Journal of Loss Prevention in the Process Industries 22 (2009) 703–709

Contents lists available at ScienceDirect

Journal of Loss Prevention in the Process Industries

journal homepage: www.elsevier.com/locate/jlp

Process monitoring and parameter estimation via unscented Kalman filtering

Cheryl C. Qu, Juergen Hahn*

Artie McFerrin Department of Chemical Engineering, Texas A & M University, College Station, TX 77843-3122, United States

ARTICLE INFO

Article history:
Received 7 May 2008
Received in revised form
18 July 2008
Accepted 29 July 2008

Keywords:
Nonlinear estimation
Unscented Kalman filter

ABSTRACT

Nonlinear estimation techniques play an important role in process monitoring since some states and most of the parameters cannot be directly measured. This paper investigates the use of several estimation algorithms such as linearized Kalman filter (LKF), extended Kalman filter (EKF), unscented Kalman filter (UKF) and moving horizon estimation (MHE) for nonlinear systems with special emphasis on UKF as it is a relatively new technique. Detailed case studies show that UKF has advantages over EKF for highly nonlinear unconstrained estimation problems while MHE performs better for systems with constraints.

1. Introduction

One important aspect of process safety is detection of abnormal operating conditions. A common process monitoring approach is to keep track of important states and parameters of a process and compare them against their upper and lower bounds. However, some of these states and most of the parameters cannot be directly measured and instead have to be inferred from plant data.

Extended Kalman filters have found widespread use for nonlinear state and parameter estimation. Unscented Kalman filters, as recently proposed by Julier and Uhlman (2004), could in theory improve upon EKF for state and parameter estimation since linearization is avoided by an unscented transformation and at least second order accuracy is provided. This last point is achieved by carefully choosing a set of sigma points, which captures the true mean and covariance of a given distribution and then passing the means and covariances of estimated states through a nonlinear transformation. As a result UKF is capable of estimating the posterior mean and covariances accurately to a high order. Despite UKF's potential for good performance for state and parameter estimation, only few applications in chemical engineering have been reported so far (Rawlings & Bakshi, 2006; Romanenko & Castro, 2004; Romanenko, Santos, & Afonso, 2004).

This paper investigates the performance of UKF in several case studies. A detailed comparison is made between several state estimation methodologies with a specific emphasis on UKF as it is a relatively new technique. The question as to what degree the filter design affects the estimation results is addressed. Detailed case studies show that UKF and EKF have a similar performance for mildly nonlinear systems and UKF outperforms EKF for strongly nonlinear systems when measurement noise levels are relatively high. However, both EKF and UKF may have limitations for constrained problems and MHE may prove to be a better suited alternative for these cases.

This paper is organized as follows: In Section 2, a brief review of nonlinear state and parameter estimation is presented along with the most widely-used EKF algorithm and the optimization-based MHE strategy. The UKF algorithm for nonlinear estimation is then presented in Section 3. Section 4 compares the performance of each filter for state and parameter estimation and concluding remarks are given in Section 5.

2. Overview of commonly used techniques

This section provides background information for state and parameter estimation and briefly reviews existing algorithms, i.e., LKF, EKF and MHE.

2.1. State estimation

A class of nonlinear systems of interest in state estimation is given by:

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k &= h(x_k, u_k, v_k) \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is a vector of the state variables, the functions f and h are differentiable functions of the state vector x, $w_k \in \mathbb{R}^n$ is a vector of plant noise, with $E[w_k] = 0$ and $E[w_k w_k^T] = Q_k$; $y_k \in \mathbb{R}^m$ is a vector of the measured variables, the function h is a differentiable function of the state vector x and $v_k \in \mathbb{R}^m$ is a vector of measurement noise, with $E[v_k] = 0$ and $E[v_k v^T{}_k] = R_k$; n is the number of states, m refers to the number of measurement variables. The distributions of w and v are

* Corresponding author. Tel.: +1 979 845 3568; fax: +1 979 845 6446.
E-mail address: hahn@tamu.edu (J. Hahn).

