Proceedings of the 41st IEEE
Conference on Decision and Control
Las Vegas, Nevada USA, December 2002

Continuous-time Errors-in-variables Filtering

Ivan Markovsky, Jan C. Willems, and Bart De Moor

ESAT-SCD (SISTA), University of Leuven
Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium

{Ivan.Markovsky,Jan.Willems}@esat.kuleuven.ac.be

Abstract

We consider estimation problems for a continuous-time linear system with a state disturbance and additive errors on the input and the output. The problem formulation and the estimation principle are deterministic. The derived filter is identical to the stochastic Kalman filter. The problem formulation with additive error on both the input and the output, however, is more symmetric then the classical Kalman filter one and allows interpretation in terms of misfit and latent variables.

Keywords: Kalman filtering, errors-in-variables, misfit, latency.

1 Introduction

Consider the continuous-time linear state space system

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}x &= Ax + Bu, \\ y &= Cx + Du, \end{aligned} \tag{1}$$

as a model of a phenomenon. Assume that we measure an input/output trajectory $(u_\mathrm{d}, y_\mathrm{d})$ on the interval $[0, t_f]$. Due to modeling and measurement errors, $(u_\mathrm{d}, y_\mathrm{d})$ can, in general, not be explained as a trajectory of the model, i.e., there is no initial condition $x(0)$ such that the response of the system (1) to the signal u_d is the signal y_d.

We further account for modeling errors by adding to the system an auxiliary input d, later called a *disturbance*. The model with the disturbance input becomes

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}x &= Ax + Bu + Gd, \\ y &= Cx + Du + Hd. \end{aligned} \tag{2}$$

The measurement errors are modeled by appending to the system the equations

$$u_\mathrm{d} = u + \tilde{u}, \qquad y_\mathrm{d} = y + \tilde{y}. \tag{3}$$

The *measurement errors* $\tilde{u}$, $\tilde{y}$ and the disturbance d are (in addition to the initial state $x(0)$) *free variables* that allow us to "explain" the measured trajectory $(u_\mathrm{d}, y_\mathrm{d})$.

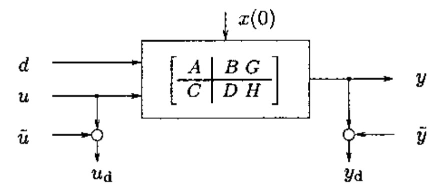

Figure 1: Errors-in-variables model

The system (2,3) is called an *errors-in-variables model* (see Figure 1).

We assume that the *true input/state/output trajectory* is generated by (1) and that the measurements deviate from the true values due to "small" initial condition, disturbance and measurement errors according to the errors-in-variables model. In the paper, we consider the problem to estimate the true input/state/output trajectory under the above assumption.

The *estimation principle* used is: find the "smallest" estimated initial condition $\hat{x}(0)$, disturbance $\hat{d}$, and measurement errors $\hat{\tilde{u}}$, $\hat{\tilde{y}}$ that "explain" the measurements by the errors-in-variables model, i.e.,

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\hat{x} &= A\hat{x} + B\hat{u} + G\hat{d}, \\ \hat{y} &= C\hat{x} + D\hat{u} + H\hat{d}, \end{aligned}$$

$$u_\mathrm{d} = \hat{u} + \hat{\tilde{u}}, \qquad y_\mathrm{d} = \hat{y} + \hat{\tilde{y}}.$$

The rationale for this deterministic estimation principle, see [Wil02], is that we assume a priori the true initial state, disturbance, and measurement errors are small, so that the "most likely" estimates are the smallest possible estimates that explain the measurements.

We interpret the estimates $\hat{x}(0)$ and $\hat{d}$ as *latent variables*, i.e., variables that modify the system equations to explain the model-data mismatch. On the other hand, $\hat{\tilde{u}}$, $\hat{\tilde{y}}$ are viewed as *misfit variables*, i.e., variables that modify the data to make it match the system equations. In the *pure latency* case, we consider the data as being error free and blame the model as being imperfect description of the reality. In the *pure misfit* case, we consider the model as being perfect and blame the measurements as being error corrupted. Thus the misfit and the latent variables are conceptually different.

In general, both misfit and latency can be considered simultaneously, which corresponds to uncertainty in both the data and the model. The terms in the cost functional, corresponding to the misfit and the latency contributions are,

$$J_{\text{misfit}} := \int_0^{t_f} \left(\|\hat{u}(t) - u_{\text{d}}(t)\|_R^2 + \|\hat{y}(t) - y_{\text{d}}(t)\|_Q^2 \right) \mathrm{d}t$$

and

$$J_{\text{latency}} := \int_0^{t_f} \|\hat{d}(t)\|_P^2 \, \mathrm{d}t + \|\hat{x}(0)\|_\Gamma^2,$$

where $R > 0$, $Q \geq 0$, $P > 0$, and $\Gamma \geq 0$ are weighting matrices, reflecting the relative importance of the terms. We introduce a variable weighting factor $\rho \in [0, 1]$, and consider the total cost

$$J := \rho J_{\text{misfit}} + (1 - \rho) J_{\text{latency}}.$$

Varying ρ from 0 to 1, allows smooth transition from a purely misfit contribution to a purely latency contribution. The general errors-in-variables estimation problem considered in the paper is

$$\min_{\hat{u},\hat{y},\hat{x},\hat{d}} J \qquad \text{s.t.} \qquad \begin{array}{rcl} \frac{\mathrm{d}}{\mathrm{d}t}\hat{x} & = & A\hat{x} + B\hat{u} + G\hat{d} \\ \hat{y} & = & C\hat{x} + D\hat{u} + H\hat{d}. \end{array} \tag{4}$$

We distinguish two types of estimation problems — smoothing and filtering. In the *smoothing problem*, one is interested in the estimates for the whole interval $[0, t_f]$ of observation. When the time-horizon t_f is increasing, one needs an estimation procedure that works in real-time. This poses the *filtering problem* — design a non-anticipating dynamical system that assumes as input the measurements and produces as an output at each time optimal estimate.

As stated, problem (4) is a smoothing problem. Our solution, however, is derived so that with no extra work we obtain the solution of the corresponding filtering problem (see Remark 1).

In Section 2, we solve an optimal control problem with cost function a general quadratic function in the state and the control. Problem (4) is a special case of this control problem. In Sections 3 and 4, we solve separately the cases of pure misfit cost, and pure latency cost. The results are stated as the solution of the corresponding smoothing problems, the finite horizon filtering problem, and the infinite horizon filtering problems. In Section 5, we solve the general problem (4). Section 6 gives conclusions.

2 Preliminaries

In this section, we solve the following optimal control problem

$$\min_{u,x} \int_0^{t_f} \begin{bmatrix} u(t) \\ x(t) \\ 1 \end{bmatrix}^T M(t) \begin{bmatrix} u(t) \\ x(t) \\ 1 \end{bmatrix} \mathrm{d}t + x^T(0)\Gamma x(0) \tag{5}$$
$$\text{s.t.} \quad \tfrac{\mathrm{d}}{\mathrm{d}t}x = Ax + Bu,$$

where

$$M(t) := \begin{bmatrix} Q_u & Q_{ux} & q_u(t) \\ * & Q_x & q_x(t) \\ * & * & q(t) \end{bmatrix},$$

is a symmetric matrix ("$*$" denotes the symmetric blocks) with $Q_u > 0$, $\left[\begin{smallmatrix} Q_u & Q_{ux} \\ * & Q_x \end{smallmatrix}\right] \geq 0$, and $\Gamma \geq 0$. We use the "completion of squares" approach [Bro70]. First we prove the following identity.

Lemma 1 *Let $K : [0, t_f] \to \mathbb{R}^{n\times n}$, $K = K^T$, and $s : [0, t_f] \to \mathbb{R}^n$ be differentiable. Then, for x and u related by $\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Bu$, the following identity holds*

$$0 = -\left(x^T K x + 2s^T x\right)\Big|_0^{t_f} + \int_0^{t_f} \begin{bmatrix} u \\ x \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & B^T K & B^T s \\ * & \frac{\mathrm{d}}{\mathrm{d}t}K + A^T K + KA & \frac{\mathrm{d}}{\mathrm{d}t}s + A^T s \\ * & * & 0 \end{bmatrix} \begin{bmatrix} u \\ x \\ 1 \end{bmatrix} \mathrm{d}t.$$

Proof: Clearly,

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} K & s \\ s^T & 0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{d}}{\mathrm{d}t}x \\ 0 \end{bmatrix}^T \begin{bmatrix} K & s \\ s^T & 0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} + \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} \frac{\mathrm{d}}{\mathrm{d}t}K & \frac{\mathrm{d}}{\mathrm{d}t}s \\ \frac{\mathrm{d}}{\mathrm{d}t}s^T & 0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} + \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} K & s \\ * & 0 \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}}{\mathrm{d}t}x \\ 0 \end{bmatrix}.$$

Substituting $Ax + Bu$ for $\frac{\mathrm{d}}{\mathrm{d}t}x$ and integrating from 0 to t_f the identity, we obtain

$$0 = \int_0^{t_f} \begin{bmatrix} u \\ x \\ 1 \end{bmatrix}^T \left(\begin{bmatrix} 0 & B^T K & B^T s \\ 0 & A^T K & A^T s \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{\mathrm{d}}{\mathrm{d}t}K & \frac{\mathrm{d}}{\mathrm{d}t}s \\ 0 & \frac{\mathrm{d}}{\mathrm{d}t}s^T & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ KB & KA & 0 \\ s^T B & s^T A & 0 \end{bmatrix} \right) \begin{bmatrix} u \\ x \\ 1 \end{bmatrix} \mathrm{d}t - \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} K & s \\ s^T & 0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \Bigg|_0^{t_f},$$

from which the lemma follows. ∎

Now we solve (5).

Lemma 2 *Let $Q_u > 0$, $\left[\begin{smallmatrix} Q_u & Q_{ux} \\ * & Q_x \end{smallmatrix}\right] \geq 0$, and $\Gamma \geq 0$. Define K as the unique solution of the Riccati differential equation*

$$0 = \tfrac{\mathrm{d}}{\mathrm{d}t}K + A^T K + KA + Q_x - (B^T K + Q_{ux})^T Q_u^{-1} (B^T K + Q_{ux}), \quad K(0) = -\Gamma,$$

and assume that $K(t_f) < 0$. Then the unique solution of problem (5) is

$$u = -Q_u^{-1}\big((B^TK + Q_{ux})x + B^Ts + q_u\big),$$

and

$$\tfrac{\mathrm{d}}{\mathrm{d}t}x = \big(A - BQ_u^{-1}(B^TK + Q_{ux})\big)x - BQ_u^{-1}B^Ts - BQ_u^{-1}q_u,$$

with final condition $x(t_f) = -K^{-1}(t_f)s(t_f)$, where s is generated by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}s = -\big(A - BQ_u^{-1}(B^TK + Q_{ux})\big)^Ts - q_x + (B^TK + Q_{ux})^TQ_u^{-1}q_u, \qquad s(0) = 0.$$

The minimum value of the cost functional is

$$\int_0^{t_f} \big(q - (B^Ts + q_u)^TQ_u^{-1}(B^Ts + q_u)\big)\,\mathrm{d}t + s^T(t_f)K^{-1}(t_f)s(t_f).$$

Proof: Adding the identity from Lemma 1 to the cost functional, we have

$$J = \int_0^{t_f} \begin{bmatrix} u(t) \\ x(t) \\ 1 \end{bmatrix}^T M(t) \begin{bmatrix} u(t) \\ x(t) \\ 1 \end{bmatrix} \mathrm{d}t - \big(x^TKx + 2s^Tx\big)\big|_0^{t_f} + x^T(0)\Gamma x(0).$$

Where $M :=$

$$\begin{bmatrix} Q_u & B^TK + Q_{ux} & B^Ts + q_u \\ * & \frac{\mathrm{d}}{\mathrm{d}t}K + A^TK + KA + Q_x & \frac{\mathrm{d}}{\mathrm{d}t}s + A^Ts + q_x \\ * & * & q \end{bmatrix}.$$

To complete the squares, we need to factor M in the form $\bar{M}^T\bar{M}$, by choosing $\frac{\mathrm{d}}{\mathrm{d}t}s$ and $\frac{\mathrm{d}}{\mathrm{d}t}K$. If we take

$$\tfrac{\mathrm{d}}{\mathrm{d}t}s + A^Ts + q_x = (B^TK + Q_{ux})^TQ_u^{-1}(B^Ts + q_u),$$

and

$$\tfrac{\mathrm{d}}{\mathrm{d}t}K + A^TK + KA + Q_x = (B^TK + Q_{ux})^TQ_u^{-1}(B^TK + Q_{ux}),$$

then $M = \bar{M}^T\bar{M}$, with

$$\bar{M} = Q_u^{-1/2}[Q_u, \quad B^TK + Q_{ux}, \quad B^Ts + q_u].$$

We have

$$J = \int_0^{t_f} \|Q_uu + (B^TK + Q_{ux})x + B^Ts + q_u\|^2_{Q_u^{-1/2}}\,\mathrm{d}t + \int_0^{t_f} \big(q - (B^Ts + q_u)^TQ_u^{-1}(B^Ts + q_u)\big)\,\mathrm{d}t - \big(x^TKx + 2s^Tx\big)\big|_0^{t_f} + x^T(0)\Gamma x(0).$$

For any input u, the following inequality holds

$$J \geq \int_0^{t_f} \big(q - (B^Ts + q_u)^TQ_u^{-1}(B^Ts + q_u)\big)\,\mathrm{d}t + 2x^T(0)s(0) + x^T(0)(K(0) + \Gamma)x(0) + s^T(t_f)K^{-1}(t_f)s(t_f) - \big(x(t_f) + K^{-1}(t_f)s(t_f)\big)^TK(t_f)\big(x(t_f) + K^{-1}(t_f)s(t_f)\big).$$

If we select $s(0) = 0$, $K(0) = -\Gamma$, and

$$u = -Q_u^{-1}\big((B^TK + Q_{ux})x + B^Ts + q_u\big),$$

then

$$x(t_f) = -K^{-1}(t_f)s(t_f), \tag{6}$$

achieves the optimal value of the cost functional. ∎

Remark 1 (Final condition) *The choice of the final condition, is what we need in filtering, since the estimate in the final moment of the time is of interest. Thus (6) gives the optimal filter as a by-product from the solution of the smoothing problem.*

3 Estimation in the pure misfit case

In this section, we consider the pure misfit case. We assume $d = 0$ and add only the misfit contribution in the cost functional. Problem (4) becomes

$$\begin{aligned} \min_{\hat{u},\hat{y},\hat{x}} \int_0^{t_f} \big(\|\hat{u}(t) - u_{\mathrm{d}}(t)\|_R^2 + \|\hat{y}(t) - y_{\mathrm{d}}(t)\|_Q^2\big)\,\mathrm{d}t \\ \text{s.t.} \quad \tfrac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + B\hat{u} \\ \hat{y} = C\hat{x} + D\hat{u}. \end{aligned} \tag{7}$$

Theorem 1 (Pure misfit smoothing) *Let $R > 0$, $Q \geq 0$, and (A, C^TQC) be observable. The unique solution of (7) is*

$$\hat{u} = -N^{-1}\big((B^TK + D^TQC)\hat{x} + B^Ts - D^TQy_{\mathrm{d}} - Ru_{\mathrm{d}}\big)$$

and

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\hat{x} = \big(A - BN^{-1}(B^TK + D^TQC)\big)^T\hat{x} - BN^{-1}B^Ts + BN^{-1}D^TQy_{\mathrm{d}} + BN^{-1}Ru_{\mathrm{d}}, \ \hat{x}(t_f) = -K^{-1}(t_f)s(t_f)$$

where $N := R + D^TQD$, s is generated by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}s = -\big(A - BN^{-1}(B^TK + D^TQC)\big)^Ts + C^TQy_{\mathrm{d}} - (B^TK + D^TQC)^TN^{-1}(D^TQy_{\mathrm{d}} + Ru_{\mathrm{d}}), \quad s(0) = 0,$$

and K is generated by

$$(B^TK + D^TQC)^TN^{-1}(B^TK + D^TQC) = \tfrac{\mathrm{d}}{\mathrm{d}t}K + A^TK + KA + C^TQC, \quad K(0) = 0. \tag{8}$$

The minimum value of the cost functional is

$$s^T(t_f)K^{-1}(t_f)s(t_f) + \int_0^{t_f} \big(y_\mathrm{d}^T Q y_\mathrm{d} + u_\mathrm{d}^T R u_\mathrm{d} \\ -(B^T s - D^T Q y_\mathrm{d} - R u_\mathrm{d})^T N^{-1}(B^T s - D^T Q y_\mathrm{d} - R u_\mathrm{d})\big)\,\mathrm{dt}.$$

Proof: The cost functional can be written as

$$\int_0^{t_f} \begin{bmatrix}\hat{u}\\ \hat{x}\\ 1\end{bmatrix}^T \begin{bmatrix} N & D^TQC & -D^TQy_\mathrm{d} - Ru_\mathrm{d} \\ * & C^TQC & -C^TQy_\mathrm{d} \\ * & * & y_\mathrm{d}^TQy_\mathrm{d} + u_\mathrm{d}^TRu_\mathrm{d}\end{bmatrix} \begin{bmatrix}\hat{u}\\ \hat{x}\\ 1\end{bmatrix} \mathrm{dt}.$$

where $N := R + D^TQD$. Then the misfit smoothing problem is a special case of (5) with

$$Q_u = N, \quad Q_{ux} = D^TQC, \quad q_u = -D^TQy_\mathrm{d} - Ru_\mathrm{d},$$
$$Q_x = C^TQC, \quad q_x = -C^TQy_\mathrm{d}, \quad q = y_\mathrm{d}^TQy_\mathrm{d} + u_\mathrm{d}^TRu_\mathrm{d},$$

and $\Gamma = 0$. The assumptions $R > 0$ and $Q \geq 0$ ensure that $N > 0$ and $\begin{bmatrix} N & D^TQC \\ * & C^TQC\end{bmatrix} \geq 0$. The observability assumption ensures that $K(t_f) < 0$. Thus we can apply Lemma 2, from which the result follows. ■

The solution of the smoothing problem contains the solution of the filtering problem.

Corollary 1 (Pure misfit finite-horizon filtering) *Under the assumptions of Theorem 1, the optimal misfit filter is*

$$\tfrac{\mathrm{d}}{\mathrm{dt}}s = -\big(A - BN^{-1}(B^TK + D^TQC)\big)^T s + C^TQy_\mathrm{d} \\ - (B^TK + D^TQC)^TN^{-1}(D^TQy_\mathrm{d} + u_\mathrm{d}), \quad s(0) = 0,$$

$$\hat{x} = -K^{-1}s,$$
$$\hat{u} = -N^{-1}\Big(\big(B^T + (B^TK + D^TQC)K\big)s - D^TQy_\mathrm{d} - Ru_\mathrm{d}\Big),$$
$$\hat{y} = C\hat{x} + D\hat{u},$$

where $N := R + D^TQD$ and K is generated by (8).

In the infinite-horizon case, assuming in addition that (A, B) is controllable, the differential Riccati equation reduces to the algebraic Riccati equation

$$0 = A^TK + KA + C^TQC \\ - (B^TK + D^TQC)^TN^{-1}(B^TK + D^TQC).$$

and the unique negative definite solution K_- is used in place of the time-function K in the optimal misfit filter.

4 Estimation in the pure latency case

In this section, we consider the pure latency case. We assume $\tilde{u} = 0$, $\tilde{y} = 0$ and add only the latency contribution in the cost functional. Problem (4) becomes

$$\min_{\hat{d},\hat{x}} \int_0^{t_f} \|\hat{d}(t)\|_P^2\,\mathrm{dt} + x^T(0)\Gamma x(0) \\ \text{s.t.} \quad \tfrac{\mathrm{d}}{\mathrm{dt}}\hat{x} = A\hat{x} + Bu_\mathrm{d} + G\hat{d} \\ y_\mathrm{d} = C\hat{x} + Du_\mathrm{d} + H\hat{d}. \tag{9}$$

This corresponds to the deterministic Kalman filter discussed in [Wil02].

Theorem 2 (Pure latency smoothing) *Let $H \in \mathbb{R}^{n_y \times n_d}$ be of full row rank with $n_y \leq n_d$, $P > 0$, and $\Gamma > 0$. The unique solution of (9) is*

$$\hat{u} = -Q_u^{-1}\big((\bar{B}^TK + H_2^TFC)\hat{x} + \bar{B}^Ts - H_2^TF(y_\mathrm{d} - Du_\mathrm{d})\big),$$

and

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\hat{x} = \big(\bar{A} - \bar{B}Q_u^{-1}(\bar{B}^TK + H_2^TFC)\big)\hat{x} - \bar{B}Q_u^{-1}\bar{B}^Ts \\ + (B - \bar{B}Q_u^{-1}H_2^TFD - G_1H_1^{-1}D)u_\mathrm{d} \\ + (G_1H_1^{-1} + \bar{B}Q_u^{-1}H_2^TF)y_\mathrm{d},$$

with final condition $x(t_f) = -K^{-1}(t_f)s(t_f)$, where s and K are generated by

$$\tfrac{\mathrm{d}}{\mathrm{dt}}s = -\big(\bar{A} - \bar{B}Q_u^{-1}(\bar{B}^TK + H_2^TFC)\big)^T s \\ + \big(C^T - (\bar{B}^TK + H_2^TFC)^TQ_u^{-1}H_2^T\big)F(y_\mathrm{d} - Du_\mathrm{d}), \; s(0) = 0,$$

and

$$(\bar{B}^TK + H_2^TFC)^TQ_u^{-1}(\bar{B}^TK + H_2^TFC) \\ = \tfrac{\mathrm{d}}{\mathrm{dt}}K + \bar{A}^TK + K\bar{A} + C^TFC, \quad K(0) = -\Gamma. \tag{10}$$

We define

$$\bar{A} := A - G_1H_1^{-1}C, \quad \bar{B} := G_2 - G_1H_1^{-1}H_2, \\ Q_u := P + H_2^TFH_2, \quad F := H_1^{-T}PH_1^{-1}. \tag{11}$$

Proof: Since H is with independent rows and wide its columns can be rearranged so that $H = [H_1\ H_2]$, with H_1 square and invertible. Let $\hat{d}^T = [\hat{d}_1^T\ \hat{d}_2^T]$ be the corresponding partitioning of $\hat{d}$. Then

$$\hat{d}_1 = H_1^{-1}(y_\mathrm{d} - C\hat{x} - Du_\mathrm{d} - H_2\hat{d}_2)$$

and we can eliminate the output equation as a constraint, by substituting $\hat{d}_1$ into the cost functional and the state equation. Let $G = [G_1\ G_2]$ corresponds to the partitioning of $\hat{d}$. The pure latency estimation problem becomes

$$\min_{\hat{d}_2,\hat{x}} \int_0^{t_f} ||H_1^{-1}(y_d(t) - C\hat{x}(t) - Du_d(t) - H_2\hat{d}_2(t))||_P^2 \, dt + \int_0^{t_f} ||\hat{d}_2||_P^2 \, dt + x^T(0)\Gamma x(0)$$

$$\text{s.t.} \quad \tfrac{d}{dt}\hat{x} = A\hat{x} + Bu_d + G_2\hat{d}_2 + G_1H_1^{-1}(y_d - C\hat{x} - Du_d - H_2\hat{d}_2).$$

The cost functional can be written as

$$\int_0^{t_f} \begin{bmatrix} \hat{d}_2(t) \\ \hat{x}(t) \\ 1 \end{bmatrix}^T M(t) \begin{bmatrix} \hat{d}_2(t) \\ \hat{x}(t) \\ 1 \end{bmatrix} dt + x^T(0)\Gamma x(0),$$

where $M :=$

$$\begin{bmatrix} P + H_2^TFH_2 & H_2^TFC & -H_2^TF(y_d - Du_d) \\ * & C^TFC & -C^TF(y_d - Du_d) \\ * & * & (y_d - Du_d)^TF(y_d - Du_d) \end{bmatrix},$$

and $F := H_1^{-T}PH_1^{-1}$. The state equation becomes

$$\tfrac{d}{dt}\hat{x} = (A - G_1H_1^{-1}C)\hat{x} + (G_2 - G_1H_1^{-1}H_2)\hat{d}_2 + [B - G_1H_1^{-1}D, \;\; G_1H_1^{-1}] \begin{bmatrix} u_d \\ y_d \end{bmatrix},$$

which is in the form $\frac{d}{dt}\hat{x} = \bar{A}\hat{x} + \bar{B}\hat{u} + v$, where v is an additional input signal

$$v := (B - G_1H_1^{-1}D)u_d + G_1H_1^{-1}y_d.$$

The pure latency estimation is a special case of the control problem (5) with

$$\begin{aligned} Q_u &= P + H_2^TFH_2, & Q_{ux} &= H_2^TFC \\ q_u &= -H_2^TF(y_d - Du_d), & Q_x &= C^TFC \\ q_x &= -C^TF(y_d - Du_d), & q &= (y_d - Du_d)^TF(\cdot), \end{aligned}$$

the substitution

$$A \leftarrow A - G_1H_1^{-1}C, \qquad B \leftarrow G_2 - G_1H_1^{-1}H_2,$$

and with added v to the right-hand side of the constraint (i.e. the state equation).

The assumption $P > 0$ ensures that $P + H_2^TFH_2 > 0$ and the assumption $\Gamma > 0$ ensures that $K(t_f) < 0$. Thus we can apply Lemma 2. To account for the additional input v, we add it to the right-hand side of the differential equation for the state estimate. ■

Corollary 2 (Pure latency filtering) *Under the assumptions of Theorem 2, the optimal latency filter is*

$$\tfrac{d}{dt}s = -\big(\bar{A} - \bar{B}Q_u^{-1}(\bar{B}^TK + H_2^TFC)\big)^Ts + \big(C^T - (\bar{B}^TK + H_2^TFC)Q_u^{-1}H_2^T\big)F(y_d - Du_d), \; s(0) = 0,$$

$$\begin{aligned} \hat{x} &= -K^{-1}s, \\ \hat{u} &= -Q_u^{-1}\Big(\big(\bar{B}^T(\bar{B}^TK + H_2^TFC)K\big)s - H_2^TF(y_d - Du_d)\Big), \\ \hat{y} &= C\hat{x} + D\hat{u}, \end{aligned}$$

where K is generated by (10), and $\bar{A}$, $\bar{B}$, Q_u, and F are defined in (11).

In the infinite-horizon case, assuming in addition that $(\bar{A}, \bar{B})$ is controllable and $(\bar{A}, C)$ is observable, the time function K is replaced by the unique negative definite solution of the algebraic Riccati equation

$$\bar{A}^TK + K\bar{A} + C^TFC = (\bar{B}^TK + H_2^TFC)^TQ_u^{-1}(\bar{B}^TK + H_2^TFC).$$

5 General case

In this section, we consider the general case (4) when both misfit and latency are taken into account.

Theorem 3 (Misfit and latency smoothing) *Let $R > 0$, $Q \geq 0$, $P > 0$, and $\Gamma > 0$. The unique solution of (4) for $\rho \in (0, 1)$ is*

$$\hat{u} = -Q_u^{-1}(Lx + \bar{B}^Ts + q_u)$$

and

$$\tfrac{d}{dt}x = (A - \bar{B}Q_u^{-1}L)x - \bar{B}Q_u^{-1}\bar{B}^Ts - \bar{B}Q_u^{-1}q_u,$$

with final condition $x(t_f) = -K^{-1}(t_f)s(t_f)$, where s and K are generated by

$$\tfrac{d}{dt}s = -(A - \bar{B}Q_u^{-1}L)^Ts + \rho C^TQy_d + LQ_u^{-1}q_u, \; s(0) = 0,$$

and

$$0 = \tfrac{d}{dt}K + A^TK + KA + \rho C^TQC - L^TQ_u^{-1}L, \quad (12)$$

$K(0) = -(1 - \rho)\Gamma$. *We define,*

$$\begin{aligned} \bar{B} &:= [B \;\; G], \\ Q_u &:= \rho \begin{bmatrix} R + D^TQD & D^TQH \\ * & \frac{(1-\rho)}{\rho}P + H^TQH \end{bmatrix}, \\ L &:= \begin{bmatrix} B^TK + \rho D^TQC \\ G^TK + \rho H^TQC \end{bmatrix}, \\ q_u &:= -\rho \begin{bmatrix} Ru_d + D^TQy_d \\ H^TQy_d \end{bmatrix}. \end{aligned} \quad (13)$$

Proof: The output equation is eliminated from the optimization problem (4) by directly substituting it in the cost function. Then the cost functional can be written as

$$\int_0^{t_f} \begin{bmatrix} \hat{u}(t) \\ \hat{d}(t) \\ \hat{x}(t) \\ 1 \end{bmatrix}^T M(t) \begin{bmatrix} \hat{u}(t) \\ \hat{d}(t) \\ \hat{x}(t) \\ 1 \end{bmatrix} \mathrm{d}t + (1-\rho)x^T(0)\Gamma x(0),$$

where $M :=$

$$\rho \begin{bmatrix} N & D^TQH & D^TQC & -(Ru_\mathrm{d} + D^TQy_\mathrm{d}) \\ * & \frac{(1-\rho)}{\rho}P + H^TQH & H^TQC & -H^TQy_\mathrm{d} \\ * & * & C^TQC & -C^TQy_\mathrm{d} \\ * & * & * & u_\mathrm{d}^TRu_\mathrm{d} + y_\mathrm{d}^TQy_\mathrm{d} \end{bmatrix}.$$

and $N := R + D^TQD$. This is a special case of problem (5) with

$$Q_u = \rho \begin{bmatrix} N & D^TQH \\ * & \frac{(1-\rho)}{\rho}P + H^TQH \end{bmatrix}, \quad Q_x = \rho C^TQC,$$

$$Q_{ux} = \rho \begin{bmatrix} D^TQC \\ H^TQC \end{bmatrix}, \quad q_u = -\rho \begin{bmatrix} Ru_\mathrm{d} + D^TQy_\mathrm{d} \\ H^TQy_\mathrm{d} \end{bmatrix},$$
$$q_x = -\rho C^TQy_\mathrm{d}, \qquad q = \rho(u_\mathrm{d}^TRu_\mathrm{d} + y_\mathrm{d}^TQy_\mathrm{d}).$$

and with the substitution

$$B \leftarrow [B \;\; G], \qquad \Gamma \leftarrow (1-\rho)\Gamma.$$

The assumption $R > 0$ ensures that $N > 0$ and $(1-\rho)\Gamma > 0$ ensures that $K(t_f) < 0$. We can apply Lemma 2, from which the theorem follows. ∎

The optimal filter is given in the following corollary.

Corollary 3 (Misfit and latency filtering) *Under the assumptions of Theorem 3, the optimal misfit and latency filter is*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}s = -(A - \bar{B}Q_u^{-1}L)^Ts + \rho C^TQy_\mathrm{d} + LQ_u^{-1}q_u, \; s(0) = 0,$$
$$\hat{x} = -K^{-1}s,$$
$$\hat{u} = -Q_u^{-1}\big((\bar{B}^T - LK^{-1})s + q_u\big),$$
$$\hat{y} = -\big(CK^{-1} + DQ_u^{-1}(\bar{B}^T - LK^{-1})\big)s - DQ_u^{-1}q_u,$$

where K is generated by (12) and $\bar{B}$, Q_u, L, q_u are defined in (13).

The infinite horizon case, assuming in addition that (A, B) is controllable and (A, C^TQC) is observable, the optimal time-invariant filter is obtained by replacing K with the matrix K_- for all t, where K_- is the unique negative definite solution of the algebraic Riccati equation

$$0 = A^TK + KA + \rho C^TQC - L^TQ_u^{-1}L.$$

6 Conclusions

We posed an estimation problem for a continuous-time linear time-invariant system with disturbance in the state and the output equations and with additive errors on the observed input/output signals. The disturbance is interpreted as a latent variable and the measurement errors, as misfit variables. The estimation problem compensates for the disturbance and the measurement errors by minimizing an appropriately defined cost functional over all trajectories of the system. The solution of the general estimation problem and its extremes, pure misfit and pure latency, lead to one and the same problem — minimization of a quadratic function of the state and the input subject to the state equation.

Acknowledgments

I. Markovsky is a research assistant and J. C. Willems is a visiting professor at the Katholieke Universiteit Leuven, Belgium. Our research is supported by grants from several funding agencies and sources:

Research Council KUL: Concerted Research Action GOA-Mefisto 666 (Mathematical Engineering), IDO (IOTA Oncology, Genetic networks), several PhD/postdoc & fellow grants;

Flemish Government: Fund for Scientific Research Flanders (several PhD/postdoc grants, projects G.0256.97 (subspace), G.0115.01 (bio-i and microarrays), G.0240.99 (multilinear algebra), G.0197.02 (power islands), G.0407.02 (support vector machines), research communities ICCoS, ANMMM), AWI (Bil. Int. Collaboration Hungary/ Poland), IWT (Soft4s (softsensors), STWW-Genprom (gene promotor prediction), GBOU-McKnow (Knowledge management algorithms), Eureka-Impact (MPC-control), Eureka-FLiTE (flutter modeling), several PhD grants);

Belgian Federal Government: DWTC (IUAP IV-02 (1996-2001) and IUAP V-10-29 (2002-2006): Dynamical Systems and Control: Computation, Identification & Modelling), Program Sustainable Development PODO-II (CP-TR-18: Sustainibility effects of Traffic Management Systems);

Direct contract research: Verhaert, Electrabel, Elia, Data4s, IPCOS;

References

[Bro70] R. W. Brockett. *Finite Dimensional Linear Systems.* John Wiley, 1970.

[Wil02] J. C. Willems. Deterministic Kalman filtering. *J. of Econometrics*, to appear, 2002.