Signal Processing 93 (2013) 1235–1242

Contents lists available at SciVerse ScienceDirect

Signal Processing

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

Decomposition based fast least squares algorithm for output error systems ☆

Feng Ding [a,b,*]

[a] Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, PR China
[b] Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122, PR China

ARTICLE INFO

Article history:
Received 19 June 2012
Received in revised form
22 October 2012
Accepted 18 December 2012
Available online 27 December 2012

Keywords:
Signal processing
Filtering
Parameter estimation
Fast least squares
Iterative algorithm

ABSTRACT

Parameter estimation methods have wide applications in signal processing, communication and system identification. This paper derives an iterative least squares algorithm to estimate the parameters of output error systems and uses the partitioned matrix inversion lemma to implement the proposed algorithm in order to enhance computational efficiencies. The simulation results show that the proposed algorithm works well.

1. Introduction

The iterative algorithms are important for finding the zeros of nonlinear functions and the solutions of linear or nonlinear matrix [1,2], e.g., the Newton iteration methods [3], the optimization and control algorithms [4–7], the Jacobi iteration and the Gauss–Seidel iterations for solving matrix equations $\boldsymbol{Ax}=\boldsymbol{b}$ [8,9], the least squares based iterative methods [10] and the hierarchical gradient based iterative methods [11] for solving coupled Sylvester matrix equations $\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{C}$ and $\boldsymbol{DX}+\boldsymbol{XE}=\boldsymbol{F}$ and general coupled matrix equations. Recently, Li et al. considered the fitting problems of nonlinear functions or nonlinear system modeling and presented a gradient based iterative algorithm and a Newton iterative algorithm to estimate the parameters of a nonlinear function from noisy data according to the negative gradient search and the Newton iteration. Furthermore, two model transformation based iterative algorithms have been developed for improving computational efficiencies [12]; a two-stage least squares based iterative estimation algorithm has been presented for CARARMA system modeling [13].

The recursive algorithms are very related to the iterative algorithms [15–17]. In general, the recursive algorithms can be used for on-line identification and the basic idea is to update the parameters of the systems by using real-time measurement information [14]. Liu et al. discussed the auxiliary model based multi-innovation estimation algorithm for multiple-input single-output systems [18] and studied the convergence properties of stochastic gradient algorithm for multivariable systems [19]; Ding et al. explored time series autoregressive modeling in the presence of missing observations by using the polynomial transformation technique [20]. Xiao et al. presented a residual based interactive least squares algorithm for a controlled autoregressive moving average (C-ARMA) model [21]; Wang et al. proposed the residual

☆ This work was supported by the National Natural Science Foundation of China (No. 61273194), the Natural Science Foundation of Jiangsu Province (China, BK2012549) and the 111 Project (B12018).
* Corresponding author at: Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122, PR China.
E-mail address: fding@jiangnan.edu.cn

0165-1684/$ - see front matter
http://dx.doi.org/10.1016/j.sigpro.2012.12.013

based interactive stochastic gradient algorithm for controlled moving average models [22].

System identification and parameter estimation methods can obtain the parameters of the systems under consideration and are basic for state estimation and filtering [23–26], and adaptive control [27–29]. The iterative algorithms can be used not only for solving matrix equations but also compute the system parameters. In the area of system identification, Ding et al. derived a least squares based and a gradient based iterative estimation methods for output error moving average systems [30,31]; similar iterative methods have been developed for Box–Jenkins systems [32]; Zhang et al. derived a hierarchical gradient based iterative algorithm for multivariable output error moving average systems [33]; Wang studied recursive and iterative algorithms for output error moving average systems [34].

Recently, Hu et al. studied an iterative least squares estimation algorithm for controlled moving average systems based on matrix decomposition [35], and a decomposition based iterative estimation algorithm for autoregressive moving average models [36]. On the basis of the work in [35,36], this paper derives an iterative least squares identification algorithm for output error systems using the information matrix decomposition and the partitioned matrix inversion lemma.

The rest of this paper is organized as follows. Section 2 gives the iterative least squares estimates for output error. Section 3 derives an iterative least squares algorithm using the partitioned matrix inversion lemma. Section 4 provides a simulation example to show the effectiveness of the proposed algorithm. Finally, Section 5 offers some concluding remarks.

2. Basic algorithms

Consider the following output error system [30]:

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

$$x(t)=-\sum_{i=1}^{n_a} a_i x(t-i)+\sum_{i=1}^{n_b} b_i u(t-i), \tag{2}$$

where $\{u(t)\}$ and $\{y(t)\}$ are the input and output sequences of the system, $\{v(t)\}$ is a white noise sequence with zero mean.

Assume that the orders n_a and n_b are known and $n:=n_a+n_b$ and $y(t)=0$, $u(t)=0$ and $v(t)=0$ for $t\leq 0$. The objective is to derive an iterative parameter estimation algorithm to estimate the unknown parameters (a_i,b_i), using on the partitioned matrix inversion lemma, from available input–output measurement data $\{u(t),y(t):t=0,1,2,\ldots,L\}$, where L denotes the data length ($L\gg n$).

Define the parameter vector $\boldsymbol{\theta}$ and the information vector $\boldsymbol{\varphi}(t)$ as

$$\boldsymbol{\theta}:=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{b}\end{bmatrix}\in\mathbb{R}^n,\quad \boldsymbol{\varphi}(t):=\begin{bmatrix}\boldsymbol{\phi}(t)\\ \boldsymbol{\psi}(t)\end{bmatrix}\in\mathbb{R}^n,$$

where

$$\boldsymbol{a}:=[a_1,a_2,\ldots,a_{n_a}]^{\mathrm{T}}\in\mathbb{R}^{n_a},$$
$$\boldsymbol{b}:=[b_1,b_2,\ldots,b_{n_b}]^{\mathrm{T}}\in\mathbb{R}^{n_b},$$
$$\boldsymbol{\phi}(t):=[-x(t-1),-x(t-2),\ldots,-x(t-n_a)]^{\mathrm{T}}\in\mathbb{R}^{n_a},$$
$$\boldsymbol{\psi}(t):=[u(t-1),u(t-2),\ldots,u(t-n_b)]^{\mathrm{T}}\in\mathbb{R}^{n_b}. \tag{3}$$

Then (2) and (1) can be written as

$$x(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{a}+\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{b}, \tag{4}$$

$$y(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{a}+\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{b}+v(t). \tag{5}$$

Define the stacked output vector $\boldsymbol{Y}$, the stacked information matrices $\boldsymbol{\Phi}$ and $\boldsymbol{U}$ as

$$\boldsymbol{Y}:=\begin{bmatrix}y(1)\\ y(2)\\ \vdots\\ y(L)\end{bmatrix}\in\mathbb{R}^L,\quad \boldsymbol{\Phi}:=\begin{bmatrix}\boldsymbol{\phi}^{\mathrm{T}}(1)\\ \boldsymbol{\phi}^{\mathrm{T}}(2)\\ \vdots\\ \boldsymbol{\phi}^{\mathrm{T}}(L)\end{bmatrix}\in\mathbb{R}^{L\times n_a},$$

$$\boldsymbol{U}:=\begin{bmatrix}\boldsymbol{\psi}^{\mathrm{T}}(1)\\ \boldsymbol{\psi}^{\mathrm{T}}(2)\\ \vdots\\ \boldsymbol{\psi}^{\mathrm{T}}(L)\end{bmatrix}\in\mathbb{R}^{L\times n_b}.$$

Note that the matrix $\boldsymbol{U}$ and the vector $\boldsymbol{Y}$ contain all the measured data $\{u(t),y(t):t=0,1,2,\ldots,L\}$, and the matrix $\boldsymbol{\Phi}$ is unknown because the true output terms (i.e., the noise-free output terms) in $\boldsymbol{\Phi}$ are the unknown inner variables.

According to (5), define a cost function

$$J(\boldsymbol{\theta}):=\sum_{t=1}^{L}[y(t)-\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{a}-\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{b}]^2=\|\boldsymbol{Y}-\boldsymbol{\Phi}\boldsymbol{a}-\boldsymbol{U}\boldsymbol{b}\|^2,$$

where the norm of the vector $\boldsymbol{x}$ is defined as $\|\boldsymbol{x}\|^2:=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

Minimizing $J(\boldsymbol{\theta})$ and letting the partial derivative of $J(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ be zero give

$$\frac{\partial J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=-2\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\\ \boldsymbol{U}^{\mathrm{T}}\end{bmatrix}[\boldsymbol{Y}-\boldsymbol{\Phi}\boldsymbol{a}-\boldsymbol{U}\boldsymbol{b}]=-2\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\\ \boldsymbol{U}^{\mathrm{T}}\end{bmatrix}\left(\boldsymbol{Y}-[\boldsymbol{\Phi},\boldsymbol{U}]\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{b}\end{bmatrix}\right)=\mathbf{0}.$$

Provided that the involved matrix is invertible, we can obtain the relation:

$$\boldsymbol{\theta}=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{b}\end{bmatrix}=\left(\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\\ \boldsymbol{U}^{\mathrm{T}}\end{bmatrix}[\boldsymbol{\Phi},\boldsymbol{U}]\right)^{-1}\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\\ \boldsymbol{U}^{\mathrm{T}}\end{bmatrix}\boldsymbol{Y}$$
$$=\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} & \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{U}\\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Phi} & \boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}\\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{Y}\end{bmatrix}. \tag{6}$$

However, since $\boldsymbol{\Phi}$ is unknown, it is impossible to calculate the parameter estimation vector $\boldsymbol{\theta}$ via the above equation directly. To solve this difficulty, the solution is based on the hierarchical identification principle [37,38].

Let $k=1,2,3,\ldots$ be an iteration variable and $\hat{\boldsymbol{\theta}}_k:=\begin{bmatrix}\hat{\boldsymbol{a}}_k\\ \hat{\boldsymbol{b}}_k\end{bmatrix}$ be the iterative estimate of $\boldsymbol{\theta}$. Let $\hat{x}_k(t-i)$ be the estimate of $x(t-i)$ at iteration k, and define the estimates:

$$\hat{\boldsymbol{\phi}}_k(t):=[-\hat{x}_{k-1}(t-1),-\hat{x}_{k-1}(t-2),\ldots,-\hat{x}_{k-1}(t-n_a)]^{\mathrm{T}}\in\mathbb{R}^{n_a},$$

$$\hat{\boldsymbol{\Phi}}_k:=\begin{bmatrix}\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(1)\\ \hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(2)\\ \vdots\\ \hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(L)\end{bmatrix}\in\mathbb{R}^{L\times n_a}. \tag{7}$$