Neural Networks 23 (2010) 396–405

Contents lists available at ScienceDirect

Neural Networks

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

A fast algorithm for AR parameter estimation using a novel noise-constrained least-squares method

Youshen Xia [a,*], Mohamed S. Kamel [b], Henry Leung [c]

[a] College of Mathematics and Computer Science, Fuzhou University, China

[b] Department of Electrical and Computer Engineering, University of Waterloo, Canada

[c] Department of Electrical and Computer Engineering, University of Calgary, Canada

ARTICLE INFO

Article history:
Received 24 April 2009
Received in revised form 18 October 2009
Accepted 9 November 2009

Keywords:
Autoregressive parameter estimation
Noise-constrained least-squares method
Fast learning algorithm
Fixed step length
Robust estimate
Mean-square error

ABSTRACT

In this paper, a novel noise-constrained least-squares (NCLS) method for online autoregressive (AR) parameter estimation is developed under blind Gaussian noise environments, and a discrete-time learning algorithm with a fixed step length is proposed. It is shown that the proposed learning algorithm converges globally to an AR optimal estimate. Compared with conventional second-order and high-order statistical algorithms, the proposed learning algorithm can obtain a robust estimate which has a smaller mean-square error than the conventional least-squares estimate. Compared with the learning algorithm based on the generalized least absolute deviation method, instead of minimizing a non-smooth linear L_1 function, the proposed learning algorithm minimizes a quadratic convex function and thus is suitable for online parameter estimation. Simulation results confirm that the proposed learning algorithm can obtain more accurate estimates with a fast convergence speed.

1. Introduction

Autoregressive (AR) parameter estimation is a standard problem in spectral estimation, detection and array signal processing (Haykin, 2001; Kay, 1988). It is also important for autoregressive moving average (ARMA) system identification since the ARMA estimation problem can be transformed to an equivalent AR estimation problem (Kizilkaya & Kayran, 2005). AR applications including forecasting, economics, speech processing, and system identification have been found (Amari, Chen, & Cichocki, 2002; Cichocki & Amari, 2002; Liu, Kadirkamanathan, & Billings, 1998; Mandic & Chambers, 2001; Schrauwen, D'Haene, & Verstraeten, 2008). This paper focuses on the optimal AR parameter estimation algorithm with prior knowledge of the model order. For model order determination, readers are referred to (Davis & Zheng, 1997).

It is well known that the most basic and commonly used estimation scheme is the least-squares (LS) method. The LS method is very appropriate for online identification and is asymptotically unbiased when the noise distribution is white. In practice, however, the measured AR signal is usually corrupted by noise which is far from white (Chen & Chen, 2000; Samonas & Petrou, 2002). As a result, the LS method gives a biased estimate of the true parameters and will be poor in the worst case. To avoid the bias problem, many significant methods for improving the LS solution, such as the instrumental variable method (Soderstrom & Stoica, 1981), the extended least-squares method (Kay, 1988), improved least-squares methods (Zheng, 1998, 1999, 2000), and generalized least-squares method (Hwang, Kim, & Lee, 2007) have been developed. These methods can be shown to be efficient under Gaussian noise environments. In order to deal with the situation in Gaussian-mixture and non-Gaussian environments, maximum-likelihood method (Park & Chois, 2008; Verbout, Ooi, Ludwig, & Oppenheim, 1999), high-order statistical approaches (Aboutajdine, Adib, & Meziane, 1996; Giannakis & Mendel, 1990) were developed. To overcome the difficulty of the noise sensitivity in estimating the AR parameter, robust estimation methods (Cadzow, 2002; Eweda, 1989; Mandic, 2004; Rojo-Alvarez, Martinez-Ramon, de Prado-Cumplido, Artes-Rodriguez, & Figueiras-Vidal, 2004) were presented. Recently, a generalized least absolute deviation (GLAD) method for the optimal AR parameter estimation was developed in Xia and Kamel (2008). Simulation results showed that the GLAD method can obtain a better AR parameter estimate in the presence of non-Gaussian measurement noise than the LAD method. Since the cost function of the GLAD method is non-smooth, the resulting GLAD algorithm will have a slow convergence rate when the number of sample points increases, and thus is not suitable for online identification.

In this paper, a novel constrained least-squares (NCLS) method is developed and a discrete-time learning algorithm for fast AR parameter estimation is proposed. It is shown that the proposed

* Corresponding author. Tel.: +86 51922865157.
E-mail address: ysxia2001@yahoo.com (Y. Xia).

0893-6080/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2009.11.004

learning algorithm is globally convergent to an optimal AR estimate under a fixed step length. Unlike the conventional LS methods (Besnerais, Bercher, & Demoment, 1999; Haykin, 2001; Shi, Jing, & Henry, 2003), the proposed NCLS method can optimize both AR parameter variables and noise error variables. The proposed NCLS method has a good performance in that the AR optimal estimate obtained by the proposed learning algorithm is robust and has a smaller mean-square error than the conventional LS estimate. Compared with the GLAD learning algorithm, which minimizes a non-smooth cost function, the proposed learning algorithm minimizes a constrained convex cost function and thus is suitable for real-time parameter estimation (Cichocki & Amari, 2002; Mandic & Chambers, 2001). Simulation results demonstrate that the proposed learning algorithm has a fast convergence speed and can get more accurate estimates than several conventional algorithms in different noise environments.

The paper is organized as follows. In Section 2, AR parameter models and the least-squares estimation methods are described and a novel constrained least-squares (NCLS) method is then presented. In Section 3, an algorithm is proposed to implement the NCLS method. In Section 4, the performance analysis of the proposed NCLS method and its algorithm are given. Computer simulations are performed to evaluate the effectiveness of the proposed NCLS method in Section 5. Section 6 gives the conclusion of this paper.

2. AR model and constrained least-squares methods

Consider the following blind linear system model:

$$x(t) = \sum_{i=1}^{p} a_i^* x(t-i) + v(t), \tag{1}$$

where p is the known order of the model, $\mathbf{a}^* = [a_1^*, \ldots, a_p^*]^T$ is the unknown AR parameter vector (non-random vector), $v(t)$ is the driving noise, and $x(t)$ is observed signal (system output) with additive measurement noise $u(t)$:

$$y(t) = x(t) + u(t), \tag{2}$$

and $u(t)$ is the measurement noise, being uncorrelated with $v(t)$. For simplicity, we denote the noisy observation vector by $\mathbf{y_t} = [y(t-1), \ldots, y(t-p)]^T$ and the measurement noise vector by $\mathbf{u_t} = [u(t-1), \ldots, u(t-p)]^T$. Then the AR observation model can be written as

$$y(t) = \mathbf{y_t}^T \mathbf{a}^* - n(t), \tag{3}$$

where $n(t) = \mathbf{u_t}^T \mathbf{a}^* - u(t) - v(t)$ is not white noise. Let $\mathbf{y} = [y(1), \ldots, y(N)]^T$, $\mathbf{n} = [n(1), \ldots, n(N)]^T$, and

$$B = \begin{pmatrix} y(0) & y(-1) & \ldots & y(1-p) \\ y(1) & y(0) & \ldots & y(2-p) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \ldots & y(N-p) \end{pmatrix},$$

(3) can be written as a linear equation in the matrix and vector form

$$B\mathbf{a}^* - \mathbf{y} - \mathbf{n} = \mathbf{0}, \tag{4}$$

where $\mathbf{n}$ is the noise vector. The problem is to estimate $\mathbf{a}^*$ from the noisy observations $\{y(t)\}_1^N$, based on (4). The conventional approach to estimate the AR parameter vector is the least-squares (LS) method. The LS method minimizes

$$E(\mathbf{a}) = \frac{1}{N} \sum_{t=1}^{N} (y(t) - \mathbf{y_t}^T \mathbf{a})^2$$

and gives the LS estimate

$$\mathbf{a}_{LS} = \left(\frac{1}{N} \sum_{t=1}^{N} \mathbf{y_t} \mathbf{y_t}^T \right)^{-1} \left(\frac{1}{N} \sum_{t=1}^{N} \mathbf{y_t} y(t) \right),$$

where $\mathbf{a} = [a_1, \ldots, a_p]^T$. There is an error between $\mathbf{a}_{LS}$ and the true parameter $\mathbf{a}^*$ (Zheng, 1998):

$$\mathbf{a}_{LS} = \mathbf{a}^* + R_y^{-1}(E[\mathbf{u_t} n(t)] - E[\mathbf{u_t}\mathbf{u_t}^T]\mathbf{a}^*),$$

where $E[\cdot]$ denotes the statistical expectation. To improve the LS estimate, many estimation methods have been developed in the past decades. Recently, a generalized least absolute deviation (GLAD) method was developed in Xia and Kamel (2008). It obtains an estimate by finding a solution to the following optimization problem:

$$\begin{aligned} \min \quad & \|B\mathbf{a} - \mathbf{y} - \mathbf{z}\|_1 \\ \text{s.t} \quad & \mathbf{a} \in R^p,\ \mathbf{z} \in \Omega_\gamma^0, \end{aligned}$$

where $\|\cdot\|_1$ denotes the l_1 norm, $\Omega_\gamma^0 = \{\mathbf{z} \in R^N \mid \gamma_1 \mathbf{m} \le \mathbf{z} \le \gamma_2 \mathbf{m}\}$, $\mathbf{m} = |E[y(t)]|\mathbf{e}$, $\mathbf{e} = [1, \ldots, 1]^T \in R^N$, and γ_1 and γ_2 are design constants. The GLAD estimate was shown to improve both the convention LS estimate and the LAD estimate in the presence of non-Gaussian measurement noise from the viewpoint of simulation. However, the cost function of the GLAD method is non-smooth and the convergence rate of the resulting GLAD algorithm becomes very slow. Moreover, it is difficult to analyze why the GLAD estimate can improve the conventional LS estimate theoretically.

To overcome the disadvantages of the GLAD method, we propose a noise-constrained least-squares (NCLS) method that obtains an optimal solution, $(\mathbf{a}_\gamma^*, \mathbf{z}^*)$, by solving the following quadratic problem:

$$\begin{aligned} \min \quad & \|B\mathbf{a} - \mathbf{y} - \mathbf{z}\|_2^2 \\ \text{s.t} \quad & \mathbf{a} \in R^p,\ \mathbf{z} \in \Omega_\gamma, \end{aligned} \tag{5}$$

where $\|\cdot\|_2$ denotes the l_2 norm, $\Omega_\gamma = \{\mathbf{z} \in R^N \mid \gamma_1(B\mathbf{q} - \mathbf{y}) \le \mathbf{z} \le \gamma_2(B\mathbf{q} - \mathbf{y})\}$, $\mathbf{q}$ is a constant vector, and the design constants γ_1 and γ_2 are obtained by using the cross-validation procedure to minimize the following prediction function:

$$\psi(\mathbf{a}, \gamma_1, \gamma_2) = \frac{\|\mathbf{a} - \mathbf{a}^*\|_2}{\|\mathbf{a}_{LS} - \mathbf{a}^*\|_2}.$$

The optimal AR estimate is then given by $\mathbf{a}_\gamma^*$.

There are three differences between the GLAD method and the NCLS method. First, the GLAD optimization problem is non-smooth and the NCLS optimization problem is quadratic. Thus the proposed NCLS method can be easily implemented by a NCLS algorithm with fast convergence rate (see later discussion). More importantly, unlike the GLAD method, the proposed NCLS method can be shown to be more accurate than the conversional least-squares (LS) method in terms of mean-square error. Next, let $\mathbf{q} = B^+(\mathbf{m} + \mathbf{y})$, where B^+ is the pseudo-inverse of B. Then Ω_γ becomes Ω_γ^0, which indicates $\Omega_\gamma^0 \subset \Omega_\gamma$. In the NCLS method, we may have set $\Omega_\gamma = \{\mathbf{n} = B\mathbf{a}^* - \mathbf{y}\}$ by taking $\gamma_1 = \gamma_2 = 1$ and $\mathbf{q} = \mathbf{a}^*$. Yet, the set $\{\mathbf{n} = B\mathbf{a}^* - \mathbf{y}\}$ is not available in the GLAD method. Third, simulation results show that the optimal design parameters can be obtained from the prediction error of the proposed NCLS method, unlike the GLAD method.

3. Discrete-time learning algorithm

There are many numerical algorithms for solving the NCLS problem (Cadzow, 2002; Gafni & Bertsekas, 1984; Grant, Boyd, & Ye, 2005; Li & Swetits, 1997; Solodov & Tseng, 1996), but they have at least a computational complexity $O(N(N+p))$ per iteration. We introduce here a discrete-time learning algorithm with a computational complexity $O(pN)$ per iteration.

3.1. Reformulation of NCLS problems

For convenience of discussion, we rewrite the NCLS problem (5) as follows: