IMPACT OF CORRELATION ERRORS ON THE OPTIMUM KALMAN FILTER GAIN IDENTIFICATION

Rafael Cardoso
CTA – Centro Técnico Aeroespacial
ITA – Instituto Tecnológico de Aeronáutica
Divisão de Engenharia Eletrônica
Departamento de Sistemas e Controle
Praça Mal. Eduardo Gomes, 50 – Vila das Acácias
12228-900 – São José dos Campos – SP, Brasil
cardoso@ieee.org

Elder Moreira Hemerly
CTA – Centro Técnico Aeroespacial
ITA – Instituto Tecnológico de Aeronáutica
Divisão de Engenharia Eletrônica
Departamento de Sistemas e Controle
Praça Mal. Eduardo Gomes, 50 – Vila das Acácias
12228-900 – São José dos Campos – SP, Brasil
hemerly@ele.ita.br

Abstract. *The impact of errors in the innovation correlation functions evaluation, related to the suboptimal filter, on the identification of the optimum steady state Kalman filter gains is investigated. An identification algorithm proposed in the literature, frequently quoted, is revisited and summarized. Then, as contributions, equations describing this impact are developed. Simulation results are presented to show this impact.*

Keywords. *Kalman filter, identification, estimation.*

1. Introduction

The Kalman filter has been widely used in many engineering applications. This is mainly due to its ability of deal with linear systems corrupted by uncertainties and provide an optimal state estimate, according to the minimum mean square error optimality criteria, in a recursive way. See Anderson et al (1979), Gelb et al (1996) and Maybeck (1979) for details. To accomplish this, the process noise covariance matrix Q and the measurement noise covariance matrix R must be known. In many practical situations, Q and R are unknown or known only approximately, thereby providing a suboptimal filter. Several authors have presented schemes for the identification of the unknown covariances or the steady state gain of the optimal filter. Among them, we have Mehra (1970) and Carew et al (1973), whom use the innovations covariance as the necessary information to identify the optimum steady state filter gain. Additionally, in Mehra (1970), we have an algorithm to identify the unknown noise covariance matrices. In this paper is presented an analysis of the impact that errors on the evaluation of the innovation correlation functions have on the identification of the steady state gain of the optimal filter, by using the approach proposed by Carew et al (1973), since it is frequently quoted and has a formal proof of convergence.

Some might claim that the linear system assumption, employed here, could be not necessary, since the extended Kalman filter could be used to deal with nonlinear systems, corrupted by uncertainties. According to Anderson et al (1979), the extended Kalman is a suboptimum filter and provides estimates that are only approximate to the optimal ones. Since we are interested on optimal estimates, the linear system assumption is necessary.

This paper is organized as follows: in section 2 the problem is presented together with the algorithm proposed by Carew et al (1973). In section 3, as contributions, equations related to the impact that errors on the evaluation of the innovations correlation functions have on the identification of the steady state gain of the optimal filter are developed. Finally, simulation results are presented in section 4.

2. Problem Formulation

Consider a discrete linear stationary multivariable stochastic system, as presented in Carew et al (1973),

$$x_{k+1} = Ax_k + Gw_k \tag{1}$$

$$y_k = Hx_k + v_k \tag{2}$$

$$\dim x_k = n,\ \dim A = n \times n,\ \dim w_k = p,\ \dim G = n \times p,\ \dim H = r \times n,\ \dim y_k = r,\ \dim v_k = r \tag{3}$$

The sequences w_k and v_k are independent stationary Gaussian white noise sequences with means and covariances

$$E\{w_i\} = 0; \quad E\{w_i w_j^T\} = Q\delta_{ij} \tag{4}$$

$$E\{v_i\} = 0; \quad E\{v_i v_j^T\} = R\delta_{ij} \tag{5}$$

$$E\{w_i v_j^T\} = 0, \quad E\{w_i x_j^T\} = 0, \quad E\{v_i x_j^T\} = 0, \quad \forall i, j \tag{6}$$

where $E\{\cdot\}$ denotes the expectation and δ_{ij} denotes the Kronecker delta function. Q and R are bounded positive definite matrices (Q>0 and R>0). Initial state x_0 is normally distributed with

$$E\{x_0\} = 0, \quad E\{x_i x_i^T\} = P_o \tag{7}$$

The system described by Eq. (1) and (2) is assumed to be stochastically controllable and observable. These conditions, together with the positive definiteness of Q and R ensure the asymptotic global stability of the Kalman filter. See Deyst et al (1968) for details. It is also assumed that the system is completely observable and controllable.

For this system, the innovations sequence $\{\upsilon_k\}$ and the estimation error $\tilde{x}_{k|k-1}$ are defined, respectively, by

$$\upsilon_k \stackrel{\Delta}{=} y_k - H\hat{x}_{k|k-1} \tag{8}$$

$$\tilde{x}_{k|k-1} \stackrel{\Delta}{=} x_k - \hat{x}_{k|k-1} \tag{9}$$

where $\hat{x}_{k|k-1}$ is the optimal linear estimate of x_k. When the filter is optimum, in steady state, the innovations sequence is a stationary white noise sequence with covariance given, according to Kailath (1968), by

$$W = H\Sigma H^T + R \tag{10}$$

where Σ is the optimal steady state error covariance, satisfying

$$\Sigma = A\Sigma A^T - A\Sigma H^T \left(H\Sigma H^T + R\right)^{-1} H\Sigma A^T + GQG^T \tag{11}$$

In steady state we also have

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + K\upsilon_k \tag{12}$$

$$K = A\Sigma H^T W^{-1} \tag{13}$$

where K is the optimum steady state filter gain.

Given data from a filter, Eq. (12), with suboptimal gain K_{sub}, which arises from incorrect Q and R covariance matrices, Carew et al (1973) proposed a recursive scheme to obtain the optimal filter gain K and the covariance W of the innovations associated to the optimal filter. The scheme is summarized in what follows.

Given Q_{sub} and R_{sub}, which are prior estimates of Q and R, by using Eq. (10), (11) and (13) we obtain the suboptimum steady state filter gain K_{sub}. Then, the suboptimal filter is given by

$$\hat{x}_{sub_{k+1|k}} = A\hat{x}_{sub_{k|k-1}} + K_{sub}\left(y_k - H\hat{x}_{sub_{k|k-1}}\right) \tag{14}$$

where $\hat{x}_{sub_{k|k-1}}$ is the suboptimal estimate of $\hat{x}_{k|k-1}$. The suboptimal filter steady state error covariance matrix is defined by

$$P_{sub} \stackrel{\Delta}{=} E\left\{\left(\hat{x}_{k+1|k} - \hat{x}_{sub_{k+1|k}}\right)\left(\hat{x}_{k+1|k} - \hat{x}_{sub_{k+1|k}}\right)^T\right\} \tag{15}$$

that, according Carew et al (1973), leads to

$$P_{sub} = (A - K_{sub}H)P_{sub}(A - K_{sub}H)^T + (K_{sub} - K)W(K_{sub} - K)^T \tag{16}$$

where W is given by Eq. (10).

The correlation functions of the innovations, in steady state, are

$$\begin{aligned} C_j &= E\{\upsilon_{sub_k}\upsilon^T_{sub_{k-j}}\} = E\{(y_k - H\hat{x}_{sub_{k|k-1}})(y_{k-j} - H\hat{x}_{sub_{k-j|k-j-1}})^T\} \\ &= \begin{cases} H(A - K_{sub}H)^{j-1}[(A - K_{sub}H)P_{sub}H^T - K_{sub}W + KW], & j \neq 0 \\ HP_{sub}H^T + W, & j = 0 \end{cases} \end{aligned} \tag{17}$$

Note that, if we have an optimal filter, then $K_{sub} = K$ and $P_{sub} = 0$. Hence, from Eq. (17), $C_j = 0$ for $j \neq 0$ and C_0 is given by Eq. (10), as expected.

In order to calculate K and W, the following iterative algorithm is proposed in Carew et al (1973),

$$W_m = C_0 - HP_{sub_m}H^T \tag{18}$$

$$K_m = (B^+D - AP_{sub_m}H^T)W_m^{-1} \tag{19}$$

$$P_{sub_{m+1}} = (A - K_{sub}H)P_{sub}(A - K_{sub}H)^T + (K_{sub} - K_m)W_m(K_{sub} - K_m)^T \tag{20}$$

In Eq. (19) B^+ is the pseudo inverse of B defined by

$$B^+ = (B^TB)^{-1}B^T \tag{21}$$

where

$$B^T = \left[H^T \vdots A^TH^T \vdots \cdots \vdots (A^T)^{n-1}H^T\right] \tag{22}$$

The matrix D is given by

$$D = \begin{bmatrix} C_1 + HK_{sub}C_0 \\ C_2 + HK_{sub}C_1 + HAK_{sub}C_0 \\ \vdots \\ C_n + HK_{sub}C_{n-1} + \cdots + HA^{n-1}K_{sub}C_0 \end{bmatrix} \tag{23}$$

and m is the iteration index.

The correlation functions C_j (j = 0, 1, 2, …, n) can be estimated experimentally, using the suboptimal innovations, by

$$C_j = \frac{1}{N}\sum_{i=1}^{N-j}\upsilon_{sub_{i+j}}\upsilon^T_{sub_i}, \quad j = (0,1,2,...,n) \tag{24}$$

where N is the number of observations. For finite N, the estimates given by Eq. (24) are biased but asymptotically they are unbiased and consistent. Carew et al (1973) show that estimates given by Eq. (19) converge to the optimum gain K, given by Eq. (13), if the correlation functions C_j are accurately known.

Now, we consider errors on the evaluation of the correlation functions given by Eq. (24) modeled by

$$C_j = C_{j_{nom}} + \Delta C_j, \quad j = (0,1,2,...,n) \tag{25}$$

where $C_{j_{nom}}$, $j=(0,1,2,...,n)$ is the nominal value of the correlation functions given by Eq. (17). From a practical point of view, a very important question is: What is the impact of the errors ΔC_j, $j=(0,1,2,..., n)$ on the identification of the optimum filter gain K? Section 3 deals with this issue.

3. Impact of Correlation Errors on the Optimum Filter Gain Identification

Suppose that $C_j = C_{j_{nom}} + \Delta C_j$, $j=(0,1,2,..., n)$ in Eq. (18) and (19). Consider that all inverses indicated exist. Due to errors ΔC_j, W_m, K_m and $P_{sub_{m+1}}$ in Eq. (18)-(20) do not converge anymore to the desired steady state values W, K and P_{sub}. Instead, they will converge to W_m^*, K_m^* and P_{m+1}^*. That is, errors ΔC_j will imply on errors in Eq. (18)-(20), modeled as $\Delta W_m, \Delta K_m$ and ΔP_{sub_m}, respectively.

Therefore, Eq. (18)-(20) will assume the form

$$W_m^* = W_m + \Delta W_m \tag{26}$$

$$K_m^* = K_m + \Delta K_m \tag{27}$$

$$P_{sub_{m+1}}^* = P_{sub_{m+1}} + \Delta P_{sub_{m+1}} \tag{28}$$

where * denotes that the variable contains errors due to ΔC_j.

By replacing Eq. (25) into Eq. (18), we get

$$W_m^* = C_{0_{nom}} + \Delta C_0 - H P_{sub_m}^* H^T \tag{29}$$

and by using Eq. (28) results

$$W_m^* = C_{0_{nom}} - H P_{sub_m} H^T + \Delta C_0 - H \Delta P_{sub_m} H^T \tag{30}$$

that can be written in the form of Eq. (26) with

$$W_m = C_{0_{nom}} - H P_{sub_m} H^T \tag{31}$$

and

$$\Delta W_m = \Delta C_0 - H \Delta P_{sub_m} H^T \tag{32}$$

In a similar way, considering Eq. (25) into Eq. (19) leads to

$$\begin{aligned} K_m^* = & \left[B^+ D - A P_{sub_m} H^T\right] W_m^{-1} - \left[B^+ D - A P_{sub_m} H^T\right] W_m^{-1} \left(W_m^{-1} + \Delta W_m^{-1}\right)^{-1} W_m^{-1} \\ & + \left[B^+ \Delta D - A \Delta P_{sub_m} H^T\right] W_m^{*\,-1} \end{aligned} \tag{33}$$

