Continuous-time Errors-in-variables Filtering

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

ESAT-SCD (SISTA), K.U.Leuven, Belgium



CDC 2002, Las Vegas, 12 December

Outline

- Errors-in-variables model
- EIV smoothing and filtering
- Misfit and latency
- Estimation in the pure misfit case
- Estimation in the pure latency case
- Conclusion

Errors-in-variables model

consider the continuous-time linear state space system

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

let $(u_{\mathrm{d}}, y_{\mathrm{d}})$ be a measured input/output trajectory on the interval $[0, t_f]$

in general, it is not a trajectory of the model, *i.e.*,
there is no $x(0)$, such that the response of (1) to the signal u_{d} is y_{d}

account for modeling errors by adding to the system an auxiliary input d

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

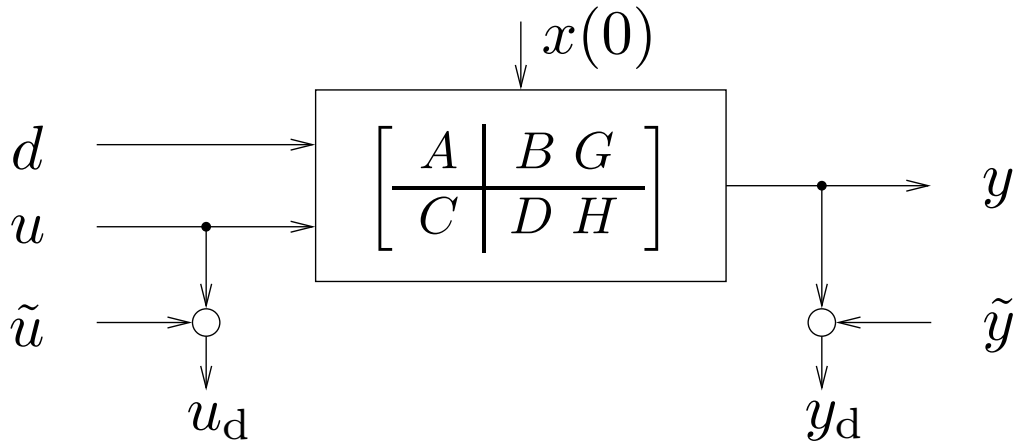
Errors-in-variables model (cont.)

measurement errors are modeled by appending to the system the eqns.

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

the measurement errors $\tilde{u}$, $\tilde{y}$ and the disturbance d are free variables that allow us to "explain" the measured trajectory $(u_{\mathrm{d}}, y_{\mathrm{d}})$

the system (2,3) is called an errors-in-variables model



EIV smoothing and filtering

problem: estimate the "true" I/S/O trajectory from the measured traj.

assumptions:

1. the true I/S/O trajectory is generated by (1)
2. the measurements are generated by the EIV model
3. the initial condition, disturbance, and measurement errors are "small"

estimation principle: find the "smallest" estimated $\hat{x}(0)$, $\hat{d}$, $\hat{\tilde{u}}$, and $\hat{\tilde{y}}$, that "explain" the measurements by the EIV model, *i.e.*,

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + B\hat{u} + G\hat{d}, \qquad \hat{y} = C\hat{x} + D\hat{u} + H\hat{d}$$

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

Misfit versus latency interpretation

$\hat{x}(0)$ and $\hat{d}$ are latent variables, *i.e.*, variables that modify the system equations to explain the model-data mismatch

$\hat{\tilde{u}}$ and $\hat{\tilde{y}}$ are misfit variables, *i.e.*, variables that modify the data to make it match the system equations

pure latency case — consider the data as being error free and blame the model as being imperfect description of the reality

pure misfit case — consider the model as being perfect and blame the measurements as being error corrupted

in general, both misfit and latency can be considered simultaneously, *i.e.*, the uncertainty is in both the data and the model

Estimation cost functional

misfit and latency cost functionals:

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

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

$R > 0$, $Q \geq 0$, $P > 0$, $\Gamma \geq 0$ — weighting matrices, reflecting the relative importance of the terms

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

varying ρ from 0 to 1, allows smooth transition from a purely misfit contribution to a purely latency contribution

Smoothing versus filtering

the general EIV estimation problem, we will consider, is

$$\min_{\hat{u},\hat{y},\hat{x},\hat{d}} \rho J_{\text{misfit}} + (1-\rho) J_{\text{latency}} \quad \text{s.t.} \quad \begin{array}{rcl} \frac{\mathrm{d}}{\mathrm{d}t}\hat{x} & = & A\hat{x} + B\hat{u} + G\hat{d} \\ \hat{y} & = & C\hat{x} + D\hat{u} + H\hat{d} \end{array} \tag{4}$$

as defined this is a smoothing problem, *i.e.*, the estimates are wanted for the whole interval $[0, t_f]$

but we will derive also the solution of the corresponding filtering problem

filtering problem — design a non-anticipating dynamical system that assumes as input the measurements and produces as output at each moment of time the optimal estimate

Common problem

we solve the following optimal control problem

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

where $M(t) := \begin{bmatrix} Q_u & Q_{ux} & q_u(t) \\ * & Q_x & q_x(t) \\ * & * & q(t) \end{bmatrix}$, $Q_u > 0$, $\begin{bmatrix} Q_u & Q_{ux} \\ * & Q_x \end{bmatrix} \geq 0$, and $\Gamma \geq 0$

completion of squares approach

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*

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

Solution of the common problem

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

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

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*

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

Estimation in the pure misfit and pure latency cases

assume $d = 0$ $\Rightarrow$ only the misfit contributes to the cost functional

$$\min_{\hat{u},\hat{y},\hat{x}} \int_0^{t_f} \left(||\hat{u}(t) - u_{\rm d}(t)||_R^2 + ||\hat{y}(t) - y_{\rm d}(t)||_Q^2 \right) {\rm d}t \quad \text{s.t.} \quad \begin{array}{r} \frac{\rm d}{{\rm d}t}\hat{x} = A\hat{x} + B\hat{u} \\ \hat{y} = C\hat{x} + D\hat{u} \end{array}$$

assume $\begin{array}{l} \tilde{u} = 0 \\ \tilde{y} = 0 \end{array}$ $\Rightarrow$ only the latency contributes to the cost functional

$$\min_{\hat{d},\hat{x}} \int_0^{t_f} ||\hat{d}(t)||_P^2 \, {\rm d}t + x^T(0)\Gamma x(0) \quad \text{s.t.} \quad \begin{array}{rcl} \frac{\rm d}{{\rm d}t}\hat{x} & = & A\hat{x} + Bu_{\rm d} + G\hat{d} \\ y_{\rm d} & = & C\hat{x} + Du_{\rm d} + H\hat{d}. \end{array}$$

this corresponds to the deterministic Kalman filter

Theorem 1. [Pure misfit smoothing] *Let* $R > 0$, $Q \geq 0$, *and* $(A, C^T Q C)$ *be observable. The unique solution of the pure misfit estimation problem is*

$$\hat{u} = -N^{-1}\big((B^T K + D^T Q C)\hat{x} + B^T s - D^T Q y_{\rm d} - R u_{\rm d}\big)$$

$$\tfrac{\rm d}{{\rm d}t}\hat{x} = \big(A - BN^{-1}(B^T K + D^T Q C)\big)^T \hat{x} - BN^{-1}B^T s + BN^{-1}D^T Q y_{\rm d} + BN^{-1} R u_{\rm d},\ \hat{x}(t_f) = -K^{-1}(t_f)s(t_f)$$

where $N := R + D^T Q D$, s *is generated by*

$$\tfrac{\rm d}{{\rm d}t}s = -\big(A - BN^{-1}(B^T K + D^T Q C)\big)^T s + C^T Q y_{\rm d} - (B^T K + D^T Q C)^T N^{-1}(D^T Q y_{\rm d} + R u_{\rm d}), \quad s(0) = 0,$$

and K *is generated by*

$$(B^T K + D^T Q C)^T N^{-1}(B^T K + D^T Q C) = \tfrac{\rm d}{{\rm d}t}K + A^T K + K A + C^T Q C,\ K(0) = 0. \quad (6)$$

Corollary 1. [Pure misfit finite-horizon filtering] *Under the assumptions of Theorem 1, the optimal misfit filter is* $\hat{x} = -K^{-1}s$, $\hat{y} = C\hat{x} + D\hat{u}$, *and*

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

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

$N := R + D^TQD$ *and* K *is generated by (6).*

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

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

Its unique negative definite solution K_- should be used (in place of K) in the optimal filter.

Theorem 2. [Pure latency smoothing] *Let $H \in \mathbb{R}^{n_y \times n_d}$ be of full rank with $n_y \leq n_d$, $P > 0$, and $\Gamma > 0$. The unique solution of the pure latency estimation problem is*

$$\hat{u} = -Q_u^{-1}\big((\bar{B}^T K + H_2^T F C)\hat{x} + \bar{B}^T s - H_2^T F(y_{\mathrm{d}} - Du_{\mathrm{d}})\big),$$

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}\hat{x} = \big(\bar{A} - \bar{B}Q_u^{-1}(\bar{B}^T K + H_2^T F C)\big)\hat{x} - \bar{B}Q_u^{-1}\bar{B}^T s \\ + (B - \bar{B}Q_u^{-1}H_2^T F D - G_1 H_1^{-1} D)u_{\mathrm{d}} + (G_1 H_1^{-1} + \bar{B}Q_u^{-1}H_2^T F)y_{\mathrm{d}},\end{aligned}$$

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

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}s = -\big(\bar{A} - \bar{B}Q_u^{-1}(\bar{B}^T K + H_2^T F C)\big)^T s \\ + \big(C^T - (\bar{B}^T K + H_2^T F C)^T Q_u^{-1} H_2^T\big) F(y_{\mathrm{d}} - Du_{\mathrm{d}}),\ \ s(0) = 0,\end{aligned}$$

$$\begin{aligned}(\bar{B}^T K + H_2^T F C)^T Q_u^{-1}(\bar{B}^T K + H_2^T F C) \\ = \tfrac{\mathrm{d}}{\mathrm{d}t}K + \bar{A}^T K + K\bar{A} + C^T F C, \quad K(0) = -\Gamma. \quad (7)\end{aligned}$$

$\bar{A} := A - G_1 H_1^{-1} C$, $\bar{B} := G_2 - G_1 H_1^{-1} H_2$, $Q_u := P + H_2^T F H_2$, $F := H_1^{-T} P H_1^{-1}$

Corollary 2. [Pure latency filtering] *Under the assumptions of Theorem 2, the optimal latency filter is* $x = -K^{-1}s$, $y = C\hat{x} + D\hat{u}$

$$u = -Q_u^{-1}\Big(\big(\bar{B}^T(\bar{B}^T K + H_2^T F C)K\big)s - H_2^T F(y_{\rm d} - Du_{\rm d})\Big),$$

$$\begin{aligned}\tfrac{\rm d}{{\rm d}t}s = &-\big(\bar{A} - \bar{B}Q_u^{-1}(\bar{B}^T K + H_2^T F C)\big)^T s \\ &+ \big(C^T - (\bar{B}^T K + H_2^T F C)Q_u^{-1}H_2^T\big)F(y_{\rm d} - Du_{\rm d}),\ \ s(0) = 0,\end{aligned}$$

where K *is generated by (7), and* $\bar{A}$, $\bar{B}$, Q_u, *and* F *are defined above.*

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

$$\bar{A}^T K + K\bar{A} + C^T F C = (\bar{B}^T K + H_2^T F C)^T Q_u^{-1}(\bar{B}^T K + H_2^T F C).$$

Misfit+latency case

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*

$$\hat{u} = -Q_u^{-1}(Lx + \bar{B}^T s + q_u), \quad \tfrac{\mathrm{d}}{\mathrm{d}t}x = (A - \bar{B}Q_u^{-1}L)x - \bar{B}Q_u^{-1}\bar{B}^T s - \bar{B}Q_u^{-1}q_u,$$

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

$$\tfrac{\mathrm{d}}{\mathrm{d}t}s = -(A - \bar{B}Q_u^{-1}L)^T s + \rho C^T Q y_{\mathrm{d}} + LQ_u^{-1}q_u, \ s(0) = 0,$$

$$0 = \tfrac{\mathrm{d}}{\mathrm{d}t}K + A^T K + KA + \rho C^T QC - L^T Q_u^{-1}L, \quad K(0) = -(1-\rho)\Gamma \tag{8}$$

$$\bar{B} := [B \ \ G], \qquad L := \begin{bmatrix} B^T K + \rho D^T QC \\ G^T K + \rho H^T QC \end{bmatrix}, \qquad q_u := -\rho \begin{bmatrix} Ru_{\mathrm{d}} + D^T Q y_{\mathrm{d}} \\ H^T Q y_{\mathrm{d}} \end{bmatrix}$$

$$Q_u := \rho \begin{bmatrix} R + D^T QD & D^T QH \\ * & \frac{(1-\rho)}{\rho}P + H^T QH \end{bmatrix}.$$

Corollary 3. [Misfit and latency filtering] *Under the assumptions of Theorem 3, the optimal misfit and latency filter is* $x = -K^{-1}s$,

$$\hat{u} = -Q_u^{-1}\big((\bar{B}^T - LK^{-1})s + q_u\big), \quad \hat{y} = -\big(CK^{-1} + DQ_u^{-1}(\bar{B}^T - LK^{-1})\big)s - DQ_u^{-1}q_u,$$

$$\tfrac{\mathrm{d}}{\mathrm{d}t}s = -(A - \bar{B}Q_u^{-1}L)^T s + \rho C^T Q y_{\mathrm{d}} + LQ_u^{-1}q_u, \;\; s(0) = 0,$$

where K *is generated by (8) and* $\bar{B}$, Q_u, L, q_u *are defined above.*

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

$$0 = A^T K + KA + \rho C^T QC - L^T Q_u^{-1} L.$$

Conclusion

- we posed an estimation problem for continuous-time LTI EIV model

  (similar results are available for the discrete-time case)

- the disturbance is interpreted as a latent variable and the measurement errors as misfit variables

- the estimation problem compensates for disturbance and 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