$$D = bT_p\phi^{p-1}. \tag{9}$$

In (9) let b be the unit matrix. Then (9) can be used to determine T_p, which is then used in (8) to generate T. Once T has been determined, then (6) is used to determine the a_j, $j = 1,\cdots,p$.

The Modified System

By taking advantage of the structure of θ and B, it is possible to represent (4) and (5) in another way. Let $z(k)$ be defined by

$$v(k) = \begin{bmatrix} z(k) \\ z(k-1) \\ \cdot \\ \cdot \\ \cdot \\ z(k-(p-1)) \end{bmatrix} \tag{10}$$

where $z(k)$ is an m vector, and therefore the process $v(k)$ is completely described by

$$z(k+1) = \sum_{j=0}^{p-1} a_{j+1}z(k-j) \tag{11}$$

$$y(k) = z(k) + w(k). \tag{12}$$

The problem of estimating the state $x(k)$ given $Y(l)$ has now been transformed to estimating the states $z(k-j)$, $j = 0,\cdots,p-1$, given $Y(l)$ of the process described by (11), where

$$\hat{x}(k/l) = T^{-1}\begin{bmatrix} \hat{z}(k/l) \\ \cdot \\ \cdot \\ \cdot \\ \hat{z}(k-(p-1)/l) \end{bmatrix}. \tag{13}$$

In using the estimators in [10] and [11] to obtain the estimates $\hat{z}(k-j/l)$, $j = 0,\cdots,p-1$, $V_z(i,j)$, $j = 0,\cdots,p-1$, $i = 0,\cdots,j$, will be necessary, where

$$V_z(i,j) = Ez(-i)z^T(-j).$$

$V_z(i,j)$ is the ijth $m \times m$ submatrix of V_v, where

$$V_v = TV_xT^T.$$

Example

For comparison purposes, consider a system described by (1), and let

$$\phi = \begin{bmatrix} 0.0 & 1.0 & 1.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & -1.0 \\ 1.0 & -1.0 & 1.0 & 1.0 \\ -4.05 & 0.9 & -5.2 & -2.8 \end{bmatrix}$$

$$D = [1.0 \quad 0.0 \quad 2.0 \quad 1.0].$$

The use of (9), (8), and (6) results in

$$a_j = -0.8, -0.5, -0.1, -0.75, \qquad j = 1,\cdots,4,$$

and therefore a fourth-order system has been transformed to an alternate state-space representation, which is a scalar system with three delay terms.

To illustrate the usefulness of the procedure, its on-line computational requirements are compared to those of the Kalman filter. The Kalman filter is applied to the system described by (1) to obtain estimates of the state $x(k)$ given $Y(l)$. Using (11) as the system description, the filtering technique developed in [10] along with (13) is applied to obtain the same estimates given the same measurements. For this example, the state estimation computing time is reduced by more than 90 percent of that required by the Kalman filter. In addition, 60 percent fewer storage locations are required by the filtering variables.

Conclusion

For optimum linear estimation of the state of a class of linear-discrete systems, a technique that requires less on-line computer time and memory than a Kalman filter has been developed. The procedure involves only one off-line computation, the determination of a linear transformation. It also involves the use of a linear filter for linear-discrete systems with time delays. A numerical example was included to demonstrate that computer time and memory requirements can be reduced. However, the larger m is, the less will be the computational advantage of this technique.

Although this procedure has been developed for time-invariant systems, an analogous technique for time-varying systems can readily be obtained.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME*, J. Basic Eng., ser. D, vol. 82, pp. 35–45, 1960.
[2] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA J.*, vol. 3, pp. 1445–1450, Aug. 1965.
[3] D. W. Heckman, "Dimension reduction of discrete Kalman filters on a class of systems having an orthogonality property," in *Proc. 2nd Asilomar Conf. Circuits and Systems*, 1968, pp. 305–308.
[4] C. F. Price, "An analysis of the divergence problem in the Kalman filter," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-13, pp. 699–702, Dec. 1968.
[5] C. H. Wells, "An approximation of the Kalman filter equations," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-13, pp. 445, Aug. 1968.
[6] C. G. Brammer, "Lower order optimal linear filtering of nonstationary random sequences," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-13, pp. 198–199, Apr. 1968.
[7] M. Aoki and J. R. Huddle, "Estimation of the state vector of a linear stochastic system with a constrained estimator," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-12, pp. 432–433, Aug. 1967.
[8] K. W. Simon and A. R. Stubberud, "Reduced order Kalman filter," *Int. J. Contr.*, vol. 11, pp. 501–509, Nov. 1969.
[9] R. A. Singer and R. G. Sea, "Increasing the computational efficiency of discrete Kalman filters," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-16, pp. 254–257, June 1971.
[10] R. Priemer and A. G. Vacroux, "Estimation in linear discrete systems with multiple time delays," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-14, pp. 384–387, Aug. 1969.
[11] ——, "On smoothing in linear discrete systems with time delays," *Int. J. Contr.*, vol. 13, pp. 299–303, 1971.
[12] M. Athans, "The relationship of alternate state-space representations in linear filtering problems," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-12, pp. 775–776, Dec. 1967.
[13] B. Ramaswami and K. Ramar, "On the transformation of time-variable systems to the phase-variable canonical form," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-14, pp. 417–419, Aug. 1969.

