INT. J. CONTROL, 1970, VOL. 12, NO. 5, 897–908

A quadratic performance index maximization problem†

B. D. O. ANDERSON

Department of Electrical Engineering,
University of Newcastle, New South Wales, 2308, Australia

[Received 23 October 1969]

The problem is considered of maximizing a positive quadratic functional of the states of a linear system, subject to a bound on the integral of the square of the control. The solution is characterized in terms of the maximum eigenvalue and associated eigenfunction of a non-negative, definite, self-adjoint, integral kernel, and computational techniques for solving the associated eigenvalue problem are discussed.

1. Introduction

The usual quadratic loss optimization problem requires the minimization of a performance index of the form:

$$V(x_0, u(\cdot)) = \int_0^T [x'(t)\,Q(t)\,x(t) + u'(t)\,R(t)\,u(t)]\,dt + x'(T)\,Ax(T) \tag{1}$$

for a linear system:

$$\dot{x} = F(t)\,x + G(t)\,x. \tag{2}$$

The matrices Q, R, F and G are assumed to have continuous entries, and u may be a vector. With an avowed aim of achieving a regulator design, $Q(t)$ and A are taken as non-negative definite or even positive definite. In order to limit the magnitude of $u(\cdot)$, the matrix $R(t)$ is taken to be positive definite.

Here, we consider instead a maximization problem; more precisely we seek to maximize:

$$V(u(\cdot)) = \int_0^T x'(t)\,Q(t)\,x(t)\,dt + x'(T)\,Ax(T), \tag{3}$$

where again $Q(t)$ and A are non-negative or positive definite and $Q(t)$ has continuous entries. As before, eqn. (2) describes the system under consideration. We impose a limit on $u(\cdot)$ by demanding that

$$\int_0^T u'(t)\,u(t)\,dt \leqslant k \tag{4}$$

for some constant k. Finally, we restrict consideration to the case $x(0) = 0$.

An example of such an optimization problem might arise in the following situation. Consider a motor driving an electromechanical system, including a flywheel; with an upper constraint on the amount of energy supplied to the system over a certain interval, it is desired to maximize the speed of the flywheel at the end of the interval, assuming the system to be initially at rest. In this case, we would have $Q = 0$, and A would consist of a diagonal matrix with but the one non-zero entry in that row and column corresponding to that entry of the state vector representing flywheel speed.

† Communicated by Professor Anderson.

As we shall see in § 2, the solution of the general problem requires the computation of eigenvalues and eigenfunctions associated with a non-negative, definite, self-adjoint, integral kernel. In § 3, we relate the eigenvalue and eigenfunction ideas to properties of the solution of a linear differential equation. In particular, we can construct eigenfunctions from the linear differential equation once we have identified the associated eigenvalue. Section 4 provides examples and remarks on computational procedures.

Related problems appear in Gill and Sivan (1969) and Quincy (1969). In Gill and Sivan (1969), the problem is considered of minimizing:

$$\int_0^T [-x'(t)\,Q(t)\,x(t)+u'(t)\,u(t)]\,dt \tag{5}$$

(with $Q(t)$ as before); conditions are sought guaranteeing a unique optimal control corresponding to absence of a conjugate point in $[0,T]$. Quincy (1969) considers the problem of maximizing the output energy of a linear, time-varying, signal transmission channel for a fixed input energy, by appropriate choice of the actual signal transmitted. It is shown that solution of the problem requires the determination of the maximum eigenvalue and associated eigenfunction of a non-negative, self-adjoint kernel, though no technique is given for this determination.

2. Derivation of the eigenvalue equation

In this section, we shall first rewrite the performance index:

$$V(u(\cdot)) = \int_0^T x'(t)Q(t)\,x(t)\,dt + x'(T)\,Ax(T)$$

in the form:

$$V(u(\cdot)) = \int_0^T\int_0^T u'(t)\,[K(t,\tau)+L(t,\tau)]\,u(\tau)\,dt\,d\tau, \tag{6}$$

where $K(\cdot,\cdot)$ and $L(\cdot,\cdot)$ are non-negative, self-adjoint, integral operators. Then we shall discuss the minimization of the index (6) subject to the constraint (5).

