Control
and Cybernetics
VOL. 4 (1975) No. 1

Application of the adaptive precision conjugate-gradient method to nonlinear planning models of national economy

by

BOLESŁAW TOŁWIŃSKI

Institute for Organization, Management and Control Sciences
Polish Academy of Sciences
Ministry of Science, Higher Education and Technology

The paper presents an algorithm for constrained dynamic optimization problems and discusses its application to the solution of nonlinear planning models, which could be used in formulation of development programs for the Polish economy. The algorithm is based on the ideas of penalty function, conjugate gradient and adaptive precision principle. The emphasis is made on the convergence properties of the algorithm, in the general case and in the case of the planning model. The corresponding theorems are formulated and proved.

1. Introduction

This paper presents an algorithm for constrained dynamic optimization problems and discusses its application to the solution of nonlinear planning models, which could be used in formulation of development programs for the Polish economy. The algorithm is based on the ideas of penalty function [2, 5], conjugate gradient [1, 7] and adaptive precision gradient method [4]. It is defined in the way to make it directly implementable on a digital computer, i.e. each iteration is made up of a finite number of arithmetical operations and approximately executed function evaluations. An emphasis is made on the convergence properties of the sequences of approximate solutions generated by the algorithm. The sufficient conditions are defined for these sequences to be critisizing or minimizing ones (in a generalized sense) and to converge in the weak and the strong sense.

The planning models considered in the paper are based on the ideas put forward by Kendrick and Taylor [3] and adapted here to the specific conditions of the Polish economy.

In Section 2 we present the adaptive precision conjugate-gradient algorithm together with theorems stating its main properties and in Section 3 we describe the model and discuss the application of the algorithm to its solution.

Some proofs are given in Appendix.

2. The adaptive precision conjugate-gradient algorithm

A dynamic optimization problem can be formulated as the problem of locating a minimizing (or maximizing) point for some real-valued functional over a subset of a Banach space.

Let us assume that E is a Banach space, f and g_i, $i=1, 2, \ldots, q$, are functionals over E and Q is a set defined by the formula:

$$Q=\{u \in E | g_i(u)=0 \quad \text{for} \quad i=1, 2, \ldots, \bar{q} \quad \text{and} \quad g_i(u) \leqslant 0 \quad \text{for} \quad i=\bar{q}+1, q+2, \ldots, q\}. \tag{2.1}$$

Now, we seek to minimize the functional f over the set Q. In particular we shall be concerned with optimal control problems, where

$$E=L_2^r(0, T); \tag{2.2}$$

$$\begin{gathered} f(u)=\int_0^T c\big(x(t), u(t), t\big)\, dt, \\ u(t) \in E^r, x(t) \in E^p \quad \text{for} \quad t \in [0, T], \end{gathered} \tag{2.3}$$

$$\begin{gathered} \frac{dx}{dt}=s\big(x(t), u(t), t\big), \\ x(0)=x_0; \end{gathered} \tag{2.4}$$

and

$$Q=\big\{u \in L_2^r(0, T) |\; a_j \leqslant u_j(t) \leqslant b_j \quad \text{for} \quad j=1, 2, \ldots, r \quad \text{and almost all} \quad t \in [0, T], \quad \text{and} \quad x_j(T)=d_j \quad \text{for some} \quad j \in \{1, 2, \ldots, p\}\big\}. \tag{2.5}$$

It is clear that the set (2.5) can be put in the form (2.1) be defining:

$$\begin{gathered} g_1(u)=\|u-\bar{u}\|_2, \text{ where} \\ \bar{u}(t)=\begin{cases} b & \text{if} \quad u(t) \geqslant b \\ u(t) & \text{if} \quad a<u(t)<b, \\ a & \text{if} \quad u(t) \leqslant a \end{cases} \end{gathered} \tag{2.6}$$

and

$$\begin{gathered} g_i(u)=|x_{ji}(T)-d_i| \quad \text{for} \quad i=2, \ldots, q \\ ji \in \{1, 2, \ldots, p\}. \end{gathered} \tag{2.7}$$

To facilitate further considerations we present in the first place an adaptive precision gradient algorithm of the Klessig-Polak type, for unconstrained minimization.

Let j be the computation precision index, $f_j(u)$ the approximate value of the functional f in the point u and $h_j(u)$ the approximate value of the gradient of the functional f in the point u.

Algorithm 1.

Step 0. Select an integer j_0, an $u_0 \in E$, and parameters $\varepsilon_0 > 0$, $\alpha, \beta \in (0, 1)$, $\lambda_{\min} \in (0, 1]$.
Set $j = j_0$, $\varepsilon = \varepsilon_0$ and $n = 0$.