On Innovation Sequence Testing of the Kalman Filter

DRAYTON D. BOOZER AND WILLIE L. MC DANIEL, JR.

***Abstract*—Sequentially proven statements are given showing that the whiteness of the innovation sequence of a steady-state Kalman filter is not a sufficient condition for the optimality of the filter. Simulation results are given which verify each of the statements. Definite conclusions are reached concerning the identification of a class of systems by using the output sequence.**

I. Introduction

In a recent paper Mehra [1] has shown that a necessary and sufficient condition for the optimality of the Kalman filter is that the innovation sequence be white (see also [2], [3]). The assumption made in arriving at this result was that the ϕ matrix of the system was known and the variances of the state and observation noise

Manuscript received May 10, 1971; revised July 14, 1971.
D. D. Boozer is with Teledyne-Brown Engineering, Military Systems Division, Huntsville, Ala.
W. L. McDaniel, Jr., is with the Department of Electrical Engineering, Mississippi State University, State College, Miss.

sequences were to be identified. If these assumptions are changed (suppose, for example, that the ϕ matrix is not known), then one reaches quite a different result. The purpose of this correspondence is to show that the whiteness of the steady-state Kalman filter innovation sequence is not a sufficient condition for optimality when the ϕ matrix of the system is not known.

II. Statement of the Problem

Consider a second-order single-output linear discrete system described by

$$x(n+1) = \phi x(n) + w(n) \tag{1}$$

$$y(n) = Ax(n) + v(n) \tag{2}$$

where $w(n)$ and $v(n)$ are white Gaussian zero-mean noise sequences whose covariance matrices are

$$W = E[w(n)w^T(n)] = \begin{bmatrix} W_{11} & 0 \\ 0 & W_{22} \end{bmatrix} \tag{3}$$

$$Q = E[v^2(n)]. \tag{4}$$

The problem is to show that for this system the whiteness of the steady-state Kalman filter innovation sequence is not a sufficient condition for optimality when ϕ is not known.

III. Solution to the Problem

The solution to the problem posed is given in a sequence of statements in a format similar to [1]. The material in this section draws heavily from [2] and much of the close detail has been omitted.

Statement 1: Consider the system described in (1)–(4). If A, W, and the autocorrelation coefficients of $y(n)$ are known, the ϕ matrix of the system can be determined only to within one linear transformation, which is given by

$$P_s = \frac{1}{A_1^2W_{11} + A_2^2W_{22}} \begin{bmatrix} A_1^2W_{11} - A_2^2W_{22} & 2A_1A_2W_{11} \\ 2A_1A_2W_{22} & A_2^2W_{22} - A_1^2W_1 \end{bmatrix} \tag{5}$$

where the only two solutions, ϕ_1 and ϕ_2, are related by

$$\phi_1 = P_s^{-1}\phi_2 P_s. \tag{6}$$

Proof: The autocorrelation coefficients of $y(n)$ are defined by

$$C_k \triangleq E[y(n)y(n-k)] \tag{7}$$

and for the preceding system are

$$C_0 = ALA^T + Q \tag{8}$$

$$C_k = A\phi_1^k LA^T, \qquad k \geq 1 \tag{9}$$

where ϕ is set to ϕ_1 and L is the steady-state covariance matrix given by

$$L \triangleq E[x(n)x^T(n)] \tag{10}$$

and for the preceding system is

$$L = \phi_1 L \phi_1^T + W. \tag{11}$$

Substituting (6) into (11) yields

$$L^* = \phi_2 L^* \phi_2^T + P_s W P_s^T \tag{12}$$

where

$$L^* = P_s L P_s^T. \tag{13}$$

Since W is known, it is required that

$$W = P_s W P_s^T. \tag{14}$$

The autocorrelation coefficients of the starred system are

$$C_0^* = AL^*A^T + Q \tag{15}$$

$$C_k^* = A\phi_2^k L^* A^T, \qquad k \geq 1. \tag{16}$$

Now from (6) it follows that

$$P_s\phi_1^m = \phi_2^m P_s \tag{17}$$

for all m. Using (13) and (17) in (15) and (16) yields

$$C_0^* = AP_sLP_s^TA^T + Q \tag{18}$$

$$C_k^* = AP_s\phi_1^k LP_s^TA^T, \qquad k \geq 1. \tag{19}$$

Since A is known, it is required that

$$A = AP_s. \tag{20}$$

Therefore

$$C_k^* = C_k, \qquad k \geq 0. \tag{21}$$

The output autocorrelation coefficients are identical for both ϕ_1 and ϕ_2 of (6) if and only if (14) and (20) are satisfied. Solving these five nonlinear algebraic equations yields two solutions: that given by (5) and the trivial case $P_s = I$.

