Systems & Control Letters 96 (2016) 110–117

Contents lists available at ScienceDirect

Systems & Control Letters

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

Primal–dual algorithm for distributed constrained optimization ☆

Jinlong Lei *, Han-Fu Chen, Hai-Tao Fang

The Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China

ARTICLE INFO

Article history:
Received 10 October 2015
Received in revised form
14 July 2016
Accepted 28 July 2016

Keywords:
Distributed constrained optimization
Primal–dual algorithm
Augmented Lagrange method
Multi-agent network

ABSTRACT

The paper studies a distributed constrained optimization problem, where multiple agents connected in a network collectively minimize the sum of individual objective functions subject to a global constraint being an intersection of the local constraints assigned to the agents. Based on the augmented Lagrange method, a distributed primal–dual algorithm with a projection operation included is proposed to solve the problem. It is shown that with appropriately chosen constant step size, the local estimates derived at all agents asymptotically reach a consensus at an optimal solution. In addition, the value of the cost function at the time-averaged estimate converges with rate $O(\frac{1}{k})$ to the optimal value for the unconstrained problem. By these properties, the proposed primal–dual algorithm is distinguished from the existing algorithms for distributed constrained optimization. The theoretical analysis is justified by numerical simulations.

1. Introduction

Distributed computation and estimation recently have received much research attention, e.g., consensus problems [1,2], distributed estimation [3], sensor localization [4], and distributed control [5,6]. In particular, distributed optimization problems have been extensively investigated in [7–19], among which the distributed subgradient or gradient algorithms [7–11] belong to the primal domain methods while [12–19] belong to the primal–dual domain methods.

The paper considers a distributed constrained optimization problem, where n agents connected in a network collectively minimize the sum of local objective functions $f(x) = \sum_{i=1}^{n} f_i(x)$ subject to a global constraint $\Omega_o = \bigcap_{i=1}^{n} \Omega_i$, where Ω_i is a convex set and $f_i(x)$ is a convex function in Ω_i. Besides, $f_i(x)$ and Ω_i are the local data known to agent i and cannot be shared with other agents. The problem is equivalent to a convex optimization problem with constraint being a linear equation constrained to a convex set.

The main contribution of the paper is to propose a distributed primal–dual algorithm with constant step size to solve the constrained optimization problem over the multi-agent network. The algorithm is derived on the basis of the gradient algorithm for finding saddle points of an augmented Lagrange function [20]. In an iteration, each agent updates its estimate only using the local data and the information derived from the neighboring agents. With appropriately chosen constant step size, the estimates derived at all agents are shown to reach a consensus at an optimal solution. Besides, it is found that the value of the cost function at the time-averaged estimate converges with rate $O(\frac{1}{k})$ to the optimal value for the unconstrained problem.

A general constrained convex optimization problem is studied in [12], where the constraint sets are assumed to be compact. The problem in the random case is investigated by [10] for non-smooth objective functions, meanwhile, the convex sets are assumed to be compact and the global constraint set is required to have a nonempty interior. Here, we study the problem in the deterministic case for smooth objective functions, while imposing weaker assumptions on the convex sets.

When there are no constraints, the problem of the paper becomes the one discussed in [7,11,15–17]. The estimates produced by the distributed gradient descent (DGD) algorithm with constant step size [7] converge to a neighborhood of the optimal solution, while our algorithm gives the accurate estimate. The estimates generated by the fast distributed gradient algorithms [11] and by EXTRA [15] converge to an optimal solution, but in [11] each cost function is assumed to be convex with gradients being bounded and Lipschitz continuous. Though it is shown by [21] that EXTRA [15] is also a saddle point method, the augmented Lagrange function used in [15] is different from ours. The primal–dual

☆ This work was supported by the NSFC under Grants 61273193, 61120106011, 61134013, 61573345, the 973 program of China under grant No. 2014CB845301, and the National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences under grant No. Y629091ZZ2.

* Corresponding author.
E-mail addresses: leijinlong11@mails.ucas.ac.cn (J. Lei), hfchen@iss.ac.cn (H.-F. Chen), htfang@iss.ac.cn (H.-T. Fang).

http://dx.doi.org/10.1016/j.sysconle.2016.07.009
0167-6911/

algorithm proposed in the paper can be seen as an extension of EXTRA [15] to constrained problems. Some continuous and discrete time algorithms are designed in [16,17] to solve the distributed unconstrained optimization problems from control perspective. Unlike [16], both unconstrained and constrained problems are discussed in the paper. For the unconstrained problem, the algorithm proposed in the paper is quite similar to that given in [16, equation (7)], but the convergence established in [16] is only for the case where the subgradients are bounded, while in this paper there is no such a restriction. Besides, the rate of convergence is also established here.

