



权重平衡有向网络下分布式约束优化的 连续时间算法设计

摘要

本文研究权重平衡有向网络下分布式约束优化问题的求解,其中网络的全局目标函数是由每个智能体的局部目标函数的和构成,全局的约束是由每个智能体的局部约束的交构成.为了分布式求解该问题的最优解,首先引入智能体的局部共轭函数将其转换为 Fenchel 对偶问题.其次,从 Fenchel 对偶问题出发,提出一类基于奇异摄动系统的分布式连续时间算法.在局部目标函数和其梯度分别满足强凸和 Lipschitz (李普希兹)连续的情况下,结合凸分析方法和 Lyapunov (李雅普诺夫)稳定性理论,结果表明所提算法能够获得原问题和对偶问题的最优值.最后,数值仿真进一步验证了所提算法的有效性.

关键词

分布式约束优化;权重平衡有向网络;Fenchel 对偶问题;奇异摄动系统

中图分类号 TP18;O224

文献标志码 A

收稿日期 2020-07-01

资助项目 南京信息工程大学人才启动经费(2020-012);国防科技创新特区项目(2019)

作者简介

朱亚楠,女,博士,讲师,研究方向为分布式优化与博弈. ynzhu@nuist.edu.cn

温广辉(通信作者),男,博士,教授,博士生导师,研究方向为分布式控制与优化. ghwen@seu.edu.cn

0 引言

在过去的几十年里,多智能体系统的分布式协调问题因其在工程、自然、社会科学等领域的广泛应用,已有大量的研究成果.分布式优化作为多智能体系统中的一个重要协调问题,其目的是多智能体通过与邻居之间的相互协作实现整个网络最优决策.在分布式优化问题的求解方面,早期的工作主要基于迭代格式的离散时间算法.比如, Nedic 等将多智能体一致机制和优化方法相结合,分别设计了用于求解无约束优化和凸约束优化的分布式次梯度算法和分布式次梯度投影算法^[1-2].另外,文献[3-8]建立了求解等式或不等式约束优化的分布式离散时间优化算法.

由于连续时间系统在算法设计和分析方面的优势,基于连续时间系统的分布式优化算法设计近年来受到了研究者的广泛关注.针对无向网络下无约束分布式优化问题,文献[9-10]分别设计了基于零梯度和一致性和基于牛顿-辛普森一致性的连续时间算法.文献[11-12]针对权重平衡有向网络下无约束的情况,从原始对偶的角度出发,分别提出了基于鞍点动力学和基于负反馈梯度流的连续时间算法.为了消除参数对全局网络信息的依赖性,文献[13]提出了一种自适应的连续时间算法.通过结合估计左特征向量的一致性协议和权重平衡有向网络下的连续时间算法^[12],文献[14]进一步给出了权重非平衡有向网络下的连续时间协调算法.相比较无约束的情况,约束优化问题在实际中的应用更加广泛^[15-20].考虑到多个局部约束交的情况,文献[15]设计了基于二阶多智能体系统的分布式优化算法.为了避免上述工作中乘子变量的无界性,文献[16]进一步提出了基于原始对偶的连续时间算法.此外,文献[17]采用神经网络动力学求解局部等式和不等式约束的分布式优化问题.最近,文献[18-20]研究了同时包含局部等式、不等式和闭凸集约束的分布式优化问题.对于更多的分布式优化算法的工作,读者可参考一些文献综述^[21-23].

从上述的工作中可以发现,分布式约束优化问题的连续时间算法设计主要针对无向网络.在有向网络下,智能体之间的信息交互往往是不对称的,这也导致无向网络下基于对称拉普拉斯矩阵的算法通常不再适用有向网络的情况.因此,有待进一步探讨有向网络下约

1 南京信息工程大学 自动化学院,南京,210044

2 东南大学 数学学院/江苏省网络群体智能重点实验室,南京,211189

束优化问题的连续时间算法设计.本文考虑一类带有局部闭凸集约束的分布式优化问题,其中网络的全局目标函数由智能体的局部目标函数的和构成.目的是在权重平衡有向网络下,智能体分布式合作找到该问题的全局最优解.对于该问题的研究,文献[24]给出了基于 Fenchel 对偶梯度的离散时间算法,该算法的详细收敛性分析在文献[25]中给出.由于获得 Fenchel 对偶梯度仍需要求解最优化问题,因此很难在迭代中直接使用 Fenchel 对偶梯度.为此,本文从不精确的 Fenchel 对偶梯度出发,结合一致性机制,提出一类基于奇异摄动系统的分布式连续时间算法.针对所提出的算法,结合凸分析和 Lyapunov 稳定性理论,对所提算法进行收敛性分析及数值仿真验证.

1 预备知识

本节给出论文中使用的数学符号、代数图论及凸分析中的相关概念和理论预备知识.

1.1 符号说明

记 \mathbf{R} 和 \mathbf{R}^n 分别表示实数域和 n 维欧式空间. $\mathbf{R}^{p \times q}$ 表示 \mathbf{R} 上的 $p \times q$ 维的矩阵空间.给定向量 $\mathbf{x} \in \mathbf{R}^n$ 和矩阵 $\mathbf{A} \in \mathbf{R}^{p \times q}$, \mathbf{x} 和 \mathbf{A} 的欧式范数分别为 $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$ 和 $\|\mathbf{A}\| = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})}$, 其中 $\lambda_{\max}(\mathbf{A}^T \mathbf{A})$ 表示矩阵 $\mathbf{A}^T \mathbf{A}$ 的最大特征值. \mathbf{I}_n 表示 n 阶单位矩阵. $\mathbf{1}_n(\mathbf{0}_n)$ 表示元素全为 1(0) 的 n 维列向量. 给定 N 个列向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$, $\text{col}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = (\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_N^T)^T$ 表示按照 \mathbf{x}_1 到 \mathbf{x}_N 顺序堆叠的列向量. $\text{diag}(a_1, a_2, \dots, a_n)$ 表示以实数 a_1, a_2, \dots, a_n 为对角元的对角矩阵. 符号 \otimes 表示克罗内克 (Kronecker) 积.

