Automatica 52 (2015) 192–201

Contents lists available at ScienceDirect

Automatica

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

Recursive algorithms for parameter estimation with adaptive quantizer ☆

Keyou You

Department of Automation, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Received 8 April 2014
Received in revised form
20 August 2014
Accepted 7 November 2014
Available online 16 December 2014

Keywords:
Parameter estimation
Recursive algorithm
Quantization
Asymptotic analysis
Convergence speed

ABSTRACT

This paper studies a parameter estimation problem of networked linear systems with fixed-rate quantization. Under the minimum mean square error criterion, we propose a recursive estimator of stochastic approximation type, and derive a necessary and sufficient condition for its asymptotic unbiasedness. This motivates to design an adaptive quantizer for the estimator whose strong consistency, asymptotic unbiasedness, and asymptotic normality are rigorously proved. Using the Newton-based and averaging techniques, we obtain two accelerated recursive estimators with the fastest convergence speed of $O(1/k)$, and exactly evaluate the quantization effect on the estimation accuracy. If the observation noise is Gaussian, an optimal quantizer and the accelerated estimators are co-designed to asymptotically approach the minimum Cramer–Rao lower bound. All the estimators share almost the same computational complexity as the gradient algorithms with un-quantized observations, and can be easily implemented. Finally, the theoretical results are validated by simulations.

1. Introduction

Quantized estimation has long been an important research topic, and bears a vast body of literature (Curry, 1970; Gray & Neuhoff, 1998; Lloyd, 1982; Max, 1960; Papadopoulos, Wornell, & Oppenheim, 2002; Widrow & Kollár, 2008). While most of the early work deals with relatively high-rate quantization, they usually cannot handle the low-rate case. Under high-rate quantization, a common approach is to model quantization errors as an extra additive white noise, thereby allowing to use the standard solutions in stochastic theory. This approach considerably simplifies the problem but is not always reasonable for coarse quantization, and more powerful techniques are needed to handle low-rate quantization. Moreover, quantized estimation is fundamental in understanding the tradeoff between communication rate and estimation performance. This work is concerned with the co-design of the fixed-rate quantizer and estimator to identify the unknown parameters of linear systems.

The application of networked systems (Baillieul & Antsaklis, 2007), such as sensor networks, micro-electromechanical systems, mobile telephony, and industrial control networks, has greatly boosted the development of quantization methods. To date, many novel methods have been developed for the networked estimation with quantized observations, including Casini, Garulli, and Vicino (2012), Chen and Varshney (2010), Fang and Li (2008), Fu and de Souza (2009), Godoy, Goodwin, Agüero, Marelli, and Wigren (2011), Guo and Zhao (2013) Li and Alregib (2007), Marelli, You, and Fu (2013), Ribeiro, Giannakis, and Roumeliotis (2006), Shen, Varshney, and Zhu (2013), Wang and Yin (2007) and Xiao, Ribeiro, Luo, and Giannakis (2006), to name a few. The key challenges include that quantization is typically a highly nonlinear operator, and the estimator is no longer able to access the raw (un-quantized) observations. This may easily render the existing algorithms dramatically deviate from the true parameter or state. Under a moderate (e.g. one or two) bit rate constraint, a quantized observation can only supply few information of the system. It is important to well use each quantized sample. To make the most of every bit rate, we should smartly co-design the quantizer and estimator in a unified approach.

In Xiao et al. (2006), distributed compression and estimation in the context of wireless sensor networks have been addressed by using quantization technique. An interesting finding is that the estimation performance is quite sensitive to the quantizer threshold. Actually, to estimate θ under binary quantization of $y = \theta + v$, where v is a white noise, an optimal way to minimize the mean square error (MSE) is to simply place the quantizer threshold on the true parameter θ. Such a threshold is obviously not implementable

☆ This work was supported by the National Natural Science Foundation of China under grant NSFC 61304038 and the Project-sponsored by SRF for ROCS, SEM. 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 address: youky@tsinghua.edu.cn.

http://dx.doi.org/10.1016/j.automatica.2014.11.018
0005-1098/

as the true parameter is unknown. Even for this simplest case, it is not straightforward to find a feasible way to design optimal quantizer thresholds. Then, a central problem is how to optimally design quantizer thresholds? A large amount of work has been devoted to studying this problem, and we elaborate those mostly relevant works below.

Under fixed quantizer thresholds, many excellent works have been done by respectively using the empirical measure (Wang & Yin, 2007; Zhao, Guo, & Zhang, 2013), kernel function (Casini et al., 2012) and so on (Chen & Varshney, 2010; Li & Alregib, 2007; Shen et al., 2013). To obtain the asymptotic efficiency of the estimator in the sense of approaching the Cramer–Rao lower bound (CRLB), it requires the system input to be periodic and sufficiently rich (Wang & Yin, 2007; Wang, Yin, Zhang, & Zhao, 2010), which however makes the algorithm inconvenient for tracking control. In addition, most of their algorithms are derived by using the exact noise probability distribution function (pdf), which lack robustness to the error of the pdf. More importantly, they cannot achieve the minimum CRLB. Note that the CRLB is quantizer-dependent, and the minimum CRLB is the one that minimizes the CRLB among all quantizers satisfying the fixed-rate constraint. This certainly implies that the bit rate in Wang and Yin (2007) is not optimally used, and there exists more efficient algorithms for quantized estimation.

