Design of Kalman filters using signal-model output statistics

Le Hung Son, B.E., and Prof. B. D. O. Anderson, Ph.D.

Indexing terms: Kalman filters, Statistics

ABSTRACT

The Kalman filter depends only on the output statistics of the message model; a technique for filter construction using only these statistics is given. The performance of the filter is considered, and, for state rather than signal estimation, the performance is found to depend on the details of the model, as distinct from its output statistics.

1 INTRODUCTION

In the usual theory and application of Kalman filtering[1-3] to discrete-time processes, one requires complete knowledge of the signal or message process from which measurements are being derived before one can construct the Kalman filter. This is in contrast to most of the Wiener filtering theory, where one usually requires knowledge of the output statistics only of a signal process.[4,5]

There appears to be some need for eliminating the need to know details of the signal-process model in designing Kalman filters. Bucy[6] has drawn attention to the existence of problems in bioengineering where the model is not available, where output statistics are available and where a filter is desired, and a second most important area, where knowledge of the signal-process model is lacking, lies in adaptive filtering.[7] Here, one initially does not even have the output statistics available, and one is required to construct a Kalman filter. In Reference 7, a 3-step procedure has been proposed, of adaptively identifying the power spectrum of the output process, passing to a model of the message process, and thence to the Kalman filter. This is complex, and suffers from the conceptual difficulty that the message process is not uniquely determined by the output-power spectrum.

In this paper, we shall show that the Kalman filter is, to a surprising degree, independent of the message process, and depends purely on the output statistics. We shall also show how to compute the filter from the output statistics alone. This technique is currently being applied to the development of adaptive filters, using some of the ideas of Reference 7, but from the output statistics. This work will be reported elsewhere.

Similar results to those of this paper have been reported for continuous time filtering.[8,9] Many of the arguments used in this paper parallel those of Reference 9, and, given the results of Reference 9, those of this paper are not particularly surprising. In one respect however, we differ significantly from the previous results. This is in a discussion in Section 4 of the spectral-factorisation problem. In the continuous-time case, the necessary background results were to be found in the literature,[10] while this is not so for discrete-time systems.

2 PRELIMINARY ANALYSIS

Consider the linear, discrete finite-dimensional system depicted in Fig. 1 and described by the state-space equations

$$\mathbf{x}(k+1) = \mathbf{F}(k)\mathbf{x}(k) + \mathbf{G}(k)\mathbf{u}(k) \tag{1a}$$

$$\mathbf{z}(k) = \mathbf{H}'(k)\mathbf{x}(k) + \mathbf{J}_1(k)\mathbf{u}(k) + \mathbf{J}_2(k)\mathbf{v}(k) \tag{1b}$$

where $k \in [k_0, k_f]$, $\mathbf{x}(\cdot)$ is an n-dimensional state vector, $\mathbf{u}(\cdot)$ is a p-dimensional input noise vector, $\mathbf{v}(\cdot)$ is an m-dimensional vector representing measurement noise, and $\mathbf{z}(\cdot)$ is an r-dimensional output vector. The matrices $\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{J}_1$ and $\mathbf{J}_2$ are of appropriate dimensions, and $\mathbf{F}(k)$ is nonsingular for all k. Henceforth, when an argument is suppressed, it takes the value k. We shall think of $\mathbf{z}$ as being the sum of a signal

$$\mathbf{y} = \mathbf{H}'\mathbf{x} \tag{2}$$

and noise

$$\mathbf{n} = \mathbf{J}_1\mathbf{u} + \mathbf{J}_2\mathbf{v} \tag{3}$$

The vectors $\mathbf{u}(\cdot)$ and $\mathbf{v}(\cdot)$ denote mutually independent vector random processes which are zero mean, Gaussian and white, with covariance

$$E[\mathbf{u}(k)\mathbf{u}'(\ell)] = \mathbf{I}_p\delta(k-\ell) \quad E[\mathbf{v}(k)\mathbf{v}'(\ell)] = \mathbf{I}_m\delta(k-\ell) \tag{4}$$

The initial state vector of the system $\mathbf{x}(k_0)$ is a zero-mean Gaussian random variable with covariance

$$E[\mathbf{x}(k_0)\mathbf{x}'(k_0)] = \mathbf{P}_0 \tag{5}$$

and is independent of $\mathbf{u}(\cdot)$ and $\mathbf{v}(\cdot)$.

We define $E[\mathbf{x}(k)\mathbf{x}'(k)] = \mathbf{P}(k)$; then it is easily shown that $\mathbf{P}(k)$ is given as the solution of the linear difference equation

$$\mathbf{P}(k+1) = \mathbf{F}\mathbf{P}(k)\mathbf{F}' + \mathbf{G}\mathbf{G}' \tag{6}$$

and, in fact, one may readily verify the following equalities:

$$\begin{aligned} E[\mathbf{y}(k)\mathbf{y}'(\ell)] = {} & \mathbf{H}'(k)\mathbf{\Phi}(k,\ell)\mathbf{P}(\ell)\mathbf{H}(\ell)1(k-\ell) \\ & + \mathbf{H}'(k)\mathbf{P}(k)\mathbf{\Phi}'(\ell,k)\mathbf{H}(\ell)1(\ell-k) \\ & + \mathbf{H}'(k)\mathbf{P}(k)\mathbf{H}(k)\delta(k-\ell) \end{aligned} \tag{7a}$$

$$E[\mathbf{y}(k)\mathbf{n}'(\ell)] = \mathbf{H}'(k)\mathbf{\Phi}(k,\ell)\mathbf{F}^{-1}(\ell)\mathbf{G}(\ell)\mathbf{J}_1'(\ell)1(k-\ell) \tag{7b}$$

$$E[\mathbf{n}(k)\mathbf{n}'(\ell)] = \mathbf{R}(k)\delta(k-\ell) \tag{7c}$$

$$\begin{aligned} E[\mathbf{z}(k)\mathbf{z}'(\ell)] = {} & \mathbf{H}'(k)\mathbf{\Phi}(k,\ell)\mathbf{K}(\ell)1(k-\ell) \\ & + \mathbf{K}'(k)\mathbf{\Phi}'(\ell,k)\mathbf{H}(\ell)1(\ell-k) + \mathbf{L}(k)\delta(k-\ell) \end{aligned} \tag{7d}$$

Here, $\mathbf{\Phi}(\cdot,\cdot)$ is the transition matrix associated with $\mathbf{x}(k+1) = \mathbf{F}(k)\mathbf{x}(k)$, and the matrices $\mathbf{R}(\cdot)$, $\mathbf{L}(\cdot)$ and $\mathbf{K}(\cdot)$ are defined by

$$\mathbf{R} = \mathbf{J}_1\mathbf{J}_1' + \mathbf{J}_2\mathbf{J}_2' \tag{8}$$

$$\mathbf{L} = \mathbf{R} + \mathbf{H}'\mathbf{P}\mathbf{H} \tag{9}$$

$$\mathbf{K} = \mathbf{P}\mathbf{H} + \mathbf{F}^{-1}\mathbf{G}\mathbf{J}_1' \tag{10}$$

Further, $1(k)$ is 1 for $k > 0$ and zero otherwise, and $\delta(k)$ is 1 for $k = 0$ and zero otherwise.

The scheme of Fig. 1 is typical of the process and noise models for which one might want to build Kalman filters. However, there exist other schemes for which one might want to construct a Kalman filter, such as that shown in Fig. 2.

