基于压缩感知的多用户分布式计算

Ali Khalesi、Sajad Daei、Marios Kountouris、Petros Elia 通信系统，EURECOM，法国索菲亚·安提波利斯 06410 电子邮件：{ali.khalesi;sajad.daei;marios.kountouris;petros.elia}@eurecom.fr

arXiv:2301.03448v1 [cs.IT] 9 Jan 2023

摘要—**本文研究多用户线性可分布式计算问题，其中 N 台服务器协助计算 K 位用户请求的实值函数；每个函数都可以表示为至多L 个（通常为非线性的）子函数的线性组合。每台服务器计算一部分 γ 子函数，随后将其计算结果的某个函数传送给部分用户，每位用户再汇总收到的数据，以恢复所需函数。我们的目标是界定所有服务器完成的计算工作负载与数据集数量之比。**

为此，我们将实值分布式计算问题重新表述为矩阵分解问题，进而转化为基本的稀疏恢复问题，其中稀疏性意味着计算节省。在此基础上，我们首先提出一种简单的子函数分配概率方案，据此可将最优归一化计算成本的上界限定为$\gamma \leq \frac{K}{N}$ 通常难以处理的 ℓ_0-最小化所给出的值。为绕过此类最优方案的不可处理性，我们证明，如果这些最优方案满足 $\gamma \leq -r\frac{K}{N}W_{-1}^{-1}\left(-\frac{2K}{eNr}\right)$（其中 $W_{-1}(\cdot)$ 为Lambert函数，r用于校准服务器与用户之间的通信），那么实际上可以通过可处理的基追踪 ℓ_1-最小化求得这些方案。这一新揭示的联系，为运用压缩感知的工具和方法设计实用的分布式计算算法带来了可能。

索引词——**分布式计算、线性可分函数、压缩感知、稀疏表示。**

I. 引言

分布式计算在加速非线性且计算困难的任务方面发挥着重要作用。随着这些任务的复杂性不断增加，人们越来越需要新颖的并行处理技术，以便高效地将计算卸载到分布式服务器集群中。复杂性的增加也带来了诸多挑战，例如计算精度 [1], [2], 可扩展性 [3]–[5], 隐私与安全性 [6]–[10], 以及延迟和掉队者缓解 [11]–[15]。此外，围绕通信与计算之间的权衡，也有大量研究进行了各种探索 [14], [16], [17], [18]。此外，受能够推导出适用于实数的方案这一应用前景的启发，其他一些有趣的相关设定也应运而生，例如实数上的拉格朗日编码秘密共享 [19], 、复矩阵上的编码分布式多项式求值 [20], [21], 以及实数或复矩阵的安全分布式乘法设定 [22],。关于相关研究的详细综述，感兴趣的读者可参阅 [23], [24]。

这种复杂度的增加也催生了多种相关框架，例如 MapReduce [25] 和 Spark [26]，它们适用于广泛的函数类别。近期研究 [27] 聚焦于有限域上的函数，提出了所谓的多用户线性可分分布式计算框架。该框架支持对符合十分广泛的线性可分格式的函数进行分布式计算，而这种格式涵盖了各种具有实际意义的线性和非线性函数。这类函数的形式为

$$F(D_1, D_2, \ldots, D_L) = \sum_{l=1}^{L} f_l g_l(D_l)$$

其中，D_1，D_2，……，D_L 是 L 输入数据集，$W_l = g_l(D_l)$ 是基子函数 $g_l(D_l)$的计算输出，f_l 是标量系数。在多用户（K 个用户和 N台服务器）场景中，每个用户请求计算自己的函数，[27]中的工作将分布式计算问题转化为有限域上的一个简单（最好是稀疏的）矩阵分解问题，随后建立了分布式计算、矩阵分解与一种新的编码理论方法之间的直接联系。具体而言，对于一个 $K \times L$需求矩阵 $\mathbf{F}$，其中每一行描述定义某个用户所请求函数的系数，该问题被转化为分解问题$\mathbf{F} = \mathbf{DE}$，其中$\mathbf{D}$和$\mathbf{E}$ 分别是解码矩阵和编码矩阵。$\mathbf{E}$规定每台服务器应计算哪些子函数，以及各服务器在传输前应如何组合计算所得的输出；而 $K \times N$ 解码矩阵$\mathbf{D}$则规定每台服务器应与哪个用户通信，以及每个用户应如何组合收到的各种信号。随后，提出了一种源自强大覆盖码类以及一种全新的所谓部分覆盖码类的解决方案。具体而言，此类码的校验矩阵发挥了$\mathbf{D}$的作用；在将$\mathbf{F}$ 的各列视为综合后，$\mathbf{E}$的各列被确定为具有最小重量的陪集首领，从而保证$\mathbf{E}$ 最为稀疏，进而使计算

本研究由欧洲研究委员会（ERC）通过欧盟“地平线2020”研究与创新计划资助，资助协议编号为725929（Project DUALITY）和101003431（Project SONATA）。

