KYBERNETIKA — VOLUME 14 (1978), NUMBER 2

Transfer-Function Solution of the Kalman-Bucy Filtering Problem

VLADIMÍR KUČERA

A novel, transfer-function solution of the Kalman-Bucy time-invariant filtering problem is presented. It is assumed that both message model and noise intensities are time invariant and that the mixture of message and noise has been observed over an infinite interval. This transfer-function approach is based on matrix fractions and spectral factorization of polynomial matrices. It offers an interesting comparison of state-variable and trasnfer-function methods, provides a deep insight into the problem discussed and, what is most important, it is computationally attractive.

INTRODUCTION

Recently, transfer-function methods have been successfully applied to solve problems in which state-variable approach used to dominate. This trend is motivated by the hope to provide a deeper insight into the problem and to obtain more efficient computational algorithms. Kalman-Bucy time-invariant filtering is just a typical problem of this kind.

The solution presented in this paper makes use of the classical notions of transfer-function matrices and spectral factorization. This mathematical machinery has been profitably used to solve the Wiener filtering problem, but it is not adequate for our purposes in its original form. The essential trick required to systematically treat unstable systems by means of transfer functions is to use the matrix fraction representation of rational matrices and algebraic minimization of inner products.

This paper is organized as follows. In the Formulation section, we begin with exact formulation of the Kalman-Bucy filtering problem to be studied here and then discuss its state-variable solution briefly. In the Solution section we proceed to transfer-function solution of the problem, the major contribution of the paper. In the Discussion section we tie the two methods together and illustrate the whole procedure on simple examples.

To begin, we shall give a precise formulation of the Kalman-Bucy filtering problem to be examined.

The message y is an m-vector random process modeled by the equations

$$\dot{x}(t) = F\,x(t) + G\,w(t)\,, \qquad y(t) = H\,x(t) \tag{1}$$

where x is an n-vector state and w is a p-vector excitation noise. It is natural to assume that system (1) is completely controllable and completely observable.

The observed mixture z of message y with an m-vector measurement noise v is modeled by the equation

$$z(t) = y(t) + v(t)\,. \tag{2}$$

The diagram of system (1), (2) is shown in Fig. 1.

We assume that w and v are uncorrelated white noise processes with zero mean and intensities Q and R, respectively. Matrices F, G, and H are constant of dimensions $n \times n$, $n \times p$, and $m \times n$ and matrices Q and R are constant symmetric positive definite of dimensions $p \times p$ and $m \times m$, respectively.

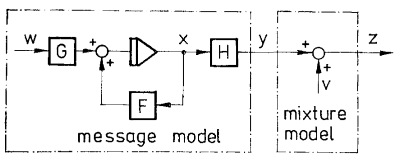

Fig. 1. Message and mixture models

Given the observed values of mixture z over the interval $(-\infty, t]$, our task is to find a linear estimate $\hat{y}(t)$ of message y at time t so as to minimize the expression

$$\mathrm{E}e'(t)\,e(t) \tag{3}$$

where $\mathrm{E}(\cdot)$ is the expected value, $e = y - \hat{y}$ is the filtering error, and the prime denotes transposition.

Compared to the original Kalman-Bucy formulation [1], we have made two additional significant assumptions: (i) both message model and noise intensities are time invariant and (ii) an arbitrarily long record of past measurements is available. The two assumptions guarantee that the optimal filter will be time invariant.

(4) **Remark.** Note that the message model (1) is not bound to be asymptotically stable. This means that the message (and hence the mixture) need not be a stationary random process but, instead, its covariance matrix may grow indefinitely. Due to this fact even our simplified steady-state formulation of the Kalman-Bucy filtering problem is more general than the classical problem solved by Wiener. Indeed, Wiener specified all random processes by their spectral-density matrices and hence a priori assumed all processes to be stationary. This is a serious limitation in many practical applications.