0950-4230/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jlp.2008.07.012

not necessarily Gaussian. The initial value x_0 may be assumed to be a Gaussian random variable with known mean and known $n \times n$ covariance matrix P_0.

The objective is to find an estimate $\hat{x}_k$ of x_k to minimize the weighted mean-squared error $E(x_k - \hat{x}_k)M(x_k - \hat{x}_k)^{\mathrm{T}}$, where M is any symmetric nonnegative definite weighting matrix. If all estimates weigh equally, the objective becomes to minimize the error covariance matrix for an unbiased estimator, given by $P = E(x_k - \hat{x}_k)(x_k - \hat{x}_k)^{\mathrm{T}}$. More specifically, the trace of P is chosen to be minimized resulting in the performance index $J = \frac{1}{2}\mathrm{Tr}[E(x_k - \hat{x}_k)(x_k - \hat{x}_k)^{\mathrm{T}}]$.

2.2. Parameter estimation

Parameter estimation involves a nonlinear mapping of the form:

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}, w_{k-1}, \theta_k) \\ y_k &= h(x_k, u_k, v_k, \theta_k) \end{aligned} \tag{2}$$

where θ_k is a vector parameterizing the nonlinear function f. The description of θ_k corresponds to a stationary process with identity state transition matrix, driven by process noise w_{k-1}.

One technique for estimating parameters is to augment the state vector with the parameters to be estimated: $z_k = [x_k^{\mathrm{T}}\,\theta_k^{\mathrm{T}}]^{\mathrm{T}}$. The estimation of both states and parameters can be done recursively by writing the state-space representation as:

$$z_k = \begin{pmatrix} f(x_{k-1}, u_{k-1}, w_{k-1}, \theta_{k-1}) \\ \theta_{k-1} + w_{k-1} \end{pmatrix} \tag{3}$$

2.3. Linearized Kalman filter

A linearized Kalman filter is the local solution for nonlinear estimation problems based on linearization about a nominal state value. The following equations define the discrete-time form of the LKF:

Prediction equations

$$\begin{aligned} \hat{x}_{k|k-1} &= A\left(\hat{x}_{k-1|k-1} - x_0\right) + x_0 + Bu_{k-1} \\ \hat{y}_k &= C\hat{x}_{k|k-1} + Du_k \end{aligned} \tag{4}$$

Update equations

$$\begin{aligned} P_{k|k-1} &= AP_{k-1|k-1}A^{\mathrm{T}} + GQG^{\mathrm{T}} \\ K_k &= P_{k|k-1}C^{\mathrm{T}}\left(CP_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}}\right)^{-1} \\ P_{k|k} &= (I - K_kC)P_{k|k-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k\left(y_k - \hat{y}_k\right) \end{aligned} \tag{5}$$

where $A \approx \frac{\partial f}{\partial x}|_{x_0}$, $B \approx \frac{\partial f}{\partial u}|_{u_0}$, $C \approx \frac{\partial h}{\partial x}|_{x_0}$, $D \approx \frac{\partial h}{\partial u}|_{u_0}$, $G \approx \frac{\partial f}{\partial w}|_{w_0}$ and $H \approx \frac{\partial h}{\partial v}|_{v_0}$ are the matrices of the linearized system model around the nominal value of the states x_0. The matrices Q and R are the tuning parameters of the Kalman filter. Q is used as a measure of confidence in the process model while R represents a measure of confidence for the sensor readings. If the process noise or uncertainties are relatively large compared to the observation noise, then Q has large values compared to R, and vice versa. The matrix P_0 provides a measure of confidence in the knowledge of the initial states x_0. The notation involving Q, R, and P_0 also applies to other estimation methods such as EKF, UKF or MHE mentioned throughout this work.