Recalling that $x(0)=0$, we observe that with $\phi(\cdot,\cdot)$ the transition matrix associated with $F(\cdot), x(\cdot)$ may be expressed in terms of $u(\cdot)$ via:

$$x(t) = \int_0^T 1(t-\tau)\,\phi(t,\sigma)\,G(\sigma)\,u(\sigma)\,d\sigma.$$

It follows that

$$\begin{aligned}
&\int_0^T x'(t)\,Q(t)\,x(t)\,dt \\
&= \int_0^T \left\{\int_0^T 1(t-\sigma)\,u'(\sigma)\,G'(\sigma)\,\phi'(t,\sigma)\,d\sigma\right\} Q(t) \left\{\int_0^T 1(t-\tau)\,\phi(t,\tau)\,G(\tau)\,u(\tau)\,d\tau\right\} dt \\
&= \int_0^T\int_0^T d\sigma\,d\tau u'(\sigma)\left\{G'(\sigma)\int_0^T \phi'(t,\sigma)\,Q(t)\,\phi(t,\tau)\,1(t-\sigma)\,1(t-\tau)\,dt G(\tau)\right\} u(\tau) \\
&= \int_0^T\int_0^T u'(t)\,K(t,\tau)\,u(\tau)\,dt d\tau,
\end{aligned} \tag{7}$$

where

$$K(t,\tau) = G'(t)\int_t^T \phi'(\lambda,t)\,Q(\lambda)\,\phi(\lambda,\tau)\,d\lambda G(\tau)\,1(t-\tau)$$

$$+G'(t)\int_\tau^T \phi'(\lambda,t)\,Q(\lambda)\,\phi(\lambda,\tau)\,d\lambda G(\tau)\,1(\tau-t). \tag{8}$$

Observe also that

$$x'(T)\,Ax(T) = \int_0^T u'(\sigma)\,G'(\sigma)\,\phi'(T,\sigma)\,d\sigma A\int_0^T \phi(T,\tau)\,G(\tau)\,u(\tau)\,dt$$

$$= \int_0^T\int_0^T u'(t)\,L(t,\tau)\,u(\tau)\,dt\,d\tau, \tag{9}$$

where

$$L(t,\tau) = G'(t)\,\phi'(T,t)\,A\phi(T,\tau)\,G(\tau). \tag{10}$$

Tying together the above two calculations, eqn. (6) is immediate.

The theory of integral operators, see e.g. Kolmogorov and Fomin (1961), gives conditions for which (6) is maximized. In brief, the kernel $K(t,\tau)+L(t,\tau)$ possesses a denumerable set of real eigenvalues $\infty>\lambda_1\geqslant\lambda_2\geqslant\ldots$ with all the λ_i non-negative. Moreover, with λ_1 the maximum eigenvalue, the maximum value of (6) subject to the constraint:

$$\int_0^T u'(t)\,u(t)\,dt\leqslant k \tag{11}$$

is:

$$V_{\max} = \lambda_1 k \tag{12}$$

and the particular $u_1(\cdot)$ which achieves the maximization satisfies the equation:

$$\int_0^T [K(t,\tau)+L(t,\tau)]\,u_1(\tau)\,d\tau = \lambda_1 u_1(t). \tag{13}$$

Further, $u_1(\cdot)$ satisfies (11) with equality, i.e.

$$\int_0^T u_1'(t)\,u_1(t)\,dt = k. \tag{14}$$

The original continuity assumptions on the entries of $F(\cdot)$, etc. can be shown to imply that $u_1(\cdot)$ is continuous. If $K(t,\tau)+L(t,\tau)$ is such that λ_1 is not a repeated eigenvalue, the eigenfunction $u_1(\cdot)$ is unique. (Note: These results of course parallel results of linear algebra: if $K+L$ is a positive definite symmetric matrix, the maximum of $x'(K+L)x$ subject to $x'x = k$ is $\lambda_1 k$, where λ_1 is the maximum eigenvalue of $K+L$. The maximizing x satisfies $(K+L)x = \lambda_1 x$.)

In the next section, results helpful for the computation of λ_1 and $u_1(\cdot)$ are discussed. We shall make particular use of the structures (8) and (10) of $K(t,\tau)$ and $L(t,\tau)$, arising from the finite-dimensional nature of the underlying systems. Related results, nevertheless different, may be found in Baggeroer (1967) and Anderson (1969).

3. Solution of the eigenvalue equation

In posing the problem of maximizing:

$$V(u(\cdot)) = \int_0^T x'(t)Q(t)x(t) + x'(T)Ax(T) \tag{3}$$

for the basic system eqn. (2), and subject to $x(0) = 0$ and

$$\int_0^T u'(t)u(t)\,dt \leqslant k, \tag{4}$$

we were led to form the integral operator $K(t,\tau)+L(t,\tau)$ (given by (8) and (9)), which has the property that its maximum eigenvalue λ_1, known to be real, possesses as its eigenfunction the control $u_1(\cdot)$ achieving the desired maximization, i.e.

$$\int_0^T [K(t,\tau)+L(t,\tau)]u_1(\tau)\,d\tau = \lambda_1 u_1(t) \tag{13}$$

and $u_1(\cdot)$ satisfies (4) with equality.

The following theorem relates the eigenvalues and eigenfunctions of (13) to the properties of a linear differential equation solution. Part 1 of the theorem in essence gives a constructive procedure for eigenvalues and eigenfunctions, while part 2 indicates that the constructive procedure in fact includes all eigenvalues and eigenfunctions rather than perhaps some subset of them.

Theorem

Consider the linear differential equation:

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} F & -\dfrac{GG'}{\lambda} \\ Q & -F' \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}. \tag{15}$$

With F an $n\times n$ matrix, the transition matrix associated with (15), which we denote by $\Psi(\cdot,\cdot,\lambda)$, is a $2n\times 2n$ matrix, parametrized by λ. Suppose it is partitioned as:

$$\Psi(t,\tau;\lambda) = \begin{bmatrix} \Psi_{11}(t,\tau;\lambda) & \Psi_{12}(t,\tau;\lambda) \\ \Psi_{21}(t,\tau;\lambda) & \Psi_{22}(t,\tau;\lambda) \end{bmatrix} \tag{16}$$

with each $\Psi_{ij}(\cdot,\cdot;\lambda)$ an $n\times n$ matrix.

(1) If $\Psi_{11}(0,T;\lambda)-\Psi_{12}(0,T;\lambda)A$ is singular for some λ, then λ is an eigenvalue of $K(\cdot,\cdot)+L(\cdot,\cdot)$. Moreover, if x_T is a vector in the nullspace of $\Psi_{11}(0,T;\lambda)-\Psi_{12}(0,T;\lambda)A$ and $p_T = -Ax_T$, and if x_T and p_T are taken as boundary conditions at time T for (15), the corresponding solution pair $x(t), p(t)$ defines an eigenfunction $u(t)$ via:

$$u(t) = -\frac{1}{\lambda}G'(t)p(t). \tag{17}$$

(2) *Conversely*, let λ be an eigenvalue of $K(\cdot,\cdot)+L(\cdot,\cdot)$. Then

$$\Psi_{11}(0,T;\lambda)-\Psi_{12}(0,T;\lambda)A$$

is singular.

Proof

(1) With x_T and $p_T = -Ax_T$ as boundary conditions for eqn. (15) at time T, it follows that

$$x(t) = \Psi_{11}(t, T; \lambda)x_T - \Psi_{12}(t, T; \lambda)Ax_T.$$

Therefore $x(0) = 0$ because x_T lies in the null-space of

$$\Psi_{11}(0, T; \lambda) - \Psi_{12}(0, T; \lambda)A.$$

Now with $u(\cdot)$ defined by (17) (but without the presumption that $u(\cdot)$ is an eigenfunction), we have from (15) and (17) that $\dot{x} = Fx + Gu$ and therefore:

$$x(t) = \int_0^T 1(t-\tau)\,\phi(t,\tau)\,G(\tau)\,u(\tau)\,d\tau, \tag{18}$$

where we have used the fact that $x(0) = 0$. In particular:

$$x(T) = \int_0^T \phi(T,\tau)\,G(\tau)\,u(\tau)\,d\tau. \tag{19}$$

Now from (15):

$$p(t) = -\int_0^T 1(\lambda - t)\,\phi'(\lambda, t)\,Q(\lambda)\,x(\lambda)\,d\lambda + \phi'(t, T)\,p(T).$$

In this expression we may replace $x(t)$ by its expression (18), and $p(T)$ by first $-Ax(T)$, and then $-A$ times the expression (19) for $x(T)$. This yields:

$$\begin{aligned} p(t) = &-\int_0^T 1(t-\tau)\left\{\int_t^T \phi'(\lambda,t)\,Q(\lambda)\,\phi(\lambda,\tau)\,d\lambda\right\}G(\tau)\,u(\tau)\,d\tau \\ &-\int_0^T 1(\tau-t)\left\{\int_\tau^T \phi'(\lambda,t)\,Q(\lambda)\,\phi(\lambda,t)\,d\lambda\right\}G(\tau)\,u(\tau)\,d\tau \\ &-\phi'(t,T)\,A\int_0^T \phi(t,\tau)\,G(\tau)\,u(\tau)\,d\tau. \end{aligned} \tag{20}$$

Equations (17) and (20) provide two relations between $u(\cdot)$ and $p(\cdot)$, constructed solely on the basis of the singularity of $\Psi_{11}(0, T; \lambda) - \Psi_{12}(0, T; \lambda)A$. However, we can now use these two relations to establish the eigenfunction property of u.

Multiply (20) on the left by $G'(t)$. The left-hand side, from eqn. (17), is simply $-\lambda u(t)$. The right-hand side, from the definitions (8) and (10) of $K(\cdot,\cdot)$ and $L(\cdot,\cdot)$, is precisely:

$$-\int_0^T [K(t,\tau) + L(t,\tau)]\,u(\tau)\,d\tau.$$

Therefore we have:

$$\lambda u(t) = \int_0^T [K(t,\tau) + L(t,\tau)]\,u(\tau)\,d\tau. \tag{21}$$

The proof of part (1) is complete.

(2) Since λ is an eigenvalue, there is an associated eigenfunction $u(\cdot)$ such that (21) holds. Define the quantities:

$$x_1(t) = \int_0^T 1(t-\tau\,\phi(t,\tau)\,G(\tau)\,u(\tau)\,d\tau \tag{22}$$

and

$$p_1(t) = -\int_0^T 1(\lambda - t)\,\phi'(\lambda, t)\,Q(\lambda)\,x_1(\lambda)\,d\lambda - \phi'(t, T)\,A x_1(T). \tag{23}$$

The expression (23) may be rewritten as:

$$\begin{aligned} p_1(t) = &-\int_0^T 1(t-\tau)\left\{\int_t^T \phi'(\lambda, t)\,Q(\lambda)\,\phi(\lambda, \tau)\,d\lambda\right\} G(\tau)\,u(\tau)\,d\tau \\ &-\int_0^T 1(\tau - t)\left\{\int_\tau^T \phi'(\lambda, t)\,Q(\lambda)\,\phi(\lambda, \tau)\,d\lambda\right\} G(\tau)\,u(\tau)\,d\tau \\ &-\phi'(t, T)\,A\int_0^T \phi(T, \tau)\,G(\tau)\,u(\tau)\,d\tau. \end{aligned}$$

Multiplying both sides on the left by $G'(t)$ and applying the definitions of $K(t, \tau)$ and $L(t, \tau)$ leads to:

$$G'(t)\,p_1(t) = -\int_0^T [K(t, \tau) + L(t, \tau)]\,u(\tau)\,d\tau.$$

Now we use the eigenfunction property of $u(\cdot)$ to conclude that

$$u(t) = -\frac{1}{\lambda} G'(t)\,p(t).$$

By substituting this result in the expression (22) for $x_1(\cdot)$, we then obtain:

$$\dot{x}_1 = F x_1 + \frac{GG'}{\lambda} p_1. \tag{24}$$

Equation (23) yields:

$$\dot{p}_1 = Q x_1 - F' p_1. \tag{25}$$

Consequently $x_1(t) = \Psi_{11}(t, T;\, \lambda)\,x_1(T) + \Psi_{12}(t, T;\, \lambda)\,p_1(T)$. But (22) and (23) also imply that $x_1(0) = 0$ and $p_1(T) = -A x_1(T)$. Therefore, we have immediately that $\Psi_{11}(0, T;\, \lambda) - \Psi_{12}(0, T;\, \lambda)\,A$ is singular.

