November, 1982 LIDS-P-1255

LINEAR ESTIMATION OF BOUNDARY VALUE STOCHASTIC PROCESSES

PART I:

THE ROLE AND CONSTRUCTION OF COMPLEMENTARY MODELS

by

Milton B. Adams†

Alan S. Willsky*

Bernard C. Levy*

Abstract: This paper presents a substantial extension of Weinert and Desai's [1] method of complementary models for minimum variance linear estimation. Specifically, the method of complementary models is extended to solve estimation problems for both discrete and continuous parameter linear boundary value stochastic processes in one and higher dimensions. A major contribution of this paper is an application of Green's identity in deriving a differential operator representation of the estimator. To clarify the development and to illustrate the range of application of this estimator, two examples are carried along throughout: one is a 1-D discrete two-point boundary value process and the other is a 2-D process governed by Poisson's equation on the unit disk.

† C. S. Draper Laboratory, 555 Technology Square, Cambridge, Ma. 02139 (also Dept. of Aero. and Astro.,MIT). The work of this author was supported by a Draper Fellowship.

* Dept. of Elec. Eng. and Comp. Sci. and L.I.D.S., MIT, Cambridge, Ma. 02139. The work of these authors was supported in part by the National Science Foundation under Grant ECS-8012668.

SECTION 1

INTRODUCTION

In this paper we present a substantial extension of Weinert and Desai's [1] method of complementary models for minimum variance linear estimation. Weinert and Desai showed that the fixed interval smoothing problem for causal one-dimensional[1] processes described by linear state equations driven by white noise could be solved by introducing the so-called complementary process. The complementary process has the property that it is orthogonal to the observations and that, when combined with the observations, contains information equivalent to the initial conditions, driving noise and measurement noise, i.e. all of the underlying variables which determine the system state. Here we generalize this idea and show how to employ the method of complementary models to solve estimation problems for both discrete and continuous parameter boundary value stochastic processes in one and higher dimensions. These processes include those defined through ordinary and partial linear differential equations and ordinary and partial linear difference equations and may be either causal or noncausal. By employing operator descriptions for these processes we are able to unify the development of the estimators for this wide variety of processes within a single framework. The major contribution of this paper is a differential operator representation for the estimator which applies to all of the cases mentioned above. A key step in the derivation of the differential operator representation of the estimator is our formulation of a differential operator representation for the complementary process.

To help clarify our presentation we carry along two examples throughout. One example is a 2-D process governed by Poisson's equation with a white noise driving function. The other is a 1-D discrete two-point boundary value process. The emphasis in this paper is on the development of the differential

[1] The terminology one-dimensional (1-D), two-dimensional (2-D) or multidimensional process is used here to indicate that the dimension of the independent variable for the process is one, two or multidimensional.

representation for the estimator. In Part II [5], we consider a 1-D continuous pararmeter boundary value stochastic process and address the details of the implementation and structure of its estimator and the computation of its error covariance.

In Section 2 we review the complementary models approach to linear estimation and motivate this approach by its application to a static problem of estimating a random vector. Following the example, the approach is generalized to second order stochastic processes by way of a restatement of the Projection theorem. Section 3 serves to introduce some notation which we employ for representing boundary value stochastic processes and their correlation functions. Utilizing this notation, we state the general form of the estimation problem for which we ultimately develop a solution. In Section 4 we present an operator representation for the complementary stochastic process associated with the general problem stated earlier. In that section we offer a proof that this operator representation satisfies the properties of a complementary process as defined in Section 2. In Section 5 a general form for the internal differential realization for the complementary process is derived. Given this realization, we are abe to formulate an internal differential realization for the estimator. Using this recipe for the operator representation of the estimator, we present differential realizations for the estimators for the 1-D and 2-D examples. Finally, some observations and concluding remarks are offered in Section 7.

SECTION 2

LINEAR ESTIMATION AND COMPLEMENTARY STOCHASTIC PROCESSES

2.1 A Static Example

Before presenting the derivation of the estimator for a general class of noncausal stochastic processes, we illustrate the application of this approach for a familiar static problem of estimating a random vector. This example provides motivation for the complementary model approach in general, and in that there are many parallels between this static example and the more general problem we will ultimately address, the example also provides insight into the structure of the operator solution we obtain later.

Let ζ be an (n+p)-dimensional, zero mean random vector partitioned into $n\times 1$ and $p\times 1$ dimensional vectors x and r respectively as

$$\zeta = \begin{bmatrix} x \\ r \end{bmatrix} \begin{matrix} \leftarrow n\times 1 \\ \leftarrow p\times 1 \end{matrix} \tag{2.1a}$$

with invertible covariance matrix

$$\Sigma_\zeta = \begin{bmatrix} \Sigma_x & 0 \\ 0 & \Sigma_r \end{bmatrix} \quad . \tag{2.1b}$$

Let y be a p-dimensional observation ($p\leq n$)

$$y = M_y \zeta \tag{2.2a}$$

where M_y is the $p\times(n+p)$ matrix partitioned into a $p\times n$ matrix H and the $n\times n$ identity matrix I as

$$M_y = [\ \underset{\substack{\uparrow \\ p\times n}}{H} \ \vdots\ \underset{\substack{\uparrow \\ p\times p}}{I}\ ] \tag{2.2b}$$

with H full rank p. That is, we have the familiar linear observation:

$$y = Hx + r \quad . \tag{2.2c}$$

The vector to be estimated, x, can also be expressed as a linear function of ζ:

$$x = M_x \zeta \tag{2.3a}$$

where

$$M_x = \left[I \vdots 0 \right] \quad . \tag{2.3b}$$

Since both x and y are defined in terms of it, ζ will be referred to as the underlying random vector.

We show below that one way to calculate the linear minimum variance estimate of x given y is by establishing a complementary random vector z which has the following properties: 1) It is a linear function of the underlying random vector ζ

$$z = M_z \zeta \quad . \tag{2.4}$$

2) It is orthogonal to the observation y

$$E[yz'] = 0 \quad . \tag{2.5}$$

3) It is complementary with respect to y in that the augmented system

$$\begin{bmatrix} y \\ z \end{bmatrix} = M\zeta \quad ; \quad M = \begin{bmatrix} M_y \\ M_z \end{bmatrix} \tag{2.6}$$

is invertible.

Define the inverse of M as N and partition it compatibly with the dimensions of the vectors y and z as

$$M^{-1} \equiv N = \left[N_y \vdots N_z \right] \tag{2.7}$$

and denote the products of these partitions with y and z as

$$\zeta_y = N_y \, y \tag{2.8a}$$

and

$$\zeta_z = N_z \, z \quad . \tag{2.8b}$$

Then ζ can be written as the sum of orthogonal components

$$\zeta = \zeta_y + \zeta_z \quad . \tag{2.8c}$$

It follows from the orthogonality of ζ_y and ζ_z and from the Projection theorem that ζ_y is the linear minimum variance estimate of ζ given y (the projection of ζ onto span(y)), and that

$$\hat{x} = M_x \zeta_y \tag{2.9}$$

is the linear minimum variance estimate of x given y. In addition, from (2.3a), (2.8c) and (2.9) the estimation error can be written as a function of ζ_z:

$$\begin{aligned} \tilde{x} &= x - \hat{x} \\ &= M_x \zeta_z \end{aligned} \tag{2.10}$$

Thus the estimation error is the linear minimum variance estimate of x given z.

All of this is of little use if the matrix M_z is not known. It is shown in the Appendix that the three conditions stated above enable us to derive the following general expression for the n×(n+p) matrix M_z:

$$M_z = T[\ I \ \vdots\ -\Sigma_x H'\Sigma_r^{-1}] \tag{2.11}$$

where T is any invertible n×n matrix (indicating the obvious fact that the complementary process is only defined uniquely up to a choice of basis). We will see in the next section that the key to extending the method of complementary models to more general stochastic processes is the interpretation of the transpose of H in (2.11) as an adjoint mapping.

By performing the augmentation and inversion indicated in (2.6) through (2.8), the underlying process can be written as a linear funtion of y and z. If we define P as the matrix

$$P = [\Sigma_x^{-1} + H'\Sigma_r^{-1}H]^{-1} \qquad , \tag{2.12a}$$

then inverting the matrix M defined by substituting (2.2b) and (2.11) into

(2.6) gives the following expression for ζ

$$\zeta = \begin{bmatrix} x \\ \\ r \end{bmatrix} = \begin{bmatrix} P \\ - \; - \\ I-HP \end{bmatrix} H'\Sigma_r^{-1} y + \begin{bmatrix} P \\ - \; - \\ -HP \end{bmatrix} \Sigma_x^{-1} T^{-1} z \quad . \tag{2.12b}$$

From (2.3b) and (2.9) the estimate of x is given by

$$\hat{x} = PH'\Sigma_r^{-1} y \quad . \tag{2.12c}$$

From (2.10) the estimation error is

$$\tilde{x} = P\Sigma_x^{-1} T^{-1} z$$

or substituting for z from (2.4) and (2.11)

$$\tilde{x} = P[\Sigma_x^{-1} x - H'\Sigma_r^{-1} r] \quad . \tag{2.12d}$$

A direct calculation from (2.12d) and the defintion of P in (2.12a) gives the error covariance as

$$E\{\tilde{x}\tilde{x}'\} = P \quad . \tag{2.12e}$$

In summary, this static example illustrates the basic concepts of the method of complementary models. We have shown that two of the key elements in the development of the estimator are (1) the knowledge of the form of M_z and (2) the ability to easily invert the augmented system (2.6) to obtain the underlying variables ζ as a function of the observations y and the complementary process z. In the remainder of this section we formalize this approach to linear estimation as a restatement of the Projection theorem. Subsequently, we apply this theorem to establish an operator form for M_z which is appropriate for the general class of noncausal processes mentioned in the introduction.

2.2 The Projection Theorem Restated

Here we consider linear estimation for second order stochastic processes. Let $L_2(dP)$ denote the Hilbert space of finite variance random variables (on some given probability space). Let I denote an index set. A second order process over I is a set of elements in $L_2(dP)$ indexed by I:

$$\phi(\alpha) \in L_2(dP), \quad \alpha \in I. \tag{2.13}$$

The closed linear span in $L_2(dP)$ of ϕ (as α ranges over I) will be denoted by $Sp(\phi)$. The space of second order processes over I will be denoted by $L_2(I;dP)$. Linear mappings between two such spaces will be called second order operators.

With these definitions we can generalize the static example as follows. Define the underlying process as a second order process over a specified index set I_ζ

$$\zeta \in L_2(I_\zeta;\ dP) \equiv \mathbf{S}_\zeta \quad . \tag{2.14}$$

The process to be estimated X, the observations Y and the complementary process Z are defined via second order linear operators acting on ζ:

$$X = M_x\zeta \ ; \quad M_x : \mathbf{S}_\zeta \rightarrow L_2(I_x;dP) \equiv \mathbf{S}_x \tag{2.15a}$$

$$Y = M_y\zeta \ ; \quad M_y : \mathbf{S}_\zeta \rightarrow L_2(I_y;dP) \equiv \mathbf{S}_y \tag{2.15b}$$

and

$$Z = M_z\zeta \ ; \quad M_z : \mathbf{S}_\zeta \rightarrow L_2(I_z;dP) \equiv \mathbf{S}_z \tag{2.15c}$$

where M_x and M_y are known and M_z must be chosen (if possible) so that the following conditions are satisfied:

Orthogonality:

$$E[Y(\alpha)Z(\beta)] = 0 \quad \text{for all } \alpha \in I_y,\ \beta \in I_z \tag{2.16}$$

Complementation:

With M defined as the augmented second order operator

$$M = \begin{bmatrix} M_y \\ M_z \end{bmatrix} \tag{2.17a}$$

the relation between ζ and $\{Y,Z\}$:

$$\begin{bmatrix} Y \\ Z \end{bmatrix} = M\zeta \tag{2.17b}$$

is invertible. This implies that $\mathbf{S}_\zeta$ and $\mathbf{S}_y \times \mathbf{S}_z$ are isomorphic.

Assume that M_z can be found so that these conditions are satisfied. Partitioning the inverse of the augmented system (2.17) as we did for the static example ($M^{-1} = N = [N_y \vdots N_z]$), we will define the components of ζ as

$$\zeta_y = N_y Y \tag{2.18a}$$

and

$$\zeta_z = N_z Z \tag{2.18b}$$

so that

$$\zeta = \zeta_y + \zeta_z \quad . \tag{2.18c}$$

Projection Theorem: Given X, Y, and ζ as defined above and given an operator M_z and corresponding process Z which satisfies the stated orthogonality and complementation conditions, then

i) the linear minimum variance estimate of ζ given Y is

$$\hat{\zeta} = \zeta_y \tag{2.19}$$

ii) the linear minimum variance estimate of X given Y is

$$\begin{aligned} \hat{X} &= M_x \zeta_y \\ &= M_x N_y Y \end{aligned} \tag{2.20}$$

iii) the estimation error is the linear minimum variance estimate of X given Z

$$\begin{aligned} \tilde{X} &= M_x \zeta_z \\ &= M_x N_z M_z \zeta \quad . \end{aligned} \tag{2.21}$$

The proof follows from the same simple arguments used for the vector case in the static example. We remark that the theorem is applicable to processes defined on multidimensional index sets. In the next section we provide an example of such a processs.

