Applications of Moving Horizon Estimation

Eric Haseltine

Department of Chemical Engineering, University of Wisconsin - Madison

1 Motivation

Often through the course of process modeling, additional information to the system dynamics and statistics can be incorporated into the state estimate. This information takes the form of equality and inequality constraints. Moving horizon estimation (MHE) provides one option to address this information.

2 MHE Overview

On-line optimization strategies estimate the state of discrete systems in the form:

$$x_{k+1} = f(x_k, u_k) + Gw_k$$
$$y_k = h(x_k) + v_k$$

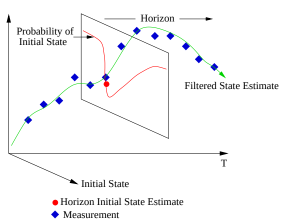
The full-information optimization solves:

$$Z_T(z) = \min_{x_o, \{w_k\}_{k=0}^{T-1}} \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \Gamma_o(x_o)$$
$$s.t. x \in \mathbf{X}, \mathbf{w} \in \mathbf{W}$$

As more data come on-line, the problem increases in size. MHE overcomes computational limitations by formulating the problem over a fixed size estimation horizon. The constrained MHE optimization is:

$$\min_{x_{T-N}, \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + Z_{T-N}(z)$$
$$s.t. \quad x \in \mathbf{X}, \mathbf{w} \in \mathbf{W}$$

The arrival cost, $Z_{T-N}(z)$, summarizes the past information up to the observer horizon. One way to account for this information is to penalize deviations from the *a priori* initial state in the horizon.



For linear systems, the Kalman filtering problem can be reconstructed with the unconstrained objective function. The effect of past data is incorporated by penalizing deviations from the *a priori* initial state in the horizon with the solution to the discrete time Riccati equation:

$$\Gamma(x_{T-N}) = (x_{T-N} - \hat{x}_{T-N})^T \Pi_{T-N}^{-1} (x_{T-N} - \hat{x}_{T-N})$$

3 Motivating Examples

Constrained Mole Fraction Example

Mole fractions provide an excellent motivating example for the need to apply constraints. From a physical standpoint, the mole fractions sum to one, and individual mole fractions fall within [0,1].

Consider a batch reactor with the reaction system:

$$A \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} B \overset{k_2}{\rightarrow} C$$

$$\dot{x} = \begin{bmatrix} -k_1 & k_{-1} & 0 \\ k_1 & -(k_{-1}+k_2) & 0 \\ 0 & k_2 & 0 \end{bmatrix} x, \quad x = \begin{bmatrix} x_A \\ x_B \\ x_C \end{bmatrix}$$

- All reactions are first order
- Total number of moles remains constant
- **Goal:** Maximize conversion of B while not exceeding “limits” for degradation product C
- **Plan:** Only A and B are measurable on-line, so observe C using the mole fraction balance

Kalman Filter Approach

- Complete system is not observable as a standard Kalman filtering problem
- Solution:
  1. Only estimate x_A and x_B
  2. Use the mole balance constraint to determine x_C: $x_C = 1 - x_A - x_B$

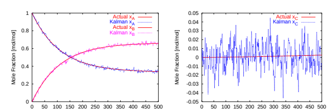


Figure 1: Kalman Filter Estimation

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 0.01^2 & \\ & 0.01^2 \end{bmatrix}, R = \begin{bmatrix} 0.02^2 & \\ & 0.02^2 \end{bmatrix}$$

- Estimate of x_C is poor with large variance

Descriptor Kalman Filter Approach

- Nikoukhah *et. al.* (1999) propose a general maximum likelihood estimate that incorporates “future” dynamics into present state values:

$$G\nu(k) + L\eta(k) = E\xi(k+1) - F\xi(k) + H\omega(k)$$
$$\mu = K\xi(0) + M\zeta$$

  * ν has only past data available
  * η is known for all times
  * ω is a set of independent $\mathcal{N}(0,1)$ vectors
  * ζ is a $\mathcal{N}(0,1)$ vector independent of ω

Descriptor Kalman Filter Approach, cont.