Step 1. Compute $h_j(u_n)$ and $\|h_j(u_n)\| E$.

Step 2. Select the direction of the minimization p_n (for example one can take $p_n = P_{E^{n(j)}} h_j((u_n))$.

Step 3. Set $\lambda = 1$.

Step 4. Set $\Delta = f_j(u_n + \lambda p_n) - f_j(u_n) + (\lambda/2) < h_j(u_n), p_n >$.

Step 5. If $\Delta \leqslant 0$ then go to step 9, else go to step 6.

Step 6. Set $\lambda = \beta\lambda$.

Step 7. If $\lambda \geqslant \lambda_{\min} \varepsilon$ then go to step 4; else go to step 8.

Step 8. Set $j = j+1$, $\varepsilon = \alpha\varepsilon$ and go to step 1.

Step 9. If $f_j(u_n + \lambda p_n) - f_j(u_n) \leqslant -\varepsilon$ then set $u_{n+1} = u_n + \lambda p_n$ and $n = n+1$; else set $j = j+1$ and $\varepsilon = \alpha\varepsilon$.

Step 10. Go to step 1.

THEOREM 2.1. Let us assume that the following conditions are satisfied:

(1) f is bounded below on the set
$W(u_0) = \operatorname{conv} \{u \in E | f(u) \leqslant f(u_0)\}$;

(2) ∇f is Lipschitz continuous on $W(u_0)$,

(3) $\{p_n\}^{\infty}_{n=1}$ give an admissible sequence of directions [1];

(4) $\exists M > 0$, $c > 0$ $\forall n \|p_n\| \leqslant M \wedge \|p_n\| \geqslant c \langle -\nabla f(u_n), p_n \rangle$;

(5) $\forall$ $u \in W(u_0) | f(u) - f_j(u) | \leqslant \delta_j$, $\|\nabla f(u) - h_j(u)\| \leqslant \delta_j$ and $\dfrac{\delta_j}{\varepsilon_j} \xrightarrow[j \to \infty]{} 0$.

Then $\{u_n\}$ is a criticizing sequence, i.e.
$\|\nabla f(u_n)\| \to 0$.

Now we shall consider the adaptive precision conjugate-gradient algorithm for unconstrained minimization, where the search along the line proceeds by means of the Algorithm 1.

Algorithm 2.

Step 0. Select integers j_0, m_0, $l_{\max}$. Select an $a \in E^{m_0}$ and parameters ε_{01}, ε_{02}, $\varepsilon_{03} > 0$, α, β_1, β_2, β_3, $\gamma \in (0, 1)$, $\lambda_{\min} \in (0, 1]$.
Select an orthonormal base $\{z_m\}_{m=1}^{\infty}$ in E.
Set $j_m = j_0$, $m = m_0$, $\varepsilon_1 = \varepsilon_{01}$, $\varepsilon_2 = \varepsilon_{02}$, $\varepsilon_3 = \varepsilon_{03}$,
$u_0 = \sum_{i=1}^{m} a^i z_i$, $n = 0$.

Step 1. Compute $h_j(u_n)$ and $\|h_j(u_n)\|_E$.

Step 2. Compute $h_j^m(u_n) = P_{E^m} \|h_j(u_n)\|$ and $\|h_j^m(u_n)\|_{E^m}$.

Step 3. If $\|h_j^m(u_n)\|_{E^m} < \gamma \|h_j(u_n)\|_{E^m}$ then set $m = m+1$ and go to step 2.

Step 4. Set $h^m c_{nj} = -h_j^m(u_n)$ and $l = 0$.
Comment $h^m c_{nj}$ is the direction of minimization.

Step 5. Set $h = h^m c_{nj} / \|h^m c_{nj}\|_{E^m}$.

Step 6. Set $x=0$.

Step 7. Set $z=u_n+xh$.

Step 8. Compute $\Theta_j'(x)=\langle h_j(z), h\rangle$.

Comment. The function $\Theta(x)$ is defined by the formula:

$$\Theta(x)=f(u_n+xh)-f(u_n).$$

It follows that $\Theta'(x)=\langle \nabla f(u_n+xh), h\rangle=\langle \nabla f(z), h\rangle$
$\Theta_j'(x)$ is taken as the approximation of $\Theta'(x)$.

Step 9. If $|\Theta_j'(x)|/\|h_j(z)\|_E\leqslant\varepsilon_2$ then go to step 17.

Comment. Start of the minimization of the function $\Theta(x)$ by means of the algorithm 1.

Step 10. Set $\lambda=1$

Step 11. Compute $\Delta=\Theta_j(x-\lambda\Theta_j'(x))-\Theta_j(x)+\frac{1}{2}\lambda\Theta_j'(x)^2$.

Step 12. If $\Delta\leqslant 0$ then go to step 16.

Step 13. Set $\lambda=\beta_1\lambda$.

Step 14. If $\lambda\geqslant\varepsilon_1\lambda_{\min}$ then go to step 11.

Step 15. Set $j=j+1$, $\varepsilon_1=\varepsilon_1\alpha$ and go to step 1.

Step 16. If $\Theta_j(x-\lambda\Theta_j'(x))-\Theta_j(x)\leqslant-\varepsilon_1$ then set $x=x-\lambda\Theta_j'(x)$ and go to step 7; else set $j=j+1$, $\varepsilon_1=\varepsilon_1\alpha$ and go to step 1.

Comment. End of the minimization along the line.

Step 17. Compute $h_j^m(z)$.

Step 18. If $\|h_j^m(z)\|_{E^m}\leqslant\gamma\|h_j(z)\|_E$ then set $m=m+1$ and go to step 17.

Step 19. Set $u_{n+1}=z$. If $l=l_{\max}$ then set $n=n+1$ and go to step 4; else set
$b_{nj}=\langle h_j(z)-h_j(u_n), h_j(z)\rangle/\|h_j(z)\|_E^2$,
$h^m c_{n+1j}=-h_j^m(z)+b_{nj}h^m c_{nj}$,
$l=l+1$, $n=n+1$.

Step 20. If $\langle -h_j(u_n), h^m c_{nj}\rangle\geqslant\varepsilon_3\|h_j(u_n)\|\cdot\|h^m c_{nj}\|$ then go to step 5; else set $\varepsilon_2=\varepsilon_2\beta_2$, $\varepsilon_3=\varepsilon_3\beta_3$ and go to step 5.

Comment. The purpose of the operations in step 20 is to find an ε_2 which is compatible with the convergence of the algorithm (see [6] p. 307).

THEOREM 2.2. Let the assumptions (1), (2) and (5) of the Theorem 2.1. be satisfied and let: (6) the set $W(u_0)$ be bounded. Then the sequence $\{u_n\}$ generated by the algorithm 2 is a criticizing one.

By using the penalty-function method it is possible to apply the Algorithm 2 to the constrained minimization, i.e. to minimize a functional f over the set Q given by (2.1). Let us define:

$$f_p(u)=f(u)+\frac{1}{2}K_p\left(\sum_{i=1}^{\bar{q}}[\max(0, g_i(u))]^2+\sum_{i=\bar{q}+1}^{q}[g_i(u)]^2\right) \tag{2.8}$$

and assume that $K_p \xrightarrow[p\to\infty]{} \infty$.

Now, the adaptive precision conjugate-gradient algorithm for constrained minimization can by defined in the following way:

Algorithm 3.

Step 0. Perform step 0 of the algorithm 2.
Select real numbers $K_0>0$ and $s>0$.
Set $K=K_0$ and $p=1$.

Step 1. Apply the operations of step 1 of the Algorithm 2 to the functional f_p.

Step 2. If $\|h_j(u_n)\|\leq 1/K$ then set $u_p=u_n$, $K=Ks$, $p=p+1$ and go to step 1.

Step 3. Perform steps 2—7 of the algorithm 2.

Step 4. If $\|h_j(z)\|\leq 1/K$ then set $u_n=z$, $K=Ks$, $p=p+1$ and go to step 1.

Step 5. Perform steps 8—20 of the algorithm 2.

THEOREM 2.3. Let us assume that the following conditions are satisfied:

(1) Sets $W_p(u_{p-1})=\operatorname{conv}\{u\in E|f(u)\leq f_p(u_{p-1})\}$.

(2) f_p are bounded below on the sets $W_p(u_{p-1})$;

(3) ∇f and ∇g_i $i=1, 2, 3, \dots, q$ are Lipschitz continuous on $\bigcup\limits_{p=1}^{\infty} W_p$;

(4) $\forall u\in W_p(u_{p-1})|f_p(u)-f_{pj}(u)|\leq\delta_j$; $\|\nabla f_p(u)-h_{pj}(u)\|\leq\delta_j$ and $\dfrac{\delta_j}{\alpha_j}\to 0$.

Then the Algorithm 3 is well defined and
$\|\nabla f_p(u_p)\|\xrightarrow[p\to\infty]{}0$.

THEOREM 2.4. If in addition to assumptions (3) and (4) of the Theorem 2.3. the set $\bigcup\limits_{p=1}^{\infty} W_p(u_{p-1})$ is bounded, f_p are bounded below on E, f and g_i, $i=1, 2, \dots, q$, are convex, and f is Lipschitz continuous in a neighbourhood of the nonempty set Q, then

$$\limsup_{p\to\infty} f(u_p)\leq \inf_{u\in Q} f(u),$$

$$\limsup_{p\to\infty} g_i(u_p)\leq 0 \quad \text{for} \quad i=1, 2, \dots, q,$$

and

$$\lim_{p\to\infty} g_i(u_p)=0 \quad \text{for} \quad q+1, \dots, q.$$

Furthermore, if for $i=1, 2, \dots, q$ the constrains defined by g_i are correct [1, 5] and $q=\bar{q}=1$, or at least one of the sets

$$Q_i=\{u|g_i(u)\leq 0\}$$

is bounded then

$$\limsup_{p\to\infty} f(u_p)=\inf_{u\in Q} f(u)$$

and

$$\rho(u_p, Q)=\inf_{u\in Q}\|u_p-u\|\xrightarrow[p\to\infty]{}0$$

(i.e. $\{u_p\}$ is a generalized minimizing sequence for f).

Let us consider the application of Algorithm 3 and Theorems 2.3. and 2.4. to the optimal control problem defined by (2.2)-(2.5). In this case to compute values of f and ∇f in any point u it is necessary to solve two systems of differential equations. It can be done approximately by means of some numerical procedure, for example Runge-Kutta or interpolation-extrapolation method [8]. According to adaptive precision principle the accuracy to the numerical integration will be variable. We shall assume that the condition (4) of the Theorem 2.3. is satisfied (of course, in practice it is possible only to a certain limit).

For the optimal control problem (2.2)-(2.5) two following lemmas are valid:

LEMMA 1. If the functions $c(.,.,.)$ and $s(.,.,.)$ are continuously differentiable in x and u and they are piecewise continuous in t, furthermore, if the matrix-valued functions $\frac{\partial}{\partial x}c(.,.,.)$, $\frac{\partial}{\partial u}c(.,.,.)$, $\frac{\partial}{\partial x}s(.,.,.)$ and $\frac{\partial}{\partial u}s(.,.,.)$ are Lipschitz continuous in x and u, piecewise continuous in t and bounded for all $u \in \bigcup_{p=1}^{\infty} W_p(u_{p-1})$, all $x \in \{x(t, u) \in C(0, T)/u \in \bigcup_{p=1}^{\infty} W_p\}$ and almost all $t \in [0, T]$, then the assumption (3) of the Theorem 2.3 is satisfied.

LEMMA 2. If for $i=1, 2, \ldots, r$ functions $a_i(t)$ and $b_i(t)$ are finite, and for some real numbers m_1, m_2

$$\int_0^T c(x(t), u(t), t)\, dt \geqslant -m_1 - m_2\|u\|_2,$$

then assumptions (1) and (2) of the Theorem 2.3 are satisfied.

In the next theorem we shall formulate sufficient conditions for $\{u_p\}$ to be a generalized minimizing sequence for the functional (2.3) over the set (2.5).

THEOREM 2.5. Let us assume that the following conditions are satisfied:

(1) $c(.,.,.)$ is continuously differentiable in x and u, piecewise continuous in t and convex in (x, u), $\frac{\partial}{\partial u}c(.,.,.)$ and $\frac{\partial}{\partial x}c(.,.,.)$ are bounded, Lipschitz continuous in x and u and piecewise continuous in t.

(2) $s(x, u, t) = A(t)x + B(t)u$, where the matrix-valued functions $A(t)$ and $B(t)$ are continuous for $t \in [0, T]$.

(3) $a_i(t)$ and $b_i(t)$ are finite for $i=1, 2, \ldots, r$ and $t \in [0, T]$.

(4) $\int_0^T c(x(t), u(t), t)\, dt \geqslant -m_1 - m_2\|u\|_2$ for some numbers m_1 and m_2.

(5) The set (2.5) is non empty.

(6) The system $\frac{dx}{dt} = Ax + Bu$ is nonsingular (i.e. if $\frac{dy}{dt} = -A^* y$ and $y \neq 0$ then $B^* y \neq 0$).

Then $\{u_p\}$ is a generalized minimizing sequence for the functional (2.3) over the set (2.5).

If we slightly modify the condition (2) of the above theorem, by assuming the strict convexity of $c(.,.,.)$ in (x, u), we shall obtain a stronger result. In this case the optimal control problem has a unique solution u^* and $\{u_p\}$ converge weakly to u^*. Furthermore, we can ensure the strong convergence of $\{u_p\}$ to u^* by assuming the uniform convexity of $c(.,.,.)$ in (x, u).

3. The planning model

Kendrick and Taylor [3] have defined a four sector, discrete planning mode. and identified its parameters for the particular case of the South Korean economyl Here the modified, continuous version of this model is considered. In computational experiments the parameters of the model have been specified in line with data concerning the Polish economy. The basic structure of the model, defined in the form of an optimal control problem, is to maximize a welfare functional over a certain period of time, subject to constraints in the form of distribution relations, production functions, capital stock accumulation equations, terminal capital stock values constraints and bounds on the investment levels.

Let n be a number of sectors. The production of each sector is described by the CES function:

$$V_i(k_i, l_i)=\gamma_i\, e^{z_i t}\,[\beta_i\, k_i^{-\rho_i}+(1-\beta_i)\, l_i^{-\rho_i}]^{-v_i/\rho_i} \qquad \text{for } i=1, 2, \ldots, n \tag{3.1}$$

where γ_i is the efficiency parameter, z_i is the rate of neutral technological changes, β_i is the distribution parameter, $\rho_i=1/\sigma_i-1$ (σ_i is the elasticity of substitution of labour for capital), v_i is the degree of returns to scale, and k_i and l_i are the capital and labour inputs. The variables $l_i(t)$ and $k_i(t)$ for $i=1, 2, \ldots, n$ are considered as control and state variables respectively. The state equations can be in the form:

$$\dot{k}_i=\delta_i(t)-h_i\, k_i(t), \tag{3.2}$$

$$k_i(0)=k_{io} \qquad \text{for } i=1, 2, \ldots, n$$

where δ_i is the investment level in sector i and h_i is the depreciation coefficient in sector i (linear case).

Or in the form:

$$\dot{k}_i=\mu_i\, k_i\left[1-\left(1+\frac{\varepsilon_i}{\mu_i}\,\frac{\delta_i}{k_i}\right)^{-1/\varepsilon_i}\right]-h_i\, k_i \tag{3.3}$$

$$k_i(0)=k_{io} \qquad \text{for } i=1, 2, \ldots, n$$

where μ_i and ε_i are some positive numbers (nonlinear case).

Now, let A be a Leontief matrix, B a capital coefficient matrix, $\delta=(\delta_1, \ldots, \delta_n)$, $V=(V_1, \ldots, V_n)$ and $c=(c_1, \ldots, c_n)$, where for $i=1, 2, \ldots, n$ c_i is the consumption level of goods produced by sector i. The consumption vector is defined by the distribution relation:

$$c=(I-A)V-B\delta. \tag{3.4}$$

We seek to maximize the welfare functional

$$f(\delta, l)=\int_0^T e^{-\lambda t}\left(\sum_{i=1}^{n} a_i\, c_i^{b_i}(t)\right) dt \tag{3.5}$$

where T is a planning horizon, $a_i \geqslant 0$, $0 \leqslant b_i \leqslant 1$ and λ is consumption discount rate, over the set Q defined as:

$$Q = \{(\delta, l) \in L_2^{2n}(0, T) | 0 \leqslant \delta \leqslant M, \; l \leqslant 0, \; \sum_{i=1}^{n} l_i(t) = L(t) \text{ and } k(T) = k^T\}. \quad (3.6)$$

Our optimal control problem is now completely stated. Using the Theorem 2.4 together with Lemmas 1 and 2 and Theorem 2.5 it is not difficult to obtain the following results:

THEOREM 3.1. If the parameters v_i in CES production functions (3.1) are less or equal zero and functionals $f_p(\delta, l)$ are defined according to (2.6), (2.7), (2.8) and (3.6) then the algorithm 3 applied to the optimal control problem defined by relations (3.1) (3.2) or (3.3), (3.4), (3.5) and (3.6) generates a sequence $\{\delta^p, l^p\}$ and

$$\|\nabla f_p(\delta^p, l^p)\| \xrightarrow[p \to \infty]{} 0.$$

THEOREM 3.2. If the assumptions of the Theorem 3.1 are satisfied then Algorithm 3 applied to the optimal control problem defined by relations (3.1), (3.2), (3.4), (3.5) and (3.6) generates a generalized maximizing sequence for the functional (3.5) over the set (3.6). Furthermore, this sequence converges weakly to the unique solution of the problem.

For computational experiments we have identified the parameters of the CES production functions for five sectors of the Polish economy 1951—1972: (1) industry, (2) construction, (3) agriculture and forestry, (4) transportation and communication, and (5) commerce. It is worth to say that for all sectors v_i have turned out to be less than 1. The matrices A and B, the coefficients h_i, the function $L(t)$ and the values of k_{i0} were evaluated on the base of data published in the year books of the Polish economy. The remaining parameters of the model have been selected arbitrarily. The computations have been carried out for models with various number of sectors and for the planning horizon T from four to twenty years. The suboptimal solutions have been obtained after a reasonable amount of the computer time, in spite of the rather small and not very up to date computer we have used. It appears that the computation time increases the most rapidly with the length of planning horizon.

Some results have been also obtained for models including the foreign trade.

4. Conclusion

We have presented here the numerical method of the dynamic optimization and stated its properties in the case of the general functional defined over a Banach space and in the particular case of the optimal control problem. The method has been tested on a rather important practical problem taken from the theory of economic growth. The results obtained demonstrate that the method is of a considerable practical interest. For example, it can be useful as a tool for the analyses of multisectoral nonlinear planning models of the national economy. It seems that further experimentation along these lines can be quite interesting.

Appendix

Our purpose here is to outline the proofs of the principal theorems given in the paper. The more detailed proofs of these theorems together with the proofs of Lemmas 1 and 2 can be found in [9].

Proof of the Theorem 2.1.

It is easy to show that from the assumption (5) we have

$$\frac{\delta_j}{f(u_n)-f(u_{n+1})} \xrightarrow[j\to\infty]{} 0 \tag{A.1}$$

and

$$\frac{\delta_j}{\|\nabla f(u_n)\|} \xrightarrow[j\to\infty]{} 0 \tag{A.2}$$

(Notice, that n and j are interdependent).

since

$$\frac{f(u_n)-f(u_n+\lambda p_n)}{\langle -\nabla f(u_n), p_n\rangle} \xrightarrow[\lambda\to\infty]{} 1, \tag{A.3}$$

(A.1) and (A.2) yield

$$\frac{f_j(u_n)-f_j(u_n+\lambda p_n)}{\langle -h_j(u_n), p_n\rangle} \xrightarrow[\lambda\to 0]{} 1, \tag{A.4}$$

and we see that the Algorithm 1 is well defined.

Now, we have that for all n

$$f_j(u_n)-f_j(u_{n+1})>0. \tag{A.5}$$

Because of the assumptions (1) and (5)

$$j \xrightarrow[n\to\infty]{} \infty, \tag{A.6}$$

and one can show, that (A.5) together with (A.6) yields

$$f(u_n)-f(u_{n+1})\geqslant 0 \tag{A.7}$$

for all n greater than some $\bar{n}$.

Because of the assumption (3) to complete the proof it is enough to show, that

$$\left\langle -\nabla f(u_n), \frac{p_n}{\|p_n\|}\right\rangle \xrightarrow[n\to\infty]{} 0. \tag{A.8}$$

(From this point the proof is similar to the proof of Theorem 4.6.2 in [1]).

If the inequality from the step 5 of the algorithm is satisfied for $\lambda=1$ then

$$f(u_n)-f(u_{n+1})\geqslant\frac{1}{2}\left[\left\langle -\nabla f(u_n), \frac{p_n}{\|p_n\|}\right\rangle \|p_n\|\right]-\left(2+\frac{1}{2}M\right)\delta_j. \tag{A.9}$$

If it is not the case there exists an integer k such that for

$$u_{n+1}=u_n+\beta^k p_n$$

and

$$u_n' = u_n + \beta^{k-1} p_n$$

we have

$$\|p_n\| \left(f_j(u_n) - f_j(u_{n+1})\right) \geqslant \|u_{n+1} - u_n\| \frac{\langle -h_j(u_n), p_n \rangle}{2} \tag{A.10}$$

and

$$\|p_n\| \left(f_j(u_n) - f_j(u_n')\right) < \|u_n' - u_n\| \frac{\langle -h_j(u_n), p_n \rangle}{2}. \tag{A.11}$$

By transforming these inequalities and using assumptions (2), (4) and (5) we finally obtain that

$$\frac{\beta}{4L} \left\langle -\nabla f(u_n), \frac{p_n}{\|p_n\|} \right\rangle^2 \leqslant f(u_n) - f(u_{n+1}) + \mathrm{const}\, \delta_j. \tag{A.12}$$

Because of (A.7) together with the assumptions (1), and the assumption (5) (A.12) yields (A.8) and the proof of the theorem is complete.

Proof of the Theorem 2.2.

Let $\{x_k\}$ be a sequence of numbers generated by the algorithm 1 applied to the minimization of the function $\Theta(x)$ with $p_k = -\Theta_k'(x_k)$. Because of Theorem 2.1 it is clear that $\Theta'(x_k) \to 0$.

Since $j \xrightarrow[n\to\infty]{} \infty$ for every $\varepsilon_2 > 0$ there exists k such that

$$|\Theta_j'(x_k)| / \|h_j(z)\| \leqslant \varepsilon_2. \tag{A.13}$$

Therefore the Algorithm 2 is well defined.

The second part of the proof can be performed by contradiction. Let us assume that there exists a positive number ε such, that

$$\|h_j(u_n)\| \geqslant \varepsilon. \tag{A.14}$$

Assumptions (2) and (6) together with (A.14) yield

$$\|b_{nj} h^m s_{nj}\|_{R^m} \leqslant D \| h_j(u_{n+1})\| \tag{A.15}$$

for a number $D > 0$.

Let for every u_n α_n be a first number x_k which satisfies (A.14). It is possible to show that because of the assumptions (1) and (2), the inequality (A.15) yields

$$\alpha_n \geqslant \zeta > 0 \text{ for a number } \zeta. \tag{A.16}$$

Now, it is easy to proove that for ε_2 sufficiently small, n sufficiently large and some $0 < \bar{\zeta} \leqslant \zeta$ we have

$$f(u_{n+1}) \leqslant f(u_n + \bar{\zeta} h^m s_{nj}). \tag{A.17}$$

Because of (A.14) and (A.17) it is possible to show, using the Taylor's formula for first-order expansions, that there exists $a > 0$ such that

$$f(u_{n+1}) - f(u_n) \leqslant -a \tag{A.18}$$

for almost all n.

It means that $f(u_n) \xrightarrow[n\to\infty]{} -\infty$ and is in contradiction with the assumption (1).

The Theorem 2.3 follows directly from the Theorem 2.2.

Proof of the Theorem 2.4.

Since the convexity of the functionals f and g_i, and the boundness of the set $\bigcup_{n=1}^{\infty} W_n$ the inequality

$$\|\nabla f_p(u_p)\| \leqslant \alpha_p \text{ yields} \tag{A.19}$$

$$f_p(u_p) \leqslant \inf_{u\in E} f_p(u) + \text{const. } \alpha_p. \tag{A.20}$$

The sequence $\{u_p\}$ generated by the Algorithm 3 satisfies (A.19) for $p=1, 2, \ldots$, with $\alpha_p \to 0$.

Now, it is easy to show that all assumptions of Theorem 11.1 in [5] are satisfied. Therefore, the Theorem 2.4 is true.

The Theorem 2.5 follows from the Theorem 2.4 and Lemmas 1 and 2.

It is also obvious that Theorem 3.1 and 3.2 are valid as conclusions from the theory presented in Section 2.

References

1. Daniel J. W., The approximate minimization of functionals. New York 1971.
2. Fiacco A. V., McCormick G. P., Nonlinear programming: sequential unconstrained minimization technique. New York 1968.
3. Kendrick D., Taylor L., Numerical solution of nonlinear planning models. *Econometrica* **38**, 3 (1970).
4. Klessig R., Polak E., An adaptive precision gradient method for optimal control. *SIAM J. Contr.* **11**, 1 (1973).
5. Levitin E. S., Polak B. T., Constrained minimization methods. (In Russian). *Zh. Vych. Mat. Mat. Fiz.* **66**, 5 (1973).
6. Polak E., Computational methods in optimization. New York 1971.
7. Polak B. T., Conjugate gradient methods for unconstrained minimization. (In Russian). *Zh. Vych. Mat. Mat. Fiz.* **66**, 9 (1973).
8. Ralston A., A first course in numerical analysis. New York 1965.
9. Tołwiński B., Numerical methods for solving dynamic optimization problems and their application to the control of large-scale systems. (In Polish). Ph. D. Th. Warszawa 1974.

Zastosowanie zasady adaptacyjnej precyzji oraz gradientu sprzężonego do analizy nieliniowego modelu planowania gospodarki narodowej

Przedstawiono algorytm rozwiązywania zadań optymalizacji dynamicznej z ograniczeniami i przedyskutowano możliwość jego zastosowania do analizy nieliniowych modeli planistycznych, które mogłyby być wykorzystane przy formułowaniu programów rozwoju gospodarki narodowej. Definicja algorytmu wykorzystuje znane pojęcia funkcji kary, gradientu sprzężonego oraz zasady adaptacyjnej precyzji obliczeń. Szczególny nacisk położono na określenie warunków zbieżności algorytmu, zarówno w przypadku ogólnym jak i przypadku modelu planistycznego. Odpowiednie twierdzenia dotyczące tego zagadnienia zostały sformułowane i udowodnione.

Применение принципа адаптивной точности и сопряженного градиента к анализу нелинейной модели планирования народного хозяйства

В статье представлен алгоритм решения задач динамической оптимизации с ограничениями и рассмотрена возможность его применения к анализу нелинейных моделей планирования, которые могут быть использованы при формулировке программ развития народного хозяйства. При определении алгоритма используются известные понятия функции штрафа, сопряженного градиента и принципа адаптивной точности вычислений. Особое внимание уделено определению условий сходимости алгоритма как в общем случае, так и в слулае модели планирования. Даны определения и доказательства соответствующих теорем, касающихся этого вопроса.

Wskazówki dla Autorów

W wydawnictwie "Control and Cybernetics" drukuje się prace oryginalne nie publikowane w innych czasopismach. Zalecane jest nadsyłanie artykułów w języku angielskim. W przypadku nadesłania artykułu w języku polskim, Redakcja może zalecić przetłumaczenie na język angielski. Objętość artykułu nie powinna przekraczać 1 arkusza wydawniczego, czyli ok. 20 stron maszynopisu formatu A4 z zachowaniem interlinii i marginesu szerokości 5 cm z lewej strony. Prace należy składać w 2 egzemplarzach. Układ pracy i forma powinny być dostosowane do niżej podanych wskazówek.

1. W nagłówku należy podać tytuł pracy, następnie imię (imiona) i nazwisko (nazwiska) autora (autorów) w porządku alfabetycznym oraz nazwę reprezentowanej instytucji i nazwę miasta. Po tytule należy umieścić krótkie streszczenie pracy (do 15 wierszy maszynopisu).

2. Materiał ilustracyjny powinien być dołączony na oddzielnych stronach. Podpisy pod rysunki należy podać oddzielnie.

3. Wzory i symbole powinny być wpisane na maszynie bardzo starannie.

Szczególną uwagę należy zwrócić na wyraźne zróżnicowanie małych i dużych liter. Litery greckie powinny być objaśnione na marginesie. Szczególnie dokładnie powinny być pisane indeksy (wskaźniki) i oznaczenia potęgowe. Należy stosować nawiasy okrągłe.

4. Spis literatury powinien być podany na końcu artykułu. Numery pozycji literatury w tekście zaopatruje się w nawiasy kwadratowe. Pozycje literatury powinny zawierać nazwisko autora (autorów) i pierwsze litery imion oraz dokładny tytuł pracy (w języku oryginału), a ponadto:

a) przy wadawnictwach zwartych (książki) — miejsce i rok wydania oraz wydawcę;

b) przy artykułach z czasopism: nazwę czasopisma, numer tomu, rok wydania i numer bieżący.

Pozycje literatury radzieckiej należy pisać alfabetem oryginalnym, czyli tzw. grażdanką.

Recomendations for the Authors

Control and Cybernetics publishes original papers which have not previously appeared in other journals. The publications of the papers in English is recommended. No paper should exceed in length 20 type written pages (210 × 297 mm) with lines spaced and a 50 mm margin on the left-hand side. Papers should be submitted in duplicate. The plan and form of the paper should be as follows:

1. The heading should include the title, the full names and surnames of the authors in alphabetic order, the name of the institution he represents and the name of the city or town. This heading should be followed by a brief summary (about 15 typewritten lines).

2. Figures, photographs tables, diagrams should be enclosed to the manuscript. The texts related to the figures should be typed on a separate page.

3. Of possible all mathematical expressions should be typewritten. Particular attention should be paid to differentiation between capital and small letters. Greek letters should as a rule be defined. Indices and exponents should be written with particular care. Round brackets should not be replaced by an inclined fraction line.

4. References should be put on the separate page. Numbers in the text identified by references should be enclosed in brackets. This should contain the surname and the initials of Christian names, of the author (or authors), the complete title of the work (in the original language) and, in addition:

a) for books — the place and the year of publication and the publisher's name;

b) for journals — the name of the journal, the number of the volume, the year of the publication, and the ordinal number.