The simple notation used to express the linear minimum variance estimate of X as a linear mapping of Y belies the complexity of the effort which may be required in (1) determining the form of the operator M_z, (2) augmenting and inverting to obtain M^{-1} and (3) implementing the solution. In the static example discussed in this section, a simple matrix form for M_z allowed for a direct inversion of the augmented system, yielding a simple form for the estimator and error equations. As one would expect, these three steps become substantially more difficult to accomplish in the more general case where we consider second order stochastic processes. For the estimation problems considered here, the mapping M_y in (2.15b) will be assumed to be of the same form as its matrix counterpart in the static example, i.e. $M_y = [H : I]$. Here I is the identity operator, and its action on the underlying process ζ produces the additive noise component to the observations. The operator H can be viewed as an input-output map describing the effect of the system dynamics. After introducing some notation and posing the general statement of the estimation problem in the next section, the first of the three issues stated above is addressed in Section 4 where a general operator representation for M_z is developed in the form of an input-output map. It will be shown that the form of the map M_z follows from that of the matrix M_z for the static example in (2.11). In particular, if one interprets the transpose of the matrix H in (2.11) as its Hilbert adjoint and appropriately interprets Σ_x and Σ_r as operators, then the operator form for M_z is of exactly the same form.

As we have just stated, a direct generalization of our static example is stated in terms of input-output representations for the observations and the complementary process. Unfortunately, working with these representations does not lead to a convenient or easily computed solution to the second step listed above, that is the augmentation of the observations with the complementary process and the inversion required to determine the estimator. However, by extending the approach taken by Levy et. al in [9], we will find that this second step is quite easily accomplished by working with internal differential realizations for the observations and complementary process. The internal realization for the observations is provided directly by the problem statement. We will see that an internal realization for the complementary process requires an internal realization for the Hilbert adjoint of H, H^*. A critical developement in our research has been the recognition that Green's identity (cf. Section 5.1) is the key to formulating an internal realization for the Hilbert adjoint map H^* in terms of the operators involved in the internal description of the observations. Given these internal realizations, we are able to perform the augmentation and inversion yielding an internal differential realization for the estimator. We feel that this representation for the estimator is an important one. In particular, if one directly applies the projection theorem to problems of the type which we consider here, the results are generally in the form of integral equations (e.g. Wiener-Hopf integral equations) which must be factored in some way in order to produce a realization for the estimator. In contrast, our solution, obtained via the method of complementary models, directly yields a differential realization of the estimator.

Much as in the case of causal processes described by finite-dimensional state equations, these realizations provide an excellent starting point for the construction of efficient algorithms for implementing the optimal estimator. This last step, determining smoothing algorithms, is only briefly discussed in this paper. However, in Part II [5] we present a detailed development of a two-filter implementation of this estimator for a noncausal one-dimensional two-point boundary value stochastic process. As can be seen in [5], it is a decidedly nontrivial step to go from an internal realization to an efficient implementation of estimators for problems of this type.

SECTION 3

MATHEMATICAL BACKGROUND AND THE GENERAL PROBLEM STATEMENT

3.0 Introduction

The noncausal stochastic processes for which we will be developing an estimator can be divided into two classes: 1) those with a continuous-valued independent variable and 2) those with a discrete-valued independent variable. More specifically, the processes in the first class are solutions of linear stochastic (partial) differential equations. Those in the second class are solutions of linear stochastic (partial) difference equations. In the first two parts of this section we introduce differential operator representations for each of the classes. By employing similar notation in the descriptions of each, we will be able to unify later discussions. An example will be provided for each of the two classes, and these examples will be carried along throughout the rest of the paper.

The purpose of the development in this section is to describe general recipes for constructing complementary processes and, more importantly, for expressing the solutions to an extremely broad class of estimation problems involving processes with independent variable of one and higher dimensions. In order to highlight the basic concepts underlying these recipes, we will not discuss in detail the technical conditions that must be satisfied in order for our most general boundary value problems to be well-posed (i.e. for existence and uniqueness of solutions to the specified stochastic differential equations) but rather we will, in effect, assume that these conditions are met. Clearly, in any application one must verify the appropriate conditions, and we will illustrate this for our two examples.

3.1 The Continuous Parameter Case

3.1.1 Differential Operators and Green's Identity

Our stochasitc differential equations are defined in terms of differential operators acting on Hilbert spaces of square-integrable functions as follows. Let Ω_N be a bounded convex region in $\mathbf{R}^N$ with smooth boundary [11]. The space of $n \times 1$ vector functions which are square-integrable on Ω_N is represented by $L_2^n(\Omega_N)$. Let L be a formal[1] differential operator defined on D(L), the subspace of sufficiently differentiable elements of $L_2^n(\Omega_N)$, so that

$$L: D(L) \to L_2^n(\Omega_N) \tag{3.1}$$

Note that D(L) is dense in $L_2^n(\Omega_N)$. With $\partial\Omega_N$ denoting the boundary of Ω_N, define a boundary condition associated with L through the mapping

$$V: D(L) \to L_2^{n_v}(\partial\Omega_N) \tag{3.2}$$

where the dimension n_v is briefly discussed below.

We will say that the pair (L,V) leads to a well-posed boundary value problem if the differential operator Λ formed by augmenting the formal differential operator L and boundary mapping V

$$\Lambda = \begin{bmatrix} L \\ V \end{bmatrix} \tag{3.3a}$$

has a unique continuous left inverse:

$$\Lambda^{\#}\Lambda = I \quad . \tag{3.3b}$$

We denote the components of the left inverse by

$$\Lambda^{\#} = [\ G_u \vdots\ G_v] \tag{3.3c}$$

where

$$G_u : L_2^n(\Omega_N) \to D(L) \quad \text{and} \quad G_v : L_2^{n_v}(\partial\Omega_N) \to D(L) \quad . \tag{3.3d}$$

[1] The term <u>formal differential operator</u> will be used to denote operators which simply represent differentiation of a function. We will reserve the term <u>differential operator</u> to denote the combined action of a formal differential operator along with an appropriate boundary condition (see (3.3a)). That is, a differential operator implicitly defines an input-output map obtained by solving a well-posed boundary value problem.

In this case, the equation

$$\Lambda x = \begin{bmatrix} u \\ v \end{bmatrix} \tag{3.4a}$$

with u and v in the domains of G_u and G_v, respectively, has a unique solution which can be written as

$$x = G_u u + G_v v \quad . \tag{3.4b}$$

Thus for a given set of inputs u and v, x is unique and varies continuously with those inputs. The value of the vector dimension n_v in (3.2) which is required for a well-posed problem depends on the type and order of the operator L and the dimensions N and n. As mentioned earlier, we will assume that we have well-posed problems here. As discussed next for the examples, equation (3.4c) is the Green's function solution of (3.3).

Example 1:(Poisson's Equation) To illustrate the preceding development we consider Poisson's equation in $\mathbf{R}^2$ with Ω_2 the unit disk and with a Dirichlet boundary condition. In this case the formal differential operator is the Laplacian ∇^2 and its domain is the space $C_2(\Omega_2)$ of scalar (i.e. n=1) functions on Ω_2 with bounded continuous second partials.

$$L: C_2(\Omega_2) \to L_2(\Omega_2) \quad ; \quad Lx = \nabla^2 x \tag{3.5a}$$

The boundary operator V is the restriction of x to its values on the boundary of Ω_2:

$$V: C_2(\Omega_2) \to L_2(\partial\Omega_2) \quad ; \quad Vx = x\Big|_{\partial\Omega_2} \tag{3.5b}$$

(here $n_v = 1$). The Green's function solution for the pair Lx=u and Vx=v, where u and v are in the ranges of L and V as specified in (3.5), is shown in [3] to be given in polar coordinates as follows. Define the kernels g_v and g_u as

$$g_v(\rho,\theta;\beta) = \frac{1}{2\pi} \cdot \frac{(1-\rho^2)}{1 - 2\rho\cos(\theta-\beta) + \rho^2} \tag{3.6a}$$

and

$$g_u(\rho,\theta;\gamma,\beta) = \frac{1}{4\pi^2} \log\left\{ \frac{\rho^2 - 2\rho\gamma\cos(\theta - \beta) + \gamma^2}{1 - 2\rho\gamma\cos(\theta - \beta) + \rho^2\gamma^2} \right\} \quad . \tag{3.6b}$$

The solution x is given by

$$x(\rho,\theta) = \int_0^{2\pi} g_v(\rho,\theta;\beta)v(\beta)d\beta + \int_0^{2\pi}\int_0^1 g_u(\rho,\theta;\gamma,\beta)u(\gamma,\beta)d\gamma d\beta \quad . \tag{3.6c}$$

Note that the boundary value contribution to the solution in (3.6c) has been written as an integral over the interval $[0,2\pi]$. This is, of course, an integral over the boundary of the unit disk, i.e. the unit circle. $\equiv$

As indicated in the previous section, Green's identity applied to the internal dynamics of the process to be estimated plays a key role in the construction of an internal realization of the complementary process and ultimately in the inversion required to solve the smoothing problem. In general, when it exists, Green's identity is obtained from integration by parts of the the N-fold integral specified by the following inner product:

$$\langle Lx,\lambda\rangle_{L_2^n(\Omega_N)} \quad .$$

This yields Green's Identity in the form

$$\langle Lx,\lambda\rangle_{L_2^n(\Omega_N)} = \langle x,L^\dagger\lambda\rangle_{L_2^n(\Omega_N)} + \text{boundary term} \quad . \tag{3.7}$$

Here $L^\dagger$ is also a formal differential operator of the same order as L and is referred to as the formal adjoint differential operator. The boundary term is in the form of an integral over the boundary $\partial\Omega_N$. This integral involves the processes x and λ and perhaps their derivatives evaluated on $\partial\Omega_N$. In general, the precise form of (3.7) is

$$\langle Lx,\lambda\rangle_{L_2^n(\Omega_N)} = \langle x,L^\dagger\lambda\rangle_{L_2^n(\Omega_N)} + \langle x_b,E\lambda_b\rangle_{H_b} \tag{3.8}$$

where x_b and λ_b are elements of a Hilbert space H_b of processes defined on $\partial\Omega_N$ and $E:H_b \to H_b$. In particular, these processes are defined through the action of an operator Δ_b:

$$\Delta_b : L_2^n(\Omega_N) \to H_b \tag{3.9a}$$

$$x_b = \Delta_b x \tag{3.9b}$$

$$\lambda_b = \Delta_b \lambda \quad . \tag{3.9c}$$

The nature of H_b, Δ_b, and E all depend upon L and Ω_N. Green's identity for ordinary differential operators can be found in [4]; for elliptic, hyperbolic and parabolic second order partial differential operators see [3]. In addition to the well-posedness assumption on (L,V), we will also restrict ourselves to operators L and regions Ω_N that admit a Green's identity.

Example 1 Continued:

Recall that in this example x is a scalar function (so that n=1), and Ω_2 is the unit disk. Performing the integration by parts on

$$\langle\nabla^2 x,\lambda\rangle_{L_2(\Omega_2)} = \iint_{\Omega_2}\left(\nabla^2 x(s,t)\right)\lambda(s,t)dsdt \tag{3.10a}$$

it can be shown that the Laplacian ∇^2 is formally self-adjoint [3], i.e.

$$\langle\nabla^2 x,\lambda\rangle = \langle x,\nabla^2\lambda\rangle + \text{boundary term} \quad . \tag{3.10b}$$

The boundary term, for this example is expressed as follows. With x_n the normal derivative of x along $\partial\Omega_N$, define the function x_b as

$$x_b = \Delta_b x = \begin{bmatrix} x\big|\partial\Omega_N \\ x_n \end{bmatrix}, \text{ or in polar coordinates } x_b(\theta) = \begin{bmatrix} x(1,\theta) \\ x_n(1,\theta) \end{bmatrix} \quad . \tag{3.11}$$

Thus, in this case x_b is an element of the Hilbert space

$$H_b = L_2^2(\partial\Omega_2)$$

with inner product

$$\langle w,z\rangle_{H_b} = (1/2\pi)\int_0^{2\pi}\left[w_1(\theta)z_1(\theta) + w_2(\theta)z_2(\theta)\right]d\theta \quad . \tag{3.13}$$

The function λ_b is defined in terms of λ in the same fashion as x_b in (3.11). Furthermore, the operator E in this case is simply the multiplication of elements of $L_2^2(\partial\Omega_2)$ by a 2×2 matrix which we also denote by E. Specifically,

$$E = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad . \tag{3.14}$$

Combining these, we have that the boundary term in Green's identity for this example is

$$\langle x_b, E\lambda_b\rangle_{L_2^2(\partial\Omega_2)} = \frac{1}{2\pi}\int_0^{2\pi}\left(\lambda(1,\theta)x_n(1,\theta) - x(1,\theta)\lambda_n(1,\theta)\right)d\theta \quad . \tag{3.15}$$

■

3.1.2 Operator Representations of Stochastic Processes

While we have so far been discussing deterministic functions, we can use the same formalism to describe stochastic processes. In particular, we will be concerned with stochastic processes defined as in (3.4b) where u and v are input stochastic processes. As is usually done, we will often represent such processes via a stochastic differential equation written as in (3.4a).

