \mathbb{R}]$ be Lipschitz near $(y_0, v_0)$. Then

$$\partial_y f(y_0, v_0) \times \partial_v f(y_0, v_0) \subset \partial f(y_0, v_0) \tag{2.9}$$

provided any of the following conditions holds:

(i) $f$ is continuously Fréchet differentiable with respect to $y$ on some neighbourhood $\mathcal{O}$ of $(y_0, v_0)$ and the derivative $\nabla_y f(y, v)$ is continuous on $\mathcal{O}$;

(ii) $f = f_1(y) f_2(v)$ and functions $f_1[Y \to \mathbb{R}]$, $f_2[V \to \mathbb{R}]$ are at $y_0$, $v_0$ regular and nonnegative, respectively (then $f$ is regular at $(y_0, v_0)$ and Incl. (2.9) holds as an equality);

(iii) $f = f_1(y)/f_2(v)$, $f_1[Y \to \mathbb{R}]$ is at $y_0$ regular and nonnegative and $f_2[V \to \mathbb{R}]$ is at $v_0$ regular and negative (then $f$ is regular at $(y_0, v_0)$ and Incl. (2.9) holds as an equality).

Proof. The proof of item (i) may be found in [8]. Items (ii) and (iii) may be easily proved in the same way as Props. 2.3.13, 2.3.14 in [2]. □

We conclude this section with two auxiliary lemmas needed in Section 4. In these lemmas functions $g_1, g_2[\mathbb{R} \to \mathbb{R}]$ are continuously differentiable.

**Lemma 2.6.** Let $f[\mathbb{R}^2 \to \mathbb{R}]$ be given by

$$f(v_1, v_2) = v_1 \max \{g_1(v_1), g_2(v_2)\} .$$

Then, $f$ is locally Lipschitz over $\mathbb{R}^2$, regular over $\mathscr{A} = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 \geqq 0\}$ and at any couple $(v_1, v_2) \in \mathscr{A}$, $\xi \in \partial f(v_1, v_2)$ provided

$$\xi = (g_1(v_1) + v_1 \nabla g_1(v_1), 0) \quad \text{if} \quad g_1(v_1) > g_2(v_2)$$

$$\xi = (g_2(v_2), v_1 \nabla g_2(v_2)) \quad \text{if} \quad g_1(v_1) < g_2(v_2)$$

$$\xi \in \text{co}\,\{(g_1(v_1) + v_1 \nabla g_1(v_1), 0), (g_1(v_1), v_1 \nabla g_2(v_2))\} \quad \text{if} \quad g_1(v_1) = g_2(v_2) .$$

Proof. The local Lipschitz continuity of $f$ is evident. We prove now that the locally Lipschitz function $\Xi[\mathbb{R}^3 \to \mathbb{R}]$ given by

$$\Xi(x_1, x_2, x_3) = x_1 \max \{x_2, x_3\}$$

is regular whenever $x_1 \geqq 0$. Directly by definition in $\mathbb{R}^n$ (cf. Thm. 2.5.1 in [2])

$$\partial \Xi(x_1, x_2, x_3) = \begin{cases} (x_2, x_1, 0) & \text{if} \quad x_2 > x_3 \\ (x_3, 0, x_1) & \text{if} \quad x_2 < x_3 \\ \text{co}\,\{(x_2, x_1, 0), (x_2, 0, x_1)\} & \text{if} \quad x_2 = x_3 . \end{cases}$$

For any $(h_1, h_2, h_3) \in \mathbb{R}^3$

$$\Xi'(x_1, x_2, x_3; h_1, h_2, h_3) =$$

$$= \lim_{\lambda \downarrow 0} [(x_1 + \lambda h_1) \max \{x_2 + \lambda h_2, x_3 + \lambda h_3\} - x_1 \max \{x_2, x_3\}] =$$

$$= \begin{cases} x_2 h_1 + x_1 h_2 & \text{if} \quad x_2 > x_3 \\ x_3 h_1 + x_1 h_3 & \text{if} \quad x_2 < x_3 \\ x_2 h_1 + x_1 \max \{h_2, h_3\} & \text{if} \quad x_2 = x_3 . \end{cases}$$

