Automatica 50 (2014) 169–174

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Robust distributed maximum likelihood estimation with dependent quantized data[☆]

Xiaojing Shen[a], Pramod K. Varshney[b,1], Yunmin Zhu[a]

[a] *Department of Mathematics, Sichuan University, Chengdu, Sichuan 610064, China*

[b] *Department of Electrical Engineering and Computer Science, Syracuse University, NY, 13244, USA*

ARTICLE INFO

Article history:
Received 4 March 2013
Received in revised form
8 June 2013
Accepted 18 September 2013
Available online 6 November 2013

Keywords:
Maximum likelihood estimation
Distributed estimation
Fisher information matrix
Wireless sensor networks

ABSTRACT

In this paper, we consider the distributed maximum likelihood estimation (MLE) with dependent quantized data under the assumption that the structure of the joint probability density function (pdf) is known, but it contains unknown deterministic parameters. The parameters may include different vector parameters corresponding to marginal pdfs and parameters that describe the dependence of observations across sensors. Since MLE with a single quantizer is sensitive to the choice of thresholds due to the uncertainty of pdf, we concentrate on MLE with multiple groups of quantizers (which can be determined by the use of prior information or some heuristic approaches) to fend off against the risk of a poor/outlier quantizer. The asymptotic efficiency of the MLE scheme with multiple quantizers is proved under some regularity conditions and the asymptotic variance is derived to be the inverse of a weighted linear combination of Fisher information matrices based on multiple different quantizers which can be used to show the robustness of our approach. As an illustrative example, we consider an estimation problem with a bivariate non-Gaussian pdf that has applications in distributed constant false alarm rate (CFAR) detection systems. Simulations show the robustness of the proposed MLE scheme especially when the number of quantized measurements is small.

Published by Elsevier Ltd

1. Introduction

Wireless sensor networks have attracted much attention with a lot of research taking place over the past several years. Many advances have been made in distributed detection, estimation, tracking and control (see e.g., Veeravalli & Varshney, 2012 and references therein). Distributed estimation and quantization problems have been considered in a number of previous studies. The parameters to be estimated are modeled as *random* and *deterministic* in different situations. We concentrate on *deterministic* parameters in this paper. For deterministic parameters, several universal distributed estimation schemes have been proposed (Xiao, Ribeiro, Luo, & Giannakis, 2006) in the presence of unknown, additive sensor noises that are bounded and identically distributed. The work in Fang and Li (2009) proposed the *vector* quantization design for distributed estimation under the assumption of an additive observation noise model.

System identification based on quantized measurements is a challenging problem even for very simple models and has been researched for a wide range of applications (see, e.g., Wang, Yin, Zhang, & Zhao, 2010). A method for the recursive identification of the nonlinear Wiener model was developed in Wigren (1995) and the corresponding convergence properties were analyzed. In Godoy, Goodwin, Aguero, Marelli, and Wigren (2011), Godoy et al. developed an MLE approach and used a scenario-based form of the expectation maximization algorithm to parameter estimation for general MIMO FIR linear systems with quantized outputs. The problem of set membership system identification with quantized measurements was considered in Casini, Garulli, and Vicino (2012). In Gustafsson and Karlsson (2009), the results from statistical quantization theory were surveyed and applied to both moment calculations and the likelihood function of the measured signal. The system identification of ARMA models using intermittent and quantized output observations was proposed in Marelli, You, and Fu (2013). The formal conditions for the asymptotic normality of the MLE to the reliability of a complex system based on a combination of full system and subsystem tests were proposed in Spall (2012).

In previous works, the MLE with quantized data is extensively used to estimate the deterministic parameters. In this paper, robust

☆ This work was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) under Grants FA9550-10-1-0263 and FA9550-10-1-0458 and in part by the NNSF of China 61273074 and IRT1273. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Er-Wei Bai under the direction of Editor Torsten Söderström.

E-mail addresses: shenxj@scu.edu.cn (X. Shen), varshney@syr.edu (P.K. Varshney), ymzhu@scu.edu.cn (Y. Zhu).