1.2 代数图论

N 个智能体之间的信息交互可表示为一个有向通信网络 $\mathcal{G}(\mathcal{V}, \mathcal{E})$, 其中 $\mathcal{V} = \{1, 2, \dots, N\}$ 是节点(智能体)的集合, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ 是节点之间的边构成的集合(假定没有自环). 边 $(i, j) \in \mathcal{E}$ 表示智能体 i 可以收到来自智能体 j 的信息. 一条有向路径是由边连接的有序节点对构成. 若任意两个节点之间都存在一条有向路径, 则称有向网络 \mathcal{G} 是强连通的. 定义 $\mathbf{A} = (a_{ij})_{N \times N}$ 是 \mathcal{G} 的权重邻接矩阵. 若 $(i, j) \in \mathcal{E}$, 则 $a_{ij} > 0$; 否则 $a_{ij} = 0$. 采用 $d_i^{\text{in}} = \sum_{j=1}^N a_{ij}$ 和 $d_i^{\text{out}} = \sum_{j=1}^N a_{ji}$ 分别表示智能体 i 的入度和出度. 相应地, \mathcal{G} 的 Laplacian (拉普拉斯矩阵) $\mathbf{L} = (l_{ij})_{N \times N}$ 定义为 $\mathbf{L} = \mathbf{D}^{\text{in}} - \mathbf{A}$, 其中

$\mathbf{D}^{\text{in}} = \text{diag}(d_1^{\text{in}}, d_2^{\text{in}}, \dots, d_N^{\text{in}})$. 由 \mathbf{L} 的定义可知, $\mathbf{L}\mathbf{1}_N = \mathbf{0}_N$. 进一步, 若 $\mathbf{1}_N^T \mathbf{L} = \mathbf{0}_N^T$, 称有向网络 \mathcal{G} 是权重平衡的.

引理 1^[12] 假定 \mathcal{G} 是权重平衡有向强连通网络, 其 Laplacian 为 \mathbf{L} . 则对称矩阵 $\frac{\mathbf{L} + \mathbf{L}^T}{2}$ 满足如下性质:

(I) $\frac{\mathbf{L} + \mathbf{L}^T}{2}$ 只有一个零特征值, 其余的特征值都是正实数, 并对特征值排序为 $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_N$;

(II) $\min_{\|\mathbf{x}\|=1} \mathbf{x}^T \frac{\mathbf{L} + \mathbf{L}^T}{2} \mathbf{x} \geq \lambda_2 \|\mathbf{x}\|^2$.

1.3 凸分析

给定集合 $\mathcal{K} \subseteq \mathbf{R}^n$, 若 $\forall \mathbf{x} \in \mathcal{K}, \forall \mathbf{y} \in \mathcal{K}$ 和 $\forall a \in (0, 1)$, 都有 $a\mathbf{x} + (1-a)\mathbf{y} \in \mathcal{K}$, 则 \mathcal{K} 为凸集. 连续可微函数 $f: \mathcal{K} \rightarrow \mathbf{R}$ 是 δ -强凸函数 ($\delta > 0$) 当且仅当 $(\mathbf{x} - \mathbf{y})^T (\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})) \geq \delta \|\mathbf{x} - \mathbf{y}\|^2, \forall \mathbf{x}, \mathbf{y} \in \mathcal{K}$. 函数 $f: \mathcal{K} \rightarrow \mathbf{R}^n$ 是 μ -Lipschitz 连续当且仅当 $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq \mu \|\mathbf{x} - \mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathcal{K}$, 其中 $\mu > 0$.

对于任意向量 $\mathbf{u} \in \mathbf{R}^n$, 定义 \mathbf{u} 在闭凸集 $\mathcal{K} \subseteq \mathbf{R}^n$ 上的投影为

$$\mathcal{P}_{\mathcal{K}}(\mathbf{u}) = \arg \min_{\mathbf{x} \in \mathcal{K}} \|\mathbf{u} - \mathbf{x}\|.$$

根据投影的定义, 可得如下不等式:

$$(\mathbf{u} - \mathcal{P}_{\mathcal{K}}(\mathbf{u}))^T (\mathbf{v} - \mathcal{P}_{\mathcal{K}}(\mathbf{u})) \leq 0, \forall \mathbf{u} \in \mathbf{R}^n, \mathbf{v} \in \mathcal{K}. \quad (1)$$

2 问题描述

考虑一组多智能体求解如下分布式优化问题:

$$\min \sum_{i=1}^N f_i(\mathbf{x}_i), \quad \text{s.t.} \quad \mathbf{x}_i \in \bigcap_{i=1}^N \Omega_i, \quad (2)$$

其中 $\mathbf{x}_i \in \mathbf{R}^n$ 是智能体享有的共同决策变量, $f_i: \mathbf{R}^n \rightarrow \mathbf{R}$ 和 $\Omega_i \subset \mathbf{R}^n, i \in \mathcal{V}$ 分别是智能体 i 的局部凸目标函数和局部闭凸集约束. 这里, 智能体只知道自己的局部目标函数和局部约束. 因此, 需要智能体分布式协作求解全局的最优决策.