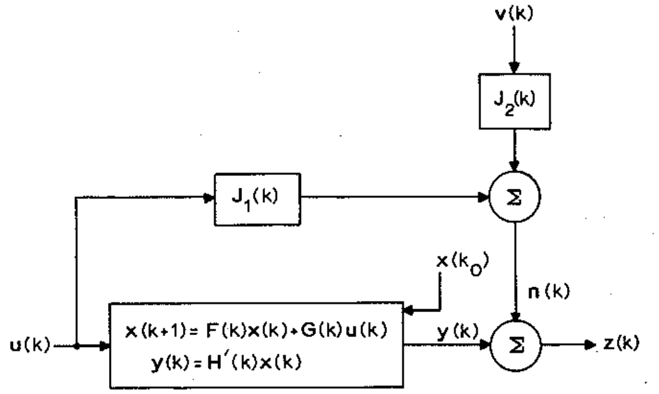

Fig. 1

General arrangement for generating signal process

Paper 6845 C, first received 13th September 1971 and in revised form 27th September 1972

Mr. Son and Prof. Anderson are with the Department of Electrical Engineering, University of Newcastle, Newcastle, NSW 2308, Australia

In Fig. 2, $\mathbf{u}(\cdot)$ and $\mathbf{n}(\cdot)$ are not independent, and we have

$$E[\mathbf{u}(k)\mathbf{n}'(\ell)] = \mathbf{J}_3(k)\delta(k-\ell) \tag{11}$$

Despite the apparent difference, the arrangement in Fig. 2 is equivalent to one of the sort depicted by Fig. 1, where $\mathbf{u}(\cdot)$ and $\mathbf{v}(\cdot)$ are independent. In precise terms, the equivalence

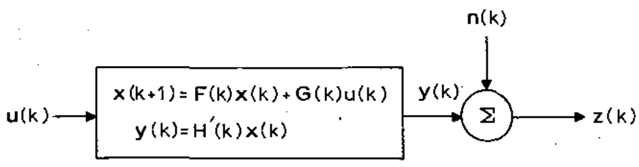

Fig. 2

Alternative arrangement for generating signal process

is defined (as direct calculation shows) by the following equations: the matrices $\mathbf{J}_1(\cdot)$ and $\mathbf{J}_2(\cdot)$ in Fig. 1 are given by

$$\mathbf{J}_1 = \mathbf{J}_3' \tag{12}$$

and $\mathbf{J}_2(\cdot)$ is any matrix such that

$$\mathbf{J}_2\mathbf{J}_2' = \mathbf{R} - \mathbf{J}_1\mathbf{J}_1' \tag{13}$$

The relationship between $\mathbf{n}(\cdot)$ in Fig. 2 and the independent processes $\mathbf{u}(\cdot)$ and $\mathbf{v}(\cdot)$ of Fig. 1 is

$$\mathbf{n} = \mathbf{J}_1\mathbf{u} + \mathbf{J}_2\mathbf{v} \tag{14}$$

In the sequel, we shall study the properties of Kalman filters for schemes of the form of Fig. 1 only. In conformity with usual practice, we shall assume that the estimation problem (stated subsequently in precise terms) is nonsingular; i.e. we shall assume that $\mathbf{R} = \mathbf{J}_1\mathbf{J}_1' + \mathbf{J}_2\mathbf{J}_2'$ is nonsingular. In other words (see eqn. 7c), we can, in loose terms, say that the additive output noise $\mathbf{n}(\cdot)$ is nonsingular.

3 INVARIANCE OF KALMAN-BUCY FILTER WITH RESPECT TO SIGNAL MODEL

The Kalman filter is, as is well known, a device for producing a minimum-error-variance online estimate of the state $\mathbf{x}(k)$ of a model of the form of Fig. 1, given the output measurements $\mathbf{z}(\cdot)$ up till time $k-1$ (or, possibly, k—the distinction here is unimportant). From an estimate $\hat{\mathbf{x}}(k)$ of $\mathbf{x}(k)$, a minimum-error-variance estimate of $\mathbf{y}(k)$ follows as $\hat{\mathbf{y}}(k) = \mathbf{H}'(k)\hat{\mathbf{x}}(k)$. In this Section, we aim to show that there are many signal-generating systems of the form of Fig. 1 which possess the same Kalman filter; in fact, we shall show that, *roughly speaking, all systems with the same output statistics possess the same filter.* (We defer to a later Section the statement of a technique allowing construction of the filter from these output statistics).

Observe that a general property of systems of the form of Fig. 1 is that

$$E[\mathbf{y}(k)\mathbf{n}'(\ell)] = 0 \quad k \leq \ell \tag{15}$$

Also, as shown in Section 2, the quantity $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$ is of the general form

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = \mathbf{A}'(k)\mathbf{B}(\ell) \quad k \geq \ell \tag{16}$$

for $k \geq \ell$ [this follows from eqn. 7a and the fact that $\Phi(k, \ell) = \Phi(k, k_0)\Phi(k_0, \ell)$ for any fixed k_0, independent of k and ℓ]. Further, as also shown in Section 2, we have

$$E[\mathbf{n}(k)\mathbf{n}'(\ell)] = \mathbf{R}(k)\delta(k-\ell) \tag{17}$$

for some positive definite symmetric $\mathbf{R}(\cdot)$. The matrices $\mathbf{A}(\cdot), \mathbf{B}(\cdot)$ and $\mathbf{R}(\cdot)$ are, as we have shown, computable from $\mathbf{F}(\cdot), \mathbf{G}(\cdot)$ etc., if these quantities are known. On the other hand, if $\mathbf{A}(\cdot), \mathbf{B}(\cdot)$ and $\mathbf{R}(\cdot)$ are known—perhaps as a result of statistical analysis of the system output—we cannot infer from them the values of $\mathbf{F}(\cdot), \mathbf{G}(\cdot)$ etc. Although the usual presentation of Kalman filtering suggests that knowledge of $\mathbf{F}(\cdot), \mathbf{G}(\cdot)$ etc. is necessary to construct the Kalman filter, this is actually not the case. Our first theorem is to the effect that $\mathbf{A}(\cdot), \mathbf{B}(\cdot)$ and $\mathbf{R}(\cdot)$ suffice to determine a minimum-error-variance filter for $\mathbf{y}(\cdot)$, in the sense that all signal-process models possessing the same $\mathbf{A}(\cdot), \mathbf{B}(\cdot)$ and $\mathbf{R}(\cdot)$, defined as above, will also possess the same Kalman filter.

Theorem 1: Suppose that, for a system of the form of Fig. 1, eqns. 15–17 are the only equations known which describe the system, with $\mathbf{A}(\cdot), \mathbf{B}(\cdot)$ and $\mathbf{R}(\cdot)$ known matrices, and with $\mathbf{R}(\cdot)$ positive definite symmetric. Then the linear system generating a minimum-error-variance unbiased estimate $\hat{\mathbf{y}}(k/k-1)$ of $\mathbf{y}(k)$ from $\mathbf{z}(\ell)$ for $k_0 \leq \ell < k$ is independent of the particular signal-process model, and depends only on $\mathbf{A}(\cdot), \mathbf{B}(\cdot)$ and $\mathbf{R}(\cdot)$.

Proof: application of the projection theorem shows easily that the impulse response $\mathbf{A_y}(k, \ell)$ of the minimum-variance unbiased filter is given by

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = \sum_{m=k_0}^{k-1} \mathbf{A_y}(k, m)E[\mathbf{z}(m)\mathbf{z}'(\ell)] \qquad k > \ell \tag{18}$$

The quantities $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$ and $E[\mathbf{z}(m)\mathbf{z}'(\ell)]$ are known (the latter using eqns. 15-17 and are independent of the particular generating signal-process model. Moreover, since $E[\mathbf{z}(m)\mathbf{z}'(\ell)]$ is positive definite for all ℓ (see Appendix 14), $\mathbf{A_y}(\cdot, \cdot)$ is uniquely determined.

The above result, although straightforward, is not quite as simple as it might initially appear. Notice that eqn. 18, defining the optimal filter, does not involve $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$ and $E[\mathbf{y}(k)\mathbf{n}'(\ell)]$ separately, but merely involves the sum of these two quantities $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$. Therefore complete knowledge of the statistics of $\mathbf{y}(\cdot)$ and $\mathbf{n}(\cdot)$ is not required, contrary to what might initially have been expected.

