Nonlinear Stochastic Modeling and State Estimation of Weakly Observable Systems: Application to Industrial Polymerization Processes

Fernando V. Lima†, James B. Rawlings† and Tyler A. Soderstrom*

†Department of Chemical and Biological Engineering
University of Wisconsin–Madison
*Core Process Control, ExxonMobil Chemical Company, Baytown, Texas

AIChE Annual Meeting
Nashville, TN
November 10, 2009
1. Introduction to State Estimation

2. Objectives and Challenges

3. Design Method for Nonlinear State Estimation
   - Background on Steps
   - Ethylene Copolymerization Example (Literature)

4. Weakly Observable Systems
   - Filtering Approach
   - Case Studies: Polymerization Processes
     - Ethylene Copolymerization Revisited
     - Industrial Ethylene Copolymerization (ExxonMobil)

5. Conclusions and Future Work

6. References
Introduction to State Estimation

- Definition: estimate the system states \( (x) \) from measurements \( (y) \)

Why estimate \( x \)?
- \( x \) is required to model the system
- Performance of an advanced feedback control system is directly affected by the quality of state estimates
- Most complex product properties are not measurable
- They must be inferred from other measurements combined with nonlinear property models
- \( x \) is corrupted with process noise \( (w) \) and \( y \) with sensor noise \( (v) \)

Challenge of state estimation
- Determine good estimates in the face of noisy and incomplete output measurements
Definition: estimate the system states ($x$) from measurements ($y$)

Why estimate $x$?
- $x$ is required to model the system
- performance of an advanced feedback control system is directly affected by the quality of state estimates
- most complex product properties are not measurable
  - they must be inferred from other measurements combined with nonlinear property models
- $x$ is corrupted with process noise ($w$) and $y$ with sensor noise ($v$)
Definition: estimate the system states ($x$) from measurements ($y$)

Why estimate $x$?
- $x$ is required to model the system
- Performance of an advanced feedback control system is directly affected by the quality of state estimates
- Most complex product properties are not measurable
  - They must be inferred from other measurements combined with nonlinear property models
- $x$ is corrupted with process noise ($w$) and $y$ with sensor noise ($v$)

Challenge of state estimation
- Determine good estimates in the face of noisy and incomplete output measurements
Moving Horizon Estimator (MHE)

- optimization-based state estimator
- robust to process disturbances and model errors
- naturally handles nonlinear models
- incorporates process constraints in its formulation

To provide high-quality estimates, MHE requires

- accurate dynamic model of the process
- knowledge on noise statistics
  - for zero-mean Gaussian sequences, covariances specify their statistics
  - Autocovariance Least-Squares (ALS) technique estimates these covariances from data
Moving Horizon Estimator (MHE)

- optimization-based state estimator
- robust to process disturbances and model errors
- naturally handles nonlinear models
- incorporates process constraints in its formulation

To provide high-quality estimates, MHE requires

- accurate dynamic model of the process
- knowledge on noise statistics
  - affecting the states and measurements
  - for zero-mean Gaussian sequences, covariances specify their statistics
  - Autocovariance Least-Squares (ALS) technique estimates these covariances from data
Objectives

- Apply a previously proposed design method for nonlinear state estimation\(^1\) to build and validate state estimators for industrial polymerization processes

\(^1\)Lima, Rawlings, and Soderstrom (2008); Lima and Rawlings (2009)
Objectives

- Apply a previously proposed design method for nonlinear state estimation \(^1\) to build and validate state estimators for industrial polymerization processes

- This design method consists of the following steps
  - nonlinear stochastic modeling
  - covariance estimation from operating data
  - estimator selection and implementation

\(^1\)Lima et al. (2008); Lima and Rawlings (2009)
Objectives

- Apply a previously proposed design method for nonlinear state estimation \(^1\) to build and validate state estimators for industrial polymerization processes

- This design method consists of the following steps
  - nonlinear stochastic modeling
  - covariance estimation from operating data
  - estimator selection and implementation

- Two processes are of particular interest
  - gas-phase ethylene copolymerization process (Literature)
  - industrial gas-phase ethylene copolymerization process (ExxonMobil)

\(^1\) Lima et al. (2008); Lima and Rawlings (2009)
Polymerization process models

- have unobservable and weakly observable modes
- are nonlinear and large dimensional (around 50 states and 20 measurements)
- product properties of interest are complex nonlinear functions of states
- stochastic disturbance structure is unknown a priori
- may be described by differential algebraic equations
Polymerization process models

- have unobservable and weakly observable modes
- are nonlinear and large dimensional (around 50 states and 20 measurements)
- product properties of interest are complex nonlinear functions of states
- stochastic disturbance structure is unknown a priori
- may be described by differential algebraic equations

In addition, industrial polymerization processes

- have asynchronous and infrequent laboratory measurements
- are operated over a wide range of conditions
Modeling approach relies on the combination of

- a nonlinear deterministic model obtained from the integration of a first principles model

\[
\frac{dx}{dt} = f(x, u) \Rightarrow x_{k+1} = F(x_k, u_k)
\]

- a stochastic component estimated from process operating data
  - discrete and available every \(\Delta_k = t_{k+1} - t_k\)
  - that provides a typical sample of measurement and process disturbances affecting the system

\[
y_k = h(x_k) + v_k
\]
Nonlinear Stochastic Modeling

- Modeling approach relies on the combination of
  - a nonlinear deterministic model obtained from the integration of a first principles model
    \[
    \frac{dx}{dt} = f(x, u) \Rightarrow x_{k+1} = F(x_k, u_k)
    \]
  - a stochastic component estimated from process operating data
    - discrete and available every $\Delta_k = t_{k+1} - t_k$
    - that provides a typical sample of measurement and process disturbances affecting the system

- Resulting nonlinear, stochastic model in discrete time (DT)
  \[
  x_{k+1} = F(x_k, u_k) + G(x_k)w_k \\
  y_k = h(x_k) + v_k
  \]

  in which $w_k \sim N(0, Q)$, $v_k \sim N(0, R)$
Consider the Nonlinear Model

\[ x_{k+1} = F(x_k, u_k) + G(x_k)w_k \]
\[ y_k = h(x_k) + v_k \]

- Noise \( w_k \) affects all the states
- Noise \( v_k \) corrupts the measurements

Idea of Autocovariances

- The state noise \( w_k \) gets propagated in time
- The measurement noise \( v_k \) appears only at the sampling times and is not propagated in time
- Taking autocovariances of data at different time lags gives covariances of \( w_k \) and \( v_k \)
Autocovariance Least-Squares (ALS) Technique

- ALS estimates covariances of system disturbances \((Q, R)\) from data \(^2\)

---

\(^2\)Odelson, Rajamani, and Rawlings (2006); Rajamani and Rawlings (2009); Rajamani, Rawlings, and Soderstrom (2007)

\(^3\)Rajamani et al. (2007)
Autocovariance Least-Squares (ALS) Technique

- ALS estimates covariances of system disturbances ($Q, R$) from data.
- For nonlinear systems, linear time-varying approach was developed.

---

2 Odelson et al. (2006); Rajamani and Rawlings (2009); Rajamani et al. (2007)
3 Rajamani et al. (2007)
Autocovariance Least-Squares (ALS) Technique

- ALS estimates covariances of system disturbances \((Q, R)\) from data \(^2\)
- For nonlinear systems, linear time-varying approach was developed \(^3\)
- Probabilistic approach for state estimation

\[
\max_{\{x_k\}} \ p(\{x_k\} \| \{y_k\})
\]

\[^2\text{Odelson et al. (2006); Rajamani and Rawlings (2009); Rajamani et al. (2007)} \]
\[^3\text{Rajamani et al. (2007)} \]
Autocovariance Least-Squares (ALS) Technique

