Control
and Cybernetics
VOL. 2 (1973) No. 3/4

Methods of mathematical programming in Hilbert space*

by

ANDRZEJ WIERZBICKI
Technical University of Warsaw
Institute of Automatics

The paper presents a unified approach to several computational techniques of dynamic optimization, based on the formulation of a basic mathematical programming problem in Hilbert space. Several known results and methods are reviewed in this unifying frame. Some new results concerning the conjugate direction methods and the variable operator methods are presented. A few applicational examples and the trends of future research are sketched in conclusions.

Introduction

Mathematical programming is a broad area of research concerned with problems of analytical and computational methods of optimalization. It includes such classic topics as linear, quadratic, nonlinear and dynamic programming, other branches of operations research, as well as more modern topics, i.e. computational approaches to control theory, both for deterministic and stochastic problems. All these problems can be put into a unifying frame when choosing a sufficiently general abstract space with appropriate mathematical structure.

Dynamic optimization problems have been treated in a quite abstract manner in Banach or even locally convex topological spaces. However, for computational purposes more mathematical structure is necessary; for example, the notions of orthogonality of scalar product are of major importance. Therefore, the chosen space can be hardly more general than the Hilbert space. On the other hand, the Hilbert space is general enough to include most of important mathematical programming problems.

---

* Presented at the Polish-Italian Meeting on "Modern applications of mathematical systems and control theory, in particular to economic and production systems", Cracow, Poland, Sept. 14–20, 1972.

1. A basic mathematical programming problem in Hilbert space

One of the simplest and most basic mathematical programming problems in Hilbert space can be stated as follows. Given two Hilbert spaces H_u and H_x, a functional Q: $H_x \times H_u \to R^1$ and an operator P: $H_x \times H_u \to H_x$, minimize $Q(x, u)$ subject to the constraint $P(x, u) = 0$—supposed the set $P = \{(x, u) \in H_x \times H_u$: $P(x, u) = 0\}$ is not empty. In many applications we can assume Q and P be twice Frechet differentiable in $H_x \times H_u$. To simplify the problem, assume also the Frechet derivative $P_x(x, u): H_x \to H_x$ be bijective for all x, u satisfying $P(x, u) = 0$. Hence there exist $P_x^{-1}(x, u)$ and, under some additional assumptions—see e.g. [4]—a transformation $S: H_u \to H_x$ implied by $P(x, u) = 0$, $x = S(u)$.

In both analytical and computational optimization we are interested in determining the gradient and the Hessian operator of the composed functional $J: H_u \to R^1$

$$J(u) = Q(S(u), u). \tag{1}$$

The transformation $S(u)$ can be shown to be also twice Frechet differentiable, hence we could easily compute the gradient and the Hessian operator in a closed form. However, when the spaces H_x and H_u are specified, it is usually too cumbersome to apply those closed forms. Most useful instead is another approach, based on the variational expansion of J

$$J(u+\delta u) = J(u)+\langle b(u), \delta u\rangle +0.5\langle \delta u, (u)\,\delta u\rangle + Ao(\|\delta u\|^2), \tag{2}$$

and the determination of the gradient $b(u) \in H_u$ and the Hessian operator $A(u): H_u \to H_u$ with help of the derivatives of an auxiliary Lagrange functional.

The assumption that $P_x^{-1}(x, u)$ exists may be called normality assumption since it corresponds to various normality assumptions in optimization problems. Under this assumption we can use the normal form of the Lagrange functional L: $H_x \times H_x \times H_u \to R^1$

$$L(\eta, x, u) = Q(x, u) + \langle \eta, P(x, u)\rangle. \tag{3}$$

We shall denote the derivatives of L by L_η, L_x, L_u, $L_{\eta x}$, $L_{\eta u}$, L_{xx} etc. without indicating the dependence on (η, x, u). Clearly, $L_\eta = P(x, u)$ and $L_{\eta\eta} = 0$; moreover, $L_{\eta x} = P_x(x, u)$ and is inveritble.

Denote $S(u+\delta u) = x+\delta x$. We have

$$x+\delta x = S(u)+Su(u)\,\delta u + o(\|\delta u\|); \quad \delta x = \delta x_1 + \delta x_2;$$
$$\delta x_1 = S_u(u)\,\delta u = -P_x^{-1}(x, u)P_u(x, u)\,\delta u = -L_{\eta x}^{-1} L_{\eta u}\delta u; \tag{4}$$
$$\delta x_2 = o(\|\delta u\|).$$

Since $P(x+\delta x, u+\delta u) = 0$, we have obviously

$$J(u+\delta u) = L(\eta, x+\delta x, u+\delta u).$$

Therefore

$$J(u+\delta u) = J(u) + \langle L_x, \delta x\rangle + \langle L_u, \delta u\rangle + o(\|\delta u\|). \tag{5}$$

The Lagrange multiplier $\eta \in H_x$ was not specified as yet. Choose η to obtain $L_x = 0$

$$\eta = -P_x^{*-1}(x, u)\, Q_x(x, u) \tag{6}$$

thus suppressing the dependence of $J(u+\delta u)$ on δx in the first-order approximation. It should be stressed that this Lagrange multiplier is slightly more general than the Lagrange multiplier related to the necessary conditions of optimality, since we have not assumed u to be optimal.

Having chosen η, we obtain the gradient

$$b(u) = J_p(u) = L_u(\eta, x, u) = Q_u(x, u) - \langle, P_u(x, u)\rangle P_x^{-1}(x, u) Q_x(x, u). \tag{7}$$

Assuming $\hat{u}$ to be optimal, we obtain the known necessary condition of optimality

$$b(\hat{u}) = L_u(\hat{\eta}, \hat{x}, \hat{u}) = 0, \tag{8}$$

where $\hat{\eta}$, $\hat{x}$ correspond to the optimal $\hat{u}$.

To investigate the second-order approximation, we introduce a variation $\delta\eta$ of η and observe that $J(u+\delta u) = L(\eta+\delta\eta, x+\delta x, u+\delta u)$. Hence:

$$\begin{aligned} J(u+\delta u) &= J(u) + \langle L_u, \delta u\rangle + \\ &+0.5(\langle \delta x, L_{xx}\delta x\rangle + \langle \delta u, L_{uu}\delta u\rangle + 2\langle \delta\eta, L_{\eta x}\delta x\rangle + 2\langle \delta\eta, L_{\eta u}\delta u\rangle + \\ &+2\langle \delta x, L_{xu}\delta u\rangle) + o(||\delta u||^2) \end{aligned} \tag{9}$$

since $L_x = 0$, $L_\eta = 0$, $L_{\eta\eta} = 0$. But $\delta x = \delta x_1 + o(||\delta u||)$ and $L_{\eta x}\delta x_1 + L_{\eta u}\delta u = 0$ by (4). Therefore, by rearranging the second-order terms

$$\begin{aligned} J(u+\delta u) &= J(u) + \langle L_u, \delta u\rangle + \\ &+0.5\langle \delta u, L_{u\eta}\delta\eta + L_{ux}\delta x_1 + L_{uu}\delta u\rangle + \\ &+0.5\langle \delta x_1, L_{x\eta}\delta\eta + L_{xx}\delta x_1 + L_{xu}\delta u\rangle + o(||\delta u||^2). \end{aligned} \tag{10}$$