It is well known that our problem has a unique solution and that the Kalman-Bucy filter generating the best linear estimate $\hat{y}$ of y is governed by the equations

$$\dot{\hat{x}}(t) = (F - KH)\,\hat{x}(t) + K\,z(t)\,, \qquad (5)$$
$$\hat{y}(t) = H\,\hat{x}(t)\,.$$

The matrix K is given by

$$K =: PH'R^{-1} \qquad (6)$$

where P is the (unique) symmetric positive-definite solution of the matrix equation

$$FP + PF' - PH'R^{-1}HP + GQG' = 0\,. \qquad (7)$$

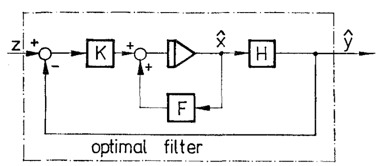

Fig. 2. Optimal filter

Note that the Kalman-Bucy filter is a feedback system obtained by taking a copy of the message model (omitting the input matrix G) as shown in Fig. 2. The matrix $F - KH$ has all eigenvalues with negative real parts and hence the filter is asymptotically stable.

In equations (5) the $\hat{x}(t)$ is an n-vector state of the filter. It is the best linear estimate of x, the state of the message model, at time t in the sense of minimizing the expression $\mathrm{E}e_x'(t)\,M\,e_x(t)$, in which $e_x = x - \hat{x}$ and M is an arbitrary symmetric positive-definite matrix. Therefore, the Kalman-Bucy filter can be used not only to separate random message from random noise but also to reconstruct the state of a system (or any linear combination of its state variables) from incomplete and noisy measurements.

The reader's attention is also drawn to the fact that the filter is optimal among linear systems only. It is apparent that we could obtain better results by nonlinear processing of the observations. On the other hand, if the noises v and w are gaussian, the Kalman-Bucy filter is optimal without any qualifications.

SOLUTION

The mathematics of the following derivations is based on real rational or polynomial matrices in complex variable s. For any rational matrix $R(s)$, let $R'(s)$, det $R(s)$ and tr $R(s)$ denote the transpose, determinant, and trace of $R(s)$, respectively. For the sake of simplicity, denote $R_*(s) =: R'(-s)$. A rational matrix $R(s)$ is said to be strictly proper if $R(\infty) = 0$.

In particular, if $P(s)$ is a polynomial matrix, we define its degree deg $P(s)$ as the highest degree among its polynomial entries and similarly $\deg_i P(s)$ for the i-th row of $P(s)$. Further denote P_H the matrix composed of the coefficients at highest powers of s in each row of $P(s)$ and call the $P(s)$ row reduced if P_H is nonsingular.

Any $m \times p$ rational matrix $R(s)$ can be written as the matrix fraction

$$R(s) = D^{-1}(s)\, N(s)$$

where $D(s)$ and $N(s)$ are left-coprime polynomial matrices of the dimensions $m \times m$ and $m \times p$, respectively, and the matrix $D(s)$ is row reduced. Then $R(s)$ is strictly proper if and only if $\deg_i N(s) < \deg_i D(s)$ for all $i = 1, 2, \ldots, m$. We remark that the condition $\deg N(s) < \deg D(s)$ is necessary but not sufficient for this purpose.

To simplify the notation we shall drop the argument s wherever convenient.

The transfer-function solution of the above specified Kalman-Bucy filtering problem can be obtained as follows. Let

$$S =: H(sI_n - F)^{-1}\, G \tag{8}$$

denote the transfer-function matrix of the message model (1) and write it in the form of the matrix fraction

$$S = A^{-1}B \tag{9}$$

where the polynomial matrices A and B of respective dimensions $m \times m$ and $m \times p$ are left coprime, A is row reduced, and $\deg_i B < \deg_i A$. Due to complete controllability and observability of (1)

$$n = \deg \det A = \sum_{i=1}^{m} \deg_i A\,.$$

The diagram of the filtering problem is shown in Fig. 3, in which W is the transfer-function matrix of the optimal filter to be found.

 The major result of the paper can be summarized in the following

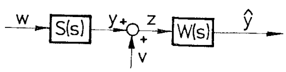

Fig. 3. Transfer-function diagram

(10) **Theorem.** The Kalman-Bucy filtering problem studied has a unique solution, which can be found as follows:

a) Calculate the real polynomial matrix C satisfying

(11) $$BQB_* + ARA_* = CRC_* ,$$

(12) $$C^{-1} \text{ analytic in } \operatorname{Re} s \geqq 0 ,$$

(13) $$C_{\mathrm{H}} = A_{\mathrm{H}} .$$

b) The transfer-function matrix of the optimal filter is then given as

(14) $$W = C^{-1}D$$

where $D =: C - A$.

(15) **Remark.** The procedure described in a) is called the spectral factorization [5]. The spectral factor C with its inverse C^{-1} analytic in $\operatorname{Re} s > 0$ always exists for any matrices A, B and Q, R provided the left-hand side of (11) is a full rank matrix. Analyticity on $\operatorname{Re} s = 0$ is then guaranteed by left coprimeness of A and B and by nonsingularity of Q and R. The spectral factor C is determined uniquely by (13).

Proof. To prove Theorem (10), rewrite expression (3) as

(16) $$\mathrm{E}e'(t)\, e(t) = \operatorname{tr} \mathrm{E}e(t)\, e'(t) = \frac{1}{2\pi \mathrm{j}} \operatorname{tr} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} \Phi_e(-s^2)\, \mathrm{d}s$$

for it is nothing else but the trace of the error covariance matrix. The spectral-density matrix Φ_e of the filtering error e can be written as

$$\Phi_e = WRW_* + (I_m - W)\, SQS_*(I_m - W)_*$$

using the definition of the error and the diagram shown in Fig. 3. It is clear, therefore, that integral (16) converges if and only if the two rational matrices W and $(I_m - W)\, S$ are both strictly proper and analytic in $\operatorname{Re} s \geqq 0$.

Rearranging,

$$\Phi_e = SQS_* - SQS_*W_* - WSQS_* + W\Phi_z W_*$$

where

$$\begin{aligned}\Phi_z &=: SQS_* + R \\ &= A^{-1}(BQB_* + ARA_*)\, A_*^{-1} \\ &= A^{-1}CRC_*A_*^{-1}\end{aligned}$$

The Φ_z can be interpreted as the spectral-density matrix of the mixture z, if it exists (i.e., if the message model is asymptotically stable). At any rate, we can write

$$\Phi_e = \Phi_0 + \Phi_y - \Phi_y\Phi_z^{-1}\Phi_y \tag{17}$$

where

$$\begin{aligned}\Phi_0 &=: (A^{-1}BQB_*C_*^{-1}R^{-1/2} - WA^{-1}CR^{1/2})\,. \\ &\quad .(A^{-1}BQB_*C_*^{-1}R^{-1/2} - WA^{-1}CR^{1/2})_*\,, \\ \Phi_y &=: SQS_*\,.\end{aligned}$$

Since the last two terms in (17) do not depend on W, it is sufficient to minimize (16) for Φ_0 instead of Φ_e. Rearranging, we obtain

$$\begin{aligned}A^{-1}BQB_*C_*^{-1}R^{-1/2} &= A^{-1}(CRC_* - ARA_*)\, C_*^{-1}R^{-1/2} \\ &= A^{-1}CR^{1/2} - A^{-1}AR^{1/2} + \\ &\quad + RC_*C_*^{-1}R^{-1/2} - RA_*C_*^{-1}R^{-1/2} \\ &= A^{-1}DR^{1/2} + RD_*C_*^{-1}R^{-1/2}\end{aligned}$$

where $D =: C - A$. Denoting

$$T =: A^{-1}DR^{1/2} - WA^{-1}CR^{1/2} \tag{18}$$

we finally arrive at

$$\Phi_0 = (RD_*C_*^{-1}R^{-1/2} + T)(RD_*C_*^{-1}R^{-1/2} + T)_*\,. \tag{19}$$

We infer from (13) that

$$\begin{aligned}\deg_i C &= \deg_i A\,, \\ \deg_i D &< \deg_i A\end{aligned} \tag{20}$$

and from convergence of integral (16) that W is a strictly proper rational matrix. Hence both rational matrices T and $R^{-1/2}C^{-1}DR$ are strictly proper. Further, convergence of integral (16) entails analyticity of $(I_m - W)S$ in $\operatorname{Re} s \geqq 0$ and hence the product

$$T\,.\,R^{-1/2}C^{-1}DR = A^{-1}(C - A)\,C^{-1}DR - WA^{-1}DR$$
$$= (I_m - W)\,A^{-1}DR - CDR$$

is analytic in Re $s \geqq 0$, too. The product being analytic in Re $s \geqq 0$ with both factors strictly proper, the following integral vanishes:

$$\frac{1}{2\pi j}\,\mathrm{tr}\int_{-j\infty}^{j\infty} TR^{-1/2}C^{-1}DR\,\mathrm{d}s = 0\,.$$

Thus (19) and the above combined give

$$\frac{1}{2\pi j}\,\mathrm{tr}\int_{-j\infty}^{j\infty} \Phi_0\,\mathrm{d}s = \frac{1}{2\pi j}\,\mathrm{tr}\int_{-j\infty}^{j\infty} (RD_*C_*^{-1}R^{-1}C^{-1}DR + TT_*)\,\mathrm{d}s\,.$$

The first term of the integrand on the right-hand side above is independent of W and hence the best we can do to minimize the integral is to set $T = 0$. The transfer-function matrix of the optimal filter then follows from (18)

$$W = A^{-1}DC^{-1}A$$
$$= A^{-1}(C - A)\,C^{-1}A$$
$$= I_m - C^{-1}A$$
$$= C^{-1}D$$

To complete the proof, it remains to check whether or not the two rational matrices

$$W = C^{-1}D\,,$$
$$(I_m - W)\,S = C^{-1}(C - D)\,A^{-1}B$$
$$= C^{-1}B$$

are strictly proper and analytic in Re $s \geqq 0$. The affirmative answer, however, can easily be deduced from (20) and (12). Existence and uniqueness of the optimal filter is a direct consequence of existence and uniqueness of the spectral factor C.

DISCUSSION

Two different approaches have been used to solve the steady-state Kalman-Bucy filtering problem and two seemingly different results have been obtained. The optimal filter produced by the state-variable method is realized as feedback system (5) around a copy of the message model. On the other hand, the transfer-function approach specifies the optimal filter by its transfer-function matrix (14).

It is instructive to make the relation between the two results explicit. The transfer-function matrix of the filter shown in Fig. 2 is

$$W = H(sI_n - F + KH)^{-1} K$$
$$= [I_m + H(sI_n - F)^{-1} K]^{-1} H(sI_n - F)^{-1} K$$

while (14) can be rearranged as follows

$$W = C^{-1} D$$
$$= (A^{-1}C)^{-1} A^{-1} D$$
$$= [A^{-1}(A + D)]^{-1} A^{-1} D$$
$$= (I_m + A^{-1}D)^{-1} A^{-1} D .$$

The identity

$$A^{-1} D = H(sI_n - F)^{-1} K$$

is then evident and it results in

(21) $$LK = D$$

where $L =: AH(sI_n - F)^{-1}$ is a polynomial matrix.

Thus if we are given the transfer-function matrix W of the optimal filter, its realization in terms of matrices H, F, and K can be obtained by determining the (unique) solution K of linear equation (21).

Of course, the optimal filter can be realized in any other way. But only the realization containing a copy of the message model has the additional property that its state is the best linear estimate of the state of the message model.

Two simple examples are included to illustrate the preceding discussions.

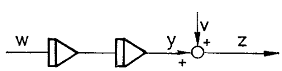

Fig. 4. Problem model

(22) **Example.** A particaı moves as a result of random force described as white noise with zero mean and unit intensity. The position of the particle was being observed for arbitrarily long period of time in the presence of additive white noise with zero mean and intensity 1/16. We are to find the best linear estimator of position.

A model of the problem is shown in Fig. 4, in which w is the acting force, y is the actual position, v is the measurement noise, and z is the observed position of the particle.

Clearly

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$H = [1 \;\; 0]$$

and

$$Q = [1], \quad R = [\tfrac{1}{16}].$$

The state-variable approach revolves around equation (7). Denoting $p_{ij} = p_{ji}$ the elements of matrix P, we are to solve the system of quadratic equations

$$\begin{aligned} 16p_{11}^2 \quad - 2p_{12} &= 0, \\ 16p_{11}p_{12} - p_{22} &= 0, \\ 16p_{12}^2 \quad - 1 &= 0. \end{aligned}$$

This system has the unique positive-definite solution

$$P = \frac{1}{8}\begin{bmatrix} \sqrt{2} & 2 \\ 2 & 4\sqrt{2} \end{bmatrix}.$$

Using (6) we calculate

$$K = \begin{bmatrix} 2\sqrt{2} \\ 4 \end{bmatrix}$$

and hence the optimal filter is realized by the matrices

$$F - KH = \begin{bmatrix} -2\sqrt{2} & 1 \\ -4 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 2\sqrt{2} \\ 4 \end{bmatrix},$$

$$H = [1 \;\; 0].$$

To apply the transfer-function approach, we first determine the transfer function (8) of the message model,

$$S = \frac{1}{s^2}.$$

Thus

$$A = s^2, \quad B = 1$$

in (9). Further we calculate the polynomial

$$BQB_* + ARA_* = 1 + \tfrac{1}{16}s^4$$

and its spectral factor satisfying (11) through (13)

$$C = s^2 + 2\sqrt{(2)}\, s + 4 .$$

Now

$$D = 2\sqrt{(2)}\, s + 4$$

and the transfer function (14) of the Kalman-Bucy filter becomes

$$W = \frac{2\sqrt{(2)}\, s + 4}{s^2 + 2\sqrt{(2)}\, s + 4} .$$

Given W, the feedback-gain matrix K needed to realize the filter by means of the message model can be found by solving equation (21):

$$[s \;\; 1]\, K = 2\sqrt{(2)}\, s + 4 .$$

(23) **Example.** Consider the situation described in Example (22) but now both the position and velocity of the particle can be observed in the presence of noise. The velocity measurement noise is a white random process with zero mean and unit intensity. What is the best linear estimate of the position and velocity?

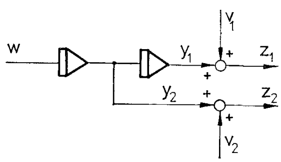

Fig. 5. Problem model

A model of the problem is shown in Fig. 5, in which w is the acting force, y_1 and y_2 is the actual position and velocity, v_1 and v_2 is the position and velocity measurement noise, and z_1 and z_2 is the observed position and velocity, respectively.

Clearly

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$Q = [1], \quad R = \begin{bmatrix} \frac{1}{16} & 0 \\ 0 & 1 \end{bmatrix}.$$

The state-variable solution is obtained by solving the set of quadratic equations (7)

$$16p_{11}^2 + p_{12}^2 - 2p_{12} = 0\,,$$
$$16p_{11}p_{12} + p_{12}p_{22} - p_{22} = 0\,,$$
$$16p_{12}^2 + p_{22}^2 - 1 = 0\,,$$

for elements $p_{ij} = p_{ji}$ of matrix P. The only positive-definite solution is

$$P = \frac{1}{20}\begin{bmatrix} 3 & 4 \\ 4 & 12 \end{bmatrix}$$

and hence

$$K = \frac{1}{5}\begin{bmatrix} 12 & 1 \\ 16 & 3 \end{bmatrix}$$

on using (6). The optimal filter is then realized by the matrices

$$F - KH = \frac{1}{5}\begin{bmatrix} -12 & 4 \\ -16 & -3 \end{bmatrix}, \quad K = \frac{1}{5}\begin{bmatrix} 12 & 1 \\ 16 & 3 \end{bmatrix},$$
$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

