Control
and Cybernetics
VOL. 5 (1976) No. 3

Dynamic optimization and sensitivity analysis of multilevel control systems

by

ASEN L. DONTCHEV
Bulgarian Academy of Sciences, Sofia

GRAŻYNA PETRICZEK, PIOTR PŁONKA, ANDRZEJ URBANIK
Technical University of Warsaw
Institute of Automatic Control

The paper is concerned with the optimization technique and the sensitivity analysis of multilevel control systems described by linear differential equations with quadratic performance index. Two computational two-level dynamic optimization methods are presented. The second-order sensitivity analysis of such systems due to various inaccuracies is presented: inaccuracies in the state equation matrix, inaccurate coordination and small time delays, which appear in the interconnections between the subprocesses. The sensitivity analysis is applied to several one and two-level optimal control structures. A suboptimal control structure, based on a prediction of the interconnection variables is proposed and its sensitivity analysis is performed. An example illustrates the application of the presented methods.

1. Introduction

Sensitivity analysis is understood here as an analysis of deviation from optimal performance caused by various inaccuracies. The inaccuracies can be related to:

(i) inaccurate determination of a mathematical model;

(ii) intentional simplifications of complex models, where practical application is related to many computational and technical problems;

(iii) inaccuracies related to optimization algorithm.

The sensitivity analysis has been developed for many years because of its practical importance. A. P. Wierzbicki [9] introduced a new concept of sensitivity, by means of which "sensitivity paradox" [10] can be explained and real control systems can be examined. This concept was further developed in [2], [10]. In this paper, the sensitivity analysis is applied to two-level optimal control structures based on decomposition of a system and coordination of subprocesses. Some typical inaccuracies characteristic of multilevel systems are distinguished. The method presented in the paper does not require large computational effort and can become an important element of the systems control technique.

In Section II, theoretical foundations of two multilevel optimization methods, presented in [5], [6] are described. In further sections, general sensitivity analysis is discussed. In Section IV, sensitivity analysis of some one- and two-level optimal control structures is presented. Computational examples are given in Section V.

2. Multilevel dynamic optimization methods

Consider a system (process) given by a linear differential equation[1]):

$$\dot{x} = Ax + Bu;\ x(0) = x_0 \tag{1}$$

where $x(t) \in R^n$ denotes state vector, $u(t) \in R^r$ denotes control vector. Assume that the performance index (2) is to be minimized

$$J(x, u) = 0.5\, x'(T)\, Fx(T) + 0.5 \int_0^T \left(u'(t)\, Ru(t) + x'(t)\, Qx(t)\right) dt \tag{2}$$

where final time T is given, F, R are matrices block diagonal in form. If the dimension of the above problem is large computational effort can be diminished by decomposition methods. The problem can be decomposed according to the structure of naturally related subprocesses (the matrix A is then close to block diagonal form), or in any other arbitrary way.

Decomposition of optimization problems consists in separation of the global problem into N subproblems and a coordination problem, such that coordinated solutions of subproblems result in the optimal solution of the global problem — see [3]. In order to decompose the problem, interaction variables $v(t) \in R^m$ are introduced[2]). These variables define connections between artificially or naturally separated subproblems, given by:

$$\dot{x}_i = A_i\, x_i + B_i\, u_i + C_i\, v_i;\ x_i(0) = x_{i0},\ i = 1, \ldots, N \text{ [3])} \tag{3}$$

where $x_i(t) \in R^{n_i}$, $u_i(t) \in R^{m_i}$ etc.

Choice of the interaction variables is natural if matrix A is close to block diagonal form. Otherwise decomposition must be performed carefully so that the new problem is sufficiently regular (existence of Lagrange multipliers).

The connections between subsystems are defined by:

$$v = Mx \tag{4}$$

where M is $m \times n$ matrix composed of 0's and 1's.

---

[1]) The content of this paper is limited to linear problems although most of the results can be generalized.

[2]) Further on it is assumed that there are no interactions of the state and the control in performance index — see [1]. Assume that local indices include positively defined matrices Q_{2i} to make the problem convex.

[3]) Matrix B is assumed to be block diagonal in form, if it is not slight modification of given equations is needed.

Performance index is also decomposed into local performance indices:

$$J_i(x_i, u_i, v_i)=0.5\, x_i'(T)\, F_i\, x_i(T)+0.5\int_0^T \left(x_i'(t)\, Q_1\, x_i(t)+\right.$$

$$\left.+u_i'(t)\, R_i\, u_i(t)+v_i'(t)\, Q_{2i}\, v_i(t)\right) dt \qquad (5)$$

where $Q=Q_1+MQ_2\, M$, $Q_1=\mathrm{diag}\,\{Q_{1i}\}_{i=1}^N$

Assume, that control $u\in L^2([0, T], R^r)$, interaction variable $v\in L^2([0, T], R^m)$ and state $x\in W_1^2([0, T], R^n)$ — Sobolev space of absolutely continuous functions with integrable squared derivative with scalar product:

$$\langle x_1, x_2\rangle=x_1'(T)\, x_2(T)+\int_0^T \dot{x}_1'\, \dot{x}_2\, dt.$$

Adjoint variables ψ and λ, respectively to x and v, belong to the adjoint spaces $\psi\in L^2([0, T], R^n)\times R^n$, $\lambda\in L^2([0, T], R^m)$. Lagrange functional of the problem (1)—(2) has the form:

$$L(x, u, v, \psi, \lambda)=J(x, u, v)+\int_0^T \left(\psi'(\dot{x}-A_1\, x-Bu-Cv)+\lambda'(v-Mx)\right) dt \qquad (6)$$

where $A_1=A-CM$, $J(x, u, v)=\sum_{i=1}^N J_i(x_i, u_i, v_i)$.

Under the assumption of strict convexity of performance functional with respect to independent variables the following relation holds:

$$\min_{(x, u, v)\in\Omega} J(x, u, v)=\max_{\psi, \lambda}\min_{x, u, v} L(x, u, v, \psi, \lambda) \qquad (7)$$

where set Ω is defined by eq. (1) and (4). Hence the minimization of performance functional is equivalent to finding a saddle point of the Lagrange functional.