Now we can choose $\delta\eta$ to satisfy

$$L_{x\eta}\delta\eta + L_{xx}\delta x_1 + L_{xu}\delta u = 0. \tag{11}$$

Hence, both $\delta x_1 = -L_{\eta x}^{-1} L_{\eta u}\delta u$ and $\delta\eta$ are linear transformations of δu

$$\delta\eta = L_{x\eta}^{-1}(L_{xx}L_{\eta x}^{-1}L_{\eta u} - L_{xu})\,\delta u. \tag{12}$$

Therefore, the term $0.5\langle \delta u, L_{u\eta}\delta\eta + L_{ux}\delta x_1 + L_{uu}\delta u\rangle$ is quadratic in δu. Thus the Hessian operator can be expressed by

$$J_{uu}(u) = A(u) = L_{u\eta}L_{x\eta}^{-1}L_{xx}L_{\eta x}^{-1}L_{\eta u} - L_{u\eta}L_{x\eta}^{-1}L_{xu} - L_{ux}L_{\eta x}^{-1}L_{\eta u} + L_{uu}. \tag{13}$$

The sufficient condition of local optimality is that $A(\hat{u})$ be positive definite. This may be stated on an equivalent way: since $\delta x_1 = -L_{\eta x}^{-1}L_{\eta u}\delta u$, the bilinear form

$$\langle \delta x_1, \hat{L}_{xx}\delta x_1\rangle + 2\langle \delta x_1, \hat{L}_{xu}\delta u\rangle + \langle \delta u, \hat{L}_{uu}\delta u\rangle \tag{14}$$

be positive for all non-zero δx_1, δu approximating linearly the constraint $P(x, u) = 0$ at $\hat{x}$, $\hat{u}$.

However, most often it is more useful to define $A(u)$ by the set of equations

$$A(u)\,\delta u = L_{u\eta}\delta\eta + L_{ux}\delta x_1 + L_{uu}\delta u, \tag{15a}$$

$$\left.\begin{aligned} 0 &= L_{x\eta}\delta\eta + L_{xx}\delta x_1 + L_{xu}\delta u \\ 0 &= \qquad\qquad L_{\eta x}\delta x_1 + L_{\eta u}\delta u \end{aligned}\right\}, \tag{15b}$$

since it is simpler to solve them in specified spaces than to determine the Hessian by (13). The equations (15b) may be called basic variational equations; we can interpret them as the variations of the equations $L_x = 0$ and $L_\eta = 0$. The equations (15a), (15b) are particularly useful when inverting the Hessian $A(u)$. Recall that if the functional $J(u)$ is quadratic

$$J(u) = c + \langle b(0), u\rangle + 0.5\langle u, Au\rangle, \tag{16}$$

and the operator A is strictly positive, then the optimal control $\hat{u}$ can be determined in the following way. Given any initial u, the gradient $b(u)$ and the operator A, we compute $\hat{u} = u+d$ by determining the gradient $J_u(\hat{u})$

$$J_u(u+d) = b(u) + Ad = 0. \tag{17}$$

Hence we get the increment d, called the Newton's step

$$d = -A^{-1}b(u), \tag{18}$$

and add it to u to obtain $\hat{u}$. If the functional is not quadratic, the Newton's step does not guarantee optimality. But it can be either used as a good iterative increment of u—in the Newton's method of computational optimization—or as a good direction of search for directional minimum of $J(u+zd)$ with respect to $z \in R^1$—in the modified Newton's method. The problem of inverting the Hessian and determining the Newton's direction of search

$$d = -A^{-1}(u)b(u) \tag{19}$$

is one of the most fundamental computational problems.

We have then to solve the set of equations

$$\begin{aligned} -b(u) &= L_{u\eta}\delta\tilde{\eta} + L_{ux}\delta\tilde{x} + L_{uu}d, \\ 0 &= L_{x\eta}\delta\tilde{\eta} + L_{xx}\delta\tilde{x} + L_{xu}d, \\ 0 &= \qquad\quad L_{\eta x}\delta\tilde{x} + L_{\eta x}d, \end{aligned} \tag{20}$$

with respect to d. Assume L_{uu}^{-1} exists and is easy to compute, as it happens in most applications. Then

$$d \doteq -L_{uu}^{-1}(L_{u\eta}\delta\tilde{\eta} + L_{ux}\delta\tilde{x} + b(u)), \tag{21a}$$

and

$$\left.\begin{aligned} 0 &= (L_{x\eta} - L_{xu}L_{uu}^{-1}L_{u\eta})\,\delta\tilde{\eta} + (L_{xx} - L_{xu}L_{uu}^{-1}L_{ux})\,\delta\tilde{x} - L_{xu}L_{uu}^{-1}b(u) \\ 0 &= -L_{\eta u}L_{vu}^{-1}L_{u\eta}\delta\tilde{\eta} + (L_{\eta x} - L_{\eta u}L_{uu}^{-1}L_{ux})\,\delta\tilde{x} - L_{\eta u}L_{uu}^{-1}b(u) \end{aligned}\right\} \tag{21b}$$

or, after obvious notational simplification

$$\begin{aligned} A_1\delta\tilde{x} + A_2\delta\tilde{\eta} &= \beta_1, \\ A_3\delta\tilde{x} + A_1^*\delta\tilde{\eta} &= \beta_2, \end{aligned} \tag{22}$$

where A_1^* is adjoint to A_1. The equations (21b) or (22) may be called canonical variational equations. They have the formal solution

$$\begin{aligned} \delta\tilde{x} &= (A_3 - A_1^* A_2^{-1}A_1)^{-1}(A_2^{-1}\beta_1 - \beta_2), \\ \delta\tilde{\eta} &= (A_2 - A_1 A_3^{-1}A_1^*)^{-1}(A_3^{-1}\beta_1 - \beta_2), \end{aligned} \tag{23}$$

which results in a formal expression for d; both are of little computational use. In more specified problems, the most important computational problem is how to solve the canonical equations effectively and thus to invert the Hessian, or how to omit the inversion either by approximating $A^{-1}(u)$ or by applying a method which does not require at all the use of the Hessian.

2. Applications to optimal control and related problems

2.1. Nonlinear programming problem

The classical nonlinear programming problem with equality constraints can be stated in the following manner: minimize $f(x, u) \in R^1$ under the constrain $g(x, u) = 0 \in R^n$, where $x \in R^n$, $u \in R^m$. The classical decision variable is actually the pair $(x, u) \in R^{n+m}$, but we split it into two parts: the independent variable u and the variable x resulting from the constraints. Assuming f_x, f_u, f_{xx}, f_{ux}, f_{uu}, g_x, g_x^{-1}, g_u, g_{xx}, g_{xu}, g_{uu} exists (g_{xx}, g_{xu} and g_{uu} are notational abbreviations, since they correspond to "three-dimensional" matrices), we get

$$L(\eta, x, u) = f(x, u) + \eta^T g(x, u), \tag{24}$$

$$L_x = 0 \Leftrightarrow \eta = -g_x^{-1} f_x, \tag{25}$$

$$b(u) = f_u - g_u^T g_x^{-1} f_x, \tag{26}$$

$$A(u) = (g_x^{-1} g_u)^T (f_{xx} - (g_x^{-1} f_x)^T g_{xx}) g_x^{-1} g_u + f_{uu} - (g_x^{-1} f_x)^T g_{uu} + \\ - (g_x^{-1} g_u)^T (f_{xu} - (g_x^{-1} f_x)^T g_{xu}) - (f_{ux} - g_{ux} g_x^{-1} f_x) g_x^{-1} g_u. \tag{27}$$

The expression (26) is often used as so called Wolfe's reduced gradient. Nobody practically attempts to compute $A(u)$ except in simple cases when the constraints are linear and $g_{xx} = 0$, $g_{xu} = 0$, $g_{uu} = 0$.

2.2. Ordinary differential optimal control problems

Consider the functional

$$Q(x, u) = \int_{t_0}^{t_1} f_0\big(x(t), u(t), t\big) dt + h\big(x(t_1)\big) \tag{28}$$

called the performance functional of Bolza type, and the constraint in the form of an ordinary differential equation, called the process equation

$$P(x, u) = 0 \Leftrightarrow \dot{x}(t) = f\big(x(t), u(t), t\big); \quad x(t_0) = x_0. \tag{29}$$

The problem of minimizing the functional (28) subject to contraints (29) is usually called continuous-time optimal control problem.

The independent variable u, called the control function, is usually assumed to be a bounded measurable function of time $t \in [t_0, t_1]$ and $u(t) \in R^m$. Introducing the norm ess sup, we have $u \in L^\infty_{[t_0, t_1]}$ but $L^\infty_{[t_0, t_1]} \subset L^2_{[t_0, t_1]}$ for any finite t_0, t_1. The resulting variable x, called the state function, is usually assumed to be an absolutely continuous function of $t \in [t_0, t_1]$ and $x(t) \in R^n$. If f is differentiable (hence continuous), x is absolutely continuous with square integrable derivative; we shall denote such a space by $W^2_{1[t_0, t_1]}$. It is a Sobolev and a Hilbert space. Therefore, we can assume $u \in L^2_{[t_0, t_1]}$, $x \in W^2_{1[t_0, t_1]}$ and perform the optimization in Hilbert space. As the result we could obtain square integrable u which would not be bounded, but such degenerate cases are easy to identify.

Assume the functions f, f_0, h are twice differentiable with respect to $x(t)$, $u(t)$, and there exists a bounded solution of (29). Define $P(x, u) = \big(\dot{x} - f(x, u, t), x(t_0) - x_0\big)$; hence P: $W^2_1 \times L^2 \to L^2 \times R^n$, whereas $L^2 \times R^n$ is isomorphic to W^2_1. We have $P_x(x, u)\,\delta x = \big(\delta\dot{x} - f_x(x, u, t)\,\delta x, \delta x(t_0)\big)$. Since for every square integrable φ and every $\delta x(t_0) \in R^n$ the equation $\delta\dot{x} - f_x(x, u, t)\,\delta x = \varphi$ has a unique solution $\delta x \in W^2_{1[t_0, t_1]}$, hence the transformation $P_x(x, u)\colon W^2_1 \to L^2 \times R^n$ has an inverse and the problem is normal.

It can be proven that the Lagrange multiplier η for the problem (28), (29) is an absolutely continuous function of time, $\eta(t) \in R^n$, $\eta \in W^2_{1[t_0, t_1]}$. Hence the Lagrange functional can be expressed in the form

$$L(\eta, x, u) = h\big(x(t_1)\big) + \int_{t_0}^{t_1} \big(\eta^T(t)\dot{x}(t) - H(\eta(t), x(t), u(t)\big)dt =$$

$$= h\big(x(t_1)\big) + \eta^T(t_1)x(t_1) - \eta^T(t_0)x(t_0) - \int_{t_0}^{t_1} \big(\dot{\eta}^T(t)x(t) + H\big(\eta(t), x(t), u(t), t\big)\big)dt \tag{30}$$

where the function H, called the Hamiltonian function

$$H\big(\eta(t), x(t), u(t)\big) \stackrel{\text{df}}{=} -f_0\big(x(t), u(t), t\big) + \eta^T(t) f\big(x(t), u(t), t\big) \tag{31}$$

is twice differentiable; we shall denote its derivatives by H_η, H_x, H_u, $H_{\eta x}$, $H_{\eta u}$, H_{xx}, H_{xu}, H_{uu} without indicating the arguments in more complicated expressions. We have

$$L_x = 0 \Leftrightarrow \dot{\eta}(t) = -H_x\big(\eta(t), x(t), u(t), t\big); \quad \eta(t_1) = -h_x\big(x(t_1)\big). \tag{32}$$

This equations is linear in η and usually called the adjoint equation. Recall that η is slighty more general than the adjoint variable in the known formulations of necessary conditions of optimality, since u was not yet assumed to be optimal. The gradient $b(u)$ has the simple form

$$b(u)(t) = -H_u\big(\eta(t), x(t), u(t), t\big). \tag{33}$$

Equations determining the Hessian $A(u)$ become

$$-A(u)\,\delta u = H_{u\eta}\delta\eta + H_{ux}\delta x_1 + H_{uu}\delta u \tag{34a}$$

$$\left.\begin{aligned} -\delta\dot{\eta} &= H_{x\eta}\delta\eta + H_{xx}\delta x_1 + H_{xu}\delta u; \quad & \delta\eta(t_1) &= -h_{xx}\delta x(t_1) \\ \delta\dot{x}_1 &= H_{\eta x}\delta x_1 + H^{\eta}_{u}\delta u; \quad & \delta x(t_0) &= 0 \end{aligned}\right\} \tag{34b}$$

Suppose the matrix of fundamental solutions of the last equation is $\Phi_x(t, t_0)$, and that of the last but one—$\Phi_\eta(t, t_0)$. We can solve the basic variational equations (34b) subsequently and get the closed form of the operator $A(u)$

$$A(u)\,\delta u(t) = -H_{uu}(t)\delta u(t) - H_{ux}(t)\int_{t_0}^{t} \Phi_x(t,\tau) H_{\eta u}(\tau)\,\delta u(\tau)\,d\tau +$$

$$+H_{u\eta}(t)\int_{t_0}^{t} \Phi_\eta(t,\tau)\left[H_{xx}(\tau)\int_{t_0}^{\tau}\Phi_x(\tau,\vartheta)H_{\eta u}(\vartheta)\delta u(\vartheta)\,d\vartheta + H_{xu}(\tau)\,\delta u(\tau)\right]d\tau +$$

$$-H_{u\eta}(t)\Phi_\eta(t_1,t)\int_{t_0}^{t}\Phi_\eta(t,\tau)\left[H_{xx}(\tau)\int_{t_0}^{\tau}\Phi_x(\tau,\vartheta)H_{\eta u}(\vartheta)\,\delta u(\vartheta)\,d\vartheta +\right.$$

$$\left. + H_{xu}(\tau)\,\delta u(\tau)\right]d\tau + H_{u\eta}(\tau)\Phi_\eta(t_1,t)h_{xx}(x(t_1))\int_{t_0}^{t_1}\Phi_x(t_1,\tau)H_{\eta u}(\tau)\,\delta u(\tau)\,d\tau \quad (35)$$

with obvious notational simplification. However, this form is of little computational use; for example, to determine $\Phi_x(t, t_0)$ and $\Phi_\eta(t, t_0)$ only, we must integrate $2n^2$ linear differential equations. Therefore, it is better to determine $A(u)\,\delta u$ simply by integrating once the basic variational equations (34b), beginning with the last, and by performing the algebraic operations indicated in (34a); we must integrate then only $2n$ linear differential equations.

To invert $A(u)$ and determine the Newton's direction $d = -A^{-1}(u)b(u)$, we assume H_{uu}^{-1} exists (almost everywhere on $[t_0, t_1]$) and get

$$d = -A^{-1}(u)\,b(u) = -H_{uu}^{-1}(H_u + H_{u\eta}\delta\tilde{\eta} + H_{ux}\delta\tilde{x}) \quad (36a)$$

and the canonical variational equations

$$\begin{aligned} \delta\dot{\tilde{x}} &= \mathscr{A}_1\delta\tilde{x} + \mathscr{A}_2\delta\tilde{\eta} + \beta_1; \quad & \delta\tilde{x}(t_0) &= 0 \\ \delta\dot{\tilde{\eta}} &= \mathscr{A}_3\delta\tilde{x} + \mathscr{A}_4\delta\tilde{\eta} + \beta_2; \quad & \delta\tilde{\eta}(t_1) &= -h_{xx}\delta\tilde{x}(t_1) \end{aligned} \quad (36b)$$

where

$$\mathscr{A}_1 = H_{\eta x} - H_{\eta u}H_{uu}^{-1}H_{ux}; \quad \mathscr{A}_2 = -H_{\eta u}H_{uu}^{-1}H_{u\eta}; \quad \mathscr{A}_3 = -H_{xx} + H_{xu}H_{uu}^{-1}H_{ux}$$

$$\mathscr{A}_4 = -\mathscr{A}_1^{T}; \quad \beta_1 = -H_{\eta u}H_{uu}^{-1}H_u; \quad \beta_2 = H_{xu}H_{uu}^{-1}H_u; \quad (37)$$

Suppose the matrices of fundamental solutions for the equations (36b) taken simultanously are $\tilde{\Phi}_{xx}(t, t_0)$, $\tilde{\Phi}_{x\eta}(t_1, t_0)$, $\tilde{\Phi}_{\eta x}(t_1, t_0)$, $\tilde{\Phi}_{\eta\eta}(t_1, t_0)$. With help of these matrices we could solve the equations, provided $\tilde{\Phi}_{\eta\eta}(t_1, t_0) + h_{xx}\tilde{\Phi}_{x\eta}(t_1, t_0)$ is nonsingular; hence we could get a closed form for the operator $A^{-1}(u)$ similar to (35). But again the closed forms is of little computational use since we had to integrate $2n^2$ linear differential equations in order to determine $\Phi_{\eta\eta}(t, t_0)$, $\Phi_{\eta x}(t, t_0)$ and additionaly $2n$ equations in order to actually solve them.

We cannot save much computational time when inverting $A(u)$. However, by setting

$$\delta\tilde{\eta}(t) = K(t)\,\delta\tilde{x}(t) + L(t), \quad (38)$$

where $K(t)\colon R^n \to R^n$ is a matrix and $L(t) \in R^n$, both K and L being absolutely continuous functions of time, we arrive at the Riccati-type equation for K

$$-\dot{K} = K\mathscr{A}_2K + K\mathscr{A}_1 - \mathscr{A}_4K - \mathscr{A}_3; \quad K(t_1) = -h_{xx} \quad (39)$$

and a linear one for L

$$-\dot{L} = (K\mathscr{A}_2 - \mathscr{A}_4)L + K\beta_1 - \beta_2; \quad L(t_1) = 0. \quad (40)$$

It remains then to solve the equation for $\delta\tilde{x}$

$$\delta\dot{\tilde{x}} = (\mathscr{A}_1+\mathscr{A}_2K)\,\delta x+\mathscr{A}_2L+\beta_1\,; \qquad \delta x(t_0) = 0. \tag{41}$$

Since K can be easily shown to be a symmetric matrix, we have to integrate $\frac{n(n+1)}{2}$ nonlinear differential equations as well as $2n$ linear ones and to perform some algebraic operations indicated in (38) and (36a) in order to determine $d = -A^{-1}(u)b(u)$. Nevertheless, the computational effort becomes a substantial one when n grows.

The sufficient conditions of invertibility of the Hessian are that H_{uu}^{-1} exists and $\Phi_{\eta\eta}(t_1, t_0)+h_{xx}\Phi_{x\eta}(t_1, t_0)$ is nonsingular or that the equation (39) has bounded solution on $[t_0, t_1]$; the last condition is slightly stronger, since it implies the first one but conversely.

2.3. The ordinary difference control problem

Consider the performance functional

$$Q(x, u) = \sum_{k=k_0}^{k_1-1} f_0(x_k, u_k, k)+h(x_{k_1}) \tag{42}$$

and the constraint in the form of a difference equations

$$P(x, u) = 0 \Leftrightarrow x_{k+1} = f(x_k, u_k, k); \qquad x_{k_0} = x_0. \tag{43}$$

The problem of minimizing (42) subject to (43) is usually called the discrete-time optimal control problem. The control sequence $u = \{u_k\}_{k_0}^{k_1-1}$, $u_k \in R^m$, and the state sequence $x = \{x_k\}_{k_0}^{k_1}$, $x_k \in R^n$, are assumed to be bounded, hence square-sumable for finite k_0, k_1. The transformation $P_x(x, u)$ can be proven nonsingular by the same type of argument we used for ordinary differential equations. We obtain also a form of the Lagrangian functional, which is analogous to (30); the integrating by parts in (30) is substituted by simple change of sumation limits. The Hamiltonian function has the form

$$H(\eta_k, x_k, u_k, k) = -f_0(x_k, u_k, k)+\eta_k^T f(x_k, u_k, k). \tag{44}$$

The condition $L_x = 0$ leads to the discrete adjoint equation

$$\eta_{k-1} = H_x(\eta_k, x_k, u_k, k); \qquad \eta_{k_1-1} = -h_x(x_{k_1}). \tag{45}$$

The form of this equation indicates strongly that a natural way of solving the adjoint equations is the reverse direction of time. The gradient $b(u)$ has the form

$$b(u) = -H_u(\eta_k, x_k, u_k, k). \tag{45}$$

The Hessian operator $A(u)$ is defined by a set of equations quite analogous to (34a), (34b). The equations which determine $d = -A^{-1}(u)b(u)$ are (36a) and, instead of (36b)

$$\begin{aligned} \delta\tilde{x}_{k+1} &= \mathscr{A}_1\delta\tilde{x}_k+\mathscr{A}_2\delta\tilde{\eta}_k+\beta_1\,; \qquad \delta\tilde{x}_{k_0} = 0\,; \\ -\delta\tilde{\eta}_{k-1} &= \mathscr{A}_3\delta\tilde{x}_k+\mathscr{A}_4\delta\eta_k+\beta_2\,; \qquad \delta\tilde{\eta}_{k_1-1} = -h_{xx}\delta x_{k_1}, \end{aligned} \tag{46}$$

where $\mathscr{A}_1, \mathscr{A}_2, \mathscr{A}_3, \mathscr{A}_4, \beta_1, \beta_2$ are of the form (37).

These discrete canonical equations are slighty more difficult to solve than the continuous ones, since they involve both $\delta\tilde{x}_{k+1}$ and $\delta\tilde{\eta}_{k-1}$; if $\mathscr{A}_1$ (and, therefore, $\mathscr{A}_4$) is singular, they cannot be solved in the same direction of time.

The Riccati-type setting becomes

$$\delta\tilde{\eta}_{k-1} = K_{k-1}\delta\tilde{x}_k + L_{k-1} \tag{47}$$

and results in

$$-K_{k-1} = \mathscr{A}_3 + \mathscr{A}_4 K_k (I - \mathscr{A}_2 K_k)^{-1}\mathscr{A}_1; \qquad K_{k_1-1} = -h_{xx} \tag{48}$$

$$-L_{k-1} = \mathscr{A}_4(K_k(I-\mathscr{A}_2K_k)^{-1}\mathscr{A}_2 + I)L_k + \mathscr{A}_4K_k(I-\mathscr{A}_2K_k)^{-1}\beta_1 + \beta_2;$$
$$L_{k_1-1} = 0 \tag{49}$$

$$\delta\tilde{x}_{k+1} = (I-\mathscr{A}_2K_k)^{-1}(\mathscr{A}_1\delta\tilde{x}_k + \mathscr{A}_2L_k + \beta_1); \qquad \delta x_{k_0} = 0. \tag{50}$$

These equations are also slightly more difficult to solve than their continuous counterparts, since we must compute and store at each step the matrix $P_k = (I-\mathscr{A}_2K_k)^{-1}$. The existence of this matrix is secured if h_{xx} is positive semi definite and H_{uu} negative definite- since then $\mathscr{A}_2$ is positive semidefinite, K_k—negative semi, definite by induction and $I-\mathscr{A}_2K_k$—positive definite.

2.4. Difference-differential optimal control problems

Consider the performance functional

$$Q(x,u) = \int_{t_0}^{t_1} f_0\big(x(t), x(t-T_1), u(t), u(t-T_2), t\big)dt + h\big(x(t_1)\big) \tag{51}$$

and the process equation

$$P(x,u) = 0 \Leftrightarrow \begin{cases} \dot{x}(t) = f\big(x(t), x(t-T_1), u(t), u(t-T_2), t\big) \\ x(t) = \varphi_1(t),\ t \in [t_0 - T_1, t_0] \\ u(t) = \varphi_2(t),\ t \in [t_0 - T_2, t_0] \end{cases} \tag{52}$$

where φ_1, φ_2 are given square-integrable functions. The problem of minimizing $Q(x,u)$ subject to $P(x,u)=0$ is often called the continuous-time delayed optimal control problem. The problem is a rather difficult one, but it can be shown that $P_x(x,u)$ is invertible. Therefore, the gradient $b(u)$ is relatively easy to compute. We have

$$L_x = 0 \Leftrightarrow \begin{cases} \eta(t_1) = -h_x\big(x(t_1)\big) \\ \dot{\eta}(t) = -H_{x_1}\big(\eta(t), x(t), x(t-T_1), u(t), u(t-T_2), t\big), t\in[t_1-T_1, t_1] \\ \dot{\eta}(t) = -H_{x_2}\big(\eta(t+T_1), x(t+T_1), x(t), u(t+T_1), u(t-T_2+ \\ \qquad\qquad + T_1), t+T_1\big) \\ \qquad - H_{x_1}\big(\eta(t_1), x(t); x(t-T_1), u(t), u(t-T_2), t\big); \\ \qquad\qquad t \in [t_0, t_1 - T_1] \end{cases} \tag{53}$$

where $H = -f_0+\eta^T f$ and H_{x_1} denotes the derivative with respect to the first argument of the functions f_0, f, whereas H_{x_2}—the derivative with respect to the second argument of these functions.

Again, the natural way of solving the adjoint equations (53) is to integrate them in the reverse direction of time. The gradient $b(u)$ has here the form

$$b(u) = \begin{cases} -H_{u_1}\big(\eta(t), x(t), x(t-T_1), u(t), u(t-T_2), t\big), & t \in [t_1-T_2 t_1] \\ -H_{u_2}\big(\eta(t+T_2), x(t+T_2), x(t-T_1+T_2), u(t+T_2), u(t), t+T_2\big) & \\ -H_{u_1}\big(\eta(t), x(t), x(t-T_1), u(t), u(t-T_2), t\big); & t \in [t_0, t_1-T_2] \end{cases} \tag{54}$$

where H_{u_1} denotes the derivative with respect to the third argument of the function f_0, f and H_{u_2}—the derivative with respect to the fourth argument.

The Hessian operator for this problem is much more difficult to determine and invert; for example, the Riccati-type equation in this case has the form of a set of partial-differential equations.

2.5. Partial-differential optimal control problem and related problems

If the operator equation $P(x, u) = 0$ is equivalent to a partial differential equation with respect to the state variable $x(t, z)$, where z is a space variable, we often call it the process equations with distributed state. The control variable can be both distributed, $u(t, z)$, or concentrated, $u(t)$. There is a large variety of partial-differential optimal control problems and, even if the gradient $b(u)$ can be determined in almost all cases, the ways of construction of the Hessian operator $A(u)$ were not sufficiently investigated yet. In a special case, when the control $u(t)$ is concentrated and the performance functional does not depend on the distributed state $x(t,z)$ but is determined according to a concentrated output variable $y(t) = R(x, u)(t)$, $R: H_x \times H_u \to H_y$, the partial-differential problem can be slightly simplified. We can eliminate the state $x(t, z)$ from the formulation of the problem and obtain

$$\bar{Q}(y, u) = Q\big(R(x, u), u\big); \quad \bar{P}(y, u) = 0 \leftrightarrow y = R(x, u), P(x, u) = 0. \tag{55}$$

Assume the process equation $\bar{P}(y, u) = 0$ is affine. It usually takes then the form

$$y(t) = \int_{t_0}^{t_1} K(t, \tau)u(\tau)\,d\tau + y_0(t) \tag{56}$$

where the kernel $K(t, \tau)$ of this integral operator can be interpreted as the matrix of output-functions corresponding to controls u described by distributions $\delta(t-\tau)$. The performance functional takes the form

$$Q(y, u) = \int_{t_0}^{t_1} \bar{f}_0\big(y(t), u(t), t\big)\,dt + \bar{h}\big(y(t_1)\big). \tag{57}$$

This integral-operator approach can simplify the difficulties related to the partial-differential optimal control problem and make the determination of the gradient and Hessian operator less cumbersome. Nevertheless, the inversion of the Hessian operator requires in all cases a comparatively large computational effort.

