

基于非光滑约束的Langevin算法

曹嘉锋¹, 李春晔^{2*}, 王 薇²

¹浙江师范大学数学科学学院, 浙江 金华

²嘉兴学院数据科学学院, 浙江 嘉兴

收稿日期: 2023年12月17日; 录用日期: 2024年1月11日; 发布日期: 2024年1月17日

摘 要

贝叶斯推断求解反问题越来越流行, 得益于它对解不确定性的诠释。在解决实际问题中, 算法的计算时间受到一定的关注, 而基于Langevin方程的算法利用了一阶导数信息, 具有采样效率高的优点, 在采样算法中被广泛使用。对于含有急剧跳跃先验信息的图像问题, 贝叶斯框架下先验的选取尤为重要, 而TV正则化方法等非光滑约束可以很好刻画先验信息。本文回顾了非光滑约束与高斯先验相结合的混合先验, 同时重述了近端Langevin算法, 并在此基础上提出了基于非光滑约束的对偶Langevin算法。最后应用于CT成像问题, 数值结果表明, 我们提出算法是有效的, 能够更好地利用非光滑约束刻画解。

关键词

贝叶斯推断, 蒙特卡罗, 全变差, 随机微分方程, 高斯测度

Langevin Algorithms Based on Non-Smooth Constraints

Jiafeng Cao¹, Chunye Li^{2*}, Wei Wang²

¹School of Mathematical Sciences, Zhejiang Normal University, Jinhua Zhejiang

²College of Data Science, Jiaying University, Jiaying Zhejiang

Received: Dec. 17th, 2023; accepted: Jan. 11th, 2024; published: Jan. 17th, 2024

Abstract

Bayesian inference for solving inverse problems is becoming increasingly popular, thanks to its interpretation of solution uncertainty. In solving practical problems, the computational time of algorithms has received certain attention, and algorithms based on Langevin equation utilize first-order

*通讯作者。

derivative information, which has the advantage of high sampling efficiency and is widely used in sampling algorithms. For image problems with sharp jump prior information, the selection of prior information in the Bayesian framework is particularly important, and non-smooth constraints such as TV regularization methods can effectively characterize prior information. This article reviews the hybrid prior combining non-smooth constraints with Gaussian prior, while reiterating the proximal Langevin algorithm, and proposes a dual Langevin algorithm based on non-smooth constraints. Finally, it is applied to the CT imaging problem, and numerical results show that our proposed algorithm is effective and can better utilize non-smooth constraints to characterize the solution.

Keywords

Bayesian Inference, Monte Carlo, Total Variation, Stochastic Differential Equation, Gaussian Measure

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

本文考虑求解不适定问题

$$Au = y,$$

其中 $A: X \rightarrow Y$ 是 Hilbert 空间 X 到 Hilbert 空间 Y 的有界线性算子。在实际应用中, 真实数据很难获得, 而常常只能得到带有噪声的扰动数据 $y = y^\dagger + \varepsilon$, ε 是服从正态分布 $N(0, \Sigma)$ 的噪声, Σ 是对称正定算子。我们关注的问题的不适定性是指解不连续依赖于右端数据。正则化方法是求解不适定问题的基本策略, 其利用具体问题的某些附加信息, 引进约束泛函给出一个逼近原问题的稳定的近似解[1] [2]。当解为不均匀介质, 包含不连续点、脉冲尖点等结构时, 经典的正则化方法因过光滑而不再适用, 一些非光滑约束如全变差约束(TV 正则化)、稀疏约束被引入[3]。

对一般凸约束 $R: X \rightarrow (-\infty, \infty]$ 构造约束泛函

$$u = \arg \min_u \left\{ \frac{1}{2} \|Au - y\|^2 + aR(u) \right\},$$

传统的优化方法只能给出反问题的“点估计”, 当物理模型或测量数据存在随机性时, 我们希望获得更多的统计信息。相应统计反演(贝叶斯推断)结合后验概率分布因其能提供更多重构解的统计信息而得到广泛关注。

