Automatica 50 (2014) 2120–2127

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Almost sure convergence rates for system identification using binary, quantized, and regular sensors[☆]

Hongwei Mei[a], Le Yi Wang[b], George Yin[a,1]

[a] Department of Mathematics, Wayne State University, Detroit, MI 48202, United States

[b] Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, United States

ARTICLE INFO

Article history:
Received 30 September 2013
Received in revised form
19 February 2014
Accepted 14 April 2014
Available online 1 July 2014

Keywords:
System identification
Quantized sensor
Almost sure convergence rate
Laws of iterated logarithms

ABSTRACT

This paper presents almost sure convergence rates for system identification under binary, quantized, and regular sensors. To accommodate practical model complexity constraints, the system under consideration is represented by a modeled part together with an unknown-but-bounded unmodeled dynamics. Under uncorrelated noise sequences, identification errors with different sensor types are studied and tight error bounds are obtained without information or constraints on noise moment conditions. The results are then extended to correlated noise sequences whose remote past and distant future are asymptotically independent. In both cases, almost sure error bounds of the laws of iterated logarithms type are derived.

1. Introduction

This paper focuses on the rate of convergence analysis for system identification algorithms under observations using binary, quantized, and regular sensors. To contain model complexity, the systems under consideration are represented by a modeled part together with unknown-but-bounded unmodeled dynamics. Our main effort is on establishing almost sure or strong convergence rates of the algorithms. Under broad conditions, error bounds of iterated logarithm types are obtained.

System identification has been studied extensively in the past several decades; see Chen and Guo (1991), Ljung (1987), Milanese and Tempo (1985), Milanese and Vicino (1993), Wang (1997), Wang and Yin (2002), and Wang and Yin (2007), Yin, Wang, and Kan (2009) among others; see also related work of Kushner and Yin (2003). One of the most recent efforts has been devoted to the study of system identification under binary and quantized sensors. In the early 2000s, we began our quest on the subject in Wang, Zhang, and Yin (2003). Since then, this line of work has been significantly expanded. Many of the recent results have been documented in Wang, Yin, Zhang, and Zhao (2010). For parameter estimation under binary or quantized observations together with unmodeled dynamics, under broad conditions, it has been shown that the estimators converge strongly (with probability one) to a neighborhood that centers on the true parameter values. In addition, scaled sequences of the estimation errors converge in distribution to a normal random variable, which characterizes convergence rates with the scaling factor together with the asymptotic covariance. The convergence is in the sense of convergence in distribution. However, there has not been any rate-of-convergence results in the almost sure sense yet. This paper aims to derive such convergence rates.

In this work, we examine system identification accuracy and convergence rates under binary and quantized observations with unmodeled dynamics. This is our first attempt to resolve the strong rate-of-convergence issues under different sensor types and unmodeled dynamics. Our aim is to obtain tight error estimates in the context of the laws of iterated logarithms. Our investigation starts with the case of independent and identically distributed (i.i.d.) noises, which enables us to obtain tight error bounds and clarify understanding of relations among data size, unmodeled dynamics, and sensor complexity. The results are further expanded to accommodate more practical correlated noise processes of mixing types whose remote past and distant future are asymptotically independent. Estimation error bounds and

☆ This research was supported in part by the Army Research Office under grant W911NF-12-1-0223. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Antonio Vicino under the direction of Editor Torsten Söderström.

E-mail addresses: hongwei.mei@wayne.edu (H. Mei), lywang@wayne.edu (L.Y. Wang), gyin@math.wayne.edu (G. Yin).

[1] Tel.: +1 313 577 2479; fax: +1 313 577 7596.

http://dx.doi.org/10.1016/j.automatica.2014.05.036
0005-1098/

strong convergence rates are derived. Impact of sensor types and quantization levels is investigated within the same framework, which characterizes sensor complexity issues. Understanding from these findings will be of utility for a wide spectrum of applications involving sensor limitations and communication channels, especially networked systems.

The rest of the paper is arranged as follows. Section 2 begins with the formulation of the system identification problem. Section 3 is devoted to the study of systems with observations using binary sensors. We first consider independent and identically distributed noise sequences, which lead to a clean representation of the asymptotic strong convergence rates. The results are then extended to include correlated noises of mixing types. Section 4 treats observations with quantized sensors. For comparison of impact of sensor limitations, in Section 5, we examine the almost sure error bounds when regular sensors are used in identification. Section 6 presents a couple of examples and issues several further remarks to conclude the paper.

2. Formulation

Consider a single-input–single-output linear time-invariant, stable, discrete-time system

$$y(t) = \sum_{i=0}^{\infty} a_i u(t-i) + d(t), \quad t = 0, 1, \ldots, \tag{1}$$

where $d(t)$ is the disturbance, $u(t)$ is the input with $u(t) = 0$ when $t < 0$; and $a = \{a_i : i \geq 0\}$ satisfies $\sum_{i=0}^{\infty} |a_i| < \infty$. In what follows, $\|\cdot\|$ denotes the ℓ_1 norm, i.e., $\|a\| = \sum_{i=0}^{\infty} |a_i|$. For simplicity, the starting time is set at $t = 0$, although this is not essential. The vector $a = \{a_i : i \geq 0\}$ represents the unknown parameter. For a selected model order n, the parameter vector is split into the n-dimensional vector $\theta = (a_0, \ldots, a_{n-1})' \in \mathbb{R}^n$, known as the modeled part, and the possibly infinite dimensional vector $\widetilde{\theta} = (a_n, a_{n+1}, \ldots)'$, which represents the unmodeled dynamics. In the above and hereafter, z' denotes the transpose of z. Our task is to identify θ under noisy observation. In addition to the modeled part, the system output is impacted by both unmodeled dynamics and observation noise

$$y(t) = \phi'(t)\theta + \tilde{\phi}'(t)\tilde{\theta} + d(t), \tag{2}$$

where $\phi'(t) = (u(t), \ldots, u(t-n+1))$ and $\tilde{\phi}'(t) = (u(t-n), u(t-n-1), \ldots)$.

