Computers and Chemical Engineering 41 (2012) 77–87

Contents lists available at SciVerse ScienceDirect

Computers and Chemical Engineering

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

Abnormal events management and process safety

Fault detection in dynamic processes using a simplified monitoring-specific CVA state space modelling approach

Shallon Stubbs, Jie Zhang*, Julian Morris

Centre for Process Analytics and Control Technology, School of Chemical Engineering and Advanced Materials, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

ARTICLE INFO

Article history:
Received 23 April 2011
Received in revised form 25 January 2012
Accepted 13 February 2012
Available online 10 March 2012

Keywords:
State space modelling
CVA
Dynamic models
Fault detection
Process monitoring

ABSTRACT

State space models have been successfully used for the modelling, control and monitoring of dynamic processes with several different approaches employed to derive the state variables of the model. Typically, state-space canonical variate analysis (CVA) modelling requires the estimation of five matrices to fully parameterize the model. This paper proposes a simpler CVA state space model defined by three matrices for the specific purpose of process monitoring. A modified definition of the past vector of inputs and output is proposed in order to facilitate efficient estimation of a reduced set of state space matrices. A sequential procedure for accurate selection of the model state vector dimension is also proposed. The proposed method is applied to the benchmark Tennessee Eastman process and the results show that the proposed method gives comparable and in some cases even better performance than the established CVA state space monitoring methods.

1. Introduction

State space models have been reported to be superior to other multivariate statistical methods for the modelling, control and monitoring of dynamic processes. In the area of system identification and predictive modelling, Juricek, Seborg, and Larimore (2005) demonstrated that subspace models based on canonical variate analysis (CVA) and numerical algorithm for subspace identification (N4SID) outperformed regression models based on partial least squares (PLS) and constraint categorical regression (CCR). They also demonstrated that, of the two subspace modelling methods, the CVA model was more accurate than its N4SID counterpart. Other comparative analysis works carried out by Simoglou, Martin, and Morris (1999a) and Negiz and Cinar (1997b) have also provided support for the superior performance of CVA based state space models.

A few variants of the state space model representations have also been explored and presented in the literature. Typically, the form of CVA based state-space representation is one that can be used in applications ranging from process modelling, control and monitoring. Such a model generally requires the estimation of five matrices to fully parameterize the model. In control system applications this representation is necessary as control of the plant is achieved via methods involving the application of calculated input signal(s) based upon the past output measurements. Thus far very little emphasis has been placed on selecting a state-space model based upon its intended application and most if not all recent papers employing state space models for process monitoring applications have resorted to this full model representation (Lee, Choi, & Lee, 2006; Odiowei & Cao, 2010; Yao & Gao, 2008).

This paper proposes an adaptation of the state space model representation and CVA based derivation for the specific purpose of process monitoring. The proposed state space model employs a significantly reduced number of parameters. The reduced dimensionality of the model, in conjunction with a slightly amended method of constructing the past vector, makes the model parameter estimation much simpler and more efficient.

The proposed model is used for process monitoring and applied to the benchmark Tennessee Eastman (TE) process under close-loop control. Process monitoring is carried out using the Hotelling's T^2 statistics and squared prediction error (SPE, also known as Q) statistics of the state and output residuals. The results are compared with the reported fault detection performance from previous publications (Russell, Chiang, & Braatz, 2000), where the same set of 21 faults are used. Russell et al. (2000) evaluated three different fault detection models: the traditional CVA state space modelling technique, standard and dynamic principal component analysis (PCA and DPCA), whereas Detroja, Gudi, and Patwardhan (2007) evaluated the detection performance of the Hotelling's T^2 statistics and Q statistics based upon a statistical method called correspondence analysis (CA). Results from these previous publications show that the traditional CVA state space model gives overall the best performance. The results of this paper demonstrates that the proposed

* Corresponding author. Tel.: +44 191 222 7240; fax: +44 191 222 5292.
E-mail addresses: shallon.stubbs@ncl.ac.uk (S. Stubbs), jie.zhang@ncl.ac.uk (J. Zhang), julian.morris@ncl.ac.uk (J. Morris).

0098-1354/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2012.02.009

CVA state space model can offer at least the same and in some cases better fault detection performance in terms of fault detection delay time compared to the traditional CVA state space model.

The paper is organised as follows: Section 2 presents the modified CVA based state space model and highlights its differences from those pioneered by Akaike (1975) and Larimore (1990). Section 3 delves into the application of several model selection criterions and how they were employed for the selection of the appropriate state vector dimension used to construct the state space model. Section 4 introduces the fault monitoring statistics employed and Section 5 provides a comparative analysis and summary of the results obtained alongside that of previous publications. Some conclusions are drawn in Section 6.

2. State space modelling and canonical variate analysis

2.1. Conventional and the proposed CVA based state space models

The well known state space model representation is given in Eq. (1). It is premised on the stochastic process exhibiting Markov properties (Akaike, 1975). In the strict sense definition of a Markov process, the future state of the process, that is, the conditional probability of future transitions should only be dependent upon the current state of the process. Hence the proposed representation given by Eq. (2) is not in contradiction to a Markovian representation and quite accurately aligns with the definition:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{e}_x; \quad \mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{u}_t + \mathbf{G}\mathbf{e}_x + \mathbf{e}_y \tag{1}$$

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{G}\mathbf{e}_y + \mathbf{e}_x; \quad \mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{e}_y \tag{2}$$

For both state space representation $\mathbf{e}_x$ is the state residuals and $\mathbf{e}_y$ is the uncorrelated output residuals. The respective residuals are of the same vector dimension as the state $\mathbf{X}_t \in \mathbb{R}^k$ and output $\mathbf{y}_t \in \mathbb{R}^{n_y}$ vectors. The proposed state space representation retains the $\mathbf{G}$ matrix but it is now incorporated in the state transition equation as opposed to the output equation. The $\mathbf{G}$ matrix is somewhat similar to the innovation term employed in Kalman filter designs (Brown & Hwang, 1992) where state estimation is iteratively improved by using the innovations or residuals of the output equation. The proposed state space representation, therefore, more closely aligns its representation with that of the Kalman filter design but makes the assumption that the covariance of the measurement data is constant.

According to Larimore (1990), accounting for the correlation between the state and output residual ensures a minimum order hidden Markov state space representation. The proposed state space representation similarly guarantees a minimum order hidden Markov model. However, the size of the state vector is determined via a cross-validation procedure using the state transition equation as opposed to the output equation as is the case for Larimore's model given in Eq. (1).

From a control system point of view the essential difference between the two representations is that the five matrix representation, Eq. (1), explicitly accounts for the input vector $\mathbf{u}_t$ and therefore finds its use in control systems applications. For the purpose of fault and disturbance detection, the proposed model, Eq. (2), would then suffice adequately and even be more desirable, given its advantages in terms of simplification of representation and stochastic estimation equations.

The state space representation Eq. (2) is more concise than Eq. (1) with the removal of the current input vector $\mathbf{u}_t$. In order to retain the information component provided by the input vector $\mathbf{u}_t$, it is proposed here to redefine Larimore's past vector representation and this will be elaborated on in the next subsection.

2.2. Canonical variate analysis and state variable extraction

The main idea behind canonical correlation analysis is to extract the relationship between two sets of variables $\mathbf{X}$ and $\mathbf{Y}$ by finding corresponding sets of linear combinations of the original variables (the canonical variates $\mathbf{U}$ and $\mathbf{V}$):

$$\mathbf{U} = \mathbf{XJ} \tag{3}$$

$$\mathbf{V} = \mathbf{YL} \tag{4}$$

The choice of transformation matrices $\mathbf{J}$ and $\mathbf{L}$ is towards maximising the correlation between the canonical variates:

$$\max_{\mathrm{J,L}} = \frac{\mathbf{J}^{\mathrm{T}}\mathbf{R}_{xy}\mathbf{L}}{\sqrt{\mathbf{J}^{\mathrm{T}}\mathbf{R}_{xx}\mathbf{J}}\sqrt{\mathbf{L}^{\mathrm{T}}\mathbf{R}_{yy}\mathbf{L}}} \tag{5}$$

where $\mathbf{R}_{xx} = E(\mathbf{X}^{\mathrm{T}}\mathbf{X})$, $\mathbf{R}_{yy} = E(\mathbf{Y}^{\mathrm{T}}\mathbf{Y})$, and $\mathbf{R}_{xy} = E(\mathbf{X}^{\mathrm{T}}\mathbf{Y})$.

