Suboptimal Design of a Nonlinear Feedback System

By

Yoshikazu NISHIKAWA*, Nobuo SANNOMIYA**, and Hidekiyo ITAKURA**

(Received March 31, 1970)

A method is developed for the approximate design of an optimal state regulator for a nonlinear system with quadratic performance index. The nonlinearity is taken to be a perturbation to the system. By making use of a power-series expansion in a small parameter, matrix equations are derived for the stepwise determination of a suboptimal feedback law. Given a polynomial nonlinearity of an arbitrary form, explicit solutions have been obtained for those matrix equations. A necessary and sufficient condition for the existence and uniqueness of the solution is also shown. Further, the performance analysis reveals the fact that the l-th order approximation in the feedback law results in the $(2l+1)$th order approximation to the optimal performance index. The method may effectively be used in a computer programmed computation.

Introduction

There have been numerous studies on the optimal feedback control of a linear dynamical system or on the optimal design of a linear regulator. On the other hand, however, relatively few works have been done on nonlinear regulator problems, because of the difficulty in determining the exact optimal feedback law.

The purpose of this paper is to present a systematic procedure for constructing a suboptimal nonlinear state regulator. The system considered contains a small nonlinearity which is characterized by an analytic function of the state. For simplicity the performance index is assumed to be quadratic. We introduce a parameter ε, called the perturbation parameter, which is associated with the nonlinearity. For a sufficiently small value of ε, a desired feedback law may be expanded and sought in a power-series form in ε. The generating solution is the unperturbed solution, which is usually obtainable by solving a Riccati type equation. Correction terms for improving the feedback law are determined in a stepwise manner by solving a sequence of linear matrix equations. Given a polynomial nonlinearity of any form, we have succeeded in obtaining definite solutions for those matrix equations. A necessary and sufficient condition for solvability is

* Department of Electrical Engineering II

** Department of Electrical Engineering

clarified in the process of the solution. In particular, it is shown that the complete controllability of the unperturbed system suffices to generate a suboptimal feedback law up to any order in ε.

Furthermore, the examination of performance clarifies the fact that, if the feedback law is optimal up to the l-th order in ε, the index of performance is minimized up to the $(2l+1)$th order. This is a generalization of the theorem given by Kokotović and Cruz[1] for linear systems.

Here some other related works should be mentioned. The method developed by Garrard et al.[2,3] is based on an approximate solution of the Hamilton-Jacobi equation, and its procedure is somewhat similar to the present one. However, since correction terms can not be determined uniquely, one is not able to obtain the best possible solution. Al'brekht[4] has developed a kind of perturbation procedure for the stepwise construction of suboptimal law, but no definite way is presented for the calculation of higher order terms. Although the method of instantaneous linearization by Pearson[5] gives the exact solution to the one-dimensional problem[6], it is generally not effective for the high dimension system.

The present method seems quite promising for the suboptimal design of a large class of nonlinear regulators. As the procedure is completely systematic and includes only linear calculations in the correcting steps, it could effectively be used in a computer programmed computation. It is also noted that the method offers an efficient way for suboptimal design of a class of large-scale systems[7].

Statement of the Problem

Consider dynamical systems governed by the equation

$$\dot{x} = A(t)x + \varepsilon f(x, t) + B(t)u \tag{1}$$

where x is the n-dimensional state vector, u the m-dimensional control vector; A and B are $n\times n$- and $n\times m$- matrices, respectively, continuous in t. The n-vector function $f(x, t)$, continuous in t, is an analytic function in x satisfying $f(0, t)=0$ for any t. ε is a small scalar parameter. Here and throughout the paper a dot over a quantity denotes differentiation with respect to time t.

The problem is to find a feedback control law $u(x, t)$ for which the quadratic index of performance

$$J(x, \tau) = \frac{1}{2}\int_{\tau}^{T} [x'Q(t)x + u'R(t)u]\,dt \tag{2}$$

is minimized. In Eq. (2), Q is a symmetric positive semidefinite matrix and R is a symmetric positive definite matrix, both continuous in t. A prime denotes transposition of a vector or a matrix.

Determination of a Suboptimal Feedback Control

It is well known that, for the unperturbed linear system, i.e., for the system (1) with $\varepsilon=0$, the optimal feedback control is given by[8]

$$u = -R^{-1}B'P(t)x \tag{3}$$

where the matrix P is the solution of the Riccati equation

$$\dot{P} = -PA-A'P+PEP-Q, \qquad P(T)=0 \tag{4}$$

with

$$E = BR^{-1}B' \qquad (\text{: symmetric}) \tag{5}$$

It is easily observed that P is symmetric.