Equation (33) can also be written in the form of Eq. (27) with

$$K_m = \left[B^+ D - A P_{sub_m} H^T\right] W_m^{-1} \tag{34}$$

and

$$\Delta K_m = \left[B^+ \Delta D - A \Delta P_{sub_m} H^T\right] W_m^{*\,-1} - K_m \left(W_m^{-1} + \Delta W_m^{-1}\right)^{-1} W_m^{-1} \tag{35}$$

where

$$
D = \begin{bmatrix} C_{1_{nom}} + HK_{sub}C_{0_{nom}} \\ C_{2_{nom}} + HK_{sub}C_{1_{nom}} + HAK_{sub}C_{0_{nom}} \\ \vdots \\ C_{n_{nom}} + HK_{sub}C_{n-1_{nom}} + \cdots + HA^{n-1}K_{sub}C_{0_{nom}} \end{bmatrix} \tag{36}
$$

and

$$
\Delta D = \begin{bmatrix} \Delta C_1 + HK_{sub}\Delta C_0 \\ \Delta C_2 + HK_{sub}\Delta C_1 + HAK_{sub}\Delta C_0 \\ \vdots \\ \Delta C_n + HK_{sub}\Delta C_{n-1} + \cdots + HA^{n-1}K_{sub}\Delta C_0 \end{bmatrix} \tag{37}
$$

Now, considering Eq. (20) and Eq. (28), we can write

$$
P^*_{sub_{m+1}} = (A - K_{sub}H)P^*_{sub_m}(A - K_{sub}H)^T + (K_{sub} - K^*_m)W^*_m(K_{sub} - K^*_m)^T \tag{38}
$$

that leads to Eq. (28) with

$$
P_{sub_{m+1}} = (A - K_{sub}H)P_{sub_m}(A - K_{sub}H)^T + (K_{sub} - K_m)W_m(K_{sub} - K_m)^T \tag{39}
$$

and

$$
\begin{aligned} \Delta P_{sub_{m+1}} = {} & (A - K_{sub}H)\Delta P_{sub_m}(A - K_{sub}H)^T + (K_{sub} - K_m - \Delta K_m)\Delta W_m(K_{sub} - K_m - \Delta K_m)^T \\ & + (K_{sub} - K_m - \Delta K_m)W_m(K_{sub} - K_m - \Delta K_m)^T - (K_{sub} - K_m)W_m(K_{sub} - K_m)^T \end{aligned} \tag{40}
$$

Equations (31), (34) and (39), together, lead to the optimum filter gain K, since the correlations involved are given by Eq. (17). Equations (32), (35) and (40) will help us to analyze the identification error ΔK due to errors in Eq. (24) modeled by Eq. (25).

As proved by Carew et al (1973), W_m and K_m converges to W and K, respectively. Then, we can write, taking Eq. (26) and Eq. (35),

$$
\Delta K_m = \left[B^+\Delta D - A\Delta P_{sub_m}H^T\right](W + \Delta W_m)^{-1} - K\left(W^{-1} + \Delta W_m^{-1}\right)^{-1}W^{-1} \tag{41}
$$

and, from Eq. (40),

$$
\begin{aligned} \Delta P_{sub_{m+1}} = {} & (A - K_{sub}H)\Delta P_{sub_m}(A - K_{sub}H)^T + (\overline{K} - \Delta K_m)\Delta W_m(\overline{K} - \Delta K_m)^T \\ & + (\overline{K} - \Delta K_m)W(\overline{K} - \Delta K_m)^T - (\overline{K})W(\overline{K})^T \end{aligned} \tag{42}
$$

where

$$
\overline{K} = K_{sub} - K \tag{43}
$$

In Eq. (41) - (43), W is given by Eq. (10) and K is given by Eq. (13).

Therefore, given ΔC_j, $j = (0,1,2,...,n)$, by using Eq. (32), (41) and (42), in a recursive manner, we can evaluate the impact of the correlation errors on the identification of the optimum steady state filter gain.

