KYBERNETIKA — VOLUME 9 (1973), NUMBER 2

On a Solution of an Optimization Problem in Linear Control Systems with Quadratic Performance Index

NGUYEN CANH

This article represents an iterative procedure for determining of the optimal control for the linear control systems, with a quadratic performance index, using special Hilbert space. The results, obtained in this paper, will be useful for solving of linear problems, in which the admissible control is a measurable, bounded function, with values in any determined convex set.

1. FORMULATION OF THE PROBLEM

We shall consider the linear control system defined by the vector differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\,\mathbf{x}(t) + \mathbf{B}(t)\,\mathbf{u}(t) \tag{1}$$

with initial condition

$$\mathbf{x}(0) = \mathbf{x}_0 . \tag{2}$$

Here $\mathbf{x}(t)$ is an n-dimensional state vector for each t, $\mathbf{u}(t)$ is an r-dimensional control vector and $\mathbf{A}(t)$ and $\mathbf{B}(t)$ are $n \times n$ and $n \times r$ matrices, which are continuous in the time t. The control $\mathbf{u}(t)$ will be called an admissible control if every component $u_i(t)$ of the control vector $\mathbf{u}(t)$ is measurable and satisfies the constraint

$$|u_i(t)| \leqq 1 , \quad i = 1, 2, \ldots, r , \tag{3}$$

for any time over $[0, T]$. We shall denote the set of admissible controls by U.

The performance index to be used below is defined as

$$J[\mathbf{u}] = \int_0^T \{\mathbf{x}^*(t)\,\mathbf{Q}(t)\,\mathbf{x}(t) + \mathbf{u}^*(t)\,\mathbf{R}\,\mathbf{u}(t)\}\,dt . \tag{4}$$

Here $\mathbf{Q}(t)$ is an $n \times n$ positive semidefinite symetric matrix which is continuous in the time t, $\mathbf{R}$ is an $r \times r$ positive definite diagonal matrix with positive constant elements, T is a fixed time and * denotes the transpose of a matrix or of a vector.

The problem is then to choose an appropriate admissible control vector $\boldsymbol{u}(t)$ so that the performance index $J[\boldsymbol{u}]$ is minimized, subject to (1) and to the additional condition.

The procedure used for solving this problem will be based on the method described in the article of T. Fujisawa and Y. Yasuda [2], in which a similar problem has been explained.

All necessary theorems and illustrative example will be mentioned.

2. FORMULATION OF THE OPTIMIZATION IN THE HILBERT-SPACE

Let H_m be a real Hilbert-space of m-dimensional functions square integrable over the time interval $[0, T]$

$$H_m = \left\{ \mathbf{y} \mid \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \quad y_k \in L_2[0, T], \quad k = 1, 2, \ldots, m \right\}. \tag{5}$$

Let us denote the inner product of two m-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ in the Hilbert-space H_m by $(\mathbf{x}, \mathbf{y})_m$, which is defined by

$$(\mathbf{x}, \mathbf{y})_m = \int_0^T \mathbf{x}^*(t)\, \mathbf{y}(t)\, \mathrm{d}t = \sum_{i=1}^m \int_0^T x_i(t)\, y_i(t)\, \mathrm{d}t\,. \tag{6}$$

$x_i(t)$, $y_i(t)$ are ith-components of the vectors $\mathbf{x}(t)$ and $\mathbf{y}(t)$, and the norm of any element $\mathbf{x} \in H_m$ by

$$\|\mathbf{x}\|_m = +\sqrt{(\mathbf{x}, \mathbf{x})_m}\,. \tag{7}$$

Lemma 1. *The set of the vector functions in the vector space defined above satisfies the following axioms of addition and scalar multiplication of inner product*

$$(\mathbf{x}, \mathbf{x})_m \geqq 0 \quad \textit{for each} \quad \mathbf{x} \in H_m\,, \tag{8}$$

$$(\mathbf{x}, \mathbf{x})_m = 0 \quad \textit{implies} \quad \mathbf{x} = 0 \quad \textit{almost everywhere}\,, \tag{9}$$

$$(\mathbf{x}, \mathbf{y})_m = (\mathbf{y}, \mathbf{x})_m\,, \tag{10}$$

$$(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2, \mathbf{y})_m = \alpha_1(\mathbf{x}_1, \mathbf{y})_m + \alpha_2(\mathbf{x}_2, \mathbf{y})_m \tag{11}$$

for any real numbers α_1 and α_2.

Proof. According to the defining above, we shall have

$$(\mathbf{x}, \mathbf{x})_m = \int_0^T \mathbf{x}^*(t)\, \mathbf{x}(t)\, dt = \sum_{i=1}^m \int_0^T x_i^2(t)\, dt .$$

The integrand of the last term of the equation is a positive definite quadratic form. Hence (8) and (9) obviously hold.

Now we shall proof the thirth property

$$\begin{aligned}(\mathbf{x}, \mathbf{y})_m &= \int_0^T \mathbf{x}^*(t)\, \mathbf{y}(t)\, dt = \sum_{i=1}^m \int_0^T x_i(t)\, y_i(t)\, dt = \\ &= \sum_{i=1}^m \int_0^T y_i(t)\, x_i(t)\, dt = \int_0^T \mathbf{y}^*(t)\, \mathbf{x}(t)\, dt = \\ &= (\mathbf{y}, \mathbf{x})_m .\end{aligned}$$

The relation (10) is proved. Finally

$$\begin{aligned}((\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2), \mathbf{y})_m &= \sum_{i=1}^m \int_0^T (\alpha_1 x_{1i} + \alpha_2 x_{2i})\, y_i\, dt = \\ &= \alpha_1 \sum_{i=1}^m \int_0^T x_{1i} y_i\, dt + \alpha_2 \sum_{i=1}^m \int_0^T x_{2i} y_i\, dt = \\ &= \alpha_1 (\mathbf{x}_1, \mathbf{y})_m + \alpha_2 (\mathbf{x}_2, \mathbf{y})_m .\end{aligned}$$

The lemma above is proved.

Then the state vector $\mathbf{x}(t)$ $(0 \leqq t \leqq T)$ will be in H_n and the control vector can be taken in H_r.

Lemma 2. *The set of admissible control U defined above is a convex bounded and strongly closed in the Hilbert space H_r.*

Proof. First, we shall prove the set U is convex. It means that for every $\mathbf{u} \in U$ and $\mathbf{v} \in U$ and for $0 \leqq \theta \leqq 1$ the points $(1 - \theta)\, \mathbf{u} + \theta \mathbf{v}$ also belong to the set U.

According to the property of the admissible control U we have

$$(12) \qquad (1 - \theta)\, u_{i\min} + \theta v_{i\min} \leqq (1 - \theta)\, u_i + \theta v_i \leqq (1 - \theta)\, u_{i\max} + \theta v_{i\max} ,$$

$$u_{i\min} = v_{i\min} = -1 ,$$

$$u_{i\max} = v_{i\max} = 1 .$$

From the relation (12) it follows

$$-1 \leqq (1 - \theta)\, u_i + \theta v_i \leqq 1 .$$

$$\|(1 - \theta)\, u_i - \theta v_i\| \leqq 1 .$$

The convexity of the set U is proved.

It is easy to see that the set U is bounded in the Hilbert space H_r.

In fact, for any $\mathbf{u} \in U$ there is

$$\|\mathbf{u}\|_r^2 = \|\sum_{i=1}^{r} \int_0^T u_i^2(t)\, dt\| \leqq rT$$

hence

$$\|\mathbf{u}\| \leqq \sqrt{(rT)}\,.$$

The set admissible control U is measurable and bounded in the Hilbert space H_r, therefore the set U is complete in H_r. It means that any sequence $\{\mathbf{u}_j\}$ satisfies the following relation

$$\|\mathbf{u}_n(t) - \mathbf{u}_m(t)\|_r < \varepsilon$$

for n, $m > N(\varepsilon)$, or in other symbols

$$\sqrt{\left\{\int_0^T \sum_{i=1}^{r} [u_{in}(t) - u_{im}(t)]^2\, dt\right\}} < \varepsilon$$

for n, $m > N(\varepsilon)$, (the sequence $\{\mathbf{u}_j\}$ converges into themselves), then $\{\mathbf{u}_j\}$ has the limit $\mathbf{u}^0 \in U$.

$$\lim_{j\to\infty} \mathbf{u}_j = \mathbf{u}^0\,.$$

The solution of (1) with the initial condition $\mathbf{x}(0)$ is given by

$$\mathbf{x}(t) = \boldsymbol{\Phi}(t)\,\mathbf{x}(0) + \boldsymbol{\Phi}(t)\int_0^t \boldsymbol{\Phi}^{-1}(\tau)\,\mathbf{B}(\tau)\,\mathbf{u}(\tau)\,d\tau \tag{13}$$

where $\boldsymbol{\Phi}(t)$, the $n \times n$ fundamental matrix satisfies the following matrix differential equation

$$\dot{\boldsymbol{\Phi}}(t) = \mathbf{A}(t)\,\boldsymbol{\Phi}(t)\,, \tag{14}$$

$$\boldsymbol{\Phi}(0) = \mathbf{I} \tag{15}$$

($\mathbf{I}$ is $n \times n$ identity matrix).

For notational simplicity, we shall denote

$$\mathbf{W}(t, \tau) = \begin{cases} \boldsymbol{\Phi}(t)\,\boldsymbol{\Phi}^{-1}(\tau)\,\mathbf{B}(\tau) & \text{for} \quad \tau \leqq t\,, \\ \mathbf{0} & \text{if otherwise}\,. \end{cases} \tag{16}$$

Clearly $\mathbf{W}(t, \tau)$ is an $n \times r$ matrix, and by the condition (16) we obtain

$$\int_0^T \mathbf{W}(t, \tau)\,\mathbf{u}(\tau)\,d\tau = \int_0^t \mathbf{W}(t, \tau)\,\mathbf{u}(\tau)\,d\tau + \int_t^T \mathbf{W}(t, \tau)\,\mathbf{u}(\tau)\,d\tau\,. \tag{17}$$

Then the solution (13) can be rewritten as follows

$$\mathbf{x}(t) = \Phi(t)\,\mathbf{x}(0) + \int_0^T \mathbf{W}(t, \tau)\,\mathbf{u}(\tau)\,\mathrm{d}\tau \,. \tag{18}$$

We define a linear integral operator $\mathbf{P}$ on H_r by

$$P(\mathbf{u}) = \int_0^T \mathbf{W}(t, \tau)\,\mathbf{u}(\tau)\,\mathrm{d}\tau \,. \tag{19}$$

Lemma 3. *The linear operator*

$$P(\mathbf{u}) = \int_0^T \mathbf{W}(t, \tau)\,\mathbf{u}(\tau)\,\mathrm{d}\tau$$

defines a bounded operator and maps space H_r into space H_n.

Proof. The matrix $\mathbf{W}(t, \tau)$ is bounded in the region $0 \leqq t \leqq T$ and $0 \leqq \tau \leqq T$, because the matrices $\Phi(t)$, $\Phi^{-1}(\tau)$, $\mathbf{B}(\tau)$ are continuous with respect to the time variable t, which satisfies the relations $0 \leqq t \leqq T$, therefore the (i, j) element W_{ij} of the matrix $\mathbf{W}(t, \tau)$ is bounded in the region $0 \leqq t \leqq T$, it is meaning that exists a positive constant K^2, which satisfies

$$\int_0^T \int_0^T |W_{ij}(t, \tau)|^2\,\mathrm{d}t\,\mathrm{d}\tau \leqq K^2 < \infty \,. \tag{20}$$

Applying Schwarz inequality to the j-th component of the term

$$\sum_{j=1}^{r} \int_0^T W_{ij}(t, \tau)\,u_j(\tau)\,\mathrm{d}\tau$$

there holds

$$\left| \int_0^T W_{ij}(t, \tau)\,u_j(\tau)\,\mathrm{d}\tau \right|^2 \leqq \int_0^T |W_{ij}(t, \tau)|^2\,\mathrm{d}\tau \int_0^T |u_j(\tau)|^2\,\mathrm{d}\tau \,. \tag{21}$$

Integrating both sides with respect to the variable t over $[0, T]$, it follows

$$\int_0^T \mathrm{d}t \left| \int_0^T W_{ij}(t, \tau)\,u_j(\tau)\,\mathrm{d}\tau \right|^2 \leqq K^2 \int_0^T |u_j(\tau)|^2\,\mathrm{d}\tau = K^2 \|u_j\|^2 \,. \tag{22}$$

Here $\| \;\; \|$ denotes the usual norm in $L_2[0, T]$. We know that the norm of the summ is less or equal to the summ of the norms of the terms in this summ. Hence

$$\left\| \sum_{j=1}^{r} \int_0^T W_{ij}(t, \tau)\,u_j(\tau)\,\mathrm{d}\tau \right\| \leqq K \sum_{j=1}^{r} \|u_j\| \tag{23}$$

and for the norm of the vector

$$\int_0^T \mathbf{W}(t, \tau)\, \mathbf{u}(\tau)\, d\tau$$

in the space H_n it holds

$$(24) \qquad \left\| \int_0^T \mathbf{W}(t, \tau)\, \mathbf{u}(\tau)\, d\tau \right\|_n \leqq \sqrt{(n)}\, K \sum_{j=1}^{r} \|u_j\| \leqq$$

$$\leqq \sqrt{(n)}\, K \sqrt{r \sum_{j=1}^{r} \|u_j\|^2} = \sqrt{(nr)}\, K \|\mathbf{u}\|_r .$$

It follows

$$(26) \qquad P(\mathbf{u}) \leqq \sqrt{(nr)}\, K \|\mathbf{u}\|_r .$$

By the same procedure mentioned above, we have the operator

$$(26) \qquad \mathbf{R}(\mathbf{u}(t)) = \mathbf{R}\, \mathbf{u}(t) \quad \text{for} \quad t \in [0, T] .$$

It defines the symetric positive linear bounded operator from H_r into H_r. There exists a positive constant $M > 0$ such that

$$(27) \qquad (\mathbf{u}, \mathbf{Ru})_r \geqq M(\mathbf{u}, \mathbf{u})_r = M \|\mathbf{u}\|_r^2 .$$

Let M be the smallest eigenvalue of the positive definite matrix $\mathbf{R}$.

The operator $\mathbf{Qx}$ is also symetric, positive linear bounded operator from H_n into H_n.

Then the performance index can be written as follows

$$(28) \qquad J[\mathbf{u}] = (\mathbf{P}(\mathbf{u}) + \mathbf{g}, \mathbf{Q}(\mathbf{g} + \mathbf{P}(\mathbf{u})))_n + (\mathbf{u}, \mathbf{Ru})_r$$

Here

$$\mathbf{g} = \boldsymbol{\Phi}(t)\, \mathbf{x}(0) .$$

This equation can be expanded to give

$$(29) \qquad J[\mathbf{u}] = (\mathbf{g}, \mathbf{Qg})_n + 2(\mathbf{Qg}, \mathbf{P}(\mathbf{u}))_n + (\mathbf{P}(\mathbf{u}), \mathbf{Q}\, \mathbf{P}(\mathbf{u}))_n + (\mathbf{u}, \mathbf{Ru})_r .$$

Now let $\mathbf{P}^*$ be the adjoint operator of $\mathbf{P}$, then $\mathbf{P}^*$ maps H_n into H_r and satisfies the relations

$$(30) \qquad \mathbf{P}^*(\mathbf{u}) = \int_\tau^T \mathbf{W}^*(t, \tau)\, \mathbf{u}(t)\, dt ,$$

$$(31) \qquad (\mathbf{x}, \mathbf{P}(\mathbf{u})) = (\mathbf{P}^*(\mathbf{x}), \mathbf{u}) ,$$

where $\mathbf{x} \in H_n$ and $\mathbf{u} \in H_r$.

Thus equation (29) can be written as

$$J[u] = (g, Qg)_n + 2(P^*(Qg), u)_r + (u, P^*(QPu))_r + (u, Ru)_r . \tag{32}$$

It can be proved as follows

$$(P^*(Qg), u)_r = (Qg, P(u))_r = \tag{33}$$

$$= \int_0^T dt\, g^*(t)\, Q(t) \int_0^T W(t, \tau)\, u(\tau)\, d\tau .$$

Changing the order of the integration in (33), it yields

$$(P^*(Qg), u)_r = \int_0^T \int_\tau^T \{W^*(t, \tau)\, Q(t)\, g(t)\}^*\, dt\, u(\tau)\, d\tau . \tag{34}$$

Equation (34) shows that

$$P^*(Qg) = \int_\tau^T W^*(t, \tau)\, Q(t)\, g(t)\, dt . \tag{35}$$

From (35) it follows that

$$P^*(Q\, P(u)) = \int_\tau^T dt\, W^*(t, \tau)\, Q(t) \int_0^t W(t, s)\, u(s)\, ds . \tag{36}$$

Changing the order of integration (36), it yields

$$P^*(Q\, P(u)) = \int_0^T \left\{ \int_{\max(\tau,s)}^T W^*(t, \tau)\, Q(t)\, W(t, s)\, dt \right\} u(s)\, ds . \tag{37}$$

Since

$$(P^*(QP))^* = P^*(QP) \tag{38}$$

the linear bounded operator $P^*(QP)$ on H_r into H_r is self-adjoint.

Moreover since

$$(P^*(Q\, P(u)), u)_r = (Q\, P(u), P(u))_r \geqq 0 \tag{39}$$

for arbitrary $u \in H_r$, the operator $P^*(QP)$ is positive.

By defining

$$\alpha = (g, Qg)_n , \tag{40}$$

$$h = 2P^*(Qg) \in H_r , \tag{41}$$

and

$$L = P^*(QP) + R , \tag{42}$$

the equation (32) can be rewritten as

$$J[\boldsymbol{u}] = (\boldsymbol{L}(\boldsymbol{u}), \boldsymbol{u})_r + (\boldsymbol{h}, \boldsymbol{u})_r + \alpha . \tag{43}$$

The operators $\boldsymbol{P}^*(\boldsymbol{QP})$ and $\boldsymbol{R}$ are self-adjoint and strictly positive, hence the operator

$$\boldsymbol{L}^* = \boldsymbol{L} , \tag{44}$$

and $\boldsymbol{L}$ is strictly positive. Therefore

$$(\boldsymbol{L}(\boldsymbol{u}), \boldsymbol{u})_r = (\boldsymbol{Q}\,\boldsymbol{P}(\boldsymbol{u}), \boldsymbol{P}(\boldsymbol{u}))_n + (\boldsymbol{Ru}, \boldsymbol{u})_r \geqq (\boldsymbol{Ru}, \boldsymbol{u})_r \geqq M\|\boldsymbol{u}\|_r^2 \tag{45}$$

which hold for any $\boldsymbol{u} \in H_r$.

The minimization problem is to find $\boldsymbol{u}^0 \in U$ such that

$$J[\boldsymbol{u}^0] = \min_{\boldsymbol{u}\in U} J[\boldsymbol{u}] . \tag{46}$$

Theorem 1. *There is one and only one element* $\mathbf{u}^0$, *which is solution of the problem demonstrated above.*

Proof. a) Prove the existence. Functional $J[\boldsymbol{u}] \geqq 0$ for any $u \in H_r$, hence exists its nonnegative infimum on U. From property of the infimum, there exists a sequence $\{\boldsymbol{u}_i\}$ of elements of U such that

$$\lim_{i\to\infty} J[\boldsymbol{u}_i] = \inf_{\boldsymbol{u}\in U} J[\boldsymbol{u}] . \tag{47}$$

Now we shall show that the sequence $\{\boldsymbol{u}_i\}$ converges strongly to an element $\boldsymbol{u}^0 \in H_r$.

In order to achieve the above desired aim, we shall use two following addition equations, which hold for arbitrarily chosen elements $\boldsymbol{u}$ and $\boldsymbol{v}$ of H_r

$$(\boldsymbol{R}(\boldsymbol{u} - \boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v})_r + (\boldsymbol{R}(\boldsymbol{u} + \boldsymbol{v}), \boldsymbol{u} + \boldsymbol{v})_r = 2\{(\boldsymbol{Ru}, \boldsymbol{u})_r + (\boldsymbol{Rv}, \boldsymbol{v})_r\} , \tag{48}$$

$$\left(\boldsymbol{Q}\left(\boldsymbol{P}\left(\frac{\boldsymbol{u}+\boldsymbol{v}}{2}\right) + \boldsymbol{g}\right), \boldsymbol{P}\left(\frac{\boldsymbol{u}+\boldsymbol{v}}{2}\right) + \boldsymbol{g}\right)_n + \left(\boldsymbol{QP}\left(\frac{\boldsymbol{u}-\boldsymbol{v}}{2}\right), \boldsymbol{P}\left(\frac{\boldsymbol{u}-\boldsymbol{v}}{2}\right)\right) = \tag{49}$$

$$= \tfrac{1}{2}\{(\boldsymbol{Q}(\boldsymbol{P}(\boldsymbol{u}) + \boldsymbol{g}), \boldsymbol{P}(\boldsymbol{u}) + \boldsymbol{g})_n + (\boldsymbol{Q}(\boldsymbol{P}(\boldsymbol{v}) + \boldsymbol{g}), \boldsymbol{P}(\boldsymbol{v}) + \boldsymbol{g})_n\} .$$

