Systems & Control Letters 112 (2018) 9–17

Contents lists available at ScienceDirect

Systems & Control Letters

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

Convergence rates of discrete-time stochastic approximation consensus algorithms: Graph-related limit bounds

Huaibin Tang [a,b], Tao Li [c,*]

[a] School of Microelectronics, Shandong University, Jinan 250100, China
[b] Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[c] Shanghai Key Laboratory of Pure Mathematics and Mathematical Practice, Department of Mathematics, East China Normal University, Shanghai 200241, China

ARTICLE INFO

Article history:
Received 9 August 2016
Received in revised form 29 November 2017
Accepted 5 December 2017

Keywords:
Consensus
Sensor network
Martingale difference sequence
Stochastic approximation
Convergence rate

ABSTRACT

In this paper, we study the convergence rates of the discrete-time stochastic approximation consensus algorithms over sensor networks with communication noises under general digraphs. Basic results of stochastic analysis and algebraic graph theory are used to investigate the dynamics of the consensus error, and the mean square and sample path convergence rates of the consensus error are both given in terms of the graph and noise parameters. Especially, calculation methods to estimate the mean square limit bounds are presented under balanced digraphs, and sufficient conditions on the network topology and the step sizes are given to achieve the fast convergence rate. For the sample path limit bounds, estimation methods are also presented under undirected graphs.

1. Introduction

Recently, consensus algorithms with stochastic disturbances in sensor networks have been widely investigated, including measurement noises, time delay, quantized data and random link failures [1–7]. For consensus problems with communication noises, the stochastic approximation (SA) consensus algorithm with decreasing step sizes is an effective method to attenuate the influence of noises. Applications of SA consensus algorithms and theoretical development in consensus problems were reported in [8–12], etc. Under fixed or time-varying topologies, the condition that the network contains or jointly contains a spanning tree to guarantee consensus is well understood. This problem has been systematically investigated by the tools of the SA theory [12,13], the quadratic Lyapunov functions [8,9], the algebraic theory [5], and the ergodicity backward product approach [10,11]. Different from the SA-type consensus algorithm with decreasing step size, Amelina et al. [14] proposed the consensus algorithm with a nonvanishing stepsize for nonlinear agent dynamics over noisy networks to achieve the approximate mean square consensus.

It is worth noting that the convergence rate of the consensus algorithm, which characterizes how fast consensus can be achieved, is an important issue from the perspective of practical applications. For the case with precise communication, the consensus error vanishes exponentially with the rate governed by the second smallest eigenvalue of the Laplacian matrix [15,16]. The problems how to characterize and optimize the convergence rate are extensively studied via the optimization of the weighted adjacency matrix [17], local node state prediction [18], and filtering techniques [19]. Recently, Olshevsky and Tsitsiklis [20] investigated the convergence time of consensus algorithm under time-varying undirected graphs and proposed a linear time average-consensus protocol under fixed undirected graphs [21]. For the SA consensus algorithm, the convergence rate problem has also attracted much attention. For the average-consensus problem under undirected graphs, Kar and Moura [22] showed that the mathematical expectation of the state vector sequence converges exponentially to the consensus value, and Dasarathan et al. [12] derived the asymptotic covariance matrix of the consensus error when the step size $a(t) = \Theta(t^{-1})$. For the case with balanced digraphs, Li and Zhang [8] obtained the sample path convergence rate of finite step mean consensus error. For the leader-following topology case, Xu et al. [13] showed that the sample path convergence rate of the consensus error is $o(a^{\delta_1}(t))$ if the step size $a(t)$ satisfies $\lim_{t\to\infty}(a(t)-a(t+1))/(a(t)a(t+1)) \geq 0$, and the mean square convergence rate of the consensus error is $o(a^{\delta_2}(t))$ if $a(t) = \Theta(t^{-\alpha})$ with $\alpha \in (0.5, 1]$, $\delta_1, \delta_2 \in (0, 1)$. Wang et al. [23,24] investigated the convergence rate in the sense of convergence in distribution for multi-scale consensus modeling

* Corresponding author.
E-mail addresses: tanghuaibin@sdu.edu.cn (H. Tang), tli@math.ecnu.edu.cn (T. Li).

https://doi.org/10.1016/j.sysconle.2017.12.002
0167-6911/

with Markovian regime switching. Compared with the case with precise communication, the convergence of SA algorithm is not exponentially fast any more and is in more complex relation to the step size $a(t)$ and network parameters. This motivates us to evaluate the impacts of the step size and network parameters on the algorithm, which is useful for the designers to improve the convergence rate.