4. Simulation Results

To evaluate the impact that covariance errors have on the identification of K, we will use an evader dynamic system, presented by Hong (1991), described by Eq. (1)-(2) with

$$
A = \begin{bmatrix} \cos(3^o) & -0.5\sin(3^o) \\ 2\sin(3^o) & \cos(3^o) \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{44}
$$

$$Q = \begin{bmatrix} 0.5^2 & 0 \\ 0 & 0.5^2 \end{bmatrix} m^2, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} m^2, \quad Q_{sub} = \begin{bmatrix} 0.25^2 & 0 \\ 0 & 0.25^2 \end{bmatrix} m^2, \quad R_{sub} = \begin{bmatrix} 1.5^2 & 0 \\ 0 & 1.5^2 \end{bmatrix} m^2 \tag{45}$$

$$x_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix} m \quad , \quad \hat{x}_{0|-1} = \hat{x}^*_{0|-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} m \tag{46}$$

$$P_{0|-1} = P^*_{0|-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} m^2 \quad , \quad \Delta W_0 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} m^2 \tag{47}$$

and only one sensor, in the central coordinate system, is considered. Note that the measurement model provides readings of the evader position on the x-y plane, corrupted by noise, and can be understood as a fictitious model of a radar. In this case, the radar, that generally provides range and azimuth measurements, provides directly the x-y values. In this system, one object moves on a two dimensional surface in a near elliptical course and the exact dynamics of the object is unknown. For more on radar measurement models, see Farina et al (1998) and Chen et al (2000).

Given the optimal statistics Q and R, the optimum steady state gain matrix, for reference, is

$$K = \begin{bmatrix} 0.3875 & 0.0075 \\ 0.0584 & 0.3949 \end{bmatrix} \tag{48}$$

The correlation functions considered ranged from 0 to 2. Their nominal values were evaluated analytically by means of Eq. (17). In Fig, 1 to 3, ΔC_j range from –50% to 100% of the nominal value $C_{j_{nom}}$, indicated in the horizontal axis of each figure, while the other ΔC_i, $i \neq j$ were kept null.

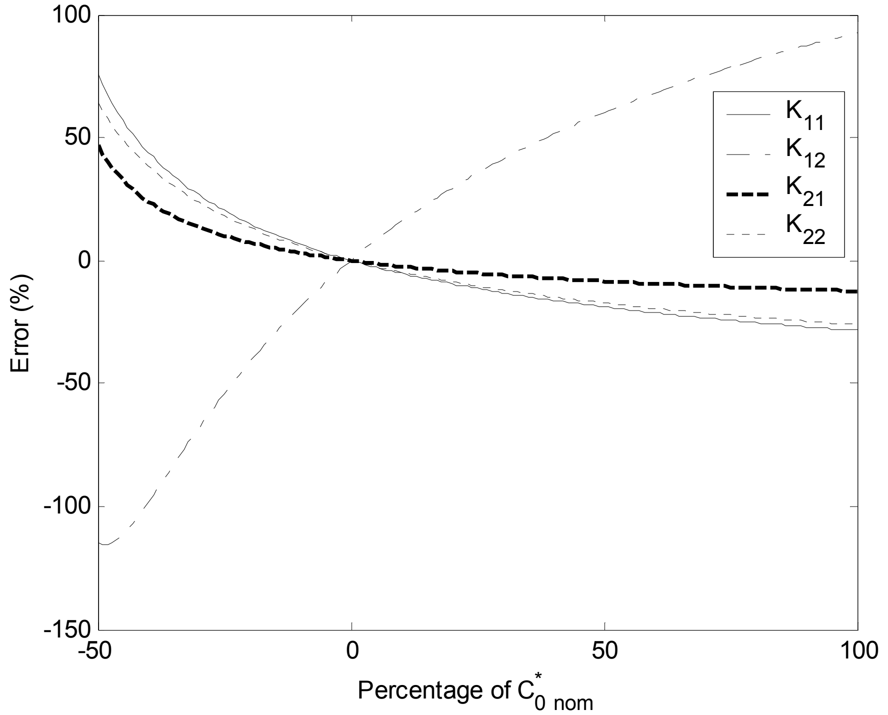

Figure 1. Gain identification errors due to error in C_0.

