Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99) Haifa, Israel - June 28-30, 1999

Kalman Bucy Filtering for Singular Output-Noise Covariance

F. Carravetta*
Ist. Anal. Sist. ed Inf. del CNR.
A. Germani†
Dip. Ing. Elettrica.
Università dell'Aquila

C. Manes‡
Dip. Ing. Elettrica.
Università dell'Aquila

Abstract

For a linear Gaussian stochastic system, the filtering problem is considered, when the covariance matrix of the observation noise is not invertible. A method that allows to build up the optimal filter in a number of cases is presented.

...

1 Introduction

Let us consider the linear stochastic system:

$$\begin{aligned} dX(t) &= AX(t)dt + FdW(t), \\ dY(t) &= CX(t)dt + GdW(t), \end{aligned} \quad 1.1) \tag{1}$$

where $W(t) \in \mathbb{R}^m$ is the standard Wiener process, $X(t) \in \mathbb{R}^n$, $Y(t) \in \mathbb{R}^q$, $m \geq q$. $X(0)$ is a Gaussian, zero mean random variable with covariance $\Psi_X(0)$, and $Y(0) = 0$.

The filtering problem for system (1.1), (1.2) was solved by Kalman and Bucy [1] in the case of a nonsingular output-noise covariance GG^T. By denoting $\widehat{X}$ the mean-square optimal state-estimate, the equations of the Kalman-Bucy filter are the following:

$$d\widehat{X}(t) = A\widehat{X}(t)dt + \Big(FG^T + P(t)C^T\Big)(GG^T)^{-1}(dY(t) - C\widehat{X}(t)dt), \tag{1.3}$$

$$\dot{P}(t) = AP(t) + P(t)A^T + FF^T - (FG^T + P(t)C^T)(GG^T)^{-1}(FG^T + P(t)C^T)^T,$$

endowed with the initial conditions:

$$\widehat{X}(0) = E\{X(0)\} = 0, \qquad P(0) = \Psi_X(0). \tag{1.4}$$

In this paper we consider the filtering problem for system (1.1), (1.2) when the covariance of observation noise, GG^T, is singular. In this case the classical Kalman-Bucy theory can not be used. Moreover, an expression of the optimal filter is up to now not available. In [2] an

*Email: *carravetta@iasi.rm.cnr.it*
†Email: *germani@ing.univaq.it*
‡Email: *manes@ing.univaq.it*

ϵ-optimal solution is given for this problem. More exactly, the output-noise covariance is replaced with $(GG^T + \epsilon I)$, and the convergence towards the optimum of the state-estimate given by the Kalman-Bucy filter, when ϵ goes to zero, is proved. However, this method is not very feasable in that it is ill-conditioned for small ϵ, so that any practical implementation would introduce large numerical errors that can not be controlled.

Nevertheless, in this paper we present a method to build up the optimal filter for singular GG^T in a number of cases. An algorithm is proposed, that gives, within a number of steps equal to the system state-dimension n, the answer about the question of whether the optimal filter can be defined using our method or not. In the case of a positive answer of the algorithm, the way to obtain the optimal filter is described.

2 Main result

Let us consider the sequence of matrices $\{\Pi_k,\ k \geq 0\}$ defined as follows:

$$\dim(\Pi_k) = \bar{q}_k \times \bar{q}_k, \quad \bar{q}_k = q - \sum_{i=0}^{k-1} q_i, \quad q_k = \mathrm{rank}(G_k) \leq \bar{q}_k, G_0 = G, \quad G_{k+1} = \Pi_k'' C_k F, C_{k+1} = \Pi_k'' C_k A, \tag{2.5}$$

Moreover, for $\bar{k} \geq 0$, let us define the matrix $\mathcal{M}_{\bar{k}}$:

$$\mathcal{M}_{\bar{k}} = \begin{bmatrix} \Pi_0' G \\ \Pi_1' G_1 \\ \vdots \\ \Pi_{\bar{k}}' G_{\bar{k}} \end{bmatrix}, \tag{2.6}$$

where it is understood that: *if for some i one has $\Pi_i' G_i = 0$, then this null block is removed.*

Theorem 2.1.

i) There exists a sequence of matrices $\{\Pi_k,\ k \geq 0\}$ that satisfies (2.5).

ii) If there exists a $\bar{k} \geq 0$ such that rank$(\mathcal{M}_{\bar{k}}) = q$, then $\bar{k} \leq n$ and there exists the optimal filter for system (1.1), (1.2).

Proof.

Let us consider the output equation (1.2), and rename $Y_0 = Y$, $C_0 = C$, $G_0 = G$. Let $\mathrm{rank}(G_0) = q_0 < q$, (remind that $\dim(G) = q \times p$). Then, there exists a matrix Π_0 such that:

$$\dim(\Pi_0) = q \times q, \quad \Pi_0 = \begin{bmatrix} \Pi_0' \\ \Pi_0'' \end{bmatrix}, \quad \mathrm{rank}(\Pi_0' G) = q_0, \quad \Pi_0'' G = 0. \tag{2.7}$$

Let us decompose $\Pi_0 Y_0$ as

$$\Pi_0 Y_0 = \begin{bmatrix} Y_0' \\ Y_0'' \end{bmatrix},$$

where $\dim(Y_0') = q_0$, and $\dim(Y_0'') = q - q_0 = \bar{q}_1$, where the definition of $\bar{q}_k$ given in (2.5) has been used. By exploiting the output equation (1.1), and the properties of Π_0 given by (2.7) one has:

$$dY_0' = \Pi_0' C_0 X dt + \Pi_0' G_0 dW, dY_0'' = \Pi_0'' C_0 X dt, \qquad 2.8$$

from which, in particular, it follows the existence of the time derivative: $\dot{Y}_0''$. Let us define

$$Y_1 \stackrel{\Delta}{=} \dot{Y}_0'' = \Pi_0'' C_0 X, \tag{2.9}$$

and let us calculate its stochastic differential:

$$dY_1 = d\dot{Y}_0'' = \Pi_0''C_0dX = \Pi_0''C_0AXdt + \Pi_0''CFdW = C_1Xdt + G_1dW, \tag{2.10}$$

where (1.1) has been used and:

$$C_1 \triangleq \Pi_0''CA, \quad G_1 \triangleq \Pi_0''CF.$$

with $\mathrm{rank}(G_1) = q_1 < \bar{q}_1$, (note that $\dim(G_1) = \bar{q}_1 \times p$).

By iterating the above described procedure, one has in general, at the k-th step:

$$dY_k' = \Pi_k'C_kXdt + \Pi_k'G_kdW, dY_k'' = \Pi_k''C_kXdt. \qquad \$$$

and

$$dY_{k+1} = C_{k+1}Xdt + G_{k+1}dW,$$

where

$$Y_{k+1} \triangleq \dot{Y}_k''$$

and

$$C_{k+1} = \Pi_k''C_kA, \quad G_{k+1} = \Pi_k''C_kF.$$

As it is immediately recognized, the above defined sequences of matrices: $\{\Pi_k\}, \{C_k\}, \{G_k\}$ agree with the ones defined by the relations (2.5). Therefore, the point i) of the theorem is proved.

Now, let us suppose that there exists a $\bar{k}$ such that $\mathrm{rank}(\mathcal{M}_{\bar{k}}) = q$. Then we can define a vector $\mathcal{Y} \in \mathbb{R}^q$:

$$\mathcal{Y} = \begin{bmatrix} Y' \\ Y_1' \\ \vdots \\ Y_{\bar{k}}' \end{bmatrix},$$

so that

$$d\mathcal{Y} = \mathcal{A}_{\bar{k}}Xdt + \mathcal{M}_{\bar{k}}dW, \tag{2.13}$$

where

$$\mathcal{A}_{\bar{k}} = \begin{bmatrix} \Pi_0' & 0 & \cdots & & 0 \\ 0 & \Pi_1'\Pi_0' & 0 & & \\ & & \ddots & & \vdots \\ & & & & \Pi_{\bar{k}-1}' \cdots \Pi_0' \end{bmatrix} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\bar{k}-1} \end{bmatrix}.$$

Let us denote $\sigma_t(\xi)$ the σ-algebra generated by a process $\{\xi(s),\ s \le t\}$ and $\sigma(X_1, ..., X_l)$ the σ-algebra generated by the random variables $X_1, ..., X_l$. Moreover, for any given algebra of Ω-subsets, namely $\mathcal{F}$, let us denote with $\sigma(\mathcal{F})$ the smallest σ-algebra containing $\mathcal{F}$. It results $\sigma(\sigma_t(\mathcal{Y}) \cup \sigma(Y(0), \dot{Y}(0), ..., Y^{(\bar{k})})) = \sigma_t(Y)$, for any t. Indeed, as it is easily recognized by the above described procedure, the map $Y \to \mathcal{Y}$ is invertible, provided that the random variables $Y(0), \dot{Y}(0) ..., Y^{(\bar{k})}(0)$ are given. Therefore there exists the optimal filter for the system (1.1), (1.2) and it results to be given by to the Kalman-Bucy filter applied to system (2.13). Moreover, noting that

$$\mathcal{M}_{\bar{k}} = \begin{bmatrix} \Pi_0' & 0 & \cdots & & 0 \\ 0 & \Pi_1'\Pi_0' & 0 & & \\ & & \ddots & & \vdots \\ & & & & \Pi_{\bar{k}-1}' \cdots \Pi_0' \end{bmatrix} \begin{bmatrix} G \\ CF \\ \vdots \\ CA^{\bar{k}-1}F \end{bmatrix}.$$

the request $\text{rank}(\mathcal{M}_{\bar{k}}) = q$ (that is $\mathcal{M}_{\bar{k}}$ is a full rank matrix) implies that the matrix $O_{\bar{k}} = [G^T\ (CF)^T\ \dots\ (CA^{\bar{k}-1}F)^T]^T$ has (column) rank equal to q. If $\bar{k} \geq n$, by the Cayley-Hamilton theorem it follows that $\text{rank}(O_n) = \text{rank}(O_{\bar{k}})$ that is, the rank no more increase for increasing $\bar{k}$. The proof of $ii)$ is then completed. •

Theorem 2.1 allows us to define the following algorithm, giving the number $\bar{k}$, and the matrices $\{\Pi_0, ..., \Pi_{\bar{k}}\}$. The algorithm is composed by the following steps:

1) let $k = 0$, $G_k = G$. 1a) If $\text{rank}(G_k) = \bar{q}_k$ and $k \leq n$ then $\bar{k} = k$, and go to step 2). Else compute Π_k such that $\text{rank}(\Pi_k' G_k) = q_k$, $\Pi_k'' G_k = 0$, compute $G_{k+1} = \Pi_k'' C_k F$, $C_{k+1} = \Pi_k'' C_k A$, increase k by one, and repeat 1a).

2) If $k > n$ then stop, the algorithm fails, else return $\bar{k}$ and $\{\Pi_0, ..., \Pi_{\bar{k}}\}$.

Theorem 2.1 assures that the algorithm terminates in a finite number of steps (less or equal to the state dimension n), giving the answer to the question of whether it is possible to build up the optimal filter (using this method) or not. In the case of a positive answer, the optimal state-estimate, $\hat{X}(t) = E(X/\sigma_t(Y))$, $t \geq 0$, of system (1.1), (1.2), can be obtained by an application of the Kalman Bucy filter to the equivalent system composed by the state equation (1.1) and the output equation (2.13). The filter should be inizialized with $\hat{X}(0) = E(X(0)/\sigma(Y(0), \dot{Y}(0), ..., Y^{(\bar{k})}(0)))$

The process $\mathcal{Y}$ can be practically obtained from process Y by using derivative devices ans following the procedure described in the proof of Theorem 2.1We stress that the output-derivatives involved in the definition of the equivalent observation $\mathcal{Y}$ are *noise-free*, and then well suited, in that they are derivable functions. This guarantees reasonable errors in the output of the derivative-device.

3 Concluding remark

The proposed filter is a continuous map $\mathcal{Y} \to \hat{X}$. This makes the filter practically implementable, provided that a good approximation is available for $\mathcal{Y}$, or in other words, a good approximation for the derivatives required by the construction procedure of $\mathcal{Y}$. This can be achieved for suitably smooth output trajectories. A comparison with the ϵ-method, mentioned in the Introduction, show that no ill-conditioning problems arise in this case (because of the continuity of the proposed filter).

It can be shown that when $\bar{k} = 1$ the filter can be implemented without approximations if the following white noise model is assumed for the output of system (1.1)

$$y(t) = CX(t) + Gn(t). \tag{3.14}$$

In this case the original Kalman-Bucy structure of the filter can be manipulated to obtain an implementation that does not need the explicit computation of the transformation $Y \to \mathcal{Y}$ of the output process.

It's well known (**??**) that.... **?**, Ch. 10), however, pointed out that...