数理解析研究所講究録
第 11 巻 1966 年 22-45

ON A SOLUTION OF AN OPTIMIZATION PROBLEM IN LINEAR SYSTEMS WITH QUADRATIC PERFORMANCE INDEX

YOSHIYUKI SAKAWA

Department of Electrical Engineering
Kyoto University, Kyoto, Japan

1. Introduction.

We consider a linear control system defined by

$$dx/dt = A(t)x(t) + B(t)u(t), \tag{1}$$

where $x(t)$ is an n-dimensional state vector, $u(t)$ is an r-dimensional control vector, and $A(t)$ and $B(t)$ are $n \times n$ and $n \times r$ matrices which are continuous in the time t. Each component $u_i(t)$ of the control vector is assumed to be constrained as

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

The control $u(t)$, $0 \leq t < \infty$, will be called an admissible control if it is measurable and it satisfies the constraints (2).

Optimization of (1), subject to the constraints (2), for a quadratic performance index has been studied by several authors [1] - [4]. Letov [1] discussed the problem using the classical calculus of variations. Wonham, Johnson and Rekasius [2] - [4] used the Hamilton-Jacobi equation for analyzing the problem. Chang [5] showed, under fairly strong conditions, that there exists a unique optimal control for any choice of the initial condition. This paper treats the problem by using a different mathematical procedure from those mentioned above. Since the state variables are expressed, by integrating the linear differential equation (1), in a linear form in the control functions, the quadratic performance index can be expressed as a quadratic functional of the control functions. Thus, we are required to minimize the quadratic functional under the constraints (2). This problem can be considered as an infinite-dimensional nonlinear programming problem. By using the generalized Kuhn-Tucker theorem in nonlinear programming, we derive a system of nonlinear integral equations as a necessary and sufficient condition for the optimal control. The existence and the uniqueness of the solution of the integral equations are studied. Successive approximations for the solution of the integral equations are shown also.

2. Formulation of the optimization problem in the Hilbert space.

The solution of (1) with initial value $x(0) = x_0$ is given by

$$x(t) = X(t)x_0 + X(t)\int_0^t X^{-1}(s)B(s)u(s)ds, \tag{3}$$

where $X(t)$, the fundamental matrix, satisfies

$$dX(t)/dt = A(t)X(t), \quad X(0) = I \text{ (identity matrix)}. \tag{4}$$

The matrix $X(t)$ is also called the transition matrix. The performance index to be used in this paper is the generalized quadratic error criterion [6]. Let $x_d(t)$ be an n-dimensional desired state vector. Let us also define the error vector to be the difference between the desired state and the actual state, i.e.

$$e(t) = x_d(t) - x(t).$$

Using (3), we can write

$$(5) \qquad e(t) = g(t) - \int_0^t W(t,s)u(s)ds,$$

where

$$(6) \qquad \begin{aligned} g(t) &= x_d(t) - X(t)x_0, \\ W(t,s) &= X(t)X^{-1}(s)B(s). \end{aligned}$$

Clearly, $W(t,s)$ is an $n \times r$ matrix.

The performance index is defined as

$$(7) \qquad I(u(t)) = \int_0^T \left\{ e^*(t)Q(t)e(t) + u^*(t)Cu(t) \right\} dt,$$

where $Q(t)$ is an $n \times n$ positive semi-definite symmetric matrix which is continuous in the time t, C 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 a vector. The matrix $Q(t)$ is usually taken to be a diagonal matrix with non-negative constant elements. The problem is then to choose an appropriate admissible control vector $u(t)$ so that the performance index is minimized.

In this paper, we use notations of functional analysis [6]-[10]. Let H_1 be a real Hilbert space of n-dimensional functions square integrable over [0,T], and H_2 be a real Hilbert space of r-dimensional functions square integrable over [0,T]. Then, the state vector $x(t)$, $0 \leq t \leq T$, will be in H_1 and the control vector $u(t)$ can be taken in H_2. Let us denote the inner product of two n-dimensional vectors x and y in the Hilbert space H_1 by $(x,y)_1$, which is defined by

$$(x,y)_1 = \int_0^T x^*(t)y(t)dt = \int_0^T y^*(t)x(t)dt.$$

In the same way, let us denote the inner product of two r-dimensional vectors u and v in H_2 by $(u,v)_2$. Then, the performance index (7) can be written as

(8) $$I(u) = (e,Qe)_1 + (u,Cu)_2.$$

We define a linear integral operator L on H_2 by

(9) $$Lu = \int_0^t W(t,s)u(s)ds \quad (0 \leq t \leq T),$$

which maps H_2 into H_1. Since $W(t,s)$ is continuous on the domain $0 \leq t,\ s \leq T$, it is obvious that the linear operator L is bounded and $Lu \in H_1$. The performance index (8) is rewritten as

(10) $$I(u) = (g-Lu,\ Qg-QLu)_1 + (u,Cu)_2.$$

Eq. (10) can be expanded to give

(11) $$I(u) = (g,Qg)_1 - 2(Qg,Lu)_1 + (Lu,QLu)_1 + (u,Cu)_2.$$

Let L^* now be the adjoint operator of L, then L^* maps H_1 into H_2 and satisfies the relation

$$(x, Lu)_1 = (L^*x, u)_2,$$

where $x \in H_1$ and $u \in H_2$. Eq. (11) can thus be written as

(12) $$I(u) = (g, Qg)_1 - 2(L^*Qg, u)_2 + (L^*QLu, u)_2 + (Cu, u)_2.$$

It can be proved, as shown in the appendix, that

(13) $$L^*Qg = \int_0^T W^*(t,s)Q(t)g(t)dt, \qquad L^*QLu = \int_0^T Y(s, \tau)u(\tau)d\tau,$$

where

(14) $$Y(s, \tau) = \int_{\max(s,\tau)}^{T} W^*(t,s)Q(t)W(t, \tau)dt.$$

Evidently, $Y(s, \tau)$ is an $r \times r$ continuous matrix and $Y^*(s, \tau) = Y(\tau, s)$. Since

$$(L^*QL)^* = L^*QL,$$

the linear bounded operator L^*QL on H_2 into H_2 is self-adjoint. Moreover, since

(15) $$(L^*QLu, u)_2 = (QLu, Lu)_1 \geq 0$$

for arbitrary $u \in H_2$, the operator L^*QL is positive.

Defining such a new operator R as

(16) $$R = L^*QL + C,$$

(12) can be written as

(17) $$I(u) = (Ru, u)_2 - 2(L^*Qg, u)_2 + (g, Qg)_1.$$

It is clear that the operator R on H_2 into H_2 is self-adjoint and positive-definite.

3. Reduction of the optimization problem to a system of integral equations.

The constraints (2) can be written as

$$1 - u_i^2(t) \geq 0 \quad (i = 1, 2, \ldots, r). \tag{18}$$

Thus, the problem is to minimize (17), the quadratic functional of $u(t)$, under the constraints (18). This problem can be considered as an infinite-dimensional nonlinear programming. For this problem, we can apply the generalized Kuhn-Tucker theorem [11] which is an extension of the Kuhn-Tucker theorem on nonlinear programming to more general topological spaces. Defining a mapping φ, which maps H_2 into H_2, by

$$\varphi(u) = \begin{bmatrix} 1 - u_1^2(t) \\ 1 - u_2^2(t) \\ \vdots \\ 1 - u_r^2(t) \end{bmatrix}, \tag{19}$$

we denote the constraints (18) as

$$\varphi(u) \geq 0. \tag{20}$$

Since the operator R on H_2 is positive-definite, it can be easily seen that the functional $I(u)$, as given by (17), is convex. It is clear that $\varphi(u)$ defined by (19) is concave. Moreover, it follows that $\varphi(0) > 0$ $(0 \in H_2)$. Therefore, from the theorem V. 3. 1. in [11], it follows that if u^0 minimizes $I(u)$ subject to $\varphi(u) \geq 0$, then there exists a non-negative r-dimensional function

(21) $$\lambda^0(t) \geq 0 \quad (\lambda^0 \in H_2)$$

such that, for the Lagrangian expression

(22) $$J(u, \lambda) = I(u) - (\lambda, \varphi(u))_2,$$

the saddle-point inequalities

(23) $$J(u, \lambda^0) \geq J(u^0, \lambda^0) \geq J(u^0, \lambda)$$

hold for all $u \in H_2$ and all $\lambda \geq 0$ $(\lambda \in H_2)$.

Conversely, from the theorem V. 1. in [11], it follows that if there exist such $u^0 \in H_2$ and $\lambda^0 \geq 0$ $(\lambda^0 \in H_2)$ that the saddle-point inequalities (23) hold for all $u \in H_2$ and all $\lambda \geq 0$ $(\lambda \in H_2)$, then

(24) $$\varphi(u^0) \geq 0$$

and, for all $u \in H_2$ satisfying $\varphi(u) \geq 0$,

(25) $$I(u^0) \leq I(u).$$

Therefore, the conditions (21) and (23) are necessary and sufficient for u^0 to be an optimal control.

Since $\lambda^0 \geq 0$ is a fixed vector in H_2, we write

$$J(u, \lambda^0) = J_0(u).$$

Let $\delta J_0(u^0;\xi)$ be the Fréchet differential of J_0 at u^0 with increment ξ ($\xi \in H_2$), which is defined by

$$\delta J_0(u^0;\xi) = \lim_{\varepsilon\to 0} \frac{J_0(u^0+\varepsilon\xi) - J_0(u^0)}{\varepsilon}, \tag{26}$$

where ε is a real number [10]. It can be shown easily that

$$J_0(u^0+\varepsilon\xi) - J_0(u^0) = \varepsilon\,\delta J_0(u^0;\xi) + \varepsilon^2 (R\xi, \xi)_2 + \varepsilon^2 \int_0^T \sum_{i=1}^{r} \lambda_i^0(t)\,\xi_i^2(t)\,dt, \tag{27}$$

where λ_i^0 and ξ_i ($i = 1,\ldots, r$) are the components of the r-dimensional functions λ^0 and ξ, respectively. Therefore, in order that the first inequality of (23), $J(u, \lambda^0) \geq J(u^0, \lambda^0)$, can be satisfied for all $u \in H_2$, it is necessary and sufficient that

$$\delta J_0(u^0;\xi) = 0 \tag{28}$$

for arbitrary $\xi \in H_2$.

Moreover, the second inequality of (23) implies that

$$(\lambda^0, \varphi(u^0))_2 \leq (\lambda, \varphi(u^0))_2$$

for all $\lambda \geq 0$ ($\lambda \in H_2$). Hence, we obtain the inequality $\varphi(u^0) \geq 0$ and

$$(\lambda^0, \varphi(u^0))_2 = 0. \tag{29}$$

Thus, the necessary and sufficient conditions for u^0 to be the optimal control are (21), (24), (28), and (29) in all. Henceforth, u^0 and λ^0 are simply written as u and λ, since no confusion may occur.

In view of the definition (26), the Fréchet differential of J_0 at u with increment ξ $(u, \xi \in H_2)$ can be evaluated as

$$\delta J_0(u; \xi) = 2(Ru, \xi)_2 - 2(L^*Qg, \xi)_2 - (\frac{\partial \varphi}{\partial u}\lambda, \xi)_2,$$

where $\partial\varphi/\partial u$ denotes an $r \times r$ diagonal matrix defined by

$$\frac{\partial \varphi}{\partial u} = \left[\frac{\partial \varphi_i}{\partial u_j} \right] = -2 \begin{bmatrix} u_1 & 0 & \dots & 0 \\ 0 & u_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & u_r \end{bmatrix}.$$

Since $\delta J_0(u; \xi)$ vanishes for arbitrary $\xi \in H_2$, it follows that

(30) $$Ru - L^*Qg - \frac{1}{2} \frac{\partial \varphi}{\partial u} \lambda = 0.$$

We set $L^*Qg = f$, then from (13) and (16)

$$Ru = \int_0^T Y(s, \tau) u(\tau) d\tau + Cu(s),$$

$$L^*Qg = \int_s^T W^*(\tau, s) Q(\tau) g(\tau) d\tau = f(s).$$

Clearly, $f(s)$ is an r-dimensional function. Write

$$Y(s, \tau) = \begin{bmatrix} y_{11}(s,\tau) & y_{12}(s,\tau) & \dots & y_{1r}(s,\tau) \\ \vdots & & & \vdots \\ y_{r1}(s,\tau) & y_{r2}(s,\tau) & \dots & y_{rr}(s,\tau) \end{bmatrix}.$$

Then, the relations (21), (29), and (30) can be written for each component as

(31) $$\lambda_i(t) \geq 0 \quad (i = 1, \ldots, r),$$

(32) $$\lambda_i(t) \left\{ 1 - u_i^2(t) \right\} = 0 \quad (i = 1, \ldots, r),$$

(33) $$\sum_{j=1}^{r} \int_0^T y_{ij}(t;s) u_j(s) ds + c_i u_i(t) + \lambda_i(t) u_i(t) = f_i(t) \quad (i = 1, \ldots, r),$$

where c_i's are the elements of the diagonal matrix C and all positive.

From (31) and (32), it follows that

$$\lambda_i(t) = 0, \quad \text{if} \quad -1 < u_i(t) < 1,$$

$$\lambda_i(t) \geq 0, \quad \text{if} \quad u_i(t) = \pm 1.$$

Hence, the relation between $u_i(t)$ and $\lambda_i(t)$ can be shown as Fig. 1 (a). The relation between $u_i(t)$ and $\lambda_i(t)u_i(t)$ and then the relation between $u_i(t)$ and $c_i u_i(t) + \lambda_i(t) u_i(t)$ can also be obtained successively from Fig. 1 (a) as shown in Fig. (b) and (c), respectively. By defining the new functions

$$v_i(t) = c_i u_i(t) + \lambda_i(t) u_i(t) \quad (i = 1, \ldots, r),$$

and denoting the relation between $u_i(t)$ and $v_i(t)$ as

$$u_i(t) = \Phi_i(v_i(t)) \quad (i = 1, \ldots, r),$$

(33) can be expressed as

$$v_i(t) + \sum_{j=1}^{r} \int_0^T y_{ij}(t,s)\,\Phi_j(v_j(s))ds = f_i(t) \tag{34}$$

$$(i = 1, 2, \ldots, r),$$

or in vector form, as

$$v(t) + \int_0^T Y(t,s)\Phi(v(s))ds = f(t). \tag{35}$$

In (34), the nonlinear function $\Phi_i(v_i)$ is shown in Fig. 2, which can be obtained from Fig. 1 (c) directly. Thus, the optimization problem has been reduced to a system of nonlinear integral equations. In other words, (34) is the necessary and sufficient condition for the optimum.

Defining such functions as

$$\left.\begin{aligned} v_i(t) - f_i(t) &= \psi_i(t) \\ \Phi_i(f_i(t) + \psi_i(t)) &= F_i(t, \psi_i(t)) \end{aligned}\right\} \quad (i = 1, \ldots, r),$$

(34) can be written as

$$\psi_i(t) + \sum_{j=1}^{r} \int_0^T y_{ij}(t,s)F_j(s, \psi_j(s))ds = 0 \tag{36}$$

$$(i = 1, \ldots, r),$$

or in vector form, as

$$\psi(t) + \int_0^T Y(t,s)F(s, \psi(s))ds = 0. \tag{37}$$

Eq. (37) is of the vector form of nonlinear integral equations of the Hammerstein type [12], [13].

4. Successive approximations for the solution of the integral equations.

Since $c_i > 0$ $(i = 1, \ldots, r)$, it is clear from Fig. 2 that the functions $F_i(t, \psi_i)$ $(i = 1, \ldots, r)$ satisfy uniformly a Lipschitz condition of the form

(38)
$$|F_i(t, \psi_i^{(1)}) - F_i(t, \psi_i^{(2)})| \leq \alpha |\psi_i^{(1)} - \psi_i^{(2)}|$$
$$(i = 1, \ldots, r),$$

where α is such a positive constant that $\alpha \geq 1/c_i$ $(i = 1, \ldots, r)$. Let us define the norm of a vector x in H_2 as

$$\|x\| = (x,x)_2^{\frac{1}{2}} = \left\{ \sum_{i=1}^{r} \int_0^T x_i^2(t)dt \right\}^{\frac{1}{2}} .$$

Moreover, let us introduce an r-dimensional function $z(t) = (z_1(t), \ldots, z_r(t))$, where the elements are defined by

$$z_i(t) = \left\{ \sum_{j=1}^{r} \int_0^T y_{ij}^2(t,s)ds \right\}^{\frac{1}{2}} \quad (i = 1, \ldots, r).$$

The functions $y_{ij}(t,s)$ $(i,j = 1, \ldots, r)$ are continuous on the domain $0 \leq t,\ s \leq T$, hence $z \in H_2$. It can be proved that if

(39)
$$\alpha \|z\| < 1,$$

then the successive approximations

(40)
$$\psi_i^{(n+1)}(t) = -\sum_{j=1}^{r} \int_0^T y_{ij}(t,s) F_j(s, \psi_j^{(n)}(s))ds$$
$$(i = 1, \ldots, r\ ;\ n = 0, 1, 2, \ldots)$$

starting, for instance, with $\psi_i^{(0)}(t) = 0$, converge to a unique solution of (36). It is obvious that the existence of a unique solution of (36) implies the existence of the unique optimal control.

In fact, from (38) it follows that

$$|\psi_i^{(n+1)}(t) - \psi_i^{(n)}(t)|$$

$$\leq \sum_{j=1}^{r} \int_0^T |y_{ij}(t,s)| |F_j(s,\psi_j^{(n)}(s)) - F_j(s,\psi_j^{(n-1)}(s))| ds$$

$$\leq \alpha \sum_{j=1}^{r} \int_0^T |y_{ij}(t,s)| |\psi_j^{(n)}(s) - \psi_j^{(n-1)}(s)| ds.$$

Furthermore, using the Schwarz inequality,

$$|\psi_i^{(n+1)}(t) - \psi_i^{(n)}(t)|$$

$$\leq \alpha \left\{ \sum_{j=1}^{r} \int_0^T y_{ij}^2(t,s) ds \right\}^{\frac{1}{2}} \left\{ \sum_{j=1}^{r} \int_0^T (\psi_j^{(n)}(s) - \psi_j^{(n-1)}(s))^2 ds \right\}^{\frac{1}{2}}$$

$$= \alpha \; z_i(t) \| \psi^{(n)} - \psi^{(n-1)} \|.$$

Thus, we obtain

$$\|\psi^{(n+1)} - \psi^{(n)}\| \leq \alpha \|z\| \|\psi^{(n)} - \psi^{(n-1)}\|. \tag{41}$$

Eq. (41) shows that the mapping defined by the right-hand side of (40) is a contraction mapping under the condition (39) [9]. Therefore, under the condition (39), we can show the existence and the uniqueness of the solution of (36).

5. Existence and uniqueness of optimal control.

In the case where $c_1 = c_2 = \ldots = c_r$, the nonlinear characteristics Φ_i $(i = 1, \ldots, r)$ shown in Fig. 2 coincide with each other. Hence, we express the characteristic as $\hat{\Phi}$. In this case, the system of nonlinear integral equations (34) can be reduced to a single integral equation with a discontinuous kernel in the basic interval $0 \le t \le rT$:

$$\hat{v}(t) + \int_0^{rT} \hat{y}(t,s)\hat{\Phi}(\hat{v}(s))ds = \hat{f}(t) \quad (0 \le t \le rT), \tag{42}$$

where

$$\hat{v}(t) = \begin{cases} v_1(t), & \text{if } 0 \le t < T \\ v_2(t-T), & \text{if } T \le t < 2T \\ \vdots & \\ v_r(t-(r-1)T), & \text{if } (r-1)T \le t \le rT, \end{cases}$$

$$\hat{f}(t) = \begin{cases} f_1(t), & \text{if } 0 \le t < T \\ f_2(t-T), & \text{if } T \le t < 2T \\ \vdots & \\ f_r(t-(r-1)T), & \text{if } (r-1)T \le t \le rT, \end{cases}$$

and

$$\hat{y}(t,s) = y_{ij}(t-(i-1)T,\ s-(j-1)T), \quad \text{if } (i-1)T \le t < iT \text{ and } (j-1)T \le s < jT \quad (i,\ j = 1,\ 2, \ldots, r). \tag{43}$$

Furthermore, defining such scalar functions as

$$\hat{v}(t) - \hat{f}(t) = \hat{\psi}(t),$$

(44)

$$\hat{\Phi}(\hat{f}(t) + \hat{\psi}(t)) = \hat{F}(t, \hat{\psi}(t)),$$

(42) can be written as

$$\hat{\psi}(t) + \int_0^{rT} \hat{y}(t,s)\hat{F}(s, \hat{\psi}(s))ds = 0 \qquad (45)$$

$$(0 \leq t \leq rT).$$

Eq. (45) is of the standard form of integral equations of the Hammerstein type [12], [13].

Hammerstein [12] proved the existence of the solution of the integral equation of the Hammerstein type, assuming that the iterated kernel

$$\hat{y}_2(t,s) = \int_0^{rT} \hat{y}(t, \tau)\hat{y}(\tau, s)d\tau$$

is continuous. However, the kernel function $\hat{y}(t,s)$ defined by (43) is not continuous, hence Hammerstein's existence theorem is not applicable to our problem. In what follows, the existence of the solution of (45) will be shown by using Krasnosel'skii's existence theorem [13], [14].

Let H be a real Hilbert space of functions square integrable over [0, rT]. The inner product is defined, as usual, by

$$(x,y) = \int_0^{rT} x(t)y(t)dt \quad (x, y \in H).$$

Let G be an operator on H defined by

$$Gx = \hat{F}(t, x(t)) \quad (x \in H), \qquad (46)$$

and K be a linear operator on H defined by

(47) $$Kx = \int_0^{rT} \hat{y}(t,s)x(s)ds \quad (x \in H).$$

Then, the integral equation (45) can be written symbolically as

(48) $$\hat{\psi} + KG\hat{\psi} = 0.$$

Since the matrix $Y(s, \tau)$ defined by (14) is continuous on the closed square domain $0 \leq s, \quad \tau \leq T$, it follows that

$$\int_0^{rT}\int_0^{rT} \hat{y}^2(t,s)dtds = \sum_{i,j=1}^{r} \int_0^{T} y_{ij}^2(t,s)dtds < \infty.$$

Therefore, the linear operator K is completely continuous [10]. From (15) it follows that

$$(Kx,x) = \int_0^{rT}\int_0^{rT} \hat{y}(t,s)x(t)x(s)dtds$$
$$= \sum_{i,j=1}^{r} \int_0^{T}\int_0^{T} y_{ij}(t,s)x(t-(i-1)T)x(s-(j-1)T)dtds \geq 0,$$

for arbitrary function $x \in H$. Hence, the operator K is positive, i.e. all its eigenvalues are positive. Moreover; the operator K is self-adjoint, i.e. $K^* = K$. Therefore, from the spectral theory of operators, the operator K can be decomposed as

(49) $$K = PP^*,$$

where P is a square root of the operator K (i.e. $P = K^{\frac{1}{2}}$) and is a positive self-adjoint completely continuous operator on H into H, and P^* is an adjoint operator of P [14]. Then, the nonlinear integral equation (48) can be written as

(50) $$\hat{\psi} + PP^*G\hat{\psi} = 0.$$

Eq. (50) is equivalent to

(51) $$\varphi + P^*GP\varphi = 0 \quad (\varphi \in H),$$

in the sense that to a solution $\varphi \in H$ of (51) there corresponds a solution $P\varphi \in H$ of (50) and, conversely, to a solution $\hat{\psi} \in H$ of (50) there corresponds a solution $P^*G\hat{\psi} \in H$ of (51). Moreover, Krasnosel'skii [14] shows that the operator $I + P^*GP$, I being an identity operator on H, is a gradient of the functional

(52) $$\Psi(\varphi) = \frac{1}{2}(\varphi, \varphi) + \int_0^{rT} dt \int_0^{P\varphi(t)} \hat{F}(t,x)dx$$

defined on H, where an operator Γ on H into H is called the gradient of the functional Ψ, if

$$\lim_{\varepsilon \to 0} \frac{\Psi(\varphi + \varepsilon\xi) - \Psi(\varphi)}{\varepsilon} = (\Gamma\varphi, \xi) \quad (\varphi, \xi \in H).$$

It is clear that the function $\hat{F}(t,x)$ satisfies the Carathéodory condition [14], i.e. it is continuous with respect to x for almost all $t \in [0, rT]$ and measurable with respect to t for all values of x.