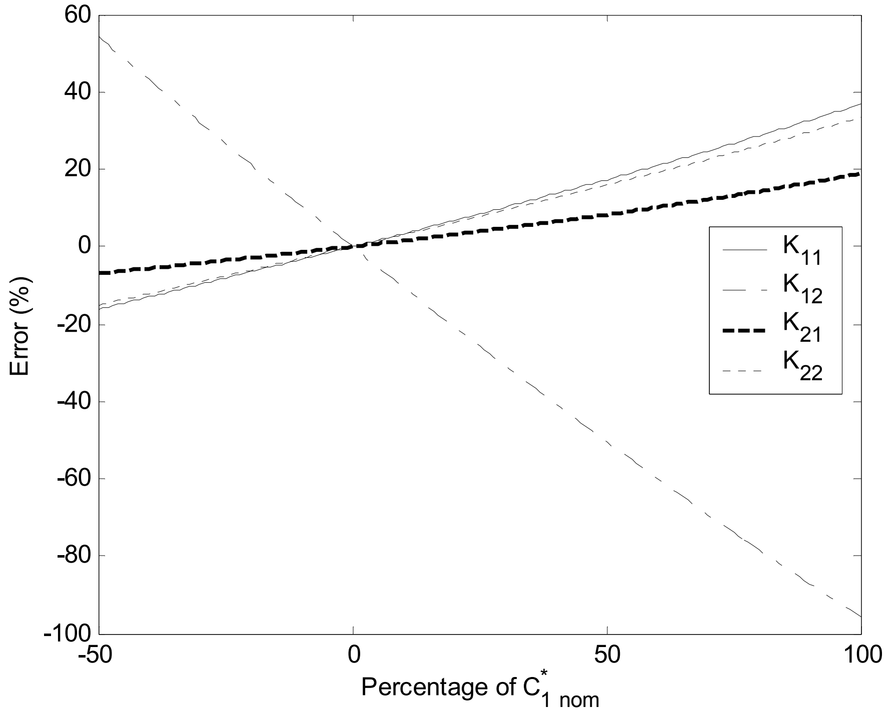

Figure 2. Gain identification errors due to error in C_1.

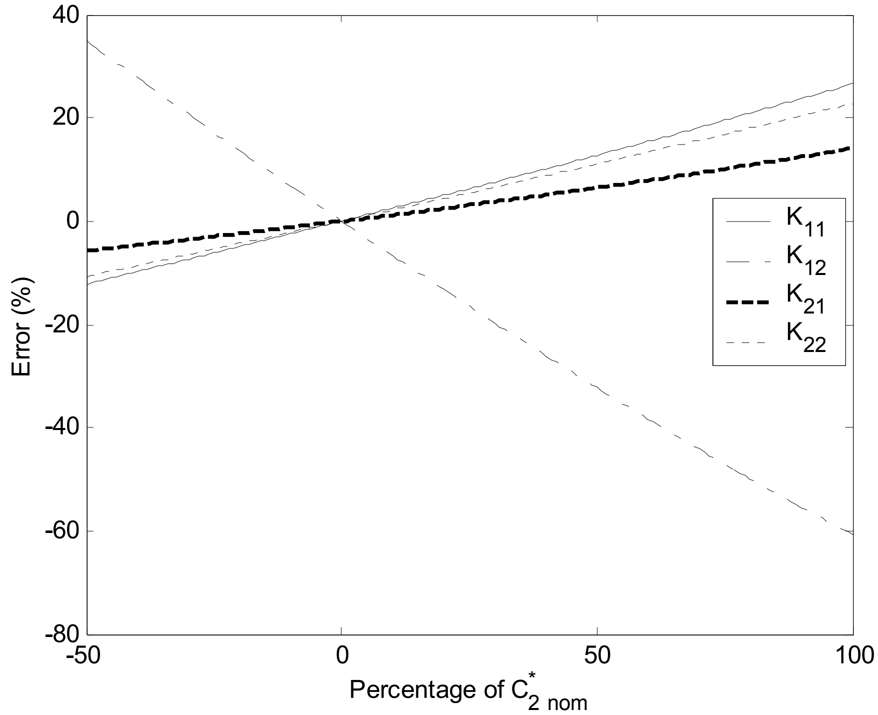

Figure 3. Gain identification errors due to error in C_2.

