Journal of Process Control 44 (2016) 120–133

Contents lists available at ScienceDirect

Journal of Process Control

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

System identification in the presence of trends and outliers using sparse optimization

Amir H. Shirdel[1], Jari M. Böling, Hannu T. Toivonen*

Process Control Laboratory, Faculty of Science and Engineering, Åbo Akademi University, FI-20500 Turku (Åbo), Finland

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form 18 January 2016
Accepted 17 May 2016
Available online 10 June 2016

Keywords:
System identification
Trend detection
Sparse optimization

ABSTRACT

In empirical system identification, it is important to take into account the effect of structural disturbances, such as outliers and trends in the data, which might otherwise deteriorate the identification accuracy. A commonly used approach is to preprocess the data to remove outliers and trends, followed by system identification using the processed data. This approach is not optimal because before a system model is available it may not be possible to separate outliers and trends in the data from excitation by the system inputs. In this study a procedure is presented for simultaneous identification of ARX and ARMAX system models and unknown structural disturbances, consisting of outliers and piece-wise linear offsets or trends. This is achieved by introducing sparse representations of the disturbances, having only a few non-zero values. The system identification problem is formulated as a least-squares problem with a sparsity constraint. The sparse optimization problem is solved using ℓ_1-regularization with iterative reweighting, which can be solved efficiently as a sequence of convex optimization problems. Simulated examples and experimental data from a pilot-plant distillation column are used to demonstrate that using the proposed method accurate system models can be identified from experimental data containing unknown trends and outliers.

1. Introduction

In empirical system identification, a system model is determined using input–output data collected from the process. Apart from stationary stochastic noise and the known system inputs, the process or the measurements are often subject to structural disturbances, such as outliers, level shifts and piecewise linear trends. It is important to detect and account for these disturbances properly, as they may otherwise deteriorate the identification accuracy. A standard approach is to apply data preprocessing to remove trends in the input and output signals system identification, while outliers can often be detected and removed by examining the residuals obtained from a preliminary identification round [1]. This method is, however, suboptimal, as it is not possible to distinguish between the effects of the system dynamics and system inputs from those of trends and outliers unless a system model is already available. Therefore, system identification and detection of outliers and trends in the data are inherently coupled, and should be performed simultaneously.

The detection of non-stationary components, such as outliers, level shifts and trends has been studied extensively in time series and signal analysis [2,3]. Applications where it is important to find an underlying trend in a time series include, among others, financial time series analysis, climate time series, astronomy, social sciences, biological and medical sciences, see for example [2,4,3] and the references therein. In industrial data analysis and process control, trend detection and characterization is an important task in process monitoring and fault detection [5,6].

Several trend filtering methods have been studied, including smoothing splines [7], exponential smoothing [8], smoothing by minimizing the sum of squares of second differences (Hodrick–Prescott filtering) [9,10], moving average filtering [11], band-pass filtering [12,13], median filtering [14], empirical mode decompositions [15], de-trending via rational square-wave filters [16], a jump process approach [17], linear programming (LP) approach with fixed kink points [18], and wavelet transform analysis [19].

One approach to estimate structural signals, such as outliers, level shifts and trends, is to exploit sparse representations of the signals. This allows the use of sparse modeling techniques [20] for signal identification. Kim et al. [4] presented an ℓ_1 trend filtering

* Corresponding author.
E-mail addresses: ashirdel@abo.fi (A.H. Shirdel), jboling@abo.fi (J.M. Böling), htoivone@abo.fi (H.T. Toivonen).
[1] Present address: Neste Jacobs, FI-20100 Turku, Finland.

http://dx.doi.org/10.1016/j.jprocont.2016.05.008
0959-1524/

method based on sparse optimization and ℓ_1 regularization. In analogy with the Hodrick–Prescott filter [9], the ℓ_1 trend filtering method determines the trend signal by minimizing the magnitude of its second differences. The Hodrick–Prescott filter generates a smoothed trend signal by constraining the sum of the squares of the second differences, while the ℓ_1 trend filtering method minimizes the sum of their absolute values, producing a piecewise linear trend estimate. ℓ_1 trend filtering has also been applied in [21] to basis function expansion in nonlinear regression when the underlying regression function has inhomogeneous smoothness.

System identification in the presence of structural disturbances has been studied in Xu et al. [22], who consider identification of a FIR model in the presence of outliers. They show that using a compressed sensing approach, the estimation error goes asymptotically to zero under certain conditions.

In this paper the ℓ_1 trend filter presented in [4] is applied to identification of ARX and ARMAX models in the presence of structural disturbances. In the proposed method, sparse optimization is applied to prediction error identification of ARX and ARMAX models and simultaneous estimation of the structural disturbances. The disturbances can be either a process disturbances or a measurement disturbances. It is assumed that the structural disturbances consist of outliers, level shifts and trends, but are otherwise unknown. The fact the structural disturbances and their first and second differences are sparse (having relatively few non-zero components) is exploited to formulate the system identification problem as a sparse optimization problem, which is solved by ℓ_1 relaxation with iterative reweighting [23,20].

The paper is organized as follows. In Section 2, the identification problem is defined. A solution based on sparse optimization is presented in Sections 3 and 4, and in Section 5 the proposed method is demonstrated using both simulated examples and experimental data from a distillation column.

