Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-20 (2016) 025–030

Dynamic Modeling of Gross Errors via Probabilistic Slow Feature Analysis Applied to a Mining Slurry Preparation Process ⋆

Chao Shang [*,**] **Biao Huang** [**] **Yaojie Lu** [**] **Fan Yang** [*] **Dexian Huang** [*]

[*] *Department of Automation, Tsinghua University, TNList, Beijing 100084, P.R. China (E-mail: shangc07@mails.tsinghua.edu.cn, {yangfan, huangdx}@tsinghua.edu.cn)*

[**] *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta T6G 2G6, Canada (E-mail: {cshang, biao.huang, yaojie}@ualberta.ca)*

Abstract: Dynamic data reconciliation and gross error detection ask for an accurate physical model, e.g. a state-space model, based on which measurement noise and gross errors can be quantitatively assessed. The model can be established based on either first-principle knowledge or process operation data. This work considers a case with limited first-principle knowledge and imperfect operation data, which is inspired by a real industrial process. We seek to develop a dynamic model using operation data contaminated by not only measurement noise but also gross errors, which conforms to known static constraints such as mass balance. Probabilistic slow feature analysis (PSFA) is adopted to describe dynamics of both nominal variations and gross errors, and model parameters are estimated by means of the expectation maximization (EM) algorithm. Data from an industrial slurry preparation process are used to demonstrate the usefulness of the proposed method.

© 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Dynamic data reconciliation, gross error detection, probabilistic slow feature analysis, Kalman filter, statistical analysis.

1. INTRODUCTION

In industrial practice, process instrumentations are prone to measurement imprecisions, which may add significant difficulties to real-time control, monitoring and optimizations. Such imprecisions are manifested in two principal forms: random measurement noises and systematic gross errors. To obtain clean data from noisy measurements, data reconciliation techniques have been well established and implemented in the past 30 years (Mah (1990); Crowe (1996)), which are typically based on the statistical properties of random noises. With a statistical description, gross errors can be effectively detected via various statistical tests and then get compensated (Narasimhan et al. (1999)). Because of their close kinship, simultaneous data reconciliation and gross error detection have further been taken into account (Tjoa et al. (1991); Soderstrom et al. (2001)), to furnish trustworthy information for control and optimization purposes.

In general, data reconciliation and gross error detection approaches can be categorized into two groups, namely the static and dynamic strategies. The principal component analysis (PCA)-based method proposed by Tong et al. (1995) is among the earliest attempts that target on the high correlations between process variables. To deal with non-isotropic measurement noise, Gonzalez et al. (2011) proposed to use factor analysis (FA) model and further estimate process variations, measurement noises and bias terms using the expectation maximization (EM) algorithm. Recently, Yuan et al. (2015) has adopted the hierarchical Bayesian framework for simultaneous gross error detection and data reconciliation in both linear and nonlinear cases.

Different from their static counterparts, dynamic data reconciliation and gross error detection approaches usually resort to a state-space model evolving over time (Albuquerque et al. (1996); Bai et al. (2006); Gonzalez et al. (2012)). It entails a sufficiently accurate physical model, which shall come from either first-principle knowledge or data-driven modeling. Unfortunately, it is common that only some steady mass balance or energy balance is available, and archived data are inevitably contaminated by gross errors. Hence in both cases, it is a great challenge to establish a physical dynamic model.

Therefore, this article is towards the development of a linear state-space model that covers nominal variations, measurement noises and gross errors using imperfect data, with steady mass balance or energy balance at hand only. It originates from an industrial slurry preparation system within the oil industry. The novelty of the proposed

⋆ This work was supported by National Natural Science Foundation of China (21276137), National Science Engineering Research Council of Canada (NSERC) and AITF. The first author is also supported by China Scholarship Council (CSC). Corresponding should be addressed to B. Huang at biao.huang@ualberta.ca

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.091

method lies in that, it not only describes the dynamics of nominal variations that conform to static constraints, but also models gross errors as time-varying, which differs from traditional dynamic alternatives that take gross errors as constants. The proposed model incorporates both a limited prior knowledge (steady contraints) and information from data that are subject to gross errors, thereby yielding a gray-box mechanism. Probabilistic slow feature analysis (PSFA), an emerging latent variable model proposed by Turner et al. (2007), is utilized to address the evolving dynamics of both process variations and gross errors. Model parameters are then adjusted using the celebrated EM algorithm. In order to enhance the optimization of EM algorithm, a sophisiticated approach is proposed to determine the initial parameters on the basis of deterministic SFA.