Necessary optimality conditions are obtained by comparing the derivatives of Lagrange functional to zero.

$$L_u=u'\, R-\psi'\, B=0, \qquad (8a)$$

$$L_v=v'\, Q_2-\psi'\, C+\lambda=0, \qquad (8b)$$

$$L_\lambda=Mx-v=0, \qquad (8c)$$

$$L_\psi=\dot{x}-A_1\, x-Bu-Cv=0. \qquad (8d)$$

Green formula is used for computing the adjoint equation.

$$L_x\, \delta x=x'(T)\, F\delta x(T)+\int_0^T \left(x'\, Q_1\, \delta x+\psi'(\delta\dot{x}-A_1\, \delta x)+\lambda'\, M\, \delta x\right) dt=$$

$$=\delta x'\left(\psi(T)+Fx(T)\right)+\int_0^T \left(\delta\dot{x}\,(\psi(t)-\int_t^T \left(A_1'\psi(\tau)-Q_1\, x(\tau)+M'\lambda(\tau)\right) d\tau+\right.$$

$$\left.-\psi(T)\right) dt.$$

A final form of adjoint equation is obtained[4])

$$\dot{\psi} = -A_1' \psi + Q_1 x - M' \lambda; \psi(T) = -Fx(T). \tag{8e}$$

Sufficient conditions of optimality can be analogously obtained by analysis of second derivatives of Lagrange functional.

THEOREM 1. If: (i) matrices R_i, Q_{2i}, $i=1, ..., N$ are positively defined, (ii) matrices F_i, Q_{1i}, $i=1, ..., N$ are semipositively defined, then eq. (8a)—(8e) define uniquely the optimal solution of the problem (1)—(2).

The proof of the above theorem can be found in [6]. Problem (1)—(2) can be solved by iteraction techniques. It can be also solved by solving the following matrix Riccati equation obtained by the Riccati substitution $\psi(t) = K(t) x(t) + L(t)$:

$$\dot{K} = -K' A - A' K - K' B R^{-1} B' K + Q; K(T) = -F. \tag{9}$$

In this case $L(t)$ equals zero. This method, however simple in computation, leads to large cost of computation when dimension of the problem increases — see Section V. That is the main reason for decomposition and two-level optimization of large scale systems.

The multilevel technique depends on independent optimization in local units and coordination of subsystems based on observation of local units in the master unit. The construction of the multilevel optimization methods relies on proper decomposition of the optimality conditions and their enforcing on particular levels. It seems that the natural way of decomposition depends on enforcing of (8a, b, d, e) in local units (optimization of subsystems) and (8c) in the master unit (coordination).

Conclusion. If the assumptions of the theorem 1 are satisfied, then for each λ there exists a unique solution of each subproblem (8a, b, d, e) $\hat{x}(\lambda)$, $\hat{u}(\lambda)$, $\hat{v}(\lambda)$, $\varphi(\lambda)$.

Moreover, the optimal value of λ can be computed by maximizing the Lagrange functional,

$$\max_{\lambda} L(\lambda) = \max_{\lambda} L(\hat{x}(\lambda), \hat{u}(\lambda), \hat{v}(\lambda), \psi(\lambda), \lambda) \tag{10}$$

whereas the Frechet derivative of Lagrange functional has the form:.

$$L_\lambda = \hat{v}(\lambda) - M\hat{x}(\lambda). \tag{11}$$

The proof results directly from the previous theorem.

The above conclusion is the basis of the price method [3] (also known as goal coordination [6] or interaction balance [5]). The price method consists in iterative solving of the subproblems with a given λ obtained from the master unit. When using the multilevel technique, the effectiveness of the coordination algorithm has the greatest influence upon the cost of computation. While evaluating the coordination algorithms it is necessary to take into consideration the following elements:

[4]) In fact, adjoint equations are obtained initially in integral form, and transformed to the differential form (8e).

a) number of iterations of the coordination algorithm necessary to achieve required accuracy (in the sense of used norm);

b) time of computation;

c) complication of the algorithm (needed memory, size of the program etc.).

In the price method the computation of the gradient (11) is very simple. Therefore various gradient methods can be easily used. The conjugate gradient methods need fairly accurate computation of the optimal step-size because of the requirement of direction conjugation. This is related to large cost of computation. The simplicity of the conjugate gradient algorithms is their obvious advantage. It is expected however, that the use of variable metric algorithms as coordination algorithms should bring better results [11]. The Fletcher — Powel — Davidon variable metric method requires also an accurate computation of the optimal step-size which raises the computation cost. Otherwise computational errors are cumulated and reinitialization is required. The Wolf—Broyden—Davidon variable metric method does not require accurate computation of the optimal step-size [11].

Discussion of an computational example is presented in the Section V.

THEOREM 2. Assume, that v, λ belong to the space of continuous functions $C([0, T], R^m)$ and mapping $\mathscr{P}(t)$; $C([0, T], R^m) \to C([0, T], R^m)$ is defined by an iterative procedure:

$$\lambda^{j+1} = -Q_2 v^j + C' \varphi^j, \tag{12a}$$

$$v^{j+1} = Mx^j, \tag{12b}$$

where x^j, ψ^j, v^j are computed from (8a, d, e) for given λ^j, v^j then there exists $t^* > 0$ such that $\|\mathscr{P}(t)\| < 1 \ \forall t < t^*$.

The proof of the above theorem is presented in [7]. The contracting mapping method [7] (also known as interaction prediction [6]) results from the above. Subproblems (8a, d, e) are solved with given λ^j, v^j. Computed x^j, ψ^j are sent to the master unit where λ^{j+1}, v^{j+1} are computed by means of (12a, b).

The Theorem 2 indicates that the algorithm is convergent to optimal solution when the horizon of optimization T is not too long. The simplicity of the coordinator is the advantage of the contracting mapping method. The dependence of the convergence on the proper choice of optimization horizon is an obvious disadvantage.

As it is shown in [7], the method can be used for somewhat longer optimization horizons when the following scheme is applied:

$$\lambda^{j+1} = k_1 (-Q_2 v^j + C' \psi^j) + k_2 \lambda^j, \tag{12c}$$

$$v^{j+1} = k_1 Mx^j + k_2 v^j, \tag{12d}$$

where $k_1 + k_2 = 1$.

Note that this method differs from the previous one by a different way of decomposition of the optimality conditions (8a—e).

3. The sensitivity analysis, basic variations

The real processes are described by more or less accurate mathematical models. Synthesis of the control system is nearly always done with the help of an inaccurate or simplified model; thus the control strictly optimal for the model is not optimal for the real process. Applying this control to the process causes deviation of the state from its optimal value resulting in the value of the performance index which is neither optimal for the model nor for the real process. The deviation of the performance index from its optimal value, that is, the difference between obtained non-optimal value of the performance index and its optimal value, is called the sensitivity measure. The sensitivity analysis used in this paper relies on the approximation of the sensitivity measure by its first and second derivatives. The sensitivity measure $S(a)$, where a represents a changeable parameter, is approximated by Taylor series:

$$S(a)=S_0+S_a\,\delta_a+0.5\,\delta a'\,S_{aa}\,\delta_a+o(\|a\|^2)$$

where $S_0=S_a=0$ see [10], S_{aa} is called the sensitivity coefficient. Thus S_{aa} only can be used in order to approximate the losses of the performance caused by various types of inaccuracies.

Computational method of sensitivity analysis consists in:

1) computing the basic variations of all variables;

2) computing the structural variations with the help of previously computed basic variations;

3) computing the sensitivity coefficient; in the presented problem it has a form:

$$\delta a'\,S_{aa}\,\delta_a=\delta x'(T)\,F\delta x(T)+\int_0^T(\delta x'\,Q\delta x+\delta u'\,R\delta u)\,dt$$

where δx, δu are corresponding structural variations;

4) computing the performance loss s:

$$s=\frac{\delta a'\,S_{aa}\,\delta a}{\hat{J}}\,100\,\%,$$

where $\hat{J}$ is the optimal value of the performance index, δa — vector of inaccuraces.

Linear parts of deviations of the values of state and control resulting from changed value of parameters of the model are called basic variations. The optimal control defined with the help of the inaccurate model is applied to the real process which differs from the model in the value of parameters. The linear parts of deviations from optimal values are called then structural variations. In order to compute structural variations, basic variations should be perviously computed.

While investigating the sensitivity of multilevel systems, the sensitivity of a computational method due to errors in model and the sensitivity of the method itself should be distinguished from the sensitivity of real control systems. In the first case the ideal sensitivity analysis [10] or numerical accuracy analysis of the method itself is used. In the second one the real sensitivity analysis is used.

In the sequel, equations defining basic variations due to three different kinds of inaccuracies, typical for the multilevel systems, shall be defined. Assume that the price method is used.

3.1. The sensitivity due to inaccuracy of parameters contained in the matrix A

Assume that the matrix A in the model is given with some inaccuracy δA. The basic variations can be found from the equations:

$$\delta\dot{x} = A\delta x + BR^{-1} B' \delta\psi + \delta A\hat{x}; \delta x(0) = 0, \tag{13a}$$

$$\delta\dot{\psi} = -A' \delta\psi + Q\delta x - \delta A' \hat{\psi}; \delta\psi(T) = -F\delta x(T), \tag{13b}$$

where $\hat{x}$, $\hat{\psi}$ are the optimal values for the model. The above equations can be solved by the Riccati substitution.

3.2. The sensitivity due to the inaccuracy of subprocesses coordination

During two-level optimization, a coordinator, i.e. an algorithm for determining the direction of improvement for coordination variables λ, can communicate with the optimizing units and observe values from the real process. This observation is understood as memorizing entire trajectories of subprocess variables (outputs) on the whole horizon of optimization. The algorithm, if convergent, leads to a solution, which is nonoptimal because of some inaccuracy of coordination. An inaccurate coordination can be caused by an inaccurate gradient computing, errors in the information transmitting, previously assumed inaccuracy of the algorithm, etc.

It can be assumed that the obtained non-optimal solution is strictly optimal for a problem in which the interconnection equation has the form:

$$v = Mx + \varepsilon \tag{14}$$

where ε represents the value of the inaccurate coordination. The difference between the strictly optimal solution ($\varepsilon = 0$) and the nonoptimal solution is approximated by basic variations, found from linearization of optimality conditions:

$$\delta\dot{x} = A_1 \delta x + B\delta u + C\delta v; \delta x(0) = 0, \tag{15a}$$

$$\delta v = M\delta x + \varepsilon, \tag{15b}$$

$$\delta\lambda = C' \delta\psi - Q_2 \delta v, \tag{15c}$$

$$\delta u = R^{-1} B' \delta\psi, \tag{15d}$$

$$\delta\dot{\psi} = -A_1' \delta\psi - M' \delta\lambda + Q_1 \delta x; \delta\psi(T) = -F\delta x(T), \tag{15e}$$

(15a—e) can be also solved by the Riccati substitution.

3.3. Sensitivity due to small time delay in the interconnection

In some applications, the influence of time delays in the interconnections can be of interest. Consider the following interconnection equation

$$v(t)=Mx(t-h) \tag{16}$$

with a known initial function $x(t)=\varphi(t)$, $t\in[-h,0]$. Such a problem is fully investigated in [2]. Basic variations are defined by:

$$\delta\dot{x}=A_1\,\delta x+B\delta u-CM\hat{x}h;\ \delta x(0)=0, \tag{17a}$$

$$\delta\dot{\psi}=-A_1'\,\delta\psi+Q\delta x-M'\,C'\,\hat{\psi}h;\psi(T)=-F\delta x(T)-M'\,C'\,\psi(T)\,h, \tag{17b}$$

$$\delta u=R^{-1}\,B'\,\delta\psi. \tag{17c}$$

The approximation accuracy depends on the initial function, see [2].

4. Control structures and structural variations

The choice and synthesis of an optimal control structures are one of the most important and difficult problems. The application of a given structure (with the exception of the simplest — the open-loop structure) is usually possible first under some additional conditions (the closed-loop structure requires the observability of the process, the trajectory tracking structure requires one-to-one correspondence between state and control etc.). It is known that the set of possible structures is uncountable and limited only by the designer's imagination.

In the sequel three structures are presented. These structures are related to the application of multilevel technique. The structural variations are computed for each structure.

4.1. The open-closed-loop structure of control (Fig. 1.)

The structure is a combination of an open-loop structure (at the coordination level) and the classical closed-loop structure (at the control level). The control law is given by

$$u_i(t)=R_i^{-1}\,B_i'\big(K_i(t)\,x_i(t)+L_i(t)\big) \tag{19}$$

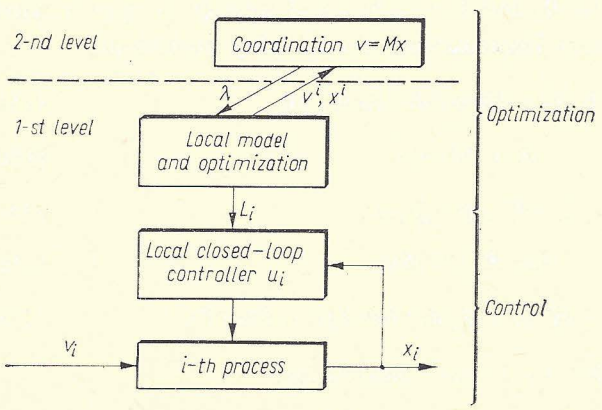

Fig. 1. The open-closed loop structure of control (opened on the coordination level, closed on the control level)

where $K_i(t)$ matrices are found as a solution of Riccati equation of i-th subprocess:

$$-\dot{K}_i(t)=K_i(t)\,A_{1i}+A'_{1i}\,K_i(t)+K_i(t)\,B_i\,R_i^{-1}\,K_i(t)-Q_i; \tag{20}$$

$$K_i(T)=-F_i$$

and $L(t)$ vector is found from:

$$\dot{L}_i(t)=-A'_{1i}\,L_i(t)-K_i(t)\,B_i\,R_i^{-1}\,B'_i\,L_i(t)-(M'\,\hat{\lambda}+K'\,C\hat{v}_i);\ L_i(T)=0. \tag{21}$$

Local analytical controllers are coordinated by the master unit by means of $\hat{\lambda}, \hat{v}$ which are computed with help of the model. If the model is inaccurate, then λ, v are not optimal for the real process and thus applied control causes deviation from the optimality. This deviation is approximated by structural variations. The structural variations of control are computed by linearization of the control law with respect to changing parameters. Some other ways of computing of structural variations are also known-see [8], [10].

The linear approximation od control deviation is

$$\delta u=R^{-1}\,B'(K\delta x+f_1) \tag{22}$$

where $K=\operatorname{diag}\{K_i\}_{i=1}^{N}$ and f_1 is found with help of equations for basic variations, namely:

(i) for the sensitivity due to inaccuracies in the matrix A:

$$\dot{f}_1=-A'_1\,f_1-K'\,BR^{-1}\,f_1-M'\,\delta\hat{\lambda}-\delta A\hat{\psi}-K'(C\delta\hat{v}-\delta A\hat{x}); \tag{23}$$

$$f_1(T)=0;$$

(ii) for the sensitivity due to coordination inaccuracy:

$$\dot{f}_1=-A'_1\,f_1-K'\,BR^{-1}\,B'\,f_1-M'\,\delta\hat{\lambda}-KC\delta\hat{v};\ f_1(T)=0; \tag{24}$$

(iii) for the sensitivity due to time delay:

$$\dot{f}_1=-A'_1\,f_1-K'\,BR^{-1}\,B'\,f_1-M'\,\delta\hat{\lambda}-KC\delta\hat{v}-M'\,\hat{\lambda}h; \tag{25}$$

$$f_1(T)=M'\,C'\,\psi(T)\,h.$$

The structural variations of the state x are computed from the linearized process equations:

$$\delta\dot{x}=A_1\,\delta x+B\delta u+C\delta v;\ \delta x(0)=0, \tag{26a}$$

$$\delta v=M\delta x. \tag{26b}$$

4.2. The on-line coordination structure of control (Fig. 2)

Assume there exist matrices K, Z and a vector f such that

$$\psi(t)=K(t)\,x(t)+Z(t)\,v(t)+f(t). \tag{27}$$

Substituting (27) into (8a—e) after some transformations the following equations are obtained

$$-\dot{K}=KA_1+A_1' K+KBR^{-1} B' K-Q_1;\ K(T)=-F, \tag{28a}$$

$$-\dot{Z}=A_1' Z+KBR^{-1} B' Z+KC;\ Z(T)=0, \tag{28b}$$

$$-\dot{f}=A_1' f+KBR^{-1} B' f+M' \hat{\lambda}-Z\hat{\dot{v}};\ f(T)=0, \tag{28c}$$

where K is block diagonal, $K=\text{diag}\,\{K_i\}_{i=1}^{N}$.

Substitution (27) can be interpreted as a two-level analytic controller. Some of the adjoint variables, ψ corresponding to interaction variables, are determined in the coordinating unit, some other corresponding to each subprocess, are determined in the local-units — see Fig. 2.

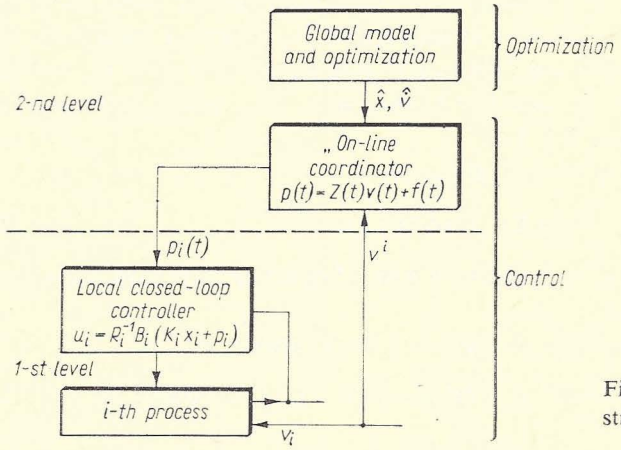

Fig. 2. The on-line coordination structure

The two-level analytic controller substitutes a closed-loop controller for the entire process. In order to construct the two-level analytic controller, N Riccati equations with dimension of each subprocess, a differential matrix equation for Z with dimension of $n\times m$, and n-dimensional equation for f based on previously computed $\hat{\dot{v}}$, $\hat{\lambda}$ should be solved. Since the dimension of interaction variables is usually much smaller than the dimension of state variables, the construction of the two-level analytic controller is more simple than the construction of a closed-loop controller for the entire process.

Structural variations are computed similarly as in the previous case. The structural variation of control is

$$\delta u=R^{-1} B' (K\delta x+Z\delta v+f_2). \tag{29}$$

The equations for f_2 are found with help of basic variations equations:

(i) for the sensitivity due to the inaccuracies in the matrix A:

$$-\dot{f}_2=A_1' f_2+KBR^{-1} B' f_2+Z\delta\hat{\dot{v}}+\delta A' \hat{\psi}+K\delta A\hat{x}; \tag{30}$$

$$f_2(T)=0;$$

(ii) for the sensitivity due to the coordination inaccuracy:

$$-\dot{f}_2=A_1' f_2+KBR^{-1} B' f_2+Z\delta\hat{\dot{v}}-M' \delta\hat{\lambda}; \tag{31}$$

$$f_2(T)=0;$$

(iii) for the sensitivity due to time delay:

$$-\dot{f}_2 = A_1' f_2 + KBR^{-1} B' f_2 + Z\delta\hat{\tilde{v}} + M' \delta\hat{\lambda} + M' C' \hat{\psi} h; \tag{32}$$

$$f_2(T) = M' K' \hat{\psi}(T) h.$$

4.3. Closed-loop one-level structure of control with the previous estimation of interaction variables — (Fig. 3)

Suppose that estimators for the optimal values of interaction variables are known (e.g. because of long system utilization). These estimators have the form of $\tilde{v} = \hat{v} + \rho$, where ρ represents the estimation inaccuracy. The optimization of sub-

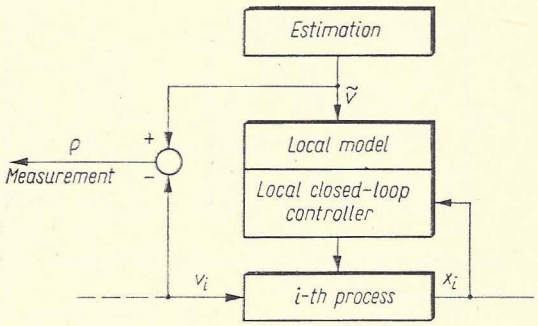

Fig. 3. The closed-loop one-level structure with the previous estimation of interaction variables

processes and the design of control one-level structure are based on these estimators. It is obvious that the applied control resulting from such a structure is not optimal for the entire system. Performance losses caused by that deviation from optimality can be evaluated by sensitivity analysis. The performance loss due to a given inaccuracy of estimation can be compared with the costs of the coordinator itself and its utilization. This comparison gives an answer to the question whether or not it is worthwhile to coordinate. Each of subprocesses is described by an equation:

$$\dot{x}_i = A_{1i} x_i + B_i u_i + (C\tilde{v})_i; \; x_i(0) = x_{0i} \tag{33}$$

and the local performance indices are as follows:

$$J_i = 0.5\, x_i'(T) F_i x_i(T) + 0.5 \int_0^T (x_i' Q_i x_i + u_i' B_i u_i)\, dt. \tag{34}$$

The optimality conditions are given by (8a, e) and (33).
Hence, the basic variations are determined by:

$$\delta\dot{x}_i = A_{1i} \delta x_i + B_i \delta u_i + (C\rho)_i; \; \delta x_i(0) = 0, \tag{35a}$$

$$\delta\dot{\psi}_i = -A_{1i}' \delta\psi_i + Q_i \delta x_i; \; \delta\psi_i(T) = -F_i \delta x_i(T), \tag{35b}$$

$$\delta u_i = R_i^{-1} B_i' \delta\psi_i. \tag{35c}$$

Structural variations are computed as follows:

$$\delta u_i = R_i^{-1} B_i' (K_i \delta x_i + f_{3i}), \tag{36}$$

$$\delta\dot{x}_i = A_{1i} \delta x_i + B_i \delta u_i + (C\delta v)_i, \tag{37}$$

where $\delta v = M\delta x$ and f_{3i} is found from the linearization of the optimal controller:

$$-\dot{f}_{3i} = (A'_{1i} + K_i B_i R_i^{-1} B'_i) f_{3i} + K_i (C)_i;$$
$$f_{3i}(T) = 0 \tag{38}$$

5. Example

The above analysis is illustrated by the sensitivity analysis of a simple example which is linear-quadratic in form (the state equation is linear (1), the performance index is quadratic (2), the final time T is fixed). Dimension of the state is 6.

The following matrices have been assumed:

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \begin{matrix} Q_1 = F = I_{6\times 6} \\ Q_2 = R = I_{3\times 3}. \end{matrix}$$

In order to decompose the system, the following interaction variables have been assumed

$$v_1 = x_6, \; v_2 = x_2, \; v_3 = x_4.$$

Moreover, the contracting mapping method has been applied to dynamic optimization while decomposing the global problem into 6 subproblems. Matrices corresponding to such decomposition have been accordingly changed.

5.1. Dynamic Optimization

Computation was programmed in Algol on ODRA—1204 computer. Integration was performed by trapezium method. Differential equations were solved by Runge Kutta algorithm of 7-th order. The global problem has been solved by Riccati substitution to compare the time of computation. Results are given in the Table 1. The norm of the gradient g in space L^2 [0, 1] has been used as a measure of coordination convergence

$$\|g\|^2 = \langle g, g \rangle = \int_0^1 g'(t)\, g(t)\, dt.$$

The scalar $\gamma = \ln \dfrac{\|g^i\|}{\|g^1\|}$ has been used as a measure of improvement in i-th iteration. The time interval has been discretized in 11 points.

5.1.1. The Contracting Mapping Method

The global problem has been solved while decomposing it into:

a) 6 one-dimensional subproblems, $v = x$,

b) 3 one-dimensional subproblems, $v = Cx$.

The curves γ versus number of iterations, are shown in the Fig. 4, where:

a) k_1 equal 1 has been substituted in (12c—d) in the case a (curve *A*),

b) k_1 equal 0.5, 0.75, 1.0 has been substituted in (12c—d) in the case b (curves *B*, *C*, *D*, *E*).

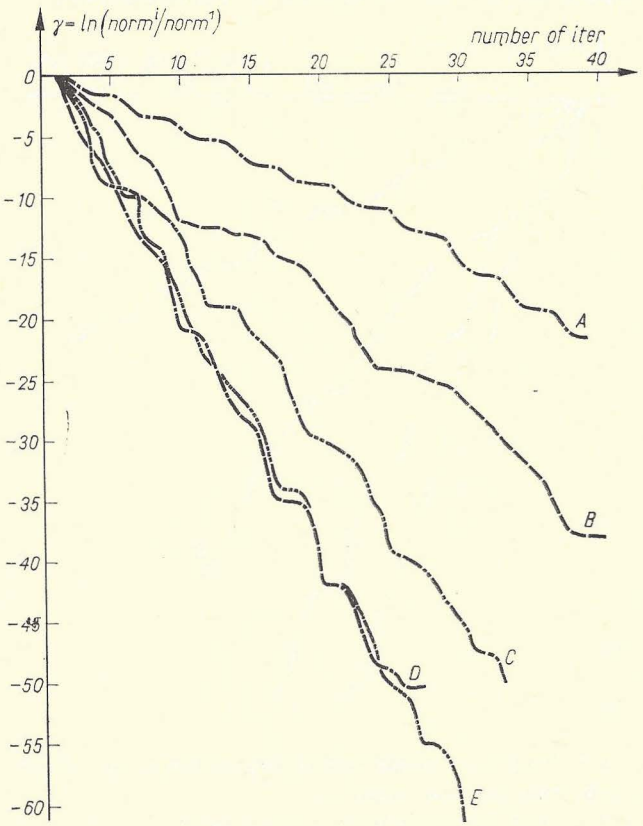

Fig. 4. The scalar γ versus number if iterations in the contracting mapping method:

$A-k_1=1.0$ decomposing into 6 subproblems, $x_0=1$,
$B-k_1=0.5$
$C-k_1=0.75$ decomposing into 3 subproblems, $x_0=1$,
$D-k_1=1.0$
$E-k_1=1.0$ decomposing into 3 subproblems, $x_0=0$

The solutions with the initial state $x_0=1$ (curves *B*, *C*, *D*) and with the initial state $x_0=0$ (curve *E*), where the solution is trivial i.e. the optimal values of the state, the control and the performance index are equal zero) has been compared in case b. It can be noted that the computational process depends on the method of decomposition (solution without decomposition, with decomposition into 6 subproblems or into 3 subproblems). In the case a, there are 6 interaction variables and subproblems are strongly coupled so that the coordination is rather difficult. There are 3 interaction variables in the case b and that is why coordination is not so difficult. The curves γ versus time of computation are shown in the Fig. 5. In the

case a, one iteration takes 65 sec. (solution of six subproblems). The time of one iteration in the case b equals 48 sec. Vertical line marks time of solving the problem without decomposition by Riccati substitution (it takes 350 sec.). Table 1 illustrates the cost of computation due to the method of decomposition.

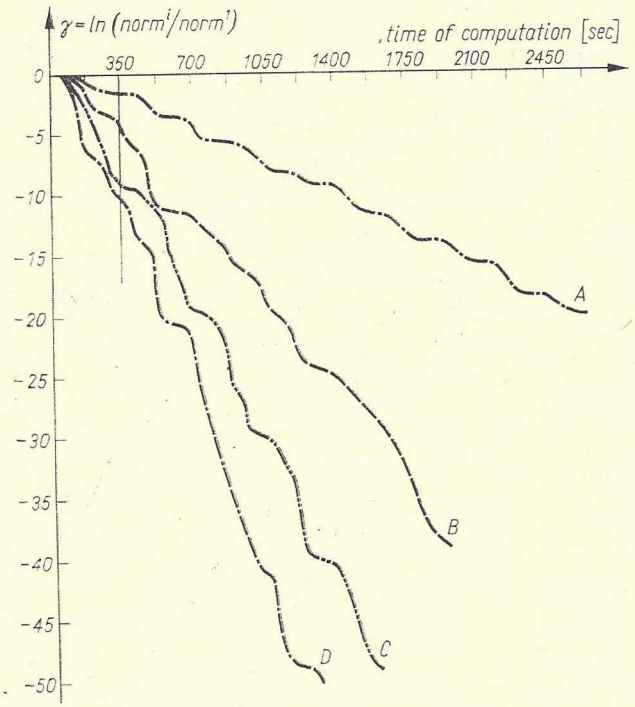

Fig. 5. The scalar γ versus time of computation in the contracting mapping method, $x_0 = 1$:
$A - k_1 = 1.0$ decomposing into 6 subproblems,
$B - k_1 = 0.5$
$C - k_1 = 0.75$ γ decomposing into 3 subproblems
$D - k_1 = 1.0$

Some differences in the values of the performance index are caused by integration with the use of the trapezium method. As it is known, the accuracy of such integration depends on the character of function and number of discretization points.

Table 1

Number of subproblems	Number of discretization points	Time of 1 iteration in sec.	Obtained performance index
—	11	—	51.41355
6	11	65	52.1540
3	11	48	51.8057
3	6	24	52.9784
3	21	95	51.5117
3	51	236	51.4292

It can be noted that the cost of computation increases as number of discretization points increases. Diagrams in the Fig. 6 represent dependence of γ on number of iterations. Curves are plotted with the length of optimization horizon treated as

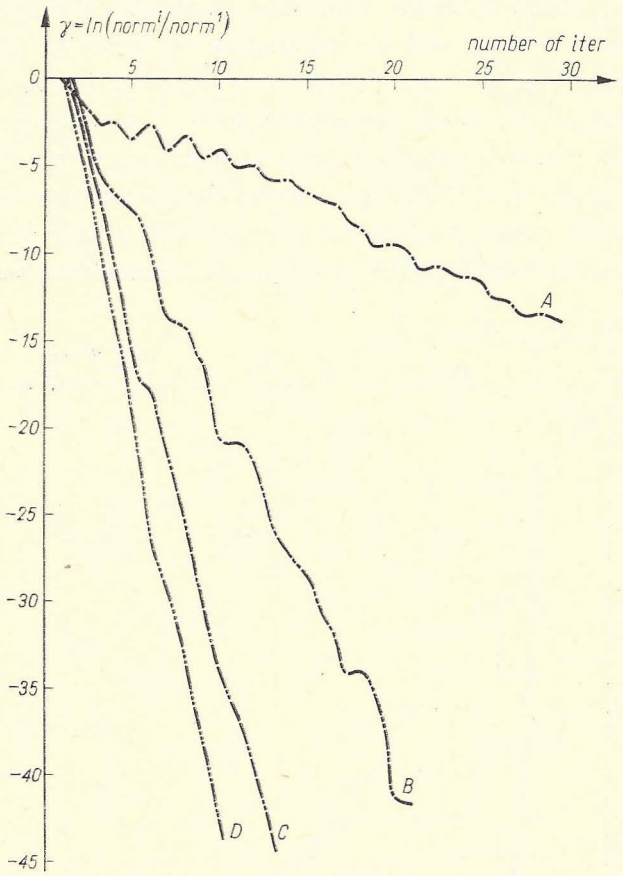

Fig. 6. The scalar γ versus number of iterations in the contracting mapping method, $x_0=1$, $k_1=1.0$.
$A-T=5.0$, $B-T=1.0$, $C-T=0.5$, $D-T=0.3$

a parmeter. The contracting mapping method is not convergent when the optimization horizon is too long ($T_{\max}>10$) which is consistent with the Theorem 2.

5.1.2. The Price Method

The price method is fully described in section II. In order to maximize functional (10) following gradient methods were applied

(i) the steepest descent,
(ii) Fletcher—Reeves conjugate gradient,
(iii) Fletcher—Powell—Davidon variable metric,
(iv) Wolfe—Broyden—Davidon variable metric.

The directional maximization is based on the Goldstein test with coefficient $\beta=0{,}4$.