After some simply changes of these above equations it follows

$$(\boldsymbol{Q}\,\boldsymbol{P}(\boldsymbol{u} - \boldsymbol{v}), \boldsymbol{P}(\boldsymbol{u} - \boldsymbol{v}))_n + (\boldsymbol{R}(\boldsymbol{u} - \boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v})_r = \tag{50}$$

$$= 2(J[\boldsymbol{u}] + J[\boldsymbol{v}]) - 4J\left[\frac{\boldsymbol{u}+\boldsymbol{v}}{2}\right].$$

We see that

$$(\boldsymbol{Q}\,\boldsymbol{P}(\boldsymbol{u} - \boldsymbol{v}), \boldsymbol{P}(\boldsymbol{u} - \boldsymbol{v}))_n + (\boldsymbol{R}(\boldsymbol{u} - \boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v})_r \geqq \tag{51}$$

$$\geqq (\boldsymbol{R}(\boldsymbol{u} - \boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v})_r \geqq M\|\boldsymbol{u} - \boldsymbol{v}\|^2 .$$

Using the relations

$$(52)\quad 0 \leqq M\|\boldsymbol{u}_j - \boldsymbol{u}_i\|^2 \leqq (\boldsymbol{Q}\,\boldsymbol{P}(\boldsymbol{u}_i - \boldsymbol{u}_j), \boldsymbol{P}(\boldsymbol{u}_i - \boldsymbol{u}_j))_n + (\boldsymbol{R}(\boldsymbol{u}_i - \boldsymbol{u}_j), \boldsymbol{u}_i - \boldsymbol{u}_j)_r = 2(J[\boldsymbol{u}_i] + J[\boldsymbol{u}_j]) - 4J\left[\frac{\boldsymbol{u}_i + \boldsymbol{u}_j}{2}\right]$$

and because of the fact that the set U is convex and the points $\boldsymbol{u}_i$, $\boldsymbol{u}_j$ belonging to U, the point $(\boldsymbol{u}_i + \boldsymbol{u}_j)/2$ also belong to U, therefore

$$(53)\quad J\left[\frac{\boldsymbol{u}_i + \boldsymbol{u}_j}{2}\right] \geqq \inf_{\boldsymbol{u}\in U} J[\boldsymbol{u}]\,.$$

From (52) and (53) we obtain

$$(54)\quad 2(J[\boldsymbol{u}_i] + J[\boldsymbol{u}_j]) - 4J\left[\frac{\boldsymbol{u}_i + \boldsymbol{u}_j}{2}\right] \leqq 2(J[\boldsymbol{u}_i] + J[\boldsymbol{u}_j]) - 4\inf_{\boldsymbol{u}\in U} J[\boldsymbol{u}]\,.$$

Let i, j tend to infinity, then the right-hand side of the above inequality tends to zero, and hence

$$(55)\quad \|\boldsymbol{u}_i - \boldsymbol{u}_j\|_r \to 0 \quad \text{for} \quad i, j \to \infty\,,$$

the set U in the space H_r is complete, there is a sequence $\{\boldsymbol{u}_i\}$ in U such that

$$(56)\quad \lim_{i\to\infty} \boldsymbol{u}_i = \boldsymbol{u}^0$$

and U is a closed set in H_r, hence $\boldsymbol{u}^0 \in H_r$ and

$$(57)\quad \lim_{i\to\infty} J[\boldsymbol{u}_i] = J[\boldsymbol{u}^0]\,,$$

$$(58)\quad J[\boldsymbol{u}^0] = \inf_{\boldsymbol{u}\in U} J[\boldsymbol{u}] = \min_{\boldsymbol{u}\in U} J[\boldsymbol{u}]\,.$$

b) Prove the uniqueness. We suppose that, $\boldsymbol{u}^0$ and $\boldsymbol{v}^0$ are two distinct optimal points. Since the strictly convexity of the functional $J[\boldsymbol{u}]$ yields, for any $\theta \in (0, 1)$

$$(59)\quad J[\theta\boldsymbol{u}^0 + (1-\theta)\,\boldsymbol{v}^0] < \theta\, J[\boldsymbol{u}^0] + (1-\theta)\, J[\boldsymbol{v}^0] = J[\boldsymbol{u}^0]$$

$$J[\boldsymbol{u}^0] = \min_{\boldsymbol{u}\in U} J[\boldsymbol{u}]$$

which is the desired contradition.

Now we shall find the necessary and the sufficient condition for an optimal control.

