Reduction of the Operator Riccati Equation

Casti, John; Ljung, Lennart

1974

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Casti, J., & Ljung, L. (1974). *Reduction of the Operator Riccati Equation*. (Technical Reports TFRT-7066). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
2

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

REDUCTION OF THE OPERATOR RICCATI EQUATION

L. LJUNG
J. ÇASTI

Report 7412(C) August 1974
Lund Institute of Tecnology
Division of Automatic Control

REDUCTION OF THE OPERATOR RICCATI EQUATION

Lennart Ljung
Division of Automatic Control
Lund Institute of Technology
S-220 07 Lund, Sweden.

John Casti
Departments of Mathematics, Systems&Industrial Engineering, University of Arizona, Tucson, Arizona 85721 and IIASA, Laxenburg, Austria.

1. INTRODUCTION.

The fundamental role of the Riccati equation for optimal control and optimal filtering of linear systems is well known. For linear distributed parameter systems it has been shown by J. L. Lions [8], and in several other papers, see e.g. [12], [13], [5], that there is an operator Riccati equation (ORE), (or, equivalently, an integro-differential equation of Riccati type for the kernel) associated with various control and filtering problems.

The solution to the ORE is an operator in the state space. In the case of an infinite dimensional state space numerical solution of the ORE may therefore be cumbersome and time consuming, or even impossible. The interesting operator for the control problem maps the state space into the space of the control vector, and it is constructed using the solution to the ORE. In [4] it has been shown that, under certain conditions, it is possible to solve for this operator directly. This leads to less extensive computations for the numerical solution if the space of the control vector is of less dimensionality than the state space. In many practical situations there are only a finite number of control variables and also only a finite number of observations. In those cases, therefore, a substantial decrease in computational effort may be achieved.

While [4] treats general control and observation spaces and distributed control, we will here consider boundary control applied at a finite number of points. Many real-life, distributed-parameter control processes seem to be of this type.

2. THE OPTIMAL CONTROL PROBLEM.

Consider the following formulation, [12]:

Let D be a connected, open domain of a r-dimensional Euclidian space E^r, and let S be the boundary of D. The spatial coordinate vector will be denoted by

$$x = (x_1, \dots, x_r) \in D$$

and the time will be denoted by $t \in T = (t_o, t_1)$. Consider a time invariant matrix differential operator

$$A_x = \sum_{i=1}^{r} \frac{\partial}{\partial x_i} \left(\sum_{j=1}^{r} A_{ij}(x) \frac{\partial}{\partial x_j} \right) + A_o(x)$$

where $A_{ij}(x)$ and $A_o(x)$ are n×n self-adjoint matrix functions, with $A_{ij} = A_{ji}$.

Introduce the derivative with respect to the co-normal of the surface S relative to the operator A_x

$$\frac{\partial}{\partial n_A} = \sum_{j=1}^{r} \left(\sum_{i=1}^{r} A_{ij}(\xi) \cos(\nu, x_i) \right) \frac{\partial}{\partial x_j}$$

where ν is the exterior normal to the surface at a point $\xi \in S$, and $\cos(\nu, x_i)$ is the i:th direction cosine of ν.

Consider a linear, distributed-parameter system described by

$$\frac{\partial z(t,x)}{\partial t} = A_x z(t,x) \; ; \qquad z(t_o,x) = z_o(x) \tag{1a}$$

where $z(t,x)$ is a n-vector state function of $t \in T$ and $x \in D$. The boundary condition is given by

$$F(\xi)z(t,\xi) + \frac{\partial z(t,\xi)}{\partial n_A} = \sum_{i=1}^{m} \delta(\xi - \xi_i) B_i u_i(t) \tag{1b}$$

where $\xi \in S$, $F(\xi)$ is an n×n matrix function and $u_i(t)$ is a ℓ×1 control vector applied at the point ξ_i at the boundary.

Introduce

$$u(t) = \mathrm{col}(u_1(t), \ldots, u_m(t)) \qquad (\ell m\times 1) \text{ matrix})$$

Assume that the system described by (1) is well posed in the sense of Hadamard. A p-dimensional output vector is observed:

$$y(t) = \int_D C(x)z(t,x)dx$$

where $C(x)$ is a $p\times n$ matrix function.

The control problem is to determine $u_i(t)$ so that the cost functional

$$J = \int_{t_o}^{t_1} [y(t)^T Q_1 y(t) + u(t)^T Q_2 u(t)]dt$$

is minimized. Q_1 and Q_2 are positive definite symmetric matrices of dimensions $p\times p$ and $\ell m\times \ell m$ respectively.

In [12] it is shown that the solution can be written

$$u(t) = - \int_D K(t,x)z(t,x)dx$$

where

$$K(t,x) = Q_2^{-1} \tilde{P}(t,x) \qquad (\ell m\times n \quad \text{matrix})$$

where

$$\tilde{P}(t,x) = \mathrm{col}(B_1^T P(t,\xi_1,x), \ldots, B_m^T P(t,\xi_m,x)) \qquad (\ell m\times n \text{ matrix})$$

and $P(t,x,x')$ is the solution to the Riccati equation

$$\begin{aligned} \frac{\partial P(t,x,x')}{\partial t} &= -A_x P(t,x,x') - [A_{x'} P(t,x,x')]^T - \\ &- C(x)^T Q_1 C(x') + \tilde{P}(t,x)^T Q_2^{-1} \tilde{P}(t,x') \end{aligned} \tag{2a}$$

with the initial and boundary conditions

$$\begin{aligned} &P(t_1,x,x') = 0 \\ &F(\xi)P(t,\xi,x) + \frac{\partial P(t,\xi,x)}{\partial n_A} = 0 \qquad \xi \in S,\ x \in D \end{aligned} \tag{2b}$$

Addition of distributed control terms leads to straightforward modifications. These are given in [12].

3. THE OPTIMAL FILTERING PROBLEM.

The problem of optimal filtering of linear distributed-parameter systems has been treated in detail and rigorously by Bensoussan [2]. Here we will follow the formulation of [13].

In most practical situations the control equipment cannot be regarded as perfect. To the right hand side of (1b) therefore should be added a noise term

$$\sum_{i=1}^{m} \delta(\xi-\xi_i)H_i \, w_i(t)$$

where $w_i(t)$ is a q×1 vector white gaussian process with covariance

$$E \, w_i(t) \, w_i(t)^T = I, \qquad E \, w_i(t) \, w_j(t)^T = 0 \qquad i \neq j$$

Furthermore, the variable y(t) cannot be observed exactly, but a noise-currupted measurement

$$\tilde{y}(t) = y(t) + e(t)$$

is obtained, where e(t) is a p×1 vector white gaussian process with covariance

$$E \, e(t) \, e(t)^T = R_2$$

An appropriate estimate, $\hat{z}(t,x)$, of the state function, z(t,x), is then obtained from $\tilde{y}(t)$ as follows:

$$\frac{\partial z(t,x)}{\partial t} = A_x \hat{z}(t,x) + \tilde{K}(t,x)[\tilde{y}(t) - \int_D C(x)\hat{z}(t,x)dx] \qquad (3)$$

$$\hat{z}(t_o,x) = z_o(x)$$

where

$$\tilde{K}(t,x)^T = R_2^{-1} \int_D C(x') \, P(t,x,x')dx'$$

and

$P(t,x,x')$ is the solution of

$$\frac{\partial P(t,x,x')}{\partial t} = A_x\, P(t,x,x') + [A_{x'}\, P(t,x,x')]^T -$$

$$- \int_D P(t,x,\xi)\, C(\xi)^T d\xi\, R_2^{-1} \int_D C(\xi)\, P(t,\xi,x')d\xi + \tag{4a}$$

$$+ \sum_{i=1}^{m} H_i\, H_i^T\, \delta(x-\xi_i)\, \delta(x'-\xi_i)$$

with initial and boundary conditions

$$P(t_o,x,x') = 0$$

$$F(\xi)\, P(t,x,\xi) + \frac{\partial P(t,x,\xi)}{\partial n_A} = 0 \qquad x \in D, \quad \xi \in S \tag{4b}$$

Clearly, the Riccati equation (4) is quite analogous to (2). It corresponds to an optimal control problem with boundary observations and distributed control applied from a finite number of sources.

4. REDUCTION OF THE RICCATI EQUATION.

The equations (2) and (4) are partial differential equations in three variables: time and two space variables. However, the sought functions $K(t,x)$ and $\tilde{K}(t,x)$ depend only on one space variable. We will derive equations directly for these functions. The derivation will be formal inasmuch as existence and uniqueness of solutions to the discussed equations will not be verified. The idea of the reduction originated from certain problems in radiative transfer, see e.g. [3] and [6] for a corresponding treatment of the finite dimensional problem.