为求解优化问题(2), 做出如下假设:

假设 1

a) 对于 $i \in \mathcal{V}, f_i: \Omega_i \rightarrow \mathbf{R}$ 是 m_i -强凸函数 ($m_i > 0$);

b) 对于 $i \in \mathcal{V}$, 梯度 $\nabla f_i: \Omega_i \rightarrow \mathbf{R}$ 是 M_i -Lipschitz 连续 ($M_i > 0$);

c) 存在一个内点 $\bar{\mathbf{x}}_0$ 使得 $\bar{\mathbf{x}}_0 \in \bigcap_{i=1}^N \Omega_i$.

注1 假设1保证了问题(2)解的存在性和唯一性.

假设2 智能体之间的有向通信网络满足强连通且权重平衡性.

3 主要结果

本节首先将问题(2)转化为等价形式,然后从等价形式的 Fenchel 对偶问题出发,设计一类基于奇异扰动系统的分布式连续时间算法.最后,对所提算法进行了详细的收敛性分析.

3.1 Fenchel 对偶问题

问题(2)等价于

$$\begin{aligned} \min f(\mathbf{x}) &= \sum_{i=1}^N f_i(\mathbf{x}_i), \\ \text{s.t. } \mathbf{x} \in \Omega &= \prod_{i=1}^N \Omega_i, \quad \mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_N, \end{aligned} \quad (3)$$

其中 $\mathbf{x} = \text{col}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \in \mathbf{R}^{Nn}$, $\prod_{i=1}^N \Omega_i \subset \mathbf{R}^{Nn}$ 是 $\Omega_i (i \in \mathcal{V})$ 的笛卡尔积.

对于 $i \in \mathcal{V}$, 定义 $f_i: \mathbf{R}^n \rightarrow \mathbf{R}$ 的共轭函数为 $f_i^*(\mathbf{w}_i) = \sup_{\mathbf{x}_i \in \Omega_i} (\mathbf{w}_i^T \mathbf{x}_i - f_i(\mathbf{x}_i))$. 根据文献[26]可知, 问题(3)的 Fenchel 对偶问题可表示为

$$\max -f^*(\mathbf{w}) = - \sum_{i=1}^N f_i^*(\mathbf{w}_i), \quad \text{s.t. } \sum_{i=1}^N \mathbf{w}_i = \mathbf{0}_n, \quad (4)$$

其中 $\mathbf{w} = \text{col}(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N)$.

在假设1下, 问题(3)和问题(4)之间没有对偶间隙. 换句话说, 两个问题的最优值相等. \mathbf{w}^* 是问题(4)的一个最优解当且仅当对于任意的 $i, j \in \mathcal{V}$, 有 $\nabla f_i^*(\mathbf{w}_i^*) = \nabla f_j^*(\mathbf{w}_j^*)$ 和 $\sum_{i=1}^N \mathbf{w}_i^* = \mathbf{0}_n$. 对于每个 $i \in \mathcal{V}$ 和 $\mathbf{w}_i \in \mathbf{R}^n$, $\bar{\mathbf{x}}_i(\mathbf{w}_i) \triangleq \arg \max_{\mathbf{x}_i \in \Omega_i} (\mathbf{w}_i^T \mathbf{x}_i - f_i(\mathbf{x}_i))$ 是唯一的且 $\nabla f_i^*(\mathbf{w}_i) = \bar{\mathbf{x}}_i(\mathbf{w}_i)$.

由文献[27]中的定理3.34可知, 对于任意的 $\mathbf{x}_i \in \Omega_i$, $\bar{\mathbf{x}}_i(\mathbf{w}_i) \in \Omega_i$ 满足:

$$(\mathbf{x}_i - \bar{\mathbf{x}}_i(\mathbf{w}_i))^T (\mathbf{w}_i - \nabla f_i(\bar{\mathbf{x}}_i(\mathbf{w}_i))) \leq 0.$$

等价地,

$$\bar{\mathbf{x}}_i(\mathbf{w}_i) = \mathcal{P}_{\Omega_i}(\bar{\mathbf{x}}_i(\mathbf{w}_i) + \mathbf{w}_i - \nabla f_i(\bar{\mathbf{x}}_i(\mathbf{w}_i))).$$

进而, 当 \mathbf{w}^* 是问题(4)的一个最优解时, 可得

$$\bar{\mathbf{x}}_1(\mathbf{w}_1^*) = \bar{\mathbf{x}}_2(\mathbf{w}_2^*) = \dots = \bar{\mathbf{x}}_N(\mathbf{w}_N^*) = \mathbf{x}_0^*$$

是问题(2)的最优解.

相应地, $\mathbf{x}^* = \bar{\mathbf{x}}(\mathbf{w}^*) = \mathbf{1}_N \otimes \mathbf{x}_0^*$ 是问题(3)的最优解且满足:

$$\mathbf{x}^* = \mathcal{P}_{\Omega}(\mathbf{x}^* + \mathbf{w}^* - \nabla f(\mathbf{x}^*)), \quad (5)$$

其中 $\nabla f(\mathbf{x}) = \text{col}(\nabla f(\mathbf{x}_1), \nabla f(\mathbf{x}_2), \dots, \nabla f(\mathbf{x}_N))$.

引理2^[24-25] 在假设1下, 对于每个 $i \in \mathcal{V}$, $\nabla f_i^*(\mathbf{w}_i) = \bar{\mathbf{x}}_i(\mathbf{w}_i)$ 是 $\frac{1}{m_i}$ -Lipschitz.

3.2 算法设计和收敛性分析

对于 $i \in \mathcal{V}$, 智能体 i 实施如下算法:

$$\begin{aligned} \varepsilon \dot{\mathbf{x}}_i &= \mathcal{P}_{\Omega_i}(\mathbf{x}_i + \mathbf{w}_i - \nabla f_i(\mathbf{x}_i)) - \mathbf{x}_i, \\ \dot{\mathbf{w}}_i &= - \sum_{j=1}^N a_{ij}(\mathbf{x}_i - \mathbf{x}_j), \end{aligned} \quad (6)$$

其中 ε 是有待确定的小参数. 在算法(6)的执行中, \mathbf{x}_i 用于逼近共轭函数 $f_i^*(\mathbf{w}_i)$ 的梯度 $\bar{\mathbf{x}}_i(\mathbf{w}_i)$, \mathbf{w}_i 用于搜索问题(4)的最优解. 当问题(4)的最优值获得时, \mathbf{x}_i 自然收敛到问题(2)的最优解. 这将在接下来的定理中给予证明. 为方便, 将算法(6)表示为如下等价形式:

$$\begin{aligned} \varepsilon \dot{\mathbf{x}} &= \mathcal{P}_{\Omega}(\mathbf{x} + \mathbf{w} - \nabla f(\mathbf{x})) - \mathbf{x}, \\ \dot{\mathbf{w}} &= -(\mathbf{L} \otimes \mathbf{I}_n)\mathbf{x}. \end{aligned} \quad (7)$$

定理1 在假设1和假设2下, 给定 $\text{col}(\mathbf{x}, \mathbf{w})$ 的初始值满足 $(\mathbf{1}_n^T \otimes \mathbf{I}_n)\mathbf{w}(0) = \mathbf{0}_n$ 和 $\mathbf{x}(0) \in \Omega$, 参数 ε 满足如下不等式:

$$\frac{1}{\varepsilon} > \frac{(M+m+1)^2 \|\mathbf{L}\|^2}{2\lambda_2 m^3} + \frac{M+1}{m^2} \|\mathbf{L}\|, \quad (8)$$

其中 m 表示 $m_i (i \in \mathcal{V})$ 中的最小值, M 表示 $M_i (i \in \mathcal{V})$ 中的最大值. 则由系统(7)产生的轨道 $\text{col}(\mathbf{x}(t), \mathbf{w}(t))$ 是有界的, 且 $\mathbf{x}(t)$ 收敛到 $\mathbf{x}^* = \mathbf{1}_N \otimes \mathbf{x}_0^*$, 其中 \mathbf{x}_0^* 是问题(2)的最优解.

证明 记 $\text{col}(\mathbf{x}^*, \mathbf{w}^*)$ 为系统(7)的一个平衡点, 则可得

$$\mathbf{x}^* = \mathcal{P}_{\Omega}(\mathbf{x}^* + \mathbf{w}^* - \nabla f(\mathbf{x}^*)), \quad (9)$$

$$(\mathbf{L} \otimes \mathbf{I}_n)\mathbf{x}^* = \mathbf{0}_{Nn}. \quad (10)$$

由式(10)可得 $\mathbf{x}^* = \mathbf{1}_N \otimes \mathbf{x}_0^*$.

注意到

$$(\mathbf{1}_n^T \otimes \mathbf{I}_n)\dot{\mathbf{w}} = (\mathbf{1}_n^T \otimes \mathbf{I}_n)(\mathbf{L} \otimes \mathbf{I}_n)\mathbf{x} = \mathbf{0}_n,$$

因此,

$$(\mathbf{1}_n^T \otimes \mathbf{I}_n)\mathbf{w}(t) = (\mathbf{1}_n^T \otimes \mathbf{I}_n)\mathbf{w}(0) = \mathbf{0}_n.$$

这也表明 $(\mathbf{1}_n^T \otimes \mathbf{I}_n)\mathbf{w}^* = \mathbf{0}_n$. 结合式(9)和3.1节的分析可得 \mathbf{x}^* 和 \mathbf{w}^* 分别是问题(3)和问题(4)的最优解.

设 $\mathbf{y} = \mathbf{x} - \bar{\mathbf{x}}(\mathbf{w})$, 其中 $\bar{\mathbf{x}}(\mathbf{w})$ 是 $f^*(\mathbf{w})$ 的梯度且满足:

$$\bar{\mathbf{x}}(\mathbf{w}) = \mathcal{P}_{\Omega}(\bar{\mathbf{x}}(\mathbf{w}) + \mathbf{w} - \nabla f(\bar{\mathbf{x}}(\mathbf{w}))),$$

则在变量 \mathbf{y} 和 \mathbf{w} 下, 可重写系统(7)为

$$\begin{aligned} \dot{\mathbf{y}} &= \frac{1}{\varepsilon} (\mathcal{P}_\Omega(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w}) + \mathbf{w} - \nabla f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w}))) - \\ &\quad \mathbf{y} - \bar{\mathbf{x}}(\mathbf{w})) - \frac{\partial \bar{\mathbf{x}}(\mathbf{w})}{\partial \mathbf{w}} \dot{\mathbf{w}}, \\ \dot{\mathbf{w}} &= -(\mathbf{L} \otimes \mathbf{I}_n)(\bar{\mathbf{x}}(\mathbf{w}) + \mathbf{y}). \end{aligned} \quad (11)$$

考虑如下的 Lyapunov 函数:

$$\begin{aligned} V &= f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})) - f(\bar{\mathbf{x}}(\mathbf{w})) - \mathbf{y}^\top \nabla f(\bar{\mathbf{x}}(\mathbf{w})) + \\ &\quad \frac{1}{2} \|\mathbf{y}\|^2 + f^*(\mathbf{w}) - f^*(\mathbf{w}^*), \end{aligned}$$

则 V 关于系统(11)的导数为

$$\begin{aligned} \dot{V} &= \left(\frac{\partial V}{\partial \mathbf{y}} \right)^\top \dot{\mathbf{y}} + \left(\frac{\partial V}{\partial \mathbf{w}} \right)^\top \dot{\mathbf{w}} = \\ &\quad \frac{1}{\varepsilon} (\nabla f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})) - \nabla f(\bar{\mathbf{x}}(\mathbf{w})) + \mathbf{y})^\top (\mathcal{P}_\Omega(\Delta) - \\ &\quad \mathbf{y} - \bar{\mathbf{x}}(\mathbf{w})) - \bar{\mathbf{x}}(\mathbf{w})^\top (\mathbf{L} \otimes \mathbf{I}_n) \bar{\mathbf{x}}(\mathbf{w}) - \\ &\quad \bar{\mathbf{x}}(\mathbf{w})^\top (\mathbf{L} \otimes \mathbf{I}_n) \mathbf{y} + \mathbf{y}^\top (\nabla^2 f(\bar{\mathbf{x}}(\mathbf{w})) + \mathbf{I}_{N_n}) \\ &\quad \frac{\partial \bar{\mathbf{x}}(\mathbf{w})}{\partial \mathbf{w}} (\mathbf{L} \otimes \mathbf{I}_n) (\bar{\mathbf{x}}(\mathbf{w}) + \mathbf{y}). \end{aligned}$$

其中

$$\begin{aligned} \Delta &= \mathbf{y} + \bar{\mathbf{x}}(\mathbf{w}) + \mathbf{w} - \nabla f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})), \\ \frac{\partial V}{\partial \mathbf{y}} &= \nabla f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})) - \nabla f(\bar{\mathbf{x}}(\mathbf{w})) + \mathbf{y}, \\ \frac{\partial V}{\partial \mathbf{w}} &= \left(\frac{\partial \bar{\mathbf{x}}(\mathbf{w})}{\partial \mathbf{w}} \right)^\top (\nabla f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})) - \nabla f(\bar{\mathbf{x}}(\mathbf{w})) - \end{aligned}$$

$$\nabla^2 f(\bar{\mathbf{x}}(\mathbf{w})) \mathbf{y}) + \bar{\mathbf{x}}(\mathbf{w}).$$

由投影不等式(1),可得:不等式:

$$\begin{aligned} (\mathcal{P}_\Omega(\Delta) - \Delta)^\top (\mathcal{P}_\Omega(\Delta) - \bar{\mathbf{x}}(\mathbf{w})) &\leq 0, \\ (\mathbf{w} - \nabla f(\bar{\mathbf{x}}(\mathbf{w})))^\top (\mathcal{P}_\Omega(\Delta) - \bar{\mathbf{x}}(\mathbf{w})) &\leq 0. \end{aligned}$$

由 f 的强凸性, ∇f 的 Lipschitz 连续性及引理 2 可分别得:

$$\begin{aligned} (\nabla f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})) - \nabla f(\bar{\mathbf{x}}(\mathbf{w})))^\top \mathbf{y} &\geq m \|\mathbf{y}\|^2, \\ \|\nabla^2 f(\bar{\mathbf{x}}(\mathbf{w}))\| &\leq M, \\ \left\| \frac{\partial \bar{\mathbf{x}}(\mathbf{w})}{\partial \mathbf{w}} \right\| &\leq \frac{1}{m}. \end{aligned}$$

注意到

$$\begin{aligned} \bar{\mathbf{x}}(\mathbf{w})^\top (\mathbf{L} \otimes \mathbf{I}_n) \bar{\mathbf{x}}(\mathbf{w}) &= \\ \bar{\mathbf{x}}(\mathbf{w})^\top \left(\frac{\mathbf{L}^\top + \mathbf{L}}{2} \otimes \mathbf{I}_n \right) \bar{\mathbf{x}}(\mathbf{w}) &= \mathbf{e}^\top \left(\frac{\mathbf{L}^\top + \mathbf{L}}{2} \otimes \mathbf{I}_n \right) \mathbf{e}, \end{aligned}$$

$$\text{其中 } \mathbf{e} = \frac{\mathbf{1}_N}{N} \otimes ((\mathbf{1}_n^\top \otimes \mathbf{I}_n) \bar{\mathbf{x}}(\mathbf{w})).$$

根据引理 1 中的(II)可知成立不等式

$$\bar{\mathbf{x}}(\mathbf{w})^\top (\mathbf{L} \otimes \mathbf{I}_n) \bar{\mathbf{x}}(\mathbf{w}) \geq \lambda_2 \|\mathbf{e}\|^2.$$

因此,可化简 \dot{V} 为

$$\begin{aligned} \dot{V} &\leq -\frac{m}{\varepsilon} \|\mathbf{y}\|^2 - \frac{1}{\varepsilon} \|\mathcal{P}_\Omega(\Delta) - \mathbf{y} - \bar{\mathbf{x}}(\mathbf{w})\|^2 - \\ &\quad \lambda_2 \|\mathbf{e}\|^2 - \mathbf{e}^\top (\mathbf{L} \otimes \mathbf{I}_n) \mathbf{y} + \\ &\quad \mathbf{y}^\top (\nabla^2 f(\bar{\mathbf{x}}(\mathbf{w})) + \mathbf{I}_{N_n}) \frac{\partial \bar{\mathbf{x}}(\mathbf{w})}{\partial \mathbf{w}} (\mathbf{L} \otimes \mathbf{I}_n) (\mathbf{e} + \mathbf{y}) \leq \\ &\quad -\frac{m}{\varepsilon} \|\mathbf{y}\|^2 - \frac{1}{\varepsilon} \|\mathcal{P}_\Omega(\Delta) - \mathbf{y} - \bar{\mathbf{x}}(\mathbf{w})\|^2 - \lambda_2 \|\mathbf{e}\|^2 + \\ &\quad \|\mathbf{L}\| \|\mathbf{e}\| \|\mathbf{y}\| + \frac{M+1}{m} \|\mathbf{L}\| \|\mathbf{e}\| \|\mathbf{y}\| + \\ &\quad \frac{M+1}{m} \|\mathbf{L}\| \|\mathbf{y}\|^2 \leq \\ &\quad -\left(\frac{m}{\varepsilon} - \frac{(M+m+1)^2}{2\lambda_2 m^2} \|\mathbf{L}\|^2 - \frac{M+1}{m} \|\mathbf{L}\| \right) \|\mathbf{y}\|^2 - \\ &\quad \frac{1}{\varepsilon} \|\mathcal{P}_\Omega(\Delta) - \mathbf{y} - \bar{\mathbf{x}}(\mathbf{w})\|^2 - \frac{\lambda_2}{2} \|\mathbf{e}\|^2. \end{aligned} \quad (12)$$

由不等式(8)可知, $\dot{V} \leq 0$.

注意集合 $\mathcal{W}_1 = \{\mathbf{w} \mid (\mathbf{1}_n^\top \otimes \mathbf{I}_n) \mathbf{w} = \mathbf{0}_n\}$ 关于系统(11)是不变集.由 LaSalle(拉萨尔)不变原理可知,轨道 $\text{col}(\mathbf{y}(t), \mathbf{w}(t))$ 收敛到最大不变集 \mathcal{M} , 其中 \mathcal{M} 是 $\mathcal{W}_1 \cap \{\dot{V} = 0\}$ 的子集.对于任意的点 $\text{col}(\bar{\mathbf{y}}, \bar{\mathbf{w}}) \in \mathcal{M}$, 可知

$$\begin{aligned} \bar{\mathbf{y}} &= \mathbf{0}_{N_n}, \mathcal{P}_\Omega(\bar{\mathbf{x}}(\mathbf{w}) + \bar{\mathbf{w}} - \nabla f(\bar{\mathbf{x}}(\mathbf{w}))) = \bar{\mathbf{x}}(\bar{\mathbf{w}}), \\ \mathbf{e} &= \mathbf{0}_{N_n}, (\mathbf{1}_n^\top \otimes \mathbf{I}_n) \bar{\mathbf{w}} = \mathbf{0}_n. \end{aligned} \quad (13)$$

因此,

$$\lim_{t \rightarrow \infty} \mathbf{y} = \mathbf{0}_{N_n}, \lim_{t \rightarrow \infty} \mathcal{P}_\Omega(\Delta) = \bar{\mathbf{x}}(\mathbf{w}), \lim_{t \rightarrow \infty} \mathbf{e} = \mathbf{0}_{N_n}.$$

(14)

再次利用 f 的强凸性和 ∇f 的 Lipschitz 连续性可得:

$$\frac{m}{2} \|\mathbf{y}\|^2 \leq f(\mathbf{y} + \bar{\mathbf{x}}(\mathbf{w})) - f(\bar{\mathbf{x}}(\mathbf{w})) -$$

$$\mathbf{y}^\top \nabla f(\bar{\mathbf{x}}(\mathbf{w})) \leq \frac{M}{2} \|\mathbf{y}\|^2.$$

结合 $f^*(\mathbf{w}) \geq f^*(\mathbf{w}^*)$, 有

$$0 \leq \frac{m}{2} \|\mathbf{y}\|^2 + f^*(\mathbf{w}) - f^*(\mathbf{w}^*) \leq V \leq$$

$$\frac{M}{2} \|\mathbf{y}\|^2 + f^*(\mathbf{w}) - f^*(\mathbf{w}^*). \quad (15)$$

因此, V 单调递减有下界.进而, V 的极限存在.由 $\dot{V} \leq 0$ 可得 $V(t) \leq V(0)$. 结合(15)的左边不等式可得 $f^*(\mathbf{w}) - f^*(\mathbf{w}^*) \leq V(0)$. 记集合 $\mathcal{W}_2 = \{\mathbf{w} \mid \mathbf{w} \in \mathcal{W}_1, f^*(\mathbf{w}) \leq V(0) + f^*(\mathbf{w}^*)\}$.

由文献[24]命题 2 可知,集合 \mathcal{W}_2 是紧集.因此存在收敛子列 $\mathbf{w}(t_k)$, 使得 $\lim_{k \rightarrow \infty} \mathbf{w}(t_k) = \bar{\mathbf{w}}$. 结合(13),

可知 $\text{col}(\mathbf{0}_{N_n}, \tilde{\mathbf{w}}) \in \mathcal{M}$, 其中 $\tilde{\mathbf{w}}$ 是问题(4)的一个最优解. 这也表明 $\lim_{t \rightarrow \infty} f^*(\mathbf{w}(t_k)) - f^*(\mathbf{w}^*) = 0$. 结合式(15)和 $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \mathbf{0}_{N_n}$, 有 $\lim_{k \rightarrow \infty} V(t_k) = 0$. 利用 V 的极限存在, 可得 V 的极限为零. 进而, $\lim_{t \rightarrow \infty} f^*(\mathbf{w}) = f^*(\mathbf{w}^*)$. 根据文献[28]中的定理 2.1.5, 并结合 $\mathbf{w}, \mathbf{w}^* \in \mathcal{W}_1$, $\nabla f^*(\mathbf{w}^*) = \mathbf{x}^*$ 和 $\nabla f^*(\mathbf{w}) = \bar{\mathbf{x}}(\mathbf{w})$, 可得:

$$f^*(\mathbf{w}) - f^*(\mathbf{w}^*) \geq (\nabla f^*(\mathbf{w}^*))^T (\mathbf{w} - \mathbf{w}^*) + \frac{m}{2} \|\nabla f^*(\mathbf{w}) - \nabla f^*(\mathbf{w}^*)\|^2 = \frac{m}{2} \|\bar{\mathbf{x}}(\mathbf{w}) - \mathbf{x}^*\|^2.$$

因此, 有 $\lim_{t \rightarrow \infty} \bar{\mathbf{x}}(\mathbf{w}) = \mathbf{x}^*$.

由 $\lim_{t \rightarrow \infty} \mathbf{y} = \mathbf{0}_{N_n} \rightarrow \lim_{t \rightarrow \infty} \mathbf{x} = \bar{\mathbf{x}}(\mathbf{w})$, 进而有 $\lim_{t \rightarrow \infty} \mathbf{x} = \mathbf{x}^*$.

综上可知, 定理 1 的结论成立.

4 数值仿真

本节给出一个数值例子验证所提算法的有效性. 考虑由 5 个智能体构成的通信网络如图 1 所示. 智能体的局部目标函数和局部闭凸集约束给定如下:

$$f_1 = (x - 1)^2 + e^{0.5x}, f_2 = (x - 4)^2, f_3 = x^2, f_4 = x^2 + e^{0.1x}, f_5 = x^2 + e^{-0.1x},$$

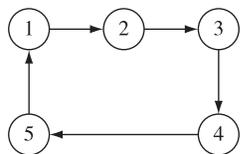


图 1 智能体之间的通信拓扑

Fig. 1 Communication topology among agents

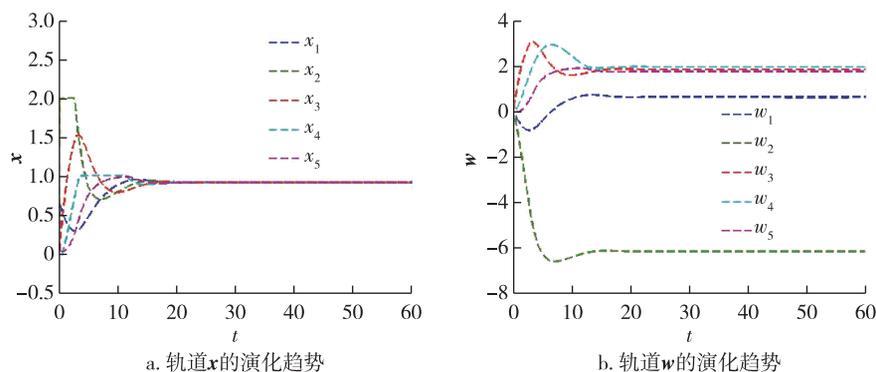


图 2 算法(6)在图 1 的通信网络下运行的结果

Fig. 2 Results generated by algorithm (6) under the communication network given in Fig.1,

(a) profile of the trajectory \mathbf{x} and (b) profile of the trajectory \mathbf{w}

$$\Omega_1 = [-1, 1], \Omega_2 = [-2, 2], \Omega_3 = [0, 2], \Omega_4 = [-5, 1], \Omega_5 = [0, 3].$$

在这个优化问题中, 可以计算出 $m = 2$ 和 $M = 2 + 0.25e^{0.5}$. 图 1 中的通信网络的 Laplacian 矩阵

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix},$$

则 \mathbf{L} 的范数和 $\frac{\mathbf{L} + \mathbf{L}^T}{2}$ 的第二特征值 λ_2 分别为

$$\|\mathbf{L}\| = \sqrt{3.618} \text{ 和 } \lambda_2 = 0.691.$$

根据表示式(8), 只要选择 $\frac{1}{\varepsilon} \geq 11.3$ 即可. 在仿真中, 选取 $\frac{1}{\varepsilon} = 12$. 图 2 给出了运行算法(6)的仿真结果. 由图 2 可以看出, 轨道 $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)^T$ 和 $\mathbf{w} = (w_1, w_2, w_3, w_4, w_5)^T$ 都是有界的. 此外, 智能体的 5 条轨道 $x_i (i = 1, 2, 3, 4, 5)$ 一致收敛到上述优化问题的最优解. 这也进一步验证了算法(6)的有效性.

5 结论与展望

为求解权重平衡有向网络下分布式凸优化问题的最优解, 本文从其等价问题的 Fenchel 对偶问题出发, 将投影方法和一致性机制相结合, 设计了基于奇异摄动的分布式连续时间算法. 然后, 结合凸分析方法和 Lyapunov 稳定性理论证明了所提算法的收敛性, 并从数值仿真进一步验证了算法的有效性. 本文仅考虑了有向网络为权重平衡的情况. 未来将进一步探讨权重非平衡网络下分布式约束优化的连续时间算法设计.