[1] Tel.: +1 315 443 1060; fax: +1 315 443 4745.

0005-1098/$ – see front matter. Published by Elsevier Ltd
http://dx.doi.org/10.1016/j.automatica.2013.09.036

distributed MLE with dependent quantized data is considered. Our work differs from previous studies in several aspects. Previous results concentrate on the problem of how to design the quantization schemes for estimating a deterministic parameter where each sensor makes *one* noisy observation. The observations are usually assumed *independent* across sensors, and they discuss the relationship between MLE performance and the number of sensors. Here, we focus on the problem of how to design estimation schemes for the unknown parameter *vector* associated with the joint pdf of the observations where the number of sensors is *fixed*. The emphasis here is on system robustness. These observations may be *dependent* across sensors. The unknown parameters may include different vector parameters corresponding to marginal pdfs and parameters that describe dependence of observations across sensors. Actually, the dependence between sensors is very important in multisensor fusion systems; for example, see the recent work on distributed location estimation with dependent sensor observations (Sundaresan & Varshney, 2011).

In this paper, we investigate the performance of MLE with multiple quantizers, since MLE with a single quantizer is sensitive to the choice of thresholds due to the uncertainty of pdf (see, e.g., Fang & Li, 2008). Our main contribution is that we analytically derive the *asymptotic efficiency* and *robustness* of a practical MLE with multiple quantizers in the context of *dependent* quantized measurements at the sensors, unknown parameter *vector* and without the knowledge of measurement models. The difficulties include the fact that due to dependence between measurements across sensors, the unknown high dimensional vector parameter estimation problem cannot be decoupled to scalar parameter estimation problems, and the quantized samples are not identically distributed due to the use of multiple different quantizers. Therefore, we have to deal with unknown *vector* parameter and unidentically distributed samples simultaneously. The asymptotic variance is derived to be the inverse of a weighted linear combination of Fisher information matrices based on J different quantizers which can be used to verify the robustness of our approach. A typical estimation problem with a bivariate non-Gaussian pdf with application to the distributed CFAR detection systems is considered. Simulations show that the new MLE scheme is robust and much better than that based on the worst quantization scheme from among the groups of quantizers. Moreover, when the number of quantized measurements is small, a surprising result is that the robust MLE has a significant and dominated advantage over the MLE with a single quantizer. It is also shown that the performance of the robust MLE is not the average performance of multiple quantizers. The rest of the paper is organized as follows. Problem formulation is given in Section 2. In Section 3, the robust MLE scheme is proposed and the asymptotic results are derived. In Section 4, numerical examples are given and discussed. In Section 5, conclusions are made.

2. Problem formulation

The basic L-sensor distributed estimation system is considered (see Fig. 1). Each sensor has k_i-dimensional observation population Y_i, $i = 1, \ldots, L$. Suppose that the joint observation population $Y \triangleq (Y_1', \ldots, Y_L')'$ has a given family of joint pdf:

$$\{p(y_1, \ldots, y_L|\theta)\}_{\theta\in\Theta\subseteq\mathbb{R}^k} \tag{1}$$

where $'$ denotes the transpose and θ is the unknown k-dimensional deterministic parameter vector which may include marginal parameters and dependence parameters. Here, we do not assume independence across sensors, knowledge of measurement models and Gaussianity of the joint pdf. Let N be independently and identically distributed (i.i.d.) sensor observation samples and joint observation samples be

$$\vec{Y}_i = (Y_{i1}, \ldots, Y_{iN}), \quad i = 1, \ldots, L; \tag{2}$$

$$\vec{Y} = (\vec{Y}_1', \ldots, \vec{Y}_L')'. \tag{3}$$

Suppose the sensors and the fusion center wish to jointly estimate the unknown parameter vector θ based on the spatially distributed observations. If there is sufficient communication bandwidth and power, the fusion center can obtain asymptotically efficient estimates with the complete observation samples based on the MLE procedure under some regularity conditions on the joint pdf.

In many practical situations, however, to reduce the communication requirement from sensors to the fusion center due to limited communication bandwidth and power, the i-th sensor quantizes the observation vector to 1 bit (it is straightforward to extend to multiple bits) by a measurable indicator quantization function:

$$I_i(y_i): \ y_i \in \mathbb{R}^{k_i} \to \{0, 1\}, \tag{4}$$

for $i = 1, \ldots, L$. Here, the quantization region of each quantizer $I_i(y_i)$ may be continuous or union of discontinuous regions. Moreover, we denote by

$$I(y) \triangleq (I_1(y_1), \ldots, I_L(y_L))' \in \mathbb{R}^L. \tag{5}$$

Once the binary quantized samples $I_i(Y_{in})$ are generated at sensor i, $i = 1, \ldots, L$, they are transmitted to the fusion center, for $n = 1, \ldots, N$. The fusion center is then required to estimate the true parameter vector θ^* based on the received quantized data. By the definition of observation samples and quantizers, we define

$$\vec{U} \triangleq (\vec{U}_1', \ldots, \vec{U}_N')', \tag{6}$$

$$\vec{U}_n \triangleq (U_{1n}, \ldots, U_{Ln})', \quad n = 1, \ldots, N, \tag{7}$$

$$U_{in} \triangleq I_i(Y_{in}), \quad n = 1, \ldots, N. \tag{8}$$

If we take $\vec{U}_n$ as the joint quantized observation sample and denote the quantized observation population by $U \triangleq I(Y) = (I_1(Y_1), \ldots, I_L(Y_L))'$, we know that U has a *discrete/categorical* distribution. Based on the pdf of Y and quantizers $I(y)$, the probability mass function (pmf) of the quantized observation population U is

$$f_U(u_1, u_2, \ldots, u_L|\theta) = \int_{\Xi_{(u_1,u_2,\ldots,u_L)}} p(y_1, y_2, \ldots, y_L|\theta)dy_1dy_2\cdots dy_L, \tag{9}$$

where

$$(u_1, u_2, \ldots, u_L) \in \mathcal{S}_u = \{(u_1, u_2, \ldots, u_L) \in \mathbb{R}^L: \ u_i = 0/1, \ i = 1, \ldots, L\}, \tag{10}$$

$$\Xi_{(u_1,u_2,\ldots,u_L)} = \{(y_1, y_2, \ldots, y_L): \ I_1(y_1) = u_1, I_2(y_2) = u_2, \ldots, I_L(y_L) = u_L\}. \tag{11}$$

Thus, the quantized observation population U has a family of joint pmf $\{f_U(u_1, u_2, \ldots, u_L|\theta)\}_{\theta\in\Theta\subseteq\mathbb{R}^k}$ which yields the following log likelihood function of samples $\vec{U}$ by (6)–(11):

$$l(\theta|\vec{U}) \triangleq \log\prod_{n=1}^{N} f_U(U_{1n}, U_{2n}, \ldots, U_{Ln}|\theta) \tag{12}$$

$$= \sum_{n=1}^{N} \log f_U(U_{1n}, U_{2n}, \ldots, U_{Ln}|\theta) \tag{13}$$

$$= \sum_{j=1}^{2^L} n_j \log f_U(\vec{u}_j|\theta) \tag{14}$$

where $n_j = \#\{(U_{1n}, U_{2n}, \ldots, U_{Ln}) = \vec{u}_j \in S_u, n = 1, \ldots, N\}$, $\sum_{j=1}^{2^L} n_j = N$; $\#\{\cdot\}$ is the cardinality of the set. The parameter vector θ is estimated by maximizing the log likelihood function (14). Let $\hat{\theta}$ denote the MLE of θ.

Based on the classical asymptotic properties of MLE (see, e.g., textbooks Casella & Berger, 2001; Van der Vaart, 2000), we have the following lemma.

Lemma 1. *Assume that* $p(y_1, y_2, \ldots, y_L|\theta)$ *and sensor quantizers* $I_1(y_1), \ldots, I_L(y_L)$ *generate the quantized samples and* $f_U(u_1, u_2, \ldots,$