To apply the preceding theorem to the optimization problem of interest requires, first, the determination of the maximum λ such that

$$\Psi_{11}(0, T;\, \lambda) - \Psi_{12}(0, T;\, \lambda)\,A$$

is singular, second, the computation of an associated eigenfunction and, third, the scaling of this eigenfunction (which of course does not affect the fact that it is an eigenfunction) by a multiplicative constant in order to satisfy the constraint on the control energy.

The computation of the maximum λ such that $\Psi_{11}(0, T;\, \lambda) - \Psi_{12}(0, T;\, \lambda)\,A$ is singular in general requires the solution of a transcendental equation, best done by trial and error. In the next section, we present two simple examples and some aids to computation.

The linear differential equation system (15) may also be associated with another technique for maximizing the performance index (3), subject to the control constraint (4). This comes about in the following way.

We pose an alternative but equivalent optimization problem to the original one, obtained by introducing a Lagrange multiplier. More precisely, we seek to *minimize*, for fixed but for the moment arbitrary positive λ, the performance index:

$$\int_0^T [-x'(t)Q(t)x(t)+\lambda u'(t)u(t)]\,dt - x'(T)Ax(T), \tag{26}$$

subject to the constraint:

$$\int_0^T u'(t)u(t)\,dt = k \tag{27}$$

and the constraint $x(0) = 0$.

In general in problems involving Lagrange multipliers, side-constraints force λ to take on restricted values or even a single value. The effect of these different λ can then be studied. This is roughly the case here too, as we shall now see. A standard procedure for minimizing (26) involves the use of the Pontryagin maximum principle; applied in this case, it yields that the control minimizing (26) (with (27) for the moment disregarded) must be given by:

$$u(t) = -\frac{1}{\lambda}G'(t)p(t), \tag{17}$$

where $p(t)$ is the adjoint vector, satisfying (together with the state vector *obtained* when using the optimal control) the equation set:

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} F & -\dfrac{GG'}{\lambda} \\ Q & -F' \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}, \tag{15}$$

subject to $x(0) = 0$ and $p(T) = -Ax(T)$.

Now *one* solution of (15) satisfying the boundary conditions is

$$x(t) \equiv 0, \quad p(t) \equiv 0,$$

which would imply that $u(t) \equiv 0$. Almost certainly, though, the optimal control minimizing (26) and which also maximizes the original performance index (3) is not identically zero. (The zero control would lead to a zero value of the original performance index.) Therefore we are led to seek other solutions to (15) which satisfy the boundary conditions, and which lead to a $u(t)$ not identically zero.

Following along similar lines to those in the proof of the theorem, it then becomes clear that only when λ is such that $\psi_{11}(0,T;\lambda) - \Psi_{12}(0,T;\lambda)A$ is singular is it possible for (15) to have other solutions.

But argument from this point on becomes much harder, since we lack the interpretation of λ as an eigenvalue of an integral operator, and we lack the notation that we must seek the maximum eigenvalue of this operator. These, however, are precisely the notions that the less standard approach leading to the main theorem makes clear.

4. Further remarks on the eigenvalue computation

In this section, we discuss two technical points which may assist in the computation of the maximum eigenvalue, as opposed to an arbitrary eigenvalue, of the appropriate integral kernel.

The first of these points establishes the fact that the maximum eigenvalue λ_1 is not such that $\Psi_{11}(\tau,T;\lambda_1) - \Psi_{12}(\tau,T;\lambda_1)A$ is singular for any τ strictly within $[0,T]$, and that as a consequence a technique for rapidly computing λ_1

can be found. The second point offers a procedure for establishing the singularity of $\Psi_{11}(0,T;\lambda_1)-\Psi_{12}(0,T;\lambda_1)A$ by solving a Riccati equation, rather than the coupled differential equation set (15).