In general for the problems of interest here the operators G_u and G_v are integral operators over Ω_N and $\partial\Omega_N$ respectively[2] with piecewise continuous kernels. That is, they are integral operators of the form

$$G:\ L_2^n(\Gamma) \to L_2^m(\Gamma)$$

$$(Gf)(\underline{t}) = \int_\Gamma g(\underline{t},\underline{s})f(\underline{s})d\underline{s} \quad ; \quad \underline{t},\underline{s} \in \Gamma \tag{3.16a}$$

where Γ may be either Ω_N or $\partial\Omega_N$ and $d\underline{s}$ is the corresponding

[2] The only exception is in the case N = 1, in which Ω_N is an interval and $\partial\Omega_N$ consists of its two endpoints. In this case G_v reduces to a matrix operating on the vector v.

infinitessimal area element. Here we have underlined the independent variables $\underline{t}$ and $\underline{s}$ to emphasize the fact that they are multidimensional indices belonging to the region Γ. The adjoint of this operator is of the same form as (3.16a) but with kernel

$$g^*(\underline{t},\underline{s}) = g'(\underline{s},\underline{t}) \quad . \tag{3.16b}$$

By extending the class of kernels to include the multidimensional Dirac delta function, $\delta(\underline{t}-\underline{s})$, over Γ, we can express products of the form

$$g(\underline{t})f(\underline{t}) \quad ; \qquad \underline{t} \in \Gamma \tag{3.17a}$$

as

$$(Gf)(\underline{t}) = \int_\Gamma g(\underline{t})\delta(t-\underline{s})f(\underline{s})d\underline{s} \quad . \tag{3.17b}$$

Note that in the case in which $\Gamma = \partial\Omega_N$, the Dirac delta function $\delta(\underline{t}-\underline{s})$ represents the measure on the manifold $\partial\Omega_N$ which is concentrated at the point t and has unit value.

In addition to the operators described above we will also encounter integral operators which map $L_2^n(\Gamma)$ into $\mathbf{R}^m$ and are of the form

$$Gf = \int_\Gamma g(\underline{s})f(\underline{s})d\underline{s} \tag{3.18}$$

and those of the form of the adjoint of (3.18) which map $\mathbf{R}^m$ into $L_2^n(\Gamma)$:

$$(G^*f)(\underline{s}) = g'(\underline{s})f \quad . \tag{3.19}$$

Each of the integral operators in (3.16) through (3.18) can be applied to mean-square continuous processes with the output being either a mean-square continuous process or a finite variance random vector. In the case of the operator defined in (3.19), if the input is a finite variance random vector, then the output is a mean-square continuous process.

We can formally extend the domain of the operators in (3.16a) and (3.18) to include white noise as follows. Consider those operators with $\Gamma = \Omega_N$. In this case, as in [12], we can define Wiener integrals with respect to multidimensional Wiener processes. Similar to the one-dimensional case, these multidimensional Wiener integrals can be represented formally by multidimensional white noise integrals. Specifically, let $w(\underline{t})$ denote an N-dimensional white noise process with

$$E[w(\underline{t})w'(\underline{s})] = Q(\underline{t})\delta(\underline{t}-\underline{s}) \quad . \tag{3.20}$$

(Here δ is again the N-dimensional delta function.) We will use the following white noise integral representation

$$\int_{\Omega_N} g(\underline{t},\underline{s})w(\underline{s})d\underline{s} \tag{3.21}$$

to denote the Wiener integral over Ω_N with respect to the kernel g. With g continuous, (3.21) defines a mean-square continuous process. Note that the critical property of such white noise integrals is the following: If we define two processes x and y as

$$x(\underline{t}) = \int_{\Omega_N} g_1(\underline{t},\underline{s})w(\underline{s})d\underline{s} \qquad \text{and} \qquad y(\underline{t}) = \int_{\Omega_N} g_2(\underline{t},\underline{s})w(\underline{s})d\underline{s} \ ,$$

then their correlation function is given by

$$E[x(\underline{t})y'(\underline{\tau})] = \int_{\Omega_N} g_1(\underline{t},\underline{s})Q(\underline{s})g_2'(\underline{\tau},\underline{s})d\underline{s} \quad . \tag{3.22}$$

In an analogous fashion, we can define white noise integrals over the smooth, closed (N-1)-dimensional manifold $\partial\Omega_N$. Specifically, if we now let $w(\underline{t})$ denote a white noise process on $\partial\Omega_N$ and if we write

$$x(\underline{t}) = \int_{\partial\Omega_N} g_1(\underline{t},\underline{s})w(\underline{s})d\underline{s} \qquad \text{and} \qquad y(\underline{t}) = \int_{\partial\Omega_N} g_2(\underline{t},\underline{s})w(\underline{s})d\underline{s} \ , \tag{3.23a}$$

then

$$E[x(\underline{t})y'(\underline{\tau})] = \int_{\partial\Omega_N} g_1(\underline{t},\underline{s})Q(\underline{s})g_2'(\underline{s},\underline{\tau})d\underline{s} \tag{3.23b}$$

where in this case $d\underline{s}$ represents an infinitessimal area element on $\partial\Omega_N$. Note that if the support of $g_1(\underline{t},\cdot)$ and $g_2(\underline{t},\cdot)$ in (3.23) are on subsets of $\partial\Omega_N$ which are homeomorphic to a subset of an (N-1)-dimensional region then the integrals in (3.23) are precisely the same as those in (3.22) except on a space of dimension (N-1). For this reason, we will often represent the covariance function of a white noise process on $\partial\Omega_N$ exactly as in (3.20) where in this case the Dirac delta function on $\partial\Omega_N$ is to be interpreted as described previously.

Applying this formalism, we will consider the differential equation in (3.3a) with white noise inputs as a formal representation for the mean-square continuous process defined via the integral representations in (3.4b) where u is now an $n\times 1$ vector white noise over Ω_N with covariance

$$E[u(\underline{t})u'(\underline{s})] = Q(\underline{t})\ \delta(\underline{t}-\underline{s}) \qquad , \qquad \underline{t},\underline{s} \in \Omega_N \tag{3.24a}$$

and in problems for which $N > 1$, v is an $n_v\times 1$ vector white noise process over $\partial\Omega_N$ with covariance

$$E[v(\underline{\sigma})v'(\underline{\tau})] = \Pi_v(\underline{\tau})\ \delta(\underline{\tau}-\underline{\sigma}) \qquad , \quad \underline{\tau},\underline{\sigma} \in \partial\Omega_N \quad . \tag{3.24b}$$

If N=1, for example when Ω_N is the interval [0,T] so that $\partial\Omega_N = \{0,T\}$, then v is simply an $n_v\times 1$ random vector with covariance matrix Π_v.

Throughout the paper we will assume that both Π_v and Q are continuous in their arguments. Thus, given the continuity assumptions for the integration kernels in the Green's function solution and the continuity of Q and Π_v, x will be mean-square continuous. For the example of Poisson's equation on the unit disk, the dynamics in (3.5a) and the Dirichlet boundary condition in (3.5b) formally represent a 2-D random field x when the input u is a 2-D white noise over the unit disk and v is a 1-D white noise on the unit circle or, equivalently, on the interval $[0,2\pi]$.

3.1.3 Correlation Operators

In this subsection we introduce some notation for representing the correlation functions of mean-square continuous stochastic processes as operators on Hilbert spaces of deterministic functions. In particular, the correlation functions of these stochastic processes can be viewed as kernels of such operators. In addition, we will also associate the correlation and cross-correlation functions of processes generated by second order mappings of the type discussed above with kernels of composite operators. Although we make use of this notation in stating the general estimation problem, its true worth will become evident later in Section 4.1 where we formulate an operator representation for complementary processes.

The correlation function of a mean-square continuous process can be viewed as the kernel of an operator as follows. Let $z(\underline{t})$ be a mean-square continuous process on Ω_N. Its correlation function

$$\Sigma_z(\underline{t},\underline{s}) = E[z(\underline{t})z'(\underline{s})] \tag{3.25}$$

is continuous on $\Omega_N \times \Omega_N$ and as such can be considered the kernel of an operator, which we will denote by Σ_z, of the same type as in (3.16a). Similarly, the correlation function for the white noise process w in (3.20),

$$\Sigma_w(\underline{t},\underline{s}) = Q(\underline{t})\delta(\underline{t}-\underline{s}) \quad , \tag{3.26}$$

can be viewed as the kernel of an operator Σ_w. In addition, the covariance matrix of a random vector can be viewed as an operator on a finite-dimensional space. Indeed, each of the operators defined in (3.16) through (3.19) can be associated with the cross-correlations of zero mean processes with each other or with random vectors.

More generally, the correlation functions of stochastic processes defined by a second order mapping of the type in (3.16) through (3.19) can be represented in terms of the composition of operators as follows. For example, let $r(\underline{s})$ be a mean-square continuous process with correlation operator Σ_r. Let the process $z(\underline{t})$ be defined by

$$z(\underline{t}) = (Gr)(\underline{t}) = \int_{\Omega_N} g(\underline{t},\underline{s})r(\underline{s})d\underline{s} \quad . \tag{3.27}$$

The correlation function of z and the cross-correlation function of z and r are the kernels of the operators

$$\Sigma_z = G\Sigma_r G^* \tag{3.28a}$$

and

$$\Sigma_{rz} = \Sigma_r G^* \quad . \tag{3.28b}$$

These are easily checked by directly computing the corresponding kernels from (3.27a). Note that these formulas are valid when r is either a second order process or a white noise process.

3.2 Discrete Parameter Stochastic Processes

In a parallel but much briefer fashion we define the class of discrete parameter stochastic processes to be considered. In order to unify later discussions to include both classes, we employ much of the same operator notation in describing the discrete process as was used to describe the continuous processes.

Let I_N denote the space of n-tuples $(i_1,\ldots,i_N)$ where each of the i_k is an integer. Let Ω_N be a bounded region in I_N and let $l_2^n(\Omega_N)$ be the space of square-summable $n\times 1$ vector sequences on Ω_N. For the discrete case, L represents a formal linear difference operator. As we will see below in the example, the support of the sequences in the range of a difference operator will be different from the support of those sequences in its domain. By properly defining the boundary set $\partial\Omega_N$, one can define the support of the sequences in the range of the formal difference operator as the union of Ω_N and $\partial\Omega_N$ (Again, the example will help clarify this point). Thus L is the mapping

$$L:\ l_2^n(\Omega_N \cup \partial\Omega_N) \rightarrow l_2^n(\Omega_N) \quad . \tag{3.29a}$$

Let V be an operator mapping[3]

$$V:\ l_2^n(\Omega_N \cup \partial\Omega_N) \rightarrow l_2^{n_v}(\partial\Omega_N) \tag{3.29b}$$

where, as in the continuous case, the value of the dimension n_v is such that the difference operator formed by combining L and V

$$\Lambda = \begin{bmatrix} L \\ V \end{bmatrix} \tag{3.29c}$$

has a unique left inverse, i.e. so that the problem is well-posed.

---

3 As in the continuous paramenter case, when N = 1, the range of the boundary operator is simply R^{n_v}.

As in the continuous case, we will consider only those L and Ω_N which admit a Green's identity of the form

$$\langle Lx,\lambda\rangle_{l_2^n(\Omega_N)} = \langle x,L^{\dagger}\lambda\rangle_{l_2^n(\Omega_N)} + \langle x_b, E\lambda_b\rangle_{H_b} \tag{3.30}$$

where $L^{\dagger}$ is the formal adjoint difference operator and the nature of E, x_b, and the Hilbert space H_b all depend on L and Ω_N. Much as in the continuous case, x_b has support on $\partial\Omega_N$, and H_b is a Hilbert space of square-summable sequences with support on $\partial\Omega_N$. In contrast to the continuous case, Green's identity is not typically employed in the solution of boundary value difference equations, and therefore, it is not usually found in texts on difference equations. However, it can be derived in the same manner as its continuous counterpart, the difference being that integration by parts is replaced with summation by parts. The Green's identities for both one- and two-dimensional difference operators are derived in [6]. Below we present the Green's identity for a particular 1-D example.

