Kalman filtering in symmetrical noise environments

Roberto Diversi, Roberto Guidorzi, Umberto Soverini

Abstract— **This paper extends the stochastic context of Kalman filtering to the presence of additive noise on input observations. This extended filter is then used to solve the problem of optimal (minimal variance) estimation of noise–corrupted input and output sequences (errors–in–variables filtering). A Monte Carlo simulation shows that the performance of this extended filtering technique leads to the expected minimal variance estimates.**

Keywords— **Optimal filtering, errors–in–variables filtering, recursive filtering, Kalman filtering.**

I. Introduction

Kalman filtering can be considered as the standard approach for recovering information from data generated by known processes and corrupted by noise with known statistical properties. Among the reasons of its wide application it is worth mentioning the applicability to time-varying processes, the robustness in presence of modelling errors and the possibility of monitoring its performance and deducing, from the observation of its innovation, *a–priori* unknown properties of the noise [1]–[4].

One of the limits of Kalman filtering is the asymmetry of its description of disturbances because it assumes, as in most identification environments, an exact knowledge of the process input since the noise acts only on the state and output. This environment is realistic in all control applications where the input is generated by a known control law but can be restrictive in other applications.

A symmetrical environment is, on the contrary, described by Errors-in-Variables (EIV) models that consider all system attributes as affected by unknown additive and correlated disturbances. EIV models, in fact, do not require any system orientation, i.e. any partition of their attributes into inputs and outputs.

As discussed in [5], Kalman filtering cannot be directly applied to EIV contexts by simply assuming as input of the system its noisy observation and by balancing the noise introduced in this way with an opposite amount on the state, at least not for the purpose of obtaining an optimal reconstruction of the process input and output.

The solution of the optimal (minimal variance) EIV interpolation and filtering problems has been recently described in [5] and [6] in both behavioural and state–space contexts. A robust and high–efficiency reformulation of the original algorithm, based on the properties of Cholesky factorization, has then been described in [7]; this version outperforms all other formulations, including the state space one.

In a deterministic context, optimal EIV filtering could also be approached as an optimization problem along the lines followed by Roorda and Heij [8], as described in [5]. An approach of this kind has been recently followed in [9].

The authors are with the Department of Electronics, Computer Science and Systems, University of Bologna, Viale del Risorgimento 2, 40136, Bologna, Italy. E-mail: {rdiversi, rguidorzi, usoverini}@deis.unibo.it

It is also possible to define a new extended filtering problem that includes as subcases both EIV filtering (optimal estimate of inputs and outputs from noisy observations) and traditional Kalman filtering (optimal estimate of output and state in presence of state and output noise). This new problem is solved in this paper by extending Kalman filtering to this more general case.

The paper is organized as follows. Section 2 contains a definition of the considered filtering problem while Section 3 develops the modified Kalman filter used for its solution. The expected performance of the filter is investigated in section 4. The Monte Carlo simulation reported in Section 5 shows the excellent agreement between expected and observed performance. Short concluding remarks are finally given in Section 6.

II. Statement of the problem

The models considered in this paper are described by the state–space relations

$$\begin{aligned} x(t+1) &= A\,x(t) + B\,\hat{u}(t) + w(t) \quad x(0) = x_0 && (1)\\ \hat{y}(t) &= C\,x(t) + D\,\hat{u}(t) && (2)\\ y(t) &= \hat{y}(t) + \tilde{y}(t) && (3)\\ u(t) &= \hat{u}(t) + \tilde{u}(t) && (4) \end{aligned}$$

where $x(t) \in \mathcal{R}^n$ is the state of the process and $\hat{y}(t) \in \mathcal{R}^m$, $\hat{u}(t) \in \mathcal{R}^r$ denote the unknown noiseless components of the observations; $w(t)$ is the state noise while $\tilde{y}(t)$ and $\tilde{u}(t)$ are the additive noise on $\hat{y}(t)$ and $\hat{u}(t)$.

Assumptions – We will assume, in the sequel, that $w(t)$, $\tilde{y}(t)$ and $\tilde{u}(t)$ are mutually uncorrelated zero mean white processes, uncorrelated with $\hat{u}(t)$ and with covariances

$$\begin{aligned} E\left[w(t)\,w^T(t-\tau)\right] &= \Sigma_w\,\delta(\tau) && (5)\\ E\left[\tilde{y}(t)\,\tilde{y}^T(t-\tau)\right] &= \tilde{\Sigma}_y\,\delta(\tau) && (6)\\ E\left[\tilde{u}(t)\,\tilde{u}^T(t-\tau)\right] &= \tilde{\Sigma}_u\,\delta(\tau), && (7) \end{aligned}$$

where $\delta(\tau)$ is the Kronecker delta function. The initial state x_0 is a random vector with mean $\bar{x}_0$ and covariance matrix P_0, uncorrelated with $w(t)$, $\tilde{y}(t)$ and $\tilde{u}(t)$, $\forall t$.

The problem under investigation can be defined as follows.

Problem 1 (Optimal filtering) – Given the model (1)–(4), the covariance matrices (5)–(7) and an increasing sequence of input–ouput observations $\{u(0), y(0), \ldots, u(t), y(t)\}$, determine the optimal (minimal variance) estimate of $\hat{u}(t), \hat{y}(t)$.

Relations (3) and (4) allow to write model (1)–(2) in the form

$$\begin{aligned} x(t+1) &= A\,x(t) + B\,u(t) - B\,\tilde{u}(t) + w(t) && (8)\\ y(t) &= C\,x(t) + D\,u(t) - D\,\tilde{u}(t) + \tilde{y}(t). && (9) \end{aligned}$$

By introducing the auxiliary processes

$$\begin{aligned} n_x(t) &= w(t) - B\,\tilde{u}(t) && (10)\\ n_y(t) &= \tilde{y}(t) - D\,\tilde{u}(t) && (11)\\ z(t) &= y(t) - D\,u(t), && (12)\end{aligned}$$

it is possible to rewrite model (8)–(9) in the classical Kalman filter form

$$\begin{aligned} x(t+1) &= A\,x(t) + B\,u(t) + n_x(t) && (13)\\ z(t) &= C\,x(t) + n_y(t), && (14)\end{aligned}$$

that will be used in the sequel for solving the filtering problem.

III. Optimal filtering

The optimal estimate $x(t|t)$ of the state $x(t)$ is given by the conditional expectation

$$x(t|t) = E\,[x(t)|z(t), z(t-1), \ldots, z(0)], \qquad (15)$$

which leads, as well–known, to the relations

$$\begin{aligned} x(t+1|t) &= \big(A - S\,R^{-1}\,C\big)\,x(t|t) + B\,u(t) + S\,R^{-1}\,z(t)\\ x(0|-1) &= \bar{x}_0 && (16)\\ x(t|t) &= x(t|t-1) + K(t)\big(z(t) - C\,x(t|t-1)\big) && (17)\\ K(t) &= P(t|t-1)\,C^T\big(CP(t|t-1)\,C^T + R\big)^{-1} && (18)\\ P(t|t) &= \big(I_n - K(t)\,C\big)\,P(t|t-1) && (19)\\ P(t+1|t) &= \big(A - S\,R^{-1}\,C\big)P(t|t)\big(A - S\,R^{-1}\,C\big)^T && (20)\\ &+ \;\; Q - S\,R^{-1}\,S^T\,, \quad P(0|-1) = P_0 \end{aligned}$$

where

$$\begin{aligned} Q &= E\left[n_x(t)\,n_x^T(t)\right] = \Sigma_w + B\,\tilde{\Sigma}_u\,B^T && (21)\\ R &= E\left[n_y(t)\,n_y^T(t)\right] = \tilde{\Sigma}_y + D\,\tilde{\Sigma}_u\,D^T && (22)\\ S &= E\left[n_x(t)\,n_y^T(t)\right] = B\,\tilde{\Sigma}_u\,D^T, && (23)\end{aligned}$$

under the assumption of nonsingularity for R.

Remark 1 – If $w(t)$, $\tilde{y}(t)$, $\tilde{u}(t)$ and x_0 are gaussian, $x(t|t)$ is the estimate that minimizes the mean square error

$$E\left[\big(x(t) - x(t|t)\big)^T\,\big(x(t) - x(t|t)\big)\right], \qquad (24)$$

otherwise $x(t|t)$ constitutes the best linear estimate (in the mean square error sense) that can be obtained from the observations.

Now, we focus our attention on the problem of obtaining optimal estimates $\hat{y}(t|t)$, $\hat{u}(t|t)$ of $\hat{y}(t)$ and $\hat{u}(t)$ starting from the knowledge of $x(t|t)$. Since

$$\begin{aligned} \hat{u}(t|t) &= u(t) - \tilde{u}(t|t) && (25)\\ \hat{y}(t|t) &= y(t) - \tilde{y}(t|t), && (26)\end{aligned}$$

these estimates can be obtained from the optimal estimates of the input and output noise.

From equation (8) it follows that

$$x(t+1|t) = A\,x(t|t) + B\,u(t) - B\,\tilde{u}(t|t) + w(t|t). \qquad (27)$$

Since

$$w(t|t) = E\,[w(t)|z(t), z(t-1), \ldots, z(0)] = 0, \qquad (28)$$

by comparing (27) with (16) we obtain

$$-S\,R^{-1}\,C\,x(t|t) + S\,R^{-1}\,z(t) = -B\,\tilde{u}(t|t), \qquad (29)$$

i.e, thanks to (23),

$$\tilde{u}(t|t) = \tilde{\Sigma}_u\,D^T\,R^{-1}\,\big(C\,x(t|t) - z(t)\big). \qquad (30)$$

Similarly, equation (9) leads to

$$z(t) = C\,x(t|t) + \tilde{y}(t|t) - D\,\tilde{u}(t|t), \qquad (31)$$

and by substituting $\tilde{u}(t|t)$ with expression (30) it is immediate to obtain

$$\tilde{y}(t|t) = \big(I_m - D\,\tilde{\Sigma}_u\,D^T\,R^{-1}\big)\,\big(z(t) - C\,x(t|t)\big). \qquad (32)$$

Finally, by using (25) and (26), the minimal variance estimates of $\hat{y}(t)$, $\hat{u}(t)$ can be written in the form

$$\begin{aligned} \hat{y}(t|t) &= y(t) - \big(I_m - D\,\tilde{\Sigma}_u\,D^T\,R^{-1}\big)\,\big(z(t) - C\,x(t|t)\big) && (33)\\ \hat{u}(t|t) &= u(t) - \tilde{\Sigma}_u\,D^T\,R^{-1}\,\big(C\,x(t|t) - z(t)\big). && (34)\end{aligned}$$

Remark 2 – When $D = 0$, the optimal estimate of the noise–free output $\hat{y}(t)$ is given by $C\,x(t|t)$, as in standard Kalman filtering, while the optimal estimate of the noise–free input $\hat{u}(t)$ coincides with its observation $u(t)$.

IV. Evaluation of the expected performance

The purpose of this section is to develop an expression for the expected performance of the filter (33)–(34), i.e. for the covariance matrices of the estimate errors

$$\begin{aligned} e_y(t) &= \hat{y}(t) - \hat{y}(t|t) && (35)\\ &= -\tilde{y}(t) + \big(I_m - D\,\tilde{\Sigma}_u\,D^T\,R^{-1}\big)\,\big(z(t) - C\,x(t|t)\big)\\ e_u(t) &= \hat{u}(t) - \hat{u}(t|t) && (36)\\ &= -\tilde{u}(t) + \tilde{\Sigma}_u\,D^T\,R^{-1}\,\big(C\,x(t|t) - z(t)\big). \end{aligned}$$

Since

$$z(t) - C\,x(t|t) = n_y(t) + C\,e_x(t), \qquad (37)$$

where

$$e_x(t) = x(t) - x(t|t), \qquad (38)$$

$e_y(t)$ and $e_u(t)$ can be expressed in the more convenient form

$$\begin{aligned} e_y(t) &= -\tilde{y}(t) + \big(I_m - D\,\tilde{\Sigma}_u\,D^T\,R^{-1}\big)\,\big(n_y(t) + C\,e_x(t)\big) && (39)\\ e_u(t) &= -\tilde{u}(t) - \tilde{\Sigma}_u\,D^T\,R^{-1}\,\big(n_y(t) + C\,e_x(t)\big). && (40)\end{aligned}$$

The terms

$$\begin{aligned} P_y(t) &= E\left[e_y(t)\,e_y^T(t)\right] && (41)\\ P_u(t) &= E\left[e_u(t)\,e_u^T(t)\right], && (42)\end{aligned}$$

will be now computed.

Since $\tilde{y}(t)$, $\tilde{u}(t)$ and $n_y(t)$ are white processes, it is possible to show, thanks to (17) and (38), that

$$\begin{aligned} E\left[\tilde{y}(t)\, e_x^T(t)\right] &= -E\left[\tilde{y}(t)\, z^T(t)\right] K^T(t) \qquad (43)\\ &= -\tilde{\Sigma}_y\, K^T(t) \\ E\left[\tilde{u}(t)\, e_x^T(t)\right] &= -E\left[\tilde{u}(t)\, z^T(t)\right] K^T(t) \qquad (44)\\ &= \tilde{\Sigma}_u\, D^T\, K^T(t), \end{aligned}$$

and hence

$$E\left[\tilde{y}(t)\left(n_y(t)+C\, e_x(t)\right)^T\right] = \tilde{\Sigma}_y\left(I_m - K^T(t)\, C^T\right) \quad (45)$$

$$E\left[\tilde{u}(t)\left(n_y(t)+C\, e_x(t)\right)^T\right] = -\tilde{\Sigma}_u\, D^T\left(I_m - K^T(t)\, C^T\right) \quad (46)$$

$$E\Big[\left(\tilde{n}_y(t)+C\, e_x(t)\right)\left(\tilde{n}_y(t)+C\, e_x(t)\right)^T\Big] = R - C\,K(t)\,R - R\,K^T(t)\,C^T + C\,P(t|t)\,C^T \quad (47)$$

From (45)–(47) we obtain

$$\begin{aligned} &P_y(t) = \qquad (48)\\ &\tilde{\Sigma}_y - \tilde{\Sigma}_y\, H^T - H\,\tilde{\Sigma}_y + \left(\tilde{\Sigma}_y - H\,R\right) K^T(t) C^T H^T \\ &+ HCK(t)\left(\tilde{\Sigma}_y - R\,H^T\right) + H\left(R + C\,P(t|t)\,C^T\right)H^T, \end{aligned}$$

where

$$H = \left(I_m - D\,\tilde{\Sigma}_u\, D^T\, R^{-1}\right). \quad (49)$$

By recalling expression (22) of R, it is immediate to verify that $HR = \tilde{\Sigma}_y$, so that

$$\begin{aligned} P_y(t) &= \tilde{\Sigma}_y - \tilde{\Sigma}_y\, H^T - H\,\tilde{\Sigma}_y \qquad (50)\\ &+ H\left(R + C\,P(t|t)\,C^T\right)H^T. \end{aligned}$$

Similar considerations lead to the expression of the covariance matrix of $e_u(t)$

$$\begin{aligned} P_u(t) &= \tilde{\Sigma}_u \qquad (51)\\ &- \tilde{\Sigma}_u D^T R^{-1}\left(I_m - C\,P(t|t)\,C^T R^{-1}\right) D\tilde{\Sigma}_u. \end{aligned}$$

Remark 3 – When the pair $(A - SR^{-1}C, C)$ is detectable and the pair $(A - SR^{-1}C, \bar{Q})$ is stabilizable, with $\bar{Q}\,\bar{Q}^T = Q - SR^{-1}S^T$, $P(t+1|t)$ converges, for $t \to \infty$, to the unique solution $\bar{P}$ of the algebraic Riccati equation

$$\begin{aligned} P \;=\; &\left(A - S\,R^{-1}\,C\right)\Big[P - P\,C^T\left(C\,P\,C^T + R\right)^{-1} C\,P\Big] \\ &\times \left(A - S\,R^{-1}\,C\right)^T + Q - S\,R^{-1}\,S^T. \qquad (52) \end{aligned}$$

Moreover, the filter (16),(17) is asymptotically stable for $t \to \infty$ [1]. In this case

$$\lim_{t\to\infty} P_y(t) = \bar{P}_y \quad (53)$$

$$\lim_{t\to\infty} P_u(t) = \bar{P}_u. \quad (54)$$

V. NUMERICAL RESULTS

The results obtained in previous sections have been numerically verified by means of a 100 runs Monte Carlo simulation performed on a one–input two–outputs model described by the matrices

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -0.3 & 0.4 & -0.2 \\ -0.1 & 0.2 & 0.4 \end{bmatrix} \qquad B = \begin{bmatrix} 0.8 \\ 0.17 \\ 1.09 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 1.7 \\ 0.51 \end{bmatrix}.$$

The number of samples is 500. The input sequence $\hat{u}(\cdot)$ has unit variance and is described in Fig. 1 (continuous line). In every run, the state, output and input noise sequences are characterized by the following covariance matrices

$$\Sigma_w = \begin{bmatrix} 0.14 & 0.09 & 0.16 \\ 0.09 & 0.08 & 0.11 \\ 0.16 & 0.11 & 0.2 \end{bmatrix}$$

$$\tilde{\Sigma}_y = \begin{bmatrix} 0.94 & 1.27 \\ 1.27 & 1.82 \end{bmatrix}$$

$$\tilde{\Sigma}_u = 0.25.$$

The percent amounts of noise, defined as one hundred times the ratios between the standard deviation of the noise and those of the noise–free signals, are equal to 50% for the input and to about 41% and 70% for the outputs. In every run, the initial state x_0 is a random vector and equations (16) and (20) have been initialized with $x(0|-1) = 0$ and $P(0|-1) = I_n$.

Figures 1–3 report the noiseless input and outputs (continuous line) and the associated noisy observations (dotted line) in a typical case of the Monte Carlo simulation (last 200 samples).

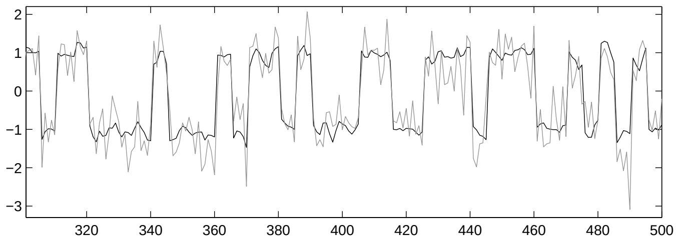

Fig. 1. Comparison between the noiseless input and its observation.

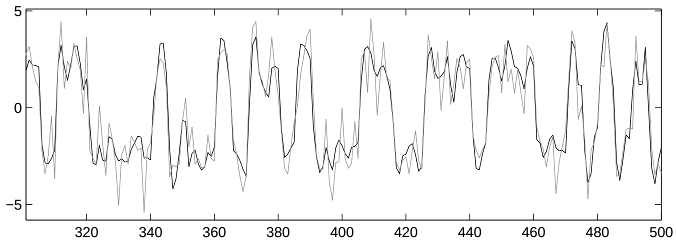

Fig. 2. Comparison between the first noiseless output and its observation.

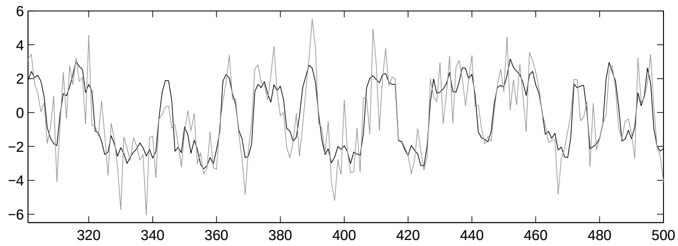

Fig. 3. Comparison between the second noiseless output and its observation.

The effectiveness of the filter can be observed, in the same typical case, in Figures 4–6, where the noiseless input and outputs (continuous line) are compared with the corresponding filtered quantities (dotted line).

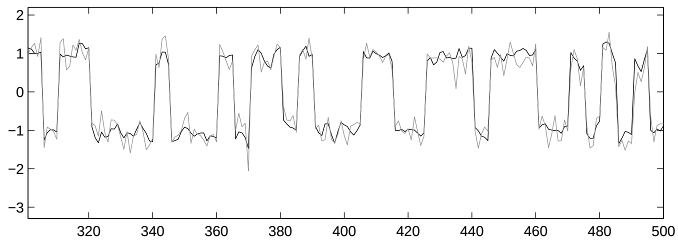

Fig. 4. Comparison between the noiseless input and its optimal estimate.

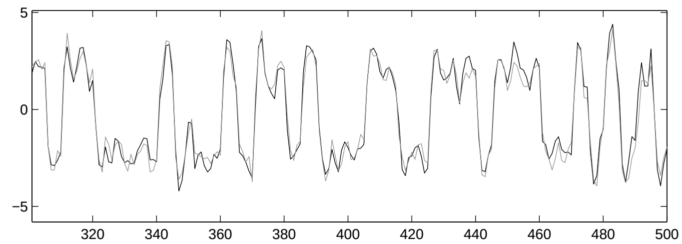

Fig. 5. Comparison between the first noiseless output and its optimal estimate.

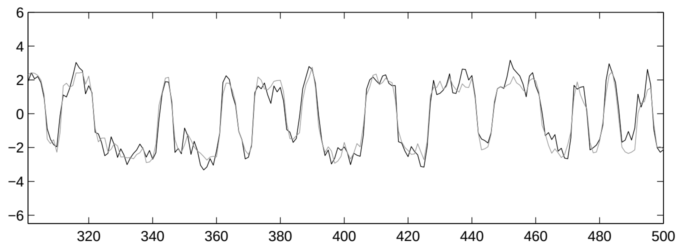

Fig. 6. Comparison between the second noiseless output and its optimal estimate.

The covariance matrices of the estimate errors, obtained by means of relations (50) and (51) for $t \to \infty$, are

$$\begin{aligned} \bar{P}_y &= \begin{bmatrix} 0.1941 & 0.1923 \\ 0.1923 & 0.2524 \end{bmatrix} \\ \bar{P}_u &= 0.0562, \end{aligned}$$

while the mean of the actual values obtained in the Monte Carlo simulation are

$$\begin{aligned} \bar{P}_y^{MC} &= \begin{bmatrix} 0.1952 \pm 0.0098 & 0.1935 \pm 0.0117 \\ 0.1935 \pm 0.0117 & 0.2536 \pm 0.0145 \end{bmatrix} \\ \bar{P}_u^{MC} &= 0.0565 \pm 0.0030. \end{aligned}$$

The results of the numerical simulation are thus in excellent agreement with the theoretical values.

VI. Conclusions

A new extended environment for Kalman filtering considering also the presence of additive noise on input observations has been defined. This environment includes as particular cases, both traditional Kalman filtering and EIV filtering. A Monte Carlo simulation has shown the effectiveness of this extended filter and the excellent agreement between its expected performance and the observed one.

References

[1] B. D. O. Anderson and J.B. Moore, *Optimal filtering*, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

[2] I. B. Rhodes, "A tutorial introduction to estimation and filtering", *IEEE Transactions on Automatic Control*, vol. 16, pp. 688–706, 1971.

[3] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering", *IEEE Transactions on Automatic Control*, vol. 15, pp. 175–184, 1970.

[4] P. R. Bélanger, "Estimation of noise covariance matrices for a linear time-varying stochastic process", *Automatica*, vol. 10, 3, pp. 267-275, 1974.

[5] R. Guidorzi, R. Diversi and U. Soverini, "Optimal Errors–in–Variables filtering", *Automatica*, vol. 39, pp. 281–289, 2003.

[6] R. Guidorzi, R. Diversi and U. Soverini, "Errors–in–variables filtering in behavioural and state–space contexts", in *Total Least Squares and Errors-in-Variables Modelling: Analysis, Algorithms and Applications*, S. Van Huffel and P. Lemmerling (Eds.), pp. 281-291, Kluwer Academic Publishers, 2002.

[7] R. Diversi, R. Guidorzi and U. Soverini, "Algorithms for optimal errors–in–variables filtering", *Systems & Control Letters*, vol.48, pp. 1–13, 2003.

[8] B. Roorda and C. Heij, "Global total least squares modeling of multivariable time series", *IEEE Transactions on Automatic Control*, vol.40, pp. 50–63, 1995.

[9] I. Markovsky and B. De Moor, "Linear dynamic filtering with noisy input and output", *Report ESAT-SISTA/TR 2002-191*, 2002.