2.4. Extended Kalman filter

Linear Kalman filters assume that a process stays close to the nominal operation point. However, the values of the states can be quite different from the nominal values due to input changes. Extended Kalman filters address this problem by linearizing the system model along its trajectory. The equations defining the discrete-time form of the EKF are summarized in the following:

Prediction equations

$$\begin{aligned} \hat{x}_{k|k-1} &= f\left(\hat{x}_{k-1|k-1}, u_{k-1}\right) \\ \hat{y}_k &= h\left(\hat{x}_{k|k-1}, u_k\right) \end{aligned} \tag{6}$$

Update equations

$$\begin{aligned} P_{k|k-1} &= A_{k-1}P_{k-1|k-1}A_{k-1}^{\mathrm{T}} + G_{k-1}QG_{k-1}^{\mathrm{T}} \\ K_k &= P_{k|k-1}C_k^{\mathrm{T}}\left(C_kP_{k|k-1}C_k^{\mathrm{T}} + H_kRH_k^{\mathrm{T}}\right)^{-1} \\ P_{k|k} &= (I - K_kC_k)P_{k|k-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k\left(y_k - \hat{y}_k\right) \end{aligned} \tag{7}$$

where $A_{k-1} \approx \frac{\partial f}{\partial x}|_{\hat{x}^-_{k-1}}$, $C_k \approx \frac{\partial h}{\partial x}|_{\hat{x}_{k|k-1}}$, $G_{k-1} \approx \frac{\partial f}{\partial w}|_{w_{k-1}}$ and $H_k \approx \frac{\partial h}{\partial v}|_{v_k}$ are the matrices of the linearized system model and evaluated at the estimated state values.

2.5. Moving horizon estimation

In contrast to EKF which is intended for unconstrained problems, moving horizon estimation is an optimization-based approach. From a perspective of Bayesian theory, the state estimation problem can be formulated as the solution of the following optimization problem:

$$\min_{x_0, \{w_k\}_{k=0}^{T-1}} \phi_T(x_0, \{w_k\}) = \min_{z, \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + \theta_{T-N}(z). \tag{8}$$

subject to

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k &= h(x_k, u_k) + v_k \\ x_k &\in X, w_k \in W, v_k \in V \end{aligned} \tag{9}$$

where the sets X, W and V can be constrained, $x_k := x(k; z, \{w_j\}_{j=T-N}^{k-1})$ denotes the solution of system (9) at time k when the initial state is z,$\{w_j\}_{j=T-N}^{k-1}$ is the process noise sequence from time $T-N$ to $k-1$ and $v_k : y_k - h(x_k, u_k)$. $\theta_{T-N}(z)$ is referred to as the arrival cost, which summarizes the effect of the data $\{y_k\}_{k=0}^{T-N-1}$ on the state $x_T - N$ and makes it possible to transform the optimization problem into one of lower dimension.

For unconstrained, linear systems, the arrival cost can be expressed explicitly since the MHE optimization simplifies to the Kalman filter and its covariance update formula can be used (Rao & Rawlings, 2002). Subject to the initial condition Π_0 and assuming the matrix Π_{T-N} is invertible, the arrival cost can then be expressed as

$$\theta_{T-N}(z) = \left(z - \hat{x}_{T-N}\right)' \Pi_{T-N}^{-1}\left(z - \hat{x}_{T-N}\right) + \phi^*_{T-N} \tag{10}$$

where $\hat{x}_{T-N}$ denotes the optimal estimate at time $T-N$ given all of the measurements y_k from time 0 to $T-N-1$, ϕ^*_{T-N} represents the optimal cost at time $T-N$ and Π_{T-N} is computed from the Kalman filter covariance update

$$\Pi_T = A\Pi_{T-1}A^{\mathrm{T}} + GQG^{\mathrm{T}} - A\Pi_{T-1}C^{\mathrm{T}}\left(C\Pi_{T-1}C^{\mathrm{T}} + HRH^{\mathrm{T}}\right)^{-1} C\Pi_{T-1}A^{\mathrm{T}} \tag{11}$$

The solution to the problem described by Eqs. (8) and (10) is the unique optimal pair $(z^*, \{\hat{w}_k^*\}_{k=T-N}^{T-1})$ which can be integrated to