An analysis of Fig. 1 to 3 indicates that the identification errors can be considerable. Additionally, the impact on theses errors is less significant as the index j of the correlation function increases. This effect may be analyzed based in Eq. (37). In this equation, as the lag index j increases, the j-lag error in the correlation matrix associated, j>1, appears n-j+1 times, providing less influence on the identification errors. Besides, as the lag index j increases, the nominal value of the correlation matrices becomes smaller. Therefore, the errors considered in the simulation, given in percentage of C_j, are also smaller.

5. Conclusions

The impact that errors on the evaluation of the correlation functions have on the identification of the steady state Kalman filter gain was investigated in this paper. The main contribution is the development of equations that quantifies the identification errors based on the correlation errors. Based on the presented results, we can conclude that in real time applications and even in simulations with pseudo random signal generators, such as those in Matlab, it will be very difficult to estimate accurately the optimum steady state filter gain. This is mainly due to experimental errors on the estimates of $C_j,\ j=(0,1,2,...,\ n)$ that will differ from the theoretical ones given by Eq. (17). Hence, in order to obtain better results on the identification of the optimum filter gain, the correlations matrices must be evaluated, based on practical data, that is, the innovations of a suboptimal filter, as accurately as possible. Although the errors in the gain identification may have less significant impact in the state estimation errors, this is only so for the single sensor environment. For multisensor applications, errors in the gain identification can degrade significantly the state estimates, as will be reported elsewhere.

6. Acknowledgements

The authors would like to thank to the Fundação de Aplicação de Tecnologias Críticas-ATECH and to the Comissão de Coordenação do Projeto de Vigilância da Amazônia-CCSIVAM for the financial support provided by means of Termo Aditivo I (ITA-Fundação ATECH-CCSIVAM) during the realization of this work.

7. References

Anderson, B. D. O., Moore, J. B., 1979, "Optimal Filtering", Englewood Cliffs, NJ, Prentice-Hall, Inc., 357 p.

Carew, B., Bélanger, P. R., 1973, "Identification of Optimum Steady-State Gain for Systems with Unknown Noise Covariances", IEEE Transactions on Automatic Control, v. AC-18, n. 6, p. 582-587.

Chen, Y. M., Huang, H. C., 2000, "Fuzzy Logic Approach to Multisensor Data Association", Elsevier Mathematics and Computers in Simulation, no. 52, p. 399-412.

Deyst, J. J., Price, C. F., 1968, "Conditions for Asymptotic Stability of the Discrete Minimum-Variance Linear Estimator", IEEE Transactions on Automatic Control, v. 13, no. 6, p. 702-705.

Farina, A., Miglioli, R., 1998, "Association of Active and Passive Tracks for Airborne Sensors", Elsevier Signal Processing, no. 69, p. 209-217.

Gelb, A., Kasper Jr., J. F., Nash Jr., R. A., Price, Charles F., Sutherland Jr., Arthur A., 1996, "Applied Optimal Estimation", Cambridge, Mass., M.I.T. Press, 374 p.

Hong, L., 1991, "Adaptive Distributed Filtering in Multicoordinated Systems", IEEE Transactions on Aerospace and Electronic Systems, v. 27, n. 4, p. 715-724.

Kailath, T., 1968, "An Innovations Approach to Least-Squares Estimation, pt. I: Linear Filtering in Additive White Noise", IEEE Transactions on Automatic Control, v. AC-13, n. 6, p. 646-655.

Maybeck, P. S., 1979, "Stochastic Models, Estimation and Control. Volume 1", New York, NY, Academic Press, Inc., 423 p.

Mehra, R., 1970, "On the Identification of Variances and Adaptive Kalman Filtering", IEEE Transactions on Automatic Control, v. AC-15, n. 2, p. 175-184.