The rest of the paper is organized as follows: In Section 2, some preliminary information about graph theory and convex analysis is provided and the problem is formulated. In Section 3, a distributed primal–dual algorithm is proposed for solving the problem, while its convergence is proved in Section 4. Two numerical examples are demonstrated in Section 5, and some concluding remarks are given in Section 6.

2. Preliminaries and problem statement

We first provide some information about graph theory, convex functions, and convex sets. Then, we formulate the distributed constrained optimization problem to be investigated.

2.1. Graph theory

Consider a network of n agents. The communication relationship among the n agents is described by an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}}, \mathcal{A}_{\mathcal{G}}\}$, where $\mathcal{V} = \{1, \dots, n\}$ is the node set with node i representing agent i; $\mathcal{E}_{\mathcal{G}} \subset \mathcal{V} \times \mathcal{V}$ is the undirected edge set, and the unordered pair of nodes $(i, j) \in \mathcal{E}_{\mathcal{G}}$ if and only if agent i and agent j can exchange information with each other; $\mathcal{A}_{\mathcal{G}} = [a_{ij}] \in \mathbb{R}^{n\times n}$ is the adjacency matrix of $\mathcal{G}$, where $a_{ij} = a_{ji} > 0$ if $(i, j) \in \mathcal{E}_{\mathcal{G}}$, and $a_{ij} = 0$, otherwise. Denote by $\mathcal{N}_i = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}_{\mathcal{G}}\}$ the neighboring agents of agent i. The Laplacian matrix of graph $\mathcal{G}$ is defined as $\mathcal{L}_{\mathcal{G}} = \mathcal{D}_{\mathcal{G}} - \mathcal{A}_{\mathcal{G}}$, where $\mathcal{D}_{\mathcal{G}} = \mathrm{diag}\{\sum_{j=1}^{n} a_{1j}, \dots, \sum_{j=1}^{n} a_{nj}\}$. For a given pair $i, j \in \mathcal{V}$, if there exists a sequence of distinct nodes $i_1, \dots, i_p$ such that $(i, i_1) \in \mathcal{E}_{\mathcal{G}}, (i_1, i_2) \in \mathcal{E}_{\mathcal{G}}, \dots, (i_p, j) \in \mathcal{E}_{\mathcal{G}}$, then $(i, i_1, \dots, i_p, j)$ is called the undirected path between i and j. We say that $\mathcal{G}$ is connected if there exists an undirected path between any $i, j \in \mathcal{V}$.

The following lemma presents some properties of the Laplacian matrix $\mathcal{L}$ corresponding to an undirected graph $\mathcal{G}$.

Lemma 2.1 (*[22]*)**.** *The Laplacian matrix $\mathcal{L}$ of an undirected graph $\mathcal{G}$ has the following properties:*

(i) *$\mathcal{L}$ is symmetric and positive semi-definite;*
(ii) *$\mathcal{L}$ has a simple zero eigenvalue with corresponding eigenvector equal to* $\mathbf{1}$*, and all other eigenvalues are positive if and only if the graph $\mathcal{G}$ is connected, where* $\mathbf{1}$ *denotes the vector of compatible dimension with all entries equal to* 1.

2.2. Gradient, projection operator and normal cone

For a given function $f : \mathbb{R}^m \to [-\infty, \infty]$, denote its domain as $\mathrm{dom}(f) \triangleq \{x \in \mathbb{R}^m : f(x) < \infty\}$. Let $f(\cdot)$ be a convex function, and let $x \in \mathrm{dom}(f)$. For a smooth (differentiable) function $f(\cdot)$, denote by $\nabla f(x)$ the gradient of the function $f(\cdot)$ at point x. Then

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle \quad \forall y \in \mathrm{dom}(f), \tag{1}$$

where $\langle x, y\rangle$ denotes the inner product of vectors x and y.

For a nonempty convex set $\Omega \subset \mathbb{R}^m$ and a point $x \in \mathbb{R}^m$, we call the point in Ω that is closest to x the projection of x on Ω and denote it by $P_\Omega\{x\}$. If $\Omega \subset \mathbb{R}^m$ is closed, then $P_\Omega\{x\}$ contains only one element for any $x \in \mathbb{R}^m$.

Consider a convex closed set $\Omega \subset \mathbb{R}^m$ and a point $x \in \Omega$. Define the normal cone to Ω at x as $N_\Omega\{x\} \triangleq \{v \in \mathbb{R}^m : \langle v, y-x\rangle \leq 0\, \forall y \in \Omega\}$. It is shown in [23, Lemma 2.38] that the following equation holds for any $x \in \Omega$:

$$N_\Omega\{x\} = \{v \in \mathbb{R}^m : P_\Omega\{x + v\} = x\}. \tag{2}$$

A set C is affine if it contains the lines that pass through any pairs of points $x, y \in C$ with $x \neq y$. Let $\Omega \subset \mathbb{R}^m$ be a nonempty convex set. We say that $x \in \mathbb{R}^m$ is a relative interior point of Ω if $x \in \Omega$ and there exists an open sphere S centered at x such that $S \cap \mathrm{aff}(\Omega) \subset \Omega$, where $\mathrm{aff}(\Omega)$ is the intersection of all affine sets containing Ω. A point $x \in \mathbb{R}^m$ is called the interior point of Ω if $x \in \Omega$ and there exists an open sphere S centered at x such that $S \subset \Omega$. A pair of vectors $x^* \in \Omega$ and $z^* \in \Psi$ is called a saddle point of the function $\phi(x, z)$ in $\Omega \times \Psi$ if

$$\phi(x^*, z) \leq \phi(x^*, z^*) \leq \phi(x, z^*) \quad \forall x \in \Omega,\ \forall z \in \Psi.$$

These definitions can be found in [20].

2.3. Problem formulation

Consider a network of n agents that collectively solve the following constrained optimization problem

$$\begin{aligned} &\text{minimize } f(x) = \sum_{i=1}^{n} f_i(x) \\ &\text{subject to } x \in \Omega_o = \bigcap_{i=1}^{n} \Omega_i, \end{aligned} \tag{3}$$

where $\Omega_i \subset \mathbb{R}^m$ is a closed convex set, representing the local constraint set of agent i, and $f_i(x) : \mathbb{R}^m \to \mathbb{R}$ is a smooth convex function in Ω_i, representing the local objective function of agent i. Assume that f_i and Ω_i are privately known to agent i. We assume that there exists at least one finite solution x^* to the problem (3). For the problem (3), denote by $f^* = \min_{x\in\Omega_o} f(x)$ the optimal value, and by $\Omega_o^* = \{x \in \Omega_o : f(x) = f^*\}$ the optimal solution set.

We use an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}}, \mathcal{A}_{\mathcal{G}}\}$ to describe the communication among agents. Let $\mathcal{L}$ denote the Laplacian matrix of the undirected graph $\mathcal{G}$.

Let us introduce the following conditions for the problem.

A1 Ω_o has at least one relative interior point.
A2 The undirected graph $\mathcal{G}$ is connected.
A3 For any $i \in \mathcal{V}$, $\nabla f_i(x)$ is locally Lipschitz continuous on Ω_i.

3. Algorithm design

We first give an equivalent form of the problem (3). Then define a distributed primal–dual algorithm with constant step size to solve the formulated problem.

3.1. An equivalent problem

Lemma 3.1 (*[14, Lemma 3]*)**.** *If* A2 *holds, then the problem (3) is equivalent to the following optimization problem*

$$\begin{aligned} &\textit{minimize } \tilde{f}(X) = \sum_{i=1}^{n} f_i(x_i) \\ &\textit{subject to } (\mathcal{L} \otimes \mathbf{I}_m)X = \mathbf{0}, \quad X \in \Omega, \end{aligned} \tag{4}$$

where $X = col\{x_1, \dots, x_n\} \triangleq (x_1^T, \dots, x_n^T)^T$, $\Omega = \prod_{i=1}^{n} \Omega_i$ denotes the Cartesian product, $\otimes$ denotes the Kronecker product, $\mathbf{I}_m$ denotes the identity matrix of size m, and $\mathbf{0}$ *denotes the vector of compatible dimension with all entries equal to* 0.