Neurocomputing 186 (2016) 251–257

Contents lists available at ScienceDirect

Neurocomputing

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

Recursive minimum component algorithms for parameter estimation of dynamic systems ☆

Huxiong Li [a], Bingxian Mu [b,*], Lei Zuo [c]

[a] School of New Media, Zhejiang University of Media and Communications, No. 998 Xueyuan Street, Xiasha Higher Education Zone, Hangzhou 310018, PR China
[b] Department of Mechanical Engineering, University of Victoria, PO Box 1700, Stn. CSC Victoria, Canada, BC V8W 2Y2
[c] School of Marine Science and Technology, Northwestern Polytechnical University, No. 127, West Youyi Road, Xi'an 710072, PR China

ARTICLE INFO

Article history:
Received 12 September 2015
Received in revised form
11 December 2015
Accepted 25 December 2015
Communicated by Hui Zhang
Available online 6 January 2016

Keywords:
System identification
Minimum component analysis
Recursive algorithm

ABSTRACT

Minimum component analysis is an elegant and promising approach to system identification. This paper focuses on developing related recursive algorithms. Two algorithms are presented. In the first approach, rank-one modification is used to get an accurate and quickly convergent algorithm. However, its drawback is low computational efficiency. Therefore, the second algorithm is proposed to realize faster computation. The simulation results verify the effectiveness of both algorithms.

© 2016 Published by Elsevier B.V.

1. Introduction

Let us first consider a general parameter estimation problem. The concerned system is described in a linear regression form [1]:

$$y(t)=\varphi^{\mathrm{T}}(t)\theta+e(t), \tag{1}$$

where $y(t)$ is the system output at time t, $e(t)$ is the external noise disturbance, and $\varphi(t)$ is a regression vector consisting of lagged outputs and inputs:

$$\varphi(t)=[-y(t-1)\ \cdots\ -y(t-n_a)\ \ u(t-1)\ \cdots\ u(t-n_b)]^{\mathrm{T}},$$

in which, $u(t)$ is the system control input, and θ is the parameter vector

$$\theta=[a_1\ \cdots\ a_{n_a}\ \ b_1\ \cdots\ b_{n_b}]^{\mathrm{T}}.$$

Many well developed system identification techniques, e.g., least squares (LS) and stochastic gradient (SG), can be applied to system (1). While doing identification with these classical methods, we have two heuristic observations:

- **Observation 1.** *Squared vertical distance versus squared orthogonal distance*? Classical methods based on LS and SG can be visualized as fitting a hyperplane to collected data points by minimizing

$$J(\theta)=\sum_{t=1}^{N}\left[y(t)-\varphi^{\mathrm{T}}(t)\theta\right]^2,$$

which is the sum of the squared vertical distance from the points to the hyperplane. A three-dimensional example is shown in Fig. 1(a). However, a question worth asking is: *Does adoption of the vertical distance in $J(\theta)$ leads to the best fitting performance*? Actually, numerous 'distances' can be defined in mathematical senses to measure fitting errors. Yet among them the orthogonal distance (see Fig. 1(b)) makes the most sense to use, since the aim is to find the hyperplane fitting 'closest' to the data points. Then the problem of system identification by minimizing the sum of squared orthogonal distance arises naturally.
- **Observation 2.** *How to handle noises*? Inherent in any practical system, noises disturb system identification. Thus, for classical methods, the noise $e(t)$ is usually assumed to have certain properties in order to guarantee theoretical convergence. For instance, recursive least squares (RLS) based identification have the following assumptions regarding the noise [2,3]: (1) $\{e_t\}$ is a martingale difference sequence adapted to a family of increasing σ-algebras $\{\mathcal{F}_t\}$; (2) $E(e_t|\mathcal{F}_{t-1})=0$, almost surely (a.s.); (3) $E\left(e_t^2|\mathcal{F}_{t-1}\right)<\infty$, a.s. However, noises can be very complex practically and may not satisfy the assumed properties. Under such circumstances, the mentioned classical identification methods

☆This research was supported by Zhejiang Province International Cooperation Foundation, 2012C24019 and Zhejiang Province Natural Science Foundation, LY13F020023.
* Corresponding author.
E-mail address: bingxianmu@gmail.com (B. Mu).

http://dx.doi.org/10.1016/j.neucom.2015.12.083
0925-2312/© 2016 Published by Elsevier B.V.

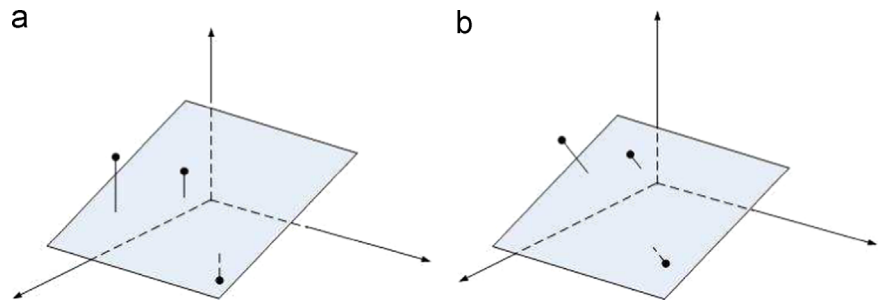

Fig. 1. (a) Fitting by minimizing the sum of the squared vertical distance. (b) Fitting by minimizing the sum of the squared orthogonal distance.

cannot be used directly. Thus it is desirable and necessary to develop noise-robust algorithms.

In fact, the above-mentioned observations are two fundamental problems in system identification. A promising way to solve them is minimum component analysis (MCA), also referred to as last principal component analysis (LPCA). Huang, for the first time, notes the significance of MCA in the field of process identification [4]. As pointed out in [4], MCA based identification has the following merits or features:

- Minimum sum of squared orthogonal distances.
- Maximum likelihood estimation property.
- Robustness against various types of noises.

Briefly, MCA is to extract the minimum component (MC), i.e., the eigenvector associated to the minimum eigenvalue, from the covariance matrix of a data set. MCA in this paper is a narrowed concept of minor component analysis, which is a powerful statistical technique in signal processing and machine learning [5–7]. Recently it is of much interest to perform minor component analysis in an adaptive or recursive manner, and a few methods have been proposed, see [8–11] and the references therein. Opposite to MCA is principal component analysis (PCA). PCA has been used a lot in fault diagnosis. For instance, Li et al. proposed two recursive PCA algorithms to achieve adaptive process monitoring [12], and the recursion is based on rank-one modification and Lanczos tridiagonalization, respectively. A survey of related research in fault diagnosis is given in [13]. A variant of PCA, dynamic PCA, finds application in subspace identification. However, the consistency cannot be ensured, as shown in [14], which thus further proposes an indirect dynamic PCA approach. A similar method widely used in time series analysis is the total least squares (TLS). Some recent works are on TLS based identification, e.g., [15–17], to name a few.

In [4], a batch-wise MCA based identification is elaborated exhaustively. However, for most adaptive schemes in control, filtering and prediction, recursive system modeling is a must [1]. Therefore, this paper will be devoted to constructing identification algorithms from recursive MCA (RMCA). A simple RMCA algorithm can be readily derived from batch-wise MCA by rank-one modification, as will be shown in Section 3, but it involves eigendecomposition that is computationally time-consuming. Thus to facilitate its implementation, it is necessary to find an alternative way to make RMCA fast.

The rest of the paper is organized as follows: Section 2 briefly describes the problem of interest. In Section 3, an accurate RMCA identification algorithm is derived intuitively, followed by the development of a fast RMCA method and its convergence analysis. The proposed algorithms are evaluated and investigated by simulation in Section 4. Finally, Section 5 offers some concluding remarks.

2. Problem formulation

Consider a single-input–single-output (SISO) linear system:

$$A(z)y_o(t) = z^{-d}B(z)u_o(t) + v(t), \tag{2}$$

where $A(z)$ and $B(z)$ are polynomials in z^{-1} (backshift operator):

$$A(z) = a_0 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a},$$

$$B(z) = b_0 + b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}.$$

The orders n_a and n_b and the system delay d are assumed to be known. No assumptions are made for the additive random noise $v(t)$ here. It will be seen later that, for any type of $v(t)$, the convergence can be ensured theoretically. Define the regression vector and parameter vector, respectively, as:

$$\phi(t) = \big[-y_o(t) \;\; -y_o(t-1) \;\; \cdots \;\; -y_o(t-n_a)$$
$$u_o(t-d) u_o(t-d-1) \;\; \cdots \;\; u_o(t-d-n_b)\big]^{\mathrm{T}},$$

$$\theta = [a_0 \;\; a_1 \;\; \cdots \;\; a_{n_a} \;\; b_0 \;\; b_1 \;\; \cdots \;\; b_{n_b}]^{\mathrm{T}}.$$

Note that θ is not unique. Any $\rho\theta$ for some non-zero and finite constant ρ would produce identical input and output measurements. In other words, any identification scheme cannot distinguish between θ and $\rho\theta$. Therefore, without loss of generality, we fix the value of θ by imposing the constraints that $\theta^{\mathrm{T}}\theta = 1$ and the first element of θ must be positive. Clearly,

$$\phi^{\mathrm{T}}(t)\theta + v(t) = 0. \tag{3}$$

To estimate θ, the batch-wise MCA can be applied [4]. First, collect all the regression vectors $\phi(t)$ for $1 \le t \le N$, and arrange them in order to form a matrix:

$$\Phi = [\,\phi(1) \quad \phi(2) \quad \cdots \quad \phi(N)\,].$$

The empirical mean vector of $\phi(t)$'s can be calculated as

$$\overline{\phi} = \frac{1}{N}\sum_{t=1}^{N} \phi(t).$$

Then subtract $\overline{\phi}$ from each column of Φ. Denote the mean-subtracted matrix by $\check{\Phi}$. Then we can get a covariance matrix:

$$R = \frac{1}{N}\check{\Phi}\check{\Phi}^{\mathrm{T}}.$$

Or equivalently, R is given by

$$R = \frac{1}{N}\sum_{t=1}^{N}\left(\phi(t) - \overline{\phi}\right)\left(\phi(t) - \overline{\phi}\right)^{\mathrm{T}}. \tag{4}$$

Eigendecompose R to obtain its eigenvalues and eigenvectors. Denote the resultant MC of R by $\underline{q}$ and the estimate of θ by $\hat{\theta}$ if no confusion can arise. Then $\hat{\theta}$ is taken as $\underline{q}$ as the MC is the estimate of the parameter vector θ.

The above is a compendious summary of the algorithm for system identification by batch-wise MCA. In the following sections, two RMCA based identification methods will be discussed: the first one is accurate, and the second is approximate but fast.

Download English Version:

https://daneshyari.com/en/article/408430

Download Persian Version:

https://daneshyari.com/article/408430

Daneshyari.com