According to Krasnosel'skii's theorem 1. 1. in chapter VI [14], if the functional (52) is increasing, i.e.

$$\lim_{\|\varphi\| \to \infty} \Psi(\varphi) = +\infty,$$

then there exists a point φ_0 in the Hilbert space H where the functional $\Psi(\varphi)$ takes on its minimum value and its gradient vanishes, i.e.

$$\varphi_0 + P^*GP\varphi_0 = 0.$$

Thus, if the functional $\Psi(\varphi)$ is increasing, then the existence of a solution of (51), and hence the existence of a solution of the fundamental equation (50) can be concluded. Since

$$\int_0^u \hat{F}(t,x)dx \le \int_0^{|u|} |\hat{F}(t,x)|dx \le |u|,$$

using the Schwarz inequality, it follows that

$$-\int_0^{rT} dt \int_0^{P\varphi(t)} \hat{F}(t,x)dx \le \int_0^{rT} |P\varphi(t)|dt \le \sqrt{rT}\,(P\varphi, P\varphi)^{\frac{1}{2}}$$
$$= \sqrt{rT}\,(K\varphi, \varphi)^{\frac{1}{2}} \le \sqrt{rT}\,M(\varphi, \varphi)^{\frac{1}{2}},$$

where

$$M = \left\{\int_0^{rT}\int_0^{rT} \hat{y}^2(t,s)dtds\right\}^{\frac{1}{2}}.$$

Consequently,

$$\Psi(\varphi) \ge \tfrac{1}{2}(\varphi, \varphi) - \sqrt{rT}\,M(\varphi, \varphi)^{\frac{1}{2}}. \tag{53}$$

Eq. (53) shows that the functional $\Psi(\varphi)$ is increasing. Thus, the existence of the solution of the nonlinear integral equation (45), and hence the existence of the optimal control has been proved.

If we assume that the positive operator K defined by (47) is further positive-definite, i.e.

$$(\varphi, K\varphi) > 0, \text{ if } \varphi \neq 0,$$

then the uniqueness of the solution of (48) can also be proved as follows. Assume that

$$\hat{\psi}_1 + KG\hat{\psi}_1 = 0, \quad \text{and} \quad \hat{\psi}_2 + KG\hat{\psi}_2 = 0. \tag{54}$$

Subtract these two equations and form the inner product with $(G\hat{\psi}_1 - G\hat{\psi}_2)$:

$$(G\hat{\psi}_1 - G\hat{\psi}_2,\ \hat{\psi}_1 - \hat{\psi}_2) + (G\hat{\psi}_1 - G\hat{\psi}_2,\ KG\hat{\psi}_1 - KG\hat{\psi}_2) = 0. \tag{55}$$

From the definition of the function $\hat{F}(t,x)$, the first term of (55) is obviously non-negative. Since the operator K is positive-definite, there is a contradiction unless $G\hat{\psi}_1 - G\hat{\psi}_2 = 0$, but in this case, from (54) we obtain $\hat{\psi}_1 = \hat{\psi}_2$. Thus, under the assumption that K is positive-definite, the uniqueness of the optimal control can be shown.

When $C = 0$ in (7), the nonlinear functions $\Phi_i (i = 1, 2, \ldots, r)$ become discontinuous as shown in Fig. 3. In Fig. 3, the vertical part of the characteristic corresponds to a singular control which takes on such continuous values as $-1 < u_i(t) < 1$. In this case, however, the existence of the solution can not be claimed since the assumed Carathéodory condition does not hold.

Acknowledgment.

The author wishes to express his gratitude to Professor C. Hayashi for his valuable suggestions.

Appendix: Derivation of Eq. (13).

