

Technical Communique

Linear dynamic filtering with noisy input and output[☆]

Ivan Markovsky*, Bart De Moor

ESAT, SCD-SISTA, K.U.Leuven, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium

Received 11 February 2004; received in revised form 4 June 2004; accepted 30 August 2004

Available online 30 October 2004

Abstract

State estimation problems for linear time-invariant systems with noisy inputs and outputs are considered. An efficient recursive algorithm for the smoothing problem is presented. The equivalence between the optimal filter and an appropriately modified Kalman filter is established. The optimal estimate of the input signal is derived from the optimal state estimate. The result shows that the noisy input/output filtering problem is not fundamentally different from the classical Kalman filtering problem.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Errors-in-variables; Kalman filtering; Optimal smoothing; Misfit; Latency

1. Introduction

The noisy input/output (I/O) estimation problem is first put forward in Guidorzi, Diversi, and Soverini (2003), where it is called errors-in-variables filtering. In Guidorzi et al. (2003), Diversi, Guidorzi, and Soverini (2003a), a transfer functions approach is used and recursive algorithms that solve the filtering problem are derived. The treatment, however, is limited to the SISO case and the obtained solution is not linked to the classical Kalman filter.

The MIMO case is addressed in Markovsky and De Moor (2003), where the equivalence with a modified Kalman filter is established. Closely related to the approach of Markovsky and De Moor (2003) is the one used in Diversi, Guidorzi, and Soverini (2003b). The continuous-time version of the noisy I/O state estimation

problem is explicitly solved in Markovsky, Willems, and De Moor (2002) by a completion of the squares approach.

In this paper, we consider the deterministic discrete-time LTI state-space system

$$x(t+1) = Ax(t) + Bu(t), \quad x(0) = x_0,$$

$$y(t) = Cx(t) + Du(t) \quad (1)$$

together with the *measurement errors model*

$$u_d = u + \tilde{u}, \quad y_d = y + \tilde{y}. \quad (2)$$

The vector of measurement errors $\tilde{w} := \text{col}(\tilde{u}, \tilde{y})$ is a white, stationary, zero mean, stochastic process with positive-definite block diagonal covariance matrix $V_{\tilde{w}} = \text{blk diag}(V_{\tilde{u}}, V_{\tilde{y}})$. We refer to model (1) together with the measurement errors model (2) as the *noisy I/O model*. The considered problem is to find the least-squares estimate of the state x from the measured I/O data $w_d := \text{col}(u_d, y_d)$. We prove that the optimal filter is the Kalman filter for the system

$$x(t+1) = Ax(t) + Bu(t) + v_1(t),$$

$$y(t) = Cx(t) + Du(t) + v_2(t), \quad (3)$$

[☆]This paper was presented at the SYSID '03 IFAC Symposium, see Markovsky and De Moor (2003). This paper was recommended for publication in revised form by Associate Editor T. Chen under the direction of Editor I. Petersen.

* Corresponding author. Fax: +32 16 321970.

E-mail addresses: Ivan.Markovsky@esat.kuleuven.ac.be (I. Markovsky), Bart.DeMoor@esat.kuleuven.ac.be (B. De Moor).

where the process noise v_1 and the measurement noise v_2 have a joint covariance matrix

$$\begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} := \text{cov} \left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right) \\ = \begin{bmatrix} -B & 0 \\ -D & I \end{bmatrix} \begin{bmatrix} V_{\tilde{u}} & \\ & V_{\tilde{y}} \end{bmatrix} \begin{bmatrix} -B & 0 \\ -D & I \end{bmatrix}^\top. \quad (4)$$

The noisy I/O state estimation problem is treated implicitly in Roorda and Heij (1995) and Fagnani and Willems (1997), where the behavioral setting is adopted. In the behavioral setting, the observations are not partitioned into inputs and outputs and in this respect all observation channels are treated symmetrically. The global total least-squares problem of Roorda and Heij (1995) has as a subproblem the computation of the closest trajectory in the behavior of a given system to a given time series. This is a deterministic estimation problem, the recursive solution of which corresponds to the Kalman filter.