Thus, for $x_1 \geqq 0$ $\Xi'(x_1, x_2, x_3; h_1, h_2, h_3) = \sup_{\xi \in \partial \Xi(x_1, x_2, x_3)} \langle \xi, (h_1, h_2, h_3) \rangle =$ $= \Xi^0(x_1, x_2, x_3; h_1, h_2, h_3)$ so that $\Xi$ is indeed regular. Then it remains merely to apply Proposition 2.2. □

**Lemma 2.7.** Let $f[\mathbb{R}^2 \to \mathbb{R}]$ be given by

$$f(v_1, v_2) = \begin{cases} c_1(g_1(v_1) - g_2(v_2)) & \text{if} \quad g_1(v_1) \geqq g_2(v_2) \\ -c_2(g_1(v_1) - g_2(v_2)) & \text{otherwise} , \end{cases}$$

where $c_1, c_2$ are nonnegative constants. Then $f$ is regular locally Lipschitz over $\mathbb{R}^2$ and $\xi \in \partial f(v_1, v_2)$ provided

$$\xi = c_1(\nabla g_1(v_1), - \nabla g_2(v_2)) \quad \text{if} \quad g_1(v_1) > g_2(v_2)$$

$$\xi = -c_2(\nabla g_1(v_1), - \nabla g_2(v_2)) \quad \text{if} \quad g_1(v_1) < g_2(v_2)$$

$$\xi \in \text{co}\,\{(c_1 \nabla g_1(v_1), - c_1 \nabla g_2(v_2)), (-c_2 \nabla g_1(v_1), c_2 \nabla g_2(v_2))\} \quad \text{if} \quad g_1(v_1) = g_2(v_2)$$

Proof. Apply Propositions 2.2, 2.3. □

## 3. MAIN RESULTS

Let us consider first the problem $(\mathscr{P})$ without state-space constraints. The following assertion holds:

**Proposition 3.1.** Let $\bar{u} \in U$ be an arbitrary control and $\bar{x} = \mu(\bar{u})$ be the corresponding trajectory. Let $(\bar{\xi}, \bar{\eta}) \in \partial J(\bar{x}, \bar{u})$ and $\lambda^*$ be a solution of the adjoint equation

$$(\nabla_x A(\bar{x}, \bar{u}))^* \lambda^* + \bar{\xi} = 0 . \tag{3.1}$$

Then

$$(\nabla_u A(\bar{x}, \bar{u}))^* \lambda^* + \bar{\eta} \in \partial \Phi(\bar{u}) , \tag{3.2}$$

provided either $J$ (or $-J$) is regular at $(\bar{x}, \bar{u})$ or $\mu$ maps every neighbourhood of $\bar{u}$ onto a set which is dense in some neighbourhood of $\bar{x}$ (in particular, if the derivative $\nabla\mu(\bar{u})$ is surjective).

Proof. $\Phi$ is clearly locally Lipschitz over $U$ so that we are entitled to compute $\partial\Phi$ at any $\bar{u} \in U$. Let us denote $F[U \to X \times U]$ the map given by

$$F: u \mapsto (\mu(u), u) .$$

This map is continuously Fréchet differentiable over $U$; hence we may utilize the relation (2.3) of Proposition 2.2 which holds as an equality due to the assumptions being imposed. We obtain

$$\partial\Phi(\bar{u}) = (\nabla F(\bar{u}))^* \partial J(\bar{x}, \bar{u}) \ni (\nabla\mu(\bar{u}))^* \bar{\xi} + \bar{\eta}$$

and it remains to express the operator $(\nabla\mu(\bar{u}))^*$ by means of derivatives $\nabla_x A(\bar{x}, \bar{u})$, $\nabla_u A(\bar{x}, \bar{u})$. Clearly, for any $u \in U$ $A(\mu(u), u) = 0$ so that

$$\nabla_x A(\bar{x}, \bar{u}) \nabla\mu(\bar{u}) + \nabla_u A(\bar{x}, \bar{u}) = 0 .$$

