Automatica 97 (2018) 82–91

Contents lists available at ScienceDirect

Automatica

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

Accuracy analysis for distributed weighted least-squares estimation in finite steps and loopy networks ☆

Tianju Sui [a], Damián Edgardo Marelli [b,c,*], Minyue Fu [d,b], Renquan Lu [b]

[a] Department of Control Science and Engineering, Dalian University of Technology, Liaoning, China
[b] School of Automation, Guandong University of Technology, Guangzhou, China
[c] French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Ocampo y Esmeralda, Rosario 2000, Argentina
[d] School of Electrical Engineering and Computer Science, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia

ARTICLE INFO

Article history:
Received 9 September 2016
Received in revised form 26 December 2017
Accepted 25 June 2018
Available online 17 August 2018

Keywords:
Distributed statistical estimation
Weighted least squares
Convergence rate

ABSTRACT

Distributed parameter estimation for large-scale systems is an active research problem. The goal is to derive a distributed algorithm in which each agent obtains a local estimate of its own subset of the global parameter vector, based on local measurements as well as information received from its neighbors. A recent algorithm has been proposed, which yields the optimal solution (i.e., the one that would be obtained using a centralized method) in finite time, provided the communication network forms an acyclic graph. If instead, the graph is cyclic, the only available alternative algorithm, which is based on iterative matrix inversion, achieving the optimal solution, does so asymptotically. However, it is also known that, in the cyclic case, the algorithm designed for acyclic graphs produces a solution which, although non optimal, is highly accurate. In this paper we do a theoretical study of the accuracy of this algorithm, in communication networks forming cyclic graphs. To this end, we provide bounds for the sub-optimality of the estimation error and the estimation error covariance, for a class of systems whose topological sparsity and signal-to-noise ratio satisfy certain condition. Our results show that, at each node, the accuracy improves exponentially with the so-called loop-free depth. Also, although the algorithm no longer converges in finite time in the case of cyclic graphs, simulation results show that the convergence is significantly faster than that of methods based on iterative matrix inversion. Our results suggest that, depending on the loop-free depth, the studied algorithm may be the preferred option even in applications with cyclic communication graphs.

1. Introduction

With the fast development of sensor networks and wireless communications, the scale of systems is becoming increasingly large. Since centralized estimation requires a fusion center to process all the information from the whole graph, the computation and communication burden increases with the system's size. Thus, the centralized estimation approach is not suitable for large-scale systems, and distributed approaches are needed. The development of distributed estimation has attracted a great deal of attention (Garin & Schenato, 2010; Gupta, Dana, Hespanha, Murray, & Hassibi, 2009; Li & Alregib, 2009; Ribeiro & Giannakis, 2006a, b). It finds applications in industrial monitoring, multi-agent systems, the smart grid, etc.

The distributed estimation problem consists of a network of interconnected nodes, each of which aims to obtain an estimate of certain vector of interest. This is achieved through an iterative procedure in which each node processes its available information, and exchange relevant information with its neighbors, in order to successively compute the required estimate as accurately as possible. The existing distributed estimation problems can be broadly classified into four classes. These classes are: static fully reconstructive, static partially reconstructive, dynamic fully reconstructive and dynamic partially reconstructive. A fully reconstructive system is one in which each node aims to obtain an estimate of the same vector. In contrast, in a partially reconstructive system, each node aims to obtain an estimate of its own partial sub-vector of interest. Also, a static system is one in which prior knowledge of the state at

☆ This work was supported by the China Postdoctoral Science Foundation, China (2017M621134). The material in this paper was presented at the 2015 IEEE 54th Annual Conference on Decision and Control (CDC), December 15–18, 2015. Osaka, Japan. This paper was recommended for publication in revised form by Associate Editor Shreyas Sundaram under the direction of Editor Christos G. Cassandras.

* Corresponding author at: School of Automation, Guandong University of Technology, Guangzhou, China.

E-mail addresses: suitj@mail.dlut.edu.cn (T. Sui), Damian.Marelli@newcastle.edu.au (D.E. Marelli), Minyue.Fu@newcastle.edu.au (M. Fu), rqlu@gdut.edu.cn (R. Lu).

https://doi.org/10.1016/j.automatica.2018.07.016
0005-1098/

a certain time is independent of the knowledge of the same state at previous times. A dynamic system refers to the complementary case. We point out that methods for dynamic estimation can be readily used for static problems, by choosing the dynamic model in a way such that the state stays constant over time.

In the static fully reconstructive problem, the most popular distributed estimation algorithm is consensus (Garin & Schenato, 2010). By running average consensus on the information vector and information matrix of each node, in view of the weighted least squares (WLS) formula, the final estimate of each node converges to the one obtained via WLS (Olfati-Saber, 2005). Although the average consensus algorithm is simple, it has two main disadvantages: First, the communication burden is large, as each node communicates $\frac{n\times(n+3)}{2}$ scalars to its neighbors, where n is the dimension of the estimated vector. Second, the convergence of average consensus requires infinite iterations, and the stopping criterion is still an open problem. To avoid these two disadvantages, many algorithms have been proposed (Ajgl & Šimandl, 2014; Calafiore & Abrate, 2009; Chen, Arambel, & Mehra, 2002; Pasqualetti, Carli, & Bullo, 2012). One of the most important works is the one in Pasqualetti et al. (2012), where using the space structure of measurements and doing kernel projection, each node achieves its minimum norm solution in a finite number of steps.

In the static partially reconstructive problem, since each node considers its own partial state, the consensus algorithm is not applicable. For the case in which the graph induced by the communication network is acyclic (i.e., without loops), an algorithm is proposed in Tai, Lin, Fu, and Sun (2013). In this algorithm, each node obtains a WLS estimate on its own state in a finite number of steps. When the graph is cyclic (i.e., with loops), Marelli and Fu (2015) gave a novel method which, based on Richardson iterations, solves the WLS estimation problem. However, it does so asymptotically, i.e., in infinite iterations. We point out that most estimation algorithms for large-scale systems are partially reconstructive, since the whole state of the system is often of very high dimension.

In the dynamic fully reconstructive problem, the consensus algorithm is also a popular option. In Matei and Baras (2012), one consensus algorithm is run at each sampling time, using the partial estimates obtained at each node, based on their local measurements. Building on this line, a study on the number of consensus iterations required at each sampling time to guarantee the stability of the estimator, under the observability condition, is done in Acikmese, Mandić, and Speyer (2014). Also, the so-called diffusion Kalman filter (Cattivelli & Sayed, 2010) runs consensus on the estimates obtained at each sensor, using local measurements as well as those from neighbors. As opposite to doing consensus on the estimates, the authors of Battistelli and Chisci (2014) found that, by running consensus on the information matrices and vectors, observability is sufficient for the estimation stability.

Concerning the dynamic partially reconstructive problem, information passing and processing methods guaranteeing a stable estimate are proposed in Farina, Ferrari-Trecate, and Scattolini (2010), Khan and Moura (2008), Zhou (2013) and Zhou (2015). Also, the authors of Haber and Verhaegen (2013) study systems with banded dynamic state transition matrices, concluding that the contribution from faraway nodes decreases with the increase of their distance. The authors also propose the moving horizon estimation approach as an approximation to the optimal state estimate.

In this paper we focus on the static partially reconstructive problem. Also, as typically done in static problems, we assume that the vector to be estimated is deterministic. More precisely, we consider the algorithm in Tai et al. (2013), which, as mentioned, yields the optimal solution in finite-time, only when the communication graph is acyclic. For cyclic graphs, this algorithm is not guaranteed to produce the optimal solution. Nevertheless, in many applications, even in the presence of loops, it delivers very good approximations to the optimal solution, in only a very few steps. For those applications, this makes the algorithm a valid alternative to the method in Marelli and Fu (2015) even for cyclic networks. This is because, while the later guarantees the optimal solution, the former one converges much faster. Motivated by this, we study the accuracy of the estimate produced by the algorithm in Tai et al. (2013), under the general setting of a cyclic graph.

For a class of systems whose topological sparsity and signal-to-noise ratio satisfy certain condition, we are able to determine the accuracy of the estimates and their associated estimation error covariances, with respect to those achievable via a centralized WLS method. Our formulas clearly show how accuracy depends on the so-called loop-free depth of each node. More precisely, the estimates and estimation error covariances approach those from the centralized solution, exponentially on the loop-free depth.

The rest of this paper is organized as follows. In Section 2, we give the problem formulation and introduce the distributed WLS algorithm under study. In Section 3, we show how to convert a given graph into other equivalent ones, which are instrumental for analyzing the behavior of the algorithm in cyclic graphs. In Section 4, we introduce our notation, as well as the definition of the *Riemannian Distance* between matrices, together with some of its properties. The accuracy of the information matrices (i.e., the inverses of the error covariances) and state estimates produced by the distributed WLS algorithm are analyzed in Sections 5 and 6, respectively. In Section 7, we provide some simulations to illustrate our results. Finally, concluding remarks are stated in Section 8. Due to space constraints, some complementary mathematical material, including most proofs and some additional lemmas, appears in the extended version (Sui, Marelli, Fu, & Lu, 2018), which is available online.

2. Problem formulation

Consider a system observed by I sensing nodes. Associated to this system, there is a deterministic vector $x^T = \left[x_1^T, x_2^T, \ldots, x_I^T\right] \in \mathbb{R}^n$, with $\sum_{i=1}^{I} n_i = n$, called the global state. For any $i = 1, \ldots, I$, node i aims to estimate the sub-vector $x_i \in \mathbb{R}^{n_i}$. There are also two kinds of measurements. The so-called *self measurements* for node i

$$z_i = C_i x_i + v_i, \tag{1}$$

and the (pair-wise) *joint measurements* between nodes i and j

$$z_{i,j} = C_{i,j} x_i + C_{j,i} x_j + v_{i,j}. \tag{2}$$

In the above, the matrices C_i, $C_{i,j}$ and $C_{j,i}$ are known, and v_i and $v_{i,j}$ are independent measurement noises with known covariances $R_i > 0$ and $R_{i,j} > 0$, respectively. Note that (1) the pair (i,j) is unordered, i.e., $(i,j) = (j,i)$; (2) $z_{i,j} = z_{j,i}$ and $v_{i,j} = v_{j,i}$; (3) It is not necessary for all nodes to have self measurements or all node pairs to have joint measurements. In fact, joint measurements are typically sparse for large graphs.

We assume that node i and node j could communicate if $z_{i,j}$ exists. Furthermore, we call node j a neighbor of node i (i.e., $j \in \mathcal{N}_i$) and node i a neighbor of node j (i.e., $i \in \mathcal{N}_j$) if there is communication between them. In view of this, communication between nodes is always two-ways; and therefore, the associated communication graph (which will be formally introduced later) is always undirected.

The target of distributed WLS estimation is to compute the WLS estimate for each x_i, and its associated estimation error covariance, using a fully distributed algorithm. The algorithm summarized in Algorithm 1, achieves this goal. In this algorithm, at iteration N, node i computes a local estimate $\hat{x}_i(N)$ of its sub-vector of interest,