Comparison of the gradient methods is difficult. It is mainly caused by numerical errors. These numerical errors result from inconsistent discretization. All the formulas have been derived with the assumption of continuity of all variables and the global problem has been discretized a posteriori (during computation). During the computational process representations of all variables in discretization points or their linear approximations between these points are only available. These errors can be avoided by discretization a priori (before computation) and use of discrete methods. It would speed up the computational process (difference equations would be solved instead of differential equations, integration would be replaced

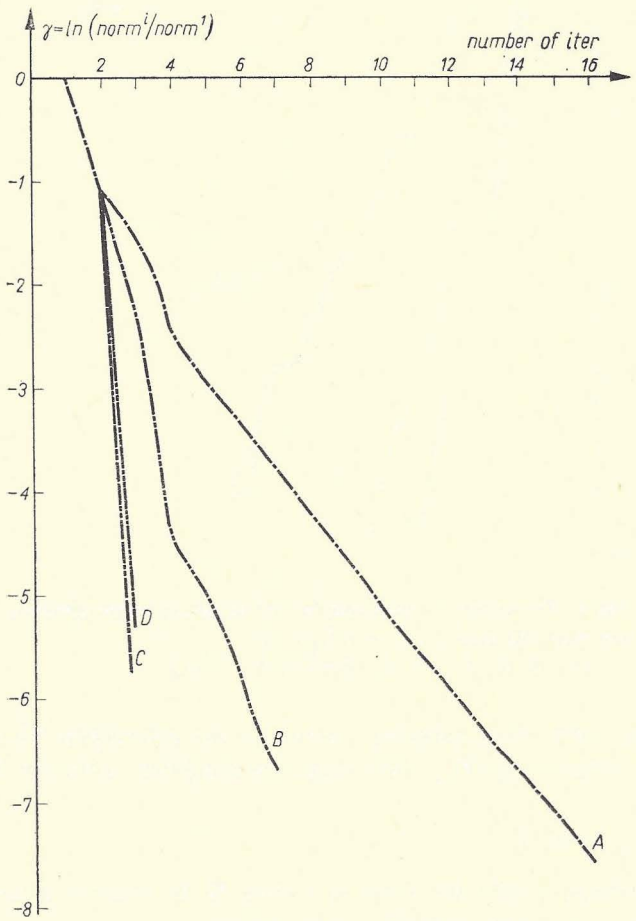

Fig. 7. The scalar γ versus number of iterations in the price method, directional maximization algorithm used, $x_0=1$. A — the steepest descent, B — Fletcher—Reeves conjugate gradient, C — Fletcher—Powell—Davidon variable metric, D — Wolfe—Broyden—Davidon variable metric

by summation) and would allow to avoid errors caused by inconsistent discretization. That problem was not investigated in this paper.

Diagram of γ versus number of iterations is presented in the Fig. 7. Computation was relatively quickly stopped because it was impossible to find maximum in properly computed direction (it has been checked by computing projection of the gradient on computed direction which has been always positive). That problem was caused by inconsistent discretization, as it was mentioned above. Further

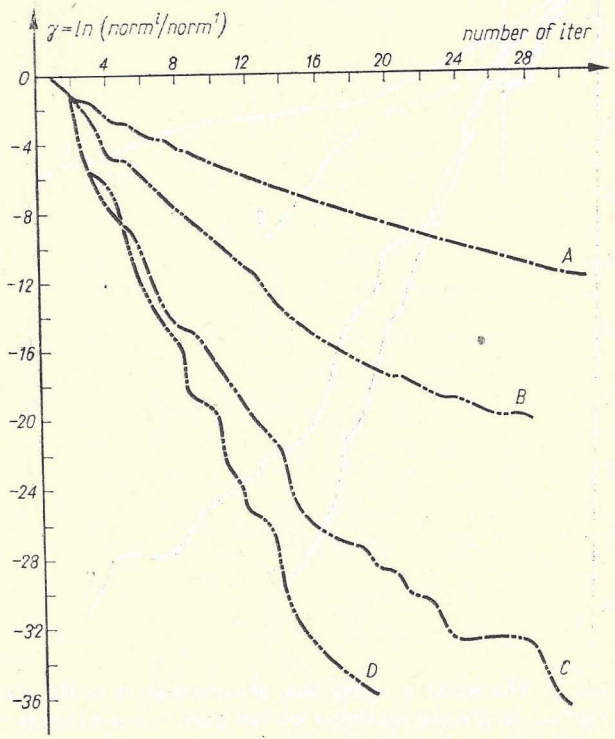

Fig. 8. The scalar γ versus number of iterations in the price method, directional maximization not used, step-size coefficient equal 1.0, $x_0 = 1$.
A — the steepest descent, B — Fletcher—Reeves conjugate gradient, C — Fletcher—Powell—Davidon variable metric, D — Wolfe—Broyden—Davidon variable metric

computations have been made with a constant step-size coefficient equal 1 since the directional maximization algorithm did not work. Further improvement was achieved in this way. Curves γ versus number of iterations are shown in the Fig. 8. The Figure 9 represents diagram of γ versus time of computation.

These numerical experiences show that it may be supposed that the price method has the property of contracting operator in the neighbourhood of optimal solution. Directional maximization leads to that neighbourhood and further improvement is obtained with a constant step-size.

Generally speaking the price method is faster than the contracting mapping method but the price method algorithm is more complicated than that of the contracting mapping method.

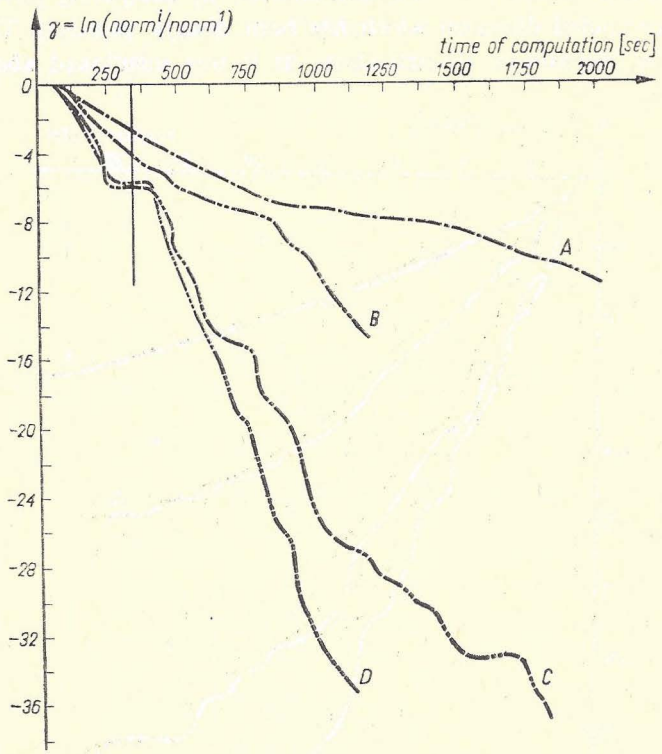

Fig. 9. The scalar γ versus time of computation in the price method, directional maximization not used, step-size coefficient equal 1.0, $x_0=1$.
A — the steepest descent, B — Fletcher—Reeves conjugate gradient, C — Fletcher—Powell—Davidon variable metric, D — Wolfe—Broyden—Davidon variable metric

