Optimal State Estimators for Linear Systems with Unknown Inputs

Shreyas Sundaram and Christoforos N. Hadjicostis

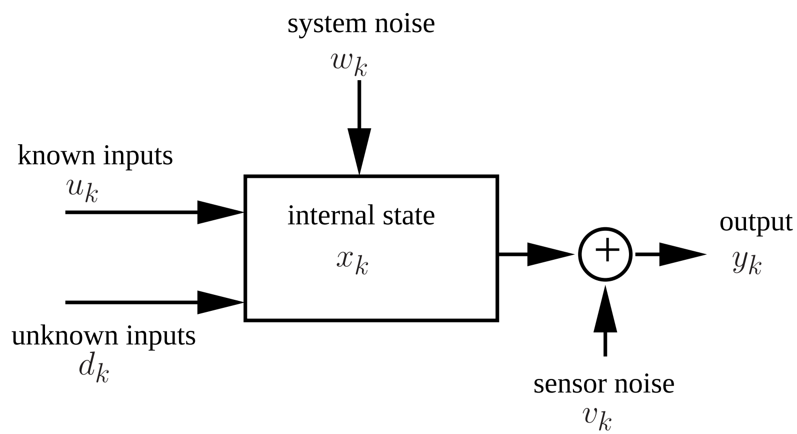
Coordinated Science Laboratory and Department of Electrical and Computer Engineering

University of Illinois at Urbana-Champaign

Overview

Many physical systems can be modeled as having:

- Known inputs: u_k (Control signals, etc.)
- Unknown inputs: d_k (Disturbances, faults, modeling uncertainties, etc.)
- System and Sensor Noise: w_k, v_k
  - $E[w_k] = 0$, $E[v_k] = 0$, $E[w_k w_j^T] = Q_k \delta_{kj}$, $E[v_k v_j^T] = R_k \delta_{kj}$
- Internal state: $x_{k+1} = Ax_k + Bd_k + Fu_k + w_k$
- Outputs: $y_k = Cx_k + Dd_k + Gu_k + v_k$



Objective: Produce an optimal estimate of the state from the outputs

Note: Known inputs are easily handled, so we will drop them in rest of discussion

Previous Work

State estimation in stochastic linear systems with unknown inputs has been studied over past two decades

- Kitanidis, Hou, Patton, Saberi, Darouach, Zasadzinski, Boutayeb, Nikoukhah, ...
- These investigations typically focus on **zero-delay** estimators
  - i.e., use $y_0, y_1, \ldots, y_k$ to estimate x_k
- Existence conditions for such estimators are quite strict
- Conditions can be relaxed by allowing delayed estimation
  - i.e., estimate x_k from $y_0, y_1, \ldots, y_k, y_{k+1}, \ldots, y_{k+\alpha}$
  - Jin and Tahk (2005): Considered delayed estimators, but estimate may not be optimal
  - Saberi, Stoorvogel and Sannuti (2000): Provided a geometric analysis, and used techniques from H_2-optimal control to design estimator

Contribution: We present an algebraic design procedure to construct optimal delayed estimators for stochastic linear systems with unknown inputs

Delayed Outputs

What information is provided by the output of the system over $\alpha + 1$ time-steps?

- System equations:

$$\begin{aligned} x_{k+1} &= Ax_k + Bd_k + w_k \\ y_k &= Cx_k + Dd_k + v_k \end{aligned}$$