We see that the last term of the right-hand side of (43) is a constant, minimizing $J[\mathbf{u}]$ is equivalent to minimizing

(60) $$I[\mathbf{u}] = J[\mathbf{u}] - \alpha = (\mathbf{L}(\mathbf{u}), \mathbf{u})_r + (\mathbf{h}, \mathbf{u})_r \,.$$

We may expand $I[\mathbf{u}]$ to

(61) $$I[\mathbf{v}] = I[\mathbf{u}] + (2\mathbf{L}(\mathbf{u}) + \mathbf{h}, \mathbf{v} - \mathbf{u})_r + (\mathbf{L}(\mathbf{v} - \mathbf{u}), \mathbf{v} - \mathbf{u})_r \,.$$

The operator $\mathbf{L}$ is positive, the third term on the right-hand side of the equation (60) is non-negative, therefore

(62) $$I[\mathbf{v}] - I[\mathbf{u}] \geqq (2\mathbf{L}(\mathbf{u}) + \mathbf{h}, \mathbf{v} - \mathbf{u})_r \,.$$

From here we shall obtain the necessary and sufficient condition for a control to be optimal.

Theorem 2. *The element $\mathbf{u}^0$ in the set U is optimal of the problem expressed above, i.e.*

(63) $$I[\mathbf{u}^0] = \min_{\mathbf{u} \in U} I[\mathbf{u}] \,,$$

if and only if the following inequality holds for any $\mathbf{u} \in U$

(64) $$(2\mathbf{L}(\mathbf{u}^0) + \mathbf{h}, \mathbf{u}^0)_r \leqq (2\mathbf{L}(\mathbf{u}^0) + \mathbf{h}, \mathbf{u})_r \,.$$

Proof. a) To prove the necessity, let $\mathbf{q}$ be an element of U, such that

(65) $$(2\mathbf{L}(\mathbf{u}^0) + \mathbf{h}, \mathbf{q})_r < (2\mathbf{L}(\mathbf{u}^0) + \mathbf{h}, \mathbf{u}^0)_r \,.$$

From the relation (61), for any point $\mathbf{u}$, which is on the line-segment between $\mathbf{q}\mathbf{u}^0$, it follows

(66) $$\begin{aligned} I[\theta \mathbf{q} + (1 - \theta)\, \mathbf{u}^0] &= I[\mathbf{u}^0 + \theta(\mathbf{q} - \mathbf{u}^0)] = \\ &= I[\mathbf{u}^0] + \theta(2\mathbf{L}(\mathbf{u}^0) + \mathbf{h}, \mathbf{q} - \mathbf{u}^0) + \\ &+ \theta^2(\mathbf{L}(\mathbf{q} - \mathbf{u}^0), \mathbf{q} - \mathbf{u}^0) \,. \end{aligned}$$

The operator $\mathbf{L}$ is bounded, the set U also is bounded hence $(\mathbf{L}(\mathbf{q} - \mathbf{u}^0), \mathbf{u} - \mathbf{u}^0)$ is bounded and therefore it exists a positive number $K > 0$ such that

(67) $$(\mathbf{L}(\mathbf{q} - \mathbf{u}^0), \mathbf{q} - \mathbf{u}^0)_r \geqq \|\mathbf{L}\| \, \|\mathbf{q} - \mathbf{u}^0\|^2 \leqq K \,.$$

By choosing the value of θ sufficiently small, the value of the right-hand side of the equation (66) may be smaller than the first term $I[\mathbf{u}^0]$, due to the convexity of the set U the point $\theta \mathbf{q} + (1 - \theta)\, \mathbf{u}^0$ belongs to the set U, which is a desired contradiction.

b) The sufficiency is evident. We shall use the inequality

$$I[u] - I[u^0] \geqq (2L(u^0) + h, u - u^0)_r \geqq 0 \,, \tag{68}$$

hence $I[u] \geqq I[u^0]$ for any $u \in U$.

3. ITERATIVE ALGORITM FOR FINDING THE OPTIMAL SOLUTION

For the first approximation $u_0 \in U$ can be chosen arbitrarily. Suppose that $u_k \in U$ of the k-th cycle has been found. Then one cycle of iteration is as follows:

(i) We compute $2L(u_k) + h$, where L and h are given by (42) and (41).

If $2L(u_k) + h = 0$, then the optimality condition is satisfied for $u^0 = u_k$. We set $u_k = u^0$ and the iteration terminates.

If otherwise go to (ii).

(ii) We shall find $v_k \in U$ such that

$$(2L(u_k) + h, v_k)_r = \min_{u \in U} (2L(u_k) + h, u)_r \,. \tag{69}$$

If the following equation holds

$$(2L(u_k) + h, u_k)_r = (2L(u_k) + h, v_k)_r \,, \tag{70}$$

then the optimality condition is satisfied and let $u_k = u^0$, the iteration terminates. If otherwise go to (iii).

(iii) We shall find θ_k $(0 \leqq \theta_k \leqq 1)$ such that

$$I[(1 - \theta_k)\, u_k + \theta_k v_k] = \min_{0 \leqq \theta \leqq 1} I[(1 - \theta)\, u_k + \theta v_k] \tag{71}$$

and calculate $u_{k+1} = (1 - \theta_k)\, u_k + \theta_k v_k$. Go back to the step (i) of the following cycle.

The sequence $\{u_j\}$ of control, which is founded above, converges to an optimal control u^0.

Now we shall show the convergence of the procedure demonstrated above. We shall prove $I[u_k] \to I[u^0]$ and $\|u^0 - u_k\| \to 0$ as $k \to \infty$, provided that the iteration does not terminate in a finite number of cycles.