5.2. The Sensitivity Analysis

The sensitivity analysis of two structures due to 10% inaccuracy in 3 parameters of the matrix A: a_{11}, a_{33}, a_{55} and inaccurate coordination $\varepsilon=0.1$ was done. These structures are as follows:

a) the open-closed-loop structure of control (Fig. 1),

b) the on-line coordination structure of control (Fig. 2).

Computation was programmed in Fortran on CDC 3170 computer. Differential equations were solved by standard Runge Kutta subroutine from the IBM Scientific Subroutine Package. Results are presented below.

5.2.1. The Open-Closed-Loop Structure of Control

The sensitivity due to inaccuracy in the matrix A

$$S_{aa}=\begin{bmatrix}2.601 & 2.579 & 2.382\\ 2.597 & 2.791 & 2.517\\ 2.382 & 2.517 & 2.701\end{bmatrix}, \quad s=0.714\%.$$

The sensitivity due to inaccurate coordination

$$S_{aa}=\begin{bmatrix}5.567 & 4.536 & 5.242\\ 4.536 & 3.859 & 4.293\\ 5.242 & 4.293 & 5.964\end{bmatrix}, \quad s=1.319\%.$$

5.2.2. The on-line coordination structure of control

The sensitivity due to inaccuracy in the matrix A

$$S_{aa}=\begin{bmatrix}2.155 & 2.581 & 3.089\\ 2.581 & 3.218 & 3.732\\ 3.089 & 3.732 & 4.509\end{bmatrix}, \quad s=0.887\%.$$

The sensitivity due to inaccurate coordination

$$S_{aa}=\begin{bmatrix}3.569 & 3.562 & 3.465\\ 3.562 & 3.586 & 3.495\\ 3.465 & 3.495 & 3.736\end{bmatrix}, \quad s=0.988\%.$$

It seems, that in that case the second structure of control is better. No general conclusion according the sensitivity of different structures of control can be achieved. It is only possible to compare control structures in particular case. The computation time of sensitivity analysis is nearly equal that of dynamic optimization of the problem.

6. Conclusions

The price method of the multilevel dynamic optimization has been investigated. The sensitivity due to different kinds of inaccuracies was researched: i.e. inaccuracy in the matrix A, inaccuracy of subprocesses coordination, small time delay in the interconnections. For the sensitivity analysis three structures of on-line multilevel systems were investigated. It was shown that the computational effort related to the sensitivity analysis is comparable with that of the dynamic optimization. The sensitivity analysis makes it possible to choose one of the structures considered such that the influence of the inaccuracy of the model of the system is decreased.

References

1. DONTCHEV A. L.: Coordination and sensitivity analysis of multilevel control system with interaction in performance index. (In Polish). *Arch. Autom. i Telemech.* **23**, 1 (1973).
2. DONTCHEV A. L.: Sensitivity of optimal control systems with small time delay. *SIAM J. Control* (to appear).
3. FINDEISEN W.: Multilevel control systems. (In Polish). Warszawa 1974.
4. LUENBERGER D. G.: Optimization by vector space method. New York 1969.
5. MESAROVIC M. D., MACKO D., TAKAHARA Y.: Theory of hierarchical multilevel systems. New York 1970.
6. PEARSON W. D.: Dynamic decomposition technique. In: Optimization method for large scale systems. Ed. D. A. Wismer. New York 1971.
7. PETRICZEK G.: Optimization methods and sensitivity analysis of optimal control systems with time delay. (In Polish). Master thesis in Institute of Automatic Control, Technical University of Warsaw. Warszawa 1974.
8. URBANIK A.: Sensitivity of multilevel dynamic optimization structures. (In Polish). Master thesis in Institute of Automatic Control. Technical University of Warsaw. Warszawa 1974.
9. WIERZBICKI A. P.: Unified approach to the sensitivity analysis of control systems. Materials of IV IFAC Congr. Warszawa 1969.
10. WIERZBICKI A. P., DONTCHEV A. L.: Basic relations in the sensitivity analysis of optimal control systems. Automatica (to appear).
11. WIERZBICKI A. P.: Coordination and sensitivity analysis of a large scale system with performance interaction. Preprints of V IFAC Congr. Paris 1972.

Received, January 1976.

Optymalizacja dynamiczna i analiza wrażliwości wielopoziomowych układów sterowania

Praca poświęcona jest problemom optymalizacji i analizy wrażliwości wielopoziomowych układów sterowania opisanych liniowymi równaniami różniczkowymi z kwadratowym wskaźnikiem jakości.

Przedstawiono dwie dwupoziomowe metody optymalizacji dynamicznej. W pracy zaprezentowano analizę wrażliwości takich układów na niedokładność macierzy stanu, niedokładną koordynację oraz małe opóźnienie w zmiennych interakcyjnych. Analizę wrażliwości zastosowano do kilku jedno i dwupoziomowych struktur sterowania optymalnego. Zaproponowano suboptymalną strukturę sterowania opartą na predykcji zmiennych interakcyjnych. Przeprowadzono analizę wrażliwości tej struktury. Wyniki zilustrowano przykładem obliczeniowym.

Динамическая оптимизация и анализ чувствительности многоуровневых сискем управления

В работе обсуждено оптимизацию и анализ чувствительности многоуровневых систем управления которые описаны линейными дифференциальными уравнениями с квадратичным критерием качества. Представлены два различные вычислительные методы двууровневой динамической оптимизации. Показано анализ чувствительности второго порядка этих методов на неточность матрици состояния, неточную координацию и невеликое опаздание интеракции между субпроцессами. Применено анализ чувствительности к некотороым одно- и двууровневым системам управления. Предложено субоптимальную структуру управления которая заключается в предикции переменных интеракционных. В работе тоже представлено анализ чувствительности этой структуры. Работа заключается вычислительном примером.