To demonstrate the transfer-function solution, we first write down the transfer-function matrix (8) of the message model,

$$S = \begin{bmatrix} \dfrac{1}{s^2} \\ \dfrac{1}{s} \end{bmatrix}$$

and its matrix fraction representation (9), e.g.

$$A = \begin{bmatrix} 0 & s \\ -s & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

We further calculate the matrix

$$BQB_* + ARA_* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}[1 \; 0] + \begin{bmatrix} 0 & s \\ -s & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{16} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & s \\ -s & 1 \end{bmatrix} =$$
$$= \begin{bmatrix} 1 - s^2 & s \\ -s & 1 - \frac{1}{16}s^2 \end{bmatrix}$$

and its spectral factor satisfying (11) through (13)

$$C = \frac{1}{5}\begin{bmatrix} 16 & 5s + 3 \\ -5s - 12 & 4 \end{bmatrix}.$$

A simple calculation yields

$$D = \frac{1}{5}\begin{bmatrix} 16 & 3 \\ -12 & -1 \end{bmatrix}$$

and the transfer-function matrix (14) of the optimal filter is found to be

$$W = \begin{bmatrix} 16 & 5s + 3 \\ -5s - 12 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 16 & 3 \\ -12 & -1 \end{bmatrix} =$$

$$= \frac{1}{5} \frac{\begin{bmatrix} 12s + 20 & s + 3 \\ 16s & 3s + 4 \end{bmatrix}}{s^2 + 3s + 4}$$

The matrix K needed to realize the filter from a copy of the message model can uniquely be determined from equation (21):

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} K = \begin{bmatrix} 16 & 3 \\ -12 & -1 \end{bmatrix}.$$

CONCLUDING REMARKS

Several authors have tried to solve the Kalman-Bucy filtering problem by means of transfer functions. However, all of them adhered to the traditional use of rational matrices. To overcome the difficulties arising from unstable message models they had to resort to generalizations of the classical notion of spectral factorization to that with prespecified poles [3]. This seems to be too complicated, computationally tedious and above all unnecessary. It proves much better to work with polynomial matrices, which can do the same job neatly and in full generality. Moreover, this approach is computationally attractive. Common methods for solving the matrix quadratic equation (7) take at least n^3 operations while iterative algorithms with linear convergence [2] to perform spectral factorization (11) through (13) do not exceed m^3n operations and one can hope for iterative algorithms featuring quadratic convergence [4] at about m^3n^2 operations. The merits of the transfer-function solution are especially pronounced when $m \ll n$. Compared to the classical Wiener solution, the need for partial fraction expansion of rational matrices is completely obviated.

(Received August 23, 1977.)

REFERENCES

[1] R. E. Kalman, R. S. Bucy: New results in linear filtering and prediction theory. Trans. ASME, J. Basic Eng. *D* 83 (1961), 95–105.

[2] N. N. Kazanjian: Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle. Ph. D. dissertation, Polytechnic Institute of New York, Farmingdale 1977.

[3] U. Shaked: A general transfer function approach to linear stationary filtering and steady-state optimal control problems. Int. J. Control *24* (1976), 6, 741–770.

[4] Z. Vostrý: New algorithm for polynomial spectral factorization with quadratic convergence. Kybernetika *12* (1976), 4, 248–259.

[5] D. C. Youla: On the factorization of rational matrices. IRE Trans. Information Theory *IT-7* (1961), 3, 172–189.

Ing. Vladimír Kučera, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.