Automatica 71 (2016) 308–313

Contents lists available at ScienceDirect

Automatica

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

Technical communique

Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model ☆

Yanjiao Wang, Feng Ding

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

ARTICLE INFO

Article history:
Received 16 April 2015
Received in revised form
22 March 2016
Accepted 6 May 2016
Available online 10 June 2016

Keywords:
Recursive identification
Least squares
Auxiliary model
Hierarchical identification
MIMO system

ABSTRACT

This communique uses the auxiliary model method to study the identification problem of a multiple-input multiple-output (MIMO) system. For such a MIMO system whose outputs are contaminated by an ARMA noise process (i.e., correlated noise), an auxiliary model based recursive least squares parameter estimation algorithm is presented through filtering input–output data. The proposed algorithm has higher estimation accuracy than the existing multivariable identification algorithm. The simulation example is given.

1. Introduction

Multiple-input multiple-output (MIMO) systems widely exist in industrial processes (Zhang, Shi, & Saadat, 2011). For example, Wang, Ding, and Zhu (2013) designed a multivariable controller for linear time-invariant MIMO systems through optimizing controller parameters. Mercèrea and Bako (2011) studied the parameterization and identification method for a MIMO canonical state-space model from data directly. Also, a hierarchical gradient-based identification algorithm was proposed for multivariable discrete-time systems (Ding & Chen, 2005).

The filtering technique has attracted much attention due to its potential for solving many problems in signal processing and analysis (Qaisar, Fesquet, & Renaudin, 2014), communication and system identification. Recently, a filtering based least squares algorithm has been developed for an equation-error MIMO system whose disturbance is an autoregressive (AR) process (Wang et al., 2013); a filtering based hierarchical stochastic gradient algorithm and two filtering based hierarchical iterative algorithms have been presented for multivariable systems (Wang & Ding, 2016a,b). On the basis of these work, this communique investigates novel parameter estimation methods for an output-error MIMO systems whose disturbance is an ARMA noise using the auxiliary model and data filtering methods. The main contributions of this work are as follows.

- This communique derives a filtering based auxiliary model recursive least squares (AM-RLS) algorithm for output-error MIMO systems with ARMA noise through filtering input–output data.
- Compared with the AM-RLS algorithm, the proposed filtering based AM-RLS algorithm has higher estimation accuracy because it uses the filtered input–output data and uses the outputs of the auxiliary model to replace the unknown variables.
- This work is based on the output-error MIMO model with ARMA noise and differs from the previous work with AR noise in Wang et al. (2013), because estimating the parameters of an AR process is a linear problem, while estimating the parameters of an ARMA process is a nonlinear one Li, Zhu, and Dickinson (1989).
- The convergence of the proposed algorithm is analyzed using the stochastic process theory.

The communique is organized as follows. Section 2 derives a filtering based auxiliary model recursive least squares (F-AM-RLS) identification algorithm for MIMO systems and analyzes its convergence. Section 3 compares the proposed algorithm with the existing algorithm. Section 4 gives a simulation example. Finally, the concluding remarks are given in Section 5.

☆ This work was supported in part by the National Natural Science Foundation of China (No. 61273194). The material in this paper was partially presented at the 9th International Symposium on Advanced Control of Chemical Processes, June 7–10, 2015, Whistler, British Columbia, Canada. This paper was recommended for publication in revised form by Associate Editor A. Pedro Aguiar under the direction of Editor André L. Tits.

E-mail addresses: yjwang12@126.com (Y. Wang), fding@jiangnan.edu.cn (F. Ding).

http://dx.doi.org/10.1016/j.automatica.2016.05.024
0005-1098/

2. The F-AM-RLS algorithm

Consider the following MIMO system with ARMA noise,

$$\boldsymbol{y}(t)=\boldsymbol{x}(t)+\boldsymbol{w}(t), \tag{1}$$

$$\boldsymbol{x}(t)=\boldsymbol{G}(z)\boldsymbol{u}(t), \tag{2}$$

where $\boldsymbol{y}(t)\in\mathbb{R}^m$ is the output vector, $\boldsymbol{u}(t)\in\mathbb{R}^r$ is the input vector, $\boldsymbol{w}(t):=H(z)\boldsymbol{v}(t)\in\mathbb{R}^m$ is an ARMA noise process, $\boldsymbol{v}(t)\in\mathbb{R}^m$ is a white noise vector with zero mean, the rational fractions $\boldsymbol{G}(z):=\boldsymbol{\beta}(z)/\alpha(z)$ and $H(z):=d(z)/c(z)$ are the transfer matrix/function, z^{-1} is a unit backward shift operator, and