Since generally the exact determination of the true optimal control law is impossible for nonlinear systems, the present aim is to obtain a suboptimal feedback control for the nonlinear system (1). For sufficiently small value of ε, it may be reasonable to assume that the optimal feedback control be analytic in ε for all $t\in[\tau, T]$. Therefore we try to find a suboptimal control in a power-series form in ε.

The Hamiltonian of the problem is given by

$$H = \frac{1}{2}x'Qx+\frac{1}{2}u'Ru+p'(Ax+\varepsilon f+Bu) \tag{6}$$

where $p(x, t)$ is the n-dimensional costate vector satisfying the differential equation

$$\dot{p} = -\frac{\partial H}{\partial x} = -Qx-A'p-\varepsilon\left(\frac{\partial f}{\partial x}\right)'p\,, \qquad p(x, T)=0 \tag{7}$$

The matrix $(\partial f/\partial x)$ is defined in such a way that the (i, j)-element of $(\partial f/\partial x)$ is $(\partial f_i/\partial x_j)$. Due to the minimum principle, a necessary condition for the optimality is that the Hamiltonian be minimum with respect to u. Hence the optimal control is given by

$$u = -R^{-1}B'p \tag{8}$$

Here we develop the costate vector $p(x, t)$ into a power series with respect to ε:

$$p(x, t) = P(t)x+g(x, t;\varepsilon) = P(t)x+\sum_{k=1}^{\infty}\varepsilon^k g^{(k)}(x, t) \tag{9}$$

Differentiation of Eq. (9) with respect to t and use of Eqs. (1) and (8) gives

$$\dot{p} = \dot{P}x + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial g^{(k)}}{\partial t} + \left[P + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial g^{(k)}}{\partial x}\right](Ax + \varepsilon f - Ep) \tag{10}$$

Substituting Eqs. (9) and (10) into (7) and equating the coefficients of the like powers of ε separately yields a sequence of equations:

$$\varepsilon^1:\ \frac{\partial g^{(1)}}{\partial t} + \frac{\partial g^{(1)}}{\partial x} Sx + S' g^{(1)} = -\left(\frac{\partial f}{\partial x}\right)' Px - Pf \tag{11-1}$$

$$\varepsilon^k:\ \frac{\partial g^{(k)}}{\partial t} + \frac{\partial g^{(k)}}{\partial x} Sx + S' g^{(k)} = -\left(\frac{\partial f}{\partial x}\right)' g^{(k-1)} - \frac{\partial g^{(k-1)}}{\partial x} f + \sum_{i=1}^{k-1} \frac{\partial g^{(k-i)}}{\partial x} E g^{(i)}$$

$$(k = 2, 3, \cdots) \tag{11-k}$$

with the boundary condition $g^{(k)}(x, T) = 0$ $(k = 1, 2, \cdots)$. S is the matrix defined by

$$S = A - EP \tag{12}$$

The zero-order equation is identical with Eq. (4).

Successive solutions of Eqs. (11–k) with increasing k may determine the functions $g^{(k)}(x, t)$. Though Eqs. (11) are linear in $g^{(k)}$, their solutions are difficult given a general form of the perturbation f. However, if f is a polynomial in x, $g^{(k)}$ are also polynomials in x and their coefficients are exactly determined under appropriate conditions.

First we can easily establish the following theorem:

Theorem 1

If $f(x, t)$ is a polynomial of the degree r in x, the coefficients being continuous in t, $g^{(k)}(x, t)$ given by Eq. (11–k) is the polynomial of the degree $k(r-1)+1$ in x whose coefficients are continuous in t. In particular, $g^{(1)}(x, t)$ is the polynomial of the same degree as $f(x, t)$.

Secondly we show important theorems to determine the polynomial $g^{(k)}(x, t)$. These theorems exhibit important features of the solution.

Theorem 2

The matrix $(\partial g^{(k)}/\partial x)$ given by Eq. (11–k) is symmetric, and consequently there exists a scalar function $v^{(k)}(x, t)$ such that

$$g^{(k)} = \frac{\partial v^{(k)}}{\partial x} \quad \text{for} \quad t \in [\tau, T] \tag{13}$$

for every $k \geq 1$. The coefficients of the polynomial $g^{(k)}$ are determined by solving Eq. (A. 24) of *Theorem A* in Appendix.