The sequence $\{I[u_k]\}$ is the monotone sequence, that is $I[u_{k+1}] < I[u_k]$ for any k.

In fact, for $0 \leqq \theta \leqq 1$ the following inequality holds

$$\begin{aligned} I[u_{k+1}] \leqq I[(1 - \theta)\, u_k + \theta v_k] &= I[u_k + \theta(v_k - u_k)] = \\ &= I[u_k] + \theta(2L(u_k) + h, v_k - u_k)_r + \\ &+ \theta^2(L(v_k - u_k), v_k - u_k)\,. \end{aligned} \tag{72}$$

As the iteration does not terminate at the k-th cycle, as above there is

$$(2\mathbf{L}(\mathbf{u}_k) + \mathbf{h}, \mathbf{v}_k - \mathbf{u}_k) < 0 . \tag{73}$$

Due to the bounded set U and the bounded operator $\mathbf{L}$ hence the value of the right-hand side of (66) can be made less than the value of the first term $I[\mathbf{u}_k]$ by choosing the value of θ sufficiently small and

$$I[\mathbf{u}_{k+1}] < I[\mathbf{u}_k] .$$

Now we show that the sequence $\{I[\mathbf{u}_k]\}$ converges to the minimum $I[\mathbf{u}^0]$.
From the relations above, we see that:

$$I[\mathbf{u}_k] - I[\mathbf{u}^0] < (2\mathbf{L}(\mathbf{u}_k) + \mathbf{h}, \mathbf{u}_k - \mathbf{u}^0) \leqq \leqq (2\mathbf{L}(\mathbf{u}_k) + \mathbf{h}, \mathbf{u}_k - \mathbf{v}_k) . \tag{75}$$

We assume that there is a positive number Δ such that

$$I[\mathbf{u}_k] - I[\mathbf{u}^0] \geqq \Delta > 0 \quad \text{for any} \quad k . \tag{76}$$

The bounded operator $\mathbf{L}$ is defined in the bounded set U, so there exists a number C such that

$$(\mathbf{L}(\mathbf{v}_k - \mathbf{u}_k), \mathbf{v}_k - \mathbf{u}_k)_r \leqq \|\mathbf{L}\| \, \|\mathbf{v}_k - \mathbf{u}_k\|^2 \leqq C . \tag{77}$$

Then, using the relations (72), (75), (76), (77), it follows

$$I[\mathbf{u}_k] - I[\mathbf{u}_{k+1}] \geqq \theta\Delta - \theta^2 C . \tag{78}$$

Therefore choosing $\theta = \theta_0 = \min(\Delta/2C, 1)$ we shall have $\theta_0 C = \min(\Delta/2, C)$; it follows

$$\Delta - \theta_0 C = \max\left(\Delta - \frac{\Delta}{2}, \Delta - C\right) = = \max(\Delta/2, \Delta - C) . \tag{79}$$

From $I[\mathbf{u}_k] - I[\mathbf{u}_{k+1}] \geqq \theta\Delta - \theta^2 C$ yields

$$I[\mathbf{u}_k] - I[\mathbf{u}_{k+1}] \geqq \theta_0(\Delta - \theta_0 C) = \theta_0 \max\left(\frac{\Delta}{2}, \Delta - C\right). \tag{80}$$

a) If $C \geqq \Delta/2$, it is that $\Delta - C \leqq C$, it follows

$$\Delta - C \leqq \Delta - \frac{\Delta}{2} = \frac{\Delta}{2}$$

and

$$\max\left(\frac{\Delta}{2}, \Delta - C\right) = \frac{\Delta}{2} = \min\left(\frac{\Delta}{2}, C\right). \tag{81}$$

b) If $C \leqq \Delta/2$, it is that $\Delta - C \geqq C$, it follows

$$\Delta - C \geqq \Delta - \frac{\Delta}{2} = \frac{\Delta}{2}$$

and

$$\max\left(\frac{\Delta}{2}, \Delta - C\right) = \Delta - C \geqq \min\left(\frac{\Delta}{2}, C\right). \tag{82}$$

From the relations (80), (81), (82) follows the inequality

$$I[\boldsymbol{u}_k] - I[\boldsymbol{u}_{k+1}] \geqq \theta_0 \min\left(\frac{\Delta}{2}, C\right) > 0\,. \tag{83}$$

The constant $\theta_0 \min(\Delta/2, C)$ is independent of k, therefore, the above inequality implies $I[\boldsymbol{u}_k] \to -\infty$ if $k \to \infty$, which is the desired contradiction.

Thus, the convergence of $\{I[\boldsymbol{u}_k]\}$ to the minimum has been shown.

According to the relations (61) and (64) it follows

$$I[\boldsymbol{u}_k] - I[\boldsymbol{u}^0] \geqq (\boldsymbol{L}(\boldsymbol{u}_k - \boldsymbol{u}^0), \boldsymbol{u}_k - \boldsymbol{u}^0)_r \geqq M\|\boldsymbol{u}^0 - \boldsymbol{u}_k\|_r^2\,. \tag{84}$$

From (84) we see that

$$I[\boldsymbol{u}_k] \to I[\boldsymbol{u}^0] \quad \text{implies} \quad \|\boldsymbol{u}^0 - \boldsymbol{u}_k\| \to 0\,.$$

This completes the proof of convergence.