有关更多信息，请参见 [28], [29]。

成本最低。例如，当$\mathbf{D}$ 由 q元域上的基本覆盖码类导出时，相应的归一化计算成本 $\gamma \in [0,1]$ ——表示每台服务器必须计算的全部子函数的比例——呈现为 $\gamma = H_q^{-1}(K/N)$，其中 $H_q^{-1}(\cdot)$是熵函数的函数逆。

不过，我们在此关注的是直接在实数域上计算函数，这确实与有限域情形有很大不同。之所以强调实数（或复数）域，是因为在有限域上计算实值问题（经过离散化后）可能不如直接在实数域上计算实用，主要是因为离散化可能带来很高的精度成本和精度损失，而且有限域计算也以比浮点运算慢得多而闻名。为此，我们将考虑定义在 L实值数据集上的实值函数（或等价地，具有 L 分量/基子函数），以及 N 个计算服务器和 K 个各自要求计算不同函数的用户。由于现在我们是在实数域上工作， [27]中的编码理论方法并不直接适用，因此需要一种新的方法。

在此，我们的方法首次建立了分布式计算与压缩感知之间的联系。首先，我们证明（命题1）存在一种可实现的方案，其归一化计算成本具有如下上界： $\gamma \leq \frac{K}{N}$。这是一个概率方案，其中$\mathbf{D}$ 从高斯系综中选取，而$\mathbf{E}$ 的相应稀疏性则由一个随机过程决定。接着，我们提出 ℓ_0-最小化方法，以$\mathbf{D}$ 和$\mathbf{F}$ 为输入，得到稀疏的$\mathbf{E}$。不过，这种最小化通常不可处理，因此我们借鉴压缩感知领域的丰富研究成果，提出一种更实用的方法。我们证明（定理 1），只要存在一种计算成本受限于$\gamma \leq -r\frac{K}{N}W_{-1}^{-1}(-\frac{2K}{eNr})$的方案（其中 $W_{-1}(\cdot)$是Lambert函数， r 是用于校准服务器与用户之间通信的参数），实际上就可以采用可处理的基追踪 ℓ_1-最小化来推导出这样的方案。

记号：对于矩阵 A 和 B， [A, B] 表示这两个矩阵的水平拼接。定义 $[n] \triangleq \{1,2,\ldots,n\}$。对于任意矩阵 $\mathbf{X} \in \mathbb{F}^{m\times n}$，$\mathbf{X}(i,j)$ $i \in [m]$, $j \in [n]$表示第 i 行、第 j 列的元素；$\mathbf{X}(i,:)$ $i \in [m]$表示第 i 行；$\mathbf{X}(:,j)$ $j \in [n]$ 表示 X 的第 j 列。对于两个索引集 $\mathcal{I} \subset [m], \mathcal{J} \in [n]$，$\mathbf{X}(\mathcal{I},\mathcal{J})$表示由 $\mathcal{I}$ 中的行和 $\mathcal{J}$中的列构成的子矩阵。我们用 $\|\mathbf{X}\|_0$ 表示矩阵（或向量）x 的非零元素个数。此外， $\otimes$ 表示克罗内克积，vec(X)表示 X 的向量化。

II. 系统模型与问题表述

我们考虑多用户线性可分布式计算场景（参见图 1），该场景由 K 个用户/客户端、 N 个活跃服务器以及一个协调服务器和用户的主节点组成。每台服务器执行的任务可能具有较高的计算复杂度，同时还受到时间限制。我们考虑这样一种场景：每台服务器 n都可以通过专用广播信道，在单次通信（一 个时隙）中向任意用户集合 $\mathcal{T}_n \subset [K]$ 发送信息。

在我们的设定中，每个用户都需要一个来自线性可分函数空间的函数（通常是非线性的）。这类函数以若干数据集 D_1、.....、 D_L 作为输入，并且由各个子函数 $g_l(D_l) \in \mathbb{R}$的线性组合构成，每个子函数都以单个数据集 D_l 作为输入。因此，用户 $k \in [K]$所需的函数 $F_k(D_1$、.....、 $D_L) \in \mathbb{R}$是一个实值函数，其形式为

$$F_k(D_1, D_2, \ldots, D_L) \triangleq f_{k,1}g_1(D_1)\ldots + f_{k,L}g_L(D_L) \quad (1)$$
$$= f_{k,1}W_1 + \ldots + f_{k,L}W_L \quad (2)$$

其中， $W_l = g_l(D_l) \in \mathbb{R}$， $l \in [L]$ 是所谓的“文件”输出，而 $f_{k,l} \in \mathbb{R}$， $k \in [K]$， $l \in [L]$ 是定义各个所需函数的线性组合系数。

A. 流程的各个阶段

该模型包含三个阶段：第一阶段是需求阶段，第二阶段是分配与计算阶段，最后一个阶段是传输与解码阶段。在需求阶段，每个用户 $k \in [K]$ 向主节点请求 $F_k(\cdot)$，主节点随后推导出式（2）所示的分解。然后，基于这些 K 所需函数，在分配与计算阶段，主节点将部分子函数分配给各个服务器 n；服务器随后计算这些子函数，并为其负责的所有子函数 $f_l(D_l)$ $l \in \mathcal{W}_n$ 生成相应的文件 $W_l = f_l(D_l)$。

在传输阶段，每台服务器 $n \in [N]$ 广播

$$z_n \triangleq \sum_{l\in[L]} e_{n,l}W_l, \ n \in [N] \quad (3)$$

一次性发送其本地计算所得输出文件的一个线性组合，并将其发送给特定的用户子集 $\mathcal{T}_n$。上述过程由所谓的编码系数 $e_{n,l} \in \mathbb{R}$ 定义，这些系数由主节点确定。

最后，在解码阶段，每个用户 k 对接收到的信号进行线性组合，如下所示

$$F'_k \triangleq \sum_{n\in[N]} d_{k,n}z_n \quad (4)$$

其中解码系数 $d_{k,n} \in \mathbb{R}$， $n \in [N]$同样由主节点确定。自然地 $d_{k,n} = 0$， $\forall k \notin \mathcal{T}_n$。最后，当对所有 $k \in [K]$ 都满足 $F'_k = F_k$ 时，我们称精确解码成功。

B. 问题表述

与有限域的情况类似，此处为了表述问题，我们使用$\mathbf{f} \triangleq [F_1, F_2, \ldots, F_K], \mathbf{f}_k \triangleq [f_{k,1}, f_{k,2}, \ldots, f_{k,L}]$， $k \in [K]$，$\mathbf{w} \triangleq [W_1, W_2, \ldots, W_L]^\intercal$，其中$\mathbf{f}$ 表示所需函数的输出向量（参见（2）），$\mathbf{f}_k$ 表示用户k 的函数系数向量（参见（2）），而$\mathbf{w}$ 表示由所有子函数的输出文件组合而成的向量。我们还有$\mathbf{e}_n \triangleq [e_{n,1}, e_{n,2}, \ldots, e_{n,L}]$， $n \in$ $^\intercal$

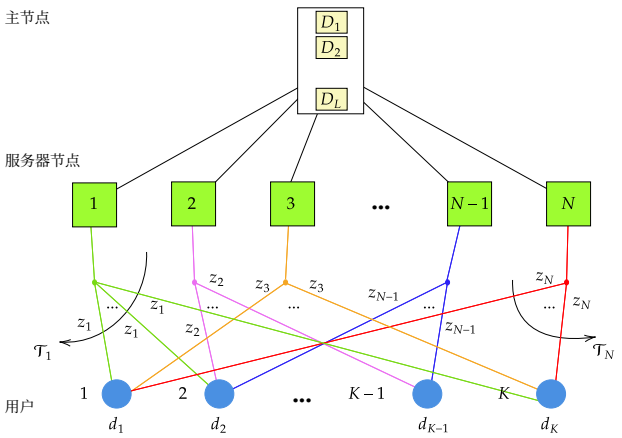

图 1。 K-用户、 N-服务器、 L-数据集的线性可分计算设置。一旦每个用户将其所需函数 $F_k(\cdot)$ 告知主节点，每个服务器 $n \in [N]$ 在 $\mathcal{W}_n \subseteq [L]$ 中计算一个子函数 $W_l = f_l(D_l) \in \mathbb{R}$ 。随后，服务器 n 向 $\mathcal{T}_n$ 中的所有用户广播一个线性组合 z_n （由本地可用的已计算文件构成）。该组合由系数 $e_{n,l}$ 定义。最后，为进行解码，每个用户 $k \in [K]$ 根据解码向量 $\mathbf{d}_k$，对其从各服务器接收到的所有信号进行线性组合。解码应为每个用户生成其所需函数 $F_k(D_1,\dots,D_L)$。

$[N]\mathbf{z} \triangleq [z_1, z_2, \dots, z_N]$ 分别表示服务器 n的编码向量，以及所有服务器的整体传输向量（参见（3））。此外，我们有 $\mathbf{d}_k \triangleq [d_{k,1}, d_{k,2}, \dots, d_{k,N}]^\intercal$, $k \in [K]\mathbf{f}' \triangleq [F'_1, F'_2, \dots, F'_K]^\intercal$

分别表示用户 k 处的解码向量，以及所有用户的已解码函数所组成的向量。此外，我们有 $\mathbf{F} \triangleq [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K]^\intercal \in \mathbb{R}^{K\times L}$, $\mathbf{E} \triangleq [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N]^\intercal \in \mathbb{R}^{N\times L}, \mathbf{D} \triangleq [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K]^\intercal \in \mathbb{R}^{K\times N}$

其中， $\mathbf{F}$ 表示所有用户的函数系数所构成的 $K \times L$ 所谓的任务矩阵； $\mathbf{E}$ 表示所有服务器的 $N \times L$ 计算与编码矩阵； 而$\mathbf{D}$ 表示所有用户的 $K \times N$ 解码矩阵 。

由（2）可直接得出:

$$\mathbf{f} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K]^\intercal \mathbf{w} \tag{5}$$

由（3）可得，整体传输向量的形式为:

$$\mathbf{z} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N]^\intercal \mathbf{w} = \mathbf{E}\mathbf{w}. \tag{6}$$

此外，由（4）可直接得出:

$$F'_k = \mathbf{d}_k^T \mathbf{z} \tag{7}$$

因此，我们有:

$$\mathbf{f}' = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K]^\intercal \mathbf{z} = \mathbf{D}\mathbf{z}. \tag{8}$$

回想一下，我们必须保证

$$\mathbf{f}' = \mathbf{f}. \tag{9}$$

将 (5)、(6) 和 (8) 代入 (9) 后，我们发现，当且仅当满足以下条件时，(9) 中的上述可行性条件才成立

$$\mathbf{DEw} = \mathbf{Fw}. \tag{10}$$

要使其对任意$\mathbf{w}$都成立，我们必须有:

$$\mathbf{DE} = \mathbf{F}. \tag{11}$$

C. 计算成本

简单回顾一下，每台服务器 n 会计算索引属于 $\mathcal{W}_n$ 的子函数；由于 $\mathcal{W}_n = \text{supp}(\mathbf{E}(n,:))$，因此在我们的情形中，归一化计算成本 自然可表示为

$$\gamma \triangleq \frac{\sum_{n=1}^{N} |\mathcal{W}_n|}{NL} = \frac{\|\mathbf{E}\|_0}{NL}. \tag{12}$$

可以看出， γ 简单地表示每台服务器必须计算的子函数的平均比例，这也就是$\mathbf{E}$中非零元素所占的比例。现在很清楚，将$\mathbf{F}$ 分解为两个矩阵$\mathbf{D}$ 和稀疏矩阵$\mathbf{E}$的乘积，可以降低计算成本，进而减少延迟。特别是，$\mathbf{E}$ 每一行中的非零元素越少，服务器完成一项任务所需的延迟就越短。

III. 结果

本节首先介绍一种用于子任务分配的基础概率方案，该方案采用高斯随机矩阵$\mathbf{D}$，并使用简单的迫零方法求解一个确定的线性系统。尽管这种方案较为基础，但它可以给出最优归一化计算成本的上界——而通过通常难以处理的 ℓ_0-最小化可得到该成本——即 $\gamma \le \frac{K}{N}$ 。

命题1. 对于包含 K 个用户、 N 台服务器和L 个数据集的多用户线性可分布式计算问题，采用随机高斯$\mathbf{D}$可保证以概率1存在一种方案，其归一化计算成本有界为 $\gamma \le K/N$，这构成了 ℓ_0-最小成本的上界。

证明. 由（11），我们有$\mathbf{F}(:\,l) = \mathbf{DE}(:\,l), \mathbf{F}(:\,l) \in \mathbb{R}^{K\times 1}$, $\forall l \in [L]$ 其中，对于每个 l，在（14）的 ℓ_0-最小化问题中，我们有$\mathbf{y} = \mathbf{F}(:\,l)$，$\mathbf{D} = \mathbf{A}$ 且$\mathbf{E}(:l) = \mathbf{z}$，这里同样有一个欠定方程组。

考虑选择 L 个任意随机子集 $\mathcal{S}_l \subset [N]$ $l \in [L]$ ，其中 $|\mathcal{S}_l| = K$。现在，对于每个 $l \in [L]$, 我们聚焦于 E 的第 $l-th$ 列 $\mathrm{E}([N], l)$，并将元素 $\mathrm{E}([N]\backslash\mathcal{S}_l, l)$ 设为 $= 0$，即第 l 列中只有索引属于 $\mathcal{S}_l$ 的元素保持非零。现在，由于 $\mathrm{F}(:\,l) = \mathrm{D}([K],[N]\backslash\mathcal{S}_l)\mathrm{E}([N]\backslash\mathcal{S}_l, l) + \mathrm{D}([K],\mathcal{S}_l)\mathrm{E}(\mathcal{S}_l, l)$，我们得到 $\mathrm{F}(:\,l) = \mathrm{D}([K],\mathcal{S}_l)\mathrm{E}(\mathcal{S}_l, l)$，这是一个定定方程组，使我们能够确定 $\mathrm{E}(\mathcal{S}_l, l)$。上述 $\mathrm{D}([K],\mathcal{S}_l)$ 是 D 中对应的 $K\times K$ 子矩阵。根据以上内容，每个这样的 $\mathrm{D}([K],\mathcal{S}_l)$ 都是一个 $K \times K$ 高斯

这意味着$\mathbf{D}$ 的每个元素都是独立同分布地从高斯分布中选取的。

子矩阵，它自然以概率 1 非奇异。因此，定定方程组对于 $\mathbf{E}(\mathcal{S}_l,\ l)\ \forall l \in [L]$ 总是有唯一解，因此该方案适用于所有 $F(:\ l)\ \ l \in [L]$。该方案保证$\|\mathbf{E}(:\ l)\|_0 \leq K$，于是 $\|\mathbf{E}\|_0 \leq KL$，从而保证 $\gamma \leq \frac{K}{N}$。采用ℓ_0最小化（如式（14）所示)–可以获得更好的性能。

□

正如我们将在第 IV 节讨论的，ℓ_0-最小化已知是 NP 困难 [30], ，因此不可处理。为了提供一种实用的解决方案，以下定理利用压缩感知的结果描述一系列实用方案，这些方案现在使用 ℓ_1-最小化来寻找——正如我们稍后将要阐明的——稀疏（且唯一的）编码矩阵$\mathbf{E}$。

定理 1. 对于多用户线性可分分布式计算问题，若有 K 个用户、N 台服务器和 L 个数据集，且存在一个方案，其采用 (κ, β) 次高斯随机矩阵$\mathbf{D}$ （参见引理 *2*），使得 ℓ_0-最小化能够得到

$$\gamma \leq -\frac{1}{r}\frac{K}{N}W_{-1}^{-1}(-\frac{2K}{erN}),\quad 0 \leq K/N \leq 12(2\beta+\kappa)/\kappa^2,$$

那么，以至少 $1-2e^{-\frac{KL}{r}}$ 的概率，可以通过基追踪 ℓ_1-最小化找到相应的（且唯一的）$\mathbf{E}$，其中 $r = 12\ (4\beta + 2\kappa)\ /\kappa^2$。

证明。 证明见下一节，该节首先简要介绍压缩感知。

□

IV. 定理 1 的证明

在开始证明之前，我们先快速介绍压缩感知的一些基本性质。

A. 压缩感知简介

我们在此简要介绍压缩感知结果 [30]–[40], ，这些结果将用于我们的分布式计算问题。我们将采用压缩感知文献中常用的符号，并在下一小节阐明这些符号与计算参数之间的联系。

如 [34], 所述，压缩感知旨在根据少量如下形式的欠定线性测量，恢复稀疏向量$\mathbf{x} \in \mathbb{R}^p$ ：

$$\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^m \tag{13}$$

其中，$\mathbf{A} \in \mathbb{R}^{m\times p}, m, p \in \mathbb{N}$ 是所谓的测量矩阵，而$\mathbf{y} = [y_1, ..., y_m]^T$是测量向量。在我们的情形中，正如我们稍后将看到的，$\mathbf{y}$ 将与我们的计算与编码矩阵$\mathbf{E}$相关联，$\mathbf{A}$ 则与通信和解码矩阵$\mathbf{D}$相关联，而$\mathbf{x}$ 将与任务矩阵$\mathbf{F}$相关联。一般方法是通过一种基本但计算上不可处理的ℓ_0-最小化来恢复最稀疏解，其形式如下

$$\min_{\mathbf{z}\in\mathbb{R}^p} \|\mathbf{z}\|_0 := \sum_{i=1}^{p} 1_{|z_i|\neq 0} \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z} \tag{14}$$

其中，$1_{|z_i|\neq 0}$ 表示指示函数。同样的优化将得到$\mathbf{E}$ 的最稀疏解，从而使 γ尽可能小。据我们所知，目前尚无研究结果能够对该最稀疏解的权重进行界定，因此我们将采用一种基本的构造性方法来界定 γ。

优化问题 **(14)** 的 NP 困难性质促使人们考虑采用 ℓ_1-范数最小化方法，也称为基追踪。该方法被认为是 **(14)** 最接近的凸可解替代方案，其形式如下

$$\min_{\mathbf{z}\in\mathbb{R}^p} \|\mathbf{z}\|_1 := \sum_{i=1}^{p} |z_i| \tag{15}$$

$$\text{s.t.}\ \ \mathbf{y} = \mathbf{A}\mathbf{z}. \tag{16}$$

在压缩感知领域，众所周知，只要满足某些条件，通过求解 **(15)** 得到的估计值$\hat{\mathbf{x}} \in \mathbb{R}^p$ 就能实现所需的唯一解$\mathbf{x}$ 。这些条件与测量矩阵的某些性质密切相关，其中一个条件是著名的限制等距性质（RIP） [41], ，它规定了 ℓ_1-范数优化算法（例如基追踪） [42], 的表现。以下结果对此进行了说明，摘自 [30, 定理 6。2],，现转载如下。

引理 1。 对于矩阵 $\mathbf{A} \in \mathbb{R}^{m\times p}$ 以及

$$\delta_s(\mathbf{A}) \triangleq \max_{\mathcal{S}\subset[N],|\mathcal{S}|\leq s} \|\mathbf{A}_{\mathcal{S}}^{*}\mathbf{A}_{\mathcal{S}} - \mathbf{I}\|_{2\to 2}^2 \tag{17}$$

作为第 s 个受限等距常数，若 $\delta_{2s}(\mathbf{A}) < \frac{1}{3}$，则每个 s-–稀疏向量 $\mathbf{x} \in \mathbb{R}^p$ 都是以下问题的唯一解

$$\underset{\mathbf{z}\in\mathbb{R}^p}{\text{minimize}} \|\mathbf{z}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{z} = \mathbf{A}\mathbf{x}. \tag{18}$$

特别地，上述引理表明，若有一个测量向量$\mathbf{y} \in \mathbb{R}^m$，且我们先验地知道它是由线性系统 $\mathbf{A}\mathbf{x}$，$\mathbf{x} \in \mathbb{R}^p, m \leq p$产生的，该系统由唯一的 s-稀疏向量$\mathbf{x}$ 诱导，若 $\delta_{2s}(\mathbf{A}) < \frac{1}{3}$，那么通过求得一个 ℓ_1-最小化解，我们可以找到$\mathbf{z}$ 作为 $\mathbf{A}\mathbf{z}$的解，其使 $\|\mathbf{z}\|_1$达到最小值；此时，上述引理保证该最小化问题的解等于 $\mathbf{x}$。上述引理表明， $\delta_{2s}(\mathbf{A}) < 1/3$ 足以保证通过 ℓ_1-最小化精确恢复所有唯一的 s-稀疏向量。它基本上说明，如果 $\mathbf{A}$ 作用于稀疏向量时的表现与正交归一矩阵相对相似，那么 ℓ_1-最小化就会起到 ℓ_0-最小化的作用。

在这里讨论的问题中，选择$\mathbf{D}$ 使$\mathbf{A}$ 满足上述性质至关重要。以下两个结果直接说明了如何做到这一点。下面的第一个结果直接改编自 [30, 定理 9.2]。

引理 2. 设$\mathbf{A} \in \mathbb{R}^{m\times p}$ 为一个独立同分布（均值为零、方差为单位值）的次高斯随机矩阵，其参数为 β 和κ，且满足

$$\mathbb{P}(|A_{i,j}| \geq t) \leq \beta \mathrm{e}^{-\kappa t^2}, \forall t > 0. \tag{19}$$

那么，对于任意满足条件的，

$$\frac{\delta_{2s}}{2^{\ -\delta^2\frac{m}{2e}}}\sqrt{m} \ \leq \delta \qquad 1-(\text{A})\text{以至少 e 的概率成立}$$

$$m \geq 2c\delta^{-2}s\ln(\frac{ep}{2s}) \tag{20}$$

对于我们的计算问题，这些条件将表现为对 γ 的上界。

其中

$$c = \frac{2(4\beta + 2k)}{3k^2}. \tag{21}$$

结合引理 2 和引理 1，并令 $\delta = 1/3$，可知只要测量数量满足 $m \geq (12(4\beta+2\kappa)/\kappa^2)s\ln(\binom{ep}{2s})$，通过式 (15) 中的 ℓ_1-最小化，就能以高概率实现对所有 s-稀疏向量的一致恢复。

B. 定理 *1* 的证明

由式 (11) 可直接得到

$$\mathrm{vec}(\mathbf{F}) = (\mathbf{D} \otimes I_{L\times L}) \times \mathrm{vec}(\mathbf{E}), \tag{22}$$

在考虑 $\mathbf{y} = \mathrm{vec}(\mathbf{F})$、$\mathbf{A} = \mathbf{D}\otimes I_{L\times L}$ 和 $\mathbf{x} = \mathrm{vec}(\mathbf{E})$ 时，这与式 (13) 中的压缩感知形式 $\mathbf{y} = \mathbf{A}\mathbf{x}$ 相符，其中 $m = KL$ 且 $p = NL$。

此外，还需注意，由 [43]，可直接得到

$$\delta_s(\mathbf{D} \otimes \mathbf{I}_{L\times L}) \leq \delta_s(\mathbf{D}). \tag{23}$$

由于$\mathbf{D}$ 的元素独立地取自均值为零、方差为单位值且参数为 β, κ 的次高斯分布（参见式 (19)），现在可以结合引理 1 和式 (23)，并在令 $\delta_{2s} < \delta = 1/3$ 后应用引理 2，得出通过基追踪驱动的 ℓ_1-最小化实现唯一$\mathbf{E}$矩阵精确恢复的阈值为

$$KL \geq r\left\|\mathbf{E}\right\|_0 \ln(\frac{eNL}{2r\left\|\mathbf{E}\right\|_0}) \tag{24}$$

其中 $r = 12(4\beta+2\kappa)/\kappa^2$。然后将两边除以 NL 进行归一化，并应用 (12)，得到

$$\frac{K}{N} \geq r\gamma \ln(\frac{e}{2r\gamma}). \tag{25}$$

现在定义 $f(x) \triangleq -r^{-1}xW_{-1}^{-1}(-\frac{2x}{e})$，并在 $x_1 = K/N$和 $x_2 = r\gamma\ln(\frac{e}{2r\gamma})$ 处将其应用于 (25) 的两边。由于 $f(x)$是定义在 $0 \leq x \leq r/2$ 和 $r/2 \geq x_1 \geq x_2 \geq 0$上的单调递增函数，因此有 $f(x_1) \geq f(x_2)$。鉴于其反函数为 $f^{-1}(x) = rx\ln(\frac{e}{2rx})$，即可得出该结论。

V. 讨论与结论

在我们的分布式计算问题中，值得注意的是，实数情形（采用了压缩感知技术）与有限域情形之间存在一些相似之处；后者 [27], 利用了覆盖码的结构，其覆盖半径可衡量解的稀疏性，适用于 $\mathbf{E}$。例如，

由此可见，对于 $g(x) \triangleq rx\frac{e}{2rx}$ $f(x) \triangleq (-r^{-1}xW_{-1}^{-1}(-2x/e))$ $g^{-1}(x) = f(x)$ ln()和()()，我们看到()()成立，这仅仅是因为 $f(g(x)) = -r^{-1}g(x)W_{-1}^{-1}(-\frac{2g(x)}{e}) = -x$ $\frac{e}{2rx}W_{-1}^{-1}((\cdot)e^{(\cdot)}) = (\cdot)$ ln() ln($-x$) $= -x$ $= x$ $e/2rx)W_{-1}^{-1}(\frac{2rx}{e}\ln(2rx/e))\ln(e/2rx)/\ln(2rx/e)$。

在一个 q 元字母表上，$H_q(x)$ 表示熵函数，其形式为：对于所有 $0 < x < 1 - 1/q$，$H_q(x) \triangleq x\log_q(q-1) - x\log_q(x) - (1-x)\log_q(1-x)$。

在 $L = q^K$的极端情况下， [27] 中的研究揭示，最优归一化计算成本的形式为 $\gamma \simeq H_q^{-1}(K/N)$；当 q趋于无穷大时，这几乎与本文推导出的 $\gamma \simeq K/N$ 相匹配。

还值得注意的是，我们在定理 1 中的结果自动满足了额外的唯一性——即式（13）中最稀疏解 $\mathbf{x}$ 的唯一性——而这一性质实际上并不是我们的分布式计算问题所必需的。在去除这一唯一性条件后，进一步探索计算成本方面的改进将会很有意义。最后可以设想，分布式计算问题的进一步改进也可能受益于 [34] 压缩感知与纠错之间所揭示的深层联系。

参考文献

[1] T. Jahani-Nezhad 和 M. A. Maddah-Ali，“Codedsketch：一种用于近似矩阵乘法分布式计算的编码方案”，*IEEE*信息论汇刊，第 67 卷，第 6 期，第 4185–4196 页，2021年。[2] J. Wang、Z. Jia 和 S. A. Jafar，“编码分布式矩阵乘法中的精度代价：一种量纲分析”，载于*2021 IEEE*信息论研讨会（*ITW*），第 1–6 页，IEEE，2021年。[3] S. Li、Q. Yu、M. A. Maddah-Ali 和 A. S. Avestimehr，“无线分布式计算的可扩展框架”，*IEEE/ACM*网络汇刊，第 25 卷，第 5 期，第 2643–2654 页，2017 年。[4] F. Haddadpour、M. M. Kamani、M. Mahdavi 和 V. Cadambe，“以通信换取冗余：加速非凸优化的分布式随机梯度下降”，载于国际机器学习会议，第 2545–2554 页，PMLR，2019 年。[5] M. Soleymani、H. Mahdavifar 和 A. S. Avestimehr，“模拟拉格朗日编码计算”，*IEEE*信息论选定领域期刊，第 2 卷，第 1 期，第 283–295 页，2021年。[6] H. Sun 和 S. A. Jafar，“隐私计算的容量”，*IEEE*信息论汇刊，第 65 卷，第 3897–3880 页，2018年，第 6 期。[7] M. Soleymani 和 H. Mahdavifar，“分布式多用户秘密共享”，*IEEE*信息论汇刊，第 67 卷，第 1 期，第 164–178 页，2020年。[8] A. Khalesi、M. Mirmohseni 和 M. A. Maddah-Ali，“分布式多用户秘密共享的容量区域”，*IEEE*信息论选定领域期刊，第 2 卷，第 3 期，第 1057–1071 页，2021年。[9] M. Soleymani、H. Mahdavifar 和 A. S. Avestimehr，“模拟域中的隐私保护分布式学习”，*arXiv*预印本 *arXiv:2007.08803*，2020年。[10] M. Soleymani、R. E. Ali、H. Mahdavifar 和 A. S. Avestimehr，“列表可译码编码计算：突破对抗容忍屏障”，*IEEE*信息论选定领域期刊，第 2 卷，第 3 期，第 867–878 页，2021年。[11] N. Raviv、I. Tamo、R. Tandon 和 A. G. Dimakis，“基于循环MDS码和扩展图的梯度编码”，*IEEE*信息论汇刊，第 66 卷，第 12 期，第 7475–7489 页，2020年。[12] K. Lee、M. Lam、R. Pedarsani、D. Papailiopoulos 和 K. Ramchandran，“利用编码加速分布式机器学习”，*IEEE*信息论汇刊，第 64 卷，第 3 期，第 1514–1529 页，2017 年。[13] M. Egger、R. Bitar、A. Wachter-Zeh 和 D. G¨und¨uz，“通过组合多臂老虎机实现高效分布式机器学习”，*arXiv*预印本 *arXiv:2202.08302*，2022年。[14] K. Wan、H. Sun、M. Ji 和 G. Caire，“分布式线性可分计算”，*IEEE*信息论汇刊，第 68 卷，第 2 期，第 1259–1278 页，2022年。[15] A. Behrouzi-Far 和 E. Soljanin，“分布式计算中用于掉队者缓解的高效复制”，*arXiv*预印本 *arXiv:2006 02318*，2020年。。

[16] K. Wan、H. Sun、M. Ji 和 G. Caire，《分布式线性可分计算中计算与通信成本之间的权衡》，IEEE通信汇刊，第69卷，第11期，第7390–7405页，2021年。[17] E. Parrinello、E. Lampiris 和 P. Elia，《通过节点协作显著提高加速因子的编码分布式计算》，载于2018年IEEE国际信息论研讨会（ISIT），第1291–1295页，2018年。[18] S. Li、M. A. Maddah-Ali、Q. Yu 和 A. S. Avestimehr，《分布式计算中计算与通信之间的基本权衡》，IEEE信息论汇刊，第64卷，第1期，第109–128页，2018年。[19] K. Tjell 和 R. Wisniewski，《基于实数秘密共享的分布式计算中的隐私》，2021年。[20] M. Soleymani、H. Mahdavifar 和 A. S. Avestimehr，《模拟拉格朗日编码计算》，IEEE信息论选定领域期刊，第2卷，第1期，第283–295页，2021年。[21] M. Soleymani、H. Mahdavifar 和 A. S. Avestimehr，《模拟秘密共享及其在隐私保护分布式学习中的应用》，IEEE信息取证与安全汇刊，第17卷，第1904页，2022年，第1893页。[22]O. Makkonen 和 C. Hollanti，《复数上的模拟安全分布式矩阵乘法》，载于2022年IEEE国际信息论研讨会（ISIT），第1211–1216页，2022年。[23] J. S. Ng、W. Y. B. Lim、N. C. Luong、Z. Xiong、A. Asheralieva、D. Niyato、C. Leung 和 C. Miao，《编码分布式计算综述》，arXiv预印本 arXiv:2008.09048，2020年。[24] S. Li 和 S. Avestimehr，《编码计算：缓解大规模分布式计算和机器学习中的基本瓶颈》，第17卷，2020年。[25] J. Dean 和 S. Ghemawat，《MapReduce：大型集群上的简化数据处理》，ACM通讯，第51卷，2008年，第113页，第1期，第107页。[26] M. Zaharia、M. Chowdhury、M. J. Franklin、S. Shenker 和 I. Stoica，《Spark：基于工作集的集群计算》，载于第二届USENIX云计算热点专题研讨会（HotCloud 10），2010年。[27] A. Khalesi 和 P. Elia，《多用户线性可分布式计算》，arXiv预印本 arXiv:2206.11119，2022年。[28] A. Mallick、S. Smith 和 G. Joshi，《用于分布式非线性计算的无速率码》，第1–5页，2021年。[29] A. Mallick 和 G. Joshi，《用于分布式非线性计算的无速率和恢复码》，[30] S. Foucart 和 H. Rauhut，《压缩感知数学导论》，2013年。[31] E. J. Candes 和 Y. Plan，《压缩感知的概率理论及其无限制等距性质版本》，IEEE信息论汇刊，第57卷，第11期，第7235–7254页，2011年。[32] E. J. Candes、J. K. Romberg 和 T. Tao，《从不完整且不准确的测量中稳定恢复信号》，《纯粹与应用数学通讯：库朗数学科学研究所期刊》，第59卷，第8期，第1207–1223页，2006年。[33] E. J. Candes，《受限等距性质及其对压缩感知的影响》，《数学报告》，第346卷，第9–10期，第589–592页，2008年。[34] E. J. Candes 和 T. Tao，《从随机投影中近最优地恢复信号：通用编码策略？》，IEEE信息论汇刊，第52卷，第12期，第5406–5425页，2006年。[35] T. T. Cai、L. Wang 和 G. Xu，《受限等距常数的新界》，IEEE信息论汇刊，第56卷，第9期，第4388–4394页，2010年。[36] T. T. Cai、L. Wang 和 G. Xu，《移位不等式与稀疏信号恢复》，IEEE信号处理汇刊，第58卷，第3期，第1300–1308页，2010年。[37] S. Foucart，《关于通过 ℓ1最小化保证稀疏恢复的注记》，《应用与计算调和分析》，第29卷，第1期，第97–103页，2010年。[38] S. Foucart 和 M.-J. Lai，《通过 ℓq最小化求解欠定线性系统的最稀疏解，适用于 $0 < q \leq 1$》，《应用与计算调和分析》，第26卷，第3期，第395–407页，2009年。[39] Q. Mo 和 S. Li，《受限等距常数 δ2k的新界》，《应用与计算调和分析》，第31卷，第3期，第460–468页，2011年。[40] S. Daei、F. Haddadi 和 A. Amini，《临近边界：全变差最小化相变的下界》，IEEE信息论汇刊，第66卷，3267，2019年5月，第3261页–。[41] E. Candes 和 T. Tao，“通过线性规划解码”，*IEEE*信息论汇刊，第51卷，4215，2005年12月，第4203页–。[42] S. S. Chen、D. L. Donoho 和 M. A. Saunders，“通过基追踪进行原子分解”，*SIAM*评论，第43卷，第1期，第129–159页，2001年。[43] S. Jokar 和 V. Mehrmann，“欠定克罗内克积系统的稀疏解”，线性代数及其应用，第431卷，第12期，第2437–2447页，2009年。