Proof

Since the right side of Eq. (11–1) is integrable with respect to x, Eq. (11–1) is rewritten as

$$\frac{\partial g^{(1)}}{\partial t}+\frac{\partial g^{(1)}}{\partial x}Sx+S'g^{(1)} = -\frac{\partial}{\partial x}(x'Pf) \tag{14–1}$$

Equation (14–1) is of the form of Eq. (A. 1) in Appendix. Due to *Corollary A*1, $(\partial g^{(1)}/\partial x)$ is symmetric. Then the right side of Eq. (11–k) for $k=2$ is integrable with respect to x, and the equation is rewritten as

$$\frac{\partial g^{(2)}}{\partial t}+\frac{\partial g^{(2)}}{\partial x}Sx+S'g^{(2)} = \frac{\partial}{\partial x}\left[-f'g^{(1)}+\frac{1}{2}g^{(1)\prime}Eg^{(1)}\right]$$

This is again of the form of Eq. (A. 1), and consequently $(\partial g^{(2)}/\partial x)$ is symmetric. Proceeding similarly, Eq. (11–k) is rewritten as

$$\frac{\partial g^{(k)}}{\partial t}+\frac{\partial g^{(k)}}{\partial x}Sx+S'g^{(k)} = \frac{\partial}{\partial x}\left[-f'g^{(k-1)}+\frac{1}{2}\sum_{i=1}^{k-1}g^{(i)\prime}Eg^{(k-i)}\right] \tag{14–k}$$

and $(\partial g^{(k)}/\partial x)$ is known to be symmetric for every $k\geqq 3$.

When $(\partial g^{(k)}/\partial x)$ is symmetric, there exists a scalar function $v^{(k)}$ satisfying Eq. (13) [refer to, e.g., Ref. 9, pp. 14–15]. Q. E. D.

Theorem 3

If the unperturbed linear system is completely controllable and time-invariant, all the eigenvalues of $S=A-BR^{-1}B'P$ have negative real parts as $t\to-\infty$. Then, for every $k\geqq 1$, $g^{(k)}(x, t)$, the polynomial function in x, is uniformly asymptotically stable as $t\to-\infty$, relative to the polynomial $g^{(k)}(x)$ such that satisfies the equation:

$$\text{For } k=1 \qquad x'S'g^{(1)} = -x'Pf \tag{15–1}$$

$$\text{For } k\geqq 2 \qquad x'S'g^{(k)} = -f'g^{(k-1)}+\frac{1}{2}\sum_{i=1}^{k-1}g^{(i)\prime}Eg^{(k-i)} \tag{15–k}$$

The polynomial function $g^{(k)}(x)$, satisfying Eq. (15–k), exists and is unique for every $k\geqq 1$. The coefficients of the polynomial can actually be calculated by solving the linear algebraic equation (A. 26) of *Corollary A*2 in Appendix.

Theorem 3 can readily be proved due to *Corollary A*2.

Since all the equations of (11) and (15) are linear in $g^{(k)}$, a direct consequence of the foregoing theorems is:

Theorem 4

If f is an analytic function in x, $g^{(k)}(x, t)$ or $g^{(k)}(x)$ is uniquely determined under appropriate conditions. If f is zero at $x=0$, so is g.

Now we have completed the procedure for calculating a suboptimal control. The solution of Eqs. (11) or Eqs. (15) up to the order l in ε gives the suboptimal control of the l-th order. It is noted that the complete controllability of the unperturbed system suffices to yield a suboptimal control of an arbitrary order.

Performance Evaluation

In this section quality of a suboptimal control law is examined. For simplicity we confine the attention to a completely controllable time-invariant system. As shown in *Theorem* 3, when $T\to\infty$ a suboptimal feedback control is free of explicit dependence on t and is given by

$$u_s(x) = -R^{-1}B'g_s(x) \tag{16}$$

Where $g_s(x)$ is a suboptimal feedback function satisfying Eqs. (15) up to a certain order.

The value of the performance index J resulting from the feedback control of Eq. (16) would be a function of $x(\tau)$ and be written as

$$J_s = \frac{1}{2}\int_\tau^\infty (x'Qx+u_s'Ru_s)\,dt = v_s[x(\tau)] \tag{17}$$

Evidently $v_s(0)=0$ if $g_s(0)=0$. Now the task is to examine the function $v_s(x)$. Differentiation of Eq. (17) with respect to τ and use of Eqs. (1) and (16) gives:

$$x'Qx+g'Eg+2\left(\frac{\partial v}{\partial x}\right)'(Ax+\varepsilon f-Eg) = 0 \tag{18}$$

In Eq. (18) all the values are estimated at $t=\tau$. The subscript s is omitted here and throughout the following part of this section, but should be understood.

We develop g and v into power series in ε:

$$g(x) = Px+\sum_{k=1}^{\infty}\varepsilon^k g^{(k)}(x) \quad \text{and} \quad v(x) = \frac{1}{2}x'Px+\sum_{k=1}^{\infty}\varepsilon^k v^{(k)}(x) \tag{19}$$

where the matrix P is the solution of

$$PA+A'P-PEP+Q = 0 \tag{20}$$

Substitution of Eqs. (19) into (18) results in a sequence of equations:

$$\varepsilon^1:\ x'S'\left(\frac{\partial v^{(1)}}{\partial x}\right) = -x'Pf \tag{21-1}$$

$$\varepsilon^k:\ x'S'\left(\frac{\partial v^{(k)}}{\partial x}\right) = -f'\left(\frac{\partial v^{(k-1)}}{\partial x}\right)-\frac{1}{2}\sum_{i=1}^{k-1}g^{(i)\prime}Eg^{(k-i)}+\sum_{i=1}^{k-1}g^{(i)\prime}E\left(\frac{\partial v^{(k-i)}}{\partial x}\right)$$
$$(\mathrm{k} = 2, 3, \cdots) \tag{21-k}$$

Then the following theorem is able to be established :

Theorem 5

If $g(x)$ is the optimal one, i.e., if $g^{(k)}=g^{*(k)}$ for every $k\geqq 1$, where $g^{*(k)}$ is the solution of Eq. (15–k), then J is given by

$$J = J^* = \frac{1}{2}x'Px + \sum_{k=1}^{\infty} \varepsilon^k v^{*(k)}[x(\tau)] \tag{22}$$

where $v^{*(k)}(x)$ is the scalar function satisfying

$$\frac{\partial v^{*(k)}}{\partial x} = g^{*(k)} \qquad \text{for every } k\geqq 1 \tag{23}$$

Proof

The theorem can be proved by induction. First comparison of Eq. (21–1) with Eq. (15–1) gives

$$x'S'\left[\frac{\partial v^{(1)}}{\partial x} - g^{*(1)}\right] = 0$$

By virtue of *Corollary A*3 in Appendix, the above equation means that $(\partial v^{(1)}/\partial x) = g^{*(1)}$. Second we assume that $(\partial v^{(i)}/\partial x)=g^{*(i)}$ holds for $i=1, 2, \cdots, k\text{–}1$. Then, by comparing Eq. (21–k) with (15–k), we have

$$x'S'\left[\frac{\partial v^{(k)}}{\partial x} - g^{*(k)}\right] = 0$$

and consequently $(\partial v^{(k)}/\partial x) = g^{*(k)}$. Q. E. D.

Further we have the following theorem of interest.

Theorem 6

If $g(x)$ is optimal up to the order l in ε, then

(i) J is equal to J^* up to the order $2l+1$ in ε.

(ii) $v^{(2l+2)}$, the $(2l+2)$th term in the expansion of J, is given by

$$x'S'\left[\frac{\partial v^{(2l+2)}}{\partial x} - g^{*(2l+2)}\right] = -\frac{1}{2}[g^{(l+1)\prime} - g^{*(l+1)\prime}]E[g^{(l+1)} - g^{*(l+1)}] \tag{24}$$

Proof

(i) Again the theorem is proved inductively. First, corresponding to $l=0$, assume that $g(x)=Px$. From Eqs. (15–1) and (21–1), it is readily observed that $v^{(1)}=v^{*(1)}$.

Second, assume that, if $g^{(i)}=g^{*(i)}$ for $i=1, 2, \cdots, l-1$, $v^{(i)}=v^{*(i)}$ for $i=1, 2, \cdots, 2l-1$. Besides if $g^{(l)}=g^{*(l)}$, comparison of Eq. (21–k) with (15–k) for $k=2l$

leads to $x'S'[(\partial v^{(2l)}/\partial x)-g^{*(2l)}]=0$. Due to *Corollary A*3, it implies that $v^{(2l)}=v^{*(2l)}$. Further, examination of Eqs. (15–k) and (21–k) for $k=2l+1$ leads to $v^{(2l+1)}=v^{*(2l+1)}$.

(ii) Comparison of Eq. (21–k) with (15–k) for $k=2l+2$ yields Eq. (24).

Q. E. D.

Remarks

(i) Especially if $g^{(l+1)}$ is zero, $v^{(2l+2)}$ is given by

$$x'S'\left[\frac{\partial v^{(2l+2)}}{\partial x}-g^{*(2l+2)}\right]=-\frac{1}{2}g^{*(l+1)\prime}Eg^{*(l+1)} \tag{25}$$

(ii) *Theorem* 6 applies also to a time-varying system with a finite T.

Conclusion

A systematic procedure has been developed for the suboptimal design of a nonlinear state regulator. When the nonlinearity is characterized by a polynomial function in the state, the definite way to determine a suboptimal feedback control is established by *Theorems* 1, 2 and *Theorem A* of Appendix. The stability property of the suboptimal feedback function is discussed in *Theorem* 3. Further *Theorem* 6 clarifies the approximation property of the suboptimal policy. The content of *Lemma* in Appendix could be used to construct a desired polynomial function, e.g., desired Liapunov function.

In a separate paper, the following topics will be discussed:

1. Strongly nonlinear systems.
2. Nonlinearity which is not represented by a finite sum of polynomials.
3. Possibility of the exact optimal design for a class of nonlinear systems.

Appendix

Consider the vector-matrix equation

$$\frac{\partial}{\partial t}g(x,t)+\left[\frac{\partial}{\partial x}g(x,t)\right]S(t)x+S'(t)g(x,t)=\frac{\partial}{\partial x}h(x,t) \tag{A.1}$$

where x is the n-vector of the state, $g(x, t)$ is an n-vector, $S(t)$ is an $n\times n$-matrix, and $h(x, t)$ is a scalar. All $g(x, t)$, $S(t)$, and $h(x, t)$ are assumed to be continuous in t. The problem is to determine the function g given a polynomial h of the degree $r+1$ in x.

The polynomial function $h(x, t)$ is written as:

$$h(x, t) = \sum_{j_1, \cdots, j_{r+1}=1}^{n} H_{j_1 j_2 \cdots j_{r+1}}(t) \, x_{j_1} x_{j_2} \cdots x_{j_{r+1}} \tag{A.2}$$

The coefficients $H_{j_1 j_2 \cdots j_{r+1}}(t)$ are $(r+1)$-index quantities, each index $j_1, j_2, \cdots, j_{r+1}$ running from 1 to n. Values of H are assumed to be, without loss of generality, symmetric with respect to any pair of indices.* Hence we have $(n+r)!/[(n-1)!(r+1)!]$ different coefficients. In what follows, we use the summation convention that a twice repeated index is to be interpreted as a summation over that index from 1 to n. By using this convention Eq. (A. 2) is simply written as

$$h(x, t) = H_{j_1 j_2 \cdots j_{r+1}}(t) \, x_{j_1} x_{j_2} \cdots x_{j_{r+1}} \tag{A.3}$$

The i-th component of $\partial h/\partial x$, the gradient of h, is given by

$$\frac{\partial h}{\partial x_i} = (r+1) \, H_{i j_1 \cdots j_r} x_{j_1} \cdots x_{j_r} \tag{A.4}$$

From Eq. (A. 1) it easily follows that the i-th component of g is a polynomial of the degree r in x. Hence the i-th component of g may be written as

$$g_i(x, t) = G^i_{j_1 \cdots j_r}(t) \, x_{j_1} \cdots x_{j_r} \tag{A.5}$$

where the coefficients G are symmetric with respect to the lower indices. Substituting Eqs. (A. 4) and (A. 5) into (A. 1) leads to

$$\begin{gathered}[\dot{G}^i_{j_1 \cdots j_r} + r S_{kj_1} G^i_{k j_2 \cdots j_r} + S_{ki} G^k_{j_1 \cdots j_r} - (r+1) H_{i j_1 \cdots j_r}] \, x_{j_1} \cdots x_{j_r} = 0 \\ (i = 1, 2, \cdots, n)\end{gathered} \tag{A.6}$$

