Digital Signal Processing 56 (2016) 15–23

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Performance analysis of multi-innovation stochastic Newton recursive algorithms

Pengfei Cao [a,*], Xionglin Luo [b]

[a] College of Electrical Engineering and Automation, Shandong University of Science and Technology, China
[b] Research Institute of Automation, China University of Petroleum, Beijing 102249, China

ARTICLE INFO

Article history:
Available online 27 May 2016

Keywords:
System identification
Recursive algorithm
Multi-innovation
Convergence analysis

ABSTRACT

The stochastic Newton recursive algorithm is studied for system identification. The main advantage of this algorithm is that it has extensive form and may embrace more performance with flexible parameters. The primary problem is that the sample covariance matrix may be singular with numbers of model parameters and (or) no general input signal; such a situation hinders the identification process. Thus, the main contribution is adopting multi-innovation to correct the parameter estimation. This simple approach has been proven to solve the problem effectively and improve the identification accuracy. Combined with multi-innovation theory, two improved stochastic Newton recursive algorithms are then proposed for time-invariant and time-varying systems. The expressions of the parameter estimation error bounds have been derived via convergence analysis. The consistence and bounded convergence conclusions of the corresponding algorithms are drawn in detail, and the effect from innovation length and forgetting factor on the convergence property has been explained. The final illustrative examples demonstrate the effectiveness and the convergence properties of the recursive algorithms.

1. Introduction

Parameter estimation methods have elicited considerable attention in system modelling, signal processing and adaptive control. (e.g. [13,17,26–32]). To improve the estimation efficiency and realize model update online, recursive identification algorithms have been developed and corresponding performance analysis has also been conducted [7,16,21,24]. In this paper, we mainly focus on the algorithm performance for time-invariant and time-varying single-rate systems. Although considerable work has been published on recursive algorithms, further research in this area is still required.

Numerous studies have emphasized several specific algorithms, such as recursive least squares (RLS) and stochastic gradient (SG) algorithms. Each algorithm has its own advantages and disadvantages with specific form [8,9]. Although numerous achievements have been obtained, research is lacking on the more extensive algorithms. The stochastic Newton recursive (SNR) algorithm is based on the gradient-descent idea and employs sample covariance matrix to control the update directions [15,33]. Compared with other recursive algorithms, the SNR algorithm has a more general form, based on which the internal relations of other algorithms can be revealed [6,12,14,15]. Concurrently, more possibilities exist for the algorithm parameters, such as the forgetting factor. The SNR algorithm shows higher research value and has been used in several areas of identification, control and signal processing [35,36]. Therefore, the SNR algorithm is studied in this paper. Following the basic form, two specific algorithms are derived for time-invariant and time-varying systems in this study.

However, when the input signal is insufficiently general and (or) the number of model parameters is large, the sample covariance matrix may be singular or approach singularity [35]. The ill-posed problem arises when inverting the covariance matrix (as shown in Eqs. (8) and (9)), which may hinder the parameter identification. A similar problem may appear in many other algorithms [5,4,11,37,38]. Certain methods have been explored to solve this problem, in which adding one positive definite matrix is the main approach [1,2,14]. A diagonal matrix is usually adopted. However, specific conditions should be satisfied for the matrix to guarantee the parameter convergence, and the determination of the matrix is not easy [1,2]. In the meanwhile, the values of the matrix would influence the convergence property of the algorithm. Ding and Chen presented multi-innovation identification theory and then proposed several multi-innovation-based algorithms [8,9,5,4,11,37,38,40]. The basic idea of multi-innovation theory is utilizing historical input–output data together with the current data at each recursion to update the parameter estimations; the parameter es-

* Corresponding author. Postal address: College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, 266590, China.
E-mail address: cpfskdzdh@sina.com (P. Cao).

http://dx.doi.org/10.1016/j.dsp.2016.05.005
1051-2004/

timation accuracy has been proven to improve [8–10]. The convergence analysis in Section 3 shows that the singular covariance matrix is avoided based on the multi-innovation method. Therefore, combined with multi-innovation theory, multi-innovation stochastic Newton recursive (MISNR) algorithms are proposed. Although this improvement is modest, the problem has been solved easily and the calculation is not excessively increased. Compared with the method of adding positive definite matrix, no additional efforts should be made to determine the matrix. Thus, the proposed algorithm is easy to operate in applications.

The remainder of the paper is organized as follows. In Section 2, two MISNR algorithms are derived. In Section 3, the consistency of the corresponding algorithm for time-invariant systems is proven. In Section 4, based on the algorithm for time-varying systems, the parameter estimations are proven boundedly convergent. In Section 5, several illustrative examples for the results in this paper are presented. Section 6 provides the concluding remarks.

2. The MISNR algorithms

In this section, the SNR algorithm is introduced first. Combined with multi-innovation theory, two MISNR algorithms are proposed for time-invariant and time-varying SISO (single-input and single-output) systems. The algorithms are then extended to MIMO (multi-input and multi-output) systems.

2.1. The MISNR algorithms for SISO systems

Consider the typical linear regression model [8,12]

$$A(z)y(k) = B(z)u(k) + v(k)$$
$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}$$
$$B(z) = b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}$$

in which z represents the forward shift operator and $z^{-1}x(k) = x(k-1)$. The model is rewritten as

$$y(k) = \varphi^\tau(k)\theta(k) + v(k) \tag{1}$$
$$\varphi^\tau(k) = \big[-y(k-1), -y(k-2), \dots, - y(k-n_a), u(k-1), u(k-2), \dots, u(k-n_b)\big]$$
$$\theta(t) = [a_1, a_2, \dots, a_{n_a}, b_1, b_2, \dots, b_{n_b}]^\tau$$

in which $y(k)$, $u(k)$ and $v(k)$ represent the output, input and noise variables, respectively; $\varphi(k) \in R^{n\times 1}$ is the information vector including historical measured data; $\theta(k) \in R^{n\times 1}$ is the parameter vector, and $n = n_a + n_b$.

The SNR algorithm is presented below [14,15,33]:

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \rho(k)R^{-1}(k)\varphi(k)\big[y(k) - \varphi^\tau(k)\hat{\theta}(k-1)\big] \\ R(k) = R(k-1) + \rho(k)\big[\varphi(k)\varphi^\tau(k) - R(k-1)\big] \end{cases} \tag{2}$$

where $e(k) = y(k) - \varphi^\tau(k)\hat{\theta}(k-1)$ is defined as innovation. $\rho(k)$ represents the forgetting factor, and $R(k)$ represents the sample covariance matrix. The first equation in Eq. (2) is the main part of the algorithm, and $R(k)$ can be substituted with certain constant matrix.

No fixed expression exists for $\{\rho(k)\}$, but $\rho(k)$ satisfies the basic condition for time-invariant systems [23,22] to ensure the consistency of the algorithm:

$$\begin{cases} 0 < \rho(k) < 1 \\ \sum_{k=1}^{\infty} \rho(k) = \infty, \ \sum_{k=1}^{\infty} \rho^2(k) < \infty \end{cases} \tag{3}$$

For time-varying systems, the algorithm of Eq. (2) is unable to track time-varying parameters for that $\rho(k)R^{-1}(k)\varphi(k)$ approaches to zero as k increases (according to Eq. (3) and Theorem 1). To improve the tracking performance of the algorithm, $\rho(k)$ is assumed equal to some positive constant λ $(0 < \lambda < 1)$ [8,9,12,14,15]; the corresponding algorithm is

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \lambda R^{-1}(k)\varphi(k)\big[y(k) - \varphi^\tau(k)\hat{\theta}(k-1)\big] \\ R(k) = R(k-1) + \lambda\big[\varphi(k)\varphi^\tau(k) - R(k-1)\big] \end{cases} \tag{4}$$

This algorithm has been proven effective for tracking time-varying parameters when adopting a constant forgetting factor [8–10]. These two recursive algorithms (Eqs. (2), (4)) are obtained by minimizing the cost function $J(\theta) := E[\|y(k) - \varphi^\tau(k)\theta\|^2]$; only the innovation scalar $e(k)$ is utilized to correct the parameter estimations. Here, $\|X\|^2 = tr(XX^\tau)$ and $tr(\bullet)$ represents the trace of the matrix. However, when $R(k)$ appears singular or approaches singularity, its inversion cannot be calculated and the identification process ceases. To solve this problem, we attempt to extend innovation scalar to the innovation vector. Correspondingly, the cost function is expressed as

$$J(\theta) := E\big[\big\|Y(k) - \Gamma^\tau(k)\theta\big\|^2\big] \tag{5}$$
$$Y(k) = \big[y(k), y(k-1), \dots, y(k-N+1)\big]^\tau \in R^N \tag{6}$$
$$\Gamma(k) = \big[\varphi(k), \varphi(k-1), \dots, \varphi(k-N+1)\big] \in R^{n\times N} \tag{7}$$

Based on the cost function, deriving the MISNR algorithm as

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \rho(k)R^{-1}(k)\Gamma(k)\big(Y(k) - \Gamma^\tau(k)\hat{\theta}(k-1)\big) \\ R(k) = R(k-1) + \rho(k)\big(\Gamma(k)\Gamma^\tau(k) - R(k-1)\big) \end{cases} \tag{8}$$

For time-invariant systems, the values of $\{\rho(k)\}$ also satisfy the basic rules in Eq. (3). If the systems are time-varying, we have the MISNR algorithm as

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \lambda R^{-1}(k)\Gamma(k)\big(Y(k) - \Gamma^\tau(k)\hat{\theta}(k-1)\big) \\ R(k) = R(k-1) + \lambda\big(\Gamma(k)\Gamma^\tau(k) - R(k-1)\big) \end{cases} \tag{9}$$

In Eqs. (8), (9), $E(k) = Y(k) - \Gamma^\tau(k)\hat{\theta}(k-1)$ is defined as innovation vector and N represents innovation length. The algorithms in Eqs. (8), (9) use both current and past data. $R(k)$ is always positive definite (Theorem 1 and Remark 1).

We have the flow diagrams for these two algorithms, see Fig. 1.

2.2. The MISNR algorithms for MIMO systems

The model for MIMO systems is expressed as

$$A(z)\mathbf{y}(k) = B(z)\mathbf{u}(k) + \mathbf{v}(k)$$
$$\mathbf{y}(k) = \big[y_1(k), y_2(k), \dots, y_m(k)\big]^\tau$$
$$\mathbf{u}(k) = \big[u_1(k), u_2(k), \dots, u_m(k)\big]^\tau$$
$$\mathbf{v}(k) = \big[v_1(k), v_2(k), \dots, v_m(k)\big]^\tau$$
$$A(z) = I + A_1 z^{-1} + A_2 z^{-2} + \cdots + A_{n_a} z^{-n_a},$$
$$B(z) = B_1 z^{-1} + B_2 z^{-2} + \cdots + B_{n_b} z^{-n_b},$$

where $y(k) \in R^{m\times 1}$, $u(k) \in R^{m\times 1}$ and $v(k) \in R^{m\times 1}$, $A(z) \in R^{m\times m}$ and $B(z) \in R^{m\times m}$ are polynomial matrices; I is unit diagonal matrix. The model can be rewritten as

$$\mathbf{y}(k) = \theta^\tau(k)\varphi(k) + \mathbf{v}(k)$$
$$\varphi(k) = \big[-\mathbf{y}^\tau(k-1), \dots, -\mathbf{y}^\tau(k-n_a), \mathbf{u}^\tau(k-1), \dots, \mathbf{u}^\tau(k-n_b)\big]^\tau \in R^{m(n_a+n_b)\times 1}$$
$$\theta^\tau(k) = [A_1, A_2, \dots, A_{n_a}, B_1, B_2, \dots, B_{n_b}] \in R^{m\times m(n_a+n_b)}$$