第 40 卷第 10 期
2023 年 10 月

控 制 理 论 与 应 用
Control Theory & Applications

Vol. 40 No. 10
Oct. 2023

基于Bandit反馈的自适应量化分布式在线镜像下降算法

谢俊如, 高文华†, 谢奕彬

(华南理工大学 数学学院, 广东 广州 510640)

摘要: 多智能体系统的在线分布式优化常用于处理动态环境下的优化问题, 节点间需要实时传输数据流. 在很多情况下, 各节点无法获取个体目标函数的全部信息(包括梯度信息), 并且节点间信息传输存在一定的通信约束. 考虑到非欧投影意义下的镜像下降算法在处理高维数据和大规模在线学习上的优势, 本文使用个体目标函数在两点处的函数值信息对缺失的梯度信息进行估计, 并且根据镜像下降算法的性质设计自适应量化器, 提出基于Bandit反馈的自适应量化分布式在线镜像下降算法. 然后分析了量化误差界和Regret界的关系, 适当选择参数可得所提算法的Regret界为$\mathcal{O}(\sqrt{T})$. 最后, 通过数值仿真验证了算法和理论结果的有效性.

关键词: 镜像下降算法; 多智能体系统; 优化; 量化; Bandit反馈

引用格式: 谢俊如, 高文华, 谢奕彬. 基于Bandit反馈的自适应量化分布式在线镜像下降算法. 控制理论与应用, 2023, 40(10): 1774 – 1782

DOI: 10.7641/CTA.2022.20152

Adaptive quantized online distributed mirror descent algorithm with Bandit feedback

XIE Jun-ru, GAO Wen-hua†, XIE Yi-bin

(School of Mathematics, South China University of Technology, Guangzhou Guangdong 510640, China)

Abstract: Online distributed optimization of multi-agent systems is often used to deal with the optimization problems in dynamic environments, and real-time data streams need to be transmitted between nodes. In many cases, each node cannot obtain all the information of the individual objective function (including gradient information), and there are communication constraints in the transmission of information between nodes. In this paper, considering the advantages of the mirror descent algorithm in the sense of non-Euclidean projection in processing high-dimensional data and large-scale online learning, the function value information of the individual objective function at two points is used to estimate the missing gradient information, and an adaptive quantizer is designed according to the property of mirror descent algorithm, and an adaptive quantized distributed online mirror descent algorithm based on the bandit feedback is proposed. Then the relationship between the quantization error bound and the regret bound is analyzed. The regret bound of the proposed algorithm can be obtained as $\mathcal{O}(\sqrt{T})$ when the parameters are chosen appropriately. Finally, the effectiveness of the algorithm and theoretical results is verified by numerical simulations.

Key words: mirror descent algorithm; multi-agent systems; optimization; quantization; Bandit feedback

Citation: XIE Junru, GAO Wenhua, XIE Yibin. Adaptive quantized online distributed mirror descent algorithm with Bandit feedback. *Control Theory & Applications*, 2023, 40(10): 1774 – 1782

1 引言

多智能体网络系统被广泛应用在当今社会的诸多领域, 如交通网络[1]、机器学习[2]以及智能电网[3]. 多智能体网络系统是由一定数量的智能体以及智能体之间的交互关系集合所构成, 其中网络系统中的每个智能体具有一定的通信能力、采集数据能力、分析数据和执行决策的能力, 且具有一定的自主性. 多智能体系统的分布式优化算法通过智能体之间的信息交互来求解整个系统的最优化问题. "去中心化"的设定可减少节点的通信负担, 并且可使系统具有更好的鲁棒性和隐私性. 上述优点使得分布式优化得到广泛关注和研究[4–5].

收稿日期: 2022−03−03; 录用日期: 2022−12−20.

†通信作者. E-mail: whgao@scut.edu.cn.

本文责任编委: 施阳.

国家自然科学基金项目(62273157), 广州市科技计划项目(202002030158)资助.

Supported by the National Natural Science Foundation of China (62273157) and the Guangzhou Science and Technology Planning Project (202002030158).

传统的梯度投影算法对于某些目标函数和约束集合难以有效计算投影. 针对这种问题, Beck和Teboulle在文献[6]中根据约束集合的特性选择合适的Bregman散度替代欧氏距离, 从而提高算法的收敛速率. Xi等人[7]基于上述思想, 将文献[5]中的分布式次梯度算法推广得到非欧投影意义下的分布式镜像下降算法. 镜像下降算法在处理高维数据、有效处理投影和大规模在线学习等方面具有优势.

考虑到实际生活中通信带宽的限制, 智能体信息的传输需要通过合适的量化器对数据进行量化处理. 而这些处理不可避免地带来量化误差, 从而很大程度上影响算法的收敛性. 最近, 已有许多文章研究带量化的优化算法的收敛速率以及量化误差对算法收敛的影响. Doan等人[8]提出了带有自适应均匀量化的分布式次梯度算法, 证明了目标函数为凸函数的情况下, 算法达到$\mathcal{O}(\ln(k)/\sqrt{k})$的收敛速率. Liu等人[9]考虑通信网络带有时滞的情况, 将文献[8]中的算法推广, 得到带有自适应量化的次梯度分布式镜像下降算法.

以上算法均为离线优化算法. 在传统的离线优化中, 通常假定在静态环境下对所有数据收集之后再进行处理, 因此离线优化面临高额的通信代价和计算代价. 而在实际生活中, 多智能体系统的网络通常面临动态环境并且需要实时处理数据流, 例如: 运动目标跟踪定位[10]、序列预测[11]以及推荐系统[12]等. 在线优化模型更适用于上述情况.

分布式在线优化已有许多结果, 比如, Yan等人[13]将文献[4]的分布式次梯度算法推广到在线学习, 提出了在线分布式次梯度算法, 并分析了网络拓扑结构和算法所具有的隐私保护特性之间的关系. Shahrampour和Jadbabaie[14]考虑一种基于镜像下降的分布式优化算法, 并对该算法的收敛结果进行了分析.

在完全信息反馈设定下, 每个节点可以知道个体目标函数的全部信息, 包括梯度信息. 但许多情况下梯度信息无法获取或者梯度难以计算, 每个节点只可能获取目标函数的部分信息, 这种情况称为Bandit反馈. 现给出有关Bandit反馈算法的研究结果. 文献[15]提出了集中式的Bandit算法. Yuan等人[16]研究了复合优化问题在完全信息反馈和Bandit信息反馈下的情况. Li等人[17]基于文献[14]提出了Bandit反馈下的分布式在线镜像下降算法.

多智能体网络中节点的通信能力是有限的, 并且节点使用量化器对数据处理后再进行信息交流可以降低通信的代价. 欧氏投影算法中所使用的量化分析方法并不能直接使用在基于非欧投影的镜像下降算法中. 为此, 受文献[9]中量化机制的启发, 本文利用了镜像下降算法的特性来设计自适应量化器, 并提出基于Bandit反馈的自适应量化分布式在线镜像下降算法 (adaptive quantized online distributed mirror descent with bandit feedback, AQODMD-B).

本文主要贡献如下:

1) 考虑基于Bandit反馈设定下的在线优化. 在Bandit反馈设定下, 节点无法获取目标函数的梯度信息. 与文献[14]中基于完全信息反馈的算法不同, AQODMD-B涉及了状态值附近邻域内函数值的计算而非梯度值的计算, 避免了计算梯度值所花费的高额计算开销.

2) 镜像下降算法利用Bregman散度替代欧氏距离, 可以根据目标函数和约束集合的特点来选择合适的Bregman散度, 从而提高算法的收敛速率. 相较于文献[14, 17]中所考虑的时不变多智能体网络, 本文所考虑的时变多智能体网络拓宽了应用范围.

3) 本文提出一种多智能体网络中存在通信约束的分布式镜像下降算法. 利用了镜像下降算法的特性来设计自适应量化器, 这种量化方法使得节点在有限比特数的量化通信下实现算法的收敛, 并且与未考虑量化通信的文献[17]中的算法有着相近的性能表现.

下面给出本文使用到的部分符号.

$\|\cdot\|$和$\|\cdot\|_\infty$分别表示向量欧氏范数和无穷范数; $[x]_i$表示$x\in\mathbb{R}^n$的第i个元素; $x\preccurlyeq z$表示对所有分量i都有$[x]_i\leqslant[z]_i$; $[A]_{ij}$表示矩阵A的第i行, 第j列元素; 给定正整数$T\geqslant 1$, $[T]$表示集合$\{1,\cdots,T\}$; 数学期望算子记为E$[\cdot]$; 记$\mathbf{1}$为n维全1列向量; $\mathrm{Proj}_{\mathcal{X}}\{x\}$表示向量$x$在集合$\mathcal{X}$上的投影; ∇f表示函数f的(次)梯度.

2 问题描述与预备知识

本文考虑由N个节点组成的时变多智能体网络$\mathcal{G}_t=(\mathcal{V},\mathcal{E}_t)$, 其中顶点集$\mathcal{V}=\{1,\cdots,N\}$, 边集$\mathcal{E}_t=\{(i,j)|$在$t$时刻, 节点$i$与节点$j$相连$\}$. 第$i$个节点的局部目标函数$f_{i,t}:\mathcal{X}\rightarrow\mathbb{R}$为非光滑凸函数且$\mathcal{X}\subseteq\mathbb{R}^n$为非空紧集. 基于在线优化设定, 每个节点$i$在$t$时刻做出决策$x_{i,t}$之后, 可以接收到环境反馈的信息. 根据反馈信息的类型可分为完全信息反馈和Bandit反馈: 1) 在完全信息反馈的框架下, 每个节点可以获得个体目标函数$f_{i,t}$的梯度信息; 2) 在Bandit反馈的框架下, 节点只可以获得$\mathcal{X}$中有限点的函数值信息, 并不能使用函数的梯度信息. 本文主要研究基于Bandit反馈的在线优化算法.

在时刻t, 全局目标函数$f_t(x)$可以用个体目标函数的求和来表示, 即$f_t(x)=\sum\limits_{i=1}^{N}f_{i,t}(x)$. 分布式在线优化考虑如下模型:

$$\min_{x\in\mathcal{X}}\sum_{t=1}^{T}\sum_{i=1}^{N}f_{i,t}(x),\tag{1}$$

其目的是让每个节点$i\in\mathcal{V}$产生的决策序列$\{x_{i,t}\}_{t=1}^{T}\subseteq\mathcal{X}$最小化平均Regret[16]. 记节点$i\in\mathcal{V}$的

平均Regret为$\mathrm{Reg}_i(T)$, 其定义如下:

$$\mathrm{Reg}_i(T) := \sum_{t=1}^{T}\sum_{j=1}^{N} f_{j,t}(x_{i,t}) - \sum_{t=1}^{T}\sum_{j=1}^{N} f_{j,t}(x^*), \quad (2)$$

其中$x^* = \arg\min\limits_{x\in\mathcal{X}} \sum_{t=1}^{T}\sum_{i=1}^{N} f_{i,t}(x)$. 如果在线算法的Regret界满足$\lim\limits_{t\to\infty} \dfrac{\mathrm{Reg}_i(T)}{T} = 0$, 则说明该算法是有效的.

2.1 图论

由N个节点构成的时变网络$\mathcal{G}_t$, 记节点i在时刻t的邻居节点集为$\mathcal{N}_i(t) = \{j\in\mathcal{V} | (i,j)\in\mathcal{E}_t\}$. 每个节点$i$对来自于$j\in\mathcal{N}_i(t)$的信息给予权重$[P(t)]_{ij} > 0$, 节点$i$和$j$在时刻$t$不相互连接当且仅当$[P(t)]_{ij} = 0$. 对于任意时刻$t$和$s$, 其中$t \geqslant s \geqslant 1$, 记状态转移矩阵为$P(t,s) = P(t)P(t-1)\cdots P(s+1)P(s)$. 同时, 定义$P(t,t) = P(t)$和$P(t,t+1) = I_n$, 其中$I_n$表示$n$维单位矩阵. 现给出以下常用于分布式网络的假设.

假设 1[4] 1) 时变网络的权重矩阵$P(t)$是双随机的, 即有$\sum_{i=1}^{N}[P(t)]_{ij} = 1$和$\sum_{j=1}^{N}[P(t)]_{ij} = 1$.

2) 时变网络$(\mathcal{V},\mathcal{E}_t)$是$B$–强连通的. 即存在正整数$B$使得对任意正整数$c$, 图$(\mathcal{V}, \bigcup_{t=cB+1}^{(c+1)B}\mathcal{E}_t)$是强连通的.

3) 存在常数$\theta\in(0,1)$, 对于$(i,j)\in\mathcal{E}_t$, $[P(t)]_{ij} \geqslant \theta$; 对于任意$i\in\mathcal{V}$, 有$[P(t)]_{ii} \geqslant \theta$.

引理 1[4] 假设1成立时, 对任意$i,j\in\mathcal{V}$, 有

$$|[P(t,s)]_{ij} - \frac{1}{N}| \leqslant \omega\gamma^{t-s},$$

其中: $t \geqslant s \geqslant 1$, $\omega = (1-\dfrac{\theta}{4N^2})^{-2}$, $\gamma = (1-\dfrac{\theta}{4N^2})^{\frac{1}{B}}$.

2.2 Bregman散度

镜像下降算法本质是使用Bregman散度[6]来替代欧氏距离. 设$h: \mathcal{X}\to\mathbb{R}$是一个$\sigma$–强凸并可微的函数, 即对任意$x,y\in\mathcal{X}$, 存在常数$\sigma\geqslant 0$使得

$$h(x) \geqslant h(y) + \langle\nabla h(y), x-y\rangle + \frac{\sigma}{2}\|x-y\|^2,$$

由h定义的Bregman散度$D_h(x,y)$如下所示:

$$D_h(x,y) := h(x) - h(y) - \langle\nabla h(y), x-y\rangle,$$

特别地, 令$h(x) = \|x\|^2$, 可得Bregman散度$D_h(x,y)$为欧氏距离$\|x-y\|^2$.

下面给出Bregman散度的常用假设和引理.

假设 2[14] 对于任意$x, y_i\in\mathcal{X}$和常数$\phi_i \geqslant 0$, 其中$\sum_{i=1}^{N}\phi_i = 1$, Bregman散度满足下式:

$$D_h(x, \sum_{i=1}^{N}\phi_i y_i) \leqslant \sum_{i=1}^{N}\phi_i D_h(x, y_i).$$

假设 3[14] 对于任意$x,y\in\mathcal{X}$, 存在$L_h\geqslant 0$, 使得$\|\nabla h(x) - \nabla h(y)\| \leqslant L_h\|x-y\|$.

引理 2[6] 设$D_h(\cdot,\cdot)$是由h定义的Bregman散度, 则对于任意$b,c,d\in\mathcal{X}$, 下式成立:

$$\langle b, \hat{x}-d\rangle \leqslant D_h(d,c) - D_h(d,\hat{x}) - D_h(\hat{x},c),$$

其中$\hat{x} = \arg\min\limits_{x\in\mathcal{X}}\{\langle b,x\rangle + D_h(x,c)\}$.

基于欧氏投影的量化算法的分析方式无法直接应用于基于非欧投影的镜像下降算法. 为此本文通过利用镜像下降算法的性质来设计自适应量化器, 接下来的引理将会在进行量化分析时用到.

引理 3[18] 对于任意的$\eta > 0$, $g_1, g_2\in\mathbb{R}^n$和$y\in\mathcal{X}$, 若x_1和x_2满足如下关系式:

$$x_1 = \arg\min_{x\in\mathcal{X}}\{\eta\langle g_1, x\rangle + D_h(x,y)\},$$
$$x_2 = \arg\min_{x\in\mathcal{X}}\{\eta\langle g_2, x\rangle + D_h(x,y)\},$$

则$\|x_2 - x_1\| \leqslant \dfrac{\eta}{\sigma}\|g_2 - g_1\|$.

2.3 Bandit反馈

目前大多凸优化算法都利用目标函数的梯度信息, 很多实际问题中无法获得函数的梯度信息, 例如: 投资组合管理[19]和在线广告投放[20]. 以投资组合管理问题为例, 投资管理者在t时刻做出决策x_t后, 股票市场给予此次决策的反馈$f_t(x_t)$. 投资管理者可能只知道此次决策的反馈$f_t(x_t)$或者是目标函数$f_t(\cdot)$在x_t某些邻近点的值. 然而, 投资管理者由于无法获得目标函数$f_t(\cdot)$在决策x_t邻域内所有点的函数值, 所以无法获取对应的梯度信息$\nabla f_t(x_t)$. 节点无法获取函数梯度信息, 只能获取部分点函数值信息的情形被称为Bandit反馈.

对于Bandit反馈下的集合$\mathcal{X}$和目标函数$f_{i,t}$, 给出如下常用假设.

假设 4[15] 1) 集合$\mathcal{X}$是非空紧集, 而且存在两个常数a和A, 其中$A > a > 0$, 使得$a\mathcal{B}\subset\mathcal{X}\subset A\mathcal{B}$.

2) 对于任意$i\in\mathcal{V}$, $t\in[T]$和$x,y\in\mathcal{X}$, 存在正数L使得

$$|f_{i,t}(x) - f_{i,t}(y)| \leqslant L\|x-y\|.$$

注 1 假设4说明存在常数$R\geqslant 0$使得任意$x,y\in\mathcal{X}$满足$\sup\limits_{x,y\in\mathcal{X}} D_h(x,y) \leqslant R^2$. 同时根据

$$\frac{\sigma}{2}\|x-y\|^2 \leqslant D_h(x,y) \leqslant R^2,$$

说明可行集$\mathcal{X}$内任意两个向量x, y满足

$$\sup_{x,y\in\mathcal{X}}\|x-y\| \leqslant \sqrt{\frac{2R^2}{\sigma}} := \kappa. \quad (3)$$

对于函数$f: \mathcal{X}\to\mathbb{R}$, 定义$f(x)$的光滑化函数为$\tilde{f}(x) := \mathrm{E}_{v\sim\mathcal{U}(\mathcal{B})}[f(x+\delta v)]$, 其中: $x\in(1-\rho)\mathcal{X}$, $\mathcal{U}(\mathcal{B})$代表了$\mathbb{R}^n$上单位球体$\mathcal{B}$的均匀分布, $\rho\in[0,1)$

为压缩系数, $\delta \in [0, a\rho]$为光滑参数. 两点采样的梯度估计值$\tilde{g}$, 其定义如下[17]:

$$\tilde{g}(x) = \frac{n}{2\delta}[f(x+\delta s) - f(x-\delta s)]s,\ s \sim \mathcal{U}(\mathcal{S}),$$

其中: $x \in (1-\rho)\mathcal{X}, \mathcal{U}(\mathcal{S})$为单位球面$\mathcal{S}$的均匀分布.

注 2 基于假设4的1)中$a\mathcal{B} \subset \mathcal{X}$. 由文献[16]的引理5: 对于任意$x \in (1-\rho)\mathcal{X}$和任意单位向量$s$, 满足$x+\delta s \in \mathcal{X}$, 其中$\delta \in [0, a\rho]$. 由该引理可知对于任意$x \in (1-\rho)\mathcal{X}$和任意单位向量$s$, $x+\delta s$在f的定义域$\mathcal{X}$内. 再由单位向量s的任意性可知$x-\delta s \in \mathcal{X}$. 因此定义$\tilde{g}(x) = \frac{n}{2\delta}[f(x+\delta s) - f(x-\delta s)]s$是合理的.

引理 4[15] 1) 光滑化函数$\tilde{f}$在可行集$(1-\rho)\mathcal{X}$上可微. 任给$x \in (1-\rho)\mathcal{X}$, 有

$$\nabla \tilde{f}(x) = \mathrm{E}_{s\sim\mathcal{U}(\mathcal{S})}[\tilde{g}(x)].$$

2) 当f为$\mathcal{X}$上的凸函数时, 光滑化函数$\tilde{f}$为$(1-\rho)\mathcal{X}$上的凸函数, 且对于任给$x \in (1-\rho)\mathcal{X}$, 有$f(x) \leqslant \tilde{f}(x)$.

3) 当f在$\mathcal{X}$上是Lipschitz连续时, 那么$\tilde{f}$在$(1-\rho)\mathcal{X}$上Lipschitz连续, 且任给$x \in (1-\rho)\mathcal{X}$, $|\tilde{f}(x) - f(x)| \leqslant \delta L$. 此外, $\|\tilde{g}(x)\| \leqslant nL$.

2.4 量化

由量化引起的量化误差有可能使得算法不收敛, 因此设计一个合适的量化器至关重要. 本文使用自适应均匀量化器对节点间传输的数据进行量化处理. 在时刻t, 对于向量$x \in \mathbb{R}^n$, 若给定量化水平参数K, 量化中间值向量$z_t \in \mathbb{R}^n$和量化区间尺寸向量$d(t) \in \mathbb{R}^n$, 定义$\hat{Q}(z_t, d(t), x) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$为量化函数, 其中$\hat{Q}(z_t, d(t), x)$的第$i$个分量$[\hat{Q}(z_t, d(t), x)]_i$定义如下:

$$[\hat{Q}(z_t, d(t), x)]_i = \begin{cases} [z_t]_i - [d(t)]_i, \\ \quad [x]_i - [z_t]_i \in (-\infty, -[d(t)]_i), \\ [z_t]_i + \dfrac{(2j-K)[d(t)]_i}{K}, \\ \quad [x]_i - [z_t]_i + [d(t)]_i \in [\dfrac{2j[d(t)]_i}{K}, \dfrac{2(j+1)[d(t)]_i}{K}), \\ [z_t]_i + [d(t)]_i, \\ \quad [x]_i - [z_t]_i \in [[d(t)]_i, +\infty), \end{cases}$$

其中$j = 0, \cdots, K-1$.

当$z_t - d(t) \preccurlyeq x \preccurlyeq z_t + d(t)$, 根据量化函数的定义, 易得量化误差$e(t) = x - \hat{Q}(z_t, d(t), x)$满足

$$\|e(t)\| \leqslant \sqrt{n}\|e(t)\|_\infty \leqslant \sqrt{n}\frac{2\|d(t)\|_\infty}{K}. \quad (4)$$

3 算法设计和分析

本节提出基于Bandit反馈的自适应量化分布式在线镜像下降算法, 即算法1.

表 1 算法1: AQODMD-B

Table 1 Algorithm 1: AQODMD-B

1 输入: 每个点$i \in \mathcal{V}$, 初始化$x_{i,1} = z_{i,1}$, 双随机矩阵$P(t)$; 步长$\alpha(t)$, 量化参数序列$\beta(t)$, 光滑参数δ和压缩系数ρ.

2 **for** $t = 1, \cdots, T$ **do**

3 **if** $t = 1$, $d(1) = \dfrac{nL\alpha(1)\beta(1)}{\sigma}\mathbf{1}$;

otherwise $d(t) = \dfrac{nL\alpha(t-1)\beta(t-1)}{\sigma}\mathbf{1}$;

4 **for** $i = 1$ to N **do**

5 发送量化状态值$\hat{Q}(z_{i,t},\ d(t),\ x_{i,t})$到节点$j \in \mathcal{N}_i(t)$; 从节点$j \in \mathcal{N}_i(t)$处接收量化状态值$\hat{Q}(z_{j,t}, d(t), x_{j,t})$;

6 $y_{i,t} = \sum\limits_{j=1}^{m}[P(t)]_{ij}\hat{Q}(z_{j,t}, d(t), x_{j,t})$;

7 $\tilde{y}_{i,t} = \operatorname*{Proj}\limits_{(1-\rho)\mathcal{X}}\{y_{i,t}\}$;

8 接收$f_{i,t}(\tilde{y}_{i,t} + \delta s_{i,t})$和$f_{i,t}(\tilde{y}_{i,t} - \delta s_{i,t})$, 计算$\tilde{g}_{i,t} = \dfrac{n}{2\delta}[f_{i,t}(\tilde{y}_{i,t} + \delta s_{i,t}) - f_{i,t}(\tilde{y}_{i,t} - \delta s_{i,t})]s_{i,t}$, 其中$s_{i,t} \sim \mathcal{U}(\mathcal{S})$;

9 $z_{i,t+1} = \arg\min\limits_{x\in(1-\rho)\mathcal{X}}\{\alpha(t)(1-\beta(t)) \times \langle\tilde{g}_{i,t}, x\rangle + D_h(x, \tilde{y}_{i,t})\}$,

$x_{i,t+1} = \arg\min\limits_{x\in(1-\rho)\mathcal{X}}\{\alpha(t)\langle\tilde{g}_{i,t}, x\rangle + D_h(x, \tilde{y}_{i,t})\}$.

10 **end for**

11 **end for**

为了分析算法1中各节点状态值$x_{i,t}$与它们平均值$\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{i,t}$的关系, 先分析量化误差$e_i(t)$和Bregman投影误差$\varepsilon_i(t)$, 即引理5–6.

引理 5 若假设4成立, 则量化误差$e_i(t)$满足

$$\|e_i(t)\| \leqslant \frac{2n^{\frac{3}{2}}L}{\sigma K}\alpha(t-1)\beta(t-1).$$

证 对于算法1步骤9产生的向量$x_{i,t}, z_{i,t} \in \mathbb{R}^n$, 对于任意分量$j \in \mathcal{V}$, 根据引理3可得

$$\|x_{i,t} - z_{i,t}\| \leqslant \frac{\|\tilde{g}_{i,t}\|\alpha(t-1)\beta(t-1)}{\sigma} \leqslant \frac{nL\alpha(t-1)\beta(t-1)}{\sigma} = [d(t)]_j. \quad (5)$$

于是, 根据$|[x_{i,t} - z_{i,t}]_j| \leqslant \|x_{i,t} - z_{i,t}\| \leqslant [d(t)]_j$, 可得$z_{i,t} - d(t) \preccurlyeq x_{i,t} \preccurlyeq z_{i,t} + d(t)$. 故$z_{i,t}$和$d(t)$使得状态值$x_{i,t} \in [z_{i,t} - d(t), z_{i,t} + d(t)]$.

因此, 由$\|d(t)\|_\infty = \dfrac{nL\alpha(t-1)\beta(t-1)}{\sigma}$, 可得

$$\|e_i(t)\| = \|\hat{Q}(z_{i,t}, d(t), x_{i,t}) - x_{i,t}\| \leqslant$$

$$\sqrt{n}\frac{2\|d(t)\|_\infty}{K}=\frac{2n^{\frac{3}{2}}L}{\sigma K}\alpha(t-1)\beta(t-1).$$

证毕.

从引理5的分析过程中可以看出, 本文通过利用引理3中Bregman散度的性质得以对量化误差$e_i(t)$进行分析. 记$E(t):=\frac{2n^{\frac{3}{2}}L}{\sigma K}\alpha(t-1)\beta(t-1)$为量化误差上界, 当$K$越大时, 量化误差上界越小.

引理 6 在假设1和假设4成立的条件下, Bregman投影误差$\varepsilon_i(t):=x_{i,t+1}-\tilde{y}_{i,t}$满足

$$\|\varepsilon_i(t)\|\leqslant\frac{nL\alpha(t)}{\sigma}.$$

证 对于算法1的第9步, 根据引理2, 令$\hat{x}=x_{i,t+1}$, $b=\alpha(t)\tilde{g}_{i,t}$, $c=d=\tilde{y}_{i,t}$, 可得

$$\begin{aligned}&\langle\alpha(t)\tilde{g}_{i,t},\tilde{y}_{i,t}-x_{i,t+1}\rangle\geqslant\\&D_h(\tilde{y}_{i,t},x_{i,t+1})+D_h(x_{i,t+1},\tilde{y}_{i,t})\geqslant\\&\sigma\|x_{i,t+1}-\tilde{y}_{i,t}\|^2,\end{aligned}\tag{6}$$

其中式(6)第2个不等式使用了函数h的σ–强凸.

根据柯西不等式, 有

$$\begin{aligned}&\langle\alpha(t)\tilde{g}_{i,t},\tilde{y}_{i,t}-x_{i,t+1}\rangle\leqslant\\&\alpha(t)\|\tilde{g}_{i,t}\|\|x_{i,t+1}-\tilde{y}_{i,t}\|.\end{aligned}\tag{7}$$

结合式(6)–(7)和引理4中的3), 可得

$$\sigma\|x_{i,t+1}-\tilde{y}_{i,t}\|^2\leqslant nL\alpha(t)\|x_{i,t+1}-\tilde{y}_{i,t}\|,\tag{8}$$

因此Bregman投影误差满足$\|\varepsilon_i(t)\|\leqslant\frac{nL\alpha(t)}{\sigma}$.

证毕.

定理 1 在假设1和假设4成立的条件下, 对于任意$i\in\mathcal{V}$, $t\in[T]$, 由算法1产生的序列$\{x_{i,t}\}$满足下式:

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{i=1}^{N}\|x_{i,t}-\bar{x}_t\|\leqslant\\&\tau+(N+\frac{N\omega}{\gamma(1-\gamma)})\sum_{t=1}^{T}(\frac{nL\alpha(t)}{\sigma}+3E(t)),\end{aligned}$$

其中$\tau=\frac{N\omega}{1-\gamma}\sum_{i=1}^{N}\|x_{i,1}\|$.

证 记投影误差$p_i(t):=\tilde{y}_{i,t}-y_{i,t}$, 则

$$\begin{aligned}&\|p_i(t)\|=\|\tilde{y}_{i,t}-y_{i,t}\|=\\&\|\tilde{y}_{i,t}-\sum_{j=1}^{N}[P(t)]_{ij}x_{j,t}-\sum_{j=1}^{N}[P(t)]_{ij}e_j(t)\|\leqslant\\&\|y_{i,t}-\sum_{j=1}^{N}[P(t)]_{ij}x_{j,t}\|+E(t)=\\&\|\sum_{j=1}^{N}[P(t)]_{ij}\hat{Q}(z_{j,t},d(t),x_{j,t})-\\&\sum_{j=1}^{N}[P(t)]_{ij}x_{j,t}\|+E(t)\leqslant 2E(t).\end{aligned}\tag{9}$$

同时, 定义$c_i(t):=\sum_{j=1}^{N}[P(t)]_{ij}(e_j(t)+p_i(t))$, 根据$x_{i,t+1}=\tilde{y}_{i,t}+\varepsilon_i(t)$和$\tilde{y}_{i,t}=y_{i,t}+p_i(t)$, 有

$$\begin{aligned}&x_{i,t+1}=\tilde{y}_{i,t}+\varepsilon_i(t)=\\&\sum_{j=1}^{N}[P(t)]_{ij}(x_{j,t}+e_j(t))+p_i(t)+\varepsilon_i(t)=\\&\sum_{j=1}^{N}[P(t)]_{ij}x_{j,t}+\sum_{j=1}^{N}[P(t)]_{ij}e_j(t)+p_i(t)+\varepsilon_i(t)=\\&\sum_{j=1}^{N}[P(t)]_{ij}x_{j,t}+c_i(t)+\varepsilon_i(t)=\\&\sum_{j=1}^{N}[P(t)]_{ij}(\sum_{k=1}^{N}[P(t-1)]_{jk}x_{k,t-1}+\\&c_j(t-1)+\varepsilon_j(t-1))+c_i(t)+\varepsilon_i(t)=\\&\sum_{j=1}^{N}[P(t)]_{ij}\sum_{k=1}^{N}[P(t-1)]_{jk}x_{k,t-1}+\\&\sum_{j=1}^{N}[P(t)]_{ij}(c_j(t-1)+\varepsilon_j(t-1))+c_i(t)+\varepsilon_i(t).\end{aligned}\tag{10}$$

$x_{i,t+1}$通过式(10)迭代可得下式:

$$\begin{aligned}x_{i,t+1}=&\sum_{s=1}^{t}\sum_{j=1}^{N}[P(t,s+1)]_{ij}(c_j(s)+\varepsilon_j(s))+\\&\sum_{j=1}^{N}[P(t,1)]_{ij}x_{j,1}.\end{aligned}\tag{11}$$

根据$\bar{x}_{t+1}$的定义, 可得

$$\bar{x}_{t+1}=\frac{1}{N}\sum_{s=1}^{t}\sum_{j=1}^{N}(c_j(s)+\varepsilon_j(s))+\frac{1}{N}\sum_{j=1}^{N}x_{j,1}.\tag{12}$$

结合式(11)–(12), 有

$$\begin{aligned}&\|x_{i,t+1}-\bar{x}_{t+1}\|\leqslant\\&\sum_{j=1}^{N}|[P(t,1)]_{ij}-\frac{1}{N}|\|x_{j,1}\|+\\&\sum_{s=1}^{t}\sum_{j=1}^{N}|[P(t,s+1)]_{ij}-\frac{1}{N}|(\|c_j(s)+\varepsilon_j(s)\|)\leqslant\\&\omega\gamma^{t-1}\sum_{j=1}^{N}\|x_{j,1}\|+\|c_i(t)+\varepsilon_i(t)\|+\\&\sum_{s=1}^{t-1}\omega\gamma^{t-s-1}\sum_{j=1}^{N}(\|c_j(s)+\varepsilon_j(s)\|)\leqslant\qquad(13a)\\&\omega\gamma^{t-1}\sum_{j=1}^{N}\|x_{j,1}\|+3E(t)+\frac{nL\alpha(t)}{\sigma}+\\&\sum_{s=1}^{t}\omega\gamma^{t-s-1}(3E(s)+\frac{nL\alpha(s)}{\sigma}),\qquad(13b)\end{aligned}$$

其中: 式(13a)使用了引理1, 式(13b)使用了引理6和下式:

$$c_i(t)\leqslant 3\sum_{j=1}^{N}[P(t)]_{ij}E(t)=3E(t),\tag{14}$$

对式(13b)关于$i \in \mathcal{V}, t \in [T]$求和, 可得

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{i=1}^{N}\|x_{i,t+1}-\bar{x}_{t+1}\| \leqslant\\ &\omega\sum_{t=1}^{T}\gamma^{t-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\|x_{j,1}\|+\\ &\sum_{t=1}^{T}\sum_{i=1}^{N}(3E(t)+\frac{nL\alpha(t)}{\sigma})+\\ &\frac{\omega}{\gamma}\sum_{i=1}^{N}\sum_{t=0}^{T-s}\gamma^{t}\sum_{s=1}^{T}(3E(s)+\frac{nL\alpha(s)}{\sigma}) \leqslant\\ &\frac{\omega}{1-\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}\|x_{j,1}\|+\sum_{t=1}^{T}\sum_{i=1}^{N}(3E(t)+\frac{nL\alpha(t)}{\sigma})+\\ &\frac{\omega}{\gamma(1-\gamma)}\sum_{i=1}^{N}\sum_{s=1}^{T}(3E(s)+\frac{nL\alpha(s)}{\sigma}) = \qquad (15a)\\ &\frac{N\omega}{1-\gamma}\sum_{j=1}^{N}\|x_{j,1}\|+N\sum_{t=1}^{T}(3E(t)+\frac{nL\alpha(t)}{\sigma})+\\ &\frac{N\omega}{\gamma(1-\gamma)}\sum_{t=1}^{T}(3E(t)+\frac{nL\alpha(t)}{\sigma}),\end{aligned} \tag{15b}$$

其中式(15a)由$\sum_{t=1}^{T}\gamma^{t-1} \leqslant \dfrac{1}{1-\gamma}$可以得到.

此外, 注意到有下式成立:

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{i=1}^{N}\|x_{i,t}-\bar{x}_{t}\| =\\ &\sum_{i=1}^{N}\|x_{i,1}-\bar{x}_{1}\|+\sum_{t=1}^{T-1}\sum_{i=1}^{N}\|x_{i,t+1}-\bar{x}_{t+1}\| \leqslant\\ &\sum_{t=1}^{T}\sum_{i=1}^{N}\|x_{i,t+1}-\bar{x}_{t+1}\|,\end{aligned} \tag{16}$$

结合式(15b)和式(16), 定理1得证.　证毕.

为了分析定理2中的Regret界, 先给出引理7和引理8.

引理 7　在假设1–3成立的条件下, 对任意$i\in\mathcal{V}$, $t \in [T]$, 由算法1产生的序列$\{x_{i,t}\}$满足下式:

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{i=1}^{N}\{\frac{1}{\alpha(t)}D_h(x^{\rho*},\tilde{y}_{i,t})-\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1})\}\leqslant\\ &3\kappa NL_h\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)}+9N(1-\xi)L_h\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}+\frac{NR^2}{\alpha(T)},\end{aligned}$$

其中$x^{\rho*}=(1-\rho)x^*$, $\rho\in(0,1)$.

证　根据$\tilde{y}_{i,t}$的定义和假设2, 可得

$$\begin{aligned}&D_h(x^{\rho*},\tilde{y}_{i,t})-D_h(x^{\rho*},x_{i,t+1}) =\\ &D_h(x^{\rho*},\sum_{j=1}^{N}[P(t)]_{ij}(x_{j,t}+e_j(t))+p_i(t))-\\ &D_h(x^{\rho*},x_{i,t+1}) \leqslant\\ &\sum_{j=1}^{N}[P(t)]_{ij}D_h(x^{\rho*},x_{j,t}+e_j(t)+p_i(t))-\\ &D_h(x^{\rho*},x_{i,t+1}).\end{aligned} \tag{17}$$

根据Bregman散度的定义, 有

$$\begin{cases}D_h(x^{\rho*},x_{j,t}+e_j(t)+p_i(t)) =\\ h(x^{\rho*})-h(x_{j,t}+e_j(t)+p_i(t))-\\ \langle\nabla h(x_{j,t}+e_j(t)+p_i(t)),\\ x^{\rho*}-x_{j,t}-e_j(t)-p_i(t)\rangle =\\ h(x^{\rho*})-h(x_{j,t})-\langle\nabla h(x_{j,t}+\xi e_j(t)+\xi p_i(t)),\\ e_j(t)+p_i(t)\rangle-\langle\nabla h(x_{j,t}+e_j(t)+p_i(t)),x^{\rho*}-\\ x_{j,t}-e_j(t)-p_i(t)\rangle,\end{cases} \tag{18}$$

式(18)中的最后一个等式是根据拉格朗日中值定理, 即存在ξ, 满足$0\leqslant\xi\leqslant 1$使得下式成立:

$$\begin{aligned}&h(x_{j,t}+e_{j,t}+p_i(t)) =\\ &h(x_{j,t})+\langle\nabla h(x_{j,t}+\xi e_j(t)+\xi p_i(t)),e_j(t)+p_i(t)\rangle.\end{aligned}$$

在式(18)中引入$\langle\nabla h(x_{j,t}),x^{\rho*}-x_{j,t}\rangle$一项, 得

$$\begin{aligned}&D_h(x^{\rho*},x_{j,t}+e_j(t)+p_i(t)) =\\ &h(x^{\rho*})-h(x_{j,t})-\langle\nabla h(x_{j,t}),x^{\rho*}-x_{j,t}\rangle-\\ &\langle\nabla h(x_{j,t}+e_j(t)+p_i(t))-\nabla h(x_{j,t}),x^{\rho*}-x_{j,t}\rangle-\\ &\langle\nabla h(x_{j,t}+\xi e_j(t)+\xi p_i(t))-\\ &\nabla h(x_{j,t}+e_j(t)+p_i(t)),e_j(t)+p_i(t)\rangle \leqslant\\ &D_h(x^{\rho*},x_{j,t})+\kappa L_h\|e_j(t)+p_i(t)\|+\\ &(1-\xi)L_h\|e_j(t)+p_i(t)\|^2 \leqslant \qquad (19a)\\ &D_h(x^{\rho*},x_{j,t})+9(1-\xi)L_hE^2(t)+3\kappa L_hE(t),\end{aligned} \tag{19b}$$

其中式(19a)根据假设3可得. 将式(19b)代入式(17)后并对$i\in\mathcal{V}$和$t\in[T]$求和, 可得

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{\alpha(t)}D_h(x^{\rho*},\tilde{y}_{i,t})-\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1}) \leqslant\\ &\sum_{t=1}^{T}\frac{1}{\alpha(t)}\sum_{i=1}^{N}\{\sum_{j=1}^{N}[P(t)]_{ij}D_h(x^{\rho*},x_{j,t})+\\ &9(1-\xi)L_hE^2(t)+3\kappa L_hE(t)-D_h(x^{\rho*},x_{i,t+1})\}\leqslant\\ &\sum_{t=1}^{T}\frac{1}{\alpha(t)}\sum_{j=1}^{N}\sum_{i=1}^{N}[P(t)]_{ij}D_h(x^{\rho*},x_{j,t})+\\ &9N(1-\xi)L_h\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}+3\kappa NL_h\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)}-\\ &N\sum_{t=1}^{T}\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1}) =\\ &N\sum_{t=1}^{T}\{\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t})-\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1})\}+\\ &9N(1-\xi)L_h\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}+3\kappa NL_h\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)} \leqslant\\ &\frac{NR^2}{\alpha(T)}+9N(1-\xi)L_h\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}+\\ &3\kappa NL_h\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)}.\end{aligned} \tag{20}$$

证毕.

引理 8 在假设1–4成立的条件下, 对任意$i\in\mathcal{V}$, $t\in[T]$, 由算法1产生的序列$\{x_{i,t}\}$满足下式:

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{E}\left[f_{i,t}(x_{i,t})-f_{i,t}(x^{\rho*})\right]\leqslant \frac{Nn^2L^2}{2\sigma}\sum_{t=1}^{T}\alpha(t)+3\kappa NL_h\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)}+ 9N(1-\xi)L_h\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}+\frac{NR^2}{\alpha(T)}+ \frac{2Nn^2L^2\omega}{\gamma(1-\gamma)\sigma}\sum_{t=1}^{T}\alpha(t)+\frac{6NnL\omega}{\gamma(1-\gamma)}\sum_{t=1}^{T}E(t)+ 9NnL\sum_{t=1}^{T}E(t)+NTL\delta+2nL\tau+ \frac{2Nn^2L^2}{\sigma}\sum_{t=1}^{T}\alpha(t).$$

证 根据引理4的2)–3), 可得

$$f_{i,t}(x_{i,t})-f_{i,t}(x^{\rho*})\leqslant \tilde{f}_{i,t}(x_{i,t})-\tilde{f}_{i,t}(x_t^{\rho*})+L\delta\leqslant \langle\tilde{g}_{i,t},x_{i,t}-x^{\rho*}\rangle+\langle\nabla\tilde{f}_{i,t}(x_{i,t})- \tilde{g}_{i,t},x_{i,t}-x^{\rho*}\rangle+L\delta, \tag{21}$$

其中式(21)是来自于$\tilde{f}_{i,t}$在定义域上的凸性. 根据$\tilde{g}_{i,t}$是对$\nabla\tilde{f}_{i,t}$的无偏估计可知

$$\mathrm{E}[\langle\nabla\tilde{f}_{i,t}(x_{i,t})-\tilde{g}_{i,t},x_{i,t}-x^{\rho*}\rangle]=0.$$

对于式(21)中的第1项, 可得

$$\langle\tilde{g}_{i,t},x_{i,t}-x^{\rho*}\rangle= \langle\tilde{g}_{i,t},\tilde{y}_{i,t}-x_{i,t+1}\rangle+\langle\tilde{g}_{i,t},x_{i,t+1}-x^{\rho*}\rangle+ \langle\tilde{g}_{i,t},x_{i,t}-\tilde{y}_{i,t}\rangle. \tag{22}$$

考虑式(22)的第1项, 根据柯西不等式有

$$\langle\tilde{g}_{i,t},\tilde{y}_{i,t}-x_{i,t+1}\rangle\leqslant \frac{1}{\alpha(t)}\frac{\sigma}{2}\|x_{i,t+1}-\tilde{y}_{i,t}\|^2+\frac{\alpha(t)}{2\sigma}\|\tilde{g}_{i,t}\|^2. \tag{23}$$

考虑式(22)的第2项, 根据引理2, 可得

$$\langle\tilde{g}_{i,t},x_{i,t+1}-x^{\rho*}\rangle\leqslant \frac{1}{\alpha(t)}D_h(x^{\rho*},\tilde{y}_{i,t})-\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1})- \frac{1}{\alpha(t)}D_h(x_{i,t+1},\tilde{y}_{i,t})\leqslant \frac{1}{\alpha(t)}D_h(x^{\rho*},\tilde{y}_{i,t})-\frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1})- \frac{1}{\alpha(t)}\frac{\sigma}{2}\|x_{i,t+1}-\tilde{y}_{i,t}\|^2. \tag{24}$$

结合不等式(23)–(24)和引理7, 有

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\{\langle\tilde{g}_{i,t},\tilde{y}_{i,t}-x_{i,t+1}\rangle+\langle\tilde{g}_{i,t},x_{i,t+1}-x^{\rho*}\rangle\}\leqslant \frac{Nn^2L^2}{2\sigma}\sum_{t=1}^{T}\alpha(t)+\sum_{t=1}^{T}\sum_{i=1}^{N}\{\frac{1}{\alpha(t)}D_h(x^{\rho*},\tilde{y}_{i,t})- \frac{1}{\alpha(t)}D_h(x^{\rho*},x_{i,t+1})\}\leqslant \frac{Nn^2L^2}{2\sigma}\sum_{t=1}^{T}\alpha(t)+\frac{NR^2}{\alpha(T)}+3\kappa NL_h\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)}+ 9N(1-\xi)L_h\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}. \tag{25}$$

对于式(22)的最后一项, 根据$\tilde{y}_{i,t}$的定义和柯西不等式, 可得

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\langle\tilde{g}_{i,t},x_{i,t}-\tilde{y}_{i,t}\rangle\leqslant \sum_{t=1}^{T}\sum_{i=1}^{N}\|\tilde{g}_{i,t}\|\|\sum_{j=1}^{N}[P(t)]_{ij}(x_{j,t}+e_j(t))+ p_i(t)-x_{i,t}\|\leqslant nL\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}[P(t)]_{ij}\|x_{j,t}-x_{i,t}\|+ nL\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}[P(t)]_{ij}\|e_j(t)\|+nL\sum_{t=1}^{T}\sum_{i=1}^{N}\|p_i(t)\|\leqslant nL\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}[P(t)]_{ij}\|x_{i,t}-x_{j,t}\|+ NnL\sum_{t=1}^{T}E(t)+2NnL\sum_{t=1}^{T}E(t)\leqslant nL\sum_{t=1}^{T}\sum_{j=1}^{N}\|x_{i,t}-x_{j,t}\|+3NnL\sum_{t=1}^{T}E(t), \tag{26}$$

将定理1的结论应用到式(26), 结合式(25)代入式(22), 可证得该引理. 证毕.

算法1引入梯度估计值$\tilde{g}_{i,t}$对梯度值进行替代. 因此, 本文将分析算法的期望–平均Regret $\mathrm{E}\,[\mathrm{Reg}_i(T)]$.

定理 2 在假设1–4成立的条件下, 对任意$i\in\mathcal{V}$, $t\in[T]$, 对于算法1产生的序列$\{x_{i,t}\}$, 若给定步长$\alpha(t)$和量化参数$\beta(t)$, 则关于算法1的期望–平均Regret$\mathrm{E}\,[\mathrm{Reg}_i(T)]$满足下式:

$$\mathrm{E}\,[\mathrm{Reg}_i(T)]\leqslant m_1\sum_{t=1}^{T}\alpha(t)+m_2\sum_{t=1}^{T}E(t)+m_3\sum_{t=1}^{T}\frac{E(t)}{\alpha(t)}+ m_4\sum_{t=1}^{T}\frac{E^2(t)}{\alpha(t)}+\kappa NT\rho L+2(n+1)L\tau+NTL\delta,$$

其中: $m_1=\frac{2NnL^2\omega}{\gamma(1-\gamma)\sigma}+\frac{2Nn^2L^2\omega}{\gamma(1-\gamma)\sigma}+\frac{Nn^2L^2}{2\sigma}+\frac{2NnL(1+nL)}{\sigma}$, $m_2=\frac{6N(n+1)L\omega}{\gamma(1-\gamma)}+9NnL+6N$, $m_3=3\kappa NL_h$, $m_4=9N(1-\xi)L_h$.

证 根据式(2), 分析下式:

$$f_t(x_{i,t})-f_t(x^*)= [f_t(x_{i,t})-f_t(\bar{x}_t)]+$$

$$[\sum_{j=1}^{N} f_{j,t}(\bar{x}_t)-\sum_{j=1}^{N} f_{j,t}(x_{j,t})]+[f_t(x^{\rho*})-f_t(x^*)]+ [\sum_{j=1}^{N} f_{j,t}(x_{j,t})-\sum_{j=1}^{N} f_{j,t}(x^{\rho*})], \tag{27}$$

对于式(27)关于$t\in[T]$求和后, 根据假设4中2)和定理1, 可得

$$\sum_{t=1}^{T} f_t(x_{i,t})-\sum_{t=1}^{T} f_t(\bar{x}_t)\leqslant \frac{3NL\omega}{\gamma(1-\gamma)}\sum_{t=1}^{T}E(t)+\frac{NnL^2\omega}{\gamma(1-\gamma)\sigma}\sum_{t=1}^{T}\alpha(t)+ L\tau+\frac{NnL}{\sigma}\sum_{t=1}^{T}\alpha(t)+3N\sum_{t=1}^{T}E(t). \tag{28}$$

对式(27)第2项运用上述同样的方法可得

$$\sum_{t=1}^{T}\sum_{j=1}^{N}(f_{j,t}(\bar{x}_t)-f_{j,t}(x_{j,t}))\leqslant \frac{3NL\omega}{\gamma(1-\gamma)}\sum_{t=1}^{T}E(t)+\frac{NnL^2\omega}{\gamma(1-\gamma)\sigma}\sum_{t=1}^{T}\alpha(t)+ L\tau+\frac{NnL}{\sigma}\sum_{t=1}^{T}\alpha(t)+3N\sum_{t=1}^{T}E(t). \tag{29}$$

对于式(27)的第3项关于$t\in[T]$求和之后, 根据式(3)可得

$$\sum_{t=1}^{T} f_t(x^{\rho*})-\sum_{t=1}^{T} f_t(x^*)\leqslant L\sum_{t=1}^{T}\sum_{j=1}^{N}\|x^{\rho*}-x^*\|\leqslant\kappa NT\rho L, \tag{30}$$

由式(28)–(30)和引理8, 可证得该定理的结果.　证毕.

推论 1　在假设1, 2, 3, 4成立的条件下, 对任意$i\in\mathcal{V}, t\in[T]$, 对于算法1产生的序列$\{x_{i,t}\}$, 选择$\delta=\frac{1}{\sqrt{T}}, \rho=\frac{\delta}{a}, \alpha(t)=\frac{1}{\sqrt{t}}$和$\beta(t)=\frac{1}{\sqrt{t}}$. 根据式(2)定义的Regret满足下式:

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{E}\,[f_{i,t}(x_{i,t})-f_{i,t}(x^*)]\leqslant m_5+m_6\sqrt{T}+\frac{2n^{\frac{3}{2}}Lm_2}{\sigma K}\ln(T-1)+\frac{2(n+1)L}{\sqrt{T}}+ \frac{2n^{\frac{3}{2}}L}{\sigma K}\ln\left|\frac{T+1-\sqrt{T}}{T-1}\right|+\frac{2n^3L^2m_4}{\sigma^2K^2}\ln\left|\frac{\sqrt{T}-1}{\sqrt{T}+1}\right|\leqslant \mathcal{O}(\sqrt{T}).$$

其中: $m_5=\frac{2n^{\frac{3}{2}}Lm_3}{\sigma K}(1+\sqrt{2})+\frac{4n^3L^2m_4}{\sigma^2K^2}(1+\sqrt{2})+\frac{4n^{\frac{3}{2}}Lm_2}{\sigma K}$, $m_6=\frac{4n^{\frac{3}{2}}Lm_3}{\sigma K}+(\kappa+1)NL+2m_1$.

注 3　文献[17]中研究了基于Bandit反馈的分布式在线镜像下降算法(ODMD-B), 并证明其Regret界是$\mathcal{O}(\sqrt{T})$. 相较于文献[17], 本文考虑带有自适应量化器的分布式在线算法, 且得所提的算法Regret界同为$\mathcal{O}(\sqrt{T})$, 这说明本文算法在有限比特率下可以得到与未考虑量化通信的ODMD-B算法相近的收敛速率. 定理2中$E(t)$表示量化误差$e_i(t)$的上界. 若设置$E(t)=0$, 将得到文献[17]中定理1的结论. 若设置$E(t)=0, \rho=\delta=0$, 可得文献[14]中定理3的结果. 根据量化状态值$\hat{Q}(z_t,d(t),x)$和量化水平参数K的关系可知, 当量化水平参数K越大时, 量化器产生的量化误差上界$E(t)$越小, 由定理2可知此时算法的Regret值也会越小, 这将在第4节的数值实验中得以验证. 此外, 定理2说明, 在本文使用梯度估计值$\tilde{g}_{i,t}$替代梯度值, 使得算法产生光滑项误差$\kappa NT\rho L$和$NTL\delta$. 但本文算法仅需要计算目标函数值, 不需要计算梯度值, 从而可节省计算成本.

4　数值实验

本节使用的实验模型来自文献[21]中“个人可穿戴追踪设备”模型, 该实验模型的目标是寻找变量和真实值之间的线性关系, 即给定数据$(a_{i,t},b_{i,t})\in\mathbb{R}^n\times\mathbb{R}$, 求解以下极小化问题:

$$\min_{x\in\mathcal{X}}\sum_{t=1}^{T}\sum_{i=1}^{N}(a_{i,t}^{\mathrm{T}}x-b_{i,t})^2, \tag{31}$$

其中$\mathcal{X}=\{x\in\mathbb{R}^n|\|x\|\leqslant A\}$. 选择$h(x)=\|x\|^2$, 可得Bregman散度$D_h(x,y)=\|x-y\|^2$满足假设2的分离凸性.

初始化$x_{i,1}=z_{i,1}$, 选择步长$\alpha(t)=\beta(t)=\frac{1}{\sqrt{t}}$, $A=1, a=0.75, \delta=\frac{1}{\sqrt{T}}, \rho=\frac{\delta}{a}$. 为了方便说明本文所提出的AQODMD-B算法中自适应量化器对算法性能的影响, 图1设定$N=20, p=0.2$, 取量化水平参数$K=5$. 比较AQODMD-B与未考虑量化通信的分布式在线镜像下降算法(ODMD)[14]和ODMD-B[17]的时间–平均Regret(time-average regret)$\frac{E(\mathrm{Reg}_i(T))}{T}$.

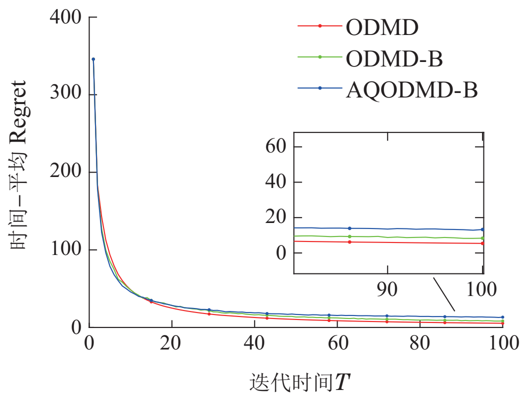

图 1　3种算法的比较

Fig. 1　Comparison of three algorithms

从图1可以看出Regret值最小的是ODMD算法, 并且AQODMD-B略微大于ODMD-B的Regret. 结果说明AQDMD-B在有限比特率的量化通信下依然有着不错的收敛性表现.

下面来比较在不同量化水平参数下算法1迭代300

步时的 $\frac{E(\mathrm{Reg}_i(T))}{T}$. 令$T=300$, $N=20$, $p=0.2$, 并分别设置量化水平参数$K=5, 10, 15, \cdots, 45$. 从图2中可以看出, 当量化水平参数越小时, 量化误差的上界$E(t)$越大, Regret的值越大, 这与定理2的结论相符合.

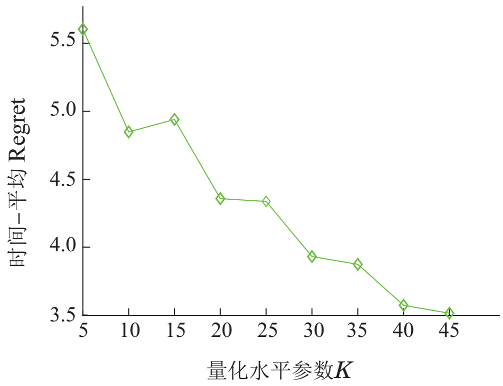

图 2 K对于AQODMD-B的影响

Fig. 2 Effect of K on AQODMD-B

5 总结

本文提出了一种Bandit反馈设定下的在线分布式镜像下降算法, 并利用镜像下降算法的性质设计了自适应量化器, 可将其用于解决时变多智能体网络中存在通信约束的分布式优化问题. 无梯度的计算方式拓展了算法的应用范围并避免了计算梯度时花费的高额计算代价. 数值仿真验证了本文所给出的定理, 并分析了不同参数对Regret值的影响.

参考文献:

[1] RAJKUMAR R, LEE I, SHA L, et al. Cyber-physical systems: The next computing revolution. *Proceedings of the 47th Design Automation Conference*. Anaheim, CA: IEEE, 2010: 731 – 736.

[2] LEE S, KIM Y K, ZHENG Y, et al. On model parallelization and scheduling strategies for distributed machine learning. *Advances in Neural Information Processing Systems*, 2014, 27(4): 2834 – 2842.

[3] DENG Z. Distributed algorithm design for aggregative games of euler-lagrange systems and its application to smart grids. *IEEE Transactions on Cybernetics*, 2021, 52(8): 8315 – 8325.

[4] NEDIĆ A, OZDAGLAR A. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 2009, 54(1): 48 – 61.

[5] NEDIĆ A, OZDAGLAR A, PARRILO P A. Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 2010, 55(4): 922 – 938.

[6] BECK A, TEBOULLE M. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 2003, 31(3): 167 – 175.

[7] XI C, WU Q, KHAN U A. Distributed mirror descent over directed graphs. *ArXiv Preprint*, 2014: arXiv:1412.5526,2014.

[8] DOAN T T, MAGULURI S T, ROMBERG J. Fast convergence rates of distributed subgradient methods with adaptive quantization. *IEEE Transactions on Automatic Control*, 2021, 66(5): 2191 – 2205.

[9] LIU J, YU Z, HO W C D. Distributed constrained optimization with delayed subgradient information over time-varying network under adaptive quantization. *ArXiv Preprint*, 2021: arXiv: 2106.08271,2021.

[10] HAZAN E, AGARWAL A, KALE S. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 2007, 69(2/3): 169 – 192.

[11] SHALEV-SHWARTZ S. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 2012,4(2): 107 – 194.

[12] SHAHRAMPOUR S, JADBABAIE A. An online optimization approach for multi-agent tracking of dynamic parameters in the presence of adversarial noise. *American Control Conference (ACC)*. Seattle, WA: IEEE, 2017: 3306 – 3311.

[13] YAN F, SUNDARAM S, VISHWANATHAN S, et al. Distributed autonomous online learning: Regrets and intrinsic privacy-preserving properties. *IEEE Transactions on Knowledge & Data Engineering*, 2013, 25(11): 2483 – 2493.

[14] SHAHRAMPOUR S, JADBABAIE A. Distributed online optimization in dynamic environments using mirror descent. *IEEE Transactions on Automatic Control*, 2018, 63(3): 714 – 725.

[15] SHAMIR O. An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *ArXiv Preprint*, 2015: arXiv: 1507.08752,2015.

[16] YUAN D, HONG Y, HO D W C, et al. Distributed mirror descent for online composite optimization. *IEEE Transactions on Automatic Control*, 2021, 66(2): 714 – 729.

[17] LI J, LI C, YU W, et al. Distributed online bandit learning in dynamic environments over unbalanced digraphs. *IEEE Transactions on Network Science and Engineering*, 2021, 8(4): 3034 – 3047.

[18] GHADIMI S, LAN G, ZHANG H. Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization. *Mathematical Programming*, 2016, 155(1): 267 – 305.

[19] CAO X, LIU K J R. Online convex optimization with time-varying constraints and bandit feedback. *IEEE Transactions on Automatic Control*, 2019, 64(7): 2665 – 2680.

[20] HAZAN E. *Introduction to Online Convex Optimization*. Foundations & Trends in Optimization, 2016, 2(3/4): 157 – 325.

[21] BRISIMI T S, CHEN R, MELA T, et al. Federated learning of predictive models from federated electronic health records. *International Journal of Medical Informatics*, 2018, 112: 59 – 67.

作者简介:

谢俊如 硕士研究生, 目前研究方向为分布式在线优化, E-mail: junru_xie@163.com;

高文华 博士, 副教授, 目前研究方向为随机时滞系统的鲁棒控制和多智能体系统的分布式优化, E-mail: whgao@scut.edu.cn;

谢奕彬 硕士, 目前研究方向为多智能体系统的分布式优化和统计过程控制, E-mail: maxieyibin@mail.scut.edu.cn.