- ALS estimates covariances of system disturbances \((Q, R)\) from data \(^2\)
- For nonlinear systems, linear time-varying approach was developed \(^3\)
- Probabilistic approach for state estimation

\[
\max_{\{x_k\}} p(\{x_k\} | \{y_k\})
\]

- Equivalent least-squares problem when noises are Gaussian

\[
\min_{x_0, x_1 \ldots x_k} \| x_0 - \bar{x}_0 \|_{P^{-1}} + \sum_{i=0}^{k-1} w_i Q^{-1} w_i^T + v_i R^{-1} v_i^T
\]

s.t.: nonlinear model and constraints

- in which \((Q, R)\) are
  - covariances of nonlinear model
  - inverse of the weights of least-squares problem

\(^2\)Odelson et al. (2006); Rajamani and Rawlings (2009); Rajamani et al. (2007)
\(^3\)Rajamani et al. (2007)
Covariance Estimation: Disturbance Modeling

- For covariance estimation, a general linear time-varying process model is proposed to
  - cope with potential plant-model mismatches
  - achieve offset-free performance for outputs

Francis and Wonham (1975, 1976); Pannocchia and Rawlings (2003); Rajamani, Rawlings, and Qin (2009)
Covariance Estimation: Disturbance Modeling

- For covariance estimation, a general linear time-varying process model is proposed to
  - cope with potential plant-model mismatches
  - achieve offset-free performance for outputs
- We augment the state vector with an integrated white noise component $d^4$
- For a general disturbance model

$$\frac{d}{dt} \begin{bmatrix} x \\ d \end{bmatrix} = f(x, u, d)$$

- The DT version of this model is obtained by integrating $f(x, u, d)$

$$x_{k+1} = F(x_k, u_k, d_k) + G(x_k)w_k$$

$$d_{k+1} = d_k + \xi_k$$

$$y_k = h(x_k, d_k) + v_k$$

and we want to estimate the covariances of $w_k$, $\xi_k$ and $v_k$.

$^4$Francis and Wonham (1975, 1976); Pannocchia and Rawlings (2003); Rajamani et al. (2009)
Ethylene Copolymerization Process (Literature)

Variables
- 41 states
- 17 measurements
- 5 inputs
- process and measurement noises

Important variables
- reactor temperature
- reactor pressure
- compositions
- production rate
- polymer properties

First case study
- noise added to temperature of recycle stream

Gani, Mhaskar, and Christofides (2007)
Simulated data are generated using published model \(^5\)

---

\(^5\)McAuley, MacGregor, and Hamielec (1990); Dadebo, Bell, McLellan, and McAuley (1997); Gani et al. (2007)
Simulated data are generated using published model \(^5\)

Assume noise sequences \(v_{\text{sim}} \sim N(0, R_{\text{sim}})\) and \(w_{\text{sim}} \sim N(0, Q_{\text{sim}})\)

with

- \(Q_{\text{sim}} = 2.80 \times 10^{-4}\);
- \(R_{\text{sim}} = 10^{-6} \times \text{diag}(5, 1, 10^4, 20, 300, 10, 200, 0.2, 0.5, 0.3, 0.04, 10^{-5}, 20, 1, 400, 300, 300)\)

\(^5\)McAuley et al. (1990); Dadebo et al. (1997); Gani et al. (2007)
Simulated data are generated using published model ⁵

Assume noise sequences \( v_{\text{sim}} \sim N(0, R_{\text{sim}}) \) and \( w_{\text{sim}} \sim N(0, Q_{\text{sim}}) \) with

- \( Q_{\text{sim}} = 2.80 \times 10^{-4} \);
- \( R_{\text{sim}} = 10^{-6} \times \text{diag}(5, 1, 10^4, 20, 300, 10, 200, 0.2, 0.5, 0.3, 0.04, 10^{-5}, 20, 1, 400, 300, 300) \)

Data sampling time is \( \Delta_k = 60s \)

⁵McAuley et al. (1990); Dadebo et al. (1997); Gani et al. (2007)
Simulated data are generated using published model \(^5\)

Assume noise sequences \(\nu_{\text{sim}} \sim N(0, R_{\text{sim}})\) and \(w_{\text{sim}} \sim N(0, Q_{\text{sim}})\) with

\[ Q_{\text{sim}} = 2.80 \times 10^{-4}; \]
\[ R_{\text{sim}} = 10^{-6} \times \text{diag}(5, 1, 10^4, 20, 300, 10, 200, 0.2, 0.5, 0.3, 0.04, 10^{-5}, 20, 1, 400, 300, 300) \]

Data sampling time is \(\Delta_k = 60s\)

Applying ALS, the following covariances of \(w_k\) and \(\nu_k\) are estimated

\[ Q_{\text{als}} = 2.84 \times 10^{-4}; \]
\[ R_{\text{als}} = 10^{-6} \times \text{diag}(5.03, 1, 9.7 \times 10^3, 20.2, 308, 10.2, 198, 0.19, 0.5, 0.3, 0.04, 9.8 \times 10^{-6}, 27.2, 0.99, 419, 289, 296) \]

\(^5\)McAuley et al. (1990); Dadebo et al. (1997); Gani et al. (2007)
Plot for the EKF using the ALS covariances

Thus, ALS estimates the covariances accurately
 Plot for the EKF using the ALS covariances

Thus, ALS estimates the covariances accurately
However, only the diagonal components of $R$ and 1 process noise were estimated
  ▶ due to an ill-conditioned full ALS problem
Plot for the EKF using the ALS covariances

Thus, ALS estimates the covariances accurately
However, only the diagonal components of $R$ and 1 process noise were estimated
  - due to an ill-conditioned full ALS problem
Poor conditioning can be reduced or eliminated by
  - scaling of process model
  - designing a reduced-order extended Kalman filter
    - to estimate only the strongly observable system states
    - for covariance estimation
Schmidt-Kalman filter (SKF) approach

- originally developed for navigation systems to improve numerical stability of KF \(^6\)
- later used to tackle weakly observable systems \(^7\)
Schmidt-Kalman filter (SKF) approach

- Originally developed for navigation systems to improve numerical stability of KF \(^6\)
- Later used to tackle weakly observable systems \(^7\)
- General idea
  - Remove weakly observable states in the KF gain calculation
  - Perform a well-conditioned calculation as only strongly observable modes are involved

\(^6\) Schmidt (1966); Brown and Hwang (1997)
\(^7\) Farrell and Barth (1998); Farrell (2008)
Weakly Observable Systems: Filtering Approach

- Schmidt-Kalman filter (SKF) approach
  - originally developed for navigation systems to improve numerical stability of KF \(^6\)
  - later used to tackle weakly observable systems \(^7\)
  - general idea
    - remove weakly observable states in the KF gain calculation
    - perform a well-conditioned calculation as only strongly observable modes are involved
  - resulting filter
    - does not estimate the removed state variables
    - still keeps track of the influences these states have on the gain applied to the other states

---

\(^6\) Schmidt (1966); Brown and Hwang (1997)

\(^7\) Farrell and Barth (1998); Farrell (2008)
Schmidt-Kalman Filter Approach

- Have system in its observability canonical form

\[
\begin{bmatrix}
\frac{x_o}{x_{no}} \\
\end{bmatrix}_{k+1} =
\begin{bmatrix}
A_o & 0 \\
A_{21} & A_{no}
\end{bmatrix}
\begin{bmatrix}
\frac{x_o}{x_{no}} \\
\end{bmatrix}_k +
\begin{bmatrix}
G_o \\
G_{no}
\end{bmatrix} w_k
\]

\[
y_k = \begin{bmatrix}
C_o & 0
\end{bmatrix}
\begin{bmatrix}
\frac{x_o}{x_{no}} \\
\end{bmatrix}_k + v_k
\]