Thus, on using the adjoint equation, one has for $k \in U$

$$\begin{aligned} \langle(\nabla\mu(\bar{u}))^* \bar{\xi}, k\rangle &= \langle -(\nabla_x A(\bar{x}, \bar{u}))^* \lambda^*, \nabla\mu(\bar{u}) k\rangle = \\ &= \langle \lambda^*, \nabla_u A(\bar{x}, \bar{u}) k\rangle = \langle(\nabla_u A(\bar{x}, \bar{u}))^* \lambda^*, k\rangle . \end{aligned}$$ □

*Remark.* If $\partial_x J(\bar{x}, \bar{u}) \times \partial_u J(\bar{x}, \bar{u}) \subset \partial J(\bar{x}, \bar{u})$, then it suffices to take $\bar{\xi} \in \partial_x J(\bar{x}, \bar{u})$, $\bar{\eta} \in \partial_u(J(\bar{x}, \bar{u})$.

The assertion below concerns $(\mathscr{P})$ with $\Phi$ replaced by $\Theta$ given by (1.3); hence problems with state-space constraints of the type (1.1) treated via exact penalty (1.2).

**Proposition 3.2.** Let $\bar{u} \in U, \bar{x} = \mu(\bar{u})$, $q$ be continuously Fréchet differentiable over $X$ and $(\bar{\xi}, \bar{\eta}) \in \partial J(\bar{x}, \bar{u})$. Let

$$\bar{\varkappa} \in (\nabla q(\bar{x}))^* \partial \operatorname{dist}_Z(-q(\bar{x}), D) \tag{3.3}$$

and $\lambda^*$ be a solution of the adjoint equation

$$(\nabla_x A(\bar{x}, \bar{u}))^* \lambda^* + \bar{\xi} + r\bar{\varkappa} = 0 . \tag{3.4}$$

Then

$$(3.5)\qquad (\nabla_u A(\bar{x}, \bar{u}))^* \lambda^* + \bar{\eta} \in \partial\Theta(\bar{u}),$$

provided $J$ is regular at $(\bar{x}, \bar{u})$.

Proof. It suffices to repeat the proof of Theorem 2.1 with $J(x, u)$ replaced by $J(x, u) + r\,\mathrm{dist}_Z(-q(x), D)$. The function $v \mapsto \mathrm{dist}_Z(v, D)$ is convex and hence regular so that we may compute elements of $\partial P(x)$ by means of Incl. (3.3) due to Proposition 2.2. The regularity of $J$ is needed to ensure that

$$(\bar{\xi} + r\bar{\varkappa}, \bar{\eta}) \in \partial(J + P)(\bar{x}, \bar{u})$$

by Proposition 2.1. □

For computing elements of $\partial P(x)$ the following assertion may be utilized.

**Lemma 3.1.** Let $Z$ be Hilbert. Then

$$(3.6)\qquad P(x) = r\|(q(x))^{D^*}\|$$

and for $z \in Z$

$$(3.7)\qquad \partial\|(z)^{D^*}\| = \begin{cases} \dfrac{(z)^{D^*}}{\|(z)^{D^*}\|} & \text{if } -z \notin D \\ B_Z^1(0) \cap D^* \cap \{z\}^\perp & \text{if } -z \in D. \end{cases}\qquad (3.8)$$

Proof. Assertion (3.6) is one of the forms of the well-known orthogonal decomposition

$$(3.9)\qquad z = z^{D^*} + z^{-D}$$

valid in Hilbert spaces. Concerning Eq. (3.7) we refer to [12], where the differentiability of the functional $\|(z)^{D^*}\|^2$ is examined. Thus, let us prove (3.8). Clearly, for $\varphi(z) = \|(z)^{D^*}\|$, $z = 0$, $h \in Z$

$$\varphi'(0; h) = \lim_{\lambda \downarrow 0} \frac{1}{\lambda} \|(\lambda h)^{D^*}\| = \|(h)^{D^*}\|$$

so that

$$\partial\varphi(0) = \{\xi \in Z \mid \langle h, \xi \rangle \leqq \|(h)^{D^*}\| \quad \forall h \in Z\}.$$

Using the decomposition (3.9), we obtain

$$\langle (h)^{D^*}, \xi \rangle + \langle (h)^{-D}, \xi \rangle \leqq \|(h)^{D^*}\|$$

which implies that $\partial\varphi(0) \subset B_Z^1(0) \cap D^*$.

Conversely, if $\xi \in B_Z^1(0) \cap D^*$, then for any $h$

$$\langle h, \xi \rangle = \langle (h)^{D^*}, \xi \rangle + \langle (h)^{-D}, \xi \rangle \leqq \langle (h)^{D^*}, \xi \rangle \leqq \|(h)^{D^*}\|$$

so that $\partial\varphi(0) = B_Z^1(0) \cap D^*$.

As $\varphi$ is a positively homogeneous function, we know from the convex analysis that for $z \neq 0$

$$\partial\varphi(z) = \{\xi \in B_Z^1(0) \cap D^* \mid \langle z, \xi \rangle = \varphi(z)\}.$$

Hence for $z \in -D$ $(\varphi(z) = 0)$ $\partial\varphi(z) = B_Z^1(0) \cap D^* \cap \{z^\perp\}$. □

## 4. ECONOMIC MODELS

We show now on two economic optimization problems of the type ($\mathscr{P}$) with state-space constraints (1.1) how to apply Proposition 3.2 to the computation of elements of generalized gradients of their augmented objectives.

The first one has been developed in [3] and concerns the optimal employment and wage policy of a monopolistic firm. Using the notation of [3], the problem attains the form

$$e^{-\vartheta T} S\, y(T) + \int_0^T e^{-\vartheta t}[c(f(y(t))) + w(t)\, y(t) + k(u(t)) + \varphi(\eta(p(t)) - f(y(t))) +$$
$$+ p(t) \max\{-\eta(p(t)), -f(y(t))\}]\, dt \to \inf$$

subj. to

(4.1) $$\dot{y}(t) = u(t) - \sigma(w(t))\, y(t) \quad \text{a.e. on} \quad [0, T], \quad y(0) \geqq 0 \quad \text{given},$$
$$y(t) \geqq 0$$
$$w(t) \geqq \bar{w}$$
$$p(t) \geqq 0 \quad \text{for all} \quad t \in [0, T],$$

where

| | |
|---|---|
| $T$ | is the given finite horizon, |
| $\vartheta$ | is an interest rate assumed to be constant and positive through time |
| $y$ | is the level of employment assumed to be homogeneous, |
| $f(y)$ | is the production function depending only on labour, |
| $c(f(y))$ | are the production costs, |
| $w$ | is the wage, |
| $u$ | is the rate of recruitment or discharge, |
| $k(u)$ | is the labour adjustment cost function, |
| $p$ | is the selling price charged by the firm, |
| $\eta(p)$ | is the demand function, |
| $\varphi(\eta(p) - f(y))$ | denotes shortage costs to be paid by the firm if $\eta(p) > f(y)$ and disposal costs if $\eta(p) < f(y)$, |
| $S$ | is the (constant) salvage value of the employment level, |
| $\sigma(w)$ | is the voluntary decrease of employment. |

In problem (4.1) we have three controls $u$, $w$ and $p$ and one state variable $y$. With respect to the system equation it seems reasonable to set $U = L_\infty[0, T, \mathbb{R}^3]$. According to [3] we will assume that $\sigma[\mathbb{R} \to \mathbb{R}]$ is continuously differentiable and non-negative over $\mathbb{R}$. Then we may set $X = C_0[0, T]$ and it can easily be shown that the system equation in (4.1) satisfies all the assumptions listed in the introduction. Furthermore, we will assume that $c, f, k, \eta$ $[\mathbb{R} \to \mathbb{R}]$ are also $C_1$ functions and

(4.2) $$\varphi(\varkappa) = \begin{cases} c_1\varkappa & \text{for} \quad \varkappa \geqq 0 \\ -c_2\varkappa & \text{for} \quad \varkappa < 0 \end{cases}$$

with $c_1$, $c_2$ given positive constants. The state-space constraint will be augmented by the exact penalty

$$(4.3) \qquad P(y) = r \int_0^T (-y(t))^+ \, \mathrm{d}t$$

(i.e. we set $Z = L_1[0, T]$). Thus, at a process $(\bar{y}, \bar{u}, \bar{w}, \bar{p})$

$$\Theta(\bar{u}, \bar{w}, \bar{p}) = \mathrm{e}^{-\vartheta T} S \, \bar{y}(T) + \int_0^T \mathrm{e}^{-\vartheta t} [c(f(\bar{y}(t)) + \bar{w}(t) \, \bar{y}(t) + k(\bar{u}(t)) +$$
$$+ \varphi(\eta(\bar{p}(t)) - f(\bar{y}(t))) + \bar{p}(t) \max \{-\eta(\bar{p}(t)), -f(\bar{y}(t))\}] \, \mathrm{d}t + r \int_0^T (-\bar{y}(t))^+ \, \mathrm{d}t \, .$$

The control constraints are clearly simple enough to be treated directly within the minimization routine. Proposition 3.2 implies the following assertion:

**Proposition 4.1.** Let $(\bar{u}, \bar{w}, \bar{p}) \in L_\infty[0, T, \mathbb{R}^3]$ with $\bar{p}(t) \geqq 0$ a.e. on $[0, T]$ and $\bar{y}$ be the corresponding trajectory. Let $s$ be the solution of the adjoint differential equation

$$(4.4) \qquad \dot{s}(t) = \sigma(\bar{w}(t)) \, s(t) +$$
$$+ \mathrm{e}^{-\vartheta t} [\nabla c(f(\bar{y}(t))) \, \nabla f(\bar{y}(t)) + \bar{w}(t) + \beta_1(\bar{y}(t), \bar{p}(t))] + \beta_2(\bar{y}(t)) \quad \text{a.e.}$$

backwards on $[0, T]$ from the terminal condition

$$(4.5) \qquad s(T) = -\mathrm{e}^{\vartheta T} S \, .$$

Functions $\beta_1[\mathbb{R}^2 \to \mathbb{R}]$, $\beta_2[\mathbb{R} \to \mathbb{R}]$ in Eq. (4.4) are given by

$$\beta_1(\bar{y}(t), \bar{p}(t)) =$$
$$= \begin{cases} -c_1 \nabla f(\bar{y}(t)) - \bar{p}(t) \nabla f(\bar{y}(t)) & \text{for} \quad t \in I_1 = \{t \in [0, T] \mid \eta(\bar{p}(t)) \geqq f(\bar{y}(t))\} \\ c_2 \nabla f(y(t)) & \text{for} \quad t \in [0, T] \setminus I_1 \, , \end{cases}$$

$$\beta_2(\bar{y}(t)) = \begin{cases} 0 & \text{for} \quad t \in I_2 = \{t \in [0, T] \mid \bar{y}(t) \geqq 0\} \\ -r & \text{for} \quad t \in [0, T] \setminus I_2 \, . \end{cases}$$

Then the function $\gamma[[0, T] \to \mathbb{R}^3]$ given by

$$(4.6) \qquad \gamma(t) = \begin{bmatrix} -1 \\ \bar{y}(t) \nabla \sigma(\bar{w}(t)) \\ 0 \end{bmatrix} s(t) + \mathrm{e}^{-\vartheta t} \begin{bmatrix} \nabla k(\bar{u}(t)) \\ \bar{y}(t) \\ \beta_3(\bar{y}(t), \bar{p}(t)) \end{bmatrix}, \quad t \in [0, T] \, ,$$

with $\beta_3[\mathbb{R}^2 \to \mathbb{R}]$ given by

$$\beta_3(\bar{y}(t), \bar{p}(t)) = \begin{cases} c_1 \nabla \eta(\bar{p}(t)) - f(\bar{y}(t)) & \text{for} \quad t \in I_1 \\ -c_2 \nabla \eta(\bar{p}(t)) - \eta(\bar{p}(t)) - \bar{p}(t) \nabla \eta(\bar{p}(t)) & \text{for} \quad t \in [0, T] \setminus I_1 \, , \end{cases}$$

belongs to $\partial \Theta(\bar{u}, \bar{w}, \bar{p})$.