A natural extension of Xiao et al. (2006) is to dynamically adjust the quantizer thresholds. In Papadopoulos et al. (2002), the quantizer thresholds are periodically selected from a given set with an equal frequency, hoping that some thresholds are close to the true parameter. This method cannot arbitrarily tune the quantizer threshold and again fails to approach the minimum CRLB. In Fang and Li (2008), an adaptive scheme is proposed through a delta modulation where the modulation size is adjusted via solving an on-line optimization per iterate. Although this algorithm indeed asymptotically approaches the minimum CRLB, it lacks a recursive form, and is hard to implement. The problem has been resolved in Marelli et al. (2013) under the maximum likelihood criterion through the expectation maximization (EM) method. The estimator is jointly computed by an iterative weighted least squares algorithm and a quasi-Newton algorithm. Again, it needs the noise pdf to compute the E-step at each iterate. Another quantized estimator in Godoy et al. (2011) is given in a scenario-based form of the EM algorithm. However, the number of scenarios has a tremendous influence on the estimation accuracy, and the larger number of scenarios used, the higher computation demand needed.

Differently, this paper proposes a recursive estimator of stochastic approximation type to minimize the MSE under any quantization process. Then, we derive a necessary and sufficient condition on its asymptotic unbiasedness for estimating parameters of linear systems. If the estimated parameters were known, a simple quantizer can be easily designed to fulfill this condition. This points us a straight way to co-design the quantizer and estimator in a unified approach. Specifically, the unknown parameter vector is replaced with its latest estimator. This results in a quantized "innovations" scheme, and is consistent with the intuition that quantizing innovations requires fewer bits than directly quantizing outputs. Similar idea has also been exploited in Marelli et al. (2013), Ribeiro et al. (2006) and You, Xie, Sun, and Xiao (2008).

Given a symmetric quantizer, we rigorously prove that the co-designed quantizer and estimator are strongly consistent, asymptotically unbiased and normal. Note that the given quantizer does not require the exact noise pdf. In addition, the proposed estimator shares the same computational complexity as the gradient algorithm with un-quantized observations, and can be easily implemented. Although our algorithm has a similar structure as Guo and Zhao (2013), their algorithm requires the noise pdf to update the iterate, and is usually not asymptotically efficient.

To increase the convergence speed, we further propose two accelerated estimators by respectively using the Newton-based and the temporal averaging techniques, and show that both can converge at the fastest convergence speed of $O(1/k)$. As in You, Song, and Qiu (2014), the main purpose of averaging is to mitigate the "large jumps" of the iterate. Under Gaussian noise and any fixed-rate constraint, both algorithms are proved to asymptotically approach the minimum CRLB. The quantization effect on the estimation accuracy can also be exactly quantified. As a by-product, the result on the Newton-based algorithm is tailored to prove the asymptotic efficiency of the SOI-KF (Ribeiro et al., 2006).

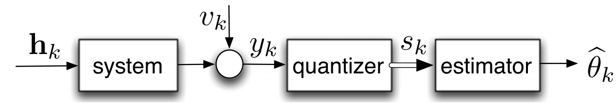

Fig. 1. Networked estimation.

In comparison, we demonstrate the advantages of our recursive algorithms from at least three folds. (1) As in the major gradient algorithm, the proposed estimator only requires the input signal, and does not need the noise pdf. (2) Their asymptotic properties hold for any sensible symmetric quantizer, and are robust to the noise pdf. (3) Under the Gaussian noise and any fixed-rate constraint, an optimal quantizer can be designed off-line for the accelerated estimators, which asymptotically approaches the minimum CRLB.

The rest of the paper is organized as follows. The problem formulation is delineated in Section 2, where we introduce the concept of symmetric quantization. In Section 3, the stochastic gradient is adopted to derive a recursive estimator. In Section 4, we co-design an adaptive quantizer and an estimator to obtain a recursive algorithm to identify the unknown parameters, and analyze its asymptotic property. In Section 5, two accelerated estimators using the weighted stochastic gradient and the averaging technique are given to achieve the fastest convergence speed. Under the Gaussian noise, their asymptotic efficiency and optimality are proved in Section 6. Simulation is performed in Section 7 to validate the theoretical results. Some concluding remarks are drawn in Section 8. To improve the readability, a technical proof is given in the Appendix.

2. Problem formulation

2.1. Quantized estimation

Consider a networked linear system as follows:

$$y_k = \mathbf{h}_k^T \boldsymbol{\theta} + v_k, \quad k = 1, 2, \ldots \tag{1}$$

where $\boldsymbol{\theta} \in \mathbb{R}^p$ is a vector of unknown parameters to be estimated, $\mathbf{h}_k \in \mathbb{R}^p$ is a sequence of known input regressors, and $v_k \in \mathbb{R}$ is a sequence of observation noises.

We are concerned with an estimation framework where the linear system and a remote estimator are connected via a digital channel. The system output y_k has to be reduced into finite precision s_k by a fixed-rate quantizer before transmitted to the estimator, see Fig. 1 for an illustration. The goal is to co-design the quantizer and estimator to identify the unknown vector $\boldsymbol{\theta}$.

To capture the essence of our key idea, we make the following assumption.

Assumption 1. The input regressors and observation noises satisfy that

(a) (Persistent excitation) $\{\mathbf{h}_k\}$ is a sequence of independent and identically distributed (i.i.d.) random vectors with a positive definite variance matrix H, i.e., $\mathbb{E}[\mathbf{h}_1\mathbf{h}_1^T] = H > 0$, and is independent of observation noise v_k. There exists some positive δ such that $\mathbb{E}[\|\mathbf{h}_1\|^{2+\delta}] < \infty$.
(b) (Symmetric noise) $\{v_k\}$ is a sequence of i.i.d. random variables with zero mean and $\mathbb{E}[v_1^2] = \sigma^2$. In addition, v_1 has an even probability density function (pdf) $p_v(\cdot)$, which is continuous at the origin.