To discuss the first point, let us suppose temporarily that instead of attempting to compute the maximum eigenvalue of the equation:

$$\lambda_1 u_1(t) = \int_0^T [K(t,\tau)+L(t,\tau)]\, u_1(\tau)\, d\tau$$

for a fixed lower bound of 0 on the integral, we examine the problems of computing the maximum eigenvalue of a similar equation with variable lower bound t_0. Since the eigenvalues and eigenfunctions are parametrized by t_0, this equation can be written as:

$$\lambda_1(t_0)\, u_1(t;\, t_0) = \int_{t_0}^T [K(t,\tau)+L(t,\tau)]\, u_1(\tau;\, t_0)\, d\tau. \tag{28}$$

The definitions of $K(\cdot,\cdot)$ and $L(\cdot,\cdot)$ are of course unaltered. Then we claim that λ_1 is a monotonically decreasing function of t_0 (Fuchs 1964). To prove this, we differentiate (28) to obtain:

$$\frac{d\lambda_1}{dt_0} u_1(t;\, t_0) + \lambda_1 \frac{\partial u_1}{\partial t_0}(t;\, t_0) = -[K(t,t_0)+L(t,t_0)]\, u_1(t_0;\, t_0)$$
$$+ \int_{t_0}^T [K(t,\tau)+L(t,\tau)] \frac{\partial}{\partial t_0} u_1(\tau;\, t_0)\, d\tau,$$

which is of the form:

$$\lambda_1 \frac{\partial u_1}{\partial t_0}(t;\, t_0) = z(t) + \int_{t_0}^T [K(t,\tau)+L(t,\tau)] \frac{\partial u_1}{\partial t_0}(\tau;\, t_0)\, d\tau, \tag{29}$$

where

$$z(t) = -\frac{d\lambda_1}{dt_0} u_1(t;\, t_0) - [K(t,t_0)+L(t,t_0)]\, u_1(t_0;\, t_0).$$

Multiplying (29) on the left by $u_1'(t;\, t_0)$, integrating with respect to t from t_0 to T and using (28) yields:

$$\int_{t_0}^T u_1'(t;\, t_0)\, z(t)\, dt = 0.$$

From this equation, the definition of $z(t)$ and from the fact that $u_1(t;\, t_0)$ is an eigenfunction, it follows that

$$\frac{d\lambda_1}{dt_0} = -\lambda_1 u_1^2(t_0;\, t_0). \tag{30}$$

Since λ_1 is positive, the desired result is immediate.

Note that because of all the earlier continuity assumptions on $F(\cdot)$, etc., λ_1 will be a differentiable function of t_0. Note also that as $t_0 \to T$, λ_1 must approach zero. These facts suggest a helpful avenue in finding λ_1 for $t_0 = 0$ and some fixed T.

Choosing an arbitrary positive λ, find the transition matrix of the differential equation (15), and find that t_0 closest to T for which $\Psi_{11}(t_0,T;\lambda)-\Psi_{12}(t_0,T;\lambda)A$ is singular. Then λ is in fact the *maximum* eigenvalue λ_1 for the kernel $K(t,\tau)+L(t,\tau)$ regarded as an operator over $[t_0, T]$. For suppose not, i.e.

suppose that $\lambda_1(t_0) > \lambda$. Then there exists $t_1 \in (t_0, T)$ such that $\lambda_1(t_1) = \lambda$, since $\lambda_1(\cdot)$ is a continuous function equal to zero at T. But with $\lambda_1(t_1) = \lambda$, it follows that λ is an eigenvalue (actually the maximum one) of $K(t,\tau)+L(t,\tau)$ regarded as an operator over $[t_1, T]$ and therefore $\Psi_{11}(t_1, T; \lambda) - \Psi_{12}(t_1, T; \lambda)A$ is singular. This contradicts the definition of t_0 as being the closest value to T for which $\Psi_{11}(t_0, T; \lambda) - \Psi_{12}(t_0, T; \lambda)A$ is singular.

Consequently, there is a comparatively straightforward method of plotting $\lambda_1(t_0)$ versus t_0; given $\lambda_1(t_0)$ but not t_0, t_0 may be found as the value of t closest to T (but less than T) for which $\Psi_{11}(t, T; \lambda_1(t_0)) - \Psi_{12}(t, T; \lambda_1(t_0))A$ is singular. The variation of $\lambda_1(t_0)$ versus t_0 being monotonic, it is also straightforward to iterate to the desired $\lambda_1(0)$.