Equations (A. 6) hold for arbitrary values of x if and only if the coefficients of $x_{j_1} x_{j_2} \cdots x_{j_r}$ vanish for any combinations of the values of indices. Hence we have $n(n+r-1)!/(n-1)!r!$ simultaneous equations for $G^i_{j_1 \cdots j_r}$:

$$\begin{aligned}&\dot{G}^i_{j_1 \cdots j_r} + S_{kj_1} G^i_{k j_2 \cdots j_r} + S_{kj_2} G^i_{j_1 k j_3 \cdots j_r} + \cdots + S_{kj_r} G^i_{j_1 \cdots j_{r-1} k} \\ &\quad + S_{ki} G^k_{j_1 \cdots j_r} - (r+1) H_{i j_1 \cdots j_r} = 0 \quad (i, j_1, \cdots, j_r = 1, 2, \cdots, n)\end{aligned} \tag{A.7}$$

The following lemma is useful for the solution of Eqs. (A.7):

Lemma

Consider the n^r algebraic equations for n^r quantities $X_{j_1 \cdots j_r}$:

---

* Values of H are said to be symmetric with respect to any pair of indices, or simply with respect to indices, if the values of H are invariant with any interchange of indices. For example, $H_{j_1=i,\, j_2=k,\, j_3 \cdots j_{r+1}} = H_{j_1=k,\, j_2=i,\, j_3 \cdots j_{r+1}}$.

$$C_{kj_1}X_{kj_2\cdots j_r}+C_{kj_2}X_{j_1kj_3\cdots j_r}+\cdots+C_{kj_r}X_{j_1\cdots j_{r-1}k}=D_{j_1\cdots j_r} \tag{A.8}$$

where C_{ij} and $D_{j_1\cdots j_r}$ are given quantities. All the indices run from 1 to n.

Equations (A.8) can be rewritten into the vector-matrix form:

$$\hat{C}^{(r)}\xi^{(r)}=\eta^{(r)} \tag{A.9}$$

where

$$\hat{C}^{(r)}=\underbrace{C'\oplus C'\oplus\cdots\oplus C'}_{r} \tag{A.10}$$

C is the $n\times n$-matrix such that the (i,j)-element is C_{ij}; $\xi^{(r)}$ and $\eta^{(r)}$ are the n^r-vectors of which the m-th components are $X_{j_1\cdots j_r}$ and $D_{j_1\cdots j_r}$, respectively, m being $\sum_{k=1}^{r}(j_k-1)n^{r-k}+1$. The symbol $\oplus$ denotes the Kronecker sum of matrices.

Before proving *Lemma*, the definitions and important properties of the Kronecker product and sum are briefly mentioned.

Definition

(i) Let A and B be $n\times n$- and $m\times m$- matrices, respectively. The $nm\times nm$-matrix whose (p, q)-element is $A_{ij}B_{kl}$, p being $m(i-1)+k$ and q being $m(j-1)+l$, is called the Kronecker product of A and B, and written $A\otimes B$.
(ii) The $nm\times nm$-matrix $A\otimes I^{(m)}+I^{(n)}\otimes B$, where $I^{(k)}$ $(k=m, n)$ is the $k\times k$-identity matrix, is called the Kronecker sum of A and B, and written $A\oplus B$.

Property

(i) The eigenvalues of $A\otimes B$ are $\lambda_i\mu_j$ $(i=1, 2, \cdots, n; j=1, 2, \cdots, m)$, where λ_i and μ_j are the eigenvalues of A and B, respectively.
(ii) The eigenvalues of $A\oplus B$ are $\lambda_i+\mu_j$.

For more details of the Kronecker product and sum, refer to Ref. 10, pp. 227–231.

Proof of Lemma

Let us proceed inductively. First, for $r=2$, Eqs. (A.8) reduce to

$$C_{kj_1}X_{kj_2}+C_{kj_2}X_{j_1k}=D_{j_1j_2} \tag{A.11}$$

By introducing the matrices X and D whose (i,j)-elements are X_{ij} and D_{ij}, respectively, Eqs. (A.11) are rewritten into

$$C'X+XC=D \tag{A.12}$$

In terms of $\xi^{(2)}$ and $\eta^{(2)}$ as defined in *Lemma*, Eq. (A.12) is equivalent to [Ref. 10, p. 231]

$$(C'\oplus C')\,\xi^{(2)} = \eta^{(2)} \tag{A.13}$$

Secondly we show that *Lemma* is true for $r+1$ if it is true for r. For $r+1$, Eqs. (A.8) are written as

$$\begin{aligned}[C_{kj_1}X_{kj_2\cdots j_{r+1}}+C_{kj_2}X_{j_1k\cdots j_{r+1}}+\cdots+C_{kj_r}X_{j_1\cdots j_{r-1}kj_{r+1}}]\\ +C_{kj_{r+1}}X_{j_1\cdots j_rk} = D_{j_1\cdots j_{r+1}}\end{aligned} \tag{A.14}$$