Example 2: (1-D Discrete Boundary Value Problem) Let x be an n×1 vector 1-D discrete process, and let D denote the 1-D delay

$$(Dx)_k = x_{k-1} \quad , \tag{3.31}$$

Consider the 1-D difference operator

$$L = (D^{-1}I - A) \quad ; \; (Lx)_k = x_{k+1} - A_k x_k \quad . \tag{3.32}$$

If we define $\Omega_1 = [0,K-1]$ and its boundary as $\partial\Omega_1 = \{0,K\}$, then the range and domain of L are properly specified by

$$L:\; l_2^n[\Omega_1 \cup \partial\Omega_1] \rightarrow l_2^n[\Omega_1]$$

or

$$L:\; l_2^n[0,K] \rightarrow l_2^n[0,K-1] \quad .$$

This example already illustrates one important point. Due to sequencing issues for discrete dynamics, it will in general be the case that $\partial\Omega_N$ is neither disjoint from nor a subset of Ω_N. The Green's identity for this example is

$$\langle Lx,\lambda\rangle_{l_2^n[0,K-1]} = \langle x,L^\dagger\lambda\rangle_{l_2^n[0,K-1]} + \langle x_b,E\lambda_b\rangle_{\mathbf{R}^{2n}} \tag{3.33}$$

where the formal adjoint difference operator is

$$L^\dagger = (I - A'D^{-1}) \qquad ; \quad (L^\dagger x)_k = x_k - A_k'x_{k+1} \quad , \tag{3.34}$$

the boundary process is

$$x_b = \Delta_b x = \begin{bmatrix} x_0 \\ x_K \end{bmatrix} \tag{3.35a}$$

and E is a 2n×2n matrix partitioned into n×n blocks as

$$E = \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} \quad . \tag{3.35b}$$

Combining these definitions, the Green's identity can be written as

$$\sum_{k=0}^{K-1}\left(x_{k+1} - A_k x_k\right)'\lambda_{k+1} = \sum_{k=0}^{K-1}\left(\lambda_k - A_k'\lambda_{k+1}\right)'x_k - x_0'\lambda_0 + x_K'\lambda_K \quad . \tag{3.35c}$$

We will assume a two-point boundary condition for this discrete example as follows. With V the n×2n matrix:

$$V = [\; V^0 \vdots\; V^K \;] \quad , \tag{3.36a}$$

the product

$$v = V\, x_b \tag{3.36b}$$

defines a two-point boundary condition (here $v \in \mathbf{R}^n$, i.e. $n_v = n$). Let $\Phi[i,k]$ denote the transition matrix for A in (3.27). Then it can be shown that the pair (L,V) is well-posed if the n×n matrix

$$F = V^{0} + V^{K}\Phi[K,0] \tag{3.36c}$$

is invertible [6]. ■

Returning to the general discrete case, a formal representation for a second order discrete stochastic process x is given by

$$\Lambda x = \begin{bmatrix} L \\ V \end{bmatrix} x = \begin{bmatrix} u \\ v \end{bmatrix} \quad ; \qquad \Lambda : l_2^n(\Omega_N \cup \partial\Omega_N) \to \left(l_2^n(\Omega_N) \times l_2^{n_v}(\partial\Omega_N)\right)$$

where u is now an n×1 vector discrete white noise process over Ω_N and v is a second order discrete process whose support is contained in $\partial\Omega_N$.

3.3 The Estimation Problem Statement

The process to be estimated is either defined on a continuous or discrete index set as described in the first two parts of this section. In stating the estimation problem, we unify the discussion to include both classes.

Let L be a formal linear differential or difference operator with range R(L) and domain D(L), where elements in each are n×1 vector functions with index set Ω_N. Let B(t) be an n×m matrix with $t \in \Omega_N$. For the continuous parameter case B(t) is assumed continuous in t. Let H_b be the Hilbert space of n_b×1 vector functions whose support is the boundary $\partial\Omega_N$. Recall that the dimension n_b is determined from Green's identity for L and Ω_N. Let V be a mapping from H_b to R(V) where the nature of the range space R(V) is determined by the well-posedness condition for the pair (L,V) (see, for example, (3.29b) and the discussion that follows this equation). As opposed to the more general form Vx = v for the boundary condition, we will restrict the boundary condition to be defined in terms of x_b as indicated in (3.37b) below. For example, in the discrete case we will consider only those operators V defined in (3.29b) which map sequences on $\partial\Omega_N$ and therefore can be thought of as having $l_2^n(\partial\Omega_N)$ as their domain (see Example 2, and in particular (3.36), for a specific illustration of such a map V).

Let u be an m×1 vector white noise on Ω_N with an invertible correlation operator Q. Let v be an n_v×1 vector second order process over $\partial\Omega_N$, uncorrelated with u and with invertible correlation operator Π_v. Then the the process to be estimated is formally defined by

$$Lx = Bu \tag{3.37a}$$

with boundary condition

$$Vx_b = v \quad . \tag{3.37b}$$

The observations are defined as follows. Let C(t) be a p×n matrix, $t \varepsilon \Omega_N$. For the continuous parameter case it is assumed that C is continuous in t. Let W be an operator mapping elements of H_b into R(W), a space of $n_w \times 1$ vector functions defined over the index set $\partial\Omega_N$. Let r be a p×1 vector white noise over Ω_N with invertible correlation operator R, and let r_b be a $n_w \times 1$ vector process with invertible correlation operator Π_b. It will be assumed that u, v, r and r_b are mutually uncorrelated. The set of observations of x is given by:

$$y = Cx + r \qquad \text{on } \Omega_N \tag{3.38a}$$

and

$$y_b = Wx_b + r_b \qquad \text{on } \partial\Omega_N \quad . \tag{3.38b}$$

We will need to make some assumptions with respect to the relationship between the operators V and W. The importance of these assumptions will become apparent later in our development of Hilbert adjoint systems in Section 5.1. As explained in the 1-D continuous case studied in Part II, one consequence of these assumptions is that no element of the boundary observation y_b can simply be absorbed into updating the boundary value v alone. That is, the boundary measurement contains information about the part of x_b not captured by Vx_b. In particular, we will assume that W and V are linearly independent, i.e. for any linear operators M_v and M_w whose range spaces are identical and whose domains are the range spaces of V and W respectively, we must have

$$M_v V + M_w W = 0 \quad \text{iff} \quad M_v = 0 \text{ and } M_w = 0 \quad . \tag{3.39a}$$

Furthermore, we will assume that if the operator obtained by augmenting V and W as

$$\begin{bmatrix} V \\ W \end{bmatrix} \tag{3.39b}$$

is not invertible, then there exists an operator W_c such that

$$\begin{bmatrix} V \\ W \\ W_c \end{bmatrix} \tag{3.39c}$$

is invertible.

Our estimation problem is to find the linear minimum variance estimate of x given the set

$$Y = \{ y, y_b \} \quad . \tag{3.40}$$

To transform this problem into notation similar to that used for the static example, let the inverse of (3.37) be denoted by

$$x = M_x \begin{bmatrix} u \\ v \end{bmatrix} \quad . \tag{3.41}$$

This is simply the Green's function solution from (3.4b) written in different notation. Recall from (3.9b) that $x_b = \Delta_b x$, so that the combined observation in (3.40) can be written as

$$\begin{bmatrix} y \\ y_b \end{bmatrix} = \begin{bmatrix} C \\ W\Delta_b \end{bmatrix} M_x \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} r \\ r_b \end{bmatrix} \quad . \tag{3.42}$$

If we define H as

$$H = \begin{bmatrix} C \\ W\Delta_b \end{bmatrix} M_x \tag{3.43}$$

and specify the underlying process as

$$\zeta = \begin{bmatrix} u \\ v \\ r \\ r_b \end{bmatrix} \quad , \tag{3.44a}$$

then the observations can be expressed in a form similar to (2.2):

$$\begin{vmatrix} y \\ y_b \end{vmatrix} = \left[\, H \; \vdots \; I \, \right] \zeta \quad . \tag{3.44b}$$

Below we illustrate the problem statement for each of our two examples. In the succeeding sections, we formulate the solution to this class of problems in a differential operator form with y and y_b as the input and boundary condition repectively and the estimate of x as an element of the output.

Examples

Continuous case: (Example 1 continued) In this case Ω_2 is the unit disk and points within the disk will be represented by index variables $s,t \in \Omega_2$. Points on the unit circle $\partial\Omega_2$ will be dented by an angle $\theta \in [0,2\pi]$. Let u be a scalar white noise over Ω_2 with continuous covariance parameter Q(s). Let v be a scalar white noise over $\partial\Omega_2$ with continuous covariance parameter $\Pi_v(\theta)$ (which, of course, is periodic with period 2π). Let B(s) be a continuous function on Ω_2 and $V(\theta)$ be a nonzero continuous function on $\partial\Omega_2$. The process to be estimated is formally defined by

$$\nabla^2 x(s) = B(s)u(s) \tag{3.45a}$$

with boundary condition (in polar coordinates)

$$V(\theta)x(1,\theta) = v(\theta) \quad . \tag{3.45b}$$

Let r be a scalar white noise over Ω_2 and r_b be a scalar white noise over $\partial\Omega_2$ with continuous covariance parameters R(s) and $\Pi_b(\theta)$ respectively. Let C(s) be a continuous function on Ω_2 and $W(\theta)$ be a nonzero continuous function on $\partial\Omega_2$. The observations are defined by

$$y(s) = C(s)\, x(s) + r(s) \qquad \text{on } \Omega_2 \tag{3.46a}$$

and

$$y_b(\theta) = W(\theta)x_n(\rho,\theta) + r_b(\theta) \quad ; \quad \rho = 1 \quad . \tag{3.46b}$$

The estimation problem is to find the least squares estimate of x given y on Ω_2 and y_b on its boundary. Note that since the augmented operator

$$\begin{bmatrix} V \\ W \end{bmatrix} x_b = \begin{bmatrix} V(\theta) & \vdots & 0 \\ - - & \vdots & - - \\ 0 & \vdots & W(\theta) \end{bmatrix} \begin{bmatrix} x(1,\theta) \\ x_n(1,\theta) \end{bmatrix}$$

is invertible, there will be no complementing operator W_c as in (3.39c). ∎

<u>Discrete case:</u> (Example 2 continued) Recall that Ω_1 is the set of integers $[0,K-1]$, and $\partial\Omega_1$ is the set $\{0,K\}$. Let u be a $m\times 1$ vector white noise over Ω_1 with nonsingular covariance matrix Q_k, $k \in \Omega_1$. Let v be a $n\times 1$ random vector with nonsingular covariance matrix Π_v. Let B_k be an $n\times m$ matrix and A_k be a $n\times n$ matrix both on Ω_1, and let V be a full rank $n\times 2n$ matrix with $n\times n$ partitions $[V^0 \vdots V^K]$. The process to be estimated is defined by the difference equation

$$x_{k+1} = A_k x_k + B_k u_k \tag{3.47a}$$

with a two-point boundary condition

$$v = V^0 x_0 + V^K x_K \quad . \tag{3.47 }$$

To define the observations, let r be a $p\times 1$ white noise over Ω_1 whose covariance matrix R_k is nonsingular on Ω_1, and let r_b be a $q\times 1$ random vector with nonsingular covariance matrix Π_b. Let C_k be a $p\times n$ matrix on Ω_1 and let W be a full rank $q\times 2n$ matrix with $q \leq n$, with the rows of W linearly independent of the rows of V and with $q\times n$ partitions: $[W^0 \vdots W^K]$. Then the observations are given by

$$y_k = C_k x_k + r_k \qquad \text{on } \Omega_1 \tag{3.48a}$$

along with the random vector

$$y_b = W^0 x_0 + W^K x_K + r_b \quad . \tag{3.48b}$$

For both examples the input processes u and v and the observation noises r and r_b are all assumed to be mutually uncorrelated. For this example, when

$q < n$, the augmented $(n+q)\times 2n$ matrix in the following equation

$$\begin{bmatrix} V \\ \\ W \end{bmatrix} x_b = \left[\begin{array}{c:c} V^0 & V^K \\ \hdashline W^0 & W^K \end{array}\right] \begin{bmatrix} x_0 \\ \\ x_K \end{bmatrix} \tag{3.49}$$

will not be invertible. Thus, to attain an invertible matrix we must choose W_c as an $(n-q)\times 2n$ matrix whose rows are linearly independent of the rows of V and W so that

$$\begin{bmatrix} V \\ W \\ W_c \end{bmatrix}$$

is invertible. As we will see in Section 5.4, we only need to actually construct such a matrix in those cases for which Π_V is singular.

SECTION 4

OPERATOR FORM FOR M_z

4.0 Introduction

Based on the matrix form for M_z established for the static example in Section 2.1, in this section we present an expression for a mapping of the underlying process defining our estimation problem, and we then prove that the resulting process satisfies the orthogonality and complementation conditions required of the complementary process. Only the continuous parameter case is addressed here; however, with a few obvious changes the same arguments can be adapted to the discrete parameter case. Indeed, since all discrete stochastic processes can be represented by a (possibly very large) random vector, the matrix representation for M_z from the static example in Section 2.1 is applicable for the discrete case.

