Advances in Applied Mathematics 应用数学进展, **2023, 12(1), 183-202**
Published Online January 2023 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2023.121022

Hans汉斯

基于梯度扰动和BB步长的迭代收缩阈值差分隐私算法

苑文丽，彭定涛*

贵州大学，数学与统计学院，贵州 贵阳

收稿日期：2022年12月28日；录用日期：2023年1月23日；发布日期：2023年1月30日

摘 要

本文研究隐私保护下带有非凸正则的经验风险极小化问题，其中损失函数是凸函数，正则项为MCP函数。提出了基于梯度扰动和Barzilar-Borwein (BB)步长的迭代收缩阈值差分隐私算法(ISTDP)。首先，基于算法每次迭代均对梯度添加高斯噪声，证明了该算法具有差分隐私保护性质。其次，基于以BB步长做试探步进行线搜索的迭代收缩阈值算法，证明了该算法可以收敛于任意给定的精度。因此，ISTDP算法是一种可以满足隐私保护要求的机器学习优化算法。

关键词

差分隐私保护，梯度扰动，收缩阈值算法，MCP 正则

Iterative Shrink Threshold Differential Privacy Algorithm Based on Gradient Perturbation and BB Step Size

Wenli Yuan, Dingtao Peng*

* 通讯作者。

文章引用: 苑文丽, 彭定涛. 基于梯度扰动和BB步长的迭代收缩阈值差分隐私算法[J]. 应用数学进展, 2023, 12(1): 183-202. DOI: 10.12677/aam.2023.121022

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: Dec. 28th, 2022; accepted: Jan. 23rd, 2023; published: Jan. 30th, 2023

Abstract

In this paper, we study the problem of empirical risk minimization with nonconvex regularization under privacy protection, where the loss function is a convex function and the regular term is the MCP function. An iterative shrinkage threshold difference privacy algorithm (ISTDP) based on gradient perturbation and Barzir-Borwein (BB) step size is proposed. First, ISTDP algorithm is proved to have the property of differential privacy protection since Gauss noise is added to the gradient for each iteration in the algorithm. Secondly, it is proved that the ISTDP algorithm can converge at any given accuracy due to the fact that ISTDP algorithm adopts an iterative shrinkage threshold algorithm together with a line search beginning with BB step size. Therefore, ISTDP algorithm is a kind of machine learning optimization algorithm which can satisfy the requirement of privacy protection.

Keywords

Differential Privacy Protection, Gradient Perturbation, Shrinkage Threshold Algorithm, MCP Regularization

1. 引言

在大数据与人工智能时代, 组织和机构收集的数据是当今信息时代的关键资源. 这些数据集

可能是众包的并包含敏感信息, 这些隐私信息的泄露对个人隐私构成了严重的威胁, 从而使得提供一个达到隐私保护的同时保持数据的可用性的算法成为一个重要的问题 [1]. 2006 年微软公司的Dwork 提出的差分隐私(Differential Privacy, 简称DP)概念 [2–4], 就是为隐私数据分析量身定制的隐私概念, 其目的是在利用隐私数据的整体信息的同时保护每个个体信息. 具体来说, 就是在一次统计查询的数据集中增加或减少一条记录, 可获得几乎相同的输出 [5]. 也就是说任何一条记录, 它在不在数据集中, 对结果的影响可忽略不计, 从而无法从结果中还原出任何一条原始的记录. 为了保证个人隐私不被泄露, 同时, 又能使得差分隐私算法的输出结果尽量准确. 差分隐私模型的构建常与经验风险极小化技术相结合, 一般的经验风险极小化模型如下 [6–8]:

$$\min_{\theta} L(\theta; D) := \frac{1}{n}\sum_{i=1}^{n} \ell(\theta, z_i)$$

其中$D = \{z_1, \cdots, z_n\}$为包含n条敏感数据$z_i = \{x_i, y_i\} \in \mathbb{R}^{d+1}$ 的数据集. $\ell(\theta, z_i)$是作用于数据z_i上的损失函数, $L(\theta; D) : \mathbb{R}^d \to \mathbb{R}$ 表示作用于整个数据集D 上的损失函数, 它是数据集D上的所有数据z_i相应损失函数$\ell(\theta, z_i)$的平均值, 不失一般性, 设它是水平有界的.

为了避免上述模型过拟合, 一般需要加入正则项, 即如下正则化的经验风险极小化模型:

$$\min_{\theta} F(\theta; D) := L(\theta; D) + R(\theta). \tag{1}$$

其中$R(\theta)$是正则函数, 常取$R(\theta) = \lambda\|\theta\|_1$或者$\lambda\|\theta\|_2^2$ [9], $\lambda > 0$ 是正则化参数.

对模型(1), 当损失函数和正则函数均为凸函数时, 满足差分隐私保护的算法目前常用的有三种：第一种是基于输出扰动的算法 [9, 10], 主要是针对数据发布中的查询结果和机器学习算法的输出结果添加噪声, 由于输出是在包含有隐私的数据上得到的, 因此输出便有可能暴露用户隐私 [1]; 第二种是基于目标扰动的算法 [11], 主要是针对差分隐私学习框架下的目标函数添加噪声, 虽可以保证隐私性, 但要在目标扰动的情况下获得最优解往往是棘手的; 第三种是基于梯度扰动的算法 [12–15], 是指在算法的迭代过程中对每一迭代步中的梯度添加噪声, 研究表明好的算法可以克服隐私泄露和计算问题 [16].

正则函数$R(\theta)$是为了防止过拟合而引入的, 但$R(\theta) = \|\theta\|_1$或者$\|\theta\|_2^2$正则问题得到的解却是有偏的, 即估计量的数学期望不等于被估计参数的真实值. 因此, 如何选择正则函数, 使得该正则化问题既能用于防止过拟合又能使结果更准确是本文考虑的重点. Fan 和Li [17] 指出一个好的正则函数应当使得产生的估计量具有下述四个性质: (1) 无偏性, (2) 稀疏性, (3) 连续性, (4) Oracle 性质. 研究者已经证明：在最小二乘损失情况下, 折叠凹(非凸)正则函数SCAD [17, 18], MCP [19]和Capped-ℓ_1 [20–25] 产生的估计量具有上述性质 [26, 27].

受 [17, 19, 20, 24, 25, 27–30] 非凸正则的启发, 本文讨论以下差分隐私保护下的带有MCP正则项

的经验风险极小化问题:

$$\min_{\theta} F(\theta; D) := L(\theta; D) + R(\theta), \tag{2}$$

其中$D = \{z_1, \ldots, z_n\} \in \mathbb{Z}^n$ 表示带有n 条记录的数据集, 其中$\mathbb{Z} \subseteq \mathbb{R}^{d+1}$称为数据的取值空间, $z_i = (x_i, y_i) \in \mathbb{R}^{d+1}$, $i \in [n] = \{1, 2, \ldots, n\}$称为数据记录, 并且每一条记录对应一个数据贡献者, 且包含该用户的所有信息, n表示数据集中包含的记录条数, $x_i \in \mathbb{R}^d$称为数据贡献者的性质信息, $y_i \in \{0, 1\} \subset \mathbb{R}$称为用户的标签. $L(\theta; D)$ 是光滑凸函数且具有梯度Lipschitz连续, 即存在一个常数$L > 0$, 使得

$$\|\nabla L(w, D) - \nabla L(u, D)\| \leq L \|w - u\|, \ \forall w, u \in R^d.$$

$R(\theta)$ 取MCP (Minimax Concave Penalty)正则函数, 即$R(\theta) = \sum_{i=1}^{d} \phi(\theta_i)$, 其中

$$\phi(t) := \begin{cases} \lambda |t| - t^2/(2\alpha), & |t| \leq \alpha\lambda, \\ \alpha\lambda^2/2, & |t| > \alpha\lambda, \end{cases} \quad \lambda > 0, \alpha > 1.$$

机器学习中许多问题都适用于以上模型, 通过选取不同的损失函数可以实现不同的学习任务, 比如常见的逻辑回归 [11]、核方法 [31] 等.

本文主要结构如下：第二部分中, 本文提出基于梯度扰动和Barzilai-Borwein (BB) 步长的迭代收缩阈值差分隐私(ISTDP)算法. 第三部分对所提出的ISTDP算法进行理论分析, 证明ISTDP算法不仅满足差分隐私保护要求, 还可以收敛到任意给定的精度. 第四部分进行简单总结.

符号说明: 以下符号贯穿全文, $\forall t \in R$, $|t|$表示t 的绝对值. $\|x\|_1$ 表示向量x的ℓ_1 范数. $\|x\|$ 表示向量x 的ℓ_2范数. $g^k = \nabla L\left(\theta^k; D\right)$ 表示损失函数在θ^k 处的梯度.θ^* 表示目标函数的最优解. Pr表示概率. sign(t)为符号函数

$$\text{sign}(t) := \begin{cases} 1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0. \end{cases}$$

2. 算法框架

本节提出一种普适的隐私保护算法——基于梯度扰动和BB 步长的迭代收缩阈值差分隐私算

法(Iterative Shrinkage Threshold Differential Privacy Algorithm based on Gradient Perturbation and Barzilar-Borwein Step Size, 简称ISTDP).

在ISTDP算法的每一迭代步, 对梯度g^k 添加服从高斯分布$N\left(0,\sigma^2 I\right)$ 的噪声, 设第k 步迭代的隐私预算为ε_k, 隐私参数为δ_k 时, 得到噪声梯度

$$\xi^k = g^k + b^k,$$

$$b^k \sim N\left(0,\sigma^2 I\right).$$

此时可保证噪声梯度ξ^k 的确定过程满足(ε_k,δ_k) −差分隐私(定义3.3).

好的步长更新策略可以大大减少线搜索花费的时间, 对算法的快速收敛至关重要. Barzilai-Borwein(BB) 步长规则 [32]已经被证明是非常有效的步长策略. 因此, ISTDP算法采用BB步长作为每次迭代的试探步, 即取

$$\alpha_k^{BB} = \frac{\left(d_\theta^k\right)^T d_\theta^k}{\left(d_\theta^k\right)^T d_g^k}$$

其中

$$d_\theta^k = \theta^k - \theta^{k-1},$$

$$d_g^k = g^k - g^{k-1}.$$

对模型(2), ISTDP算法通过求解下述子问题来更新θ:

$$\theta^{k+1} = \arg\min_\theta \left\{ Q(\alpha_k^{BB},\theta^k,\theta) := L\left(\theta^k;D\right) + \langle \xi^k, \theta-\theta^k\rangle + \frac{1}{2\alpha_k^{BB}}\left\|\theta-\theta^k\right\|^2 + R\left(\theta\right)\right\}. \tag{3}$$

问题(3)等价于以下邻近点问题：

$$\theta^{k+1} = \arg\min_\theta \left\{ \frac{1}{2}\left\|\theta-u^k\right\|^2 + \alpha_k^{BB} R\left(\theta\right)\right\}, \tag{4}$$

其中

$$u^k = \theta^k - \alpha_k^{BB}\xi^k.$$

注意到$\|\cdot\|^2$和$R(\theta)$的可分性, 问题(4)可以等价地分解为d个独立的单变量优化问题：

$$\theta_i^{k+1} = \arg\min_{\theta_i}\{h(\theta_i,u_i^k) := \frac{1}{2}(\theta_i-u_i^k)^2 + \alpha_k^{BB}\phi_i(\theta_i)\}, \quad i=1,2,\cdots,d.$$

其中$u_i^k=\theta_i^k-\alpha_k^{BB}\xi_i^k$. 这$d$个单变量优化问题具有相同的结构, 为了简化符号, 我们通过删除上下标来整理成如下统一的形式：

$$t(s)=\arg\min_t\{h(t,s):=\frac{1}{2}(t-s)^2+\gamma\phi(t)\}.$$

根据 [30], 上述问题具有如下闭式解：

$$t(s)=\begin{cases}t_1, & h_1(t_1,s)\le h_2(t_2,s),\\ t_2, & h_1(t_1,s)>h_2(t_2,s).\end{cases}$$

其中

$$t_1=\arg\min_t\left\{h_1(t,s):=\frac{1}{2}(t-s)^2+\lambda\gamma|t|-\frac{t^2}{2\alpha}\quad\text{s.t. }|t|\le\alpha\lambda\right\},$$

$$t_2=\arg\min_t\left\{h_2(t,s):=\frac{1}{2}(t-s)^2+\frac{\alpha(\lambda\gamma)^2}{2}\quad\text{s.t. }|t|\ge\alpha\lambda\right\}.$$

更具体的, 通过计算可以得到

$$t_1=\text{sign}(s)z,$$

$$t_2=\text{sign}(s)\max(\alpha\lambda,|s|),$$

这里 $z=\arg\min_{t\in\mathcal{C}}\left\{\frac{1}{2}(t-|s|)^2+\lambda\gamma t-\frac{t^2}{2\alpha}\right\}$, 其中

$$\mathcal{C}=\begin{cases}\left\{0,\alpha\lambda,\min\left(\alpha\lambda,\max\left(0,\frac{\alpha(|s|-\gamma\lambda)}{\alpha-1}\right)\right)\right\}, & \text{当}\alpha\ne1\text{ 时}\\ \{0,\alpha\lambda\}, & \text{否则.}\end{cases}$$

习惯上也称$t(s)$为收缩阈值函数.

在ISTDP算法中, 除了用BB步长做试探步外, 建议接受的步长需满足以下下降水平 [30]：

$$E[F(\theta^{k+1};D)]\le E[F(\theta^k;D)]-\frac{\beta\alpha_k^{BB}}{2}E[\|\theta^{k+1}-\theta^k\|^2],\ \beta\in(0,1),\tag{5}$$

易见, 在此步长准则之下, 目标函数值的期望是单调递减的.

求解问题(2)的ISTDP算法框架如下：

算法 1 (ISTDP算法)

- **输入:** 数据集$D=\{z_1,z_2,\ldots,z_n\}$.
- **选择参数:** $T>0,\eta\in(0,1),\beta\in(0,1)$和满足$0<\alpha_{\min}<\alpha_{\max}$的$\alpha_{\min},\alpha_{\max}$.
- **初始化:** 选取$\theta^0,\theta^1\in\mathbb{R}^d,\ \theta^0\neq\theta^1$, 计算$g^0=\nabla_\theta L\left(\theta^0;D\right)$, $k=1$;
- **外循环:** while $k<T$, do

1.1. 计算梯度: $g^k=\nabla_\theta L\left(\theta^k;D\right)$, 给梯度添加噪声: $\xi^k=g^k+b^k$, $b^k\sim N(0,\sigma^2I)$;

1.2. 计算BB步长: $d_\theta^k=\theta^k-\theta^{k-1},d_g^k=g^k-g^{k-1}$,

$$\alpha_k^{BB}=\frac{\left(d_\theta^k\right)^T d_\theta^k}{\left(d_\theta^k\right)^T d_g^k},\quad \alpha_k^{BB}:=\min\{\max\{\alpha_k^{BB},\alpha_{\min}\},\alpha_{\max}\};$$

1.3. 内循环: 更新θ^{k+1}

1.3.1. 计算: $\hat{\theta}^{k+1}=\arg\min_\theta\{Q(\alpha_k^{BB},\theta^k,\theta)=L(\theta^k;D)+\langle\xi^k,\theta-\theta^k\rangle+\frac{1}{2\alpha_k^{BB}}\|\theta-\theta^k\|^2+R(\theta)\}$;

1.3.2. 如果$\hat{\theta}^{k+1}$满足:

$$E(F(\theta^k;D)-F(\hat{\theta}^{k+1};D))\geq\frac{\beta\alpha_k^{BB}}{2}E\left(\|\hat{\theta}^{k+1}-\theta^k\|^2\right),$$

则$\theta^{k+1}=\hat{\theta}^{k+1}$, 停止内循环, 转步1.4; 否则, $\alpha_k^{BB}=\eta\alpha_k^{BB}$, 转步1.3.1;

1.4. 更新迭代步: $k:=k+1$, 转步1.1;

- **end** while;
- **输出:** $\theta^{priv}=\theta^T$.

3. 理论分析

本节分析ISTDP算法的理论性质, 将证明它不仅是一个满足差分隐私要求的算法, 同时也是收敛的, 它的输出结果可以满足任何给定的精度要求.

3.1. 算法1的隐私保护性

下面先介绍差分隐私保护的基本概念和结论, 再研究算法1的差分隐私保护性质.

定义 3.1 [6,33] 给定数据集$D,D'\in\mathbb{Z}^n$, 若$|(D\setminus D')\cup(D'\setminus D)|=1$, 则称$D,D'\in\mathbb{Z}^n$为相邻数据集, 记为$D\sim D'$.

注: 若数据集$D,D'\in\mathbb{Z}^n$是相邻的, 则数据集D'可以通过在数据集D上添加或者删除一条记录得到.

定义 3.2 [5,33] 给定任意两个相邻数据集$D,D'\in\mathbb{Z}^n$, 设$\mathcal{M}$为在数据集$D,D'\in\mathbb{Z}^n$上运行的一个随机机制, P_M为$\mathcal{M}$的所有可能的输出结果构成的集合. 若对任意的子集$S\subseteq P_{\mathcal{M}}$皆满足

$$\Pr\left[\mathcal{M}\left(D\right)\in S\right]\leq e^{\varepsilon}\Pr\left[\mathcal{M}\left(D'\right)\in S\right],$$

其中ε称为隐私预算, 则称机制$\mathcal{M}$满足纯差分隐私, 记为ε-*Differential Privacy*, 简称$\varepsilon-\mathrm{DP}$.

差分隐私可以保证数据隐私攻击者无法区分相邻数据集, 进而实现在数据处理过程中不泄露任何有关数据集中的任意一个用户信息的目的.

定义 3.3 [8] 给定任意两个相邻数据集$D, D' \in \mathbb{Z}^n$, 设$\mathcal{M}$为在数据集$D, D' \in \mathbb{Z}^n$ 上运行的一个随机机制, $P_{\mathcal{M}}$为$\mathcal{M}$的所有可能的输出结果构成的集合. 若对任意的子集$S \subseteq P_n$ 皆满足

$$\Pr\left[\mathcal{M}\left(D\right) \in S\right] \leq e^{\varepsilon} \Pr\left[\mathcal{M}\left(D'\right) \in S\right] + \delta,$$

其中ε称为隐私预算, $\delta > 0$称为纯差分隐私的违反程度*(一般取接近于0的值)*, 则称机制$\mathcal{M}$满足近似差分隐私, 记为(ε, δ)-*Differential Privacy*, 简称$(\varepsilon, \delta)-\mathrm{DP}$.

近似的差分隐私概念是对纯差分隐私的推广, 且具有更广泛应用.

定义 3.4 [1] 给定函数$f: \mathbb{Z}^n \to \mathbb{R}^d$, f的L_i模敏感度$\Delta_i\left(f\right)(i=1,2)$ 定义如下

$$\Delta_i\left(f\right) = \max_{D \sim D'} \left\| f\left(D\right) - f\left(D'\right) \right\|_i,$$

其中$D, D' \in \mathbb{Z}^n$为任意的相邻数据集.

事实上, 敏感性刻画了函数f作用在任何相邻数据集下的输出值之间的最大变化. 在不致引起混乱的情况下, 将敏感性写成$\Delta\left(f\right)$. 研究者常常基于函数的敏感度$\Delta\left(f\right)$和隐私预算ε生成噪声, 进而建立相应的差分隐私保护机制.

定理 3.1 给定作用于数据集$D \subset \mathbb{Z}^n$的函数$f: D \to \mathbb{R}^d$, $\forall \varepsilon \in (0,1)$, 若随机机制$\mathcal{M}$满足

$$\mathcal{M}\left(D\right) = f\left(D\right) + \mathcal{N}\left(0, \sigma^2 I\right),$$

其中σ满足

$$\sigma \geq \frac{\Delta_2\left(f\right)}{\varepsilon} \sqrt{2\log\left(\frac{1.25}{\delta}\right)}$$

$\Delta_2\left(f\right)$为函数f的L_2敏感度, $\mathcal{N}\left(0, \sigma^2 I\right)$表示$d$维高斯分布, 则机制$\mathcal{M}$满足$(\varepsilon, \delta)-DP$.

证明. 先讨论$d=1$的情形. 根据模型(2)可知, 此时对实值函数$f: \mathbb{Z}^n \to \mathbb{R}$ 添加符合标准正态分布得噪声. 根据敏感度的定义3.4, $\Delta\left(f\right) = \Delta_1\left(f\right) = \Delta_2\left(f\right) = \max_{D \sim D'} \left|f\left(D\right) - f\left(D'\right)\right|$. 设噪声$x \sim N(0, \sigma^2)$, 根据文献 [33], 此时隐私损失为

$$\left|\log\frac{e^{\left(-1/2\sigma^2\right)x^2}}{e^{(-1/2\sigma^2)(x+\Delta f)^2}}\right| = \left|\log e^{\left(-1/2\sigma^2\right)\left[x^2-(x+\Delta f)^2\right]}\right|$$
$$= \left|-\frac{1}{2\sigma^2}\left[x^2-\left(x^2+2x\Delta f+\Delta f^2\right)\right]\right|$$
$$= \left|-\frac{1}{2\sigma^2}\left[2x\Delta f+(\Delta f)^2\right]\right|.$$

当$|x|<\dfrac{\sigma^2\varepsilon}{\Delta f}-\dfrac{\Delta f}{2}$时, $\left|\log\frac{e^{\left(-1/2\sigma^2\right)x^2}}{e^{(-1/2\sigma^2)(x+\Delta f)^2}}\right|\le\varepsilon$. 为确保$\Pr\left[\left|\log\frac{e^{\left(-1/2\sigma^2\right)x^2}}{e^{(-1/2\sigma^2)(x+\Delta f)^2}}\right|\le\varepsilon\right]\ge 1-\delta$, 需要

$$\Pr\left[|x|\ge\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right]<\delta, \tag{6}$$

从而需要

$$\Pr\left[x\ge\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right]<\frac{\delta}{2}.$$

设$0<\varepsilon<1$, 取$t=\dfrac{\sigma^2\varepsilon}{\Delta f}-\dfrac{\Delta f}{2}$, 则

$$\Pr[x\ge t]=\int_t^\infty\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx\le\frac{\sigma}{\sqrt{2\pi}t}e^{-\frac{t^2}{2\sigma^2}}.$$

因此, 只需

$$\frac{\sigma}{\sqrt{2\pi}}\frac{1}{t}e^{-\frac{t^2}{2\sigma^2}}<\frac{\delta}{2},$$

这等价于需要

$$\frac{\sigma}{t}e^{-\frac{t^2}{2\sigma^2}}<\frac{\sqrt{2\pi}\delta}{2}\Leftrightarrow\frac{t}{\sigma}e^{\frac{t^2}{2\sigma^2}}>\frac{2}{\sqrt{2\pi}\delta}\Leftrightarrow\log(\frac{t}{\sigma})+\frac{t^2}{2\sigma^2}>\log(\frac{2}{\sqrt{2\pi}\delta}).$$