3. Computational methods in Hilbert space

3.1. General convergence properties

Many of the computational methods is Hilbert space can be put into a unifying frame by the definition of a direction of improvement. Given a set $C \subset H$ and a Frechet differentiable functional $J: C \to R^1$, with the gradient $b(u)$, a mapping $d: C \to H$ is called a direction of improvement mapping for the functional J if $\langle d(u), b(u)\rangle < 0$ and, given $\varepsilon > 0$, there exists a $\delta > 0$ such that $-\langle d(u), b(u)\rangle < \delta$ implies $\|b(u)\| < \varepsilon$ for all $u \in C$. A direction of improvement method consist of minimizing the functional J along subsequent direction of improvement, $J(u_{i+1}) = J(u_i + zd(u_i))$. The directional minimization need not to be precise. In fact, the following theorem holds.

THEOREM 1. (Goldstein). Suppose $C = \{u \in H: J(u) \leqslant J(u_0)\}$ and $b(u_0) \neq 0$. Suppose z_i is chosen according to the following rule. If $\langle d(u_i), b(u_i)\rangle = 0$, set $z_i = 0$; otherwise define the functional

$$\Delta(u, z) = \frac{J(u+zd(u)) - J(u)}{z\langle d(u), b(u)\rangle}. \tag{59}$$

If $\Delta(u_i, 1) \geqslant \sigma$, where $\sigma \in (0, 0.5)$ is an arbitrairly chosen constant, set $z_i = 1$. Otherwise chose z_i to satify $\sigma < \Delta(u_i, z_i) \leqslant 1-\sigma$. Set $u_{i+1} = u_i + z_i d(u_i)$.

(a) If C is bounded or J is bounded from below, then the sequence $\{b(u_i)\}$ converges to 0 while $\{J(u_i)\}$ converges downward to a limit J. If C is weakly compact, then every weak cluster point of $\{u_i\}$ is a zero of $b(u)$.

(b) If J is strictly convex and $A(u) > 0$ for all u in C, then $\{u_i\}$ converges to $\hat{u}$ such that $J(\hat{u}) < J(u)$ for all $u \neq \hat{u}$, $u \in C$.

3.2. Newton's method

The Newton's method is conceptionally the oldest one and in many simpler cases one of the most efficient. However, the computational effort in this method is large or even prohibitive in more complicated cases. The original method consists of the iteration

$$u_{i+1} = u_i - A^{-1}(u_i) b(u_i). \tag{60}$$

A modification of this method, called Picard's method

$$u_{i+1} = u_i - A^{-1}(u_0) b(u_i) \tag{61}$$

can save some computational effort, but has a smaller radius of convergence in highly nonlinear cases. An increased radius of convergence has the modified Newton's method

$$d_i = A^{-1}(u_i) b(u_i); \; J(u_i + \hat{z}_i d_i) = \min_{z \in R^1} J(u_i + zd_i); \quad u_{i+1} = u_i + \hat{z}_i d_i. \tag{62}$$

The directional minimization in (61) need not to be precise. In fact, the following theorem holds:

THEOREM 2 (Goldstein). Assume the Hessian $A(u)$ has the spectral bounds $m > 0$ and M

$$m\|\delta u\|^2 \leqslant \langle A(u)\,\delta u,\, \delta u\rangle \leqslant M\|\delta u\|^2 \tag{63}$$

in a convex set $C = \{u \in H: J(u) \leqslant J(u_0)\}$. Apply the Newton's direction $d(u_i) = -A(u_i)b(u_i)$ and choose z_i as in theorem 1, $u_{i+1} = u_i + z_i d_i$. Then

(a) there exists a number N such that $z_i = 1$ for $i > N$;

(b) there is a unique optimal $\hat{u}$ in C and $\{u_i\}$ converges to $\hat{u}$ faster than any geometric progression.

In other words, the convergence of the modified Newton's method (as also of the original Newton's method) is superlinear. The Picard's method has a linear convergence (at a rate of a geometric progression).

3.3. Methods of conjugate directions

A set of directions $\{d_1, \ldots, d_j, \ldots, d_k\}$, $d_j \in H$ is called A-conjugate or A-orthogonal if $\langle d_i, Ad_j\rangle = 0$ for $i \neq j$. Let the functional J be quadratic

$$J(u) = c + \langle b_0, u\rangle + 0.5\,\langle u, Au\rangle \tag{64}$$

where $b_0 = b(0)$ and A is assumed strictly positive. Consider $u = u_1 + \sum_{j=1}^{k} (\xi_j d_j)$, $\xi_j \in R^1$. Since

$$J(u) = J(u_1) + \sum_{j=1}^{k} (\xi_j \langle b(u_1), d_j\rangle + 0.5\,\xi_j^2 \langle d_j, Ad_j\rangle) \tag{65}$$

we can minimize J independently along each direction.

The set of conjugate directions is usually constructed by the following family of algorithms:

$$d_1 = -b(u_1);\; d_{i+1} = -b(u_{i+1}) + \beta_i d_i \tag{66}$$

where $u_{i+1} = u_i + \hat{z}_i d_i$ and $\hat{z}_i$ minimizes $J(u_i + zd_i)$. If d_i are constructed in such a manner, the following lemma holds.

LEMMA 3.

(a) If $d_1, \ldots, d_k$ are conjugate, then $b(u_{k+1})$ is orthogonal to all $b(u_1), \ldots, b(u_k)$.

(b) If $d_1, \ldots, d_k$ are conjugate and β_k is chosen to obtain the conjugacy of d_{k+1} to d_k, then d_{k+1} is also conjugate to all previous $d_1, \ldots, d_{k-1}$.