The rest of this work is devided into the following sections. Section 2 reviews basics of the PSFA model in brief. Section 3 proposes a state-space model that simultaneously describes the dynamics of nominal variations and gross errors, as well as the measurement noise. Section 4 gives the EM-algorithm for parameter estimation, along with an effective initialization strategy to improve the optimization performance. In Section 5, empirical results on data from an oil sand slurry preparation system are provided. Finally, conclusions are provided.

2. PROBABILISTIC SLOW FEATURE ANALYSIS

In this section, we first review the PSFA model presented in Shang et al. (2015a), a simplified state-space model for unsupervised feature learning from time series data. Assume that observations are denoted as $\mathbf{x}(t) \in \mathbb{R}^m$ and latent variables are denoted as $\mathbf{s}(t) \in \mathbb{R}^q$, and most information of $\mathbf{x}(t)$ can be explained by $\mathbf{s}(t)$. Different from traditional latent variable models such as principal component analysis (PCA), $\mathbf{s}(t)$ is assumed to be temporally correlated in a Markov fashion, which is described as:

$$\begin{aligned}\mathbf{s}(t) &= \mathbf{F}\mathbf{s}(t-1) + \mathbf{e}(t)\\ \mathbf{x}(t) &= \mathbf{H}\mathbf{s}(t) + \boldsymbol{\epsilon}(t)\end{aligned} \tag{1}$$

where $\mathbf{e}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma})$, $\boldsymbol{\epsilon}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$. Matrices $\mathbf{F}$ and $\boldsymbol{\Gamma}$ are governed by a group of transition parameters $\{\lambda_j,\ 1 \le j \le q\}$:

$$\begin{aligned}\mathbf{F} &= \operatorname{diag}\{\lambda_1, \cdots, \lambda_q\},\\ \boldsymbol{\Gamma} &= \operatorname{diag}\left\{1-\lambda_1^2, \cdots, 1-\lambda_q^2\right\}.\end{aligned} \tag{2}$$

Every transition parameter must satisfy $0 \le \lambda_j < 1$. It can be readily verified that each slow feature $s_j(t)$ is an independent first-order auto-regressive (AR(1)) process, of which the stationary mean and variance are specified as, respectively, zero and unity. $\boldsymbol{\Sigma} = \operatorname{diag}\{\sigma_1^2, \cdots, \sigma_m^2\}$ denotes the covariance matrix of measurement noise that is assumed to be diagonal but non-isotropic.

PSFA can be perceived as a special form of linear dynamical system (LDS) in which independence assumptions are made over latent states. Its physical implications are clear that observations $\mathbf{x}(t)$ are intrinsically driven by a series of independent AR(1) processes with different dynamic behaviors. Another attractive feature is that for parameter learning of PSFA using EM algorithm, their initial values can be desirably set by deterministic SFA, as to be clarified in the next section.

3. DYNAMIC MODEL WITH EVOLVING GROSS ERRORS AND KNOWN CONSTRAINTS

In the presence of measurement noise and gross errors, we assume that observations $\mathbf{x}(t)$ can be, in general, partitioned into two parts

$$\mathbf{x}(t) = \tilde{\mathbf{x}}(t) + \mathbf{b}(t), \tag{3}$$

where $\tilde{\mathbf{x}}(t)$ denotes the nominal variations of a process that conform to a known constraint $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{0}$, and $\mathbf{b}(t)$ incorporates both measurement noise and gross errors that are at variance with the constraint. In the preceding section, PSFA is introduced to depict nominal variations of process data without any specific postulations. Next, along the same line with PSFA, we develop dynamic formulations to describe how $\tilde{\mathbf{x}}(t)$ and $\mathbf{b}(t)$ evolve over time.

3.1 Dynamic Model for Mass/Energy Balances

For nominal variations that agree with the constraint, they can be described in a dynamic sense as follows:

$$\begin{aligned}\mathbf{v}(t) &= \mathbf{F}_v\mathbf{v}(t-1) + \mathbf{e}_v(t)\\ \tilde{\mathbf{x}}(t) &= \mathbf{H}_v\mathbf{v}(t) + \mathbf{x}_0\end{aligned} \tag{4}$$

where $\mathbf{H}_v\mathbf{v}(t)$ denotes the zero-mean nominal variations of process driven by p independent slow features $\mathbf{v}(t) \in \mathbb{R}^p$, and $\mathbf{x}_0$ represents the static working point of $\mathbf{x}$. To satisfy the constraint $\mathbf{A}\tilde{\mathbf{x}}(t) = 0$, it is assumed *a priori* that $\mathbf{A}\mathbf{H}_v = \mathbf{0}$ and $\mathbf{A}\mathbf{x}_0 = \mathbf{0}$ such that both dynamic variations and static point meet the constraint. Dynamic parameters are specified as

$$\begin{aligned}\mathbf{F}_v &= \operatorname{diag}\{\lambda_1, \cdots, \lambda_p\},\\ \mathbf{e}_v(t) &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_e),\\ \boldsymbol{\Gamma}_e &= \operatorname{diag}\left\{1-\lambda_1^2, \cdots, 1-\lambda_p^2\right\},\end{aligned} \tag{5}$$

in a similar fashion as PSFA.

3.2 Dynamic Model for Gross Errors and Measurement Noise

Gross errors and measurement noises are absorbed into $\mathbf{b}(t)$, and its dynamics can then be formulated as

$$\begin{aligned}\mathbf{w}(t) &= \mathbf{F}_w\mathbf{w}(t-1) + \mathbf{e}_w(t),\\ \mathbf{b}(t) &= \mathbf{H}_w\mathbf{w}(t) + \mathbf{b}_0 + \boldsymbol{\epsilon}(t),\end{aligned} \tag{6}$$

where $\mathbf{H}_w\mathbf{w}(t)$ stands for the zero-mean nominal variations of gross errors that transgress the constraint, and $\mathbf{w}(t) \in \mathbb{R}^q$ denotes underlying slow features. $\mathbf{b}_0$ is the steady operation point of gross errors, considered as a random variable with Gaussian uncertainties $\mathbf{b}_0 \sim \mathcal{N}(\mathbf{b}, \boldsymbol{\Sigma}_b)$. $\boldsymbol{\epsilon}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ stands for the measurement noise term. Dynamic parameters are specified as $\mathbf{F}_w = \operatorname{diag}\{\lambda_{p+1}, \cdots, \lambda_{p+q}\}$ and $\mathbf{e}_w(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_w)$, $\boldsymbol{\Gamma}_w = \operatorname{diag}\left\{1-\lambda_{p+1}^2, \cdots, 1-\lambda_{p+q}^2\right\}$.

3.3 The Entire Model

If we further combine models of nominal variations and bias together based on the preceding assumptions, observations $\mathbf{x}(t)$ can be finally decomposed as five parts in a dynamic sense:

$$\mathbf{x}(t) = \mathbf{H}_v\mathbf{v}(t) + \mathbf{x}_0 + \mathbf{H}_w\mathbf{w}(t) + \mathbf{b}_0 + \boldsymbol{\epsilon}(t). \tag{7}$$

Download English Version:

https://daneshyari.com/en/article/5002977

Download Persian Version:

https://daneshyari.com/article/5002977

Daneshyari.com