We now turn to a discussion of the second point mentioned at the start of the section. We make the observation that in terms of the partitioned blocks of the transition matrix $\Psi(\cdot, \cdot; \lambda)$ given in the theorem of the last section, the solution $P(t)$ of:

$$-\dot{P} = PF + F'P - PG\lambda^{-1}G'P - Q, \quad P(T) = -A, \tag{31}$$

may be expressed as:

$$P(t) = [\Psi_{21}(t, T; \lambda) - \Psi_{22}(t, T; \lambda)A][\Psi_{11}(t, T; \lambda) - \Psi_{12}(t, T; \lambda)A]^{-1}, \tag{32}$$

provided that $[\Psi_{11}(\tau, T; \lambda) - \Psi_{12}(\tau, T; \lambda)A]^{-1}$ exists for all τ in $[t, T]$, see e.g. Levin (1959).

Suppose now that we solve (31) directly for some fixed λ, rather than obtain its solution via the formula (32). The matrix $P(t)$ will be found by iterating backwards in time from $t = T$. At a certain time, $t = t_0$ say, the solution of (31) will become unbounded. From (32), we see that t_0 will be that value of time closest to T such that $\Psi_{11}(t_0, T; \lambda) - \Psi_{12}(t_0, T; \lambda)A$ is singular. Therefore the value of λ for which (31) is solved is the maximum eigenvalue of the integral operator $[K(t,\tau)+L(t,\tau)]$ considered as an operator over $[t_0, T]$. In other words, we have another technique involving the direct solution of (31) for computing $\lambda_1(t_0)$ versus t_0.

5. Examples

The first example illustrates a situation where the solution of a transcendental equation is *not* required. For the system:

$$\dot{x} = bx + u, \quad x(0) = 0,$$

with x a scalar and b a constant, we are required to maximize:

$$[x(T)]^2$$

subject to:

$$\int_0^T u^2\,dt \leqslant 1.$$

The linear differential equation from which we shall study the appropriate integral kernel eigenvalue is:

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} b & -\dfrac{1}{\lambda} \\ 0 & -b \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix},$$

which has the transition matrix:

$$\Psi(t,\tau;\lambda)=\begin{bmatrix}\exp[b(t-\tau)] & \frac{1}{2b\lambda}\{\exp[b(t-\tau)]-\exp[-b(t-\tau)\} \\ 0 & -\exp([-b(t-\tau)]\end{bmatrix}.$$

The eigenvalue equation becomes:

$$\exp(bT)-\frac{1}{2b\lambda}[\exp(bT)-\exp(-bT)]=0,$$

from which:

$$\lambda=\frac{2b\exp(bT)}{\exp(bT)-\exp(-bT)}.$$

An associated eigenfunction is found by computing the differential equation solution with $x(T)=1$, $p(T)=-1$. These boundary conditions lead to:

$$p(t)=\exp[-b(t-T)]$$

and so:

$$u(t)=-\frac{1}{\lambda}\exp[-b(t-T)].$$

However, $u(t)$ can be scaled and still retain its eigenfunction property. In order that

$$\int_0^T u^2\,dt=1,$$

we must have:

$$u(t)=\sqrt{(2b)}\exp[-b(t-T)].$$

The more general problem of maximizing $x'(T)Ax(T)$ given the system $\dot{x}=Fx+Gu, x(0)=0$ with a bound on

$$\int_0^T u'(t)\,u(t)\,dt$$

also leads to a readily solvable, non-transcendental equation for λ.

For a second example, we take as the basic system:

$$\dot{x}=u,\quad x(0)=0$$

and seek to maximize:

$$\int_0^T x^2\,dt+[x(T)]^2,$$

subject to:

$$\int_0^T u^2\,dt\leqslant 1.$$

We form:

$$\begin{bmatrix}\dot{x}\\ \dot{p}\end{bmatrix}=\begin{bmatrix}0 & -\frac{1}{\lambda}\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ p\end{bmatrix},$$

which has transition matrix:

$$\Psi(t,\tau;\lambda)=\begin{bmatrix}\cos\frac{1}{\sqrt{\lambda}}(t-\tau) & -\frac{1}{\sqrt{\lambda}}\sin\frac{1}{\sqrt{\lambda}}(t-\tau)\\ \sqrt{\lambda}\sin\frac{1}{\sqrt{\lambda}}(t-\tau) & \cos\frac{1}{\sqrt{\lambda}}(t-\tau)\end{bmatrix}.$$