Proof. The integrand in the objective possesses two nonsmooth terms and for $\bar{p}(t) \geqq 0$ a.e. on $[0, T]$ both are regular. Hence, the whole integrand is regular and Proposition 2.4 may be utilized with Incl. (2.7) attained as an equality. Points belonging to generalized gradients of two nonsmooth terms in the integrand of the cost as well as a point from the generalized gradient of the integrand of the penalty may be computed with the help of Lemmas 2.6, 2.7. Finally, it remains to apply Proposition 3.2, using the well-known tricks to construct the adjoint equation (4.4), (4.5), cf. e.g. [11]. □

The second economic problem is a generalization of the production-inventory model from [4]. With the notation taken partially from (4.1) it attains the form

$$e^{-\vartheta T} S\, x(T) + \int_0^T e^{-\vartheta t}[c(v(t)) + \psi(x(t)) + \varphi(\eta(p(t)) - o(t)) - p(t)\, o(t)]\, dt \to \inf$$

subj. to

$$(4.7)\qquad \begin{aligned} &\dot{x}(t) = v(t) - o(t) \quad \text{a.e. on} \quad [0, T], \quad x(0) \geqq 0 \quad \text{given},\\ &x(t) \geqq 0\\ &0 \leqq v(t) \leqq b\\ &p(t) \geqq 0\\ &0 \leqq o(t) \leqq \eta(p(t)) \quad \text{for all} \quad t \in [0, T], \end{aligned}$$

where

| | |
|---|---|
| $x$ | is the inventory, |
| $v$ | is the production rate, |
| $b$ | is the upper bound for the production rate, |
| $c(v)$ | are the production costs (inclusive wages), |
| $\psi(x)$ | are the holding costs, |
| $o$ | is the actual supplied output, |
| $\varphi(\eta(p) - o)$ | denotes the shortage costs to be paid by the firm if $\eta(p) > o$, |
| $po$ | is the return, |
| $S$ | is the (constant) salvage value of the inventory level, |

and all other symbols have the same meaning as in (4.1).

In problem (4.7) we have basically three controls $v$, $p$ and $o$ and one state variable $x$ if only one commodity is produced. However, the model remains valid if $n$ commodities are produced, of course with $3n$ control and $n$ state variables.

*Remark.* For a non-monopolistic firm it could be dangerous to apply intentionally the control $o < \eta(p)$ because of the competition on the matket. This risk is not included in model (4.7).

We will assume that $c[\mathbb{R} \to \mathbb{R}]$ is continuously differentiable,

$$(4.8)\qquad \psi(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ h_1 x & \text{for} \quad 0 \leqq x < x_1 \\ h_2 x - (h_2 - h_1)\, x_1 & \text{for} \quad x \geqq x_1 \end{cases}$$

with $h_1$, $h_2$ given positive constant and

$$(4.9)\qquad \varphi(\varkappa) = d(\varkappa)^+$$

with $d$ being a given positive constant. The piecewise linear form of the holding costs was taken from the paper [4] and corresponds to a certain warehousing constraint $x_1$. For inventory levels $x > x_1$ an additional space has to be rented at a unit cost of $h_2 > h_1$.

It remains to augment suitably the state-space constraint and the nontrivial control constraint $o(t) \leqq \eta(p(t))$, $t \in [0, T]$ which e.g. in the code of Cl. Lemaréchal written according to [6] cannot be treated directly. Fortunately, we may utilize the form of $\psi$ and $\varphi$ and perform the augmentation just by replacing $\psi$ by $\tilde{\psi}$ and $\varphi$ by $\tilde{\varphi}$ with $\tilde{\psi}$, $\tilde{\varphi}$ given by

$$(4.10) \qquad \tilde{\psi}(t, x) = \begin{cases} -r\, e^{\vartheta t} x & \text{for } x < 0 \\ h_1 x & \text{for } 0 \leqq x < x_1 \\ h_2 x - (h_2 - h_1)\, x_1 & \text{for } x_1 \leqq x, \end{cases}$$

$$(4.11) \qquad \tilde{\varphi}(t, \varkappa) = \begin{cases} -r\, e^{\vartheta t} \varkappa & \text{for } \varkappa < 0 \\ d\varkappa & \text{for } \varkappa \geqq 0. \end{cases}$$