From the definition (9) of the operator L, it follows that

$$(56) \qquad (L^*Qg,u)_2 = (Qg,Lu)_1 = \int_0^T dt\ g^*(t)Q(t)\int_0^t W(t,s)u(s)ds.$$

The region of integration (56) is given by $0 \le t \le T$, $0 \le s \le t$ which is equivalent to $0 \le s \le T$, $s \le t \le T$. Then, changing the order of the integration (56) yields

$$(57) \qquad (L^*Qg,u)_2 = \int_0^T \int_s^T \left\{W^*(t,s)Q(t)g(t)\right\}^* dt\ u(s)ds.$$

Eq. (57) shows that

$$(58) \qquad L^*Qg = \int_s^T W^*(t,s)Q(t)g(t)dt.$$

From (58), it follows that

$$(59) \qquad L^*QLu = \int_s^T dt\ W^*(t,s)Q(t)\int_0^t W(t,\tau)u(\tau)d\tau.$$

The region of integration (59) is given by $s \le t \le T$, $0 \le \tau \le t$ which is equivalent to $0 \le \tau \le T$, $\max(s,\tau) \le t \le T$, where $\max(s,\tau)$ denotes the maximum of s and τ. Changing the order of the integration (59) yields

$$(60) \qquad L^*QLu = \int_0^T \left\{\int_{\max(s,\tau)}^T W^*(t,s)Q(t)W(t,\tau)dt\right\} u(\tau)d\tau$$

$$= \int_0^T Y(s,\tau)u(\tau)d\tau.$$

References

(1) A. M. Letov, Analytical design of controllers, Avtomat. i Telemeh., 21 (1960), pp. 561-568.

(2) W. M. Wonham and C. D. Johnson, Optimal bang-bang control with quadratic performance index, Trans. ASME, J. Basic Engrg., 86 (1964), pp. 107-115.

(3) C. D. Johnson and W. M. Wonham, On a problem of Letov in optimal control, Trans. ASME, J. Basic Engrg., 87 (1965), pp. 81-89.

(4) Z. V. Rekasius and T. C. Hsia, On a inverse problem in optimal control, IEEE Trans. on Automatic Control, AC-9 (1964), pp. 370-375.

(5) A. Chang, An optimal regulator problem, Jour. SIAM Control, Series A, 2 (1965), pp. 220-233.

(6) H. C. Hsieh, Synthesis of optimum multivariable control systems by the method of steepest descent, IEEE Trans. on Applications and Industry, 82 (1963), pp. 125-130.

(7) A. V. Balakrishnan, An operator theoretic formulation of a class of control problems and a steepest descent method of solution, Jour. SIAM Control, series A, 1 (1963), pp. 109-127.

(8) A. V. Balakrishnan and H. C. Hsieh, Function space methods in control systems optimization, presented at the Optimum System Synthesis Conference, Dayton, Ohio, 1962.

(9) A. N. Kolmogorov and S. V. Formin, Elements of the Theory of Functions and Functional Analysis, Graylock Press, Rochester, N.Y., 1957.

(10) L. V. Kantrovich and G. P. Akirov, Functional Analysis in Normed Space, Pergamon Press, London, 1964.

(11) L. Hurwicz, Programming in linear spaces, in K. J. Arrow, L. Hurwicz and H. Uzawa, Studies in Linear and Non-linear Programming, Stanford University Press, Calif., pp. 38-102.

(12) F. G. Tricomi, Integral Equations, Interscience Publishers, Inc., New York, 1957.

(13) C. L. Dolph and G. J. Minty, On nonlinear integral Equations of the Hammerstein type, in P. M. Anselone, Ed., Nonlinear Integral Equations, Univ. of Wisconsin Press, Madison, Wis., 1964.

(14) M. A. Krasnosel'skii, Topological Methods in the Theory of Nonlinear Integral Equations, Pergamon Press, London, 1964.

