Recursive Algorithms for linear least squares estimation problems have been based mainly on state-space models. Recently, some new recursive solutions were obtained for processes classified in terms of their "index of nonstationarity" or equivalently—the displacement rank of their covariance functions. While this definition provides a natural explanation of the properties of constant-coefficient state-space models, it is not satisfactory for time-variant models. However a modified definition of the displacement rank makes it possible to embed time-varying state-space models in the more general input-output framework. In
So doing, we are able to show the mutual relationships of the Kalman filter Riccati equation, the time-varying Chandrasekhar equations and the Krein-Levinson equations.
A MODIFIED DISPLACEMENT RANK AND SOME APPLICATIONS†

B. Friedlander ‡ T. Kailath and M. Morf
Information Systems Laboratory
Stanford University, CA 94305

Abstract

Recursive algorithms for linear least squares estimation problems have been based mainly on state-space models. Recently, some new recursive solutions were obtained for processes classified in terms of their "index of nonstationarity" or equivalently—the displacement rank of their covariance functions. While this definition provides a natural explanation of the properties of constant-coefficient state-space models, it is not satisfactory for time-varying models. However, a modified definition of the displacement rank makes it possible to embed time-varying state-space models in the more general input-output framework. In so doing, we are able to show the mutual relationships of the Kalman filter Riccati equation, the time-varying Chandrasekhar equations and the Krein-Levinson equations.

I. Introduction

In [1]—[5] we have developed recursive estimation algorithms using input-output models (e.g., covariance functions) instead of state-space models. A central idea in our approach was that of the displacement rank $\alpha$ (an "index of nonstationarity") of the covariance functions of the signal and observation processes. Using this notion certain Sobolev- and Krein-Levinson-type differential equations were developed for the optimal smoother and optimal filter, leading to computational algorithms whose complexity depends on the displacement rank $\alpha$.

By imposing constant-parameter state-space structure on the covariance functions, we then showed in [1] how the Krein-Levinson equations led to the Chandrasekhar equations for the computation of the Kalman gain. The successful imbedding of the state-space case in the input-output framework was due to the fact that the covariance functions of constant parameter state-space models have a (relatively) small displacement rank. However, processes associated with time varying models will not have small $\alpha$ and may, in fact, have infinite displacement ranks. This fact prevented us from treating time varying models in our input-output framework.

† This work was supported by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract AF44-620-74-C-0068, and in part by the National Science Foundation under Contract NSF-Eng-75-18952 and the Joint Services Electronics Program under Contract N00014-75-C-0061.

‡ Now with Systems Control, Inc., Palo Alto, CA 94304.

In Section II we shall show that this difficulty can be circumvented by introducing a more general definition that (i) leads to small values of $\alpha$ both for time-varying and time-invariant state-space models, (ii) coincides with the previous definition of $\alpha$ in the time-invariant case. In Sec. III, we note briefly how the general results of [1]—[2] will be modified with the new definition of displacement rank—the changes are minor. Then in Sec. IV we shall show how the imposition of state-space structure leads to the time-variant Chandrasekhar equations of [6]. A direct derivation of these equations was first given in [7]—[8] (see also [9]—[10]). Unfortunately, as noted in [8]—[9] the time-varying version is just a set of two-point boundary value equations, which is not especially easy to solve. In fact, a standard approach to two-point equations is via the Riccati equation, and we shall show that this can be done here as well. Of course, the Riccati equation could also have been directly obtained from the state-space models. (as in the usual Kalman filter). The contribution here is that we show how to deduce the Riccati equation from a more general set of equations applicable when no state-space models are available.

Most proofs are omitted in this short paper, but may be found in [11]; there we also indicate the analogous results for discrete-time estimation.

II. The displacement rank of covariance functions

The displacement rank of a kernel $K(t,s)$ is defined in [1] as the smallest integer $\alpha$ such that we can write

$$
\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial s^2} K(t,s) = K(t,s) K(t,s) + D(t) A D'(s)
$$

where $K$, $D$, and $A$ are matrix functions with dimensions $p \times p$, $p \times \alpha$ and $\alpha \times \alpha$ respectively. The functions $D$ and $A$ need not be unique, though it will often be simplest to assume $A$ to be diagonal.

The reasons for choosing this definition of the displacement rank and its application in solving Fredholm and Wiener-Hopf type integral equations are discussed in detail in [1], [2] and will not be repeated here.

Instead we shall focus on the special case of processes generated by lumped state-space models

$$
x(t) = F(t)x(t) + G(t)u(t), \quad x(t) = x_0
$$

$$
y(t) = H(t)x(t) + v(t), \quad t > T
$$

where $x(0)$ is mx1, $u(\cdot)$ is mx1, and $y(\cdot)$ scalar,

$$
E[u(t)]u'(s) = Q(t) \delta(t-s), \quad E[v(t)v'(s)] = 0 \delta(t-s)
$$

and

$$
E[u(t)v'(s)] = 0 \delta(t-s),
$$

and

where $\delta(t-s)$ is Dirac's delta.

This work was supported by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract AF44-620-74-C-0068, and in part by the National Science Foundation under Contract NSF-Eng-75-18952 and the Joint Services Electronics Program under Contract N00014-75-C-0061.
When the model parameters \( F(\cdot), G(\cdot), H(\cdot), Q(\cdot) \) are constant then it was shown in [11] that in problems where there is no dependence on \( \tau \) (say a \( K(t,s) \) defined for \( 0 \leq t, s \leq T \)), the new definition can still be used by introducing \( \tau \) artificially, say by \( K(t,s) = K(t-\tau, s-\tau) \), so that

\[
\frac{\partial}{\partial \tau} K(t-\tau, s-\tau) \bigg|_{\tau = 0} = \frac{\partial}{\partial s} + \frac{\partial}{\partial t} K(t,s),
\]

which coincides with the original definition.

III. The modified Sobolev- and Levinson-Krein-type equations

Let us now consider the problem of estimating a stochastic process \( x(\cdot) \) (n-dimensional) from observations of a related process \( y(\cdot) \) (p-dimensional), using the knowledge of their covariance functions,

\[
E(y(t)y'(s)) = I \delta(t-s) + K(t,s), \quad 0 \leq t, s \leq T
\]

\[
E(x(t)y'(s)) = K_{xy}(t,s)
\]

It is well known that the optimal smoother \( K_{xy}(t,s) \) and optimal filter \( h_{xy}(t,s) \) for the process \( x(\cdot) \) can be obtained as the solution of certain integral equations. For example, \( h_{xy}(t,s) \) obeys a Wiener-Hopf equation of the second kind. The notion of the displacement rank was useful in reducing the solution of such equations to that of the generalized Krein-Levinson equations (cf., [1], [2]). To do this, it was necessary to make the following structural assumption about the cross covariance function \( K_{xy}(\cdot,\cdot) \),

\[
(\frac{\partial}{\partial t} + \frac{\partial}{\partial s})K_{xy}(t,s) = K_{xy}(t-t)K(t,s) + D_{xy}(t)sD'(s)
\]

where \( D, A \) are as defined earlier, and \( D_{xy} \) is defined by the equation above. Using our new definition for the displacement rank we shall replace (9) by

\[
-\frac{\partial}{\partial t} K_{xy}(t,s) = K_{xy}(t-t)K(t,s) + D_{xy}(t)A D'(s).
\]

It can be shown that with the new definitions, we can obtain a very similar set of equations for \( h_{xy}, K_{xy} \) to those presented in [1]. In fact, the only changes that have to be made are replacing \( \frac{\partial}{\partial t} + \frac{\partial}{\partial s} \) by \( (-\frac{\partial}{\partial t}) \) and using the new values of \( D, D_{xy}, A, \alpha \).

We give here only a subset of the analogs of the equations of [1], in fact only those necessary for the analysis in Sec. IV:

\[
-\frac{\partial}{\partial t} h_{xy}(t,s) = B_{xy}(t-t)A B'(t;s)
\]

where

\[
E x x' = P(t), \quad \text{a given n x n matrix.}
\]
IV. Embedding the state-space case in the input-output framework.

Let us now introduce some further assumptions about the covariance functions \( K, K_{xy} \), which will impose a state-space type structure on the processes \( x(t) \) and \( y(t) \). We assume that there exist functions \( F(\cdot) \) (\( n \times m \)) and \( H(\cdot) \) (\( p \times n \)) such that \( K(t,s) = H(t)K_{xy}(t,s) \), and

\[
\frac{\partial}{\partial t} K_{xy}(t,s) = F(t)K_{xy}(t,s),
\]

where \( h_{xy}(t,.) \) can be identified as the Kalman gain. The significance of this fact is that, at least in the constant-parameter case, it was shown [1], [7] that the Chandrasekhar-type equations could be used to compute \( h_{xy}(t,.) \), instead of having to compute \( h_{xy}(t,s) \) for all \( t \leq s \leq t \) as would be required when no state-space structure is available.

We shall now show what are the corresponding equations for the time-varying case. With the assumptions (13), we see from (12a,b) that we can identify

\[
B(t;s) = H(s)B_{xy}(t;s)
\]

and therefore, we can write (11) as

\[
-\frac{\partial}{\partial t} h_{xy}(t,.) = B_{xy}(t;.)A B_{xy}(t;.)H'(t)
\]

With a little calculation (see [11] and [1]), it can also be shown that

\[
\frac{\partial}{\partial t} B_{xy}(t;.) = (F(t) - h_{xy}(t,.)H(t))B_{xy}(t;.)H(t).
\]

Equations (14), (15) are the time-varying version of the Chandrasekhar-type equations that were derived using a completely different approach in [8], [9]. Note that (14), (15) can not be solved directly since the time arguments "do not fit". That is, because of the opposite directions of evolution of these equations, at any intermediate point the values of \( B_{xy}(\cdot,.') \) needed to solve for \( h_{xy}(\cdot,') \), will not be available. This difficulty is circumvented in the time-invariant model case, because now we can reverse the direction of time in (14) since in this case

\[
\frac{\partial}{\partial t} h_{xy}(t,.) = -\frac{\partial}{\partial t} h_{xy}(t,.)
\]

In the time-variant case, the Chandrasekhar equations (first obtained in [8]) have to be regarded as a general set of two-point boundary-value equations, with all the attendant difficulties. It is well known that the Riccati equation enables us to replace the two-point Hamiltonian equations of control and estimation theory by an initial-value equation, and this can be done here as well.

V. Concluding Remarks

We presented a new definition for the displacement rank of covariance functions that makes it possible to embed both time-invariant and time-varying state-space models in a more general input-output framework. This approach provides insight into the relationship between various solutions to the estimation problem and clarifies the role of the state-space structure in simplifying the estimation algorithms. For processes with measure \( \alpha \) of "distance from stationarity" recursive estimation algorithms of the Levinson-type can be derived [1], [4]. If additional (state space type) structure is added to the problem, alternative algorithms become available. In the time-varying state-space case the Kalman-filter and the Riccati equation are naturally obtained, while in the constant-parameter case the more efficient Chandrasekhar equations may be used. Of course the general input-output recursions have to be used when state-space models are not readily available.

References