4.1 An Operator Representation for the Complementary Process

It will be convenient to partition the underlying process ζ into two parts denoted by ζ_1 and ζ_2. The first part ζ_1 corresponds to the boundary value and input process, and the second part ζ_2 represents the additive noise on the observations:

$$\zeta = \begin{bmatrix} \zeta_1 \\ -- \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} u \\ v \\ -- \\ r \\ r_b \end{bmatrix} . \tag{4.1}$$

The covariance parameters of the elements of ζ are assumed to be continuous and the covariance parameters and covariance matrices are all assumed invertible. As discussed previously the second order statistics of ζ can be defined by way of a correlation operator. The range and domain of this operator are identical and are defined by the following spaces:

$$\mathbf{S}_1 = L_2^n(\Omega_N) \times L_2^{n_v}(\partial\Omega_N) \quad , \tag{4.2a}$$

$$\mathbf{S}_2 = L_2^p(\Omega_N) \times L_2^q(\partial\Omega_N) \quad \text{and} \quad \mathbf{S} = \mathbf{S}_1 \times \mathbf{S}_2$$

As discussed earlier in Section 3.1, when $\partial\Omega_N$ is finite (i.e. when N = 1), the L_2 spaces of functions over $\partial\Omega_N$ should be replaced by the Euclidian spaces R^{n_v} and R^q. The correlation operator Σ_ζ is the self-adjoint invertible mapping

$$\Sigma : \mathbf{S} \rightarrow \mathbf{S} \tag{4.3a}$$

which we will express in partitioned form as

$$\Sigma_\zeta = \begin{bmatrix} \Sigma_{\zeta_1} & 0 \\ 0 & \Sigma_{\zeta_2} \end{bmatrix} \quad . \tag{4.3b}$$

The observations are defined via the operator M_y

$$M_y : \mathbf{S} \rightarrow \mathbf{S}_2 \tag{4.4a}$$

where from (3.40b)

$$M_y = [\ H \ \vdots \ I\ ] \quad . \tag{4.4b}$$

Thus the observation Y is given by the second order mapping

$$\begin{aligned} Y &= M_y\zeta \\ &= H\zeta_1 + \zeta_2 \quad . \end{aligned} \tag{4.5}$$

The following theorem establishes an operator representation of the complementary process for this set of observations.

<u>Theorem:</u> (Complementary Process) Let M_z be the mapping

$$M_z = \left[-I \vdots H^* \right] \Sigma_\zeta^{-1} \quad ; \quad M_z : \mathbf{S} \rightarrow \mathbf{S}_1 \tag{4.6a}$$

where

H^* is the Hilbert adjoint of H in (3.40),

I is the identity on $\mathbf{S}_1$,

and

Σ_ζ^{-1} is the inverse of Σ_ζ in (4.2).

Then the stochastic process given by the second order mapping

$$Z = M_z \ \zeta \tag{4.6b}$$

is the complementary process for the observations Y in (4.4), i.e. Z in (4.6b) satisfies both the orthogonality and complementation conditions as prescribed in our restatement of the projection theorem.

<u>Proof:</u>

<u>Orthogonality:</u> This condition requires that the correlation between elements of Y and Z are zero, or equivalently that the kernel of the correlation operator

$$\Sigma_{yz} = M_y \ \Sigma_\zeta \ M_z^* \tag{4.7a}$$

is identically zero. Substituting from (4.4) and (4.6), Σ_{yz} can be written as

$$\begin{aligned} \Sigma_{yz} &= \left[H \vdots I \right] \Sigma_\zeta \ \Sigma_\zeta^{-1} \begin{bmatrix} -I \\ H \end{bmatrix} \\ &= 0 \quad , \end{aligned} \tag{4.7b}$$

verifying the orthogonality condition.

Complementation: In order to prove the complementation condition, we will need the following lemma [7] which is the multidimensional version of a basic result from the theory of Fredholm integral equations.

> Lemma: Let Ω_N be a multidimensional index set. Let $\underline{s}$ and $\underline{t}$ be index variables in Ω_N. Let $g(\underline{t},\underline{s})$ be a symmetric $\left(g(\underline{t},\underline{s}) = g(\underline{s},\underline{t}) \right)$ continuous kernel defining an integral operator
>
> $$G: L_2(\Omega_N) \to L_2(\Omega_N)$$
>
> with G having no negative eigenvalues. Then the operator $(I + G)$ has a unique inverse of the same form: $(I + G)^{-1} = (I + K)$, where the kernel of the integral operator K is also symmetric and continuous.

To establish the complementation condition it is sufficient [6] to show that the augmented map M given by

$$M = \begin{bmatrix} M_y \\ \\ M_z \end{bmatrix} \quad ; \qquad M: \mathbf{S} \to \mathbf{S} \tag{4.8a}$$

is invertible. Substituting for M_y and M_z we have the explicit representation for M

$$M = \left[\begin{array}{c:c} H & I \\ \hdashline -\Sigma_{\zeta_1}^{-1} & H^*\Sigma_{\zeta_2}^{-1} \end{array}\right] . \tag{4.8b}$$

Assuming the existence of the inverses in its partitions, the following operator can be shown by direct calculation to be the inverse of M

$$M^{-1} = \left[\begin{array}{c|c} \Sigma_{\zeta_1} H^*\left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} & -\left(\Sigma_{\zeta_1}^{-1} + H^*\Sigma_{\zeta_2}^{-1}H\right)^{-1} \\ \hline \Sigma_{\zeta_2}\left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} & H\left(\Sigma_{\zeta_1}^{-1} + H^*\Sigma_{\zeta_2}^{-1}H\right)^{-1} \end{array}\right] \tag{4.9}$$

i.e. $\quad M\,M^{-1} = M^{-1}\,M = I$.

The lemma is invoked to establish the invertiblity of the operators

$$\left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right) \quad \text{and} \quad \left(\Sigma_{\zeta_1}^{-1} + H^*\Sigma_{\zeta_2}^{-1} H\right)$$

by the following argument. Rewrite these two operators as

$$\Sigma_{\zeta_2}^{1/2}\left(I + \Sigma_{\zeta_2}^{-1/2} H\Sigma_{\zeta_1} H^*\Sigma_{\zeta_2}^{-1/2}\right)\Sigma_{\zeta_2}^{1/2} \tag{4.10a}$$

and

$$\Sigma_{\zeta_1}^{-1/2}\left(I + \Sigma_{\zeta_1}^{1/2} H^*\Sigma_{\zeta_2}^{-1} H\Sigma_{\zeta_1}^{1/2}\right)\Sigma_{\zeta_1}^{-1/2} \quad . \tag{4.10b}$$

If we choose symmetric continuous kernels for the square roots in (4.10a,b), then the kernels of the operators in the parentheses will be symmetric, positive and continuous, and the lemma ensures the aforementioned invertibility.

4.2 An Operator Representation for the Estimator

Recall from the Projection theorem that the estimator for x (2.20) is given by

$$\hat{x} = M_x \; M^{-1} \begin{bmatrix} Y \\ 0 \end{bmatrix} \quad . \tag{4.11}$$

By substituting for M^{-1} from (4.9), we obtain an explicit operator representation for the estimator:

$$\hat{x} = M_x \begin{bmatrix} \Sigma_{\zeta_1} H^* \\ - - \\ \Sigma_{\zeta_2} \end{bmatrix} \left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} Y \quad . \tag{4.12}$$

Similarly, it can be shown that the estimation error (2.21) can be expressed as a linear function of the underlying process ζ

$$\tilde{x} = M_x \begin{bmatrix} \Sigma_{\zeta_1} H^* \\ - - \\ \Sigma_{\zeta_2} \end{bmatrix} \left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} \left[\; -I \;\vdots\; H \;\right]\Sigma_{\zeta}^{-1}\zeta \quad . \tag{4.13}$$

A direct implementation of the estimator in (4.12) requires a realization of the indicated inverse. As an alternative, in Section 5 we obtain a realization for the estimator without explicitly performing this inversion. As we indicated in Section 2, Green's identity plays a critical role in this formulation. In particular, by invoking Green's identity we will be able to formulate a differential realization of the Hilbert adjoint H^*, yielding a differential realization of the complementary process Z. We will find that augmenting the differential operator representation for Y with that for Z results in a system which is easily inverted to give a differential operator representation of the estimator and estimation error.

4.3 M_z For Σ_{ζ_1} Singular

There are cases of interest for which there may be a singular correlation operator for either the input process or boundary condition. For example, in Section 6.3 we consider the example of a 1-D periodic process for which the boundary condition is, by the nature of the process, known without error. In this section we state a form for the operator M_Z which does not require the invertibility of Σ_{ζ_1}. An outline of the proof of the orthogonality and complementation conditions is given. We remark that although the form for M_Z presented earlier in (4.6a) could be derived directly from the one given below in this section, we have deliberately separated the two. As we will see later, the estimator for the case when Σ_{ζ_1} is singular is somewhat more complex than that for the case when it is nonsingular.

The action of the following operator on the underlying process (cf. (4.6b)) defines the complementary process and does not require the invertibility of Σ_{ζ_1}:

$$M_z = [\ -I \ \vdots \ \Sigma_{\zeta_1} H^* \Sigma_{\zeta_2}^{-1}] \qquad (4.14)$$

The orthogonality condition for the complementary process formed in this way is easily established by the same approach taken in (4.7a,b). The complementation condition is proved by showing that the inverse of

$$M = \begin{bmatrix} M_y \\ M_z \end{bmatrix} = \left[\begin{array}{c|c} H & I \\ \hline -I & \Sigma_{\zeta_1} H^* \Sigma_{\zeta_2}^{-1} \end{array}\right] \tag{4.15}$$

is given by

$$M^{-1} = \left[\begin{array}{c|c} \Sigma_{\zeta_1} H^*\left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} & -\left(I + \Sigma_{\zeta_1} H^* \Sigma_{\zeta_2}^{-1} H\right)^{-1} \\ \hline \Sigma_{\zeta_2}\left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} & H\left(I + \Sigma_{\zeta_1} H^* \Sigma_{\zeta_2}^{-1} H\right)^{-1} \end{array}\right] \tag{4.16}$$

The existence of the inverse of the operators in the left hand column of (4.16) is established in (4.10a). The existence of the inverse required in the right hand column is proved by invoking the operator version of the matrix inversion lemma [13] to write:

$$\left(I + \Sigma_{\zeta_1} H^* \Sigma_{\zeta_2}^{-1} H\right)^{-1} = I - \Sigma_{\zeta_1} H^*\left(\Sigma_{\zeta_2} + H\Sigma_{\zeta_1} H^*\right)^{-1} H \quad . \tag{4.17}$$

Note that the form of M_z in (4.6a) is obtained by operating on the left of (4.14) by the inverse of Σ_{ζ_1}.

SECTION 5

A DIFFERENTIAL OPERATOR REPRESENTATION FOR THE ESTIMATOR

5.0 Introduction

In this section we derive a differential operator representation for the estimator. The key to its derivation is the formulation of a differential operator representation for the complementary process whose I/O map is given in (4.6). It is in the formulation of this differential representation for the complementary process that the Green's Identity introduced in Section 2 plays an important role. With differential representations for both the process to be estimated and the corresponding complementary process, we will find that the augmentation and inversion steps (cf. Section 2.3) required in the formulation of the estimator become trivial.

5.1 The Hilbert Adjoint System

In the previous section we proved that the complementary process is given by

$$Z = \left[\, -I \;\vdots\; H^{*} \right] \Sigma_{\zeta}^{-1} \zeta \quad . \tag{5.1a}$$

Substituting from (4.1) and (4.3) into (5.1a), we can view the complementary process as an output signal plus noise:

$$Z = H^{*} \Sigma_{\zeta_2}^{-1} \zeta_2 - \Sigma_{\zeta_1}^{-1} \zeta_1 \quad . \tag{5.1b}$$

Our objective in this section is to formulate an internal realization for the input-output map H^*. The internal process in this realization is defined by a differential operator whose input process and boundary condition are the inputs to H^*.

To determine an internal differential realization for H^*, we temporarily leave the stochastic setting. That is, throughout the rest of this subsection all processes should be considered as elements of Hilbert spaces of deterministic functions rather than stochastic processes.

The internal realization for the input-output map H is

$$Lx = Bu \tag{5.2a}$$

$$Vx_b = v \tag{5.2b}$$

$$\Phi = \begin{bmatrix} \phi \\ \phi_b \end{bmatrix} = \begin{bmatrix} Cx \\ Wx_b \end{bmatrix} \quad \text{i.e.} \quad \Phi = H\begin{bmatrix} u \\ v \end{bmatrix} \quad . \tag{5.3}$$

Each of the maps L, B, V, C and W has been defined in Section 3.3. It will be convienient to define the spaces containing u and v as D_u and D_v respectively so that the domain of H can be written as $D(H) = D_u \times D_v$. Similarly, define the range spaces containing the output elements ϕ and ϕ_b as R_ϕ and R_{ϕ_b} so that the range of H is $R(H) = R_\phi \times R_{\phi_b}$.

The Hilbert adjoint of H is defined to be that operator which maps from the range of H into the domain of H

$$H^*: R(H) \to D(H) \tag{5.4a}$$

and for which the inner product identity

$$\langle H\xi, \eta \rangle_{R(H)} = \langle \xi, H^*\eta \rangle_{D(H)} \tag{5.4b}$$

is satisfied for arbitrary ξ and η in D(H) and R(H) respectively [8].

The first step in determining an internal realization for H^* is to rewrite (5.4b) in a more convenient form. Since the input u in (5.2a) enters only through the action of B, we can decompose H as

$$H = \tilde{H} \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix} \tag{5.5a}$$

If we denote the range of B by R_B, then $\tilde{H}$: $(R_B \times D_v) \to R(H)$. Given this decomposition of H, its adjoint H^* can be decomposed as

$$H^* = \begin{bmatrix} B^* & 0 \\ 0 & I \end{bmatrix} \tilde{H}^* \quad . \tag{5.5b}$$

If we denote the input process by

$$\xi = \begin{bmatrix} u \\ v \end{bmatrix} \tag{5.6}$$

and if we partition η, which is an element of $R(H) = R_\phi \times R_{\phi_b}$ and denote its partitions as

$$\begin{bmatrix} u_\lambda \\ v_\lambda \end{bmatrix} \equiv \eta \quad ; \quad \begin{matrix} u_\lambda \in R_\phi \\ v_\lambda \in R_{\phi_b} \end{matrix} \quad , \tag{5.7}$$

then substituting (5.5b) and (5.7) into the right hand side of (5.4b) gives

$$\begin{aligned} \langle H\xi,\eta\rangle &= \langle \xi, \begin{bmatrix} B^* & 0 \\ 0 & I \end{bmatrix} \tilde{H}^* \begin{bmatrix} u_\lambda \\ v_\lambda \end{bmatrix} \rangle \\ &= \langle \begin{bmatrix} Bu \\ v \end{bmatrix} , \tilde{H}^* \begin{bmatrix} u_\lambda \\ v_\lambda \end{bmatrix} \rangle \quad . \end{aligned} \tag{5.8}$$

Similarly, defining the partitions of $\tilde{H}^*\eta$ which is an element of $D(\tilde{H})$, as

$$\begin{bmatrix} \lambda \\ \psi_b \end{bmatrix} \equiv \tilde{H}^* \begin{bmatrix} u_\lambda \\ v_\lambda \end{bmatrix} \quad ; \quad \begin{matrix} \lambda \in R_B \\ \psi_b \in D_v \end{matrix} \tag{5.9}$$

and substituting for Bu and v from (5.2a) and (5.2b), (5.8) becomes

$$\begin{aligned} \langle H\xi,\eta\rangle &= \langle Lx,\lambda\rangle + \langle Vx_b,\psi_b\rangle \\ &= \langle Lx,\lambda\rangle + \langle x_b, V^*\psi_b\rangle \quad . \end{aligned} \tag{5.10}$$

Finally, by noting from (5.3) that

$$H\xi = \begin{bmatrix} \phi \\ \phi_b \end{bmatrix} = \begin{bmatrix} Cx \\ Wx_b \end{bmatrix} \quad ,$$

we can rewrite the left hand side of (5.10) so that the inner product identity in (5.4b) can be expressed as

$$\langle x, C^* u_\lambda \rangle + \langle x_b, W^* v_\lambda \rangle = \langle Lx, \lambda \rangle + \langle x_b, V^* \psi_b \rangle \quad . \tag{5.11}$$

Up to this point we have simply combined some new notation along with that for the internal representation for H to re-express the inner product identity (5.4b). The next step is more substantial and is a key one in the development of the internal realization for H^*. In particular, we employ Green's identity from (3.18) to replace $\langle Lx,\lambda\rangle$ in (5.11). Then (5.4b) can be written in terms of the formal adjoint differential (difference) operator $L^\dagger$:

$$\langle x, \left[C^* u_\lambda - L^\dagger \lambda\right] \rangle = \langle x, \left[E\lambda_b + V^* \psi_b - W^* v_\lambda\right] \rangle \quad . \tag{5.12}$$

Although the Hilbert adjoint H^* is a unique map, there exists a family of equivalent internal differential realizations. Using the notation introduced above, we will verify one internal realization for H^* with input η and output $\Psi = \{\psi, \psi_b\}$ by showing that it satisfies (5.12).

Let W_c be one of the family of operators which complements V and W (see equation (3.35c)), in that

$$\Gamma \equiv \begin{bmatrix} -V \\ W_c \\ W \end{bmatrix} \tag{5.13}$$

is invertible (as will become clear shortly, we have included a minus sign in defining Γ for convenience). Employing the inverse of Γ and the operator E in the boundary term of Green's identity (3.13), define the partitioned operator

$$\begin{bmatrix} W_\lambda \\ V_{\lambda c} \\ V_\lambda \end{bmatrix} \equiv \left(\Gamma^*\right)^{-1} E \quad . \tag{5.14}$$

This leads to an expression for E that will be useful later:

$$E = \left[-V^* \vdots\ W_c^* \vdots\ W^* \right] \begin{bmatrix} W_\lambda \\ V_{\lambda c} \\ V_\lambda \end{bmatrix} = -V^* W_\lambda + W_c^* V_{\lambda c} + W^* V_\lambda \quad . \tag{5.15}$$

The following theorem establishes an internal differential realization for H^*.

<u>Theorem:</u> (Hilbert Adjoint System) An internal differential realization for the input-output map

$$\Psi = \begin{bmatrix} \psi \\ \psi_b \end{bmatrix} = H^* \begin{bmatrix} u_\lambda \\ v_\lambda \end{bmatrix} \tag{5.16a}$$

is given by an internal process λ satisfying

$$L^\dagger \lambda = C^* u_\lambda \tag{5.16b}$$

with boundary condition

$$\begin{bmatrix} V_\lambda \\ V_{\lambda c} \end{bmatrix} \lambda_b = \begin{bmatrix} v_\lambda \\ 0 \end{bmatrix} \tag{5.16c}$$

and output map

$$\Psi = \begin{bmatrix} \psi \\ \psi_b \end{bmatrix} = \begin{bmatrix} B^* \lambda \\ W_\lambda \lambda_b \end{bmatrix} \quad . \tag{5.16d}$$

<u>Proof:</u> With the dynamics of λ given by (5.16b), the left hand side of (5.12) is zero. To show that the right hand side is also zero, we employ (5.16d) and the first row of (5.16c) to rewrite the right hand side of (5.12) as

$$<x_b, \left[E\lambda_b + V^* \psi_b - W^* v_\lambda \right]> = <x_b, \left[E + V^* W_\lambda - W^* V_\lambda \right] \lambda_b> \quad . \tag{5.17a}$$

Substituting for E from (5.15) gives

$$\langle x_b, [E\lambda_b + V^*\psi_b - W^*v_\lambda]\rangle = \langle x_b, W_c^* V_{\lambda c}\lambda_b\rangle \quad , \tag{5.17b}$$

A further substitution from the second row of (5.16c) completes the proof:

$$\begin{aligned}\langle x_b, [E\lambda_b + V^*\psi_b - W^*v_\lambda]\rangle &= \langle x_b, 0\rangle \\ &= 0 \quad .\end{aligned} \tag{5.17c}$$

Thus (5.12) is satisfied with both the left and right hand sides identically zero. Although this differential realization is not unique due to the degrees of freedom in choosing W_c, we will show that the estimator itself is invariant with respect to the choice of W_c, as it must be.

5.2 Augmentation and Inversion

The internal differential realization for H^* in (5.16) defines a representation for the complementary stochastic process in (5.1a) as:

$$\begin{bmatrix} z \\ z_b \end{bmatrix} = [\ -I \ \vdots \ H^*]\Sigma_\zeta^{-1}\zeta \tag{5.18}$$

In this subsection we augment the internal realization for (5.18) with that for the observations to get an internal differential realization for the combined system:

$$\begin{bmatrix} Y \\ Z \end{bmatrix} = M\zeta \quad ; \quad M = \begin{bmatrix} M_Y \\ M_Z \end{bmatrix} \quad . \tag{5.19}$$

We then invert this realization to obtain an internal differential realization for the estimator.

The differential form for the augmented system in (5.19) is

$$\begin{bmatrix} L & 0 \\ 0 & L^\dagger \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} B & 0 \\ 0 & C^*R^{-1} \end{bmatrix}\begin{bmatrix} u \\ r \end{bmatrix} \tag{5.20a}$$

with boundary condition

$$\begin{bmatrix} v \\ \hline 0 \\ \Pi_b^{-1} r_b \end{bmatrix} = \begin{bmatrix} V & \vdots & 0 \\ \hline 0 & \vdots & V_{\lambda c} \\ 0 & \vdots & V_\lambda \end{bmatrix} \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} \tag{5.20b}$$

and outputs

$$\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & B^* \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} + \begin{bmatrix} 0 & I \\ -Q^{-1} & 0 \end{bmatrix} \begin{bmatrix} u \\ r \end{bmatrix} \qquad \text{on } \Omega_N \tag{5.20c}$$

$$\begin{bmatrix} y_b \\ z_b \end{bmatrix} = \begin{bmatrix} W & 0 \\ 0 & W_\lambda \end{bmatrix} \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & -\Pi_v^{-1} \end{bmatrix} \begin{bmatrix} v \\ r_b \end{bmatrix} \qquad \text{on } \partial\Omega_N. \tag{5.20d}$$

As indicated by (5.19), the inverse system we seek is one with $\{Y,Z\} = \{y,y_b,z,z_b\}$ as input and $\zeta = \{u,v,r,r_b\}$ as output. To this end, following the approach taken by Levy et al for the 1-D causal case in [9], we first solve for the elements of ζ by inverting the output equations (5.20c) and (5.20d):

$$\begin{bmatrix} u \\ r \end{bmatrix} = \begin{bmatrix} -Q & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} + \begin{bmatrix} 0 & QB^* \\ -C & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} \tag{5.21a}$$

$$\begin{bmatrix} v \\ r_b \end{bmatrix} = \begin{bmatrix} -\Pi_v & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} z_b \\ y_b \end{bmatrix} + \begin{bmatrix} 0 & \Pi_v W_\lambda \\ -W & 0 \end{bmatrix} \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} \quad . \tag{5.21b}$$

Substituting these expressions into the dynamics and boundary conditions in (5.20a) and (5.20b) yields an internal differential realization of the inverse system with dynamics:

$$\begin{bmatrix} L & \vdots & -BQB^* \\ \hline C^*R^{-1}C & \vdots & L^\dagger \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -BQ & \vdots & 0 \\ \hline 0 & \vdots & C^*R^{-1} \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} \tag{5.22}$$

and with boundary condition:

$$\begin{bmatrix} z_b \\ 0 \\ \Pi_b^{-1} y_b \end{bmatrix} = \left[\begin{array}{c:c} -\Pi_v^{-1} V & W_\lambda \\ \hdashline 0 & V_{\lambda c} \\ \hdashline \Pi_b^{-1} W & V_\lambda \end{array}\right] \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} \quad . \tag{5.23}$$

This boundary condition can be simplified so that its dependence on W_c, V_λ, and $V_{\lambda c}$ is eliminated. Recalling the relation between these operators and E in Green's identity from (5.15), it can be shown that operating on the left of (5.23) by $\left[-V^*\vdots W_c{}^*\vdots W^*\right]$ gives the boundary condition as

$$\left[W^* \Pi_b^{-1} y_b - V^* z_b \right] = \left[W^* \Pi_b^{-1} W + V^* \Pi_v^{-1} V \;\vdots\; E \right] \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} \quad . \tag{5.24}$$

The estimator is the solution of (5.22) and (5.24) projected onto Sp(Y), i.e. the solution with $Z = \{z, z_b\} = 0$:

$$\left[\begin{array}{c:c} L & -BQB^* \\ \hdashline C^* R^{-1} C & L^\dagger \end{array}\right] \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} 0 \\ \hdashline C^* R^{-1} y \end{bmatrix} \tag{5.25a}$$

$$W^* \Pi_b^{-1} y_b = \left[W^* \Pi_b^{-1} W + V^* \Pi_v^{-1} V \;\vdots\; E \right] \begin{bmatrix} \hat{x}_b \\ \hat{\lambda}_b \end{bmatrix} \tag{5.25b}$$

The estimates of the elements of the underlying process ζ, if desired, can be computed form the output equations (5.21a) and (5.21b) evaluated at the solution of (5.25) and with z and z_b equal to zero. Note that since L and $L^\dagger$ are of the same order, the order of the estimator is twice that of L. Also note the remarkable fact that in addition to the original problem statement, we only need to know E and $L^\dagger$ from Green's identity in (3.13) to completely define the differential realization for the estimator.

5.3 The Estimation Error

The estimation error

$$\tilde{x} = x - \hat{x} \tag{5.26}$$

is obtained as the solution of (5.22) and (5.24) projected onto Sp(Z) rather than Sp(Y). Here we formulate a differential realization of the estimation error which is driven by ζ whose probability law is known. The second order statistics of the estimation error can be computed from those of ζ using this relation.

Recall from the restatement of the Projection Theorem in Section 2.3 that

$$x = M_x M^{-1} \begin{bmatrix} 0 \\ Z \end{bmatrix} = M_x M^{-1} \begin{bmatrix} 0 \\ M_z \end{bmatrix} \zeta \quad . \tag{5.27}$$

That is, Z has been replaced by its representation given in terms of ζ. Consider the boundary condition (5.24) projected onto Sp(Z), i.e. (5.24) evaluated with y_b equal to zero:

$$- V^* z_b = [\, W^* \Pi_b^{-1} W + V^* \Pi_v^{-1} V \;\vdots\; E \,] \begin{bmatrix} \tilde{x}_b \\ \tilde{\lambda}_b \end{bmatrix} \tag{5.28}$$

Substituting for z_b from (5.23b)

$$z_b = W_\lambda \lambda_b - \Pi_v^{-1} v \tag{5.29a}$$

and using the basic definition:

$$\tilde{\lambda}_b = \lambda_b - \hat{\lambda}_b \quad , \tag{5.29b}$$

(5.28) becomes

$$V^* \Pi_v^{-1} v = [\, W^* \Pi_b^{-1} W + V^* \Pi_v^{-1} V \;\vdots\; E + V^* W_\lambda \;:\; V^* W_\lambda \,] \begin{bmatrix} \tilde{x}_b \\ \tilde{\lambda}_b \\ \hat{\lambda}_b \end{bmatrix} \quad . \tag{5.30}$$

To eliminate the dependence on W_λ, V_λ and $V_{\lambda c}$ as we had done for the estimator, recall from (5.15) and (5.16b) that

$$E + V^* W_\lambda = W^* V_\lambda + W_c^* V_{\lambda c} \quad , \tag{5.31a}$$

$$V_\lambda \lambda_b = V_\lambda (\tilde{\lambda}_b + \hat{\lambda}_b) = \Pi_b^{-1} r_b \tag{5.31b}$$

and

$$W_c^* V_{\lambda c} \lambda_b = W_c^* V_{\lambda c} (\tilde{\lambda}_b + \hat{\lambda}_b) = 0 \quad . \tag{5.31c}$$

From these three equations we can write

$$[\, E + V^* W_\lambda] \tilde{\lambda}_b = [\, W^* V_\lambda + W_c^* V_\lambda] \; [\lambda_b - \hat{\lambda}_b]$$

$$= W^* \Pi_b^{-1} r_b - [W^* V_\lambda + W_c^* V_{\lambda c}] \hat{\lambda}_b \quad , \tag{5.32}$$

and substituting into (5.28), the boundary condition becomes

$$[\, V^* \Pi_v^{-1} v - W^* \Pi_b^{-1} r_b] = [\, W^* \Pi_b^{-1} W + V^* \Pi_v^{-1} V \; \vdots \; E \,] \begin{bmatrix} \tilde{x}_b \\ -\hat{\lambda}_b \end{bmatrix} \quad . \tag{5.33}$$

We have chosen $-\hat{\lambda}_b$ instead of $\hat{\lambda}_b$ to highlight the similarity between the structure of the boundary condition for the estimation error in (5.33) and that of the estimator in (5.25).

The projection of (5.22) onto Sp(Z) gives the error dynamics as

$$\begin{bmatrix} L & \vdots & -BQB^* \\ - & & - \\ C^* R^{-1} C & \vdots & L^* \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} -BQz \\ - \\ 0 \end{bmatrix} \quad . \tag{5.34}$$

Replacing z from (5.20c)

$$z = B^* \lambda - Q^{-1} u \quad , \tag{5.35a}$$

employing

$$\tilde{\lambda} = \lambda - \hat{\lambda} \tag{5.35b}$$

and recalling from (5.20a) that the dynamics of λ are given by

$$L^{\dagger}\lambda = C^{*}R^{-1}r \qquad , \tag{5.35c}$$

(5.34) can be rewritten as

$$\begin{bmatrix} L & \vdots & -BQB^{*} \\ \cdots & \vdots & \cdots \\ C^{*}R^{-1}C & \vdots & L^{\dagger} \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \hat{} \\ -\lambda \end{bmatrix} = \begin{bmatrix} Bu \\ \cdots \\ -C^{*}R^{-1}r \end{bmatrix} \quad . \tag{5.36}$$

Thus (5.33) and (5.36) completely define the estimation error in terms of $\zeta = \{u,v,r,r_b\}$ whose probability law is known. In addition, the dynamics and boundary conditions of the estimation error have been shown to be similar to those of the estimator. One should be able to take advantage of these similarities when computing the estimate and its error covariance. For example, see the discussion of the implementation of the estimator and the computation of the error covariance for the 1-D noncausal process in Part II of this paper [5].

5.4 Special Case: Π_V Singular

In Section 4.3 we presented a model for the complementary process which did not require the invertiblity of the covariance parameters Q and Π_V (i.e. the invertibility of Σ_{ζ_1}). In this section we define the estimator for the case when Π_V is singular.

By augmenting with the complementary process defined through (4.14) and inverting, it can be shown that we arrive at the same dynamics for the estimator as obtained previously in (5.22). However, the boundary condition for the inverted system is slightly different than that in (5.23). In particular, the boundary condition in this case is:

$$\begin{bmatrix} \Pi_v z_b \\ \cdots \\ 0 \\ \cdots \\ \Pi_b^{-1} y_b \end{bmatrix} = \begin{bmatrix} -V & \vdots & \Pi_v W_\lambda \\ \cdots & \vdots & \cdots \\ 0 & \vdots & V_{\lambda c} \\ \cdots & \vdots & \cdots \\ \Pi_b^{-1} W & \vdots & V_\lambda \end{bmatrix} \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} \quad . \tag{5.37}$$

As we had done for (5.23), we will rewrite this boundary condition in terms of the operator E found in Greeen's Identity. Let Ψ be the partitioned operator:

$$\Psi = \begin{bmatrix} \Pi_v & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \tag{5.38a}$$

with the partitions compatible with those of Γ in (5.13). Let Θ be the operator

$$\Theta = \Psi \, \Gamma^{*-1} \quad . \tag{5.38b}$$

Then recalling (5.15), we can write the boundary condition (5.37) as

$$\begin{bmatrix} \Pi_v z_b \\ \hline 0 \\ \hline \Pi_b^{-1} y_b \end{bmatrix} = \left[\begin{array}{c|c} -V & \\ \hline 0 & \\ \hline \Pi_b^{-1} W & \end{array}\right] \Theta E \begin{bmatrix} x_b \\ \lambda_b \end{bmatrix} \tag{5.39}$$

Projecting (5.39) onto Sp(Y) gives the boundary condition for the estimator as

$$\begin{bmatrix} 0 \\ \hline 0 \\ \hline \Pi_b^{-1} y_b \end{bmatrix} = \left[\begin{array}{c|c} -V & \\ \hline 0 & \\ \hline \Pi_b^{-1} W & \end{array}\right] \Theta E \begin{bmatrix} \hat{x}_b \\ \hat{\lambda}_b \end{bmatrix} \tag{5.40}$$

Following arguments similar to those in Section 5.3 and applying (5.38), it can be shown that the estimation error dynamics for this case are the same as those in (5.36), while the boundary condition is

$$\begin{bmatrix} -v \\ \hline 0 \\ \hline -\Pi_b^{-1} r_b \end{bmatrix} = \left[\begin{array}{c|c} -V & \\ \hline 0 & \\ \hline \Pi_b^{-1} W & \end{array}\right] \Theta E \begin{bmatrix} \tilde{x}_b \\ -\hat{\lambda}_b \end{bmatrix} \quad . \tag{5.41}$$

The boundary conditions in (5.40) and (5.41) may appear to be in as simple a form as their counterparts for the case when Π_V is invertible. However, the requirement to invert Γ^* in order to compute Θ in (5.38b) makes the boundary conditions derived in this subsection considerably more complex. Furthermore, in this formulation of the boundary condition, the operator W_C, one of the partitions of Γ, has not been eliminated. Since the estimator and error equations must be independent of the choice of W_C it should be possible to reduce our formulation to one which is independent of this choice. This remains an open question at this time. Nevertheless, we will see in an example of a periodic process in Section 6.3 that the form of the boundary condition presented here is useful in deriving the estimator for a particular case when Π_V is singular.

SECTION 6

THE ESTIMATOR FOR THE TWO EXAMPLES

6.0 Introduction

By considering the two examples introduced earlier, we demonstrate the ease with which one can apply (5.25) to obtain an internal differential representation for the estimator of a noncausal stochastic process. We show that the estimator for the process governed by Poisson's equation takes the form of a fourth order biharmonic equation. In the case of the 1-D discrete boundary value process, it will be shown that a special case of the solution we obtain from (5.25) is a well-known form of the solution for the fixed-interval smoother for 1-D discrete causal processes [10]. In addition, in Section 6.3 we apply the solution for the estimator when Σ_{ζ_1} is singular to a discrete 1-D periodic process.

6.1 2-D Continuous Case: Poisson's Equation (Example 1)

The problem statement has been given in Section 3.3 by equations (3.45) and (3.46). As we have done previously for this example, s denotes an index variable representing elemtents of the unit disk, and θ represents elements of $[0,2\pi]$ which we have identified with $\partial\Omega_2$. Substituting for $L^{\dagger}$ in the estimator solution (5.25) from (3.10a), we obtain the estimator dynamics as

$$\left[\begin{array}{c:c} \nabla^2 & -B^2(s)Q(s) \\ \hdashline C^2(s)R^{-1}(s) & \nabla^2 \end{array}\right] \begin{bmatrix} \hat{x}(s) \\ \hat{\lambda}(s) \end{bmatrix} = \begin{bmatrix} 0 \\ \hdashline C(s)R^{-1}(s) \end{bmatrix} y(s) \quad . \tag{6.1}$$

Note from (3.45b) and (3.46b) that the boundary condition and boundary obsevation can be expressed by functions on $[0,2\pi]$ as

$$(Vx_b)(\theta) = [\ V(\theta) \ \vdots \ 0\ ]x_b(\theta) \tag{6.2a}$$

and

$$(Wx_b)(\theta) = [\ 0 \ \vdots \ W(\theta)\ ]x_b(\theta) \tag{6.2b}$$

where we recall that $x_b{}'(\theta) = [x(1,\theta),x_n(1,\theta)]$. Using this expression and substituting for E from (3.14), it can be shown that the boundary condition for (6.1) is (in polar coordinates evaluated at $\rho = 1$)

$$\begin{bmatrix} 0 \\ - & - & - \\ W(\theta)\Pi_b^{-1}(\theta)y_b(\theta) \end{bmatrix} = \begin{bmatrix} \nabla^2(\theta)\Pi_v^{-1}(\theta) & \vdots & 0 \\ - & \vdots & - \\ 0 & \vdots & 1 \end{bmatrix} \begin{bmatrix} \hat{x}(\rho,\theta) \\ \hat{\lambda}(\rho,\theta) \end{bmatrix} + \begin{bmatrix} 0 & \vdots & -1 \\ - & \vdots & - \\ W^2(\theta)\Pi_b^{-1}(\theta) & \vdots & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_n(\rho,\theta) \\ \hat{\lambda}_n(\rho,\theta) \end{bmatrix} . \quad (6.3)$$

When $B^2(s)Q(s) > 0$ for all s, we can solve for $\hat{\lambda}$ in (6.1) as

$$\hat{\lambda}(s) = \left[B^2(s)Q(s)\right]^{-1}\nabla^2 \hat{x}(s) \quad . \qquad (6.4)$$

Substituting (6.4) back into (6.1), we find that the estiator dynamics are given by the biharmonic equation:

$$\left\{\nabla^2\left[B^2(s)Q(s)\right]^{-1} + C^2(s)R^{-1}(s)\right\}\nabla^2 \hat{x}(s) = C(s)R^{-1}(s)y(s) \quad . \qquad (6.5)$$

With $\partial/\partial n$ denoting the normal derivative and substituting from (6.4), the boundary condition in (6.3) can be rewritten as

$$0 = \Pi_v^{-1}(\theta)\hat{x}(\rho,\theta) - (\partial/\partial n)\left\{\left[B^2(\rho,\theta)Q(\rho,\theta)\right]^{-1}\nabla^2 \hat{x}(\rho,\theta)\right\} \qquad (6.6a)$$

and

$$W(\theta)y_b(\theta) = W(\theta)^2\hat{x}_n(\rho,\theta) + \Pi_b(\theta)\left[B^2(\rho,\theta)Q(\rho,\theta)\right]^{-1}\nabla^2 \hat{x}(\rho,\theta) \qquad (6.6b)$$

evaluated at $\rho = 1$. We have not investigated analytical or numerical solutions for this biharmonic equation. In practice, one could employ one of the many available numerical techniques such as finite difference approximations [14].

6.2 1-D Discrete Case: Two-Point Boundary Value Process (Example 2)

Again we simply substitute from the problem statement in equations (3.47) and (3.48) and from (3.34) and (3.35b) for $L^\dagger$ and E into (5.25) to obtain the estimator dynamics

$$\begin{bmatrix} \hat{x}_{k+1} \\ \hat{\lambda}_k \end{bmatrix} = \begin{bmatrix} A_k & \vdots & -B_kQ_kB_k' \\ - & \vdots & - \\ -C_k'R_k^{-1}C_k & \vdots & A_k' \end{bmatrix} \begin{bmatrix} \hat{x}_k \\ \hat{\lambda}_{k+1} \end{bmatrix} + \begin{bmatrix} 0 \\ - \\ C_k'R_k^{-1} \end{bmatrix} y_k \qquad (6.7a)$$

and boundary condition

$$\begin{bmatrix} W^{0'} \\ W^{K'} \end{bmatrix} \Pi_b^{-1} y_b = \left[\begin{array}{c:c} V^{0'}\Pi_v^{-1}V^0 + W^{0'}\Pi_b^{-1}W^0 & -I \\ \hdashline V^{K'}\Pi_v^{-1}V^0 + W^{K'}\Pi_b^{-1}W^0 & 0 \end{array}\right] \begin{bmatrix} \hat{x}_0 \\ \hat{\lambda}_0 \end{bmatrix} + \left[\begin{array}{c:c} V^{0'}\Pi_v^{-1}V^K + W^{0'}\Pi_b^{-1}W^K & 0 \\ \hdashline V^{K'}\Pi_v^{-1}V^K + W^{K'}\Pi_b^{-1}W^K & I \end{array}\right] \begin{bmatrix} \hat{x}_K \\ \hat{\lambda}_K \end{bmatrix} . \tag{6.7b}$$

If we consider the special case of <u>no</u> boundary observation y_b (i.e. $W^0 = W^K = 0$) and an <u>initial</u> condition for x (i.e. $V^0 = I$, $V^K = 0$), then the boundary condition in (6.7b) becomes

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} I & -\Pi_v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \hat{\lambda}_0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}_K \\ \hat{\lambda}_K \end{bmatrix} . \tag{6.8}$$

This boundary condition along with the dynamics in (6.7a) is recognized as the well-known solution for the fixed-interval smoother for causal discrete 1-D stochastic processes [10].

6.3 1-D Discrete Case: A Periodic Process

To illustrate the case when Π_v is singular, we consider a 1-D discrete periodic process. The dynamics of this process are the same as those introduced earlier except for the following periodicity constraints:

$$A_{k+K} = A_k \quad , \tag{6.9a}$$

$$B_{k+K} = B_k \tag{6.9b}$$

and

$$u_{k+K} = u_k \quad . \tag{6.9c}$$

Along with (6.9), the following boundary condition guarantees that x repeats itself with a period of K:

$$0 = x_0 - x_K \tag{6.10a}$$

i.e.

$$V = [\, I \;\vdots\; -I \,] \tag{6.10b}$$

and

$$\Pi_v = 0 \quad . \tag{6.10c}$$

We also assume for the moment that there is no boundary observation.

As discussed earlier in Section 5.4, the estimator dynamics for the case when Π_V is singular are the same as those in (6.7a) above. From (5.40) it can be shown that in this case the boundary conditions for the estimator can always be put in the following form (i.e. for any admissable choice of W_C)

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \hat{\lambda}_0 \end{bmatrix} + \begin{bmatrix} -I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \hat{x}_K \\ \hat{\lambda}_K \end{bmatrix} \quad . \tag{6.11}$$

If there is some additional a priori information about the zero mean random variable x_0, then it can be included as a boundary measurement as follows. Let Π_0 be the a priori variance of x_0. Then the boundary measurement

$$y_b = 0 = [\ I \ \vdots \ 0\ ]x_b - r_b \quad , \tag{6.13a}$$

$$\text{i.e.} \quad W = [\ I \ \vdots \ 0\ ] \ , \tag{6.13b}$$

where the variance of r_b is Π_0, will account for this information in the estimator. When y_b in (6.13) is included, the boundary condition for the estimator of the periodic process becomes (from (5.40)):

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ -\Pi_0^{-1} & I \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \hat{\lambda}_0 \end{bmatrix} + \begin{bmatrix} -I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \hat{x}_K \\ \hat{\lambda}_K \end{bmatrix} \quad . \tag{6.14}$$

As shown in [6], solutions for discrete two-point boundary value problems representing the estimators in these last two sections can be implemented via a two-filter form similar to two-filter forms of the smoother for causal processes.

SECTION 7

CONCLUSIONS

Through an extension of the method of complementary models [1], we have developed a procedure for writing the estimator for both discrete and continuous parameter linear boundary value stochastic processes in a differential operator form. The two major steps in the development of the estimator have been (1) the formulation of an input-output operator representation for the complementary process in Section 4 and (2) the use of Green's identity in Section 5 in the derivation of an internal differential realization for this input-output map. We emphasize that at no point in our derivations have we required a Markov representation for the process to be estimated. The variety of problems for which our estimator solution is applicable has been illustrated through two examples: a 1-D discrete parameter process and a 2-D continuous parameter process.

The major advantage in having a differential realization for the estimator is that this form of representation provides an excellent starting point for the development of methods for implementing the estimator. This is in contrast to estimators derived by a direct application of the projection theorem, which usually leads to integral equations (e.g. Wiener-Hopf) requiring factorization in order to obtain an implementation. Furthermore, we have also derived an internal differential realization for the estimation errors in a form which is nearly identical to that for the estimator. In Part II of this paper [5] we apply the estimator solution formulated in this paper to a continuous 1-D two-point boundary value stochastic process and develop a stable, recursive implementation for the resulting differentential form of the estimator. In addition, by following the same procedures as used to obtain the recursive estimator implementation, we develop recursions for the computation of the smoothing error covariance.

There remain a variety of interesting open issues pertaining to our estimator solution. Thus far, very little has been attempted in the area of developing efficient methods for implementing the estimator in multi-D cases. In this regard, one might hypothesize that, as in the 1-D case (see [5]), two-filter forms may also exist for the implementation of our estimator in 2-D cases. For instance, one might consider implementing the solution derived for Example 1 of this paper as the linear combination of one process obtained by filtering the observations radially outward from the center of the unit disk and another obtained by filtering radially inward from the boundary of the disk. In addition to questions of implementation, there are also unanswered questions which relate to the boundary conditions for multi-D problems. For example, recall from (5.25b) that the boundary condition for our estimator is defined in terms of the operator adjoints V^* and W^* and the inverses of the correlation operators Π_v and Π_b. In our 2-D example we have tacitly avoided any complications which might arise in determining these adjoints and inverses by choosing v and r_b as white noise and by choosing V and W as a simple scaling of the process on the boundary (see (6.2a,b)). It would be of interest to investigate the estimator's boundary condition for this 2-D example when the boundary value v is, for instance, a 1-D periodic stochastic process on the unit circle. Another open issue concerning to the estimator boundary condition has already been raised in Section 5.4. That is, it should be possible to further simplify the expression in (5.41) for the estimator's boundary condition in the case when Π_v is singular.

In summary, we feel that this paper presents an extremely useful and broadly applicable method for deriving optimal estimators for noncausal processes in several dimensions. Given this valuable tool, one is then in a position to focus one's attention on the problem of implementing the optimal estimator in an efficient fashion. As mentioned previously, this is precisely what is done in [5] for the case of 1-D continuous parameter processes. In the more general, multi-dimensional case many open questions remain, but the results in this paper bring us significantly closer to answering them.

REFERENCES

[1] Weinert, H.L. and Desai, U.B., "On Complementary Models and Fixed-Interval Smoothing," IEEE Trans. Aut. Cont., V. AC-26, Aug. 1981.

[2] Adams, M.B., "Estimation of Noncausal Processes in One and Two Dimensions," Doctoral Thesis Proposal, C.S. Draper Lab. Rept. CSDL-T-770, Feb. 1982.

[3] Greenberg, M.D., Applications of Green's Functions in Science and Engineering, Prentice-Hall, Englewood Cliffs, NJ, 1976.

[4] Coddington, E.A. and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill Book Co., NY, 1955.

[5] Adams, M.B., Willsky, A.S. and Levy, B.C., "Linear Estimation of Boundary Value Stochastic Processes, Part II: 1-D Smoothing Problems," submitted to IEEE Trans. Aut. Cont.

[6] Adams, M.B., "Linear Estimation of Boundary Value Stochastic Processes," Ph.D. Dissertation, Dept. of Aero and Astro, MIT, Cambridge, MA, in preparation.

[7] Smithies, F., Integral Equations, Cambridge University Press, Cambridge, England, 1958.

[8] Kreyszig, E., Introductory Functional Analysis with Applications, John Wiley and Sons, NY, 1978.

[9] Levy, B.C., et al., "A Scattering Framework for Decentralized Estimation Problems," LIDS-P-1075, MIT Dept. of Elec. Eng. and Comp. Sci., Cambridge, MA, Mar. 1981, to appear in Automatica.

[10] Kailath, T.K., "A View of Three Decades of Linear Filtering Theory," IEEE Trans. Inf. Th., IT-20, No. 2, pp. 146-181, Mar. 1974.

[11] Courant, R. and Hilbert, D., Methods of Mathematical Physics, Vol. 2, Wiley Interscience, NY, 1962.

[12] Ito, K., "Multiple Wiener Integrals," J. Math. Soc.(Japan), Vol. 3, pp. 157-169, 1951.

[13] Grenander, U. and Rosenblatt, M., Statistical Analysis of Stationary Time Series, John Wiley and Sons, NY, 1957.

[14] Lapidus, L. and Pinder, P.F., Numerical Solution of Partial Differential Equations in Science and Engineering, John Wiley and Sons, NY, 1982.

APPENDIX

DERIVATION OF M_z FOR THE STATIC EXAMPLE

In Section 2.1 it was stated that the complementary process associated with the observation vector y in (2.2a) is the n-dimensional random vector

$$z = M_z \zeta \tag{A.1}$$

where, with T any invertible n×n matrix, M_z is given by

$$M_z = T[\ I \ \vdots \ -\Sigma_x H' \Sigma_r^{-1}\ ] \quad . \tag{A.2}$$

In this appendix we derive this general form for the matrix M_z.

Let M_z be denoted by

$$M_z = [Z_x \ \vdots \ Z_r] \tag{A.3}$$

where the partitions are compatible with the partitions of ζ in (2.1a). With M_z given by (A.3), M in (2.6) is

$$M = \begin{bmatrix} M_y \\ M_z \end{bmatrix} = \left[\begin{array}{c:c} H & I \\ \hdashline Z_x & Z_r \end{array}\right] \quad . \tag{A.4}$$

Now use the orthogonality and complementation conditions stated in Section 2.1 to find expressions for Z_x and Z_r.

<u>Complementation</u>

This condition implies that the (n+p)×(n+p) matrix M is invertible or that for k an (n+p)×1 vector

$$Mk = 0 \implies k = 0 \quad . \tag{A.5}$$

If we denote partitions of k by

$$k = \begin{bmatrix} k_x \\ k_r \end{bmatrix} \begin{array}{l} \leftarrow n\times 1 \\ \leftarrow p\times 1 \end{array} \quad , \tag{A.6}$$

then Mk = 0 implies that

$$Hk_x = -k_r \tag{A.7a}$$

and

$$Z_x k_x = -Z_r k_r \quad . \tag{A.7b}$$

Substituting (A.7a) into (A.7b) we get the following condition which is equivalent to (A.5)

$$(Z_x - Z_r H)k_x = 0 \implies k_x = 0 \quad . \tag{A.7c}$$

Orthogonality

This condition requires that the elements of y and z are uncorrelated

$$E[yz'] = 0$$

$$= [H \ \vdots \ I] \begin{bmatrix} \Sigma_x & 0 \\ 0 & \Sigma_r \end{bmatrix} \begin{bmatrix} Z'_x \\ Z'_r \end{bmatrix} \tag{A.8}$$

Thus

$$H\Sigma_x Z'_x = -\Sigma_r Z'_r \tag{A.9a}$$

or

$$Z_r = -Z_x \Sigma_x H' \Sigma_r^{-1} \quad . \tag{A.9b}$$

Combining (A.7c) and (A.9b) we have

$$Z_x (I + \Sigma_x H' \Sigma_r^{-1} H) k_x = 0 \implies k_x = 0 \quad . \tag{A.9c}$$

Since

$$(I + \Sigma_x H' \Sigma_r^{-1} H)$$

is invertible, the statement (A.9c) is true if and only if Z_x is invertible. Therefore, let Z_x equal some invertible matrix T:

$$Z_x = T \quad . \tag{A.10a}$$

Then from (A.9b)

$$Z_r = -T\Sigma_x H'\Sigma_r^{-1} \quad , \tag{A.10b}$$

and from (A.3) we obtain the expression we seek

$$M_z = T\left[I \;\vdots\; -\Sigma_x H'\Sigma_r^{-1}\right] \quad . \tag{A.10c}$$

As mentioned in Section 2.1, different choices for T simply represent different bases for the complementary process. Below we consider two values for T which suggest the operator forms for M_z used in Sections 4.1 and 4.3

Two Special Cases

(1) If we let (Here we assume Σ_x is nonsingular)

$$T = -\Sigma_x^{-1} \quad ,$$

then

$$M_z = \left[-I \;\vdots\; H'\right]\Sigma_\zeta^{-1} \quad ,$$

which leads to the operator form hypothesized in Section 4.1

(2) Another simple form is obtained if we set

$$T = -I \quad ,$$

then

$$M_z = \left[-I \;\vdots\; \Sigma_x H'\Sigma_r^{-1}\right] \quad .$$

This suggests the operator form employed in (4.14) which is applicable when Σ_x is nonsingular.