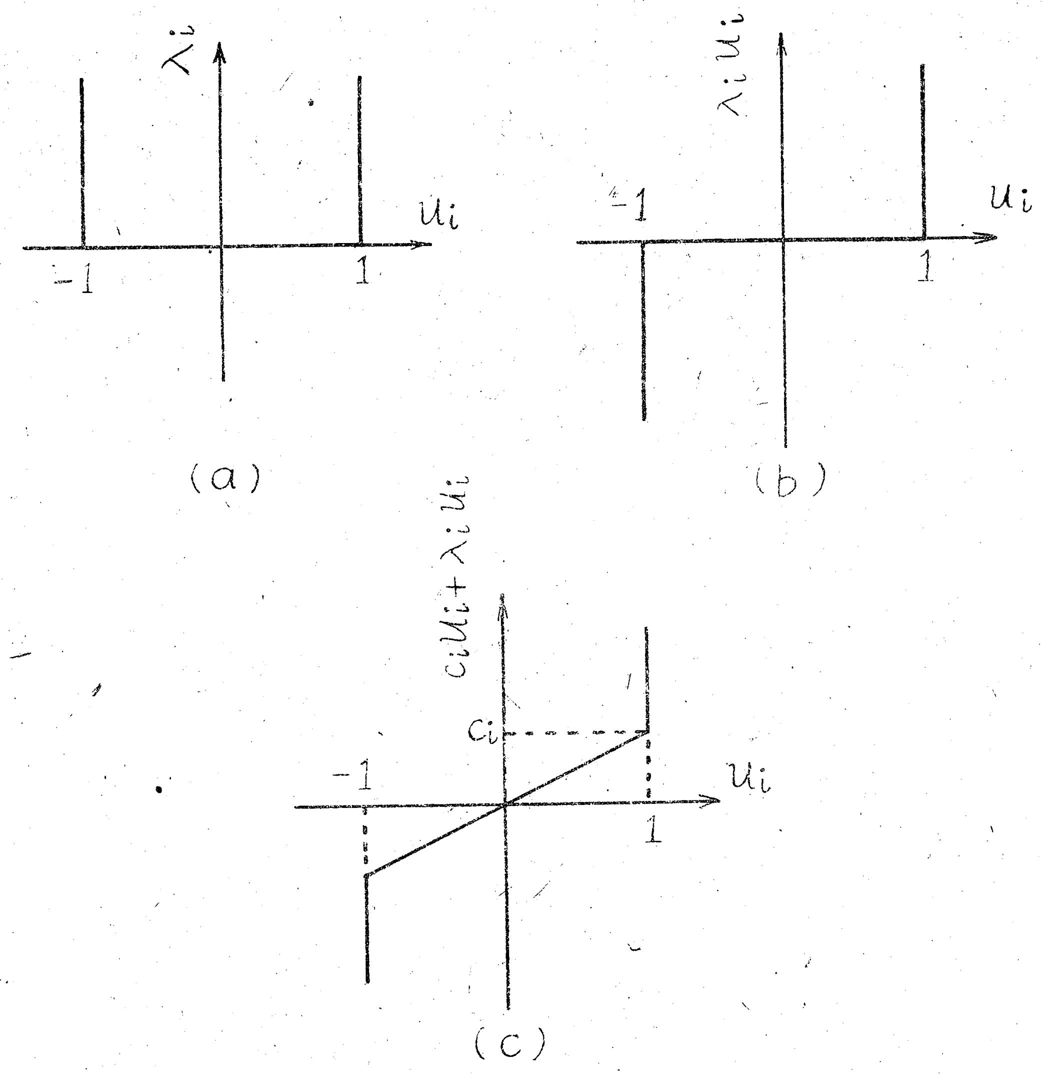

Fig. 1 Relations between the variables.

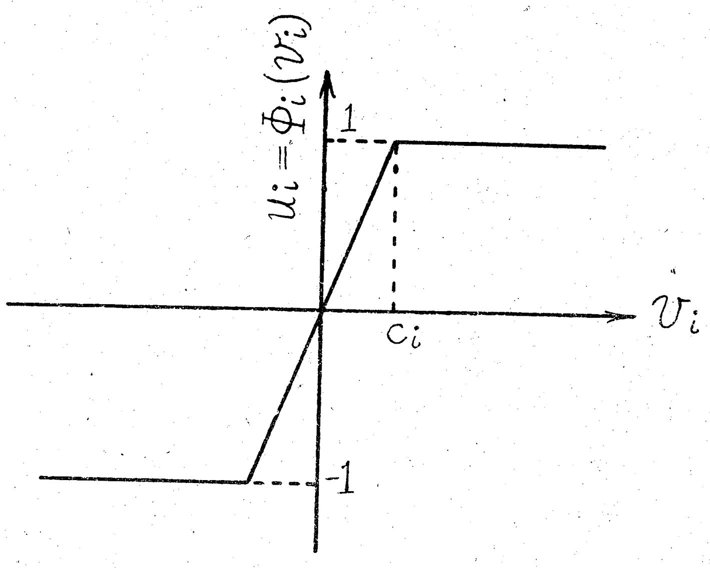

Fig. 2 Nonlinear characteristic.

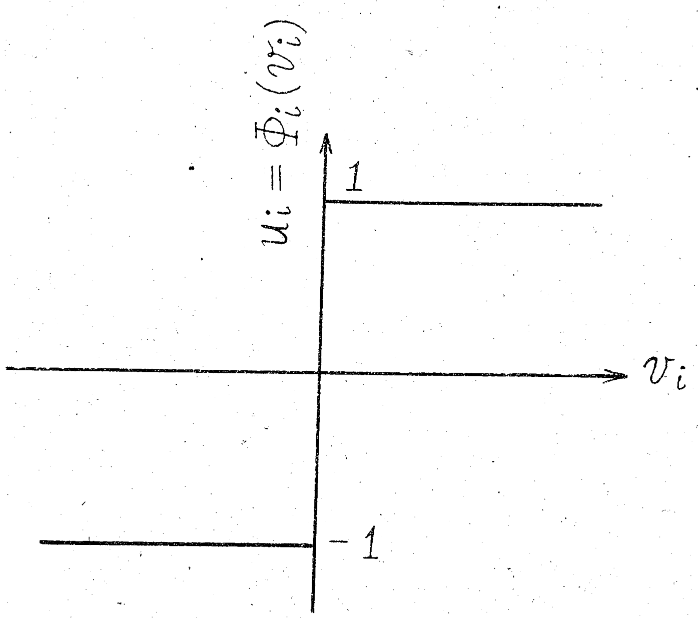

Fig. 3 Discontinuous nonlinear characteristic.