- This method solves problems of the form:

$$Dx_{k+1} = Ax_k + Bu_k + Gw_k$$
$$y_k = Cx_k + v_k$$

- Kalman filtering can be formulated in this way:

$$G\nu_k + L\eta_k = E\xi_{k+1} - F\xi_k + H\omega_k$$
$$\begin{bmatrix} 0 \\ I \end{bmatrix} y_{k+1} + \begin{bmatrix} -I \\ 0 \end{bmatrix} u_k =$$
$$\begin{bmatrix} -I \\ C \end{bmatrix} x_{k+1} - \begin{bmatrix} -A \\ 0 \end{bmatrix} x_k + \begin{bmatrix} Q^{0.5} & \\ & R^{0.5} \end{bmatrix} \begin{bmatrix} w_k \\ r_{k+1} \end{bmatrix}$$

$$\mu = K\xi_0 + M\zeta$$
$$\begin{bmatrix} x(0) \\ y_0 \end{bmatrix} = \begin{bmatrix} I \\ C \end{bmatrix} x_0 + \begin{bmatrix} P^{0.5} & \\ & R^{0.5} \end{bmatrix} \zeta$$

- Using a stochastic shuffle algorithm, the problem becomes regular and solvable recursively
- Now we can specify that x_C has smaller state fluctuations than x_A and x_B

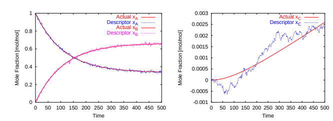


Figure 2: Descriptor Kalman Filter Estimation

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.01^2 & & \\ & 0.01^2 & \\ & & 0.0001^2 \end{bmatrix}, \quad R = \begin{bmatrix} 0.02^2 & \\ & 0.02^2 \end{bmatrix}$$

- Estimate of x_C is significantly improved but negative mole fractions are possible

MHE Approach

- Constrained optimization:
  1. All mole fractions $\in [0,1]$
  2. Sum of mole fractions is 1
- Penalize initial state with descriptor covariance

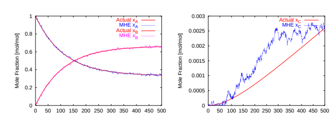
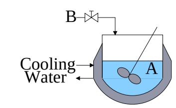


Figure 3: Constrained MHE

- Estimate of x_C obeys model assumptions
- Estimate eventually converges to the descriptor Kalman filter results

Calorimetry Example

Consider the semibatch, exothermic reaction:



$$A + 2B \rightarrow C$$

- Use calorimetry to estimate the rate of reaction

$$\frac{dV}{dt} = F$$
$$\frac{dC_A}{dt} = r - \frac{F}{V} C_A$$
$$\frac{dC_B}{dt} = 2r - \frac{F}{V}(C_{Bf} - C_B)$$
$$\frac{dT}{dt} = \frac{\Delta H_r}{\rho C_p} r + \frac{F}{V}(T_f - T) + \frac{UA}{\rho C_p V}(T_c - T)$$
$$\frac{dr}{dt} = 0$$

- $\mathcal{N}(0,1)$ noise affects temperature measurements
- Use the following constraints to overcome unobservability problems:
  1. Reaction rate constants are nonnegative
  2. Concentrations are nonnegative

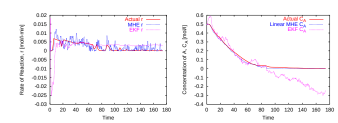


Figure 4: MHE and EKF Comparison

4 Concluding Remarks

As seen through the motivating examples, appropriate use of MHE can significantly improve the state estimate over conventional techniques, such as the Kalman filter. Potential situations to use constrained MHE include:

1. Implementation of inequality and equality constraints
   - Bounds on parameters
   - Material and energy balances
   - Emissions
2. Means to overcome lack of observability
3. Method for handling multi-rate systems
4. Smoothing strategy