2. Identification in the presence of structural disturbances

We consider a linear discrete-time system described by

$$A(q^{-1})y(k) = B(q^{-1})u(k-l) + C(q^{-1})e(k) + F(q^{-1})d(k) \tag{1}$$

where is $y(k)$ is the output variable, $u(k)$ is the input, $e(k)$ is random zero-mean white noise with variance σ^2, and $d(k)$ is a structural disturbance. The operators $A(\cdot), B(\cdot), C(\cdot), F(\cdot)$ are polynomials in the backward shift operator q^{-1} ($q^{-1}y(k) = y(k-1)$),

$$\begin{aligned} A(q^{-1}) &= 1 - a_1 q^{-1} - \cdots - a_{n_A} q^{-n_A} \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \cdots + b_{n_B} q^{-n_B} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_{n_C} q^{-n_C} \\ F(q^{-1}) &= 1 + f_1 q^{-1} + \cdots + f_{n_F} q^{-n_F} \end{aligned}$$

and l is a non-negative time delay. It is assumed that the disturbance $d(k)$ is composed of spikes (outliers) $d_0(k)$, piecewise constant offsets $d_1(k)$, and piecewise linear trends $d_2(k)$,

$$d(k) = d_0(k) + d_1(k) + d_2(k). \tag{2}$$

In the time-series literature, the term structural time series has been used to refer to time series which can be decomposed as in (2) [24,25]. Here, we will call the disturbance $d(k)$ a structural disturbance.

The outliers occur at discrete time instants $k_{0,i}$, so that

$$d_0(k) = \begin{cases} d_0(k_{0,i}), & k = k_{0,i}, \quad i = 1, \ldots, M_0 \\ 0, & otherwise. \end{cases} \tag{3}$$

The piecewise constant offsets are described by

$$d_1(k) = d_1(k_{1,i}), \quad k_{1,i} \le k < k_{1,i+1}, \quad i = 1, \ldots, M_1 \tag{4}$$

with discontinues at time instants $k_{1,i}$. The piecewise linear trends are modeled by

$$d_2(k) = d_2(k-1) + \beta_i, \quad k_{2,i} \le k < k_{2,i+1}, \quad i = 1, \ldots, M_2 \tag{5}$$

where β_i is the slope of the trend in the indicated time interval, and $k_{2,i}$ are the kink points, i.e., the time instants at which the slope changes [4].

The problem studied in this paper is to identify the system model (1) in the presence of structural disturbance $d(k)$ from a sequence $\{y(k), u(k), k = 1, \ldots, N\}$ of system inputs $u(k)$ and measured outputs $y(k)$. If the non-stationary structural disturbances are large, they will deteriorate the identification accuracy, and need to be taken into account.

A commonly applied way to eliminate the effect of unknown trends is to preprocess the input and output sequences by fitting linear trends by least squares prior to system identification [1]. In this approach it is, however, hard to distinguish between effects due to the known inputs and the unknown trends, and it is difficult to generalize it to cases where the trends may change slopes during the experiment. It is therefore well motivated to perform estimation of the system model and the structured disturbances simultaneously, and to estimate the disturbances even when the main objective may be to estimate the model parameters a_i, b_i, c_i. Observe that neither the time instants $k_{m,i}$ of the disturbance discontinuities nor their values are known. However, it is assumed that the number of discontinuities M_0, M_1 and M_2 is small compared to the number N of data points. The identification problem therefore consists of finding the statistically significant disturbances (3)–(5), which are not described by the stochastic noise.

In many cases a structural disturbance, such as an outlier or trend, can be traced to the measurement device. In this case the disturbance enters at the measured output, and the system is described by

$$\begin{aligned} A(q^{-1})y_0(k) &= B(q^{-1})u(k-l) + C(q^{-1})e(k) \\ y(k) &= y_0(k) + d(k) \end{aligned} \tag{6}$$

where $y_0(k)$ is the disturbance-free output. The system description (6) is a special case of (1) with $F(q^{-1}) = A(q^{-1})$. As the focus in this paper is on identifying the system parameters in the presence of structural disturbances rather than the disturbance dynamics, we will focus on the cases where the disturbance dynamics are either ignored, i.e., $F(q^{-1}) = 1$, or it enters at the output as in (6), i.e., $F(q^{-1}) = A(q^{-1})$. Introducing the parameter vectors

$$\begin{aligned} \theta_a &= [a_1 \ \cdots \ a_{n_A}]^T \\ \theta_b &= [b_1 \ \cdots \ b_{n_B}]^T \\ \theta_c &= [c_1 \ \cdots \ c_{n_C}]^T \\ \theta_f &= [f_1 \ \cdots \ f_{n_F}]^T \end{aligned} \tag{7}$$

and the variable vectors

$$\begin{aligned} \varphi_y(k) &= [y(k-1) \ \cdots \ y(k-n_A)]^T \\ \varphi_u(k) &= [u(k-1) \ \cdots \ y(u-n_B)]^T \\ \varphi_e(k) &= [e(k-1) \ \cdots \ e(u-n_C)]^T \\ \varphi_d(k) &= [d(k-1) \ \cdots \ d(k-n_F)]^T \end{aligned} \tag{8}$$

system (1) can be written as

$$y(k) = \theta_a \varphi_y(k) + \theta_b \varphi_u(k) + \theta_c \varphi_e(k) + \theta_f \varphi_d(k) + d(k) + e(k). \tag{9}$$