参考文献

References

- [1] Nedic A, Ozdaglar A. Distributed subgradient methods for multi-agent optimization [J]. IEEE Transactions on Automatic Control, 2009, 54 (1) : 48-61
- [2] Nedic A, Ozdaglar A, Parrilo P A. Constrained consensus and optimization in multi-agent networks [J]. IEEE Transactions on Automatic Control, 2010, 55 (4) : 922-938
- [3] Zhu M H, Martinez S. On distributed convex optimization under inequality and equality constraints [J]. IEEE Transactions on Automatic Control, 2012, 57 (1) : 151-164
- [4] Chang T H, Nedić A, Scaglione A. Distributed constrained optimization by consensus-based primal-dual perturbation method [J]. IEEE Transactions on Automatic Control, 2014, 59 (6) : 1524-1538
- [5] Yuan D M, Ho D W C, Xu S Y. Regularized primal: dual subgradient method for distributed constrained optimization [J]. IEEE Transactions on Cybernetics, 2016, 46 (9) : 2109-2118
- [6] Mateos-Núñez D, Cortés J. Distributed saddle-point subgradient algorithms with laplacian averaging [J]. IEEE Transactions on Automatic Control, 2017, 62 (6) : 2720-2735
- [7] Lee S, Zavlanos M M. Approximate projection methods for decentralized optimization with functional constraints [J]. IEEE Transactions on Automatic Control, 2018, 63 (10) : 3248-3260
- [8] You K Y, Tempo R, Xie P. Distributed algorithms for robust convex optimization via the scenario approach [J]. IEEE Transactions on Automatic Control, 2019, 64 (3) : 880-895
- [9] Lu J, Tang C Y. Zero-gradient-sum algorithms for distributed convex optimization: the continuous-time case [J]. IEEE Transactions on Automatic Control, 2012, 57 (9) : 2348-2354
- [10] Varagnolo D, Zanella F, Cenedese A, et al. Newton-raphson consensus for distributed convex optimization [J]. IEEE Transactions on Automatic Control, 2016, 61 (4) : 994-1009
- [11] Gharesifard B, Cortés J. Distributed continuous-time convex optimization on weight-balanced digraphs [J]. IEEE Transactions on Automatic Control, 2014, 59 (3) : 781-786
- [12] Kia S S, Cortés J, Martínez S. Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication [J]. Automatica, 2015, 55 : 254-264
- [13] Li Z H, Ding Z T, Sun J Y, et al. Distributed adaptive convex optimization on directed graphs via continuous-time algorithms [J]. IEEE Transactions on Automatic Control, 2018, 63 (5) : 1434-1441
- [14] Zhu Y N, Yu W W, Wen G H, et al. Continuous-time coordination algorithm for distributed convex optimization over weight-unbalanced directed networks [J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2019, 66 (7) : 1202-1206
- [15] Liu Q S, Wang J. A second-order multi-agent network for bound-constrained distributed optimization [J]. IEEE Transactions on Automatic Control, 2015, 60 (12) : 3310-3315
- [16] Zeng X L, Yi P, Hong Y G. Distributed continuous-time algorithm for constrained convex optimizations via nonsmooth analysis approach [J]. IEEE Transactions on Automatic Control, 2017, 62 (10) : 5227-5233
- [17] Liu Q S, Yang S F, Wang J. A collective neurodynamic approach to distributed constrained optimization [J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28 (8) : 1747-1758
- [18] Yang S F, Liu Q S, Wang J. A multi-agent system with a proportional-integral protocol for distributed constrained optimization [J]. IEEE Transactions on Automatic Control, 2017, 62 (7) : 3461-3467
- [19] Zhu Y N, Yu W W, Wen G H, et al. Continuous-time distributed subgradient algorithm for convex optimization with general constraints [J]. IEEE Transactions on Automatic Control, 2019, 64 (4) : 1694-1701
- [20] Zhu Y N, Yu W W, Wen G H, et al. Projected primal: dual dynamics for distributed constrained nonsmooth convex optimization [J]. IEEE Transactions on Cybernetics, 2020, 50 (4) : 1776-1782
- [21] 衣鹏, 洪奕光. 分布式合作优化及其应用 [J]. 中国科学: 数学, 2016, 46 (10) : 1547-1564
YI Peng, HONG Yiguang. Distributed cooperative optimization and its applications [J]. Scientia Sinica (Mathematics), 2016, 46 (10) : 1547-1564
- [22] 谢佩, 游科友, 洪奕光, 等. 网络化分布式凸优化算法研究进展 [J]. 控制理论与应用, 2018, 35 (7) : 918-927
XIE Pei, YOU Keyou, HONG Yiguang, et al. A survey of distributed convex optimization algorithms over networks [J]. Control Theory & Applications, 2018, 35 (7) : 918-927
- [23] Yang T, Yi X, Wu J, et al. A survey of distributed optimization [J]. Annual Reviews in Control, 2019, 47 : 278-305
- [24] Wu X Y, Lu J. Fenchel dual gradient methods for distributed convex optimization over time-varying networks [C]. // 2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2017 : 2894-2899
- [25] Wu X Y, Lu J. Fenchel dual gradient methods for distributed convex optimization over time-varying networks [J]. IEEE Transactions on Automatic Control, 2019, 64 (11) : 4629-4636
- [26] Bertsekas D P. Nonlinear programming [M]. Belmont, MA, USA: Athena Scientific, 1999
- [27] Ruszczyński A P. Nonlinear optimization [M]. Boca Raton, FL: Princeton University Press, 2006
- [28] Nesterov Y. Introductory lectures on convex optimization [M]. Boston, MA: Springer US, 2004

Continuous-time algorithm design for distributed constrained optimization over weight-balanced directed networks

ZHU Yanan¹ WEN Guanghui²

1 School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044

2 Jiangsu Provincial Key Laboratory of Networked Collective Intelligence, School of Mathematics, Southeast University, Nanjing 211189

Abstract This paper investigates a distributed convex optimization with local constraint sets over weight-balanced directed networks, where the global objective function is described as a sum of some agents' local objective functions. To solve this problem in a distributed way, the problem is transformed into a Fenchel dual problem by introducing local conjugate functions. Then, for the Fenchel dual problem, a distributed continuous-time algorithm is proposed based on the singular perturbation system. When the local objective functions are strongly convex and their gradients are Lipschitz continuous, it is shown that the primal and dual optimality can be both achieved by using the tools from convex analysis and Lyapunov stability. Finally, simulation results are given to illustrate the effectiveness of the proposed algorithm.

Key words distributed constrained optimization; weight-balanced directed networks; Fenchel dual problem; singular perturbation system