The convergence in the almost sure sense has been established in our previous work; see Wang et al. (2010, Chapter 3) for a summary. This paper aims to obtain convergence rate result in the almost sure sense. To illustrate the desired result, we consider the binary sensor without unmodeled dynamics for simplicity. Under suitable assumptions, $\|\widehat{\theta}_K - \theta\| \to 0$ w.p.1 as $K \to \infty$. This strong consistency assertion is more or less a law of large number result. The rate of convergence of the sequence of parameter estimates can be considered for the scaled sequence of centered estimation errors $\sqrt{K}(\widehat{\theta}_K - \theta)$, which converges in the sense of in distribution to a normal random variable. So one way of treating the rate of convergence is to use the asymptotic normality together with the scaling factor $\sqrt{K}$ and the asymptotic covariance, which indicates how the estimation errors depend on the scaling factor and how much variation the estimates have. The law of iterated logarithm to be obtained in this paper is a sharp result and describes the magnitude of the fluctuations in the almost sure sense. It is easily seen that although $\sqrt{K}(\widehat{\theta}_K - \theta)$ convergence in distribution, it is divergent in the almost sure sense. A question then is: What is the appropriate scaling factor Δ_K so that $\Delta_K \sqrt{K}(\widehat{\theta}_K - \theta)$, the scaled sequence of error will be convergent in the almost sure sense. This paper provides a very precise bound in a pathwise sense. It demonstrates that $\Delta_K = (\log\log K)^{-1/2}$.

Throughout this paper we assume that the input $\{u(t)\}$ is n-periodic and denote its $n \times n$ Toeplitz matrix by

$$\Phi = \begin{pmatrix} u(0) & u(n-1) & \ldots & u(2) & u(1) \\ u(1) & u(0) & \ldots & u(3) & u(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u(n-2) & u(n-3) & \ldots & u(0) & u(n-1) \\ u(n-1) & u(n-2) & \ldots & u(1) & u(0) \end{pmatrix}.$$

Assuming that $u(t)$ is full rank implies that Φ is invertible. The unmodeled dynamics is unknown but bounded by $\|\tilde{\theta}\| \leq \varepsilon$ for some $\varepsilon > 0$. Since $u(t)$ is periodic, it is bounded with $\|u\|_\infty \leq c_0 < \infty$ for some $c_0 > 0$. Consequently,

$$\sup_t |\tilde{\phi}'(t)\tilde{\theta}| \leq \|u\|_\infty \|\tilde{\theta}\|_1 \leq c_0 \varepsilon := \varepsilon_1.$$

For simplicity and without loss of generality, in this paper we normalize the input to $\|u\|_\infty = 1$, leading to $c_0 = 1$ and $\varepsilon_1 = \varepsilon$. Write $\log x = \ln(x \vee e)$, where $a \vee b = \max\{a, b\}$ for two real numbers a and b. Define $\{\widetilde{u}(i) : -n \leq i \leq n-1\}$ as

$$\widetilde{u}(i-n) = \widetilde{u}(i) = u(i) \quad \text{when } 0 \leq i \leq n-1.$$

For $0 \leq j \leq n-1$ and $k \geq 0$, set

$$\alpha_j := \sum_{l=0}^{\infty} a_{ln+j}, \qquad \beta_j := \sum_{i=0}^{n-1} \widetilde{u}(j-i)\alpha_i$$

$$\begin{aligned} \gamma_j^k &:= \beta_j - \sum_{i=0}^{\infty} a_i u(kn+j-i) \\ &= \beta_j - \sum_{i=0}^{n-1} \widetilde{u}(j-i) \sum_{l=0}^{i+ln \leq kn+j} a_{ln+i}. \end{aligned}$$

Then we have

$$\begin{aligned} &\beta = \Phi\alpha, \quad \text{where } \beta = (\beta_0, \ldots, \beta_{n-1})' \\ &\alpha = (\alpha_0, \ldots, \alpha_{n-1})'. \end{aligned} \tag{3}$$

Henceforth, the notation $O(y)$ denotes a vector-valued function of y such that $\|O(y)\|/\|y\| \leq \kappa$ for some $\kappa > 0$, and $o(y)$ denotes a function such that $\lim_{y \to b}(\|o(y)\|/\|y\|) = 0$ with b being either 0 or ∞, where $\|\cdot\|$ can be any vector norm. Also, the phrase *almost surely* will be abbreviated to *a.s.* We also denote $M = \|u\|_\infty \times \|a\|$.

3. Binary sensors

In this section, we consider the case when output observations are measured by a binary-valued sensor

$$s(t) = I(y(t) \leq C) = \begin{cases} 1 & \text{if } y(t) \leq C, \\ 0 & \text{else.} \end{cases}$$

3.1. Independent and identically distributed noise sequences

In this section we assume: The disturbance $\{d(t)\}$ is an i.i.d sequence of random variables. Its distribution function is F that has a density function f being continuous and positive on $[C - M - \sigma, C + M + \sigma]$ for some $\sigma > 0$. As a result, we deduce that f and $(d/dx)F^{-1}$ are bounded on $[C-M, C+M]$ and $[F(C-M), F(C+M)]$, respectively. In what follows, let K be the number of periods so that the discrete time elapsed is $k = nK$.