Statement 2: If a Kalman filter with any assumed ϕ matrix is connected to the system of Satement 1, the autocorrelation coefficients of the innovation sequence will be identical for system ϕ matrices related by the transformation P_s.

Proof: This assertion follows since Kailath [3] has shown that the Kalman filter innovation sequence contains the same information as the system output sequence, only in a less correlated form. The extension of Statement 1 to the innovation sequence is immediate.

Statement 3: The whiteness of the innovation sequence is a necessary, but not sufficient, condition for the optimality of the steady-state Kalman filter.

Proof: The necessary condition is shown by examination of the equations for the autocorrelation coefficients of the innovation sequence when the ϕ matrix is not known. Define

$$L \triangleq E[x(n)x^T(n)] \tag{22}$$

$$P \triangleq E[\hat{x}^n_{n+1} - x(n+1)][x^T(n+1)] \tag{23}$$

$$M \triangleq E[\hat{x}^n_{n+1} - x(n+1)][\hat{x}^n_{n+1} - x(n+1)]^T \tag{24}$$

$$\phi^* \triangleq \phi + \Delta\phi \tag{25}$$

$$\gamma \triangleq \phi^*[I - KA] \tag{26}$$

where ϕ^* is the Kalman filter ϕ matrix and K is the corresponding Kalman gain.

For the system of Statement 1 with a Kalman filter attached, these covariance matrices for the steady-state case are

$$L = \phi L \phi^T + W \tag{27}$$

$$P = \gamma P \phi^T + \Delta\phi L \phi^T - W \tag{28}$$

$$M = \gamma M \gamma^T + \phi^* KQK^T \phi^{*T} + W + \Delta\phi L(\Delta\phi)^T + \gamma P(\Delta\phi)^T + [\gamma P(\Delta\phi)^T]^T. \tag{29}$$

The autocorrelation coefficients for the innovation sequence $\nu(n)$ are defined by

$$C_k^f \triangleq E[\nu(n)\nu(n-k)] \tag{30}$$

and can be shown to be (see [2])

$$C_0^f = AMA^T + Q \tag{31}$$

$$C_k^f = A(\phi^*)^k MA^T + A[(\phi^*)^k - \phi^k]P^TA^T - A\sum_{p=0}^{k-1} (\phi^*)^{p+1} KC^f_{k-p-1}, \qquad k \geq 1. \tag{32}$$

TABLE I

Kalman Filter Error Signal Correlation Coefficients

$$\phi^* = \begin{bmatrix} 0.500 & 0.816 \\ -0.600 & 0.400 \end{bmatrix} \quad W = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\underline{A} = [1 \quad 0]$$

	EXAMPLE 1	EXAMPLE 2
ϕ	$\begin{bmatrix} 0.500 & 0.816 \\ -0.600 & 0.400 \end{bmatrix}$	$\begin{bmatrix} 0.500 & -0.816 \\ 0.600 & 0.400 \end{bmatrix}$
c_0^f	3.1780	3.1787
c_1^f	-0.0002	0.0004
c_2^f	-0.0001	0.0000
c_3^f	0.0000	0.0000
c_4^f	0.0000	0.0000
c_5^f	0.0000	0.0000

The necessary condition for the case in which ϕ is known and the noise variances are unknown was shown by Mehra [1]. The negative assertion follows from Statement 2 since the filter ϕ matrix can be equated to ϕ_1, one of the two possible ϕ matrices generating the output sequence. The innovation sequence will be white, since the system is optimal. However, Statement 1 provides another ϕ matrix ϕ_2, which yields the same output as ϕ_1. Therefore, from Statement 2 the innovation sequence must remain white.

IV. Simulation Results and Conclusion

In order to verify Statement 3, (27)–(32) were coded. If $A_2 = 0$, the transformation matrix of Statement 1 is

$$P_s = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{33}$$

This transformation changes the sign of the off-diagonal elements of the ϕ matrix. In Table I are listed the first six autocorrelation coefficients for the cases $\phi = \phi^*$ and $\phi = P_s^{-1}\phi^* P_s$. Notice that both cases yield white-noise sequences (small numbers were caused by the convergence criteria used in solving (27)–(29) iteratively). Only the system of Example 1 is optimal with

$$\text{tr}\,[M] = 3.6163$$

whereas

$$\text{tr}\,[M] = 9.3342$$

for Example 2.