There are possible many choices of β_i in (66), all mutually equivalent in the case of a quadratic functional J. However, the methods of conjugate directions can be effectively applied also to nonquadratic functionals, and the choice of β_i influences the efficency of a computational algorithm. Some of the algorithms are

$$\beta_i = \frac{\langle b(u_{i+1}), Ad_i\rangle}{\langle d_i, Ad_i\rangle} \text{ (proposed by Pagurek and Woodside in } H) \tag{67}$$

$$\beta_i = \frac{\langle b(u_{i+1}) - b(u_i), b(u_{i+1})\rangle}{\langle b(u_i), b(u_i)\rangle} \text{ (proposed by Polak and Ribière in } R^n) \tag{68}$$

$$\beta_i = \frac{\langle b(u_{i+1}), b(u_{i+1})\rangle}{\langle b(u_i), b(u_i)\rangle} \quad \text{(proposed by Fletcher and Reeves in } R^n\text{, Lasdon and Mitter in } H) \tag{69}$$

The convergence of the conjugate direction methods follows from the Theorem 1. Stronger results were obtained only recently:

THEOREM 4 (Winnicki):

(a) Suppose the functional J is quadratic and the Hessian A is strictly positive with the spectral bounds m and M. Then the methods of conjugate directions converge linearly:

$$||u_{i+1}-u|| \leqslant \sqrt{\frac{J(u_1)-J(\hat{u})}{m}}\left(\frac{M-m}{M+m}\right)^i. \tag{70}$$

If, additionally, $A = A_1+A_2$ where A_1 is a compact operator and $A_2 = \lambda I$, $\lambda > 0$, then the convergence is superlinear.

(b) Suppose the functional J is not quadratic but convex and twice Frechet differentiable. Then the algorithms (67) and (68) result in at least linear convergence.

Fortuna [17] has recently generalized the above result by proving the superlinear convergence for the case when $A = A_1+A_2$, A_1 being compact and A_2^n being a linear combination of A_2^i, $i = 0, \ldots, n-1$. This property of the Hessian operator posses all dynamic optimization problems described by differential, difference and difference—differential equations, if $H_{uu} = -A_2$ is a constant matrix.

If should be pointed out that the results concerning convergence are valid under the assumption of a precise directional minimization. Nevertheless, the vast computational experience with conjugate directions methods in R^n is quite positive; there are less known, but also positive computational applications in infinite dimensional spaces.

3.4. Variable operator methods

Consider a quadratic functional J of the form (64) with a strictly positive Hessian A. Denote $s_i = u_{i+1}-u_i$, $y_i = b(u_{i+1})-b(u\,)$. The relations

$$y_i = As_i;\ s_i = A^{-1}y_i \tag{71}$$

can be used in order to approximate the operator A and its inverse. The variable operator methods—which are, in fact, a generalisation into R^n and infinite dimensional spaces of the secant method in R^1—construct a sequence of self-adjoint operators $V_{i+1} = V_i+\Delta V_i$ such that

$$s_j = V_{i+1}y_j,\ j = 1, \ldots, i. \tag{72}$$

It is easy to construct V_{i+1} such that

$$s_i = V_{i+1}y_i;\ \Delta V_i y_i = s_i - V_i y_i. \tag{73}$$

In fact, there is an entire subspace of linear operators $\Delta V_i: \mathscr{H} \to \mathscr{H}$ satisfying (73). However, we are interested in formulae satisfying (72) also for $j = 1, \ldots, i-1$;

we say that such operators have the property of hereditary approximation. If we construct a sequence of linearly independent s_j, then the sequence $\{y_j\}$ has also linearly independent elements; if (72) holds, then the operators V_{i+1} are equivalent to A^{-1} on expanding subspaces of H spanned by $\{y_j\}_1^i$. Hence the hereditary approximation implies the equivalence of V_{i+1} and A^{-1} on these expanding subspaces.

For optimization purposes a quasi-Newton's direction is determined

$$d_{i+1} = -V_{i+1} b(u_{i+1}) \tag{74}$$

and the step $s_{i+1} = u_{i+2} - u_{i+1}$ is usually determined by the minimization of J along this direction.

The operator ΔV_i is usually constructed with the help of outer products in H. Recall that an operator $B: H \to H$ denoted by $B = a\rangle\langle b$ is called the outer product of $a \in H$ and $b \in H$ if for all $u \in H$

$$Bu = a\rangle\langle bu \overset{\text{df}}{=} a\langle b, u\rangle. \tag{75}$$

There are several variable operator method in R^n (called usually variable metric methods). Two of them were generalised and applied in H

$$\Delta V_i = \frac{s_i\rangle\langle s_i}{\langle s_i, y_i\rangle} - \frac{V_i y_i\rangle\langle V_i y_i}{\langle V_i y_i, y_i\rangle} \tag{76}$$

by Davidon, later by Fletcher and Powell in R^n and by Horwitz and Sarachik in $\mathscr{H}$, and

$$\Delta V_i = \frac{s_i - y_i V_i\rangle\langle s_i - y_i V_i}{\langle s_i - y_i V_i, y_i\rangle} \tag{77}$$

by Davidon, Broyden or Wolfe in R^n and by the author in $\mathscr{H}$ [15]. The algorithm (76) has been commonly used in R^n through last five years and is considered one of the most effective computational methods. However, the hereditary approximation property (72) for this algorithm results from the conjugacy of the directions d_i generated by (74); therefore, the algorithm requires a rather precise directional minimization for a good approximation of A^{-1} by V_{i+1}. The convergence of this algorithm in the case of a quadratic functional is given by the theorem 4, part a, since the algorithm is equivalent in this case to conjugate direction methods. In case of a nonquadratic functional, the algorithm has a better computational efficency than the conjugate direction methods. The convergence of the algorithm results from a general convergence consideration, of Theorem 1. Stronger results in R^n were obtained by Powell, see [13], who has proven the superlinear convergence under reasonable assumptions; there are no stronger results in $\mathscr{H}$ available as yet.

The algorithm (77) was not much used in R^n until the thorough investigation of Murtagh and Sargent [7]. It has one basic advantage over the algorithm (76): the hereditary approximation property (72) holds even if the directional minimization is not precise and the directions (74) are not conjugate, as long as the expression (77) is well defined. It has some other drowbacks—see [7], [15]—but this one advantage is of particular importance in several special problems of optimization, like coordination of large-scale problems [15] etc., where a precise directional minimization is unreasonable.

The following theorem on convergence of the algorithm (77) was presented for R^n by Goldfarb and recently generalised into H:

THEOREM 5 (Winnicki). Suppose the functional J is twice Frechet differentiable and the Hessian $A(u)$ has for all $u \in H$ the spectral bounds m and M, where $0 < m \leqslant \leqslant M < \infty$. Suppose, moreover, that either $A(u+h) \geqslant A(u)$ $\big(A(u+h)-A(u)$ is positive definite$\big)$ for all $h \in \{h \in H: J(u+h) \leqslant J(u)\}$ and $A^{-1}(u_1) \leqslant V_1 \leqslant KI$, where $\frac{1}{m} \leqslant K < \infty$, or $A(u+h) \leqslant A(u)$ for all $h \in \{h \in H: J(u+h) < J(u)\}$ and $kI \leqslant V_1 \leqslant A^{-1}(u_1)$. If the algorithm (77) is applied to determine the direction (74) and the sequence $\{u_i\}$ is determined by directional minimization of J, then

$$||u_{i+1}-\hat{u}|| \leqslant \sqrt{\frac{2}{m}(J(u_1)-J(\hat{u}))}\left(\sqrt{1-\frac{m^2}{M}}\right)^i \tag{78}$$

and the algorithm is at least linearly convergent.