Now we turn to state estimation, which can be thought of as the true aim of the Kalman filter. Before one can consider the problem sensibly, it is necessary to specify precisely the co-ordinate basis for the state space of the generating system. We shall do this by demanding that, in lieu of eqn. 16, we have

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = \mathbf{H}'(k)\Phi(k, \ell)\mathbf{K}(\ell) \qquad k \geq \ell \tag{19}$$

with $\mathbf{H}(\cdot)$, $\Phi(\cdot, \cdot)$ and $\mathbf{K}(\cdot)$ separately known, and with $\mathbf{K}(\cdot)$ identifiable given $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$, $\mathbf{H}(\cdot)$ and $\Phi(\cdot, \cdot)$. In other words, we assume known in the arrangement of Fig. 1 the matrices $\mathbf{F}(\cdot)$ and $\mathbf{H}(\cdot)$, and we assume that $[\mathbf{F}(\cdot), \mathbf{H}(\cdot)]$ is completely observable, so that, if $\mathbf{H}'(k)\Phi(k, \ell)\mathbf{K}_2(\ell) = \mathbf{H}'(k)\Phi(k, \ell)\mathbf{K}_2(\ell)$ for all k and ℓ, then $\mathbf{K}_1(\cdot) = \mathbf{K}_2(\cdot)$. The complete-observability constraint is a natural one to impose when posing the problem of identifying the state of a system given output measurements, whether the latter are noisy or not.

Theorem 2: Suppose that, for a system of the form of Fig. 1, $\mathbf{F}(\cdot)$ and $\mathbf{H}(\cdot)$ are known, with $[\mathbf{F}, \mathbf{H}]$ completely observable, and that quantities $\mathbf{R}(\cdot)$ in eqn. 17 and $\mathbf{K}(\cdot)$ in eqn. 19 are known, with $\mathbf{R}(\cdot)$ positive definite symmetric. Then the linear system generating a minimum-error-variance unbiased estimate $\hat{\mathbf{x}}(k/k-1)$ of $\mathbf{x}(k)$ from $\mathbf{z}(\ell)$, $k_0 \leq \ell < k$ is independent of the particular signal-process model, and depends only on $\mathbf{F}(\cdot), \mathbf{H}(\cdot), \mathbf{K}(\cdot)$ and $\mathbf{R}(\cdot)$.

Proof: The impulse response $\mathbf{A_x}(\cdot, \cdot)$ of the minimum-variance unbiased filter is given by

$$E[\mathbf{x}(k)\mathbf{z}'(\ell)] = \sum_{m=k_0}^{k-1} \mathbf{A_x}(k, m)E[\mathbf{z}(m)\mathbf{z}'(\ell)] \qquad k > \ell \tag{20}$$

The positive definiteness of $E[\mathbf{z}(m)\mathbf{z}'(\ell)]$ again guarantees existence and uniqueness of a solution to eqn. 20, at least if $E[\mathbf{x}(k)\mathbf{z}'(\ell)]$ is available.

We know that $E[\mathbf{z}(m)\mathbf{z}'(\ell)]$ is independent of the particular signal model, and is obtainable from $\mathbf{F}(\cdot), \mathbf{H}(\cdot), \mathbf{K}(\cdot)$ and $\mathbf{R}(\cdot)$.

For the system depicted in Fig. 1, it is also straightforward to show that

$$E[\mathbf{x}(k)\mathbf{z}'(\ell)] = \Phi(k, \ell)\mathbf{K}(k) \qquad k > \ell \tag{21}$$

This means that $E[\mathbf{x}(k)\mathbf{z}'(\ell)]$ is determinable from $\mathbf{F}(\cdot)$ and $\mathbf{K}(\cdot)$ alone.

Notice that the theorem gives no explicit construction for the Kalman filter. Direct solution of eqn. 20, based on inversion of $E[z(m)z'(\ell)]$, would, in theory, be possible. However, this would not yield, at least in a straightforward fashion, the Kalman filter in its usual finite-dimensional format. In later Sections, we shall study the question of efficiently obtaining the Kalman filter from $\mathbf{F}(\cdot), \mathbf{H}(\cdot), \mathbf{K}(\cdot)$ and $\mathbf{R}(\cdot)$ only. Notice also that the filter is determined by $\mathbf{F}(\cdot), \mathbf{H}(\cdot)$ and $\mathbf{K}(\cdot)$ or $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$, which is the sum of $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$ and $E[\mathbf{y}(k)\mathbf{n}'(\ell)]$. The individual summands are not required in defining the filter.

4 SPECTRAL-FACTORISATION RESULT

Having established that the Kalman filter depends only on the output statistics of a signal-process model, and not on the model itself, we now seek a procedure for obtaining the filter in finite-dimensional form from these statistics. As a preliminary to this, we study, in this Section, a problem of spectral factorisation. In this context, this is the problem of determining a linear, discrete finite-dimensional system, such that, when excited by white noise, its output is a sample function of a specified covariance. In Section 5, we shall use the spectral-factorisation result to obtain a computational algorithm for the Kalman filter.

A summary of this Section is as follows: in theorem 3, we establish the existence of a spectral factor, or system with the requisite properties. In theorem 4, we establish something about the form of this system, and in theorem 5 we establish further properties. These are used to establish theorem 6, dealing with the existence of the solution of a discrete-time Riccati equation. Theorem 6 is the ultimate goal of this Section.

Theorem 3: Let

$$\mathbf{R}(k,\ell) = \mathbf{H}'(k)\Phi(k,\ell)\mathbf{K}(\ell)1(k-\ell) + \mathbf{K}'(k)\Phi'(\ell,k)\mathbf{H}(\ell)1(\ell-k) + \mathbf{L}(k)\delta(k-\ell) \quad (22)$$

be a positive definite covariance on $[k_0, k_f]$. Suppose that $\Phi(k,\ell)$ is nonsingular for all k and ℓ. Then there exists an impulse-response matrix $\mathbf{w}(\cdot,\cdot)$, causal, in the sense that $\mathbf{w}(k,\ell) = 0$ for $\ell > k$, so that a system excited by white noise of variance $I\delta(k-\ell)$ has output covariance $\mathbf{R}(\cdot,\cdot)$, i.e.

$$\mathbf{R}(k,\ell) = \sum_{m=k_0}^{k_f} \mathbf{w}(k,m)\mathbf{w}'(\ell,m) \quad (23)$$

Proof: Define $\mathcal{R}$ as the matrix

$$\mathcal{R} = \begin{bmatrix} \mathbf{R}(k_0,k_0) & \cdot\;\cdot\;\cdot & \mathbf{R}(k_0,k_f) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{R}(k_f,k_0) & & \mathbf{R}(k_f,k_f) \end{bmatrix} \quad (24)$$

Positive definiteness of the covariance $\mathbf{R}(\cdot,\cdot)$ implies positive definiteness of the matrix $\mathcal{R}$. Therefore there exists a lower triangular matrix $\mathcal{W}$ so that[11]

$$\mathcal{R} = \mathcal{W}\mathcal{W}' \quad (25)$$

Partition the matrix $\mathcal{W}$ as for $\mathcal{R}$ to define blocks $\mathbf{w}(k,\ell)$ by

$$\mathcal{W} = \begin{bmatrix} \mathbf{w}(k_0,k_0) & 0 & \cdot\;\cdot\;\cdot & 0 \\ \mathbf{w}(k_0+1,k_0) & \mathbf{w}(k_0+1,k_0+1) & & \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \mathbf{w}(k_f,k_0) & & & \mathbf{w}(k_f,k_f) \end{bmatrix} \quad (26)$$

Notice that $\mathbf{w}(k,\ell) = 0$ for $k < \ell$. Eqn. 23 is immediate. Notice that the positive definiteness of $\mathbf{R}$ implies that each $\mathbf{w}(i,i)$ is nonsingular.

We now show that $\mathbf{w}(\cdot,\cdot)$ is the impulse response of a finite-dimensional system.

Theorem 4: Let $\mathbf{w}(k,\ell)$ be defined as in theorem 3; then $\mathbf{w}(k,\ell)$ has the form

$$\mathbf{w}(k,\ell) = \hat{\mathbf{L}}(k)\delta(k-\ell) + \mathbf{H}'(k)\Phi(k,\ell)\mathbf{M}(\ell)1(k-\ell) \quad (27)$$

for certain matrices $\hat{\mathbf{L}}(k)$ and $\mathbf{M}(\ell)$.

Proof: Let $\mathbf{V}$ be a matrix, such that $\mathcal{V} = \mathcal{W}^{-1}$; then $\mathcal{V}$ is a lower triangular matrix like $\mathcal{W}$. Let $\mathcal{V}$ be partitioned like $\mathcal{W}$. Now, $\mathcal{R} = \mathcal{W}\mathcal{W}'$, and so

$$\mathcal{W} = \mathcal{R}\mathcal{V}' \quad (28)$$

The matrix multiplication is carried out and the form of $\mathbf{R}(k,\ell)$ is used, and the triangular nature of $\mathbf{V}$ leads to

$$\mathbf{w}(k,\ell) = [\mathbf{L}(k)\mathbf{v}'(k,k) + \mathbf{H}'(k)\sum_{m=k_0}^{k} \Phi(k,m)\mathbf{K}(m)\mathbf{v}'(k,m)]\delta(k-\ell) + \mathbf{H}'(k)\Phi(k,\ell)[\mathbf{K}(\ell)\mathbf{v}'(\ell,\ell) + \sum_{m=k_0}^{\ell} \Phi(\ell,m)\mathbf{K}(m)\mathbf{v}'(\ell,m)]1(k-\ell) \quad (29)$$

Eqn. 29 implies eqn. 27.

Notice that $\hat{\mathbf{L}}(k) = \mathbf{w}(k,k)$ is nonsingular for all k. Further properties of $\hat{\mathbf{L}}(\cdot)$ and $\mathbf{M}(\cdot)$ are as follows

Theorem 5: With $\mathbf{M}(\cdot)$ and $\hat{\mathbf{L}}(\cdot)$ defined as in theorem 4, let $\mathbf{P}_g(k)$ be defined as follows

$$\mathbf{P}_g(k+1) = \mathbf{F}[\mathbf{P}_g(k) + \mathbf{M}\mathbf{M}']\mathbf{F}' \qquad \mathbf{P}_g(k_0) = 0 \quad (30)$$

Then

$$\mathbf{M} = [\mathbf{K} - \mathbf{P}_g\mathbf{H}][\hat{\mathbf{L}}']^{-1} \quad (31)$$

and

$$\hat{\mathbf{L}}\hat{\mathbf{L}}' = \mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H} \quad (32)$$

Proof: from the definition of $\mathbf{P}_g(k)$, we have

$$\sum_{m<k,\,\ell} \Phi(k,m)\mathbf{M}(m)\mathbf{M}'(m)\Phi'(\ell,m) = \Phi(k,\ell)\mathbf{P}_g(\ell)1(k-\ell) + \mathbf{P}_g(k)\Phi'(\ell,k)1(\ell-k) + \mathbf{P}_g(k)\delta(k-\ell)$$

Direct substitution of eqn. 27 into eqn. 23 and use of this identity yields

$$\mathbf{R}(k,\ell) = [\hat{\mathbf{L}}(k)\hat{\mathbf{L}}'(k) + \mathbf{H}'(k)\mathbf{P}_g(k)\mathbf{H}(k)]\delta(k-\ell) + \mathbf{H}'(k)\Phi(k,\ell)[\mathbf{M}(\ell)\hat{\mathbf{L}}'(\ell) + \mathbf{P}_g(\ell)\mathbf{H}(\ell)]1(k-\ell) + [\hat{\mathbf{L}}(k)\mathbf{M}'(k) + \mathbf{H}'(k)\mathbf{P}_g(k)]\Phi'(\ell,k)\mathbf{H}(\ell)1(\ell-k) \quad (33)$$

Comparison of eqns. 33 and 22 establishes eqns. 31 and 32.

The key to what follows is the matrix $\mathbf{P}_g(\cdot)$, introduced in theorem 5. This theorem, with the earlier ones, suffices to establish the existence of $\mathbf{P}_g(\cdot)$, for, recall that, from $\mathbf{R}(\cdot,\cdot)$, we established the existence of $\mathbf{w}(\cdot,\cdot)$, from $\mathbf{w}(\cdot,\cdot)$, the existence of $\mathbf{M}(\cdot)$, and, in theorem 5, from $\mathbf{M}(\cdot)$, the existence of $\mathbf{P}_g(\cdot)$. Other than existence however, theorem 5, in essence, provides another computational algorithm for obtaining $\mathbf{P}_g(\cdot)$. Eqns. 30-32 allow elimination of $\mathbf{M}(\cdot)$ and $\hat{\mathbf{L}}(\cdot)$ to yield

$$\mathbf{P}_g(k+1) = \mathbf{F}\{\mathbf{P}_g(k) + [\mathbf{P}_g(k)\mathbf{H} - \mathbf{K}][\mathbf{L} - \mathbf{H}'\mathbf{P}_g(k)\mathbf{H}]^{-1} \times [\mathbf{P}_g(k)\mathbf{H} - \mathbf{K}]'\}\mathbf{F}' \qquad \mathbf{P}_g(k_0) = 0 \quad (34)$$

This is a recursive equation for $\mathbf{P}_g(\cdot)$; in fact, it is a discrete Riccati equation. As we have shown, $\mathbf{P}_g(\cdot)$ always exists, and is nonnegative definite on account of eqn. 30. The inverse in eqn. 34 always exists, since, as noted above, $\hat{\mathbf{L}}(k)$ is nonsingular, and the matrix being inverted in eqn. 34 is precisely $\hat{\mathbf{L}}(k)\hat{\mathbf{L}}'(k)$. Accordingly, we have proved the following theorem.

Theorem 6: Let $R(k,\ell)$, defined in eqn. 22, be positive definite. Then the solution $\mathbf{P}_g(\cdot)$ of eqn. 34 exists, is nonnegative definite symmetric, with the matrix $\mathbf{L}(k) - \mathbf{H}'(k)\mathbf{P}_g(k)\mathbf{H}(k)$ positive definite for all k.

The result of this Section was first obtained by Moore and Colebatch.[12] Their technique for establishing the result rested heavily on a result on discrete-time quadratic minimisation problems,[13] and was quite different to the method of this Section.

5 FILTER COMPUTATION WITHOUT SIGNAL-PROCESS MODEL

In this Section, we shall proceed as follows:

(a) We shall obtain, using the results of Section 4, a linear finite-dimensional system, such that, when excited by white noise, the output covariance is $E[z(k)z'(\ell)]$. This system will be a member of the general class depicted in Fig. 1.

(b) Using the measurement process $z(\cdot)$, we shall obtain a state estimate for this special linear system. Because of its special properties, estimation is easy and achievable with zero error by feeding $\mathbf{z}(\cdot)$ into a linear finite-dimensional system whose output is $\mathbf{x}(k)$. Since the optimal filter is known to be unique, the second system must then, by the results of Section 3, be the optimal filter for any linear system generating the prescribed output statistics and possessing the required $\mathbf{F}(\cdot)$ [or $\Phi(\cdot,\cdot)$] and $\mathbf{H}(\cdot)$.

As in Section 3, we assume that we are given the quantities $\mathbf{F}(\cdot), \mathbf{H}(\cdot), \mathbf{K}(\cdot)$ and $\mathbf{R}(\cdot)$ associated with the output statistics.

Theorem 7: With quantities as defined earlier, and with $\mathbf{P}_g(\cdot)$ the matrix defined in theorem 6, a system with output covariance $\mathbf{R}(\cdot,\cdot)$ is provided by

$$\mathbf{x}_g(k+1) = \mathbf{F}\mathbf{x}_g(k) + \mathbf{F}[\mathbf{K} - \mathbf{P}_g\mathbf{H}][\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H}]^{-1/2}\mathbf{u}(k)$$
$$\mathbf{x}_g(k_0) = 0$$
$$\mathbf{z}(k) = \mathbf{H}'\mathbf{x}_g(k) + [\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H}]^{1/2}\mathbf{u}(k) \qquad (35)$$

with $\mathbf{u}(\cdot)$ zero mean, Gaussian and white with covariance $I\delta(k-\ell)$.

The proof of this theorem will be omitted, being a straightforward application of the analysis results of Section 2, and making use of eqn. 34 for $\mathbf{P}_g(\cdot)$. It turns out that $\mathbf{P}_g(k)$ is precisely $E[\mathbf{x}_g(k)\mathbf{x}'_g(k)]$.