The simulation results confirm the fact that the whiteness of the Kalman filter innovation sequence is not a sufficient condition for optimality, although a sufficient condition has been stated for the case in which the ϕ matrix is known and suboptimality is caused by incorrect variances.

References

[1] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering", *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 175–184, Apr. 1970.
[2] D. D. Boozer, "Linear stochastic system identification using correlation techniques", Ph.D. dissertation, Dep. Elec. Eng., Mississippi State Univ., State College, Aug. 1970.
[3] T. Kailath, "An innovations approach to least-squares estimation: Part I—Linear filtering in additive white noise," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 646–655, Dec. 1968.

New Linear Smoothing Formulas

BRIAN D. O. ANDERSON and S. CHIRARATTANANON

Abstract—**Formulas are given expressing the smoothed estimate of the state of a noisy linear system in terms of filtered estimates of the state, for both continuous and discrete time.**

We begin by defining the various quantities of interest. We are given the system

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)u(t) \tag{1}$$

$$y(t) = H'(t)x(t) + v(t) \tag{2}$$

with $u(\cdot)$ and $v(\cdot)$ zero-mean white Gaussian processes such that

$$E[u(t)u'(\tau)] = Q(t)\delta(t-\tau)$$

$$E[v(t)v'(\tau)] = R(t)\delta(t-\tau)$$

and with $R(t)$ nonsingular for all t. It is assumed that (1) and (2) apply for $t \geq t_0$, that $x(t_0)$ is Gaussian with mean $\bar{x}_0$ and covariance P_0, and that $x(t_0)$, $u(\cdot)$, and $v(\cdot)$ are mutually independent.

As is well known, the quantity $\hat{x}(t/t) = E\{x(t)/y(\tau), t_0 \leq \tau < t\}$ may be computed as follows. Define $P(t)$ by

$$\dot{P} = PF' + FP - PHR^{-1}H'P + GQG', \qquad P(t_0) = P_0 \tag{3}$$

and $K(t)$ by

$$K = PHR^{-1}. \tag{4}$$

Then

$$\begin{aligned}\dot{\hat{x}} &= F\hat{x} - K(H'\hat{x} - y)\\ &= (F - KH')\hat{x} + Ky, \qquad \hat{x}(t_0/t_0) = \bar{x}_0.\end{aligned} \tag{5}$$

Suppose that $b > t$. Then in [1, eq. (28)] it is shown that the quantity $\hat{x}(t/b) = E\{x(t)/y(\tau), t_0 \leq \tau < b\}$ is given by

$$\hat{x}(t/b) = \hat{x}(t/t) + P(t)\int_t^b \Phi_K'(\tau,t)H(\tau)R^{-1}(\tau)\nu(\tau)\,d\tau \tag{6}$$

where $\Phi_K(\cdot,\cdot)$ is the transition matrix associated with $F - KH'$, and the innovation process $\nu(t)$ is defined by

$$\nu(t) = y(t) - H'(t)\hat{x}(t/t). \tag{7}$$

Equations (6) and (7) imply that a smoothed estimate $\hat{x}(t/b)$ can be regarded as a linear functional of the measurements $y(\tau)$ for $t \leq \tau < b$ and the filtered estimate $\hat{x}(\tau/\tau)$ for $t \leq \tau < b$. The content of our main result is that $\hat{x}(t/b)$ can be regarded as a linear functional of only the filtered estimate $\hat{x}(\tau/\tau)$ for $t \leq \tau \leq b$.

Theorem: With all quantities as defined previously, suppose that $P(\tau)$ is nonsingular for $t \leq \tau \leq b$. Then the following formula expresses $\hat{x}(t/b)$ as a linear functional of $\hat{x}(\tau/\tau)$, $t \leq \tau \leq b$:

$$\begin{aligned}\hat{x}(t/b) &= P(t)\Phi_K'(b,t)P^{-1}(b)\hat{x}(b/b)\\ &\quad + P(t)\int_t^b \Phi_K'(\tau,t)P^{-1}(\tau)G(\tau)Q(\tau)G'(\tau)P^{-1}(\tau)\hat{x}(\tau/\tau)\,d\tau.\end{aligned} \tag{8}$$

Proof: From the filter equation (5) and the definition (7) of $\nu(\cdot)$ we see that

$$\dot{\hat{x}}(t/t) - F(t)\hat{x}(t/t) = K(t)\nu(t) = P(t)H(t)R^{-1}(t)\nu(t).$$

Manuscript received March 29, 1971; revised June 28, 1971. This work was supported by the Australian Research Grants Committee.
The authors are with the Department of Electrical Engineering, University of Newcastle, Newcastle, N.S.W. 2308, Australia.