In this paper, we consider the discrete-time SA consensus algorithm under general digraphs corrupted by martingale difference sequence communication noises. Different from [9–11] which focused on the consensus conditions of the SA consensus algorithm, the main goal of this paper is focused on the convergence rate analysis in relation to the step size and network graph parameters. For a class of typical step sizes, we apply basic results of stochastic analysis and algebraic graph theory to investigate the consensus error dynamic equation. Compared with the continuous-time SA consensus algorithm in [25], there is no Itô formula as the effective tool and we develop more technical tools of inequality theory to handle with the closed-form of the consensus error. Our contribution mainly includes the following three aspects.

- For the case of fixed topologies, we show that if the step size $a(t) = \Theta(t^{-\gamma}), \gamma \in (0.5, 1)$, then the mean square convergence rate of the consensus error is $O(t^{-\gamma})$; especially, for the case with balanced digraphs, the convergence rate is exactly $\Theta(t^{-\gamma})$. Furthermore, both upper and lower limit bounds of $t^{\gamma}E(\|\delta(t)\|^2)$ are explicitly given in terms of the noise intensity, the number of nodes, the smallest and the largest nonzero eigenvalues of the Laplacian matrix of the symmetrized graph.
- If $a(t) = \Theta(t^{-1})$, intuitively, the mean square convergence rate of the consensus error might be $O(t^{-1})$ and higher than the case with $a(t) = \Theta(t^{-\gamma}), \gamma \in (0.5, 1)$. Interestingly, we found that this is not always true, and the mean square convergence rate is $O(t^{-1})$ only if the Laplacian eigenvalues of the network topology graph satisfy certain conditions. It is observed that the fast convergence rate $O(t^{-1})$ depends on the step size $a(t)$ and eigenvalues of the Laplacian matrix. Especially, for the case with balanced graphs, choosing $a(t) = \Theta(t^{-1})$ with $\lambda_2(\widehat{L_{\mathcal{G}}})\liminf_{t\to\infty}(ta(t)) \gtrsim 1$ will achieve the convergence rate $O(t^{-1})$, where $\lambda_2(\widehat{L_{\mathcal{G}}})$ is the algebraic connectivity of the symmetrized graph. For the case with undirected graphs, the condition on the step size $a(t)$ can be relaxed to $\lambda_2(L_{\mathcal{G}})\liminf_{t\to\infty}(ta(t)) > 1/2$, where $\lambda_2(L_{\mathcal{G}})$ is the algebraic connectivity of the graph.
- We study the sample path behavior of the consensus error under undirected graphs. It is observed that the consensus error has a convergence rate slightly slower than $O(t^{-\gamma/2})$ almost surely. The upper limit bound of the sample path of the consensus error is calculated.

Compared with the existing related works [12,8,13,22,26,24], we systematically analyze the stochastic convergence rates of the distributed SA consensus algorithm in the sense that both the network topology and the class of step size are more general. Besides, the explicit limit bounds of the stochastic convergence rates are provided, which clearly show the impacts of various kinds of system parameters on the convergence rates, i.e. the number of nodes, the variance of noises, the maximal weight, the eigenvalues of the Laplacian matrix, etc. Also, sufficient conditions are given to achieve fast convergence rate $O(t^{-1})$. These will be all helpful for developing efficient and practical distributed algorithms over large scale sensor networks by designing the step sizes and network parameters.

This paper is organized as follows. In Section 2, we formulate the problem to be investigated. In Section 3, we investigate the dynamic consensus error equation and give the mean square and sample path convergence rates for the SA consensus algorithm. Numerical simulations to corroborate our analytical findings are presented in Section 4, and concluding remarks are given in Section 5. For the sake of conciseness, all the proofs are put in Appendix.

In this paper, we adopt the following notations. $\mathbf{1}_{N\times 1}$ and $\mathbf{0}_{N\times 1}$ denote $N \times 1$ column vectors with all ones and all zeros, respectively. For a given vector or matrix A, A^T denotes its transpose, and $\|A\|$ denotes its 2-norm. For any given complex number λ, $Re(\lambda)$ denotes its real part and $Im(\lambda)$ denotes its imaginary part. We denote $f(t) = o(g(t))$ if $\lim_{t\to\infty}|f(t)/g(t)| = 0$; $f(t) = O(g(t))$ if $\limsup_{t\to\infty}|f(t)/g(t)| < \infty$; $f(t) = \Omega(g(t))$ if $\liminf_{t\to\infty}|f(t)/g(t)| > 0$; and $f(t) = \Theta(g(t))$ if both $f(t) = O(g(t))$ and $f(t) = \Omega(g(t))$. For a differentiable function $f(t)$, $f^{(k)}(t)$ denotes its kth derivative and $f^{(0)}(t) = f(t)$.