Notice that the system defined by eqn. 35 and depicted in Fig. 3 has the requisite form of Fig. 1; it possesses the correct $\mathbf{F}(\cdot)$ and $\mathbf{H}(\cdot)$ matrices. The matrix $\mathbf{J}_1$ is $[\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H}']^{1/2}$, which exists and is nonsingular, because $\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H}$ is positive definite. The matrix $\mathbf{J}_2$ is zero, so that the noise $\mathbf{n}(\cdot)$ is entirely due to the input noise $\mathbf{u}(\cdot)$. Notice also that $\mathbf{J}_1\mathbf{J}_1 + \mathbf{J}_2\mathbf{J}_2'$ is nonsingular, as required.

Now, let us consider the estimation of $\mathbf{x}_g(k)$. This is trivial because the system given by eqn. 35 is invertible, and also the initial state $\mathbf{x}_g(k_0)$ of eqn. 35 is known. Consider the system

$$\hat{\mathbf{x}}(k+1/k) = \mathbf{F}[\mathbf{I} - (\mathbf{K} - \mathbf{P}_g\mathbf{H})(\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H})^{-1}\mathbf{H}']\hat{\mathbf{x}}(k/k-1)$$
$$+ \mathbf{F}(\mathbf{K} - \mathbf{P}_g\mathbf{H})(\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H})^{-1}\mathbf{z}(k)$$
$$\hat{\mathbf{x}}(k_0/k_0 - 1) = 0 \qquad (36)$$

As is evident, the input to this system is the measurement process $z(k)$, and $\hat{\mathbf{x}}(k+1/k)$ is a zero-error estimate of $\mathbf{x}_g(k+1)$. To see this, observe that eqns. 35 and 36 imply that

$$\hat{\mathbf{x}}(k+1/k) - \mathbf{x}_g(k+1) = \mathbf{F}[\mathbf{I} - (\mathbf{K} - \mathbf{P}_g\mathbf{H})(\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H})^{-1}\mathbf{H}'] \times [\hat{\mathbf{x}}(k/k-1) - \mathbf{x}_g(k)]$$

Since $\hat{\mathbf{x}}(k_0/k_0 - 1) = 0 = \mathbf{x}_g(k_0)$, it follows that $\hat{\mathbf{x}}(k/k-1) = \mathbf{x}_g(k)$ for all $k \geq k_0$, as required.

The linear system, eqn. 36, is depicted in Fig. 4 in several different formats. In Fig. 4, we show also the estimate $\hat{\mathbf{y}}(k/k-1)$ of $\mathbf{y}_g(k)$; the estimate is, of course, given by $\mathbf{H}'(k)\hat{\mathbf{x}}(k/k-1)$.

Taking into account the results of Section 3, explaining that the filter is independent of the details of the signal process, we can state the following theorem. This theorem explains how to obtain a Kalman filter from $\mathbf{F}(\cdot), \mathbf{H}(\cdot)$ and the output statistics only.

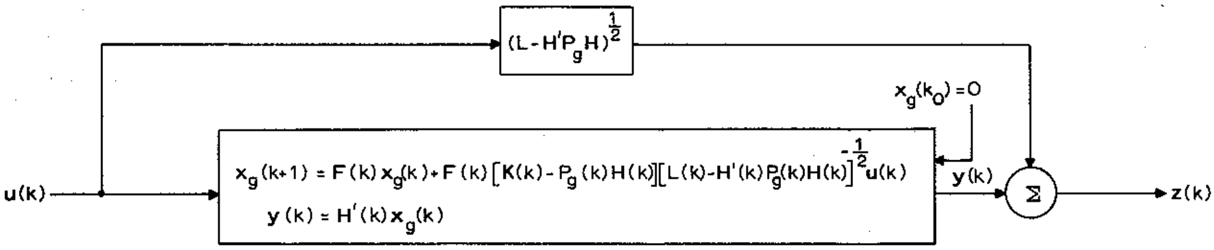

Fig. 3

System derived by spectral factorisation having prescribed output covariance

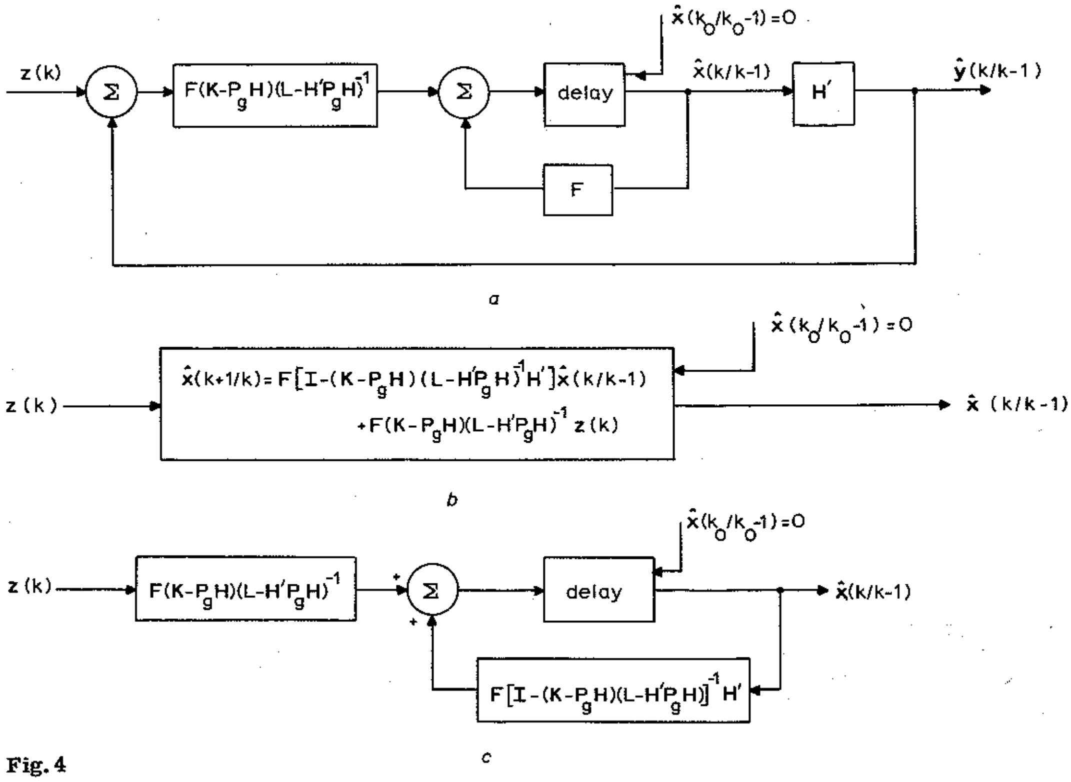

Fig. 4

Various arrangements of filter for system of Fig. 3

Theorem 8: Assume the same hypotheses as in theorem 2. Then, with $\mathbf{P}_g(k)$ defined as in eqn. 34, the optimal filter providing a minimum-error-variance estimate $\hat{\mathbf{x}}(k/k-1)$ is defined by eqn. 36.

We reiterate that the filter is constructed solely from the quantities $\mathbf{F}(\cdot), \mathbf{H}(\cdot), \mathbf{K}(\cdot)$ and $\mathbf{R}(\cdot)$. The filter also provides an estimate of the signal $\mathbf{y}(k) = \mathbf{H}'(k)\mathbf{x}(k)$. If we are interested in providing an estimate of the signal only, we can, of course, use any $\mathbf{F}(\cdot)$ and $\mathbf{H}(\cdot)$ compatible with the decomposition

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = \mathbf{H}'(k)\mathbf{\Phi}(k, \ell)\mathbf{K}(\ell) \qquad k \geq \ell \tag{19}$$

In particular, if we are given $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$ in the form $\mathbf{A}'(k)\mathbf{B}(\ell)$ for $k \geq \ell$ (see eqn. 16), we can take $\mathbf{H} \equiv \mathbf{A}, \mathbf{K} \equiv \mathbf{B}$ and $\mathbf{F} \equiv 0$.

A filter which produces $\hat{\mathbf{x}}(k/k)$, a minimal-error-variance estimate of $\mathbf{x}(k)$, given measurements $\mathbf{z}(\ell)$ for $\ell \leq k$, is easily obtained from that of Fig. 4 by use of the relation $\hat{\mathbf{x}}(k+1/k) = \mathbf{F}(k)\hat{\mathbf{x}}(k/k)$.

6 RAPPROCHEMENT WITH STANDARD FILTER DERIVATION

In this Section, our aim is to check that the filter described in theorem 8 is the same as that computed by standard procedures. For this purpose, we need to assume complete knowledge of the signal-process model, which we shall take here to be

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{u}(k) \\ \mathbf{z}(k) &= \mathbf{H}'\mathbf{x}(k) + \mathbf{R}^{1/2}\mathbf{v}(k) \end{aligned} \tag{37}$$

As before, $\mathbf{x}(k_0)$, $\mathbf{u}(\cdot)$ and $\mathbf{v}(\cdot)$ are Gaussian, of zero mean and are mutually independent. The covariances are

$$E[\mathbf{u}(k)\mathbf{u}'(\ell)] = \mathbf{I}\delta(k-\ell) \qquad E[\mathbf{v}(k)\mathbf{v}'(\ell)] = \mathbf{I}\delta(k-\ell)$$
$$E[\mathbf{x}(k_0)\mathbf{x}'(k_0)] = \mathbf{P}_0 \tag{38}$$

Notice that the measurement noise $\mathbf{R}^{1/2}(k)\mathbf{v}(k)$ here is independent of the input noise $\mathbf{u}(k)$.

It is well known that the Kalman filter is defined using the solution $\mathbf{P}_K(k)$ of

$$\mathbf{P}_K(k+1) = \mathbf{F}\mathbf{P}_K\mathbf{F}' - \mathbf{F}\mathbf{P}_K\mathbf{H}(\mathbf{H}'\mathbf{P}_K\mathbf{H} + \mathbf{R})^{-1}\mathbf{H}'\mathbf{P}_K\mathbf{F}' + \mathbf{G}\mathbf{G}'$$
$$\mathbf{P}_K(k_0) = \mathbf{P}_0 \tag{39}$$

and is

$$\hat{\mathbf{x}}(k+1/k) = \mathbf{F}[\mathbf{I} - \mathbf{P}_K\mathbf{H}(\mathbf{H}'\mathbf{P}_K\mathbf{H} + \mathbf{R})^{-1}\mathbf{H}']\hat{\mathbf{x}}(k/k-1)$$
$$+ \mathbf{F}\mathbf{P}_K\mathbf{H}(\mathbf{H}'\mathbf{P}_K\mathbf{H} + \mathbf{R})^{-1}\mathbf{z}(k) \qquad \hat{\mathbf{x}}(k_0/k_0-1) = 0 \tag{40}$$

Eqns. 36 and 40 will be the same if, and only if,

$$\mathbf{P}_K\mathbf{H}(\mathbf{H}'\mathbf{P}_K\mathbf{H} + \mathbf{R})^{-1} = (\mathbf{K} - \mathbf{P}_g\mathbf{H})(\mathbf{L} - \mathbf{H}'\mathbf{P}_g\mathbf{H})^{-1} \tag{41}$$

Recall now the analysis results of Section 2, which predict that

$$\mathbf{K} = \mathbf{P}\mathbf{H} \qquad \mathbf{L} = \mathbf{R} + \mathbf{H}'\mathbf{P}\mathbf{H} \tag{42}$$

where

$$\mathbf{P}(k+1) = \mathbf{F}\mathbf{P}(k)\mathbf{F}' + \mathbf{G}\mathbf{G}' \qquad \mathbf{P}(k_0) = \mathbf{P}_0 \tag{43}$$

Here, $\mathbf{P}(k)$ is the covariance $E[\mathbf{x}(k)\mathbf{x}'(k)]$ of the system given by eqn. 37. Then we can write eqn. 41 as

$$\mathbf{P}_K\mathbf{H}(\mathbf{H}'\mathbf{P}_K\mathbf{H} + \mathbf{R})^{-1} = (\mathbf{P} - \mathbf{P}_g)\mathbf{H}[\mathbf{H}'(\mathbf{P} - \mathbf{P}_g)\mathbf{H} + \mathbf{R}]^{-1} \tag{44}$$

The equality holds if it is true that

$$\mathbf{P} - \mathbf{P}_g = \mathbf{P}_K \tag{45}$$

This relationship is easily established from the defining equations for $\mathbf{P}$, $\mathbf{P}_g$ and $\mathbf{P}_K$ by showing that $\mathbf{P} - \mathbf{P}_g$ satisfies the same equation as $\mathbf{P}_K$, with the same initial condition.

7 STATIONARY PROBLEMS

Eqn. 45 also provides the key to what happens in filtering a stationary covariance. In this case, $|\lambda_i(\mathbf{F})| < 1$ and $\mathbf{P}(k)$ is independent of k. The matrix $\mathbf{P}_K(k)$, if initialised at some finite k_0 with the value $\mathbf{P}$, approaches a limiting value exponentially fast;[2] consequently, $\mathbf{P}_g$ must also approach a limiting value exponentially fast. Further, because the limiting values of $\mathbf{P}$ and $\mathbf{P}_K$ both satisfy steady-state versions of the equations for the transient $\mathbf{P}(k)$ and $\mathbf{P}_K(k)$, the limiting $\mathbf{P}_g$ is readily seen to satisfy a steady-state version of the transient equation for $\mathbf{P}_g(k)$.

8 ERROR VARIANCES

We showed in Section 3 that there exists an infinity of different signal-process models possessing the same optimal filter. In this Section, we consider the performance of the optimal filter. For state estimation, we find that the error variance depends in detail on the signal model, as well as on the output statistics. For signal estimation, however, the output statistics alone determine the error variance. This is the simpler situation, and so we consider it first.

Theorem 9: Assume the same hypotheses as for theorem 1. Then $E\{[\mathbf{y}(k) - \hat{\mathbf{y}}(k/k-1)][\mathbf{y}(k) - \hat{\mathbf{y}}(k/k-1)]'\}$ depends only on $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$, $E[\mathbf{y}(k)\mathbf{n}'(\ell)]$ and $E[\mathbf{n}(k)\mathbf{n}'(\ell)]$.

The proof of this theorem is standard, and will be omitted. Note that slightly more data are required to determine the performance of the filter than to determine the filter itself. Performance determination requires separate knowledge of $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$ and $E[\mathbf{y}(k)\mathbf{n}'(\ell)]$, while filter determination requires knowledge only of the sum of these quantities.

Let us now look at the error variance in state estimation. For convenience, let us assume that the noise $\mathbf{n}(\cdot)$ is uncorrelated with the input noise $\mathbf{u}(\cdot)$ in Fig. 1; i.e. $\mathbf{J}_1 \equiv 0$. Now, it is well known[1,2] that

$$E\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k/k-1)][\mathbf{x}(k) - \hat{\mathbf{x}}(k/k-1)]'\} = \mathbf{P}_K(k) \tag{46}$$

where $\mathbf{P}_K$ is as defined in the previous Section in eqn. 39. Also, as we know from eqn. 45, $\mathbf{P}_K = \mathbf{P} - \mathbf{P}_g$. Now, $\mathbf{P}_g(\cdot)$ depends only on the output statistics; $\mathbf{P}(\cdot)$, on the other hand, being $E[\mathbf{x}(k)\mathbf{x}'(k)]$, depends very much on the particular model. Accordingly, $\mathbf{P}_K(\cdot)$ also depends on the particular model.

It is interesting to note that

$$\begin{aligned} E\{[\mathbf{y}(k) - \hat{\mathbf{y}}(k/k-1)][\mathbf{y}(k) - \hat{\mathbf{y}}(k/k-1)]'\} \\ = \mathbf{H}'(k)E\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k/k-1)][\mathbf{x}(k) - \\ - \hat{\mathbf{x}}(k/k-1)]'\}\mathbf{H}(k) = \mathbf{H}'(k)\mathbf{P}_K(k)\mathbf{H}(k) \end{aligned} \tag{47}$$

Accordingly, even though $\mathbf{P}_K$ depends on the model, $\mathbf{H}'\mathbf{P}_K\mathbf{H}$ will only depend on the output statistics.

9 SMOOTHING PROBLEM

We consider now what data are required to obtain the optimal smoothed estimate of $\mathbf{y}(k)$, given measurements $\mathbf{z}(\ell)$, for $\ell = k_0, \ldots, k+N$, where N is a fixed but arbitrary positive integer. The problem of finding a smoothed estimate of the state $\mathbf{x}(k)$ of a signal-process model is also studied.

Theorem 10: Assume the same hypotheses as for theorem 1. Then the linear finite-dimensional system generating an optimal smoothed estimate of the signal process $\mathbf{y}(k)$ from $\mathbf{z}(\ell)$ $(\ell = k_0, \ldots, k+N)$ depends on $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$, $E[\mathbf{y}(k)\mathbf{n}'(\ell)]$ and $E[\mathbf{n}(k)\mathbf{n}'(\ell)]$, but is otherwise independent of the particular model generating $\mathbf{y}(\cdot)$.

Proof: The impulse response $\mathbf{A}_\mathbf{y}(k+N, \ell)$ of the optimal smoother is given by

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = \sum_{m=k_0}^{k-1+N} \mathbf{A}_\mathbf{y}(k+N, m)E[\mathbf{z}(m)\mathbf{z}'(\ell)]$$
$$k_0 \leq \ell < k+N \tag{48}$$

The positive definiteness of $E[\mathbf{z}(m)\mathbf{z}'(\ell)]$ guarantees the existence and uniqueness of a solution to this equation, assuming availability of $E[\mathbf{y}(k)\mathbf{z}'(\ell)]$ for the appropriate values of k and ℓ.

For $k > \ell$, we have

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = E[\mathbf{y}(k)\mathbf{y}'(\ell)] + E[\mathbf{y}(k)\mathbf{n}'(\ell)] \tag{49}$$

while, for $k \le \ell$, $E[\mathbf{y}(k)\mathbf{n}'(\ell)] = 0$, and so

$$E[\mathbf{y}(k)\mathbf{z}'(\ell)] = E[\mathbf{y}(k)\mathbf{y}'(\ell)] \tag{50}$$

In view of the symmetry of $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$, knowledge of this quantity for $k < \ell < k + N$ is equivalent to knowledge of this quantity for $\ell < k < \ell + N$, or simply $\ell < k$ if N is arbitrary. Therefore knowledge of the quantities on the left-hand side of eqns. 49 and 50 is equivalent to knowledge of $E[\mathbf{y}(k)\mathbf{y}'(\ell)]$ and $E[\mathbf{y}(k)\mathbf{n}'(\ell)]$ separately. This proves the theorem.

No such result holds, however, for smoothing of the state variable. The Wiener-Hopf equation applicable is

$$E[\mathbf{x}(k)\mathbf{z}'(\ell)] = \sum_{m=k_0}^{k-1+N} \mathbf{A_x}(k+N, m)E[\mathbf{z}(m)\mathbf{z}'(\ell)]$$

$$k_0 \le \ell < k + N \tag{51}$$

It is easily verified that, for $k < \ell$,

$$E[\mathbf{x}(k)\mathbf{z}'(\ell)] = [\mathbf{H}'(\ell)\Phi(\ell, k)\mathbf{P}(k)]' \tag{52}$$

It is evident from eqn. 52 that $\mathbf{P}$ itself, rather than, say, $\mathbf{P}(k)\mathbf{H}(k)$ or $\mathbf{PH} + \mathbf{F}^{-1}\mathbf{GJ}_1'$ must be known to compute $E[\mathbf{x}(k)\mathbf{z}'(\ell)]$, assuming that $\Phi(\cdot,\cdot)$ and $\mathbf{H}(\cdot)$ are known. As we know, $\mathbf{P}$ depends on the particular model of the signal process. Therefore we must know the detailed signal-process model to design the associated smoother.

10 NUMERICAL EXAMPLE

In this Section, we give a simple numerical example to show how the Kalman filter can be constructed using output statistics only. We shall check that this filter is the same as that computed by standard procedures. We also show that two different signal-process models generating the same output statistics possess the same optimal filter.

Consider a 1st-order system of the type depicted in Fig. 1. The system parameters are, with $E[u(k)u'(\ell)] = Q(k)\delta(k - \ell)$,

$$F(k) = 0{\cdot}8, G(k) = 1{\cdot}0, H(k) = 1{\cdot}0, J_1(k) = 0{\cdot}5$$

$$J_2(k) = 1{\cdot}0, Q(k) = 10 \text{ and } R(k) = 10$$

The stationary output covariance can be computed to be

$$c(\ell) = H'PH + J_1QJ_1' + J_2RJ_2' \qquad \ell = 0$$

$$c(\ell) = H'F^{\ell}(PH + F^{-1}GQJ_1') \qquad \ell > 0 \tag{53}$$

where $P = E[x(k)x'(k)]$ is obtained from $P - FPF' = GQG'$.

Note that this covariance does not allow deduction from it of G, assuming knowledge of F, H, and $J_1QJ_1' + J_2RJ_2'$. Nevertheless, the covariance allows the determination of the Kalman-filter gain, which can be defined as (see Fig. 4)

$$M = F(K - P_gH)(L - H'P_gH)^{-1} \tag{54}$$

Here, K and L are given by eqns. 9 and 10, respectively, and P_g is the steady-state value of $P_g(k)$, defined as in eqn. 34.

It is easily shown that $L = c(0) = 40{\cdot}2778$, $K = c(1)(HF)^{-1} = 34{\cdot}0278$ and $P_g = 18{\cdot}9500$, whence $M = 0{\cdot}565564$. The optimal filter is therefore

$$\hat{y}(k + 1/k) = 0{\cdot}234436\hat{x}(k/k - 1) + 0{\cdot}565564\, z(k)$$

$$\hat{y}(k_0/k_0 - 1) = 0 \tag{55}$$

Further, the covariance provides enough information to compute the performance of the filter viewed as a signal estimator. (Note that the error variance in estimating the state is the same in this particular case). Let P_y be the error variance in estimating the signal; then

$$P_y = H'PH - H'P_gH \tag{56}$$

where

$$H'PH = c(0) - (J_1QJ_1' + J_2RJ_2') \tag{57}$$

so that

$$P_y = 8{\cdot}82782$$

Of course, the Kalman gain and error performance could have been obtained from the original data, the relevant equations being

$$P_K = F\{P_K - (P_KH + F^{-1}GQJ_1')(H'P_KH + J_1QJ_1' + J_2RJ_2')^{-1}$$
$$(P_KH + F^{-1}GQJ_1')'\}F' + GQG' \tag{58}$$

$$\hat{x}(k + 1/k) = F\hat{x}(k/k - 1) + M[z(k) - H'\hat{x}(k/k - 1)]$$
$$\hat{x}(k_0/k_0 - 1) = 0 \tag{59}$$

$$M = (FP_KH + GQJ_1')(H'P_KH + J_1QJ_1' + J_2RJ_2')^{-1} \tag{60}$$

and

$$P_y = H'P_KH \tag{61}$$

Solution of eqn. 58 leads to

$$P_y = 8{\cdot}82782 \text{ and } M = 0{\cdot}565564$$

Of course, the filter constructed using the output covariance is the same as that computed using eqns. 58-60.

Let us now consider a second signal-process model, described by the following equations

$$x(k + 1) = 0{\cdot}8x(k) + 2{\cdot}5u(k) \tag{62}$$
$$z(k) = 0{\cdot}4x(k) + 0{\cdot}5u(k) + 1{\cdot}0v(k)$$

where $u(\cdot)$ and $v(\cdot)$ are independent zero-mean Gaussian white-noise processes with covariances $10{\cdot}0$.

It can easily be verified by simple calculation that the process $z(\cdot)$ of eqn. 62 has the same output covariance as the first system, with

$$c(0) = 40{\cdot}2778,\ c(1) = 27{\cdot}2222$$

so that the optimal filter is still given by eqn. 55, and the error variance is as before. Alternatively, one could use eqns. 58-61; in this case, eqn. 58 would contain different numbers than for the first case, although the same filter would be obtained.

11 CONCLUSION

Our main result has been to show that the Kalman filter, regarded as a state estimator, is designable using output statistics of the signal process alone, once a state-space co-ordinate basis has been fixed. We believe this result will bring about an important advance in adaptive filter design. Our other results have been connected with signal, as distinct from state, estimation, with the performance of optimal filters, and with the smoothing problem. We showed that, for signal estimation, output statistics alone determine the optimal filter. We showed also that the performance of the optimal filter, although depending only on the output statistics when used as a signal estimator, depends on the detailed signal-process model when used as a state estimator. In this connection, one can speculate as to the existence of signal-process models with independent input and output noise possessing the lowest possible associated error-variance matrix. Such are known in the continuous-time case,[14] but their properties can only be studied with difficult singular spectral-factorisation ideas.

It would also be of theoretical interest to study the extent to which the ideas of this paper would carry over to the situation of a singular measurement-noise covariance.

12 ACKNOWLEDGMENT

This work was supported by the Australian Research Grants Committee.

13 REFERENCES

1 KALMAN, R. E.: 'A new approach to linear filtering and prediction problems', Trans. ASME Ser. D, 82, pp. 35-45

2 JAZWINSKI, A. H.: 'Stochastic processes and filtering theory' (Academic Press, 1970)

3 BUCY, R. S., and JOSEPH, P. D.: 'Filtering for stochastic processes and applications to guidance' (Interscience, 1968)
4 KUO, B. C.: 'Analysis and synthesis of sampled-data control systems' (Prentice-Hall, 1963)
5 WIENER, N.: 'Extrapolation, interpolation and smoothing of stationary time series' (MIT Press, 1949)
6 BUCY, R. S.: 'Linear and nonlinear filtering', Proc. Inst. Elec. Electron. Eng., 1970, **58**, pp. 854-864
7 MEHRA, R. K.: 'On-line identification of linear dynamic systems with application to Kalman filtering', IEEE Trans., 1971, **AC-16**, pp. 12-21
8 ANDERSON, B. D. O., and MOORE, J. B.: 'State estimation via the whitening filter', Proceedings of the Joint Automatic Control Conference, Ann Arbor, Michigan, USA, June 1968, pp. 123-129
9 ANDERSON, B. D. O., and MOORE, J. B.: 'The Kalman-Bucy filter as a true time-varying Wiener filter', IEEE Trans., 1971, **SMC-1**, pp. 119-127
10 ANDERSON, B. D. O., MOORE, J. B., and LOO, S. G.: 'Spectral factorization of time-varying covariance functions', ibid., 1969, **IT-15**, pp. 550-557
11 GANTMACHER, F. R.: 'The theory of matrices' (Chelsea, New York, 1959)
12 MOORE, J. B., and COLEBATCH, P. M.: 'The simulation of nonstationary discrete stochastic processes', University of Newcastle, NSW, Australia, 1968, Technical report EE-6813
13 MOORE, J. B., and COLEBATCH, P. M.: 'Extensions of quadratic minimization theory: the discrete time case', Proc. IEE, 1970, **117**, (1), pp. 219-222
14 ANDERSON, B. D. O., and KAILATH, T.: 'The choice of signal-process models in Kalman-Bucy filtering', J. Math. Anal. & Appl., 1971, **35**, pp. 659-668

14 APPENDIX

Positive-definite nature of $\mathbf{R}(k, \ell) = E[\mathbf{z}(k)\mathbf{z}'(\ell)]$

We say that the covariance $\mathbf{R}(\cdot, \cdot)$ is positive definite over $[k_0, k_f]$ if

$$\sum_{k, \ell=k_0}^{k_f} \mathbf{m}'(k)\mathbf{R}(k, \ell)\mathbf{m}(\ell) > 0 \tag{63}$$

for all $\mathbf{m}(\cdot)$ for which $\mathbf{m}(k)$ is not identically zero. In this Appendix, we shall prove that, for the system of Fig. 1, the output covariance is positive definite, so long as the matrix $\mathbf{R}(k) = \mathbf{J}_1(k)\mathbf{J}_1'(k) + \mathbf{J}_2(k)\mathbf{J}_2'(k)$ is positive definite in the usual sense.

Since $\mathbf{u}(\cdot)$ and $\mathbf{v}(\cdot)$ are independent processes, we have

$$\mathbf{R}(k, \ell) = \mathbf{R}_1(k, \ell) + \mathbf{R}_2(k, \ell) \tag{64}$$

where $\mathbf{R}_1(\cdot, \cdot)$ is the covariance that would result at the output if $\mathbf{J}_2(k)$ were identically zero, and $\mathbf{R}_2(\cdot, \cdot)$ is the covariance that would result if $\mathbf{J}_1(k)$ and $\mathbf{H}(k)$ were identically zero.

Suppose $\mathbf{R}(k, \ell)$ is not positive definite; then there exists a sequence $\mathbf{m}(\cdot)$, such that

$$\sum_{k, \ell} \mathbf{m}'(k)\mathbf{R}_1(k, \ell)\mathbf{m}(\ell) = 0 \tag{65}$$

and

$$\sum_{k, \ell} \mathbf{m}'(k)\mathbf{R}_2(k, \ell)\mathbf{m}(\ell) = 0 \tag{66}$$

Since $\mathbf{R}_2(k, \ell) = \mathbf{J}_2(k)\mathbf{J}_2'(k)\delta(k-\ell)$ for all k and ℓ, it follows that, for all k,

$$\mathbf{m}'(k)\mathbf{J}_2(k) = 0 \tag{67}$$

Let $\mathbf{y}(\cdot, \cdot)$ denote the impulse response of the system of Fig. 1, with $\mathbf{J}_2(\cdot)$ identically zero. Evidently, $\mathbf{y}(k, k) = \mathbf{J}_1(k)$. Let $\mathcal{R}_1$ be the matrix

$$\mathcal{R}_1 = \begin{bmatrix} \mathbf{R}_1(k_0, k_0) & & \mathbf{R}_1(k_0, k_f) \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \mathbf{R}_1(k_f, k_0) & \cdots & \mathbf{R}_1(k_f, k_f) \end{bmatrix} \tag{68}$$

and let $\mathcal{Y}$ be the lower triangular matrix

$$\mathcal{Y} = \begin{bmatrix} \mathbf{y}(k_0, k_0) & & 0 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{y}_1(k_f, k_0) & \cdots & \mathbf{y}_1(k_f, k_f) \end{bmatrix} \tag{69}$$

Because of the way $\mathbf{y}(\cdot, \cdot)$ and $\mathbf{R}_1(\cdot, \cdot)$ are related, it is evident that

$$\mathcal{R}_1 = \mathcal{Y}\mathcal{Y}' \tag{70}$$

Eqns. 65, 68 and 70 establish that

$$[\mathbf{m}'(k_0) \cdots \mathbf{m}'(k_f)]\mathcal{Y} = 0 \tag{71}$$

or

$$\begin{gathered} \mathbf{m}'(k_f)\mathbf{J}_1(k_f) = 0 \\ \mathbf{m}'(k_f-1)\mathbf{J}_1(k_f-1) + \mathbf{m}'(k_f)\mathbf{y}(k_f, k_f-1) = 0 \\ \mathbf{m}'(k_f-2)\mathbf{J}_1(k_f-2) + \mathbf{m}'(k_f-1)\mathbf{y}(k_f-1, k_f-2) \\ + \mathbf{m}'(k_f)\mathbf{y}(k_f, k_f-2) = 0 \text{ etc.} \end{gathered} \tag{72}$$

Eqn. 67 for $k = k_f$ and the first of eqns. 72 imply that

$$\mathbf{m}'(k_f)[\mathbf{J}_1(k_f)\mathbf{J}_1'(k_f) + \mathbf{J}_2(k_f)\mathbf{J}_2'(k_f)] = 0$$
$$\mathbf{m}(k_f) = 0$$

because $\mathbf{R}(k)$ is assumed positive definite. Then, eqn. 67, for $k = k_f - 1$, the second of eqns. 72 and the fact that $\mathbf{m}(k_f) = 0$, imply that $\mathbf{m}(k_f - 1) = 0$. Continuing in this way, we establish that

$$\mathbf{m}(k) \equiv 0 \tag{73}$$

This is a contradiction. Therefore $\mathbf{R}(\cdot, \cdot)$ must be positive definite.