- Output of system over $\alpha + 1$ time-steps ($\alpha \in \mathbb{N}$):

$$\underbrace{\begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+\alpha} \end{bmatrix}}_{\mathbf{y}_{k:k+\alpha}} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\alpha} \end{bmatrix}}_{\Theta_\alpha} x_k + \underbrace{\begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-1}B & CA^{\alpha-2}B & \cdots & D \end{bmatrix}}_{M_\alpha} \underbrace{\begin{bmatrix} d_k \\ d_{k+1} \\ \vdots \\ d_{k+\alpha} \end{bmatrix}}_{\mathbf{d}_{k:k+\alpha}}$$

$$+ \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ C & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-1} & CA^{\alpha-2} & \cdots & C \end{bmatrix}}_{M_{w,\alpha}} \underbrace{\begin{bmatrix} w_k \\ w_{k+1} \\ \vdots \\ w_{k+\alpha-1} \end{bmatrix}}_{\mathbf{w}_{k:k+\alpha-1}} + \underbrace{\begin{bmatrix} v_k \\ v_{k+1} \\ \vdots \\ v_{k+\alpha} \end{bmatrix}}_{\mathbf{v}_{k:k+\alpha}}$$

Colored Noise

- Use delayed outputs as "new" output of system:

$$x_{k+1} = Ax_k + Bd_k + w_k$$

$$\mathbf{y}_{k:k+\alpha} = \Theta_\alpha x_k + M_\alpha \mathbf{d}_{k:k+\alpha} + M_{w,\alpha}\mathbf{w}_{k:k+\alpha-1} + \mathbf{v}_{k:k+\alpha}$$

- However, $\mathbf{v}_{k:k+\alpha} \equiv \begin{bmatrix} v_k^T & v_{k+1}^T & \cdots & v_{k+\alpha}^T \end{bmatrix}^T$
  - $\mathbf{v}_{k:k+\alpha}, \mathbf{v}_{k+1:k+1+\alpha}, \ldots, \mathbf{v}_{k+\alpha:k+2\alpha}$ are correlated
  - Also true for $\mathbf{w}_{k:k+\alpha-1}, \mathbf{w}_{k+1:k+\alpha}, \ldots, \mathbf{w}_{k+\alpha-1:k+2(\alpha-1)}$
- In other words, system is affected by **colored** noise
- Increase dimension of system model in order to handle colored noise [Anderson & Moore, 1979]

Augmented System

- Original system equations:

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bd_k + w_k \\
y_k &= Cx_k + Dd_k + v_k
\end{aligned}
$$

- Rewrite system as

$$
\underbrace{\left[\begin{array}{c} x_{k+1} \\ w_{k+1} \\ \vdots \\ w_{k+\alpha-2} \\ w_{k+\alpha-1} \\ \hline v_{k+1} \\ v_{k+2} \\ \vdots \\ v_{k+\alpha-1} \\ v_{k+\alpha} \end{array}\right]}_{\bar{x}_{k+1}}
=
\underbrace{\left[\begin{array}{ccccc|ccccc}
A & I & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & I & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & I & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\hline
0 & 0 & 0 & \cdots & 0 & 0 & I & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & I & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & I \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0
\end{array}\right]}_{\bar{A}}
\underbrace{\left[\begin{array}{c} x_k \\ w_k \\ w_{k+1} \\ \vdots \\ w_{k+\alpha-2} \\ \hline v_k \\ v_{k+1} \\ \vdots \\ v_{k+\alpha-2} \\ v_{k+\alpha-1} \end{array}\right]}_{\bar{x}_k}
+
\underbrace{\left[\begin{array}{c} B \\ 0 \\ \vdots \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{array}\right]}_{\bar{B}}
d_k +
\underbrace{\left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ I & 0 \\ \hline 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & I \end{array}\right]}_{\bar{B}_n}
\underbrace{\left[\begin{array}{c} w_{k+\alpha-1} \\ v_{k+\alpha} \end{array}\right]}_{n_k}
$$

$$
y_k = \left[\begin{array}{ccccc|ccccc} C & 0 & 0 & \cdots & 0 & I & 0 & 0 & \cdots & 0 \end{array}\right] \bar{x}_k + Dd_k
$$

Optimal Estimator

- Augmented state equation: $\bar{x}_{k+1} = \bar{A}\bar{x}_k + \bar{B}d_k + \bar{B}_n n_k$
- Output of augmented system over $\alpha + 1$ time-steps is:

$$\mathbf{y}_{k:k+\alpha} = \underbrace{\left[\begin{array}{cccc|cccc} C & 0 & \cdots & 0 & I & 0 & \cdots & 0 \\ CA & C & \cdots & 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-1} & CA^{\alpha-2} & \cdots & C & 0 & 0 & \cdots & I \\ CA^{\alpha} & CA^{\alpha-1} & \cdots & CA & 0 & 0 & \cdots & 0 \end{array}\right]}_{\bar{\Theta}_\alpha} \bar{x}_k + M_\alpha \mathbf{d}_{k:k+\alpha} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ C & I \end{bmatrix}}_{M_{n,\alpha}} n_k$$

- Note: noise n_k is no longer colored
- Consider estimator of the form $\hat{x}_{k+1} = \bar{A}\hat{x}_k + K_k(\mathbf{y}_{k:k+\alpha} - \bar{\Theta}_\alpha \hat{x}_k)$
  - Estimation error: $e_k \equiv \hat{x}_k - \bar{x}_k$
  - Optimality: Estimator gain K_k must be chosen so that
    - Estimator is unbiased: $E[e_k] = 0$ for all k
    - Trace of error covariance matrix $E[e_k e_k^T]$ is minimized

Strategy: Examine estimation error and choose K_k to achieve above objectives

Unbiased Estimation: $E[e_k] = 0$

- Estimation error:
$$\begin{aligned} e_{k+1} &\equiv \hat{x}_{k+1} - \bar{x}_{k+1} \\ &= \left(\bar{A} - K_k \bar{\Theta}_\alpha\right) e_k + \left(K_k M_\alpha - \begin{bmatrix} \bar{B} & 0 & \cdots & 0 \end{bmatrix}\right) \mathbf{d}_{k:k+\alpha} \\ &\quad + \left(K_k M_{n,\alpha} - \bar{B}_n\right) n_k \end{aligned}$$

- Unbiased estimation: $K_k M_\alpha - \begin{bmatrix} \bar{B} & 0 & \cdots & 0 \end{bmatrix} = 0$

Theorem: There exists a matrix K_k satisfying above equation if and only if

$$\text{rank}\,[M_\alpha] - \text{rank}\,[M_{\alpha-1}] = \text{rank} \begin{bmatrix} B \\ D \end{bmatrix}$$

- This is the Massey-Sain condition for system inversion with delay α (1969)
  - We must invert the inputs in order to obtain unbiased estimation
- The larger the delay, the better the chance of satisfying the condition
- Upper bound on inversion delay provided by Willsky (1974) as $\alpha = n - \text{nullity}[D] + 1$

Next step: Choose K_k to obtain both unbiased and minimum variance estimation

Parameterizing the Gain

Idea: Use some portion of K_k to obtain unbiased estimation, and use remaining freedom to minimize trace of $E[e_k e_k^T]$

- Unbiased estimation: $K_k M_\alpha = \begin{bmatrix} \bar{B} & 0 & \cdots & 0 \end{bmatrix}$
- We show that there exists a matrix $\mathcal{N}$ such that $K_k = \begin{bmatrix} L_k & \bar{B} \end{bmatrix} \mathcal{N}$
  - L_k is remaining freedom in K_k after decoupling unknown inputs
- Substitute parameterization back into error expression to get

$$\begin{aligned} e_{k+1} &= \left(\bar{A} - K_k \bar{\Theta}_\alpha\right) e_k + (K_k M_{n,\alpha} - \bar{B}_n) n_k \\ &\equiv (\mathcal{A} - L_k \Phi)\, e_k + (\mathcal{B} + L_k \Psi)\, n_k \end{aligned}$$

for some constant matrices $\mathcal{A}$, $\mathcal{B}$, Φ, Ψ

Next step: Choose L_k to minimize trace of error covariance matrix

Minimum Variance Estimation

- Estimation error: $e_{k+1} = (\mathcal{A} - L_k\Phi)\, e_k + (\mathcal{B} + L_k\Psi)\, n_k$
- Define $\Pi_k \equiv E[n_k n_k^T] = \begin{bmatrix} Q_{k+\alpha-1} & 0 \\ 0 & R_{k+\alpha} \end{bmatrix}$
- Error covariance matrix:

$$\begin{aligned} \Sigma_{k+1} &\equiv E[e_{k+1}e_{k+1}^T] \\ &= (\mathcal{A} - L_k\Phi)\, \Sigma_k\, (\mathcal{A} - L_k\Phi)^T + (\mathcal{B} + L_k\Psi)\, \Pi_k\, (\mathcal{B} + L_k\Psi)^T \end{aligned}$$

- To minimize trace of Σ_{k+1}, take gradient of above expression w.r.t. L_k and set equal to zero:

$$L_k = \left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)\left(\Phi\Sigma_k\Phi^T + \Psi\Pi_k\Psi^T\right)^{-1}$$

- Substitute optimal L_k into covariance update equation:

$$\Sigma_{k+1} = \mathcal{A}\Sigma_k\mathcal{A}^T + \mathcal{B}\Pi_k\mathcal{B}^T - \left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)\left(\Phi\Sigma_k\Phi^T + \Psi\Pi_k\Psi^T\right)^{-1}\left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)^T$$

- Obtain optimal estimator gain as $K_k = \begin{bmatrix} L_k & \bar{B} \end{bmatrix} \mathcal{N}$

Example (1)

Consider the system

$$x_{k+1} = \underbrace{\begin{bmatrix} 0.1 & 1 \\ 0 & 0.2 \end{bmatrix}}_{A} x_k + \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{B} d_k + w_k$$

$$y_k = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}}_{C} x_k + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}_{D} d_k + v_k$$

$$Q_k \equiv E[w_k w_k^T] = 0.01 I_2, \quad R_k \equiv E[v_k v_k^T] = 0.04 I_2$$

Step 1: Find minimum delay α satisfying $\text{rank}[M_\alpha] - \text{rank}[M_{\alpha-1}] = \text{rank}\begin{bmatrix} B \\ D \end{bmatrix}$

- $\alpha = 0$: $\text{rank}[M_0] = \text{rank}[D] = 1 \neq \text{rank}\begin{bmatrix} B \\ D \end{bmatrix}$
- $\alpha = 1$: $\text{rank}[M_1] - \text{rank}[M_0] = \text{rank}\begin{bmatrix} D & 0 \\ CB & D \end{bmatrix} - 1 = 1 \neq \text{rank}\begin{bmatrix} B \\ D \end{bmatrix}$
- $\alpha = 2$: $\text{rank}[M_2] - \text{rank}[M_1] = \text{rank}\begin{bmatrix} D & 0 & 0 \\ CB & D & 0 \\ CAB & CB & D \end{bmatrix} - 2 = 2 = \text{rank}\begin{bmatrix} B \\ D \end{bmatrix}$ $\surd$

Example (2)

Step 2: Form augmented system

$$\underbrace{\begin{bmatrix} x_{k+1} \\ w_{k+1} \\ v_{k+1} \\ v_{k+2} \end{bmatrix}}_{\bar{x}_{k+1}} = \underbrace{\begin{bmatrix} A & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\bar{A}} \underbrace{\begin{bmatrix} x_k \\ w_k \\ v_k \\ v_{k+1} \end{bmatrix}}_{\bar{x}_k} + \underbrace{\begin{bmatrix} B \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\bar{B}} d_k + \underbrace{\begin{bmatrix} 0 & 0 \\ I & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix}}_{\bar{B}_n} \underbrace{\begin{bmatrix} w_{k+1} \\ v_{k+2} \end{bmatrix}}_{n_k}$$

$$\underbrace{\begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}}_{\mathbf{y}_{k:k+2}} = \underbrace{\begin{bmatrix} C & 0 & I & 0 \\ CA & C & 0 & I \\ CA^2 & CA & 0 & 0 \end{bmatrix}}_{\bar{\Theta}_2} \bar{x}_k + \underbrace{\begin{bmatrix} D & 0 & 0 \\ CB & D & 0 \\ CAB & CB & D \end{bmatrix}}_{M_2} \mathbf{d}_{k:k+2} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ C & I \end{bmatrix}}_{M_{n,2}} n_k$$

Step 3: Parameterize K_k to solve $K_k M_2 = \begin{bmatrix} \bar{B} & 0 & 0 \end{bmatrix}$

$$K_k = \begin{bmatrix} L_k & \bar{B} \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1.8 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}}_{\mathcal{N}}$$

Example (3)

Step 4: Substitute parameterization into error expression

$$\begin{aligned} e_{k+1} &= \left(\bar{A} - K_k \bar{\Theta}_2\right) e_k + (K_k M_{n,2} - \bar{B}_n) n_k \\ &\equiv (\mathcal{A} - L_k \Phi)\, e_k + (\mathcal{B} + L_k \Psi)\, n_k \end{aligned}$$

Step 5: Calculate optimal gain at each time-step k:

$$L_k = \left(\mathcal{A}\Sigma_k \Phi^T - \mathcal{B}\Pi_k \Psi^T\right)\left(\Phi\Sigma_k\Phi^T + \Psi\Pi_k\Psi^T\right)^{-1}$$

$$\Sigma_{k+1} = \mathcal{A}\Sigma_k\mathcal{A}^T + \mathcal{B}\Pi_k\mathcal{B}^T - \left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)\left(\Phi\Sigma_k\Phi^T + \Psi\Pi_k\Psi^T\right)^{-1}\left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)^T$$

$$K_k = \begin{bmatrix} L_k & \bar{B} \end{bmatrix} \mathcal{N}$$

Step 6: Optimal estimator for $\bar{x}_k$:

$$\hat{x}_{k+1} = \bar{A}\hat{x}_k + K_k\left(\mathbf{y}_{k:k+\alpha} - \bar{\Theta}_\alpha \hat{x}_k\right)$$

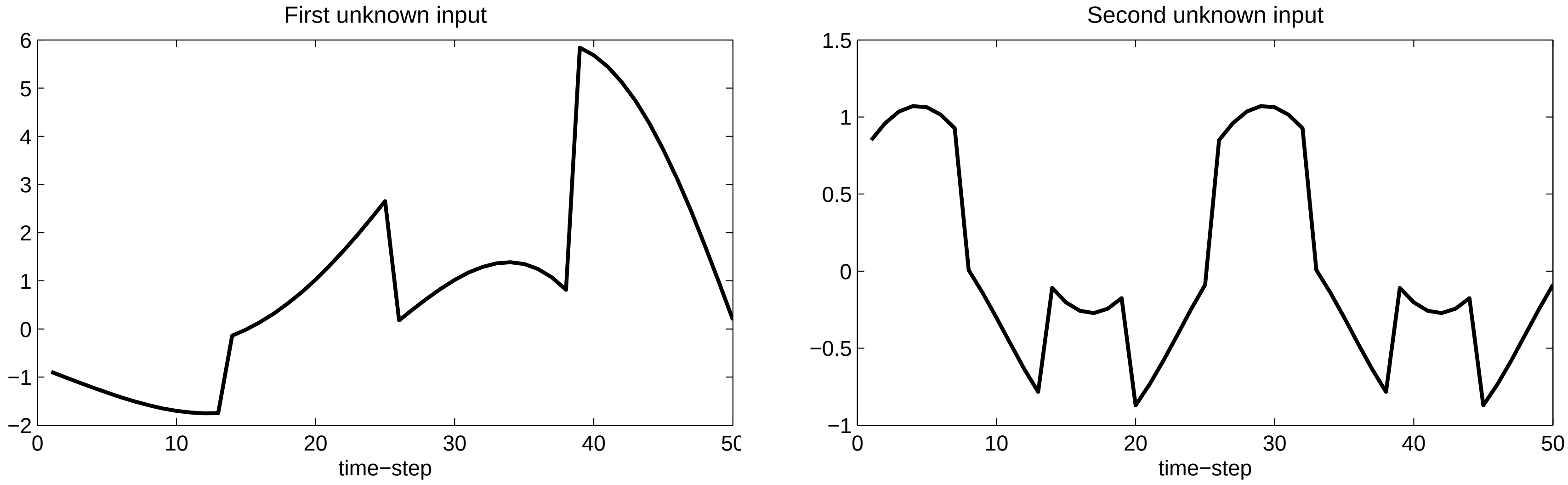
Step 7: Obtain estimate of x_k as $\begin{bmatrix} I_2 & 0 \end{bmatrix} \hat{x}_k$