Application of the noisy I/O estimation problem for fault detection and data reconciliation is presented in Mourot, Maquin, and Ragot (1999), Ragot, Kratz, and Maquin (1999).

2. Problem formulation

A signal variable, without time index, denotes the column vector obtained by stacking consecutive signal samples, e.g., over the time horizon $0, 1, \dots, t_f - 1$, $u := \text{col}(u(0), \dots, u(t_f - 1))$, and $x := \text{col}(x(0), \dots, x(t_f))$. We denote by $\mathbf{V}_{\tilde{u}}$ and $\mathbf{V}_{\tilde{y}}$ the covariance matrices of \tilde{u} and \tilde{y} , and define $\mathbf{V} := \text{blk diag}(\mathbf{V}_{\tilde{u}}, \mathbf{V}_{\tilde{y}})$.

Definition 1 (Noisy I/O smoothing problem). The noisy I/O smoothing problem is defined by

$$\min_{\hat{x}, \hat{w} = \text{col}(\hat{u}, \hat{y})} (\hat{w} - w_d)^\top \mathbf{V}^{-1} (\hat{w} - w_d) \quad \text{s.t.} \\ \hat{x}(t+1) = A\hat{x}(t) + B\hat{u}(t), \quad \hat{y}(t) = C\hat{x}(t) + D\hat{u}(t) \quad (5)$$

for $t = 0, 1, \dots, t_f - 1$, and the noisy I/O smoothed state estimate $\hat{x}(\cdot, t_f)$ is the optimal solution of (5).

Definition 2 (Noisy I/O filtering problem). The noisy I/O filtering problem is to find a dynamical system,

$$z(t+1) = A_f(t)z(t) + B_f(t)w_d(t), \\ \hat{x}(t) = C_f(t)z(t) + D_f(t)w_d(t), \quad (6)$$

such that $\hat{x}(t) = \hat{x}(t, t+1)$, where $\hat{x}(\cdot)$ is the solution of (6), i.e., the noisy I/O filtered state estimate, and $\hat{x}(\cdot, t+1)$ is the noisy I/O smoothed state estimate with a time horizon $t+1$.

Note 1: The noisy I/O filtering problem is defined as a state estimation problem. When the input is measured with additive noise, an extra step is needed to find the filtered I/O signals from the state estimate. This is explained in Note 5.

Note 2 (Initial conditions): Definition 1 implicitly assumes there is no prior information about $x(0)$. Another possibility is $x(0)$ is exactly known. The standard assumption is $x(0) \sim \mathcal{N}(x_0, P_0)$, i.e., $x(0)$ normally distributed with mean x_0 and covariance P_0 . Exactly known initial conditions correspond to $P_0 = 0$ and unknown initial conditions correspond to $P_0 = \infty \cdot I$. The latter is a singular case that calls for the information matrix $P_0^{-1} = 0$. We have chosen the initial condition assumption that results in the simplest derivation.

3. Solution of the smoothing problem

We represent the I/O dynamics of system (1), over the time horizon $0, \dots, t_f - 1$, explicitly as $y = \Gamma x_0 + T u$, where

$$\Gamma := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t_f-1} \end{bmatrix}, \quad T := \begin{bmatrix} H_0 & 0 & \cdots & 0 \\ H_1 & H_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ H_{t_f-1} & H_{t_f-2} & \cdots & H_0 \end{bmatrix},$$

and $H_0 = D$, $H_t = CA^{t-1}B$, for $t > 0$, are the Markov parameters of (1). Using this representation, (5) becomes a classical Weighted Least Squares (WLS) problem

$$\min_{\hat{x}_0, \hat{u}} \left\| \sqrt{\mathbf{V}^{-1}} \left(\begin{bmatrix} u_d \\ y_d \end{bmatrix} - \begin{bmatrix} 0 & I \\ \Gamma & T \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \hat{u} \end{bmatrix} \right) \right\|^2. \quad (7)$$

Alternatively, we represent the input/state/output dynamics of the system over the time horizon $0, \dots, t_f - 1$, as $\bar{y} = \mathcal{A}x + \mathcal{B}u$, where

$$\bar{y} := \begin{bmatrix} y(0) \\ 0 \\ y(1) \\ 0 \\ \vdots \\ y(t_f-1) \\ 0 \end{bmatrix}, \quad \mathcal{A} := \begin{bmatrix} C & 0 \\ A & -I \\ & C & 0 \\ & A & -I \\ & & \ddots & \ddots \\ & & & C & 0 \\ & & & A & -I \end{bmatrix}, \\ \mathcal{B} := \begin{bmatrix} D \\ B \\ & D \\ & B \\ & & \ddots \\ & & & D \\ & & & B \end{bmatrix}.$$

Substituting $y_d - \bar{y}$ for y and $u_d - \tilde{u}$ for u (see (2)), we have

$$\bar{y}_d + \mathcal{B}u_d = \mathcal{A}x + \mathcal{B}\tilde{u} + \mathcal{C}\tilde{y}, \quad (8)$$

where \bar{y}_d is defined analogously to \bar{y} and $\mathcal{C} := \text{blk diag} \left(\begin{bmatrix} I \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} I \\ 0 \end{bmatrix} \right)$. Using (8) and defining $\Delta w := \text{col}(\Delta u, \Delta y)$, (5) is equivalent to the following problem:

$$\min_{\hat{x}, \Delta w} \Delta w^\top \mathbf{V}^{-1} \Delta w \\ \text{s.t. } \bar{y}_d + \mathcal{B}u_d = \mathcal{A}\hat{x} + [\mathcal{B} \ \mathcal{C}]\Delta w, \quad (9)$$

which is a minimum-norm-type problem, so that its solution can be computed in closed form.

Next we show a recursive solution for the case when $x(0) = x_0$ is given and $D = 0$. The more general case, $x(0) \sim \mathcal{N}(x_0, P_0)$ and $D \neq 0$ leads to similar but heavier result.

Define the value function $V_\tau: \mathbb{R}^n \rightarrow \mathbb{R}$, for $\tau = 0, 1, \dots, t_f$, where $n := \dim(x_0)$ as follows: $V_\tau(z)$ is the minimum value of (5) over $t = \tau, \dots, t_f - 1$ with $\hat{x}(\tau) = z$. Then $V_0(x_0)$ is the optimum value of the I/O smoothing problem. By the dynamic programming principle, we have

$$V_\tau(z) = \min_{\hat{u}(\tau)} (\|\sqrt{V_{\hat{u}}^{-1}}(\hat{u}(\tau) - u_d(\tau))\|^2 + \|\sqrt{V_{\hat{y}}^{-1}}(Cz - y_d(\tau))\|^2 + V_{\tau+1}(Az + B\hat{u}(\tau))). \quad (10)$$

The value function V_τ is quadratic for all τ , i.e., there are $P(\tau) \in \mathbb{R}^{n \times n}$, $s(\tau) \in \mathbb{R}^n$, and $v(\tau) \in \mathbb{R}$, such that $V_\tau(z) = z^\top P(\tau)z + 2s^\top(\tau)z + v(\tau)$. This allows us to solve (10).

Theorem 3 (Recursive smoothing). *The solution of the noisy I/O smoothing problem with given $x(0) = x_0$ and $D = 0$ is*

$$\hat{u}(t) = -(B^\top P(t+1)B + V_{\hat{u}}^{-1})^{-1}(B^\top P(t+1)A\hat{x}(t) + B^\top s(t+1) - V_{\hat{u}}^{-1}u_d(t)), \quad (11)$$

$\hat{x}(t+1) = A\hat{x}(t) + B\hat{u}(t)$, with $\hat{x}(0) = x_0$, and $\hat{y}(t) = C\hat{x}(t)$ for $t = 0, \dots, t_f - 1$, where

$$P(t) = -A^\top P(t+1)B(B^\top P(t+1)B + V_{\hat{u}}^{-1})^{-1} \times B^\top P(t+1)A + A^\top P(t+1)A + C^\top V_{\hat{y}}^{-1}C \quad (12)$$

for $t = t_f - 1, \dots, 0$, with $P(t_f) = 0$ and

$$s(t) = -A^\top P(t+1)B(B^\top P(t+1)B + V_{\hat{u}}^{-1})^{-1} \times (B^\top s(t+1) - V_{\hat{u}}^{-1}u_d(t)) + A^\top s(t+1) - C^\top V_{\hat{y}}^{-1}y_d(t) \quad (13)$$

for $t = t_f - 1, \dots, 0$, with $s(t_f) = 0$.

P and s are obtained from the backward-in-time recursions (12) and (13), and the estimates \hat{u} , \hat{x} , and \hat{y} are obtained from the forward-in-time recursion (11).

Note 3 (Suboptimal smoothing): With (C, A) observable, (12) has a steady-state solution \bar{P} that satisfies the algebraic Riccati equation (ARE)

$$\bar{P} = -A^\top \bar{P}B(B^\top \bar{P}B + V_{\hat{u}}^{-1})^{-1}B^\top \bar{P}A + A^\top \bar{P}A + C^\top V_{\hat{y}}^{-1}C. \quad (14)$$

In a suboptimal solution, \bar{P} can be substituted for $P(t)$ in (13) and (11). This is motivated by the typically fast convergence of $P(t)$ to \bar{P} . Then the smoothing procedure becomes:

1. solve the algebraic Riccati equation (14),
2. simulate the LTI system (13) with $P(t) \equiv \bar{P}$,
3. simulate the LTI system (11) with $P(t) \equiv \bar{P}$.

4. Solution of the filtering problem

Analogous to the derivation of (8) in Section 3, we now derive an equivalent to the noisy I/O model (1–2) representation in form (3). Substitute $u_d - \tilde{u}$ for u and $y_d - \tilde{y}$ for y (see (2)) in (1)

$$x(t+1) = Ax(t) + Bu_d(t) - B\tilde{u}(t), \\ y_d(t) = Cx(t) + Du_d(t) - D\tilde{u}(t) + \tilde{y}(t),$$

and define a “fake” process noise v_1 and measurement noise v_2 by $v_1 := -B\tilde{u}$ and $v_2 := -D\tilde{u} + \tilde{y}$. The resulting system

$$x(t+1) = Ax(t) + Bu_d(t) + v_1(t), \\ y_d(t) = Cx(t) + Du_d(t) + v_2(t), \quad (15)$$

is in form (3), where Q , S , and R are given in (4).

The Kalman filter corresponding to the modified system (15) with covariance (4) is

$$z(t+1) = A_{\text{kf}}(t)z(t) + B_{\text{kf}}(t)w_d(t), \\ \hat{x}(t) = C_{\text{kf}}(t)z(t) + D_{\text{kf}}(t)w_d(t), \quad (16)$$

where

$$A_{\text{kf}}(t) = (A - K(t)C), \quad B_{\text{kf}}(t) = [B - K(t)D, K(t)], \\ C_{\text{kf}}(t) = I - P(t)C^\top(CP(t)C^\top + R)^{-1}C, \\ D_{\text{kf}}(t) = P(t)C^\top(CP(t)C^\top + R)^{-1}[-D \ I], \\ K(t) = (AP(t)C^\top + S)(CP(t)C^\top + R)^{-1}, \quad (17)$$

and

$$P(t+1) = AP(t)A^\top - (AP(t)C^\top + S) \\ \times (CP(t)C^\top + R)^{-1}(AP(t)C^\top + S)^\top + Q.$$

We call (16) the *modified Kalman filter*. It recursively solves (8) (which is equivalent to (15)) for the last block entry of the unknown x . The solution is in the sense of the WLS problem

$$\min_{\hat{x}, \hat{e}} \hat{e}^\top ([\mathcal{B} \ \mathcal{C}]V[\mathcal{B} \ \mathcal{C}]^\top)^{-1} \hat{e} \quad \text{s.t. } \bar{y}_d + \mathcal{B}u_d = \mathcal{A}\hat{x} + \hat{e},$$

which is an equivalent optimization problem to the noisy I/O smoothing problem (9). Therefore, the noisy I/O filtering problem is solved by the modified Kalman filter.

Theorem 4. *The solution of the noisy filtering problem is $A_f = A_{\text{kf}}$, $B_f = B_{\text{kf}}$, $C_f = C_{\text{kf}}$, and $D_f = D_{\text{kf}}$, defined in (17).*

Note 4 (Suboptimal filtering): One can replace the time-varying Kalman filter with the (suboptimal) time-invariant

filter, obtained by replacing $P(t)$ in (16) with the steady-state solution \bar{P} of the ARE

$$\bar{P} = A\bar{P}A^\top - (A\bar{P}C^\top + S)(C\bar{P}C^\top + R)^{-1} \\ \times (A\bar{P}C^\top + S)^\top + Q.$$

Note 5 (Optimal estimation of the input/output signals): Up to now we were interested in the optimal filtering in the sense of state estimation. The optimal estimates of the input and the output, however, can be derived from the modified Kalman filter. The state estimate \hat{x} , the one-step-ahead prediction $z(t+1)$, and the optimal input estimate \hat{u} satisfy the equation

$$z(t+1) = A\hat{x}(t) + B\hat{u}(t). \quad (18)$$

Then we can find \hat{u} exactly from \hat{x} and $z(t+1)$ obtained from the modified Kalman filter (16). In fact, (18) and the Kalman filter equations imply that

$$\hat{u}(t) = E(t)z(t) + F(t)w_d(t), \quad (19)$$

where $E(t) := -V_{\hat{u}}D^\top(CP(t)C^\top + R(t))^{-1}C$ and

$$F(t) := [I - V_{\hat{u}}D^\top(CP(t)C^\top + R)^{-1}D, \\ V_{\hat{u}}D^\top(CP(t)C^\top + R)^{-1}].$$

The optimal output estimate is

$$\hat{y}(t) = (CC_{\text{kf}}(t) + DE(t))z(t) \\ + (CD_{\text{kf}}(t) + DF(t))w_d(t). \quad (20)$$

Appending the output equation of the noisy I/O filter (6) with Eqs. (19) and (20), we have an explicit solution of the errors-in-variables filtering problem of Guidorzi et al. (2003) as a (modified) Kalman filter.

Note 6 (Misfit/latency): More general estimation problems occur when the signals u, y are generated by the stochastic model (3) with a noise covariance matrix $V_v := \text{cov}(\text{col}(v_1, v_2))$, and the signals u_d, y_d , available for estimation, are generated by the measurement error model (2). The noisy I/O smoothing and filtering problems can be defined in this case analogously to Definitions 1 and 2 and the results of the paper can be repeated mutatis mutandis for the new problems. The final result is the equivalence of the noisy I/O filter to the modified Kalman filter (16, 17), with the only difference that now

$$\begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} = V_v + \begin{bmatrix} -B & 0 \\ -D & I \end{bmatrix} \begin{bmatrix} V_{\hat{u}} & \\ & V_{\hat{y}} \end{bmatrix} \begin{bmatrix} -B & 0 \\ -D & I \end{bmatrix}^\top.$$

The more general setup is attractive because the noises v_1, v_2 have different interpretation from that of \tilde{u}, \tilde{y} . In the former model the *latency* contribution and the latter model the *misfit* contribution, see Lemmerling and De Moor (2001), Markovsky et al. (2002).

Table 1

Comparison of the absolute errors of the state, input, and output estimates for all methods. (MKF—modified Kalman filter)

Method	$\ \hat{x} - x\ ^2$	$\ \hat{u} - u\ ^2$	$\ \hat{y} - y\ ^2$
Optimal smoothing	75.3981	29.2195	15.5409
Optimal filtering	75.7711	35.5604	16.4571
Time-varying MKF	75.7711	35.5604	16.4571
Time-invariant MKF	76.1835	35.7687	16.5675
Noisy data	116.3374	42.4711	41.2419

5. Numerical example

We illustrate numerically the results of the paper. The particular system used is $A = \begin{bmatrix} 0.6 & -0.45 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $C = [0.48429 \quad -0.45739]$, and $D = 0.5381$. The time horizon is $t_f = 100$, the initial state is $x_0 = 0$, the input signal is a normal white noise sequence with unit variance, and $V_{\hat{u}} = V_{\hat{y}} = 0.4$.

The estimate of the noisy I/O filter is computed directly from the definition, i.e., we solve a sequence of smoothing problems with increasing time horizon. Every smoothing problem is as a WLS problem (7). The last block entries of the obtained sequence of solutions form the noisy I/O filter state estimate.

We compare the noisy I/O filter estimate with the estimate of the modified Kalman filter (16). The state estimate \hat{x}_{KF} obtained by the modified Kalman filter is up to the numerical errors equal to the state estimate \hat{x}_f obtained by the noisy I/O filter, $\|\hat{x}_{\text{KF}} - \hat{x}_f\| < 10^{-14}$. This is the desired numerical verification of the theoretical result. The absolute errors of estimation $\|\hat{x} - x\|^2$, $\|\hat{u} - u\|^2$, $\|\hat{y} - y\|^2$ for all estimation methods discussed in the paper are given in Table 1.

6. Conclusions

We considered optimal noisy I/O estimation problems for discrete-time LTI systems. A recursive solution for the smoothing problem is derived. The filtering problem is solved via a modified Kalman filter. The equivalence between the noisy I/O filter and the modified Kalman filter is established algebraically using explicit state-space representation of the system. The optimal estimate of the input is a linear function of the optimal state estimate, so that it is obtained by an extra output equation of the modified Kalman filter. The results are extended to the case when the system is driven by a measured and an unobserved input.

Acknowledgements

Ivan Markovsky is a research assistant and Dr. Bart De Moor is a full-time professor at K.U. Leuven, Belgium. Our research is supported by Research Council KUL: GOA-Mefisto 666, several PhD/postdoc and fellow grants; Flemish Government: FWO: PhD/postdoc

grants, projects, G.0240.99 (multilinear algebra), G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), research communities (ICCoS, ANMMM); AWI: Bil. Int. Collaboration Hungary/Poland; IWT: PhD Grants, Soft4s (softsensors), Belgian Federal Government: DWTC (IUAP IV-02) (1996–2001) and IUAP V-22 (2002–2006), PODO-II (CP/40: TMS and Sustainability); EU: CAGE; ERNSI; Eureka 2063-IMPACT; Eureka 2419-FlITE; Contract Research/agreements: Data4s, Electrabel, Elia, LMS, IPCOS, VIB.

References

- Diversi, R., Guidorzi, R., & Soverini, U. (2003a). Algorithms for optimal errors-in-variables filtering. *Systems & Control Letters*, *48*, 1–13.
- Diversi, R., Guidorzi, R., & Soverini, U. (2003b). Kalman filtering in symmetrical noise environments. In: *Proceedings of 11th IEEE Mediterranean conference on control and automation*.
- Fagnani, F., & Willems, J. C. (1997). Deterministic Kalman filtering in a behavioral framework. *Systems & Control Letters*, *32*, 301–312.
- Guidorzi, R., Diversi, R., & Soverini, U. (2003). Optimal errors-in-variables filtering. *Automatica*, *39*, 281–289.
- Lemmerling, P., & De Moor, B. (2001). Misfit versus latency. *Automatica*, *37*, 2057–2067.
- Markovsky, I., & De Moor, B. (2003). Linear dynamic filtering with noisy input and output. In *Proceedings of the 13th IFAC symposium on system identification* (pp. 1749–1754).
- Markovsky, I., Willems, J. C., & De Moor, B. (2002). Continuous-time errors-in-variables filtering. In *Proceedings of the CDC* (pp. 2576–2581).
- Mourot, G., Maquin, D., & Ragot, J. (1999). Simultaneous state and parameter estimation. In *14th IFAC World congress*. Beijing, China.
- Ragot, J., Kratz, F., & Maquin, D. (1999). Finite memory observer for input–output estimation. In *IFAC symposium*. Budapest, Hungary.
- Roorda, B., & Heij, C. (1995). Global total least squares modeling of multivariate time series. *IEEE Transactions on Automatic Control*, *40*(1), 50–63.