- In this structure
  - \(x_o\) are the observable modes
  - \(x_{no}\) are the unobservable/weakly observable modes
- Calculate KF gain for observable system, \(L_o\), using \(A_o\), \(C_o\) and \(G_o\)
Schmidt-Kalman Filter Approach

- Have system in its observability canonical form
\[
\begin{bmatrix}
\frac{x_o}{x_{no}}
\end{bmatrix}_{k+1} = \begin{bmatrix}
A_o & 0 \\
A_{21} & A_{no}
\end{bmatrix}
\begin{bmatrix}
\frac{x_o}{x_{no}}
\end{bmatrix}_k + \begin{bmatrix}
G_o \\
G_{no}
\end{bmatrix} w_k
\]

\[
y_k = \begin{bmatrix}
C_o & 0
\end{bmatrix}
\begin{bmatrix}
\frac{x_o}{x_{no}}
\end{bmatrix}_k + v_k
\]

in which \((A_o, C_o)\) is observable

- In this structure
  - \(x_o\) are the observable modes
  - \(x_{no}\) are the unobservable/weakly observable modes

- Calculate KF gain for observable system, \(L_o\), using \(A_o, C_o\) and \(G_o\)

- Augment \(L_o\) with zeros as \(\bar{L} = \begin{bmatrix} L_o \\ 0 \end{bmatrix}\)

- Physical states
\[
\hat{x}_{k+1|k+1} = F(\hat{x}_{k|k}, u_k, d_k) + L[y_{k+1} - h(\hat{x}_{k+1|k}, d_{k+1})]
\]

in which \(L = T\bar{L}\) and \(T\) is a similarity transformation matrix
Now consider the case in which 3 process noises are added to
- hydrogen concentration
- reactor temperature
- cooler cooling water temperature
Now consider the case in which 3 process noises are added to
- hydrogen concentration
- reactor temperature
- cooler cooling water temperature

The following noise covariances are assumed for data generation
- $Q_{\text{sim}} = \text{diag}(5 \times 10^{-11}, 2 \times 10^{-4}, 0.5)$
- $R_{\text{sim}} = \text{diag}(0.02, 5 \times 10^{-4}, 6 \times 10^4, 0.3, 80, 0.07, 40, 1.5 \times 10^{-5}, 2 \times 10^{-4}, 5 \times 10^{-5}, 1 \times 10^{-6}, 1 \times 10^{-13}, 0.2, 6 \times 10^{-4}, 100, 50, 55)$
Now consider the case in which 3 process noises are added to:
- hydrogen concentration
- reactor temperature
- cooler cooling water temperature

The following noise covariances are assumed for data generation:
- \( Q_{\text{sim}} = \text{diag}(5 \times 10^{-11}, 2 \times 10^{-4}, 0.5) \)
- \( R_{\text{sim}} = \text{diag}(0.02, 5 \times 10^{-4}, 6 \times 10^4, 0.3, 80, 0.07, 40, 1.5 \times 10^{-5}, 2 \times 10^{-4}, 5 \times 10^{-5}, 1 \times 10^{-6}, 1 \times 10^{-13}, 0.2, 6 \times 10^{-4}, 100, 50, 55) \)

And the ALS estimated values are:
- \( Q_{\text{als}} = \text{diag}(5 \times 10^{-11}, 1.8 \times 10^{-4}, 0.43) \)
- \( R_{\text{als}} = \text{diag}(0.03, 5.1 \times 10^{-4}, 6 \times 10^4, 0.3, 84, 0.07, 46, 1.9 \times 10^{-5}, 2.4 \times 10^{-4}, 5.9 \times 10^{-5}, 1 \times 10^{-6}, 1.5 \times 10^{-13}, 0.5, 6.1 \times 10^{-4}, 114, 60, 59) \)
Ethylene Copolymerization Process (Literature): Estimation Results & Conclusions

- Plot for EKF with statistics determined by ALS

Thus, after applying model scaling and SKF approach, better conditioned state and covariance estimation problems are obtained. ALS is now able to estimate multiple process noises. Good results were also obtained for previous case with 1 process noise.
Ethylene Copolymerization Process (Literature): Estimation Results & Conclusions

- Plot for EKF with statistics determined by ALS

\[ \text{Ethylene Concentration (mol/L)} \]

\[ \text{Time (hours)} \]

\[ \text{Reactor Temperature (K)} \]

\[ \text{Time (hours)} \]

Thus, after applying model scaling and SKF approach

- better conditioned state and covariance estimation problems are obtained
- ALS is now able to estimate multiple process noises
- good results were also obtained for previous case with 1 process noise
Variables

- 44 states
- 16 measurements
- 13 inputs – 5 manipulated
- process and measurement noises
Variables

- 44 states
- 16 measurements
- 13 inputs – 5 manipulated
- process and measurement noises

Industrial data sets are provided by ExxonMobil Chemical Company
Industrial Ethylene Copolymerization Process (ExxonMobil)

- Variables
  - 44 states
  - 16 measurements
  - 13 inputs – 5 manipulated
  - process and measurement noises

- Industrial data sets are provided by ExxonMobil Chemical Company

- Currently implemented state estimation technology needs improvement
  - model is solved one step ahead with some initial guess and a few empirical calculations are performed
  - disturbance models are limited
Industrial Ethylene Copolymerization Process (ExxonMobil)

- Variables
  - 44 states
  - 16 measurements
  - 13 inputs – 5 manipulated
  - process and measurement noises

- Industrial data sets are provided by ExxonMobil Chemical Company

- Currently implemented state estimation technology needs improvement
  - model is solved one step ahead with some initial guess and a few empirical calculations are performed
  - disturbance models are limited

- Project goals
  - have model predictions match process measurements
  - implement state estimators (EKF/MHE) and estimate their noise covariances
Industrial Ethylene Copolymerization: Results

- Nonlinear process model (Gas-phase PE) successfully converted from Fortran to Octave/Matlab
Nonlinear process model (Gas-phase PE) successfully converted from Fortran to Octave/Matlab

Data analysis
- inputs were filtered
- data outliers were removed
Industrial Ethylene Copolymerization: Results

- Nonlinear process model (Gas-phase PE) successfully converted from Fortran to Octave/Matlab

- Data analysis
  - inputs were filtered
  - data outliers were removed

- Plots for original model predictions
Industrial Ethylene Copolymerization: Results After Disturbance Modeling

- Proposed approach to improve predictions
  - Disturbance modeling
    - output disturbance added to hydrogen concentration
    - input disturbance added to catalyst flow rate
  - SKF estimator designed to remove weakly observable states in KF gain calculation

Plots for model predictions after disturbance modeling and KF implementation
Industrial Ethylene Copolymerization: Results After Disturbance Modeling

- Proposed approach to improve predictions
  - Disturbance modeling
    - output disturbance added to hydrogen concentration
    - input disturbance added to catalyst flow rate
  - SKF estimator designed to remove weakly observable states in KF gain calculation

- Plots for model predictions after disturbance modeling and KF implementation
Future Work

- perform covariance estimation to handle disturbances and parameter fluctuations
- implement MHE to systems described by differential algebraic equations
Future Work

- perform covariance estimation to handle disturbances and parameter fluctuations
- implement MHE to systems described by differential algebraic equations

The main contributions of this work are in

- development of optimization-based moving horizon estimators for weakly observable systems
- nonlinear estimation using physical models
- nonlinear covariance estimation from data
- building low complexity disturbance models for nonlinear systems
Future Work

- perform covariance estimation to handle disturbances and parameter fluctuations
- implement MHE to systems described by differential algebraic equations

The main contributions of this work are in

- development of optimization-based moving horizon estimators for weakly observable systems
- nonlinear estimation using physical models
- nonlinear covariance estimation from data
- building low complexity disturbance models for nonlinear systems

Acknowledgments

- NSF DDDAS grant #CNS–0540147
- PRF grant #43321–AC9
- TWCCC participants