2. Problem formulation

For a weighted digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$, $\mathcal{V} = \{1, \dots, N\}$ denotes the set of N nodes, $\mathcal{E}$ denotes the set of edges, and $\mathcal{A} = [a_{ij}] \in \mathbf{R}^{N\times N}$ denotes the weighted adjacency matrix. The pair $(j, i) \in \mathcal{E} \Leftrightarrow$ node j can send information to node i directly. Then j is called the parent of i. Node i is called a source if it has no parent. The neighborhood of the ith node is denoted by $\mathcal{N}_i = \{j \in \mathcal{V}|(j, i) \in \mathcal{E}\}$. For any given $i, j \in \mathcal{V}$, $a_{ij} \geq 0$, and $a_{ij} > 0$ if and only if $j \in \mathcal{N}_i$. $L_{\mathcal{G}} = \mathcal{D}-\mathcal{A}$ is called the Laplacian matrix of $\mathcal{G}$, where $\mathcal{D} = diag(\sum_{j=1}^N a_{1j}, \dots, \sum_{j=1}^N a_{Nj})$. The digraph $\mathcal{G}$ is balanced, if $\sum_{j=1}^N a_{ji} = \sum_{j=1}^N a_{ij}$ for all $i \in \mathcal{V}$. A directed tree is a digraph, where every node except the root has exactly one parent and the root is a source. A spanning tree of $\mathcal{G}$ is a directed tree whose node set is $\mathcal{V}$ and whose edge set is a subset of $\mathcal{E}$. If the digraph $\mathcal{G}$ contains a spanning tree, then $L_{\mathcal{G}}$ has a unique zero eigenvalue and all other $N-1$ eigenvalues have positive real parts. We denote $\lambda_1 = 0$ and all its distinct non-zero eigenvalues by $\lambda_2, \dots, \lambda_l$. We denote $\lambda_2^* = \min\{Re(\lambda_m), 2 \leq m \leq l\}$. It is known that there exists a unique probability measure π^T which is the left eigenvector of $L_{\mathcal{G}}$ associated with λ_1, i.e., $\pi^T L_{\mathcal{G}} = \mathbf{0}_{N\times 1}$. If the digraph $\mathcal{G}$ is balanced, then $\pi^T = (1/N)\mathbf{1}_{N\times 1}$.

Consider the discrete-time SA consensus algorithm for a N nodes network

$$x_i(t+1) = x_i(t) + a(t)\sum_{j\in\mathcal{N}_i} a_{ij}(y_{ji}(t) - x_i(t)),\ t \geq 0,\ i \in \mathcal{V}, \quad (1)$$

here the step size $a(t) > 0$; $x_i(t) \in \mathbf{R}$ is the ith node's state, and the initial state $x_i(0)$ is deterministic; $y_{ji}(t)$ is the received information of the ith node from the jth node:

$$y_{ji}(t) = x_j(t) + \omega_{ji}(t),\ j \in \mathcal{N}_i, \quad (2)$$

where $\{\omega_{ji}(t), t \geq 0, i, j \in \mathcal{V}\}$ are the communication noises.

Denote $X(t) = [x_1(t), \dots, x_N(t)]^T$. Eq. (1) can be rewritten as follows:

$$X(t+1) = (I_N - a(t)L_{\mathcal{G}})X(t) + a(t)\Sigma_{\mathcal{G}}W(t). \quad (3)$$

Here, $W(t) = [w_1^T(t), \dots, w_N^T(t)]^T$, $w_i(t) = [\omega_{1i}(t), \dots, \omega_{Ni}(t)]^T$ and $\Sigma_{\mathcal{G}} = diag(\alpha_1^T, \dots, \alpha_N^T)$ is an $N \times N^2$ dimensional block diagonal matrix with α_i^T being the ith row of the weighted adjacency matrix $\mathcal{A}$.

It was proved in [9] that if the digraph $\mathcal{G}$ contains a spanning tree and $a(t)$ satisfies the standard conditions $\sum_{t=0}^{\infty} a(t) = \infty$, $\sum_{t=0}^{\infty} a^2(t) < \infty$, then the SA consensus algorithm (1)–(2) can achieve both mean square and almost sure consensus, i.e., $E|x_i(t)|^2 < \infty$, and there exists a random variable x^* such that $\lim_{t\to\infty} E|x_i(t) - x^*|^2 = 0$ and $\lim_{t\to\infty} x_i(t) = x^*$ a.s., for all $i \in \mathcal{V}$. Hereinafter, to measure the disagreement among the nodes, we denote $J = \mathbf{1}_{N\times 1}\pi^T$ and the dynamic consensus error by $\delta(t) = (I_N - J)X(t)$.

In this paper, we will study the stochastic convergence rate of $\delta(t)$ and our main goal includes two aspects: