

切换网络分布式次梯度优化算法

李甲地 李德权

(安徽理工大学数学与大数据学院 安徽 淮南 232001)

摘要 研究了切换网络的多个体分布式次梯度优化算法。在有向切换网络是周期强连通的且对应的邻接矩阵是随机的而非双随机条件下,利用非二次李雅普诺夫函数方法证明了所提多个体分布式次梯度优化算法的收敛性。最后,通过仿真实例验证了所提算法的有效性。

关键词 多个体网络,分布式优化,有向切换网络,非二次李雅普诺夫函数,次梯度算法

中图分类号 TP301 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.01.040

Distributed Subgradient Optimization Algorithm for Multi-agent Switched Networks

LI Jia-di LI De-quan

(School of Mathematics and Big Data, Anhui University of Science and Technology, Huainan, Anhui 232001, China)

Abstract This paper studied the distributed subgradient algorithm for multi-agent optimization problem over switched networks. By using the non-quadratic Lyapunov function method, we proved that the convergence of the proposed distributed optimization algorithm can still be guaranteed under the condition that the directed switched network is periodically strongly connected and the corresponding adjacency matrix is stochastic rather than doubly stochastic. Finally, a simulation example was given to demonstrate the effectiveness of the proposed optimization algorithm.

Keywords Multi-agent network, Distributed optimization, Directed switched network, Non-quadratic Lyapunov function, Subgradient algorithm

1 引言

多个体网络是大量自主个体通过局部信息传递耦合而成的大规模网络化系统,网络中每个个体都具有自主决策、计算等能力并通过和周围邻居个体进行信息传递,协调地完成复杂任务。目前,多个体网络已被广泛应用于人类生活的各个方面,如公交车的调度问题、因特网网络系统、集群式无人飞行器的编队飞行、卫星群的协调等。

因此,研究多个体网络可以解决各类大型、复杂的现实问题,从而实现网络系统的正常运转。目前关于多个体网络的研究主要分为集中式方法和分布式方法,本文研究将采用分布式方法。

分布式凸优化问题是关于多个体网络的一类重要研究课题,如文献[1]中所讨论的无约束凸优化、带约束凸优化、博弈和 Nash 均衡等优化问题都可转化成分布式凸优化问题。一类型典型的分布式凸优化问题具有以下特点:关于整个网络的优化问题目标函数可表示为网络中所有个体各自的目标函数之和,每个个体仅知道自己的目标函数,且只能通过与其邻居个体进行信息交换协同地解决这个复杂问题。但由于实际网络中每个个体的凸函数不一定全部可微,因此文献[2]中无梯度方法将不再适用,次梯度方法是解决这类凸优化问题基

本的有效方法之一。网络中所有个体的状态变量最终趋于相同,则称所有个体达成一致^[3-5],且个体间的局部信息交换是所有个体能够达成一致的首要条件。次梯度优化算法由于涉及到个体间的局部信息交换,因此与一致性问题密切相关。在许多实际应用中,个体间的信息交换不一定是相互或双向的,即当个体 i 发送信息给其邻居个体 j 时,并不意味着个体 j 一定要发送信息给个体 i ,这就导致了通信网络通常是有向的,因此其所对应的邻接矩阵一般是随机的而未必是双随机的,如文献[6]在邻接矩阵不是双随机的条件下研究了带约束的一致性。需要指出的是:邻接矩阵是双随机的则意味着网络中的个体处于相同的地位。这个假设非常苛刻,因为这在实际应用中会带来实施方面的问题。此外,在分布式的情况下,尤其当网络在恶劣环境下运行时,由于数据的丢包或者连边的失效,或者不同个体具有不同的感知范围,使得个体间的平衡或对称信息交换变得非常困难,网络拓扑通常是时变的或切换的。因此,研究非平衡有向切换网络的多个体分布式次梯度优化算法有极强的应用背景,也极为必要。但是对这一重要问题,目前尚缺乏系统性的结论。

文献[7]最早给出了网络目标函数可分解为所有个体自身凸目标函数之和的分布式原始优化问题模型,并提出了一种分布式次梯度方法来解决这个问题,同时研究了该算法的

到稿日期:2016-12-06 返修日期:2017-03-18 本文受国家自然科学基金项目(61472003),安徽省高校学科(专业)拔尖人才学术资助重点项目(gxbjZD2016049),安徽省学术和技术带头人及后备人选科研活动项目(2016H076)资助。

李甲地(1991—),男,硕士,主要研究方向为分布式优化理论与应用,E-mail:1378475744@qq.com;李德权(1973—),男,博士,教授,硕士生导师,主要研究方向为分布式优化理论、分布式协调控制。

收敛性。文献[8]则利用分布式次梯度算法和一致性算法解决了一类凸优化问题,并分析了算法的优劣性以及收敛速度。文献[9]研究了个体间交互的信息被量化后对分布式次梯度算法收敛速度等性能的影响。文献[10]研究了随机网络的次梯度算法的优化问题。文献[11]则进一步研究了基于量化次梯度算法和数据交流速率分析的分布式优化问题,并针对切换或者固定拓扑网络,研究了如何在最小化量化水平数或数据率的情况下仍能实现网络优化目标的问题。但是文献[7-11]通常都假定信息通信网络所对应的邻接矩阵是双随机,即只考虑了信息通信网络所对应的图是无向图或者是平衡有向图,而未考虑所对应的图是非平衡有向图的情形。

最近,文献[12]研究了具有敌对节点时有向网络的分布式优化及一致性问题。受文献[12]启发,并作为对文献[7]研究的完善,本文研究了非平衡有向切换网络的多个体分布式次梯度算法。但与文献[7]不同的是,本文考虑的信息通信网络所对应的邻接矩阵是随机的而未必是双随机的。进一步地,基于非二次李雅普诺夫函数方法,本文证明了只要个体状态的最大值与最小值之差有上界,则所提出的多个体分布式次梯度优化算法就收敛。

2 问题描述

本文考虑如下的分布式优化问题:

$$\begin{aligned} \min_x f(x) &= \sum_{i=1}^n f_i(x) \\ \text{s. t. } x &\in R^m \end{aligned} \quad (1)$$

其中, $f(x)$ 为网络目标函数,等于每个个体的局部目标函数之和。

不可微凸函数:

$$f_i(x): R^m \rightarrow R, i=1, 2, \dots, n$$

为个体 i 的局部目标函数,且仅为个体 i 所知。

函数 $f(x)$ 的定义域定义为: $\text{dom}(f) = \{x \in R^m \mid f(x) < \infty\}$ 。若对于 $\bar{x} \in \text{dom}(f)$, $f(x) \geq f(\bar{x}) + d^T(\bar{x})(x - \bar{x})$, $\forall x \in \text{dom}(f)$ 成立,则称 $d(\bar{x})$ 为函数 $f(x)$ 在 \bar{x} 处的次梯度。次梯度是可微函数梯度概念的推广。若 $f_i(x)$ 满足 $f_i(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f_i(x_1) + (1-\lambda)f_i(x_2)$, $\forall \lambda \in [0, 1], x_1, x_2 \in \text{dom}(f_i)$, 则称 $f_i(x)$ 为凸函数。

已有研究表明:通过网络个体间的局部信息交换和计算,能协同地解决优化问题(1)。

假设 1^[7] 问题(1)是可解的且最优解集是非空的,并记最优解集为 X^* ,即:

$$X^* = \{x^* \in R \mid \sum_{i=1}^n f_i(x^*) = f(x^*)\}$$

其中, $f(x^*)$ 为最优值, $x^* \in X^*$ 为最优解。

假设 2 每个个体的局部目标函数的次梯度是有界的,即存在一个正整数 μ 使得:

$$\|d_i(x)\|_2 \leq \mu, \forall d_i(x) \in \partial f_i(x), x \in \text{dom}(f)$$

3 图论的相关知识

有向切换网络中个体间的信息交换可建模成一个有向图 $G=(V, E(t), A(t))$,其中 $V=\{1, 2, \dots, n\}$ 表示个体集合, n 为个体总数, $E(t)=\{(j, i) \mid i, j \in V\}$ 表示网络中所有有向边构成的集合, $A(t)=(a_{ij}(t))_{n \times n}$ 为网络所对应的邻接矩阵。有

向边 $(j, i) \in E(t)$ 表示在 t 时刻个体 j 向个体 i 发送信息,此时有向边 (j, i) 的边权 $a_{ij}(t) > 0$,且称个体 j 是个体 i 的入度邻居,否则 $a_{ij}(t) = 0$ 。由此定义个体 i 的入度邻居集合为 $N_i = \{j \mid (j, i) \in E(t), j \in V, i \neq j\}$ 。类似地,可以定义个体 i 的出度邻居集合。对于无向图而言,入度邻居集合和出度邻居集合是相同的,但有向图则不然。若图 G 的有序节点序列 (i_1, i_2, \dots, i_r) 满足 $i_{j,j+1} \in E(t)$,其中 $j \in \{1, \dots, r-1\}$,则称这个有序节点序列 (i_1, i_2, \dots, i_r) 为有向图 G 的一条有向路径或强路径。若有向图 G 中任意不同的两个有序节点之间都存在一条强路径,则称 G 是强连通的。

若矩阵 $A(t)=(a_{ij}(t))_{n \times n}$ 的元素 $a_{ij}(t)$ 满足 $a_{ij}(t) \geq 0$,则称矩阵 $A(t)$ 为非负的。如果非负矩阵 A 每一行的行和均为 1,则称矩阵 $A(t)$ 为随机矩阵。进一步地,如果 $A^T(t)$ 也是随机的,则 $A(t)$ 称为双随机的。因此, $A(t)$ 为双随机的当且仅当所对应的加权有向图 G 是平衡图,而 $A(t)$ 为随机的则意味着其所对应的加权有向图 G 是非平衡图。显然,有向网络拓扑包含平衡拓扑和非平衡拓扑两种情形。

本文将着重研究所有个体均具有自环的多个体非平衡有向网络的分布式次梯度优化算法,目的是设计适当的分布式优化算法,使得网络中所有个体的状态最终达成一致,即最终实现 $x_1 = x_2 = \dots = x_n = x^*$,且整个网络的目标函数在 x^* 达到最优,即 $\min_x f(x) = f(x^*)$ 。

4 切换网络分布式次梯度优化算法

在非平衡加权有向图的情况下,针对优化问题(1),本文提出如下的线性分布式次梯度优化算法:

$$x_i(t+1) = a_{ii}(t)x_i(t) + \sum_{j \in N_i} a_{ij}(t)x_j(t) - a_t d_i(t) \quad (2)$$

其中, $i \in V$, $a_{ij}(t)$ ($j \in \{i\} \cup N_i(t)$) 为随机矩阵 A 的第 i 行第 j 列的元素; $d_i(t)$ 为不可微凸函数 $f_i(x)$ 的次梯度; a_t 是 t 时刻的步长。算法(2)实际上包括两部分:1)每个个体与其邻居节点进行信息交换以及自身状态更新的一致性运算;2)每个个体沿其目标函数次梯度方向的迭代运算。

在实际应用中,网络的不可靠通信会造成网络拓扑是动态切换的。对于非平衡加权有向图,有如下假设:

假设 3(周期强连通性) 存在一个正数 $B \geq 1$,对于任意 $t \geq 0$,有向联合图 $(G, E(t) \cup E(t+1) \cup \dots \cup E(t+B-1))$ 是强连通的,所对应的邻接矩阵 A 是随机的。

假设 3 意味着在一个有限周期内,有向图的联合图是强连通的,从而确保个体间的信息能够得到有效的交换。

假设 4 邻接矩阵 $A(t)$ 是一个具有正对角元素的随机矩阵,即存在一个正常数 ρ ,使得对于所有的 $i \in V$ 和 $t \geq 0$, $a_{ii}(t) = 1 - \sum_{j \in N_i(t)} a_{ij}(t) \geq \rho > 0$ 一致成立;同时对于 $\forall i \neq j$, $a_{ij}(t)$ 满足 $a_{ij}(t) \in \{0\} \cup (\rho, 1]$ 。

5 分布式次梯度优化算法的收敛性分析

由于通信网络所对应的是非平衡有向图,因此文献[7]中的证明方法对本文所提算法的收敛性将不再适用。受文献[12]研究的启发,下文将利用非二次李雅普诺夫函数方法^[3-4]以及文献[5,13]中的相关研究结论来证明非平衡有向图的次梯度优化算法的收敛性。

为此,定义:

$$M(t) = \max_{i \in V} x_i(t) \quad m(t) = \min_{i \in V} x_i(t)$$

$$\Delta r_{\max}(t) = \max_{i \in V} \Delta r_i \quad \Delta r_{\min}(t) = \min_{i \in V} \Delta r_i$$

并进一步定义:

$$D(t) = M(t) - m(t) \quad \Delta R(t) = \Delta r_{\max}(t) - \Delta r_{\min}(t)$$

显然,对于任意的 $t \geq 0$, $D(t) = \max_{i,j \in V} |x_i(t) - x_j(t)| \geq 0$ 成立。进一步地,若 $\lim_{t \rightarrow \infty} D(t) = 0$, 则算法(2)达成一致。因此,非二次李雅普诺夫函数 $D(t)$ 刻画了一致性误差,从而可以用来定量分析所提算法的收敛性^[3-4]。

此外,由假设4可知 $a_{ii}(t) = 1 - \sum_{j \in N_i} a_{ij}(t) > 0 (i \in V)$, 则算法(2)可进一步写成:

$$x_i(t+1) = a_{ii}(t)x_i(t) + \sum_{j \in N_i} a_{ij}(t)x_j(t) - \alpha_i d_i(t)$$

$$= (1 - \sum_{j \in N_i} a_{ij}(t))x_i(t) + \sum_{j \in N_i} a_{ij}(t)x_j(t) - \alpha_i d_i(t)$$

$$= x_i(t) - \sum_{j \in N_i} a_{ij}(t)x_i(t) + \sum_{j \in N_i} a_{ij}(t)x_j(t) - \alpha_i d_i(t)$$

$$= x_i(t) + \sum_{j \in N_i} a_{ij}(t)(x_j(t) - x_i(t)) - \alpha_i d_i(t)$$

$$= x_i(t) + \sum_{j \in N_i} a_{ij}(t)(x_j(t) - x_i(t)) + \Delta r_i(t) \quad (3)$$

其中, $t=0, 1, 2, \dots; \Delta r_i(t) = -\alpha_i d_i(t)$ 。

另外,由假设2可得 $\Delta r_{\max}(t) \leq \alpha_i \mu$, $\Delta r_{\min}(t) \geq -\alpha_i \mu$, 因此可得如下关系式:

$$\Delta R(k) = \Delta r_{\max}(t) - \Delta r_{\min}(t) \leq 2\alpha_i \mu \quad (4)$$

对假设3作进一步的解读,对于任意的 $s \geq 0$, 限定个体 $v \in V$, 并令 $V_0 = \{v\}$, 显然若假设3成立, 一定存在一个非空集合 $V_1 = V/V_0$, 使得在时间间隔 $[s, s+(B-1)]$ 内, 任意的个体 $i \in V_1$ 和个体 v 至少交换一次信息。同理, 根据数学归纳法可知, $V_{l+1} = V/(V_0 \cup V_1 \cup \dots \cup V_l)$ 表示的是这样的集合: 对于任意的个体 $j \in V_{l+1}$, 一定存在某个个体 $i \in (V_0 \cup V_1 \cup \dots \cup V_l)$, 使得在时间间隔 $[s+lB, s+((l+1)B-1)]$ 内个体 i 和个体 j 至少交换一次信息; 且由假设3可知, 只要 $V/(V_0 \cup V_1 \cup \dots \cup V_l) = \emptyset$, 则必然有 $V_{l+1} \neq \emptyset$ 。因此, 集合 V_0, V_1, \dots, V_L 是个体集合 V 的一个分割, 这里 $L \leq n-1$ 。

基于以上知识, 可以发现优化算法(2)的另外一种形式(3)和文献[13]所提出的一阶动态一致性算法(简称 FO-DAC)具有完全相同的形式。当假设3和假设4成立时, 文献[13]中引理3.1和定理3.1的所有假设条件都满足。因此借鉴文献[13]中的引理3.1和定理3.1的相关结论, 给出如下引理及证明。

引理1 假定假设3和假设4成立, 正整数 $s \geq 0$, 且固定 $v \in V$, 集合 V_0, V_1, \dots, V_L 是个体集合 V 的一个分割, 则对于任意的 $l \in [1, 2, \dots, L] (L \leq n-1)$, 一定存在一个 $\eta_l \geq 0$, 其中 $\eta_0 = \rho^{nB-1}$, 使得对于正整数 $p \in [lB, lB+1]$ 和 $i \in V_l$, 当 $t = s+p$ 时, 如下不等式成立。

$$x_i(t) \geq m(s) + \sum_{q=0}^{p-1} \Delta r_{\min}(s+q) + \eta_l (x_v(s) - m(s)) \quad (5)$$

$$x_i(t) \leq M(s) + \sum_{q=0}^{p-1} \Delta r_{\max}(s+q) - \eta_l (M(s) - x_v(s)) \quad (6)$$

证明: 这里仅考虑 $s=0$ 的情况, s 的其他情况与此类似。

对于给定的某一个体 i , 令 $t=0$, 由式(3)可知, 如下不等式成立。

$$x_i(1) = x_i(0) + \sum_{j \in N_i} a_{ij}(0)(x_j(0) - x_i(0)) + \Delta r_i(0)$$

$$= (1 - \sum_{j \in N_i} a_{ij}(0))x_i(0) + \sum_{j \in N_i} a_{ij}(0)x_j(0) + \Delta r_i(0)$$

$$\geq (1 - \sum_{j \in N_i} a_{ij}(0))m(0) + \sum_{j \in N_i} a_{ij}(0)m(0) + \Delta r_{\min}(0)$$

$$= m(0) + \Delta r_{\min}(0) \quad (7)$$

类似地, 由式(7)可得:

$$x_i(1) \geq m(0) + \Delta r_{\min}(0)$$

$$x_i(2) \geq m(1) + \Delta r_{\min}(1)$$

$$\dots$$

$$x_i(t) \geq m(t-1) + \Delta r_{\min}(t-1)$$

把不等式(8)分别相加, 可得:

$$x_i(t) \geq m(0) + \sum_{p=0}^{t-1} \Delta r_{\min}(p) \quad (9)$$

注意到, 对于任意的 $t \geq 0$, 有 $\sum_{j=1}^n a_{kj}(t) = 1$ 成立; 同时利用式(9), 得:

$$x_k(t+1) - m(0) - \sum_{p=0}^t \Delta r_{\min}(p)$$

$$= \sum_{j=1}^n a_{kj}(t)x_j(t) - \Delta r_k(t) - m(0) - \sum_{p=0}^t \Delta r_{\min}(p)$$

$$= \sum_{j=1}^n a_{kj}(t)(x_j(t) - m(0) - \sum_{p=0}^{t-1} \Delta r_{\min}(p)) - \Delta r_k(t) - \Delta r_{\min}(t)$$

$$\geq a_{kk}(t)(x_k(t) - m(0) - \sum_{p=0}^{t-1} \Delta r_{\min}(p))$$

$$\geq \rho(x_k(t) - m(0) - \sum_{p=0}^{t-1} \Delta r_{\min}(p)) \quad (10)$$

对于任意的正整数 $p \in [0, LB+B-1]$, 令 $t=p$, 对式(10)中的括号部分进行类似推导, 可得:

$$x_k(t) - m(0) - \sum_{q=0}^{p-1} \Delta r_{\min}(q)$$

$$\geq \rho^{p-1}(x_k(1) - m(0) - \Delta r_{\min}(0))$$

$$\geq \rho(x_k(0) - m(0))$$

$$\geq \eta_0(x_k(0) - m(0)) \quad (11)$$

其中, $\eta_0 = \rho^{nB-1}$ 。式(10)利用了式(9)和 $x_k(0) - m(0) \geq 0$ 。

下面用数学归纳法证明式(5)。对于给定的 $0 \leq l < L$, 假设式(5)成立, 则给定某一个体 $i \in V_{l+1}$, 一定存在某个个体 $j \in (V_0 \cup V_1 \cup \dots \cup V_l)$, 使得在时间间隔 $t' \in [lB, lB+B-1]$ 内 $a_{ij}(t') \neq 0$ 。由归纳假设可知, 对于任意的正整数 $p \in [lB, lB+B-1]$, 令 $t=p$, 有以下不等式成立。

$$x_j(t) - m(0) - \sum_{q=0}^{t-1} \Delta r_{\min}(q) \geq \eta_l (x_k(0) - m(0)) \quad (12)$$

进一步地, 把式(12)代入式(10)可得:

$$x_i(t'+1) - m(0) - \sum_{q=0}^{t'} \Delta r_{\min}(q)$$

$$\geq a_{ij}(t')(x_j(t') - m(0) - \sum_{q=0}^{t'-1} \Delta r_{\min}(q))$$

$$\geq \rho \eta_l (x_k(0) - m(0)) \quad (13)$$

再把式(13)代入式(10), 以此类推, 并令 $t=p$, 有:

$$x_i(t+1) - m(0) - \sum_{q=0}^p \Delta r_{\min}(q) \geq \eta_{l+1} (x_k(0) - m(0))$$

其中, $p \in [(l+1)B, lB+B-1]$, $\eta_{l+1} = \rho^{(n-l)B} \eta_l$ 。则式(5)对 $i \in V_{l+1}$ 成立。由数学归纳法可知, 式(5)成立。式(6)的证明类似。引理1证毕。

进一步地,令 $\eta = \rho^{\frac{1}{2}n(n+1)B-1}$, 则对于任意的 $l \in [1, 2, \dots, L]$ ($L \leq n-1$), $\eta \leq \eta_l$ 成立。现用 k 和 $k_1 = k + LB + B - 1$ 分别代替式(5)和式(6)中的 s 和 t , 则有:

$$\begin{aligned} m(k_1) &= \min_{l \in \{1, 2, \dots, L\}} \min_{i \in V_l} x_i(k_1) \\ &\geq m(k) + \sum_{q=k}^{k_1-1} \Delta r_{\min}(q) + \min_{l \in \{1, 2, \dots, L\}} \eta_l (x_v(k) - m(k)) \\ &\geq m(k) + \sum_{q=k}^{k_1-1} \Delta r_{\min}(q) + \eta (x_v(k) - m(k)) \quad (14) \end{aligned}$$

同理可得:

$$\begin{aligned} M(k_1) &= \max_{l \in \{1, 2, \dots, L\}} \max_{i \in V_l} x_i(k_1) \\ &\leq M(k) + \sum_{q=k}^{k_1-1} \Delta r_{\max}(q) + \max_{l \in \{1, 2, \dots, L\}} \eta_l (M(k) - x_v(k)) \\ &\leq M(k) + \sum_{q=k}^{k_1-1} \Delta r_{\max}(q) - \eta (M(k) - x_v(k)) \quad (15) \end{aligned}$$

由式(14)和式(15)以及 $D(t)$ 的定义可得如下定理 1。

定理 1 假定假设 3 和假设 4 成立。令 $\eta = \rho^{\frac{1}{2}n(n+1)B-1}$, 对于任意的 $k \geq 0$, 令 $k_1 = k + LB + B - 1$, 则有:

$$D(k_1) \leq (1-\eta)D(k) + \sum_{q=k}^{k_1-1} \Delta R(q) \quad (16)$$

证明: 由式(14)、式(15)可得如下不等式:

$$\begin{aligned} M(k_1) - m(k_1) &\leq M(k) - m(k) + \sum_{q=k}^{k_1-1} \Delta r_{\max}(q) - \sum_{q=k}^{k_1-1} \Delta r_{\min}(q) - \eta (M(k) - x_v(k)) - \eta (x_v(k) - m(k)) \\ D(k_1) &\leq D(k) + \sum_{q=k}^{k_1-1} (\Delta r_{\max}(q) - \Delta r_{\min}(q)) - \eta (M(k) - x_v(k) + x_v(k) - m(k)) \\ &\leq D(k) + \sum_{q=k}^{k_1-1} \Delta R(q) - \eta D(k) \\ &\leq (1-\eta)D(k) + \sum_{q=k}^{k_1-1} \Delta R(q) \end{aligned}$$

上述不等式的证明利用了非二次李雅普诺夫函数的定义。此外, 对于正整数 $h \geq 1$, 定义 $h(nB-1) = T_h$ 。进而对于任意的正整数 $t \geq 1$, 令 s 是满足 $s(nB-1) = T_s \leq t < (s+1)(nB-1)$ 的最大整数, 从而由式(16)可得:

$$\begin{aligned} D(t) &\leq (1-\eta)D(T_s) + \sum_{q=T_s}^{t-1} \Delta R(q) \\ &\leq (1-\eta)^s D(0) + (1-\eta)^{s-1} \sum_{q=0}^{T_s-1} \Delta R(q) + \dots + (1-\eta) \sum_{q=T_{s-1}-1}^{T_s-1} \Delta R(q) + \sum_{q=T_s}^{t-1} \Delta R(q) \\ &= (1-\eta)^{s+1} \frac{1}{1-\eta} D(0) + \Omega(t) \\ &= (1-\eta)^{\frac{(s+1)(nB-1)}{nB-1}} \frac{1}{1-\eta} D(0) + \Omega(t) \\ &\leq (1-\eta)^{\frac{t}{nB-1}} \frac{1}{1-\eta} D(0) + \Omega(t) \\ &\leq (1-\eta)^{\frac{t-1}{nB-1}} \frac{1}{1-\eta} D(0) + \Omega(t) \\ &\leq (1-\eta)^{\frac{t-nB}{nB-1}} D(0) + \Omega(t) \quad (17) \end{aligned}$$

上述不等式利用了 $0 < 1-\eta < 1$ 和 $t < (s+1)(nB-1)$, 且

其中 $\Omega(t)$ 为:

$$\begin{aligned} \Omega(t) &= (1-\eta)^{s-1} \sum_{q=0}^{T_s-1} \Delta R(q) + \dots + (1-\eta) \sum_{q=T_{s-1}-1}^{T_s-1} \Delta R(q) + \sum_{q=T_s}^{t-1} \Delta R(q) \\ &\leq 2(nB-2)\alpha_i\mu [(1-\eta)^{s-1} + \dots + (1-\eta) + 1] \\ &\leq 2(nB-2) \frac{1}{\eta} \alpha_i\mu \\ &\leq 2(nB-2) \frac{1}{\rho^{\frac{1}{2}n(n+1)B-1}} \alpha_i\mu \quad (18) \end{aligned}$$

该不等式的证明利用了式(4)。

进一步地, 在本文所提算法(2)的作用下, 对优化问题(1)有如下主要结论。

定理 2 设个体的初始条件状态值 $x_i(0)$ ($i \in V$) 满足 $\|x(0)\|_{\infty} = \max_{1 \leq i \leq n} |x_i(0)| \leq C_{\delta}$, $D(0) = M(0) - m(0) \leq C_{\delta}$, 其中 C_{δ} 是已知的非负常数, 则以下不等式成立。

$$D(t) \leq (1-\eta)^{\frac{t-nB}{nB-1}} C_{\delta} + 2(nB-2) \frac{1}{\rho^{\frac{1}{2}n(n+1)B-1}} \alpha_i\mu \quad (19)$$

证明: 由定理 1 和上述条件, 式(17)可进一步写为式(19), 定理 2 证毕。

定理 2 揭示了 $D(t)$ 有上界, 且该上界值与有向图的周期 B 、个体初始条件状态值、算法的步长 α_i 及次梯度的有界值 μ 有关。定理 2 还表明: 在所提多个体分布式次梯度优化算法(2)的作用下, 非平衡有向切换网络中的个体状态最终达到一致, 即可求得最优解 x^* , 不仅满足 $x^* = x_1 = x_2 = \dots = x_n$, 且使得 $f(x)$ 在 x^* 处达到最优值 $f(x^*)$ 。

6 仿真分析

下面对本文结论分析进行仿真实验。考虑如图 1 所示的具有 3 个个体的有向切换网络 $G(t) = (V, E(t), A(t))$, 其中个体集合为 $V = \{1, 2, 3\}$, 有向联合图 $E(3t) \cup E(3t+1) \cup E(3t+2) = G(a) + G(b) + G(c)$, 即网络切换周期 $B=3$ 。求最优解 x^* , 使得目标函数 $f(x^*) = \sum_{n=1}^3 f_n(x^*)$ 最优。其中网络中各子图有向边上的数字表示相应边上的权重 $a_{ij}(t)$ 。各个子图对应的随机邻接矩阵 $A(t)$ 分别为:

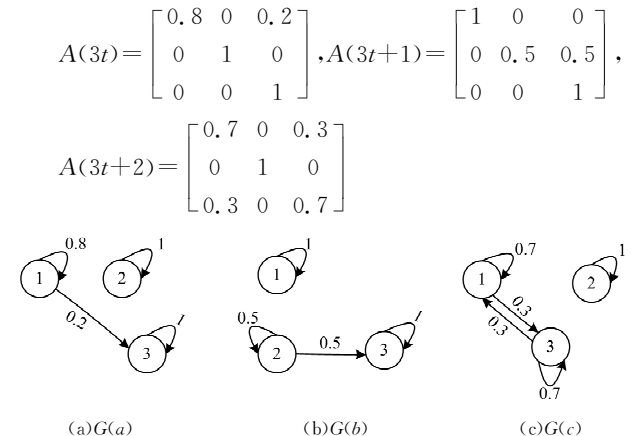


图 1 具有 3 个节点的有向切换网络

Fig. 1 A switching directed network with three nodes

设分布式优化问题(1)的目标函数为实数集上的二次函数, 则系统中所对应的目标函数为 $f_i = (x - a_i)^2$, $a_i \in R, i =$

1, 2, 3, 其中 a_i 为随机选取的实数。算法中梯度和步长都是时变的, 即 $d_i = 2(x_i(t) - a_i)$, 步长 $\alpha_i = \frac{1}{\sqrt{t}}$ 。利用经典微分理论求出优化问题(1)的最优解为:

$$x^* = \frac{a_1 + a_2 + a_3}{3}$$

则每个个体的局部目标函数的最优值为 $f_i(x^*)$, 将其代入可得问题(1)的全局目标函数最优值为:

$$f(x^*) = f_1(x^*) + f_2(x^*) + f_3(x^*)$$

图2中横轴表示迭代时间 t , 纵轴表示个体的状态值 $x_i(t)$ 。图2表明: 所有个体状态最终达成一致。图3中横轴表示迭代时间 t , 纵轴表示目标函数 $f(x_i(t))$ 。图3则进一步表明: 3个个体的目标函数 $f_i(t)$ 都收敛到最优值的邻域内。图4给出了在图1的网络拓扑情形下问题(1)的最优目标函数值 $f(x^*)$ 分别在双随机分布式次梯度优化算法和随机分布式次梯度优化算法作用下的变化情况。由图4可知, $f(x^*)$ 在这两种算法的运算下都收敛到最优值的邻域内; 但明显地, 邻接矩阵是双随机的分布式次梯度优化算法的收敛速度稍快; 而邻接矩阵是随机的分布式次梯度优化算法的具有更小的迭代误差。

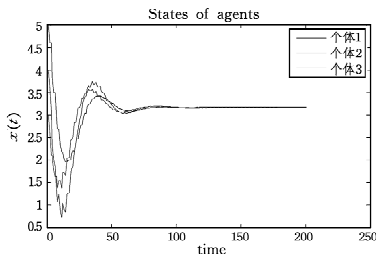


图2 3个个体的状态值 $x(t)$ 的轨迹图

Fig. 2 The trajectories of the state value $x(t)$ of three individuals

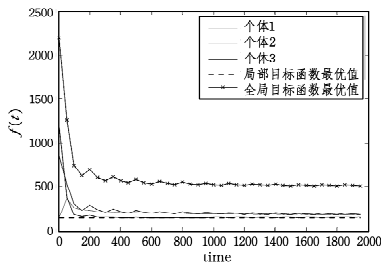


图3 3个个体的局部目标函数值 $f_i(t)$ 以及全局目标函数值 $f(x^*)$ 的轨迹图

Fig. 3 The trajectories of three individuals' local objective function value $f_i(t)$ and the global objective function value $f(x^*)$

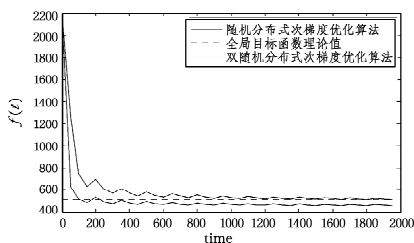


图4 $f(x^*)$ 在双随机分布式次梯度优化算法和随机分布式次梯度优化算法作用下的变化趋势

Fig. 4 The variation trend of $f(x^*)$ under double stochastic distributed subgradient optimization algorithm and stochastic distributed subgradient optimization algorithm

结束语 在文献[5, 12-13]研究的基础上, 本文证明了若有向切换网络所对应的邻接矩阵是随机的, 则所提出的分布式次梯度优化算法是收敛的。在证明过程中, 主要利用了非二次李雅普诺夫函数刻画算法的收敛性误差, 并给出仿真实例验证了理论分析的有效性。本文只是一个初步的研究, 还有许多问题有待解决, 例如非平衡有向切换网络的量化次梯度优化算法将是下一步的研究工作。

参考文献

- [1] HONG Y G, ZHANG Y Q. Distributed optimization: algorithm design and convergence analysis[J]. Control Theory and Applications, 2014, 31(7): 850-857. (in Chinese)
洪奕光, 张艳琼. 分布式优化: 算法设计和收敛性分析[J]. 控制理论与应用, 2014, 31(7): 850-857.
- [2] LI D Q, CHEN P. Distributed Random Projection Gradient-Free Optimization Algorithm for Multi-Agent Networks[J]. Journal of Frontiers of Computer Science and Technology, 2016(11): 1564-1570. (in Chinese)
李德权, 陈平. 多个体网络分布式随机投影无梯度优化算法[J]. 计算机科学与探索, 2016(11): 1564-1570.
- [3] TSITSIKLIS J N, ATHANS M. Convergence and asymptotic agreement in distributed decision problems[J]. IEEE Transactions on Automatic Control, 1984, 29(1): 42-50.
- [4] TSITSIKLIS J N. Problems in decentralized decision making and computation[D]. Cambridge: Massachusetts Institute of Technology, MA, 1984.
- [5] LI D Q. On quantized and robust consensus for multi-agent systems with directed network topologies[D]. Shanghai: Shanghai Jiao Tong University, 2012. (in Chinese)
李德权. 有向网络多个体系统的量化与鲁棒一致性研究[D]. 上海: 上海交通大学, 2012.
- [6] LIN P, REN W. Distributed Constrained Consensus in the Presence of Unbalanced Switching Graphs and Communication Delays[C]// 51st IEEE Conference on Decision and Control. Maui, Hawaii, USA, 2012: 2238-2243.
- [7] NEDIC A, OZDAGLAR A. Distributed subgradient methods for Multi-Agent optimization [J]. IEEE Transactions on Automatic Control, 2009, 54(1): 48-61.
- [8] JOHANSSON B, KEVICZKY T, JOHANSSON M, et al. Subgradient methods and consensus algorithms for solving convex optimization problems[C]// IEEE Conference on Decision and Control (CDC). 2008.
- [9] NEDIC A, OZDAGLAR A, TSITSIKLIS J N. Distributed Subgradient Methods and Quantization Effects[C]// IEEE Conference on Decision and Control. 2008: 9-11.
- [10] LOBEL I, OZDAGLAR A. Distributed subgradient methods for convex optimization over random networks[J]. IEEE Transactions on Automatic Control, 2011, 56(6): 1291-1306.
- [11] YI P, HONG Y G. Quantized Subgradient Algorithm and Data-Rate Analysis for Distributed Optimization[J]. IEEE Transactions on Control of Network Systems, 2014, 1(4): 380-392.
- [12] SUNDARAM S, GHARESI B. Distributed Optimization Under Adversarial Nodes[J]. arXiv:1606.08939v1, 29.06.2016.
- [13] ZHU M H, MARTINEZ S. Discrete-time dynamic average consensus[J]. Automatic, 2010, 46(2): 322-329.