4. ILLUSTRATIVE EXAMPLE

We shall consider the linear control system defined by the system differential equation

$$\begin{aligned} \dot{x}_1(t) &= x_2(t)\,, \\ \dot{x}_2(t) &= u(t) \end{aligned} \tag{85}$$

with initial state

$$\begin{aligned} x_1(0) &= 1\,, \\ x_2(0) &= 1 \end{aligned}$$

and – for terminal time $T = 4$ – terminal state

$$\begin{aligned} x_1(T) &= 0\,, \\ x_2(T) &= 0\,. \end{aligned}$$

It is required to choose an appropriate admissible control $u(t)$ so that the performance index,

(86) $$J[u] = \tfrac{1}{2} \int_0^T [x_1^2(t) + u^2(t)] \, dt$$

is minimized, subject to the constraint $|u(t)| \leqq 1$ for $0 \leqq t \leqq 4$.

For the system (85), the additional conditions, we have

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$\mathbf{x}(t) = \mathbf{\Phi}(t) \, \mathbf{x}(0) + \mathbf{\Phi}(t) \int_0^t \mathbf{\Phi}^{-1}(\tau) \, \mathbf{B}(\tau) \, u(\tau) \, d\tau \, .$$

where

$$\mathbf{\Phi}(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{\Phi}^{-1}(\tau) = \begin{bmatrix} 1 & -\tau \\ 0 & 1 \end{bmatrix}.$$

For the performance index, we have

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{R} = \tfrac{1}{2} \, .$$

According to the equations (16) and (30) we have respectively

$$\mathbf{W}(t, \tau) = \begin{bmatrix} -\tau + t \\ 1 \end{bmatrix},$$

$$\mathbf{W}^{\mathrm{T}}(\tau, t) = [\tau - t, 1] \, .$$

According to the equations (40) ÷ (43), the performance index is of the form for $r = 1$

$$J[u] = (L(u), u)_1 + (h, u)_1 + \alpha$$

where

$$L(u) = (P^*(\mathbf{Q}P) + R)(u) \, .$$

$$P^*(\mathbf{Q} \, P(u)) = \int_t^4 (t_1 - t) \int_0^{t_1} \tfrac{1}{2} (t_1 - \tau) \, u(\tau) \, d\tau \, dt_1 \, ,$$

$$L(u) = \frac{1}{2} \int_t^4 (t_1 - t) \int_0^{t_1} (t_1 - \tau) \, u(\tau) \, d\tau \, dt_1 + \tfrac{1}{2} \, u(t) \, ,$$

$$h = 2P^*(\mathbf{Q}\mathbf{g}) \, ,$$

$$g = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix},$$

$$h = \frac{t^3}{6} + \frac{t^2}{2} - 12t + \frac{88}{3},$$

$$\alpha = 20,6666.$$

Table 1.

Results of iterations

N^0 of cycles	$J[u_k]$	$(2Lu^n + h,\ u_k - v^n)$
1	20,666666	42,666666
2	6,718735	4,531717
3	4,009219	2,224769
4	3,383828	1,076131

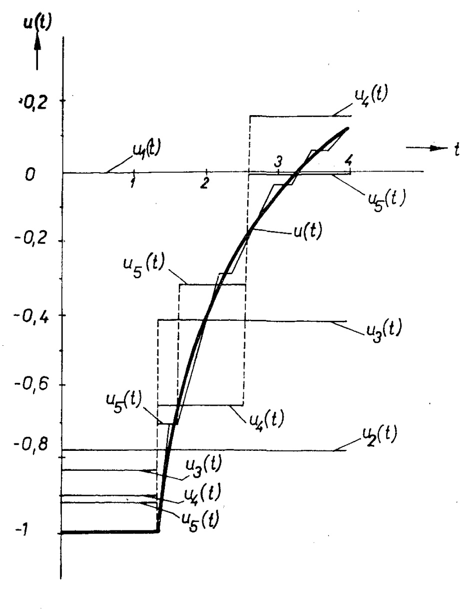

Fig. 1.

The computational results obtained after four cycles are shown in the table 1 and Fig. 1.

Table 1 indicates the minimum value $J[u] = 3{,}383828$ and the error $J[u_k] - J[u^0]$ after four iterations.

The derived sequence of controls $u_k(t)$ is plotted in Fig. 1.

(Received April 11, 1972.)

REFERENCES

[1] W. M. Wonham and C. D. Johnson: Optimal bang-bang control with quadratic performance index. Joint automatic control conference, Minnessota. 1963, 101—112.

[2] T. Fujisawa and Y. Yasuda: Optimization of linear control systems with amplitude constraint and quadratic performance index. J. of Math. analysis and applications *19* (1967), 386—400.

[3] Y. Sakawa: On a solution of an optimization problem in linear systems with quadratic performance index. J. Siam control *4* (1966), 2, 382—395.

[4] M. Athans and P. L. Falb: Optimal control. An introduction to the Theory and Its applications. Mc. Graw-Hill Book Co., New York, St. Louis, San Francisco, Toronto, London, Sydney 1966.

[5] A. Tuzar: Numerické řešení úloh optimálního řízení soustav s kvadratickým integralním kritériem. Report ÚTIA — ČSAV 1970.

Ing. Nguyen Canh, Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Vyšehradská 49, 128 48 Praha 2. Czechoslovakia.