由于$t=\dfrac{\sigma^2\varepsilon}{\Delta f}-\dfrac{\Delta f}{2}$, 上式又等价于需要

$$\log\left[\left(\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right)\frac{1}{\sigma}\right]+\left[\left(\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right)^2\frac{1}{2\sigma^2}\right]>\log\left(\frac{2}{\sqrt{2\pi}\delta}\right)=\log\left(\sqrt{\frac{2}{\pi}}\frac{1}{\delta}\right).$$

为此只需要下述两式同时成立即可

$$\log\left[\left(\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right)\frac{1}{\sigma}\right]>0;\quad\left[\left(\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right)^2\frac{1}{2\sigma^2}\right]\ge\log\left(\sqrt{\frac{2}{\pi}}\frac{1}{\delta}\right). \tag{7}$$

记$c=\dfrac{\varepsilon\sigma}{\Delta f}$, 则$\sigma=\dfrac{c\Delta f}{\varepsilon}$. 由于

$$\begin{aligned}\frac{1}{\sigma}\left(\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right)&=\frac{1}{\sigma}\left[\frac{(c\Delta f)^2}{\varepsilon^2}\frac{\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right]\\&=\frac{1}{\sigma}\left[c^2\left(\frac{\Delta f}{\varepsilon}\right)-\frac{\Delta f}{2}\right]\\&=\frac{\varepsilon}{c\Delta f}\left[c^2\left(\frac{\Delta f}{\varepsilon}\right)-\frac{\Delta f}{2}\right]\\&=c-\frac{\varepsilon}{2c},\end{aligned}$$

注意到$0<\varepsilon\leq 1$, 因此, 当$c\geq\frac{3}{2}$ 时, $c-\frac{\varepsilon}{2c}\geq\frac{7}{6}$, 此时, $\log\left[\left(\dfrac{\sigma^2\varepsilon}{\Delta f}-\dfrac{\Delta f}{2}\right)\dfrac{1}{\sigma}\right]>0$满足.

另一方面,

$$\frac{1}{2\sigma^2}\left(\frac{\sigma^2\varepsilon}{\Delta f}-\frac{\Delta f}{2}\right)^2=\frac{1}{2}\left(c-\frac{\varepsilon}{2c}\right)^2=\frac{1}{2}\left(c^2-\varepsilon+\frac{\varepsilon^2}{4c^2}\right).$$

当$c\geq\dfrac{3}{2}$且$0<\varepsilon\leq 1$ 时, 由于$\left(c^2-\varepsilon+\dfrac{\varepsilon^2}{4c^2}\right)'_c>0$, 所以$c^2-\varepsilon+\dfrac{\varepsilon^2}{4c^2}\geq c^2-\dfrac{8}{9}$. 因此, 要使$\left[\left(\dfrac{\sigma^2\varepsilon}{\Delta f}-\dfrac{\Delta f}{2}\right)^2\dfrac{1}{2\sigma^2}\right]\geq\log\left(\sqrt{\frac{2}{\pi}}\frac{1}{\delta}\right)$, 只需

$$c^2-\frac{8}{9}>2\log\left(\sqrt{\frac{2}{\pi}}\frac{1}{\delta}\right),$$

即只需

$$\begin{aligned}c^2&>2\log\left(\sqrt{\frac{2}{\pi}}\right)+2\log\left(\frac{1}{\delta}\right)+\log\left(e^{\frac{8}{9}}\right)\\&=\log\left(\frac{2}{\pi}\right)+\log\left(e^{\frac{8}{9}}\right)+2\log\left(\frac{1}{\delta}\right)\\&=\log\left(\frac{2}{\pi}e^{\frac{8}{9}}\right)+2\log\left(\frac{1}{\delta}\right),\end{aligned}$$

注意到$\sqrt{\dfrac{2}{\pi}e^{\frac{8}{9}}}<1.25$, 从而只需$c^2>2\log\left(\dfrac{1.25}{\delta}\right)$, 即只需$\sigma\geq\dfrac{\Delta_2(f)}{\varepsilon}\sqrt{2\log\left(\dfrac{1.25}{\delta}\right)}$.

令$R_1=\{x\in R:|x|\leq\frac{c^2\Delta f}{\varepsilon}-\frac{\Delta f}{2}\}$, $R_2=\{x\in R:|x|>\frac{c^2\Delta f}{\varepsilon}-\frac{\Delta f}{2}\}$, 则$R=R_1\cup R_2$. $\forall S\subseteq R$, 定义

$$S_1=\{f(D)+x\mid x\in R_1\},\quad S_2=\{f(D)+x\mid x\in R_2\},$$

则

$$\begin{aligned}\Pr_{x\sim\mathcal{N}(0,\sigma^2)}[f(D)+x\in S]=&\Pr_{x\sim\mathcal{N}(0,\sigma^2)}[f(D)+x\in S_1]+\Pr_{x\sim\mathcal{N}(0,\sigma^2)}[f(D)+x\in S_2]\\ \le&\Pr_{x\sim\mathcal{N}(0,\sigma^2)}[f(D)+x\in S_1]+\delta\\ \le&e^{\varepsilon}\left(\Pr_{x\sim\mathcal{N}(0,\sigma^2)}[f(D')+x\in S_1]\right)+\delta,\end{aligned}$$

从而得证对一维情况下高斯机制满足(ε,δ)-DP.

再讨论$d>1$的情形. 对于$f:D\to R^d$和任意相邻数据集D,D', 记$\Delta f=\Delta_2(f)$和$v=f(D)-f(D')$ (它代表添加的梯度扰动所掩盖的信息), 则向量v满足$\|v\|\le\Delta f$. 设噪声向量$x\sim\mathcal{N}(0,\sigma^2I)$, 则根据文献 [33], 隐私损失为

$$\left|\log\frac{e^{\left(\frac{-1}{2\sigma^2}\right)\|x-\mu\|^2}}{e^{\left(\frac{-1}{2\sigma^2}\right)\|x+v-\mu\|^2}}\right|=\left|\log e^{\left(\frac{-1}{2\sigma^2}\right)\left[\|x-\mu\|^2-\|x+v-\mu\|^2\right]}\right|=\left|\frac{1}{2\sigma^2}(\|x\|^2-\|x+v\|^2)\right|,$$

其中$\mu=(0,\ldots,0)^T\in R^d$. 令$a_1=\dfrac{v}{\|v\|}$, 并将$a_1$扩充成$R^d$中的一组标准正交基$a_1,\ldots,a_d$, 则存在$\beta_1,\cdots,\beta_d\in R$且$\beta_i\sim\mathcal{N}(0,\sigma^2)$ $(i=1,\cdots,d)$, 使得$x=\sum_{i=1}^m\beta_ia_i$. 记$x^{[i]}=\beta_ia_i$. 注意到, 由基的正交性可得$\langle x^{[i]},v\rangle=0$, $i=2,\ldots,d$, 进而$\langle\sum\limits_{i=2}^m x^{[i]},v\rangle=0$. 因此,

$$\begin{aligned}\|x\|^2&=\sum_{i=1}^m\left\|x^{[i]}\right\|^2=\left\|x^{[1]}\right\|^2+\sum_{i=2}^m\left\|x^{[i]}\right\|^2,\\ \|x+v\|^2&=\left\|x^{[1]}+\sum_{i=2}^m x^{[i]}+v\right\|^2=\left\|v+x^{[1]}\right\|^2+\sum_{i=2}^m\left\|x^{[i]}\right\|^2.\end{aligned}$$

由$x^{[1]}=\beta_1b_1=\beta_1\dfrac{v}{\|v\|}$, 得$\left\|v+x^{[1]}\right\|^2=\left(\|v\|(1+\dfrac{\beta_1}{\|v\|})\right)^2=(\|v\|+\beta_1)^2$. 因此,

$$\|x+v\|^2-\|x\|^2=\|v+x^{[1]}\|^2-\|x^{[1]}\|^2=(\|v\|+\beta_1)^2-\beta_1^2=\|v\|^2+2\beta_1\|v\|.$$

因$\|v\|\le\Delta f$, 故

$$\left|\frac{1}{2\sigma^2}(\|x\|^2-\|x+v\|^2)\right|=\left|\frac{1}{2\sigma^2}(\|v\|^2+2\beta_1\|v\|^2)\right|\le\left|\frac{1}{2\sigma^2}[2\beta_1\Delta f+(\Delta f)^2]\right|.$$

由于$\beta_1\sim\mathcal{N}(0,\sigma^2)$, 上式表明隐私损失依赖于一维随机变量$\beta_1$, 即又回到了一维情形. 由前述一维情形结论可得, 维数$d>1$情况下高斯机制依然满足(ε,δ)-DP. □

定理 3.2 设机制$\mathcal{M}_1(\cdot)$满足$(\epsilon,\delta)-DP$, $\mathcal{M}_2$是任意不消耗隐私预算的算法, 则$\mathcal{M}_1$和$\mathcal{M}_2$的复合$\mathcal{M}_2(\mathcal{M}_1(\cdot))$满足$(\epsilon,\delta)-DP$.

证明. 设D 和D'是任意两个相邻数据集. 令Range$(\mathcal{M}_1)=\mathcal{S}$ 表示$\mathcal{M}_1$的值域集合.

(i)若$\mathcal{S}$是离散集合, $\forall t\in$ Range $(\mathcal{M}_2)$, 有

$$\begin{aligned}\Pr\left[\mathcal{M}_2\left(\mathcal{M}_1(D)\right)=t\right] &= \sum_{s\in\mathcal{S}}\Pr\left[\mathcal{M}_1(D)=s\right]\Pr\left[\mathcal{M}_2(s)=t\right]\\ &\leq \sum_{s\in\mathcal{S}}\left(e^{\varepsilon}\Pr\left[\mathcal{M}_1\left(D'\right)=s\right]+\delta\right)\Pr\left[\mathcal{M}_2(s)=t\right]\\ &= e^{\varepsilon}\Pr\left[\mathcal{M}_2\left(\mathcal{M}_1\left(D'\right)\right)=t\right]+\delta\end{aligned}$$

(ii) 若$\mathcal{S}$是连续集合, $\forall t\in$ Range $(\mathcal{M}_2)$,
$\Pr\left[\mathcal{M}_2\left(\mathcal{M}_1(D)\right)=t\right]=\int_{s\in\mathcal{S}}\Pr\left[\mathcal{M}_1(D)=s\right]\Pr\left[\mathcal{M}_2(s)=t\right]\mathrm{d}s$, 采用与上述证明类似的逻辑, 可证明结论成立. □

差分隐私的一个重要特征是它的隐私保护性在多重复合下可以累积. (ε,δ)-DP 的累积定理如下.

引理 3.1 [5] $\forall\varepsilon>0,\delta>0,\delta'>0$, 则$(\varepsilon,\delta)-$DP 机制在$k$ 重自适应复合(即算法的每次执行结果都是由之前结果的自适应所得)下满足$\left(\sqrt{2k\log\left(1/\delta'\right)}\varepsilon+k\varepsilon\left(e^{\varepsilon}-1\right),k\delta+\delta'\right)-$DP.

这一性质使差分隐私适用于模块化的算法设计和分析: 当一个复杂的算法包含一系列满足差分隐私的步骤时, 可以确定该算法整体仍满足差分隐私性.

定理 3.3 设$\nabla_\theta L(\theta;D)$在$\mathbb{Z}^n$上一致有上界$U$, 给定隐私预算$\varepsilon$及参数$\delta$, 若第$k$次迭代的隐私预算$\varepsilon_k=\min\{\frac{\varepsilon}{2\sqrt{2T\log(2/\delta)}},\frac{1}{2}\sqrt{\frac{\varepsilon}{T}}\}$, 参数$\delta_k=\frac{\delta}{2T}$, 噪声水平$\sigma=\frac{2U}{\varepsilon_k}\sqrt{2\log\left(\frac{1.25}{\delta_k}\right)}$, 则算法1 满足$(\varepsilon,\delta)-$DP.

证明. 在算法1的第k步迭代中, 仅有噪声梯度ξ^k 的确定需要用到隐私预算ε_k及参数δ_k. 在噪声服从的高斯分布$\mathcal{N}\left(0,\sigma^2I\right)$ 中选取标准差σ 满足$\sigma=\frac{2U}{\varepsilon_k}\sqrt{2\log\left(\frac{1.25}{\delta_k}\right)}\geq\frac{\Delta_2\left(g^k\right)}{\delta_k}\sqrt{2\log\left(\frac{1.25}{\delta_k}\right)}$, 则由定理3.1可知, 噪声梯度的产生满足$(\varepsilon_k,\delta_k)-$DP.

第k步迭代是噪声梯度的生成与迭代点更新算法的复合, 根据定理3.2, 第k步迭代也满足$(\varepsilon_k,\delta_k)-$DP.

由于算法整体最多迭代T步, 并且每步迭代中噪声梯度都由上一步迭代点来确定, 因此算法满足T重自适应复合. 根据引理3.1, 取$\delta_k=\frac{\delta}{2T}$, $\delta'=\frac{\delta}{2}$, $\varepsilon_k=\min\{\frac{\varepsilon}{2\sqrt{2T\log(2/\delta)}},\frac{1}{2}\sqrt{\frac{\varepsilon}{T}}\}$ 时, 有

$$\sqrt{2T\log\left(1/\delta'\right)}\varepsilon_k+T\varepsilon_k\left(e^{\varepsilon_k}-1\right)\leq\frac{\varepsilon}{2}+2T\varepsilon_k^2\leq\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,\quad T\delta_k+\delta'=\delta,$$

因此, 整个算法的迭代过程满足$(\varepsilon,\delta)-$DP. □

3.2. 算法1的收敛性

本节分析ISTDP算法的收敛性, 将证明, 随着迭代次数的增加它的输出结果可以满足任意给定的精度要求.

下面首先证明对每个k, 内循环1.3有限步即可满足.

引理 3.2 *对每个$k \geq 0$, 内循环停止准则(步1.3.2的不等式)至多$\left\lceil \frac{\log(\beta\alpha_{\max}^2+\alpha_{\max}L)}{-\log\eta}+1 \right\rceil$ 步就会满足.*

证明. 由步1.3.1可知, $Q(\alpha_k^{BB},\hat{\theta}^{k+1},\theta^k) \leq Q(\alpha_k^{BB},\theta^k,\theta^k)$, 即

$$\begin{aligned} & L(\theta^k;D)+\langle g^k+b^k,\hat{\theta}^{k+1}-\theta^k\rangle+\frac{1}{2\alpha_k^{BB}}\|\hat{\theta}^{k+1}-\theta^k\|^2+R(\hat{\theta}^{k+1}) \\ \leq\ & L(\theta^k;D)+R(\theta^k). \end{aligned} \tag{8}$$

因为$\nabla L(\theta;D)$ is L-Lipschitz 连续的, 故

$$L(\hat{\theta}^{k+1};D) \leq L(\theta^k;D)+\langle g^k,\hat{\theta}^{k+1}-\theta^k\rangle+\frac{L}{2}\|\hat{\theta}^{k+1}-\theta^k\|^2. \tag{9}$$

由(8)和(9), 得

$$\begin{aligned} & L(\hat{\theta}^{k+1};D)+R(\hat{\theta}^{k+1})+\langle b^k,\hat{\theta}^{k+1}-\theta^k\rangle+\frac{1}{2}(\frac{1}{\alpha_k^{BB}}-L)\|\hat{\theta}^{k+1}-\theta^k\|^2 \\ \leq\ & L(\theta^k;D)+R(\theta^k). \end{aligned} \tag{10}$$

由

$$F(\hat{\theta}^{k+1};D)=L(\hat{\theta}^{k+1};D)+R(\hat{\theta}^{k+1}),$$

$$F(\theta^k;D)=L(\theta^k;D)+R(\theta^k),$$

和不等式(10), 得

$$E(F(\theta^k;D)-F(\hat{\theta}^{k+1};D)) \geq \frac{1}{2}(\frac{1}{\alpha_k^{BB}}-L)E\left(\|\hat{\theta}^{k+1}-\theta^k\|^2\right). \tag{11}$$

因此, 当$\frac{1}{\alpha_k^{BB}}-L \geq \beta\alpha_k^{BB}$(即$\frac{1}{\alpha_k^{BB}}-\beta\alpha_k^{BB} \geq L$)时, 步1.3.2 满足. 由于$\frac{1}{\alpha}-\alpha\beta$关于$\alpha>0$是单调递减函数且$\lim_{\alpha\to 0+}(\frac{1}{\alpha}-\alpha\beta)=+\infty$, 而当步1.3.2不满足时, $\alpha_k^{BB}=\eta\,\alpha_k^{BB}$是按比例$\eta$ $(0<\eta<1)$ 压缩α_k^{BB}, 故至多有限步之后步1.3.2即可满足. 令$\bar{\alpha}_k^{BB}$ 表示第k步迭代α_k^{BB} 的最终值, 则下面两式同时

成立

$$\frac{1}{\bar{\alpha}_k^{BB}} - \beta\bar{\alpha}_k^{BB} \geq L, \quad \frac{\eta}{\bar{\alpha}_k^{BB}} - \frac{\beta\bar{\alpha}_k^{BB}}{\eta} < L.$$

令m_k 表示第k次外循环时内循环(步1.3)的迭代次数, 该内循环初始步长为α_k^{BB}, 则$\bar{\alpha}_k^{BB} = \alpha_k^{BB}\eta^{m_k}$, 且

$$\frac{1}{\alpha_k^{BB}\eta^{m_k-1}} - \beta\alpha_k^{BB}\eta^{m_k-1} < L.$$

由于$\frac{1}{\alpha} - \alpha\beta$关于$\alpha > 0$是单调递减函数且$\alpha_{\min} \leq \alpha_k^{BB} \leq \alpha_{\max}$, 故

$$\frac{1}{\alpha_{\max}\eta^{m_k-1}} - \beta\alpha_{\max}\eta^{m_k-1} < L.$$

由$0 < \eta < 1$, 得

$$\frac{1}{\alpha_{\max}\eta^{m_k-1}} \leq \beta\alpha_{\max}\eta^{m_k-1} + L \leq \beta\alpha_{\max} + L.$$

因此, $m_k \leq \left\lceil \frac{\log(\beta\alpha_{\max}^2 + \alpha_{\max}L)}{-\log\eta} + 1 \right\rceil$, 证毕. □

注1: 根据引理3.2可知, 在整个算法中步长$\{\alpha_k^{BB}\}$ 均有正的下界和上界:

$$0 < \underline{\alpha}_{\min} := \alpha_{\min}\eta^{\left\lceil \frac{\log(\beta\alpha_{\max}^2 + \alpha_{\max}L)}{-\log\eta} + 1 \right\rceil} \leq \alpha_k^{BB} \leq \alpha_{\max}.$$

注2: 根据内循环1.3可知, 第k此外循环结束时, 下述两式同时成立:

$$\theta^{k+1} = \arg\min_{\theta}\{Q(\alpha_k^{BB}, \theta^k, \theta) = L(\theta^k; D) + \langle \xi^k, \theta - \theta^k\rangle + \frac{1}{2\alpha_k^{BB}}\|\theta - \theta^k\|^2 + R(\theta)\};$$

$$E(F(\theta^k; D) - F(\theta^{k+1}; D)) \geq \frac{\beta\alpha_k^{BB}}{2} E\left(\|\theta^{k+1} - \theta^k\|^2\right).$$

第二式意味着$\{E(F(\theta^k; D)\}_k$ 是单调减少的.

引理 3.3 [34] 设可微函数$f : R^n \to R$的定义域$domf = R^n$, 且梯度$L-$ *Lipschitz* 连续, 则下述不等式成立:

$$f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{L}{2}\|y - x\|^2, \ \forall x, y \in domf.$$

定理 3.4 对模型(2), 算法1满足$E\left[F\left(\theta^{priv};D\right)-F\left(\theta^{*};D\right)\right]\leq\frac{\left\|\theta^{1}-\theta^{*}\right\|^{2}}{2T\underline{\alpha}_{min}}$, 其中$T$ 表示算法1的最大迭代值, $\underline{\alpha}_{\min}$表示$\alpha_k^{BB}$ 的下界, $\theta^{*}\in R^{d}$是$\{\theta^{k}\}$的任一聚点.

证明. 因$\{E\left(F\left(\theta^{k};D\right)\right)\}$ 单调递减有下界, 故$\{E\left(F\left(\theta^{k};D\right)\right)\}$必收敛. 由于$F(\theta;D)$是水平有界的, $\{\theta^{k}\}$必有聚点, 设θ^{*}是$\{\theta^{k}\}$的任意一个聚点, 则当$k\to\infty$时, $E\left(F(\theta^{k};D)\right)\to E\left(F(\theta^{*};D)\right)$.

根据模型(2)可以得出

$$\begin{aligned}
& E\left[F(\theta^{k+1};D)-F(\theta^{*};D)\right]\\
= \;& E\left[L(\theta^{k+1};D)-L(\theta^{*};D)+R(\theta^{k+1})-R(\theta^{*})\right]\\
\overset{(10.1)}{\leq}\; & E\left[L(\theta^{k};D)+\left\langle g^{k},\theta^{k+1}-\theta^{k}\right\rangle+\frac{L}{2}\left\|\theta^{k+1}-\theta^{k}\right\|^{2}-L(\theta^{*};D)+R(\theta^{k+1})-R(\theta^{*})\right]\\
\overset{(10.2)}{\leq}\; & E\left[\left\langle g^{k},\theta^{k}-\theta^{*}\right\rangle+\left\langle g^{k},\theta^{k+1}-\theta^{k}\right\rangle+\frac{L}{2}\left\|\theta^{k+1}-\theta^{k}\right\|^{2}+R(\theta^{k+1})-R(\theta^{*})\right]\\
\overset{(10.3)}{=}\; & E\left[\left\langle \xi^{k},\theta^{k}-\theta^{*}\right\rangle+\left\langle g^{k},\theta^{k+1}-\theta^{k}\right\rangle+\frac{L}{2}\left\|\theta^{k+1}-\theta^{k}\right\|^{2}+R(\theta^{k+1})-R(\theta^{*})\right]. \qquad (12)
\end{aligned}$$

其中不等式(10.1)基于$L(\cdot;D)$ 的强光滑性, 不等式(10.2)基于$L(\cdot;D)$ 的凸性, 等式(10.3) 是因为$E(b^{k})=0$. 由于

$$\begin{aligned}
\theta^{k+1} &= \arg\min_{\theta}\{L(\theta^{k};D)+\langle\xi^{k},\theta-\theta^{k}\rangle+\frac{1}{2\alpha_k^{BB}}\|\theta-\theta^{k}\|^{2}+R(\theta)\}\\
&= \arg\min_{\theta\in\mathbb{R}^{d}}\{\frac{1}{2}\left\|\theta-\theta^{k}+\alpha_k^{BB}\xi^{k}\right\|^{2}+\alpha_k^{BB}R(\theta)\},
\end{aligned}$$

故存在次梯度$\nu^{k+1}\in\partial R(\theta^{k+1})$ [35], 使得 $\theta^{k+1}-\theta^{k}+\alpha_k^{BB}(\xi^{k}+\nu^{k+1})=0$. 根据次梯度的性质, 有

$$R(\theta^{*})-R(\theta^{k+1})\geq\left\langle\nu^{k+1},\theta^{*}-\theta^{k+1}\right\rangle.$$

从而,

$$\begin{aligned}
& R(\theta^{*})-R(\theta^{k+1})+\langle\frac{1}{\alpha_k^{BB}}(\theta^{k+1}-\theta^{k})+\xi^{k},\theta^{*}-\theta^{k+1}\rangle\\
& \geq\langle\frac{1}{\alpha_k^{BB}}(\theta^{k+1}-\theta^{k})+\xi^{k}+\nu^{k+1},\theta^{*}-\theta^{k+1}\rangle=0,
\end{aligned}$$

得

$$R(\theta^{*})-R(\theta^{k+1})+\langle\xi^{k},\theta^{*}-\theta^{k+1}\rangle\geq\frac{1}{\alpha_k^{BB}}\langle\theta^{k}-\theta^{k+1},\theta^{*}-\theta^{k+1}\rangle.$$

进而,

$$\begin{aligned}
&\left\langle \xi^k, \theta^k-\theta^*\right\rangle + R(\theta^{k+1}) - R(\theta^*)\\
=\ &\left\langle \xi^k, \theta^k-\theta^{k+1}\right\rangle + \left\langle \xi^k, \theta^{k+1}-\theta^*\right\rangle + R(\theta^{k+1}) - R(\theta^*)\\
\le\ &\left\langle \xi^k, \theta^k-\theta^{k+1}\right\rangle - \frac{1}{\alpha_k^{BB}}\left\langle \theta^{k+1}-\theta^k, \theta^{k+1}-\theta^*\right\rangle\\
=\ &\left\langle \xi^k, \theta^k-\theta^{k+1}\right\rangle + \frac{\left\langle \theta^k-\theta^{k+1}, \theta^k-\theta^*\right\rangle}{2\alpha_k^{BB}} + \frac{\left\langle \theta^{k+1}-\theta^*, \theta^k-\theta^*\right\rangle}{2\alpha_k^{BB}} - \frac{\left\langle \theta^{k+1}-\theta^k, \theta^{k+1}-\theta^*\right\rangle}{2\alpha_k^{BB}}\\
&-\frac{\left\langle \theta^k-\theta^*, \theta^{k+1}-\theta^*\right\rangle}{2\alpha_k^{BB}} - \frac{\left\langle \theta^{k+1}-\theta^k, \theta^{k+1}-\theta^*\right\rangle}{2\alpha_k^{BB}} - \frac{\left\langle \theta^{k+1}-\theta^k, \theta^*-\theta^k\right\rangle}{2\alpha_k^{BB}}\\
=\ &\left\langle \xi^k, \theta^k-\theta^{k+1}\right\rangle + \frac{\left\langle \theta^k-\theta^*, \theta^k-\theta^*\right\rangle}{2\alpha_k^{BB}} - \frac{\left\langle \theta^{k+1}-\theta^*, \theta^{k+1}-\theta^*\right\rangle}{2\alpha_k^{BB}} - \frac{\left\langle \theta^{k+1}-\theta^k, \theta^{k+1}-\theta^k\right\rangle}{2\alpha_k^{BB}}\\
=\ &\left\langle \xi^k, \theta^k-\theta^{k+1}\right\rangle + \frac{\left\|\theta^k-\theta^*\right\|^2}{2\alpha_k^{BB}} - \frac{\left\|\theta^{k+1}-\theta^*\right\|^2}{2\alpha_k^{BB}} - \frac{\left\|\theta^{k+1}-\theta^k\right\|^2}{2\alpha_k^{BB}}. \qquad (13)
\end{aligned}$$

结合(12)和(13)两个不等式, 有

$$\begin{aligned}
&E\left[F\left(\theta^{k+1};D\right)-F\left(\theta^*;D\right)\right]\\
\le\ &E\left[\left\langle \xi^k-g^k, \theta^k-\theta^{k+1}\right\rangle - \frac{1-\alpha_k^{BB}L}{2\alpha_k^{BB}}\left\|\theta^{k+1}-\theta^k\right\|^2 + \frac{\left\|\theta^k-\theta^*\right\|^2-\left\|\theta^{k+1}-\theta^*\right\|^2}{2\alpha_k^{BB}}\right]\\
\overset{(12.1)}{\le}\ &E\left[\frac{\alpha_k^{BB}}{2(1-\alpha_k^{BB}L)}\left\|\xi^k-g^k\right\|^2 + \frac{\left\|\theta^k-\theta^*\right\|^2-\left\|\theta^{k+1}-\theta^*\right\|^2}{2\alpha_k^{BB}}\right]\\
=\ &E\left[\frac{\left\|\theta^k-\theta^*\right\|^2-\left\|\theta^{k+1}-\theta^*\right\|^2}{2\alpha_k^{BB}}\right] \overset{(12.2)}{\le} E\left[\frac{\left\|\theta^k-\theta^*\right\|^2-\left\|\theta^{k+1}-\theta^*\right\|^2}{2\underline{\alpha}_{min}}\right], \qquad (14)
\end{aligned}$$

其中不等式(12.1)由$\left\|\frac{\alpha_k^{BB}}{1-\alpha_k^{BB}L}(\xi^k-g^k)-(\theta^k-\theta^{k+1})\right\|^2 \ge 0$ 的展开式整理所得; 由引理3.1 的注1可知, α_k^{BB}有下界$\underline{\alpha}_{\min}>0$, 从而不等式(12.2) 成立.

由不等式(14)结合引理3.1的注2 ($\{E[F\left(\theta^{k+1};D\right)]\}$的单调递减性质), 有

$$\begin{aligned}
&E\left[F\left(\theta^T;D\right)-F\left(\theta^*;D\right)\right]\\
\le\ &E\left[\sum_{k=0}^{T-1}\frac{F\left(\theta^{k+1};D\right)}{T}-F\left(\theta^*;D\right)\right] = E\left[\frac{1}{T}\sum_{k=0}^{T-1}\left[F(\theta^{k+1};D)-F\left(\theta^*;D\right)\right]\right]\\
\le\ &E\left[\frac{\left\|\theta^1-\theta^*\right\|^2-\left\|\theta^{k+1}-\theta^*\right\|^2}{2T\underline{\alpha}_{min}}\right] \le E\left[\frac{\left\|\theta^1-\theta^*\right\|^2}{2T\underline{\alpha}_{min}}\right] = \frac{\left\|\theta^1-\theta^*\right\|^2}{2T\underline{\alpha}_{min}},
\end{aligned}$$

即

$$E\left[F\left(\theta^{priv};D\right)-F\left(\theta^*;D\right)\right] = E\left[F\left(\theta^T;D\right)-F\left(\theta^*;D\right)\right] \le \frac{\left\|\theta^1-\theta^*\right\|^2}{2T\underline{\alpha}_{min}}. \qquad \square$$

注: 根据定理3.4, 给定精度$\epsilon > 0$, 要达到该精度算法迭代次数T应满足$T > \left\lceil \frac{m^2(\theta^1)}{2\alpha_{\min}\epsilon} \right\rceil$, 其中$m(\theta^1)$ 表示水平集$L(\theta^1) := \{\theta \in R^d : F(\theta; D) \leq F(\theta^1; D)\}$的直径.

4. 总结

本文研究了基于梯度扰动的差分隐私保护下带有非凸正则的经验风险极小化问题. 基于以BB步长为试探步的线搜素和迭代收缩阈值算法提出了ISTDP算法, 并证明了ISTDP算法既满足差分隐私保护要求又可以收敛到任意给定的精度, 是一种可以实现隐私保护的机器学习优化算法. 下一步将通过数值实验和算例进一步检验算法的实际效果.

基金项目

国家自然科学基金项目(11861020, 12261020)、贵州省高层次留学人才创新创业择优资助重点项目([2018]03)、贵州省科技计划项目(ZK[2021]009, [2018]5781)、贵州省青年科技人才成长项目([2018]121).

参考文献

[1] Li, N., Lyu, M., Su, D. and Yang, W. (2016) Differential Privacy: From Theory to Practice. In: *Synthesis Lectures on Information Security Privacy and Trust*, Springer, Cham, 1-138. https://doi.org/10.1007/978-3-031-02350-7

[2] Dwork, C., McSherry, F., Nissim, K. and Smith, A (2016) Calibrating Noise to Sensitivity in Private Data Analysis. *Proceedings of the Third Conference on Theory of Cryptography*, **7**, 17-51. https://doi.org/10.29012/jpc.v7i3.405

[3] Jiang, B., Li, J. and Yue, G. (2021) Differential Privacy for Industrial Internet of Things: Opportunities, Applications, and Challenges. *IEEE Internet of Things Journal*, **8**, 10430-10451. https://doi.org/10.1109/JIOT.2021.3057419

[4] El Ouadrhiri, A. and Abdelhadi, A. (2022) Differential Privacy for Deep and Federated Learning: A Survey. *IEEE Access*, **10**, 22359-22380. https://doi.org/10.1109/ACCESS.2022.3151670

[5] Dwork, C., Rothblum, G. and Vadhan, S. (2010) Boosting and Differential Privacy. *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, Las Vegas, NV, 23-26 October 2010, 23-26. https://doi.org/10.1109/FOCS.2010.12

[6] Girgis, A., Data, D. and Diggavi, S. (2021) Shuffled Model of Differential Privacy in Federated Learning. *International Conference on Artificial Intelligence and Statistics, PMLR*, **130**, 2521-2529.

[7] Adnan, M., Kalra, S. and Cresswell, J. (2022) Federated Learning and Differential Privacy for Medical Image Analysis. *Scientific Reports*, **12**, Article No. 1953. https://doi.org/10.1038/s41598-022-05539-7

[8] Abadi, M., Andy, C. and Goodfellow, L. (2016) Deep Learning with Differential Privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, Vienna, 24-28 October 2016, 308-318. https://doi.org/10.1145/2976749.2978318

[9] Chaudhuri, K. and Monteleoni, C. (2011) Differentially Private Empirical Risk Minimization. *Journal of Machine Learning Research*, **12**, 1069-1109.

[10] Zhang, J., Zheng, K., Mou, W. and Wang, L. (2017) Efficient Private ERM for Smooth Objectives. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 3922-3928. https://doi.org/10.24963/ijcai.2017/548

[11] Chaudhuri, K. and Monteleoni, C. (2009) Privacy-Preserving Logistic Regression. *Advances in Neural Information Processing Systems*, **21**, 289-296.

[12] Wang, Y., Wang, Y. and Singh, A. (2015) Differentially Private Subspace Clustering. *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems*, **28**, 1000-1008.

[13] Bassily, R., Smith, A. and Thakurta, A. (2014) Private Empirical Risk Minimization: Efficient Algorithms and Tight Error Bounds. *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, Philadelphia, PA, 18-21 October 2014, 464-473. https://doi.org/10.1109/FOCS.2014.56

[14] Beimel, A., Kasiviswanathan, S. and Nissim, K. (2010) Bounds on the Sample Complexity for Private Learning and Private Data Release. In Micciancio, D., Ed., *Theory of Cryptography. TCC 2010. Lecture Notes in Computer Science*, Vol. 5978, Springer, Berlin, Heidelberg, 437-454. https://doi.org/10.1007/978-3-642-11799-2_26

[15] Williams, O. and Mcsherry, F. (2010) Probabilistic Inference and Differential Privacy. *Advances in Neural Information Processing Systems*, **23**, 2451-2459.

[16] Wang, D., Ye, M. and Xu, J. (2017) Differentially Private Empirical Risk Minimization Revisited: Faster and More General. *31st Advances in Neural Information Processing Systems*, **30**, 2722-2731.

[17] Fan, J. and Li, R. (2001) Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties. *Journal of the American Statistical Association*, **96**, 1348-1360. https://doi.org/10.1198/016214501753382273

[18] 罗孝敏, 彭定涛, 两个绝对值优化问题解的等价性[J]. 重庆工商大学学报: 自然科学版, 2020, 37(5): 37-42.

[19] Zhang, C. (2010) Nearly Unbiased Variable Selection under Minimax Concave Penalty. *Annals of Statistics*, **38**, 894-942. https://doi.org/10.1214/09-AOS729

[20] Ong, C. and An, L. (2013) Learning Sparse Classifiers with Difference of Convex Functions Algorithms. *Optimization Methods and Software*, **28**, 830-854. https://doi.org/10.1080/10556788.2011.652630

[21] Peleg, D. and Meir, R. (2008) A Bilinear Formulation for Vector Sparsity Optimization. *Signal Processing*, **88**, 375-389. https://doi.org/10.1016/j.sigpro.2007.08.015

[22] Thi, H., Dinh, T., Le, H. and Vo, X. (2015) DC Approximation Approaches for Sparse Optimization. *European Journal of Operational Research*, **244**, 26-46. https://doi.org/10.1016/j.ejor.2014.11.031

[23] Zhang, T. (2013) Multi-Stage Convex Relaxation for Feature Selection. *Bernoulli*, **19**, 2277-2293. https://doi.org/10.3150/12-BEJ452

[24] 彭定涛, 唐琦, 张弦, 组稀疏优化问题精确连续Capped-L1松弛[J]. 数学学报, 2022, 65(2): 243-262.

[25] Zhang, X. and Peng, D. (2022) Solving Constrained Nonsmooth Group Sparse Optimization via Group Capped-L1 Relaxation and Group Smoothing Proximal Gradient Algorithm. *Computational Optimization and Applications*, **83**, 801-844. https://doi.org/10.1007/s10589-022-00419-2

[26] Bian, W. and Chen, X. (2017) Optimality and Complexity for Constrained Optimization Problems with Nonconvex Regularization. *Mathematics of Operations Research*, **42**, 1063-1084. https://doi.org/10.1287/moor.2016.0837

[27] Chartrand, R. and Staneva, V. (2008) Restricted Isometry Properties and Nonconvex Compressive Sensing. *Inverse Problems*, **24**, Article 035020. https://doi.org/10.1088/0266-5611/24/3/035020

[28] Peng, D. and Chen, X. (2020) Computation of Second-Order Directional Stationary Points for Group Sparse Optimization. *Optimization Methods and Software*, **35**, 348-376. https://doi.org/10.1080/10556788.2019.1684492

[29] 罗孝敏, 彭定涛, 张弦, 基于MCP正则的最小一乘回归问题研究[J]. 系统科学与数学, 2021, 41(8): 2327-2337.

[30] Gong, P. and Zhang, C. (2013) A General Iterative Shrinkage and Thresholding Algorithm for Non-Convex Regularized Optimization Problems. *Proceedings of the 30th International Conference on International Conference on Machine Learning*, **28**, 37-45.

[31] Jain, P. and Thakurta, A. (2013) Differentially Private Learning with Kernels. *Proceedings of the 30th International Conference on Machine Learning*, **28**, 118-126.

[32] Barzilai, J. and Borwein, J.M. (1988) Two-Point Step Size Gradient Methods. *IMA Journal of Numerical Analysis*, **8**, 141-148. https://doi.org/10.1093/imanum/8.1.141

[33] Dwork, C. and Roth, A. (2014) The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science*, **9**, 211-407. https://doi.org/10.1561/0400000042

[34] 刘浩洋, 户将, 李勇锋, 文再文. 最优化: 建模、算法与理论[M]. 北京: 高教出版社, 2021.

[35] 高岩. 非光滑优化[M]. 北京: 科学出版社, 2008.