Recently, Ostryhanski [18] has shown that the results on superlinear convergence of variable metric algorithm for nonquadratic functions in $R^n_?$, obtained by Powell, can be easily extended to cover the algorithm (77).

3.5. Other metheds

There are several other methods of mathematical programming in Hilbert space, based either on the classical Ritz technique, or on minimization of a distance to a given set, on orthogonalization procedures, on generating a random set of directions, on contraction mapping algorithms for the approximation of A^{-1}, etc. Some of them are quite effective, but they are usually less general than the algorithms presented above.

4. Applications and conclusions

The Hilbert space approach to mathematical programming unifies computational algorithms for various optimal control problems. Moreover, it provides for effective algorithms of solving several complicated problems, when the determination and inversion of the Hessian operator is difficult, such as the difference-differential of partial differential optimal control problems, large scale coordination problems etc.

There are several application of this approach. Recently, in the Institute of Automatics of the Technical University of Warsaw, the following problems are being investigated in the unifying frame of this approach:

(a) the control of a natural gas supply system;

(b) the control of a steel furnace

(c) comparison of algorithms for difference-differential optimal control problems.

Several interesting areas of research are stil open. They include

(1) Problems related to the determination and inversion of the Hessian operator for partial differential optimal control problems etc.

(2) Investigation of properties of conjugate direction and variable operator methods under inadequate directional minimization.

(3) Comparative discussion of various computational methods of mathematical programming in Hilbert space, including those mentioned on 3.4.

References

1. FLETCHER R., A new approach to variable metric algorithms. *Computer J.* **13** (1970) 317–322.
2. FLETCHER R., POWELL M., A rapidy convergent desumt method for minimization. *Computer J.* **6** (1963) 163–168.
3. FLETCHER R., REEVES C., Function minimization by conjugate gradients. *Computer J.* **7** (1964) 149–154.
4. GOLDFARB D., Sufficient conditions for the convergence of a variable metric algorithm. In: Optimization, by R. Fletcher. New York 1969.
5. GOLDSTEIN A. A. Constructive real analysis. New York 1967.
6. HORWITZ L. B., SARACHIK P. E., Davidon's method in Hilbert space. *SIAM J. Appl. Math* **16** (1968) 676–695.
7. LASDON L. S., MITTER S. K., WARREN A. D., The conjugate gradient method for optimal control problems. *TEEE Trans. Antom. Contr.* **12** (1967) 132–138.
8. MURTAGH B. A., SARGENT R. W., A constrained minimization problem with quadratic convergence. In: Optimization, by R. Fletcher. New York 1969.
9. MURTAGH B. A., SARGENT R. W., Computational experience with quadratic convergent minimization methods. *Computer J.* **13** (1970) 185.
10. PAGUREK B., WOODSIDE C. M., The conjugate gradient method for optimal control problems with bounded control variables. *Automatica* **4** (1968) 337–349.
11. PEARSON J. D., Variable metric methods of minimization. *Computer J.* **12** (1969) 171–178.
12. PIERSON B. L., RAJTOLA S. G., Computational experience with Davidon method applied to optimal control problems. *IEEE Trans. Syst. Sci. Cyber.* (1970) 240–242.
13. POLAK E., Computational methods in optimization. New York 1971.
14. POWELL M. J., A survey of numerical method for unconstrained minimization. *SIAM Rev.* **12** (1970) 79–97.
15. WIERZBICKI A. P., Coordination and sensitivity analysis of a large scale system with performance interaction. Proc. V IFAC Congr. Paris 1972.
16. WINNICKI A., Methods of variable operator in Hilbert space and theis application to dynamic optimization problems. Master thesis. Technical University of Warsaw, Institute of Automatics. Warsaw 1972.
17. FORTUNA Z., On the superlinear convergence of conjugate gradient algorithms. *SIAM J. Contr.* (to appear).
18. OSTRYHANSKI R., Comparison of several gradient optimization techniques in R^n. Master thesis. Technical University of Warsaw, Institute of Automatics. Warsaw 1972.

Metody programowania matematycznego w przestrzeniach Hilberta

Przedstawiono jednolite podejście do pewnych metod obliczeniowych optymalizacji dynamicznej na podstawie sformułowania podstawowego zadania programowania matematycznego w przestrzeni Hilberta. W tym jednolitym ujęciu przedstawiono szereg znanych wyników i metod. Przedstawiono również pewne nowe wyniki dotyczące metod kierunków sprzężonych i metod zmiennego operatora. We wnioskach omówiono kilka przykładów zastosowań i naszkicowano kierunki dalszych badań.

Методы математического программирования в гильбертовых пространствах

Представлен единый подход к некоторым численным методам динамической оптимизации исходя из формулировки основной задачи математического программирования в гильбертовом пространстве. При этом подходе представлен ряд известных результатов и методов. Приведены также некоторые новые результаты касающиеся метода сопряженных направлений и метода переменного оператора. В части выводов рассмотрено несколько примеров применений и намечены направления дальнейших исследований.

Cena 50,— zł.

Warunki prenumeraty

Cena prenumeraty krajowej

rocznie zł 100.—
półrocznie zł 50.—

Instytucje państwowe, społeczne, zakłady pracy, szkoły itp. mogą zamawiać prenumeratę wyłącznie w miejscowych Oddziałach i Delegaturach RSW „Prasa—Książka—Ruch".

Prenumeratorzy indywidualni mogą opłacać w urzędach pocztowych i u listonoszy, lub dokonywać wpłat na konto PKO Nr 1-6-100020 RSW „Prasa—Książka—Ruch", Centrala Kolportażu Prasy i Wydawnictw, ul. Towarowa 28, 00-958 Warszawa (w terminie do 10 dnia miesiąca poprzedzającego okres prenumeraty).

Prenumeratę ze zleceniem wysyłki za granicę, która jest o 40% droższa od prenumeraty krajowej, przyjmuje RSW „Prasa—Książka—Ruch", Biuro Kolportażu Wydawnictw Zagranicznych, ul. Wronia 23, 00-840 Warszawa, konto PKO Nr 1-6-100024.

Bieżące i archiwalne numery można nabyć lub zamówić we Wzorcowni Wydawnictw Naukowych PAN—OSSOLINEUM—PWN, Pałac Kultury i Nauki (wysoki parter) 00-901 Warszawa oraz w księgarniach naukowych „Domu Książki".

Sprzedaż egzemplarzy zdezaktualizowanych, na uprzednie pisemne zamówienie prowadzi RSW „Prasa—Książka—Ruch", Centrala Kolportażu Prasy i Wydawnictw, ul. Towarowa 28, 00-958 Warszawa, skr. poczt. 12.

A subscription order stating the period of time, along with the subsyiber's name and address can be sent to your subsrription agent or directly to Foreign Trade Enterprise Ars Polona-Ruch, 00-068 Warszawa, 7 Krakowskie Przedmieście, P.O. Box 1001, POLAND Please send payments to the account of Ars Polona—Ruch in Bank Handlowy S.A., 7 Traugutt Street, 00-067 Warszawa, POLAND.

Indeks 35615