The terms in the bracket are identical, for a fixed value of j_{r+1}, with those on the left side of Eqs. (A.8). Hence, with use of $\xi^{(r+1)}$ and $\eta^{(r+1)}$ Eqs. (A.14) are rewritten into

$$[I^{(n)}\otimes\hat{C}^{(r)}+C'\otimes I^{(n^r)}]\xi^{(r+1)} = \eta^{(r+1)} \tag{A.15}$$

Due to the foregoing *Definition* (ii), Eq. (A.15) is equivalent to

$$\hat{C}^{(r+1)}\xi^{(r+1)} = \eta^{(r+1)} \tag{A.16}$$

Q. E. D.

Let T be a nonsingular $n\times n$-matrix which diagonalizes the matrix C', i.e.,*

$$TC'T^{-1} = \Lambda \quad \text{or} \quad C' = T^{-1}\Lambda T \tag{A.17}$$

where

$$\Lambda = \text{diag}\,(\lambda_i) \tag{A.18}$$

λ_i are the eigenvalues of C' or equivalently of C. By defining the $n^r\times n^r$-matrix U such that

$$U = \underbrace{T\otimes T\otimes\cdots\otimes T}_{r} \tag{A.19}$$

one is able to transform Eq. (A.9) into

$$U^{-1}MU\xi^{(r)} = \eta^{(r)} \tag{A.20}$$

where

$$M = \text{diag}\,(\lambda_{j_1}+\lambda_{j_2}+\cdots+\lambda_{j_r}) \tag{A.21}$$

As $\xi^{(r)}$ is uniquely given by Eq. (A.20) if and only if the matrix M is nonsingular, the following corollaries are established:

---

* For simplicity we assume that the diagonalization is possible. If C' has multiple eigenvalues and if C' can not be reduced to a diagonal form, we must consider a transformation into the Jordan canonical form. With a slight modification, *Corollary L1* applies also to the case of the Jordan form.

Corollary L1

If and only if any sum of r eigenvalues of C, $\lambda_{j_1}+\lambda_{j_2}+\cdots+\lambda_{j_r}(j_1, j_2, \cdots, j_r = 1, 2, \cdots, n)$, is not zero, $\xi^{(r)}$ is uniquely determined; $(\xi^{(r)})_m$, the m-th component of $\xi^{(r)}$, is given by

$$(\xi^{(r)})_m = U_{mi}^{-1} M_{ij}^{-1} U_{jk} (\eta^{(r)})_k = U_{mi}^{-1} M_{ii}^{-1} U_{ik} (\eta^{(r)})_k \tag{A.22}$$

where U_{ij} and M_{ij} denote the (i, j)-elements of U and M, respectively. In terms of X and D of Eqs. (A.8), Eq. (A.22) is rewritten as

$$X_{j_1 \cdots j_r} = \frac{T_{j_1 i_1}^{-1} T_{j_2 i_2}^{-1} \cdots T_{j_r i_r}^{-1} T_{i_1 k_1} T_{i_2 k_2} \cdots T_{i_r k_r} D_{k_1 \cdots k_r}}{\lambda_{i_1}+\lambda_{i_2}+\cdots+\lambda_{i_r}} \tag{A.23}$$

where T_{ij} and T_{ij}^{-1} denote the (i, j)-elements of T and T^{-1}, respectively.

Corollary L2

If, in particular, all the eigenvalues of C have negative real parts, $X_{j_1 \cdots j_r}$ is uniquely given by Eq. (A.23).

Corollary L3

If $D_{j_1 \cdots j_r}$ is symmetric with respect to the indices, $X_{j_1 \cdots j_r}$ is also symmetric.

Now we return to Eqs. (A.7). By virtue of the foregoing *Lemma*, we can establish the following theorem.

Theorem A

Equations (A.7) are rewritten into the vector-matrix form:

$$\dot{\xi}(t)+\hat{S}^{(r+1)}(t)\xi(t) = \eta(t) \tag{A.24}$$

where $\hat{S}^{(r+1)}$ is the Kronecker sum of the $r+1$ identical matrices S'; ξ and η are the n^{r+1}-vectors such that the i-th components are $G_{j_1 \cdots j_r}^{j_0}$ and $H_{j_0 j_1 \cdots j_r}$ respectively, i being $\sum_{k=0}^{r} (j_k-1) n^{r-k}+1$.

The statement of *Theorem A* is a direct consequence of *Lemma* and could be understood without proof. As Eq. (A.24) is a linear equation of the simple form, the solution ξ could be obtained by using a conventional integration technique. The following corollaries summarize several facts of interest.

Corollary A1

The matrix $(\partial g/\partial x)$ is symmetric.

Proof

Since the coefficients H are considered to be, without any loss of generality, symmetric with respect to all the indices, the coefficients G are also symmetric with respect to all the indices, including the upper and the lower ones. This is due to *Corollary L*3. It implies that $(\partial g_i/\partial x_j)=(\partial g_j/\partial x_i)$ for any pair of i and j.

Q. E. D.

By virtue of symmetry of $(\partial g/\partial x)$, Eq. (A.1) is equivalent to

$$\frac{\partial}{\partial t}g(x, t)+\frac{\partial}{\partial x}[x'S'g(x, t)-h(x, t)] = 0 \tag{A.25}$$

Corollary A2

If S and η are time-invariant, and if all the eigenvalues of S have negative real parts, $\xi(t)$ is uniformly asymptotically stable relative to ξ_0 as $t\rightarrow-\infty$, where ξ_0 is the solution of the linear algebraic equation

$$\hat{S}^{(r+1)}\xi_0 = \eta \tag{A.26}$$

Equation (A.26) is uniquely solvable under the given condition.

*Corollary A*2 implies that, in Eq. (A.1) if S reduces to a stable constant matrix and h reduces to be free of explicit dependence on t as $t\rightarrow-\infty$, then the function $g(x, t)$ is uniformly asymptotically stable as $t\rightarrow-\infty$, relative to the polynomial $g_0(x)$ such that satisfies the equation:

$$x'S'g_0(x) = h(x) \tag{A.27}$$

The coefficients $G^i_{j_1\cdots j_r}$ of the polynomial $g_0(x)$ are given by Eq. (A.26).

Corollary A3

If the equation

$$x'S'g = 0 \quad \text{or} \quad g'Sx = 0 \tag{A.28}$$

holds for any value of x with a stable S, and if $(\partial g/\partial x)$ is known to be symmetric, then g is the identically zero vector.

References

1) Kokotović, P.V., Cruz, J.B., and D'Ans, G., "ε-Coupling Method for Near-Optimum Design of Large-Scale Linear Systems," Proc. IEE, Vol. 116, pp. 889–892 (1969).
2) Garrard, W.L., McClamroch, N.H., and Clark, L.G., "An Approach to Sub-Optimal Feedback Control of Non-Linear Systems," Int. J. of Control, Vol. 5, pp. 425–435 (1967).
3) Garrard, W.L., "Additional Results on Sub-Optimal Feedback Control on Non-Linear Systems," Int. J. of Control, Vol. 10, pp. 657–663 (1969).
4) Al'brekht, E.G., "On the Optimal Stabilization of Nonlinear Systems," J. Appl. Math. Mech. (PMM) Vol. 25, pp. 1254–1266 (English translation of Russian original, 1961).
5) Pearson, J.D., "Approximation Methods in Optimal Control, I. Sub-Optimal Control," J. Electronics and Control, Vol. 13, pp. 453–469 (1962).

6) Sawaragi, Y., Ono, T., and Matsui, Y., "A Study on the Sub-Optimal Control of Nonlinear Systems by the Method of Instantaneous Linearization," J. Japan Assoc. Aut. Cont. Eng., Vol. 11, pp. 131–138 (in Japanese, 1967).
7) Nishikawa, Y., Sannomiya, N., Itakura, H., and Fukuda, T., "Methods for Suboptimal Control of Large-Scale Systems," to be presented at IFAC Symp. on Systems Engineering Approach to Computer Control, Kyoto, August 1970.
8) Kalman, R.E., "Contributions to the Theory of Optimal Control," Bol. Soc. Math. Mex., Vol. 5, pp. 102–119 (1960).
9) Morse, P.M. and Feshbach, H., Methods of Theoretical Physics, McGraw-Hill, New York, 1953.
10) Bellman, R., Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.