Consider eq. (2). Differentiate it with respect to t:

$$P_{tt} = -A_x\, P_t - [A_{x'}\, P_t]^T + \tilde{P}_t^T\, Q_2^{-1}\, \tilde{P} + \tilde{P}^T\, Q_2^{-1}\, \tilde{P}_t \tag{5a}$$

where

$$P_t = \frac{\partial}{\partial t} P$$

and the arguments have been suppressed.

If $\tilde{P}$ were known, eq. (5a) could be regarded as a linear equation for P_t with initial and boundary conditions

$$P_t(t_1 x,x') = -C(x)^T Q_1 C(x')$$

$$F(\xi) P_t(t,\xi,x) + \frac{\partial}{\partial n_A} P_t(t,\xi,x) = 0 \qquad \xi \in S, \quad x \in D \tag{5b}$$

Consider the equation

$$\frac{\partial}{\partial t} L(t,x) = -A_x L(t,x) + \tilde{P}(t,x)^T Q_2^{-1} \bar{L}(t) \tag{6a}$$

where

$$\bar{L}(t) = \mathrm{col}(B_1^T L(t,\xi_1),\ldots,B_m^T L(t,\xi_m)) \qquad (\ell m \times p \text{ matrix})$$

with initial and boundary conditions

$$L(t_1,x) = C(x)^T \sqrt{Q_1} \qquad (n\times p \text{ matrix})$$

$$F(\xi) L(t,\xi) + \frac{\partial L(t,\xi)}{\partial n_A} = 0 \;; \qquad \xi \in S \tag{6b}$$

It is straightforward to see that

$$P_t(t,x,x') = -L(t,x) L(t,x')^T$$

with $L(t,x)$ defined by (6) satisfies eq. (5).

Since

$$K(t,x) = Q_2^{-1} \tilde{P}(t,x)$$

and

$$\tilde{P}_t(t,x) = \mathrm{col}(B_1^T L(t,\xi_1) L(t,x)^T,\ldots,B_m^T L(t,\xi_m) L(t,x)^T) =$$

$$= -\bar{L}(t) L(t,x)^T$$

the equations

$$\frac{\partial}{\partial t} K(t,x) = -Q_2^{-1} \bar{L}(t) L(t,x)^T$$
$$\frac{\partial}{\partial t} L(t,x) = -A_x L(t,x) + K(t,x)^T \bar{L}(t) \tag{7a}$$

with initial and boundary conditions

$$K(t_1,x) = 0$$
$$L(t_1,x) = C(x)^T \sqrt{Q_1} \tag{7b}$$
$$F(\xi) L(t,\xi) + \frac{\partial}{\partial n_A} L(t,\xi) = 0 \qquad \xi \in S$$

determine the sought gain function K(t,x) directly without first determining P(t,x,x')

Analogously, eq. (4) can be reduced to

$$\frac{\partial}{\partial t} \tilde{K}(t,x)^T = R_2^{-1} \int_D C(x') \tilde{L}(t,x')dx' \tilde{L}(t,x)^T$$
$$\frac{\partial}{\partial t} \tilde{L}(t,x) = A_x \tilde{L}(t,x) - \tilde{K}(t,x) \int_D C(x') \tilde{L}(t,x')dx' \tag{8a}$$

with intial and boundary conditions

$$\tilde{K}(t_o,x) = 0 \qquad (n\times p \text{ matrix})$$
$$\tilde{L}(t_o,x) = (H_1 \, \delta(x-\xi_1),\ldots,H_m \, \delta(x-\xi_m)) \qquad (n\times qm \text{ matrix}) \tag{8b}$$
$$F(\xi) \tilde{L}(t,\xi) + \frac{\partial}{\partial n_A} \tilde{L}(t,\xi) = 0 \qquad \xi \in S$$

5. NUMERICAL SOLUTION.

Numerical solution of the ORE has been discussed in several papers. The following three methods seem to be predominating:

Finite difference approximations
Galerkin methods
Eigenfunction expansion

The best choice of solution method probably depends on the problem formulation, and it is difficult to give any general rules. We will here briefly discuss how the mentioned methods can be used to solve the reduced Riccati equation (7) and how the computational effort is reduced.

Finite difference approximation has been discussed in e.g. [1]. When it is applied to eq. (2) with, say, $n = 1$ and $r = 1$, a three-dimensional space must be discretized. For stability reasons the time step must be kept quite small, and consequently numerical solution will be time consuming. When the same method is applied to the reduced equation (7), the computation required is of an order of magnitude less, since one space variable has been eliminated.