Example: Simulation (1)

- Suppose initial system state has $E[x_0] = 0$, $E[x_0 x_0^T] = I$
- Initial augmented state $\bar{x}_0 = \begin{bmatrix} x_0^T & w_0^T & v_0^T & v_1^T \end{bmatrix}^T$ has

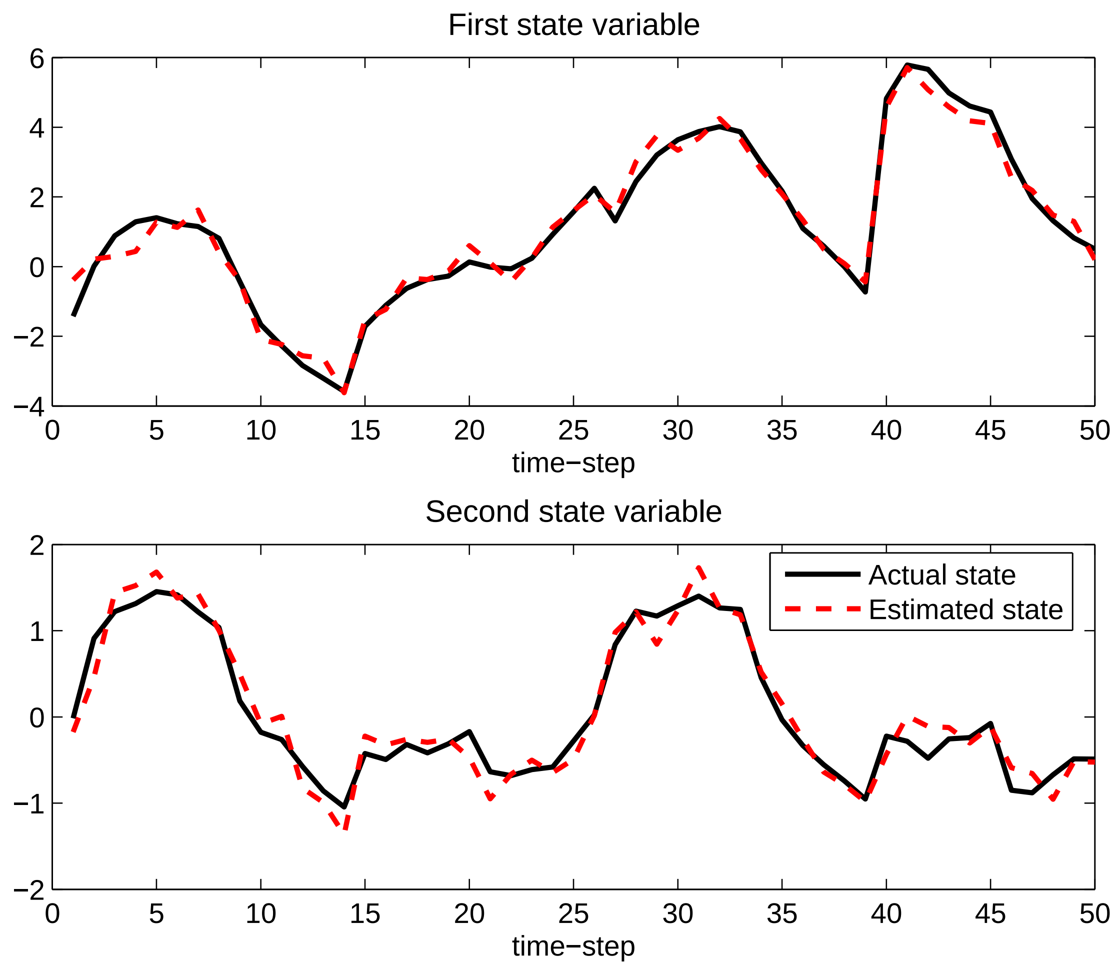
$$E[\bar{x}_0] = 0, \quad E[\bar{x}_0 \bar{x}_0^T] = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & Q_0 & 0 & 0 \\ 0 & 0 & R_0 & 0 \\ 0 & 0 & 0 & R_1 \end{bmatrix}$$

- Initialize estimator with $\hat{x}_0 = E[\bar{x}_0] = 0$, $\Sigma_0 = E[\bar{x}_0 \bar{x}_0^T]$
- Unknown inputs to system:



Example: Simulation (2)

State estimates:



Note: Estimated state should be delayed by $\alpha = 2$ time-steps, but it is shifted forward for purposes of comparison

Summary of Design Procedure

1. Find smallest α such that $\text{rank}[M_\alpha] - \text{rank}[M_{\alpha-1}] = \text{rank}\begin{bmatrix} B \\ D \end{bmatrix}$
2. Form augmented system: $\bar{x}_{k+1} = \bar{A}\bar{x}_k + \bar{B}d_k + \bar{B}_n n_k$
3. Parameterize K_k as $K_k = \begin{bmatrix} L_k & \bar{B} \end{bmatrix} \mathcal{N}$ to solve equation $K_k M_\alpha = \begin{bmatrix} \bar{B} & 0 & \cdots & 0 \end{bmatrix}$
4. Substitute parameterization into error expression

$$\begin{aligned} e_{k+1} &= \left(\bar{A} - K_k \bar{\Theta}_\alpha\right) e_k + (K_k M_{n,\alpha} - \bar{B}_n) n_k \\ &\equiv (\mathcal{A} - L_k \Phi)\, e_k + (\mathcal{B} + L_k \Psi)\, n_k \end{aligned}$$

5. Calculate optimal gain:

$$\begin{aligned} L_k &= \left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)\left(\Phi\Sigma_k\Phi^T + \Psi\Pi_k\Psi^T\right)^{-1} \\ \Sigma_{k+1} &= \mathcal{A}\Sigma_k\mathcal{A}^T + \mathcal{B}\Pi_k\mathcal{B}^T - \left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)\left(\Phi\Sigma_k\Phi^T + \Psi\Pi_k\Psi^T\right)^{-1}\left(\mathcal{A}\Sigma_k\Phi^T - \mathcal{B}\Pi_k\Psi^T\right)^T \\ K_k &= \begin{bmatrix} L_k & \bar{B} \end{bmatrix} \mathcal{N} \end{aligned}$$

6. Optimal estimator for $\bar{x}_k$: $\hat{x}_{k+1} = \bar{A}\hat{x}_k + K_k\left(\mathbf{y}_{k:k+\alpha} - \bar{\Theta}_\alpha \hat{x}_k\right)$
7. Obtain optimal estimate of x_k from estimate of $\bar{x}_k$

Conclusions and Future Work

- Provided a design procedure for optimal delayed estimators for stochastic linear systems with unknown inputs
  - System must be invertible for unbiased estimation
    - Characterized minimum delay for unbiased estimation
  - Dimension of estimator was increased to handle colored noise induced by delays
  - Decoupled unknown inputs from estimation error, and used remaining freedom in gain matrix to minimize mean square error

- Future work:
  - Analyze convergence and stability of estimator
  - Study minimum dimension estimators
  - Prove global optimality of estimator (i.e., is it optimal over all linear estimators?)