$$\alpha(z):=1+\alpha_1z^{-1}+\alpha_1z^{-2}+\cdots+\alpha_nz^{-n},$$

$$\boldsymbol{\beta}(z):=\boldsymbol{\beta}_1z^{-1}+\boldsymbol{\beta}_2z^{-2}+\cdots+\boldsymbol{\beta}_nz^{-n},$$

$$c(z):=1+c_1z^{-1}+c_2z^{-2}+\cdots+c_{n_c}z^{-n_c},$$

$$d(z):=1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d}.$$

Without loss of generality, suppose that the structure parameters n, n_c and n_d are predetermined and $\boldsymbol{y}(t)=\boldsymbol{0}$, $\boldsymbol{u}(t)=\boldsymbol{0}$, $\boldsymbol{w}(t)=\boldsymbol{0}$ and $\boldsymbol{v}(t)=\boldsymbol{0}$ for $t\leqslant 0$. The available input–output data are $\{\boldsymbol{u}(t),\boldsymbol{y}(t)\}$.

Define the parameter vectors/matrices and the information vectors/matrices:

$$\boldsymbol{\theta}^{\mathrm{T}}:=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n]\in\mathbb{R}^{m\times(nr)},$$

$$\boldsymbol{\alpha}:=[\alpha_1,\alpha_2,\ldots,\alpha_n]^{\mathrm{T}}\in\mathbb{R}^n,$$

$$\boldsymbol{\rho}:=[c_1,c_2,\ldots,c_{n_c},d_1,d_2,\ldots,d_{n_d}]^{\mathrm{T}}\in\mathbb{R}^{n_c+n_d},$$

$$\boldsymbol{\varphi}(t):=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\ldots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}\in\mathbb{R}^{nr},$$

$$\boldsymbol{\zeta}(t):=[-\boldsymbol{x}(t-1),-\boldsymbol{x}(t-2),\ldots,-\boldsymbol{x}(t-n)]\in\mathbb{R}^{m\times n},$$

$$\begin{aligned}\boldsymbol{\xi}(t):=[&-\boldsymbol{w}(t-1),-\boldsymbol{w}(t-2),\ldots,-\boldsymbol{w}(t-n_c),\\&\boldsymbol{v}(t-1),\boldsymbol{v}(t-2),\ldots,\boldsymbol{v}(t-n_d)]\in\mathbb{R}^{m\times(n_c+n_d)}.\end{aligned}$$

Then, Eqs. (1)–(2) can be rewritten as

$$\boldsymbol{x}(t)=\boldsymbol{\zeta}(t)\boldsymbol{\alpha}+\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t), \tag{3}$$

$$\boldsymbol{w}(t)=\boldsymbol{\xi}(t)\boldsymbol{\rho}+\boldsymbol{v}(t), \tag{4}$$

$$\boldsymbol{y}(t)=\boldsymbol{\zeta}(t)\boldsymbol{\alpha}+\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)+\boldsymbol{w}(t). \tag{5}$$

Remark 1. For the identification model in (3)–(5), the input–output data $\boldsymbol{u}(t)$ and $\boldsymbol{y}(t)$ are available. That is, only $\boldsymbol{y}(t)$ and $\boldsymbol{\varphi}(t)$ are available but $\boldsymbol{\zeta}(t)$ and $\boldsymbol{\xi}(t)$ are unavailable and unknown.

Remark 2. The difficulty of identification is that the information matrix $\boldsymbol{\zeta}(t)$ contains the unknown inner variable $\boldsymbol{x}(t-i)$. The solution here is to construct an auxiliary model using the measured data $\boldsymbol{u}(t)$ and $\boldsymbol{y}(t)$, and to replace the unknown $\boldsymbol{x}(t-i)$ in the identification algorithm with the output $\boldsymbol{x}_{\mathrm{a}}(t-i)$ of the auxiliary model $\boldsymbol{G}_{\mathrm{a}}(z)$ in Fig. 1.

Remark 3. From (5), we can see that the output $\boldsymbol{y}(t)$ contains the correlated noise $\boldsymbol{w}(t)$, which results in biased estimates. In this work, we use the filtering technique and combine the auxiliary model for investigating a novel identification method. The details are as follows.