This is equivalent to solving the following optimization problem:

$$\max_{\mathrm{J,L}}\varnothing = \mathbf{J}^{\mathrm{T}}\mathbf{R}_{xy}\mathbf{L} + \lambda_x(\mathbf{I}_x - \mathbf{J}^{\mathrm{T}}\mathbf{R}_{xx}\mathbf{J}) + \lambda_y(\mathbf{I}_y - \mathbf{L}^{\mathrm{T}}\mathbf{R}_{yy}\mathbf{L}) \tag{6}$$

where $\mathbf{I}_x$ and $\mathbf{I}_y$ are identity matrices of appropriate dimensions.

The solution is given by:

$$SVD(\mathbf{R}_{xx}^{-1/2}\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1/2}) = \hat{\mathbf{J}}\mathbf{S}\hat{\mathbf{L}}^{\mathrm{T}} \tag{7}$$

$$\mathbf{J} = \mathbf{R}_{xx}^{-1/2}\hat{\mathbf{J}}; \quad \mathbf{L} = \mathbf{R}_{yy}^{-1/2}\hat{\mathbf{L}} \tag{8}$$

The main diagonal of the $\mathbf{S}$ matrix contains the correlation coefficients. The combined operation of Eqs. (7) and (8) is referred to as the generalized singular value decomposition (GSVD) of $\mathbf{R}_{xy}$.

For our application, the states are derived as the canonical variates between two sets of variables, one set being the past vector $\mathbf{P}$ and the other being the future vector $\mathbf{F}$, which are traditionally defined as follows:

$$\mathbf{P}_t^{\mathrm{T}} = [\mathbf{y}_{t-1}^{\mathrm{T}}; \mathbf{y}_{t-2}^{\mathrm{T}}, \ldots, \mathbf{y}_{t-l_y}^{\mathrm{T}}, \mathbf{u}_{t-1}^{\mathrm{T}}, \mathbf{u}_{t-2}^{\mathrm{T}}, \ldots, \mathbf{u}_{t-l_u}^{\mathrm{T}}]^{\mathrm{T}} \tag{9}$$

$$\mathbf{F}_t^{\mathrm{T}} = \left[\mathbf{y}_t^{\mathrm{T}}; \mathbf{y}_{t+1}^{\mathrm{T}}, \ldots, \mathbf{y}_{t+f}^{\mathrm{T}}\right]^{\mathrm{T}} \tag{10}$$

where l_y, l_u, f are, respectively, the numbers of lags in the output, input, and the number of lead elements of the output samples in the future vector.

The state vector $\mathbf{x}_t$ is computed from the canonical variate transform $\mathbf{J}$ of the past vector:

$$\mathbf{x}_t = \mathrm{J}\mathbf{P}_t^{\mathrm{T}}; \quad \mathrm{GSVD}(\mathbf{R}_{pf}) = \mathbf{JSL}^{\mathrm{T}} \tag{11}$$

subject to $\mathbf{J}^{\mathrm{T}}\mathbf{R}_{pp}\mathbf{J} = \mathbf{I}_m$ and $\mathbf{L}^{\mathrm{T}}\mathbf{R}_{ff}\mathbf{L} = \mathbf{I}_q$, where $\mathbf{R}_{pp} = \mathbf{P}^{\mathrm{T}}\mathbf{P}$, $\mathbf{R}_{pf} = \mathbf{P}^{\mathrm{T}}\mathbf{F}$, and $\mathbf{R}_{ff} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$.

To account for the removal of the $\mathbf{u}_t$ input in the proposed state space representation, the following definition of the past vector $\mathbf{P}$ is proposed in this paper:

$$\mathbf{P}_t^{\mathrm{T}} = [\mathbf{y}_{t-1}^{\mathrm{T}}; \mathbf{y}_{t-2}^{\mathrm{T}}, \ldots, \mathbf{y}_{t-l_y}^{\mathrm{T}}, \mathbf{u}_t^{\mathrm{T}}, \mathbf{u}_{t-1}^{\mathrm{T}}, \ldots, \mathbf{u}_{t-l_u}^{\mathrm{T}}]^{\mathrm{T}} \tag{12}$$

The subtle amendment is the inclusion of the $\mathbf{u}_t$ vector in the past matrix definition such that the process of deriving the states would retain what information that is contained by the input vector at the current time $\mathbf{u}_t$.

2.3. Estimating parameters in the state space model

Larimore's stochastic estimation procedure is summarised by Eqs. (13)–(16). The stochastic algorithms first derives estimates for the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ and then proceeds to simultaneously