Galerkin methods have been applied to the ORE in [11], and for more general control problems for parabolic systems in [10]. Then the kernel $P(t,x,x')$ is expanded in some coordinate function system:

$$P(t,x,x') = \sum_{i,j} P_{ij}(t)\, \Psi_i(x)\, \Psi_j(x')$$

The expansion is truncated at some suitable number N, and ordinary differential equations are derived for $P_{ij}(t)$, so that the orthogonal projection of the residual on the subspace spanned by $\Psi_i(x)$, $i=1,\dots,N$ is zero. The resulting equations have some resemblance with the matrix Riccati equation. The same technique can be applied to the reduced equation (7), which leads to $N(\ell m+p)$ ordinary differential equations instead of $N(N+1)/2$.

The eigenfunction method is discussed in e.g. [12], [13], [5] and [14]. It can be regarded as a special case of Galerkin's method with the coordinate funtions being eigenfunctions of the operator A_x. Consequently, the same reduction of the computational effort is obtained as for this method. In [9] a similar reduction is discussed for the case when the operator A_x is diagonal.

6. A NUMERICAL EXAMPLE.

Consider a heat rod of length Λ (45cm) with diffusion constant κ (1.16 cm^2/sec), conductivity constant μ (3.8 W/cm ^oC) and cross section area S(1.54 cm^2). The heat flow at the left endpoint is controlled, while the right endpoint is isolated. At either endpoint the heat flow is disturbed by white gaussian noise, $w_1(t)$ and $w_2(t)$, with variances r_1 ($3.4\cdot10^{-3}$ W^2). The output variable is the temperature at $\Lambda/4$ from the right endpoint, and the temperature measurements are corrupted by white gaussian noise with variance r_2 $(4 \cdot 10^{-8}\ {}^{o}C^2)$.

The numerical values given in paranthesis apply for a laboratory process at the Division of Automatic Control, Lund Institute of Technology, see e.g. [7].

Let the temperature at distance x from the left endpoint at time t be denoted by z(t,x). Then

$$\frac{\partial z(,x)}{\partial t} = \kappa \frac{\partial^2 z(t,x)}{\partial x^2}$$

$$z(0,x) = z_o(x) \qquad \text{(known)}$$

$$S\mu \frac{\partial z(t,0)}{\partial x} = u(t) + \sqrt{r_1}\ w_1(t)$$

$$S\mu \frac{\partial z(t,\Lambda)}{\partial x} = \sqrt{r_1}\ w_2(t)$$

$$y(t) = z(t,3\Lambda/4) + e(t)$$

where

$$E\ w_i(t)\ w_j(s) = \delta_{ij}\ \delta(t-s)$$

$$E\ e(t)\ e(s) = r_2\ \delta(t-s)$$

$$E\ w_i(t)\ e(s) = 0$$

Let the objective of the control be to minimize

$$E \int_0^{t_1} [q_1 u^2(t) + q_2 z^2(t,3\Lambda/4)]dt$$

According to Sections 2-4 the solution is given by

$$u(t) = \int_0^{\Lambda} K(t,x)\ \hat{z}(t,x)dx$$

where

$$\frac{\partial}{\partial t}\,\hat{z}(t,x) = \kappa \frac{\partial^2}{\partial x^2}\,\hat{z}(t,x) + \tilde{K}(t,x)[y(t) - \hat{z}(t,3\Lambda/4)]$$

$$\hat{z}(0,x) = z_o(x)$$

$$\hat{z}(t,0) = u(t)$$

$$\hat{z}(t,\Lambda) = 0$$

The function K(t,x) is the solution of

$$\frac{\partial}{\partial t}\,K(t,x) = -1/q_1\;L(t,0)\;L(t,x)$$

$$\frac{\partial}{\partial t}\,L(t,x) = -\kappa \frac{\partial}{\partial x^2}\,L(t,x) + L(t,0)\;K(t,x)$$

$$K(t_1,x) = 0$$

$$L(t_1,x) = \delta(x-3\Lambda/4)\cdot\sqrt{q_2}$$

$$\frac{\partial}{\partial x}\,L(t,0) = \frac{\partial}{\partial x}\,L(t,\Lambda) = 0$$

and the function K(t,x) is the solution of

$$\frac{\partial}{\partial t}\,\tilde{K}(t,x) = r_2^{-1}[\tilde{L}_1(t,3\Lambda/4)\;\tilde{L}_1(t,x) + \tilde{L}_2(t,3\Lambda/4)\;\tilde{L}_2(t,x)]$$

$$\frac{\partial}{\partial t}\,\tilde{L}_1(t,x) = \kappa \frac{\partial^2}{\partial x^2}\,\tilde{L}_1(t,x) - \tilde{L}_1(t,3\Lambda/4)\;\tilde{K}(t,x)$$

$$\frac{\partial}{\partial t}\,\tilde{L}_2(t,x) = \kappa \frac{\partial^2}{\partial x^2}\,\tilde{L}_2(t,x) - \tilde{L}_2(t,3\Lambda/4)\;\tilde{K}(t,x)$$

$$\tilde{K}(0,x) = 0$$

$$\tilde{L}_1(0,x) = \sqrt{r_1}\;\delta(x)$$

$$\tilde{L}_2(0,x) = \sqrt{r_1}\;\delta(x-\Lambda)$$

$$\frac{\partial}{\partial x}\,\tilde{L}_i(t,0) = \frac{\partial}{\partial x}\,\tilde{L}_i(t,\Lambda) = 0 \qquad i = 1,2$$

These equations have been solved using a straightforward difference scheme with 100 grid points in the x coordinate. The numerical values of the constants are those given in paranthesis. Furthermore, $q_1 = 1$ (W^{-2}), $q_2 = 8.5\cdot 10^{-4}$ $(^{\circ}C^{-2})$ and $t_1 = 1000$ sec. The solution is shown in fig. 1.

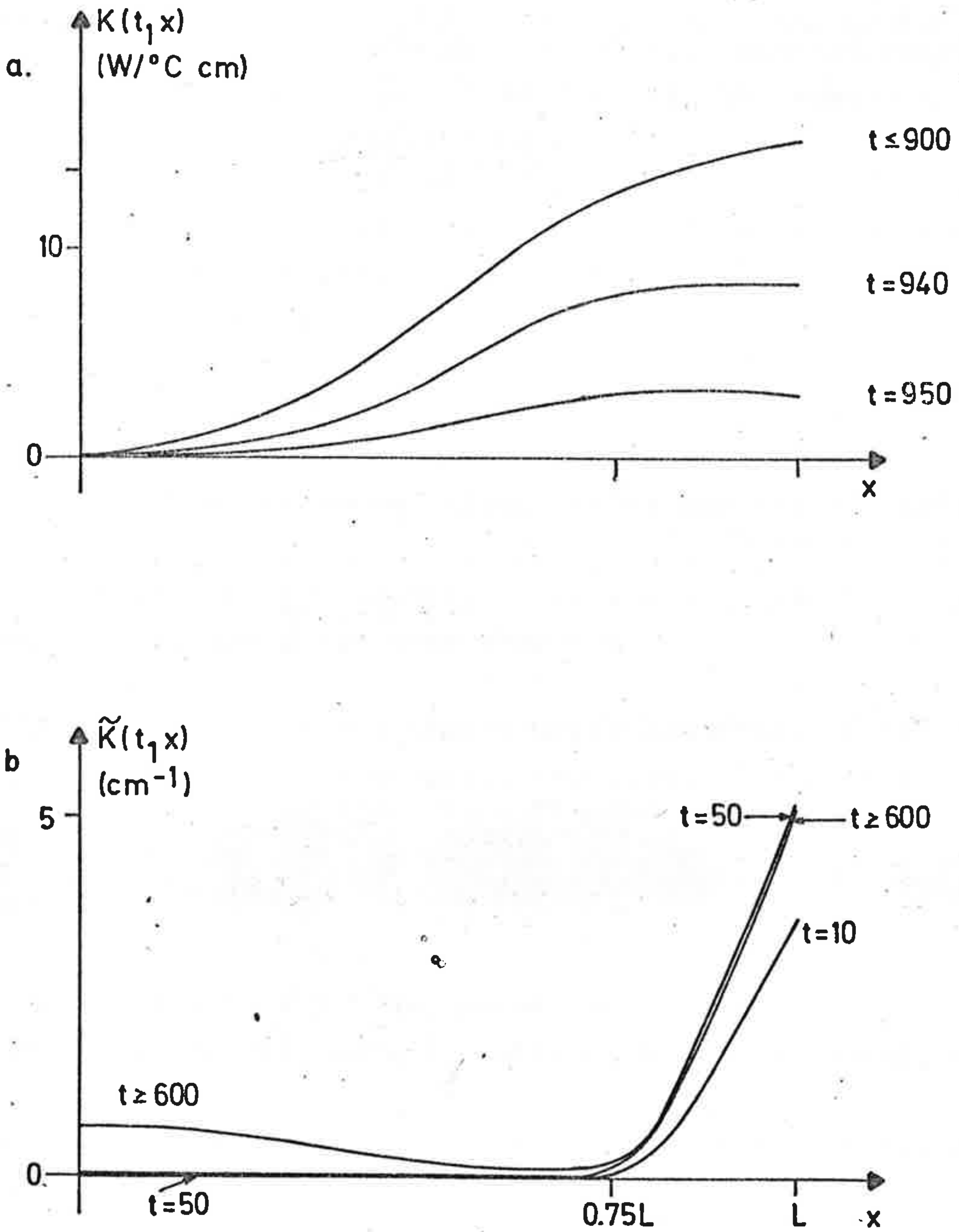

Fig. 1. a) Optimal feedback function K(t,x) for the heat regulation problem for various t. The final time t_1 is 1000 sec, and the numerical values of the constants are given in Section 6.

b) Optimal gain function $\tilde{K}(t,x)$ for the Kalman filter associated with the heat regulation problem stated in Section 6.

7. CONCLUSIONS.

It has been shown that the operator Riccati equation can, under certain, frequently occuring conditions be reduced to equations that contain fewer independent variables.

Therefore, numerical solution of the reduced equation requires less computation effort than solution of the original equation.

A case with boundary control at a finite number of points has been considered. Combining this with the results of [4] for distributed control, the general case with both distributed and boundary control can be treated straightforwardly.

It should be noticed that the reduction has been shown only for the initial condition $P(t_o,x,x') = 0$, corresponding to no terminal cost for the control problem or known initial state for the filtering problem.

There seems to be no possibility to extend the results to general initial conditions, using the same approach.

However, when computing the steady-state solutions, which are frequently used in the control strategies, the initial condition can as well be taken as zero.

REFERENCES.

[1] Alvardo, F., R., Mukundan, An optimization problem in distributed parameter systems, Int. J. Control, 9, No. 6 (1969), 665-677.

[2] Bensoussan, A., Filtrage Optimal des Systèmes Linéaires, Dunod, Paris 1971.

[3] Casti, J., Matrix Riccati equations, dimensionality reduction and generalized X-Y functions, to appear Utilitas Mathematica, 1974.

[4] Casti, J., L. Ljung, Some new analytic and computational results for operator Riccati equations, Report 7344, December 1973, Lund Institute of Technology, Division of Automatic Control. To appear in SIAM J. Control.

[5] Erzberger, H., M. Kim, Optimum boundary control of distributed parameter systems, Information and Control 9, (1966) 265-278.

[6] Kailath, T., Some new algorithms for recursive linear estimation in constant linear systems, IEEE Trans. on information theory IT-19, Nov., (1973).

[7] Leden, B., M. H. Hamza, M. A. Sheirah, Different methods for estimation of thermal diffusivity of a heat diffusion process, Proc. of the 3rd IFAC Symposium on Identification and System Parameter Estimation, the Hague (1973).

[8] Lions, J. L., The optimal control of systems governed by partial differential equations, Springer-Verlag, New York, 1971.

[9] Mayhew, M. J. E., A. J. Pritchard, Reduction of the Riccati equation for distributed parameter systems, Elec. Letters 7, No. 20, (1971), 628-629.

[10] McKnight, R. S., W. E. Bosarge, The Ritz-Galerkin procedure for parabolic control problems, SIAM J. Control 11, No. 3 (1973), 510-524.

[11] Prabhu, S. S., I. McCausland, Optimal control of linear diffusion processes with quadratic error criteria, Automatica 8 (1972), 299-308.

[12] Sakawa, Y., A matrix Green's formula and optimal control of Linear distributed-parameter systems, J, of Optimization Theory and Applications 10, No. 5 (1972), 290-299.

[13] Sakawa, Y., Optimal filtering in linear distributed-parameter systems, Int. J. Control 16, No. 1 (1972), 115-127.

[14] Tzafestas, S. G., J. M. Nightingale, Optimal control of a class of linear stochastic distributed-parameter systems, proc. IEE 115, No. 8 (1968).