It corresponds to using of exact penalties with $Z = L_1[0, T]$ and the penalty parameter $r$ for both constraints. The remaining control constraints are simple enough to be treated within the minimization routine. As in the preceding problem we set $U = L_\infty[0, T, \mathbb{R}^3]$, $X = C_0[0, T]$ and apply Prop. 3.2 which provides us with the following assertion:

**Proposition 4.2.** Let $(\bar{v}, \bar{p}, \bar{o}) \in L_\infty[0, T, \mathbb{R}^3]$ and $\bar{x}$ be the corresponding trajectory. Let $s$ be the solution of the adjoint differential equation

$$(4.12) \qquad \dot{s}(t) = e^{-\vartheta t}\, \tilde{\beta}_1(t, \bar{x}(t)) \quad \text{a.e.}$$

backwards on $[0, T]$ from the terminal condition (4.5). The function $\tilde{\beta}_1[[0, T] \times \times \mathbb{R} \to \mathbb{R}]$ in Eq. (4.12) is given by

$$\tilde{\beta}_1(t, \bar{x}(t)) = \begin{cases} h_2 & \text{for } t \in N_2 = \{t \in [0, T] \mid \bar{x}(t) \geqq x_1\} \\ h_1 & \text{for } t \in N_1 = \{t \in [0, T] \mid 0 \leqq \bar{x}(t) < x_1\} \\ -er^{\vartheta t} & \text{for } t \in [0, T] \setminus (N_1 \cup N_2). \end{cases}$$

Then the function $\tilde{\gamma}[[0, T] \to \mathbb{R}^3]$ given by

$$(4.13) \qquad \tilde{\gamma}(t) = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} s(t) + e^{-\vartheta t} \begin{bmatrix} \nabla c(\bar{v}(t)) \\ \\ \tilde{\beta}_2(t, \bar{p}(t), \bar{o}(t)) \end{bmatrix}, \quad t \in [0, T]$$

with $\tilde{\beta}_2[[0, T] \times \mathbb{R}^2 \to \mathbb{R}^2]$ given by

$$\tilde{\beta}_2(t, \bar{p}(t), \bar{o}(t)) =$$

$$= \begin{cases} (d\, \nabla\eta(\bar{p}(t)) - \bar{o}(t), -d - \bar{p}(t)) & \text{for } t \in M_1 = \{t \in [0, T] \mid \eta(\bar{p}(t)) \geqq \bar{o}(t)\} \\ (-r\, e^{\vartheta t}\, \nabla\eta(\bar{p}(t)) - \bar{o}(t), r\, e^{\vartheta t} - \bar{p}(t)) & \text{for } t \in [0, T] \setminus M_1, \end{cases}$$

belongs to $\partial\Theta(\bar{v}, \bar{p}, \bar{o})$.

The proof may be performed along the same lines as the proof of Proposition 4.1, using the obvious generalization of Lemma 2.7 for the computation of a generalized gradient of $\tilde{\psi}$.

Thus, for the numerical solution of both above problems, a numerical method of the bundle type could be used provided we replace $U$ by a finite-dimensional subspace. This corresponds clearly also to the economic reality. From the numerical point of view it is then recommended to choose a consistent difference scheme and a quadrature formula for the system equation and the objective, respectively, and modify appropriately Proposition 4.1, 4.2 for the computation of the generalized gradients of $\Theta$ for the discretized problems. Propositions 3.1, 3.2 form again the theoretical basis for these modifications.

*Remark.* In both above problems $L_1$-exact penalties have been used for the augmentation of state-space and nontrivial control constraints. However, there is an encouraging experience with using Hilbert space norms, actually with $Z = L_2$ or $H^1$ (if possible), cf. [10].

*Remarks.* Models (4.1) and (4.7) may be easily connected, arriving in such a way at a rather complex employment-production-inventory model

$$e^{-\vartheta T}(S_1 x(T) + S_2 y(T)) + \int_0^T e^{-\vartheta t} [c(f(y(t))) + w(t) y(t) + k(u(t)) + \psi(x(t)) + \\ + \varphi(\eta(p(t)) - o(t)) - p(t) o(t)] \, dt \to \inf$$

subj. to

$$\begin{aligned} &\dot{x}(t) = f(y(t)) - o(t) \quad \text{a.e. on} \quad [0, T], \quad x(0) \geqq 0 \quad \text{given}, \\ &\dot{y}(t) = u(t) - \sigma(w(t)) y(t) \quad \text{a.e. on} \quad [0, T], \quad y(0) \geqq 0 \quad \text{given}, \\ &x(t) \geqq 0 \\ &y(t) \geqq 0 \\ &w(t) \geqq \bar{w} \\ &p(t) \geqq 0 \\ &0 \leqq o(t) \leqq \eta(p(t)) \quad \text{for all} \quad t \in [0, T], \end{aligned} \tag{4.14}$$

with four controls and two state variables. This model will be studied elsewhere.

## 5. CONCLUSION

Propositions 3.1, 3.2 are in fact an application of the chain rule II of Clarke (Prop. 2.2) to extremal problems of the type ($\mathscr{P}$). Unfortunately, unlike from necessary optimality conditions, we need equality in Incl. (2.3) and both appropriate assumptions (regularity of $J(-J)$ and surjectivity of $\nabla\mu$) are rather restrictive. Nevertheless, there is a lot of practical problems in which either of these assumptions can be satisfied as documented in Section 4. In [3], [4] the authors made the complete synthesis for problem (4.1) and a simplified version of problem (4.7). This can be done practically only in the case of one state variable. Our approach allows to find approximately optimal controls independently on the number of state-variables (hence it can be applied also to model (4.14)) and also the number of kinks in the objective may be increased.

(Received November 22, 1985.)

## REFERENCES


[1] F. H. Clarke: The maximum principle under minimal hypotheses. SIAM J. Control *14* (1976), 1078—1091.

[2] F. H. Clarke: Optimization and Nonsmooth Analysis. Wiley, New York 1983.

[3] G. Feichtinger and M. Luptáčik: Optimal Employment and Wage Policies of a Monopolistic Firm. Research Report No. 62, Inst. for Econometrics and Oper. Res., TU Vienna 1983.

[4] R. F. Hartl and S. P. Sethi: Optimal control problems with differential inclusions: Sufficiency conditions and an application to a production-inventory model. Optimal Control Appl. and Methods *5* (1984), 289—307.

[5] A. D. Ioffe: Necessary and sufficient condition for a local minimum. 2: Conditions of Levitin-Miljutin-Osmolovskii type. SIAM J. Control Optim. *17* (1979), 251—265.

[6] Cl. Lemaréchal, J. J. Strodiot and A. Bihain: On a Bundle Algorithm for Nonsmooth Optimization. NPS 4, Madison 1980.

[7] J. V. Outrata: On a class of nonsmooth optimal control problems. Appl. Math. Optim. *10* (1983), 287—306.

[8] J. V. Outrata: Minimization of nonsmooth nonregular functions: Application to discrete-time optimal control problems. Problems Control Inform. Theory *13* (1984), 6, 413—424.

[9] J. V. Outrata and Z. Schindler: On some nondifferentiable problems in optimal control. In: Nondifferentiable Optimization: Motivations and Applications — Proceedings, Sopron, Hungary, 1984 (Demyanov, Pallaschke, eds.), Lecture Notes in Econ. Math. Systems 255, Springer-Verlag, Berlin—Heidelberg—New York—Tokio 1985, pp. 118—128.

[10] J. V. Outrata and J. Jarušek: Exact penalties with Sobolev norms in optimal control. Preprints of the Workshop "Mathematical Programming — Theory and Applications" held at Wartburg, November 1984.

[11] E. Polak: Computational Methods in Optimization. Academic Press, New York 1971.

[12] E. H. Zarantonello: Projections on convex sets in Hilbert space and spectral theory. In: Contributions to Nonlinear Functional Analysis (E. H. Zarantonello, ed.), Academic Press, New York 1971.

[13] J. Zowe: Nondifferentiable optimization — a motivation and a short introduction into the subgradient — and the bundle concept. In: ASI Proceedings on Computational Mathematical Programming (to appear).



*Ing. Jiří V. Outrata, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Prague 8, Czechoslovakia.*