The eigenvalue equation is:

$$\cos\frac{1}{\sqrt{\lambda}}T+\frac{1}{\sqrt{\lambda}}\sin\frac{1}{\sqrt{\lambda}}T=0$$

or

$$1+\frac{1}{\sqrt{\lambda}}\tan\frac{T}{\sqrt{\lambda}}=0.$$

No analytic solution exists for the maximum λ (or, indeed, any λ) satisfying this equation. For the case $T=1$, the maximum value of λ turns out to be 0·597, obtainable graphically or by numerical iteration.

With boundary condition $x(T)=1, p(T)=-1$, the solution $p(t)$ becomes:

$$0{\cdot}357\sin 2{\cdot}80(t-1)-\cos 2{\cdot}80(t-1)$$

and this is also the unscaled value of $u(t)$. The scaled value of $u(t)$ may of course be found readily. The maximum value of the performance index is 0·597.

6. Concluding remarks

We have posed a maximization problem of optimal control and solved it, in the sense of presenting a procedure for finding a numerical solution in a given instance. Though a quadratic performance index is used, it differs from the standard regulator problem index in that maximization of the index is sought rather than minimization, and no constraint on the control vector is indirectly imposed by the index itself. A specific constraint on control energy is however used. Further, the initial state vector is assumed zero.

If a non-zero initial state vector is used, the problem becomes much harder, and perhaps even intractable. Some idea of the additional complexity can be gained by comparing the following two problems (corresponding to $x(0)=0$ and $x(0)\neq 0$) in a finite-dimensional linear space.

Problem 1: Given a positive definite matrix Q, maximize $x'Qx$ subject to $x'x=1$.

Problem 2: Given a positive definite matrix Q and a constant vector a, maximize $x'Qx+2a'x$ subject to $x'x=1$.

Though the second problem appears to be an innocuous extension to the first, this is not really the case.

In the context of the optimization problems considered, there may be little point in retaining the same performance index for $x(0)$ non-zero. As a measure of the advantage gained by using a non-zero control as opposed to the zero control, the performance index:

$$\int_0^T [x'(t)-x'(0)\phi'(t,0)]Q(t)[x(t)-\phi(t,0)x(0)]\,dt$$
$$+[x'(T)-x'(0)\phi'(T,0)]A[x(T)-\phi(T,0)x(0)]$$

may be more appropriate. Maximizing this performance index for non-zero $x(0)$ is of course equivalent to the original problem.

The solution of the optimization problem, i.e. the value of the optimal performance index and an expression for the optimal control, may be described in the first instance in terms of the maximum eigenvalue and associated eigenfunction of a certain non-negative, self-adjoint, integral kernel. The special form of this kernel, arising from the finite-dimensional nature of the systems considered, allows a characterization of the eigenvalues and eigenfunctions in terms of the transition matrix of a set of linear differential equations (15).

The maximum eigenvalue in particular can be characterized by further properties of this transition matrix, which are apparently helpful in the actual computation process. In general, analytic determination of the maximum eigenvalue is impossible, though for one special class of problems analytical solution is possible.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Grants Committee. Discussions with J. Moore, P. Moylan and with T. Kailath regarding the calculation of the integral kernels in §2 are gratefully acknowledged.

REFERENCES

ANDERSON, B. D. O., 1969, *Proc. of the* 1969 *Hawaii International Conference on System Sciences*, 779. (Hawaii: Conference Committee).

BAGGEROER, A. B., 1967, *Technical Report* 459 (Research Laboratory of Electronics, Massachusetts Institute of Technology).

FUCHS, W. H. J., 1964, *J. math. Analysis Applic.*, **9**, 317.

GILL, A., and SIVAN, R., 1969, *I.E.E.E. Trans. autom. Control*, **AC–14**, 83.

KOLMOGOROV, A. N., and FOMIN, S. V., 1961, *Functional Analysis*, Vol. 2 (New York: Graylock, Albany).

LEVIN, J. J., 1959, *Proc. Am. math. Soc.*, **10**, 519.

QUINCY, E. A. 1969, *I.E.E.E. Trans. Inform. Theory*, **IT-15**, 325.