We use a linear filter $L(z):=H^{-1}(z)$ to filter the input–output data $\boldsymbol{u}(t)$ and $\boldsymbol{y}(t)$, and $\boldsymbol{\zeta}(t)$ and $\boldsymbol{\varphi}(t)$, leading to the filtered variables $\boldsymbol{y}_{\mathrm{f}}(t):=L(z)\boldsymbol{y}(t)\in\mathbb{R}^m$, $\boldsymbol{u}_{\mathrm{f}}(t):=L(z)\boldsymbol{u}(t)\in\mathbb{R}^r$, $\boldsymbol{\zeta}_{\mathrm{f}}(t):=L(z)\boldsymbol{\zeta}(t)\in\mathbb{R}^{m\times n}$ and $\boldsymbol{\varphi}_{\mathrm{f}}(t):=L(z)\boldsymbol{\varphi}(t)\in\mathbb{R}^{nr}$. Because $H(z)$ is to be identified and unknown, the proposed algorithm has to be implemented through the recursive/iterative scheme using the estimate of $L(z)$. Multiplying both sides of (5) by $L(z)$, we obtain a filtered identification model,

$$\boldsymbol{y}_{\mathrm{f}}(t)=\boldsymbol{\zeta}_{\mathrm{f}}(t)\boldsymbol{\alpha}+\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{f}}(t)+\boldsymbol{v}(t). \tag{6}$$

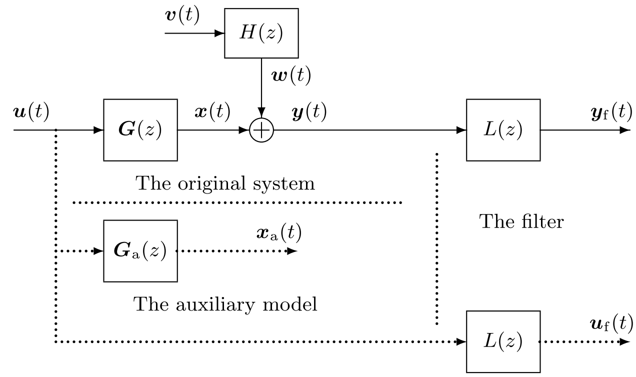

Fig. 1. The MIMO system with an auxiliary model.

Remark 4. Because the filtered output $\boldsymbol{y}_{\mathrm{f}}(t)$ in (6) involves only the white noise $\boldsymbol{v}(t)$, we will derive a new algorithm for estimating $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ using the filtered input–output data $\boldsymbol{u}_{\mathrm{f}}(t)$ and $\boldsymbol{y}_{\mathrm{f}}(t)$ (i.e., $\boldsymbol{\zeta}_{\mathrm{f}}(t)$ and $\boldsymbol{\varphi}_{\mathrm{f}}(t)$)—see Fig. 1.

The filtered identification model in (6) contains both a parameter vector $\boldsymbol{\alpha}$ and a parameter matrix $\boldsymbol{\theta}$. One solution is to derive the estimation algorithm of the parameter vector $\boldsymbol{\alpha}$ for fixed $\boldsymbol{\theta}$, and to derive the estimation algorithm of the parameter matrix $\boldsymbol{\theta}$ for fixed $\boldsymbol{\alpha}$ using the hierarchical identification principle in Ding and Chen (2005).

Let $\hat{\boldsymbol{\alpha}}(t)$, $\hat{\boldsymbol{\theta}}(t)$ and $\hat{\boldsymbol{\rho}}(t)$ be the estimates of $\boldsymbol{\alpha}$, $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$ at time t, respectively. Based on the identification model in (6), for fixed $\boldsymbol{\theta}$ (that is $\boldsymbol{\theta}$ is regarded as being known in this step), defining and minimizing a quadratic criterion function, we can obtain the following recursive least squares relation for $\hat{\boldsymbol{\alpha}}(t)$ (Goodwin & Sin, 1984; Ljung, 1999):

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)+\boldsymbol{P}_{\alpha}(t)\boldsymbol{\zeta}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{e}_{\alpha}(t), \tag{7}$$

$$\boldsymbol{e}_{\alpha}(t)=\boldsymbol{y}_{\mathrm{f}}(t)-\boldsymbol{\zeta}_{\mathrm{f}}(t)\hat{\boldsymbol{\alpha}}(t-1)-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{f}}(t), \tag{8}$$

$$\boldsymbol{P}_{\alpha}^{-1}(t)=\boldsymbol{P}_{\alpha}^{-1}(t-1)+\boldsymbol{\zeta}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\zeta}_{\mathrm{f}}(t). \tag{9}$$

Similarly, for fixed $\boldsymbol{\alpha}$, defining and minimizing a quadratic criterion function, we can obtain the following recursive least squares relation for $\hat{\boldsymbol{\theta}}(t)$:

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{P}_{\theta}(t)\boldsymbol{\varphi}_{\mathrm{f}}(t)\boldsymbol{e}_{\theta}^{\mathrm{T}}(t), \tag{10}$$

$$\boldsymbol{e}_{\theta}(t)=\boldsymbol{y}_{\mathrm{f}}(t)-\boldsymbol{\zeta}_{\mathrm{f}}(t)\boldsymbol{\alpha}-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}_{\mathrm{f}}(t), \tag{11}$$

$$\boldsymbol{P}_{\theta}^{-1}(t)=\boldsymbol{P}_{\theta}^{-1}(t-1)+\boldsymbol{\varphi}_{\mathrm{f}}(t)\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t). \tag{12}$$

Based on the identification model in (4), we can obtain the recursive relation for $\hat{\boldsymbol{\rho}}(t)$:

$$\begin{aligned}\hat{\boldsymbol{\rho}}(t)&=\hat{\boldsymbol{\rho}}(t-1)+\boldsymbol{P}_{\rho}(t)\boldsymbol{\xi}^{\mathrm{T}}(t)[\boldsymbol{w}(t)-\boldsymbol{\xi}(t)\hat{\boldsymbol{\rho}}(t-1)]\\&=\hat{\boldsymbol{\rho}}(t-1)+\boldsymbol{P}_{\rho}(t)\boldsymbol{\xi}^{\mathrm{T}}(t)\\&\quad\times[\boldsymbol{y}(t)-\boldsymbol{\zeta}(t)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)-\boldsymbol{\xi}(t)\hat{\boldsymbol{\rho}}(t-1)],\end{aligned} \tag{13}$$

$$\boldsymbol{P}_{\rho}^{-1}(t)=\boldsymbol{P}_{\rho}^{-1}(t-1)+\boldsymbol{\xi}^{\mathrm{T}}(t)\boldsymbol{\xi}(t). \tag{14}$$

$\boldsymbol{P}_{\alpha}(t)$, $\boldsymbol{P}_{\theta}(t)$ and $\boldsymbol{P}_{\rho}(t)$ are the covariance matrices. However, Eqs. (7)–(14) cannot be implemented because their right-hand sides contain the unknown variables $\boldsymbol{y}_{\mathrm{f}}(t)$, $\boldsymbol{\zeta}_{\mathrm{f}}(t)$, $\boldsymbol{\varphi}_{\mathrm{f}}(t)$, $\boldsymbol{w}(t)$, $\boldsymbol{\xi}(t)$, $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$.

Since the filter $L(z)$ and the inner variable $\boldsymbol{x}(t)$ are unknown, so are the filtered variables $\boldsymbol{y}_{\mathrm{f}}(t)$, $\boldsymbol{\zeta}_{\mathrm{f}}(t)$ and $\boldsymbol{\varphi}_{\mathrm{f}}(t)$. To solve this problem, referring to Fig. 1, we use $\boldsymbol{G}_{\mathrm{a}}(z)$ to generate the estimate $\boldsymbol{x}_{\mathrm{a}}(t)$ of $\boldsymbol{x}(t)$ by means of the auxiliary model or reference model method, and use the estimate $\hat{L}(t,z)$ of $L(z)$ at time t to generate the estimates $\hat{\boldsymbol{y}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{\zeta}}_{\mathrm{f}}(t)$ and $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$ of $\boldsymbol{y}_{\mathrm{f}}(t)$, $\boldsymbol{\zeta}_{\mathrm{f}}(t)$ and $\boldsymbol{\varphi}_{\mathrm{f}}(t)$:

$$\hat{\boldsymbol{y}}_{\mathrm{f}}(t):=\hat{L}(t,z)\boldsymbol{y}(t),\qquad \hat{\boldsymbol{u}}_{\mathrm{f}}(t):=\hat{L}(t,z)\boldsymbol{u}(t), \tag{15}$$