Computers and Chemical Engineering 94 (2016) 28–44

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Maximum likelihood estimation of noise covariance matrices for state estimation of autonomous hybrid systems

Vinay A. Bavdekar, Naresh N. Nandola, Sachin C. Patwardhan *

Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India

ARTICLE INFO

Article history:
Received 11 November 2015
Received in revised form 4 April 2016
Accepted 3 July 2016
Available online 20 July 2016

Keywords:
Nonlinear state estimation
Unscented Kalman filter
Ensemble Kalman filter
Autonomous hybrid systems
Probability density function estimation

ABSTRACT

A critical aspect of developing Bayesian state estimators for hybrid systems, that involve a combination of continuous and discrete state variables, is to have a reasonably accurate characterization of the stochastic disturbances affecting their dynamics. Recently, Bavdekar et al. (2011) have proposed a maximum likelihood (ML) based framework for estimation of the noise covariance matrices from operating input–output data when an EKF is used for state estimation. In this work, the ML framework is extended to estimation of the noise covariance matrices associated with autonomous hybrid systems, and, to a wider class of recursive Bayesian filters. Under the assumption that the innovations generated by an estimator form a white noise sequence, the proposed ML framework computes the noise covariance matrices such that they maximize the log-likelihood function of the estimator innovations. The efficacy of the proposed scheme is demonstrated through the simulation and experimental studies on the benchmark three-tank system.

1. Introduction

Dynamic systems that involve continuous state variables, discrete states and/or logical states or parameters, are common in process industries. For example, in a fermentation reaction, the rate of growth of the organism is limited by its physiological capabilities of absorbing and processing the nutrients and oxygen. Alternatively, the organism may have different pathways for processing nutrients, depending on their availability (Chaudhary et al., 2007). The modelling of such phenomena introduces discontinuities in the system models (modelled as saturation constraints or using logic variables). Also a continuous process controlled using a programmable logic controller (PLC), which employs a combination of continuous and logic based control algorithms, is a classic example of a hybrid system. If such a process is modelled together with the PLC, then the resulting model consists of both continuous states as well as discrete states. A special class of hybrid systems, namely *autonomous* hybrid systems is of interest in this work. In an autonomous hybrid system, the discrete states are explicit functions of the continuous states.

The problem of state estimation of autonomous hybrid systems has begun receiving attention in the literature, only in the recent past. The main challenge in state estimation of autonomous hybrid systems is the presence of discontinuities introduced because of the discrete variables. While the extended Kalman filter (EKF) is widely used for state estimation of nonlinear systems, its major drawback is that it requires computation of the Jacobian of the nonlinear state and measurement equations at every sampling instant. Due to the use of discrete or logic variables in modelling such systems, it is not possible to compute the Jacobian of the nonlinear model at such points. Hence, sampling-based derivative-free filters have to be employed for state estimation of autonomous hybrid systems (Vachhani et al., 2006; Prakash et al., 2010a; Straka et al., 2011; Stano et al., 2013). Prakash et al. (2010b) have proposed a modified version of the unscented Kalman filter (UKF) and the ensemble Kalman filter (EnKF), for state estimation of autonomous hybrid systems. These Bayesian filters do not require the Jacobian of the system equations and the statistical properties of the state estimates are computed using sample statistics. The point of discontinuity gets straddled by the samples of the states and disturbances drawn from their respective distributions, thereby allowing the UKF and EnKF to approximate the effect of the discontinuity. Prakash et al. (2010b), however, assume that the characteristics of the stochastic signals influencing the system dynamics are known accurately. In practice, however, the latter assumption can prove to be a bottleneck in the implementation of these filters. Juloski et al. (2003) have proposed a particle filter based approach for state estimation of such systems. In the recent past, Bemporad et al. (1999)

* Corresponding author.
E-mail address: sachinp@iitb.ac.in (S.C. Patwardhan).

http://dx.doi.org/10.1016/j.compchemeng.2016.07.010
0098-1354/

and Ferrari-Trecate et al. (2002) have proposed moving horizon estimation (MHE) formulations for state estimation of autonomous hybrid systems.

Processes have to operate in the presence of various unmeasured disturbances. Accurate characterization of the unmeasured signals is a critical step in the development of Bayesian state estimators (Fitzgerald, 1976). When the state and the measurement noise are assumed to be zero mean Gaussian white noise processes, the problem of estimating the characteristics of these unmeasured signals translates to estimation of the covariance matrices of the Gaussian probability density functions. This choice of distribution can be justified on the basis of the central limit theorem, which states that an ensemble of sufficiently large number of independent random variables, each with a well-defined mean and finite variance, is normally distributed regardless of the underlying distributions (Papoulis and Pillai, 2002). If the unmeasured disturbances acting on a system are viewed as combined effects of multiple independent physical causes, then it is reasonable to model their total effect as Gaussian variables by invoking the central limit theorem (Soderstrom, 2002). The problem of estimating the state and measurement noise covariance matrices, $\mathbf{Q}$ and $\mathbf{R}$ respectively, for Kalman filter (KF) has been well studied in the literature. These approaches can be classified as least-squares approaches (Mehra, 1970; Odelson et al., 2006) and maximum likelihood estimates (MLE) approaches (Zagrobelny and Rawlings, 2015). However, relatively much less work has been reported in the literature on estimating the probability density functions of the noise associated with nonlinear dynamic systems. Valappil and Georgakis (2000) present a systematic method to estimate the process noise covariance matrix, associated with the EKF. Their approach assumes a structured uncertainty in the model, which can be modelled as variations in the parameters. Prior knowledge about the parameter variances is fused with the state estimator, using parameter sensitivity matrices. Lima and Rawlings (2011) have presented a time-varying lagged autocovariance least-squares formulation to estimate the noise covariance matrices from operating data. A least squares formulation is developed using the linearized state space model obtained at every sampling instant, to estimate the values of the state and measurement noise covariance matrices. Recently, Bavdekar et al. (2011) proposed to employ a constrained optimization formulation under the MLE framework for identification of the state and measurement noise covariance matrices associated with the use of the extended Kalman filter (EKF) for state and parameter estimation. This approach uses operating input–output data and yields optimal estimates of the noise covariance matrices, such that they maximize the likelihood function of the innovations sequence generated by the EKF. A numerical solution to the resulting nonlinear optimization problem can be obtained using standard gradient based optimization methods.

In this work, it is proposed to extend the approach developed by Bavdekar et al. (2011) for estimation of the noise covariance matrices associated with a more general class of nonlinear systems, which may include discontinuities and to a wider class of recursive Bayesian state estimators other than the EKF. In order to accommodate the discontinuities, modified versions of the UKF and EnKF (Prakash et al., 2010b) and their respective constrained versions (Vachhani et al., 2006; Prakash et al., 2010a) are used for state estimation. Under the assumption that the random unmeasured process disturbances arise from physical sources and the measurement noise arises from independent channels, the problem of estimating the noise covariance matrices is formulated as a constrained optimization problem in which a suitable objective function of the innovations sequence is minimized. Using this assumption, the covariance matrices can be parametrized as diagonal matrices, which helps in reducing the number of parameters to be identified. The proposed approach is able to deal with measurements that are irregularly sampled at multiple rates. The efficacy of the proposed scheme is demonstrated through simulation studies on the benchmark hybrid three-tank system. The proposed approach is further validated using experimental data generated from the hybrid three tank setup available at Automation Laboratory, Department of Chemical Engineering, I.I.T. Bombay.

The paper is organized as follows. The process modelling for simulation and state estimation and the problem formulation for identifying the noise covariance matrices are described in detail in Section 2. The results obtained from simulations of the benchmark hybrid three-tank setup and studies on the experimental prototype are presented in Section 3.

2. Problem formulation

2.1. Process model

Consider an autonomous hybrid system given by the following mechanistic model

$$\begin{aligned} \frac{d\mathbf{z}}{dt} &= \mathbf{f}(\mathbf{z}, \boldsymbol{\xi}, \mathbf{m}, \mathbf{d}_m, \mathbf{d}_u, \mathbf{p}, t) \\ \boldsymbol{\xi} &= \mathbf{g}(\mathbf{z}) \\ \mathbf{y}_T(t) &= \mathbf{h}(\mathbf{z}) \end{aligned} \tag{1}$$

where $\mathbf{z} \in \mathbb{R}^n$ denotes the continuous states of the process, $\boldsymbol{\xi} \in \mathbb{R}^{n_d}$ denotes the discrete states, $\mathbf{m} \in \mathbb{R}^m$ denotes the manipulated inputs, $\mathbf{p} \in \mathbb{R}^p$ denotes the parameters of the process and $\mathbf{y}_T \in \mathbb{R}^r$ denotes the true signal corresponding to the obtained measurements. The function $\mathbf{g}(\cdot)$ is expressed as a combination of logic variables such as OR, AND, XOR, etc. $\mathbf{d}_m \in \mathbb{R}^{d_m}$ represents the measured disturbances and $\mathbf{d}_u \in \mathbb{R}^{d_u}$ denotes the unmeasured disturbances. The function $\mathbf{h}(\cdot)$ is a map between the states and the measurements. In the present work, the discrete states, $\boldsymbol{\xi}$, are assumed to be solely a function of the continuous states $\mathbf{x}$.

It may be noted that the mechanistic model available involves continuous time signals and derivatives, whereas the methods chosen for nonlinear state estimation are in discrete time. Thus, it is necessary to introduce additional assumptions to use the mechanistic model for state estimation in discrete time. To arrive at such a model, the following simplifying assumptions are made:

Assumption 1. The measurements are available as samples obtained after every few time instants. The smallest sampling time is T and the measurements can now be represented as

$$\mathbf{y}_k = \mathbf{h}(\mathbf{z}_k) + \mathbf{v}_k \tag{2}$$

where $\mathbf{v}_k \in \mathbb{R}^r$ represents the random measurement noise. The measurement noise is assumed to be a zero-mean white noise signal with a Gaussian distribution, i.e. $\mathbf{v}_k \sim \mathcal{N}(\bar{\mathbf{0}}, \mathbf{R})$, where $\mathbf{R}$ is the covariance matrix. In case where some of the measurements are sampled at irregular intervals, the measurement equation is written as

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{z}_k) + \mathbf{v}_k \tag{3}$$

where the dimension of the time-varying function vector $\mathbf{h}_k(\cdot)$ and the measurement noise vector, varies between 0 and r, depending on the number of measurements available at the kth sampling instant. In such a scenario, the measurement noise characteristic is given by $\mathcal{N}(\bar{\mathbf{0}}, \mathbf{R}_k)$, where the dimension of the covariance matrix $\mathbf{R}_k$ changes with time. It is assumed that all measurements are sampled at time intervals that are integral multiples of the smallest sampling time, T.

Assumption 2. The manipulated inputs are piece-wise constants over the smallest sampling interval, T

$$\mathbf{m}(t) = \mathbf{m}_k \quad \text{for} \quad t_k \le t < t_k + T \tag{4}$$