在贝叶斯视角下, 统计反演问题中对反演解的高斯先验、TV 先验、Besov 先验与 Tikhonov 型正则化方法的平方约束、TV 约束、稀疏约束相呼应[4], 反演解作为随机变量可以获得后验分布的采样。对于后验分布的诠释有两种: 最大后验(MAP)估计和条件均值(CM)估计。MAP 估计指选取最可能的值, 即密度函数取值最大的地方, 等价于优化问题的求解。CM 估计选取后验概率分布的均值, 其对应于一个积分问题, 实际应用受限于高维积分的庞大计算量, 通常通过马尔可夫链蒙特卡罗(MCMC)方法来实现。

在处理高维问题, 甚至函数空间时, 相应的快速算法与理论分析是研究的热点。文献[5]提出了一系列函数空间的 MCMC 方法, 包括预处理 Crank-Nicolson (pCN)算法和基于 Langevin 方程的预处理 Crank-Nicolson (pCNL)算法, 这两个算法的优势在于结果与问题的离散维数无关, 且 pCNL 算法进一步

引入导数信息,加快了 Markov 链的收敛速度。这些方法更多限定在 Gauss 先验,而涉及一般非光滑约束的算法讨论不多。对全变差(TV)先验,文献[6]指出随着离散化维度的提高 TV 先验产生的后验分布不具有边缘保持性,不会收敛到一个定义良好的无限维测度。相应地,文献[4]提出了以高斯测度作为参考测度并与 TV 项相混合的 TG 先验,并构造了分裂 pCN 算法。这些方法通过定义提议分布和接受率来构造满足细致平衡条件的转移核,引入接受率控制采样来消除渐近误差,相应的 Langevin 算法被称作 MALA 算法。其中,接受率一定程度上会影响采样效率。

回顾涉及 TV 先验的 MCMC 算法,TV 约束并未出现在提案(proposal)中,仅仅在接受率的计算中有所体现[6]。而文献[7][8]基于 Langevin 随机微分方程对一般非光滑凸约束,给出了相应的 Langevin 系统,并引入 Moreau 近似获得近似的后验,使得约束部分在提案中有所体现。特别地,基于 Langevin 随机方程,构造不使用接受率控制采样点的蒙特卡罗方法,称之为 ULA 算法。文献[9]理论分析了 ULA 算法的收敛性。因为 ULA 算法不涉及接受-拒绝步,因此在一定的收敛条件下,ULA 算法相比 MALA 算法具有采样效率更高的优点。

本文基于非光滑先验综述现有的采样算法,主要关注 ULA 算法。特别地,基于凸分析理论,我们将目光聚焦在随机微分方程的构造,提出一种新的 ULA 算法,期望获得基于非光滑约束快速收敛的 Langevin 算法,使其适用于高维问题,并在求解 CT 成像问题验证方法的优势。

2. 基于非光滑约束的 Langevin 算法

在实际问题中,如在实际的地震波形反演中,由于地下介质的复杂,常常会遇到解为不连续函数或含尖点的函数,为此为了使重构解具有这些特征,一些非光滑先验被引入。如 Ω 为 \mathbb{R}^d 上的有界区域,全变分约束

$$R(u) := \int_{\Omega} |Du| := \sup \left\{ \int_{\Omega} u(\operatorname{div} f) \, dw : f \in C_0^1(\Omega; \mathbb{R}^N), \|f\|_{L^\infty(\Omega)} \leq 1 \right\},$$

可以很好地反演出跳跃性较大的不连续参数部分。

记 $\mu_0 = N(0, C)$ 为 X 上的 Gauss 测度,其中 C 为对称正定的协方差算子。结合非光滑约束 $R(u)$ 先验,引入混合先验 μ_{pr} , 有

$$\frac{d\mu_{pr}}{d\mu_0}(u) \propto \exp(-\alpha R(u)),$$

相应的后验分布 μ_{post} 的关于 μ_0 的 Randon-Nikodym 导数为

$$\frac{d\mu_{post}}{d\mu_0}(u) \propto \exp(-\Phi(u) - \alpha R(u)),$$

其中 $\Phi(u) = \frac{1}{2} \|Au - y\|_{\Sigma}^2 = \frac{1}{2} \left\| \Sigma^{-\frac{1}{2}}(Au - y) \right\|_2^2$, $\alpha > 0$ 。

本文主要关心求解高维问题 $Au = y$ ($A: \mathbb{R}^N \rightarrow \mathbb{R}^M$) 下基于过阻尼 Langevin 方程的相关算法。对于目标分布

$$\pi(u) \propto \exp(-J(u)),$$

若 $J: \mathbb{R}^N \rightarrow \mathbb{R}$ 是光滑的,则有相应的 Langevin 随机微分方程

$$du = \nabla \log \pi(u) dt + \sqrt{2} dw = -\nabla J(u) + \sqrt{2} dw, \quad (1)$$

其中 $w \sim N(0, I_N)$, I_N 表示 N 维单位矩阵,下同。在 $\int_{\mathbb{R}^N} \exp(-J(u)) du < \infty$ 的条件下, $\pi(u)$ 是关于(1)式

的唯一不变分布[8], 于是当 $t \rightarrow \infty$, 可对目标分布进行采样。对时间 t 应用 Euler 离散, 可得马氏链

$$u_{k+1} = u_k - \delta \nabla J(u_k) + \sqrt{2\delta} w_{k+1}, \quad (2)$$

其中 $w_{k+1} \sim N(0, I_N)$, 下同。可以看到(1)式的构造基于 J 的光滑性, 当面临 TV 等非光滑约束时并不适用。

近端 Langevin 算法

涉及 TV 等非光滑约束时, 文献[10]引入近端算子 R^γ 来近似函数 R

$$R^\gamma(u) = \min_v \left\{ R(v) + \frac{1}{2\gamma} \|u - v\|^2 \right\}, \quad (3)$$

其中 $\gamma > 0$ 。当 $\gamma \rightarrow 0$ 时, R^γ 点收敛到 R , 并且 R^γ 连续可微

$$\nabla R^\gamma(u) = \frac{1}{\gamma} (u - \text{prox}_R^\gamma(u)),$$

这里

$$\text{prox}_R^\gamma(u) = \arg \min_v \left\{ R(v) + \frac{1}{2\gamma} \|u - v\|^2 \right\}$$

表示 R 的近端算子。考虑由混合先验下近似的后验概率密度

$$\pi^\gamma(u) \propto \exp\left(-\Phi(u) - \frac{1}{2} \|u\|_C^2 - \alpha R^\gamma(u)\right),$$

其中 $\|u\|_C^2 = \left\| C^{-\frac{1}{2}} u \right\|_2^2$ 。结合对称正定算子 K 的 Langevin 随机微分方程[5]

$$du = -K(C^{-1}u + \nabla\Phi(u) + \alpha\nabla R^\gamma(u))dt + \sqrt{2K}dw,$$

取 $K = C$ 并对 t 使用 Euler 方法离散, 设步长为 δ , 可得马氏链 $(u_k)_{k \geq 0}$

$$u_{k+1} = (1 - \delta)u_k - \delta C(\nabla\Phi(u_k) + \alpha\nabla R^\gamma(u_k)) + \sqrt{2\delta C}w_{k+1}. \quad (4)$$

综上, 基于 Euler 离散的近端 Langevin (PxEuL)算法具体流程如下:

算法 1: 基于 Euler 离散的近端 Langevin (PxEuL)算法

步骤 1: 设时刻 k , 有 u_k , 随机抽取 $w_{k+1} \sim N(0, I_N)$;

步骤 2: 计算 $\text{prox}_R^\gamma(u_k)$;

步骤 3: 依据(5)式计算 u_{k+1} ;

步骤 4: 重复步骤 1~3 直至产生足够多的样本。

3. 对偶 Langevin 算法

这一节, 我们基于凸分析理论再次考虑非光滑约束 R 与 Gauss 先验的混合约束, 构造新的随机微分方程, 进而获得新型 ULA 算法。令约束

$$f(u) = R(u) + \frac{1}{2\eta} \|u\|_2^2,$$

注意到 R 非光滑, 因而 f 是非光滑的。当 R 为一般凸函数时, f 具有强凸性, 其 Fenchel 共轭

$$f^*(\xi) = \sup_u \{ \langle u, \xi \rangle - f(u) \}$$

是 Frechet 可导, 且导数 ∇f^* 满足 Lipschitz 连续。按照凸分析理论可知

$$u = \nabla f^*(\xi) \Leftrightarrow u = \arg \min_u \{ f(u) - \langle \xi, u \rangle \}。 \quad (5)$$

实际上, 对 $f(u) = R(u) + \frac{1}{2\eta} \|u\|_2^2$, 我们有

$$\nabla f^*(\xi) = \arg \min_u \left\{ \frac{1}{2\eta} \|u - \eta\xi\|_2^2 + R(u) \right\},$$

对比近端近似算子有 $\nabla f^*(\xi) = \text{prox}_R^\eta(\eta\xi)$ [8]。近些年, 基于非光滑强凸函数成功的结合到迭代正则化方法[11] [12], 渐近正则化方法中[13] [14], 利用凸分析的理论基础证明了这些方法的收敛性与正则性, 数值结果表明 TV 等非光滑约束成功地刻画了反演解的特征。

受文献[13]的启发, 我们针对约束问题

$$\min \{ f(u) : Au = y, u \in X \}, \quad (6)$$

定义 Lagrangian 函数

$$L(u, \lambda) = f(u) - \langle \lambda, Au - y \rangle,$$

及其对偶函数

$$V(\lambda) = f^*(A^*\lambda) - \langle \lambda, y \rangle,$$

进而关注分布

$$\pi^f(\lambda) \propto \exp(-V(\lambda))。$$

假设 $V: \mathbb{R}^N \rightarrow \mathbb{R}$ 是连续可微的。考虑 $\pi^f(\lambda)$ 对应的 Langevin 随机微分方程, 取 $K = I$ 有

$$d\lambda = -\nabla V(\lambda)dt + \sqrt{2}dw,$$

实际上 $V(\lambda)$ 可导, 且 $\nabla V(\lambda) = A\nabla f^*(A^*\lambda) - y$, 故有

$$d\lambda = -\left(A\nabla f^*(A^*\lambda) - y\right)dt + \sqrt{2}dw。$$

我们将 Langevin 算法与对偶梯度流方法[13]相结合, 提出了新的随机微分方程

$$\begin{cases} d\lambda = -(Au - y)dt + \sqrt{2}dw \\ u = \nabla f^*(A^*\lambda) \end{cases}, \quad (7)$$

应用 Euler 离散, 设步长为 δ , 则有马氏链 $(\lambda_k)_{k \geq 0}$

$$\begin{cases} \lambda_{k+1} = \lambda_k - \delta(Au_k - y) + \sqrt{2\delta}w_{k+1} \\ u_{k+1} = \nabla f^*(A^*\lambda_{k+1}) \end{cases}。 \quad (8)$$

综上, 基于非光滑约束的对偶 Langevin (DuEuL)算法具体流程如下:

算法 2: 基于非光滑约束的对偶 Langevin (DuEuL)算法

步骤 1: 设时刻 k , 有 λ_k 和 u_k , 随机抽取 $w_{k+1} \sim N(0, I_N)$;

步骤 2: 依据(9)式计算 λ_{k+1} ;

步骤 3: 计算 $u_{k+1} = \nabla f^*(A^* \lambda_{k+1})$;

步骤 4: 重复步骤 1~3 直至产生足够多的样本。

4. CT 成像数学模型及其离散

4.1. CT 成像问题

CT 成像是现代医学一种常见的影像扫描技术, 它通过获取人体头颅、心肺、腹部等内部器官的二维断层图像来反映病人的病症, 具有密度分辨率高, 无创伤等特点, 因而被广泛用于临床。传统的 CT 重建算法是以检测目标为中心, 通过 X 射线旋转扫描, 采集目标在不同视角下的投影数据, 利用图形重建算法还原出检测目标的断层图像。X 射线穿过组织体, 射线强度会发生衰退, 这一过程可以用 Radon 变换来描述。记矩形区域 $\Omega \subset \mathbb{R}^2$, 相应的非负衰减系数 $u: \Omega \rightarrow [0, +\infty)$, 当组织体包含在矩形区域中, 即 $\Omega_1 \subset \Omega$, 当 $x \in \Omega \setminus \Omega_1$ 时, $f(x) = 0$, 有

$$A: L^2(\Omega) \rightarrow L^2(D),$$

$$(Au)(\theta, s) = \int_{L(\theta, s)} u(l) dl,$$

其中 $L(\theta, s) = \{u \in \mathbb{R}^2 : u_1 \cos \theta + u_2 \sin \theta = s\}$ 为 X 射线路径, $D = \{(\theta, s) : -r \leq s \leq r, 0 \leq \theta \leq \pi\}$ 为投影空间。

在数值模拟中, 我们使用在 1 度到 180 度之间的 30 个均匀分布的投影数据, 每个投影包含 362 行光束, 所探测的图像离散后为 256×256 。我们借助文献[15]的 MATLAB 包 AIRTOOL 中函数 `parallelomo` 来实现问题的离散, 进而获得不适定线性问题 $Au = y$, 其中 A 是 $M \times N$ 的稀疏矩阵, $M = 10860$ 和 $N = 66536$ 。设图像 u 在两方向上的离散维数分别为 n 和 m , 则涉及的非光滑约束 TV 约束的定义为

$$\|u\|_{TV} = \sum_{i=1}^n \sum_{j=1}^m |\nabla_1 u| + |\nabla_2 u|,$$

其中 $(\nabla_1 u, \nabla_2 u)$ 表示

$$(\nabla_1 u)_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j}, & i = 1, \dots, n-1; j = 1, \dots, m-1 \\ u_{n,j+1} - u_{n,j}, & i = n; j = 1, \dots, m-1 \end{cases};$$

$$(\nabla_2 u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & i = 1, \dots, n; j = 1, \dots, m-1 \\ u_{i+1,m} - u_{i,m}, & i = 1, \dots, n-1; j = m \end{cases}.$$

4.2. 数值算例

在数值模拟中, 真实图像选取为脑部图, 如图 1 所示, 对投影数据添加高斯噪声 $y = y^\dagger + \varepsilon$, 其中: $\varepsilon \sim N(0, \sigma_1^2 I_N)$, 其中: σ_1 取测量数据 y 中最大元素的 0.5%。PxEuL 算法混合先验中 $C = \sigma_2^2 I_N$ 。

在算法的实现过程当中, PxEuL 算法涉及求解

$$\text{prox}_R^\gamma(u) = \arg \min_v \left\{ R(v) + \frac{1}{2\gamma} \|u - v\|^2 \right\}$$

以及 DuEuL 算法涉及求解

$$\nabla f^*(\xi) = \arg \min_u \left\{ \frac{1}{2\eta} \|u - \eta\xi\|_2^2 + R(u) \right\} = \text{prox}_R^\eta(\eta\xi)$$

的部分, 通过主对偶混合梯度方法[10] [16]进行有效求解。

我们基于非光滑约束对比贝叶斯推断方法

- PxEuL-TV 算法: $\gamma=5$, $\delta=1 \times 10^{-4}$, $\alpha=1$, $\sigma_2=0.1$ 。
- DuEuL-TV 算法: α 固定为 1, 分别取不同的 η , 并调整步长 δ 的大小。

其中 PxEuL 为已有的近端算法, 我们只选取了一种参数, 而本文的重点为 DuEuL 算法, 实验中共选取了四组参数, 需要注意的是, 为了使马氏链达到平稳状态, 随着 η 的变大, 相应的步长 δ 应变小。

为了衡量图像的重构效果, 我们使用了相对误差(RelErr)和峰值信噪比(PSNR)等指标, 设 u 是重构得到的解, u^* 是真解, 则它们的定义如下

$$\text{RelErr} = \frac{\|u - u^*\|_2}{\|u^*\|_2},$$

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}} [\text{dB}],$$

其中 MSE 为重构解 u 的均方误差。除了 RelErr 和 PSNR 指标外, 我们还使用结构相似性(SSIM)指标来衡量算法结果和原始图像的相似性。

实际算例中, PxEuL 和 DuEuL 算法的初始值 u_0 和 λ_0 选取为 N 维零向量, 马氏链长为 3×10^4 , 取后 1×10^4 个点作为平稳后的采样点。贝叶斯方法给出了条件均值(CM)估计和最大后验(MAP)估计, 记 $k_0 = 2 \times 10^4$ 以及 $k_1 = 3 \times 10^4$, 数值实验中它们的求解如下

$$u_{CM} = \frac{1}{k_1 - k_0} \sum_{k=k_0+1}^{k_1} u_k,$$

$$u_{MAP} = \arg \max_{k \in \{k_0+1, \dots, k_1\}} \pi(u_k),$$

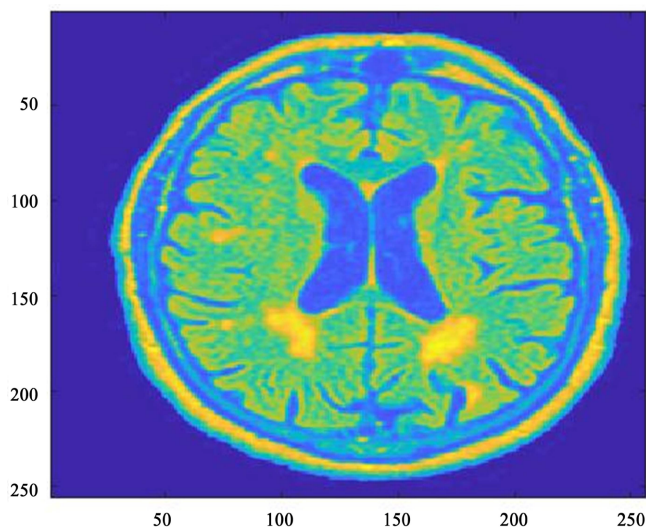


Figure 1. Original image

图 1. 原始图像

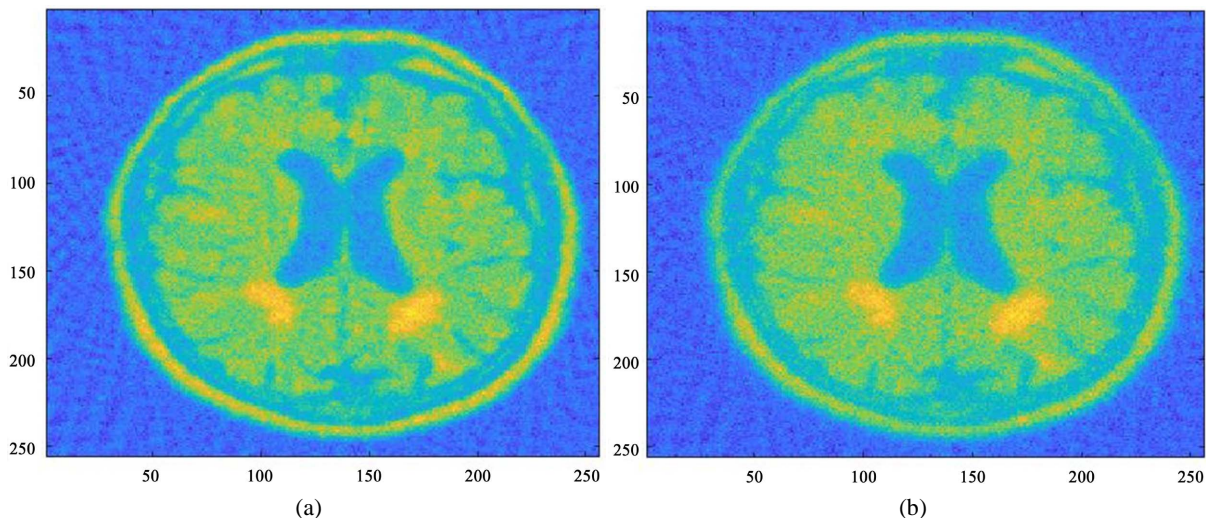


Figure 2. PxEuL algorithm results: MAP estimation (left) and CM estimation (right), $\gamma = 5$

图 2. PxEuL 算法重构结果: MAP 估计(左)和 CM 估计(右), $\gamma = 5$

图 2、图 3 给出了两种算法的重构结果, 包括 MAP 估计和 CM 估计。从单个算法来看, CM 估计比 MAP 估计更优、更为清晰; 从两个算法来看, DuEuL 算法给出的重构解相比于 PxEuL 算法有着更清晰的边界, 且背景更颜色更深更接近原始图像。两种算法一定程度上都可以反演出脑部结构, 但相比于原始图像, 存在背景颜色更浅、反演结果存在伪影的问题。

表 1 给出了两种算法求解的 MAP 估计和 CM 估计对应的相对误差、峰值信噪比和结构相似性等结果。从单个算法来看, CM 估计的结果通常优于 MAP 估计, 这说明通过贝叶斯方法给出的 CM 估计是相当有必要的; 从两个算法比较来看, 我们提出的 DuEuL 算法在三个指标上大大优于 PxEuL 算法, 说明 DuEuL 算法中 TV 约束更能够刻画反演解的特征。表 2 给出了 DuEuL 算法取不同 η 时的结果, 由 MAP 估计和 CM 估计结果可知, 随着 η 的变大, 算法的重构结果更好。

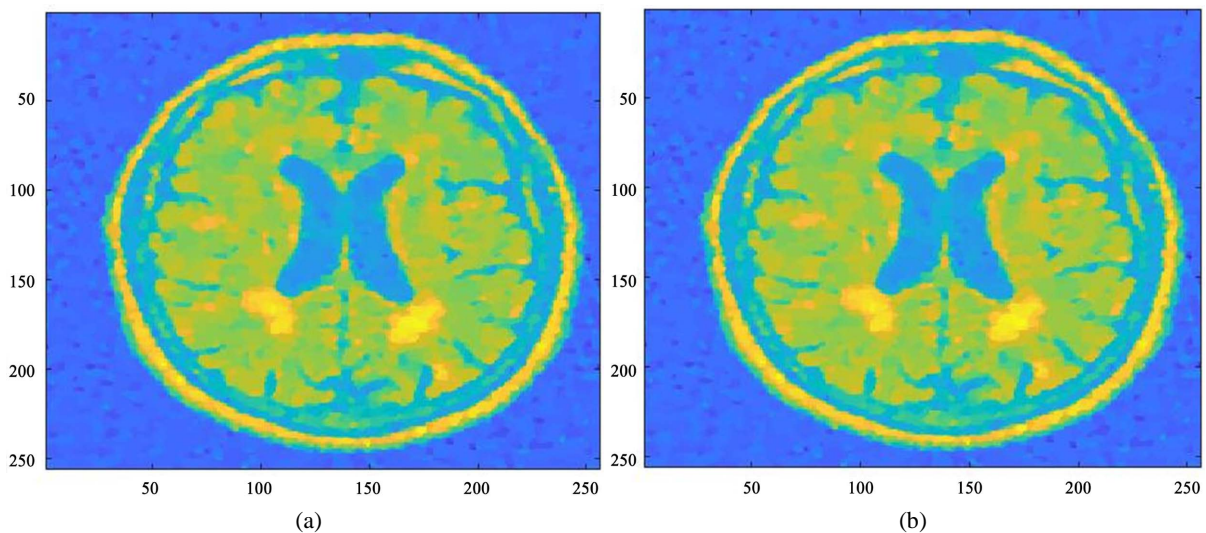


Figure 3. DuEuL algorithm results: MAP estimation (left) and CM estimation (right), $\eta = 5$

图 3. DuEuL 算法结果: MAP 估计(左)和 CM 估计(右), $\eta = 5$

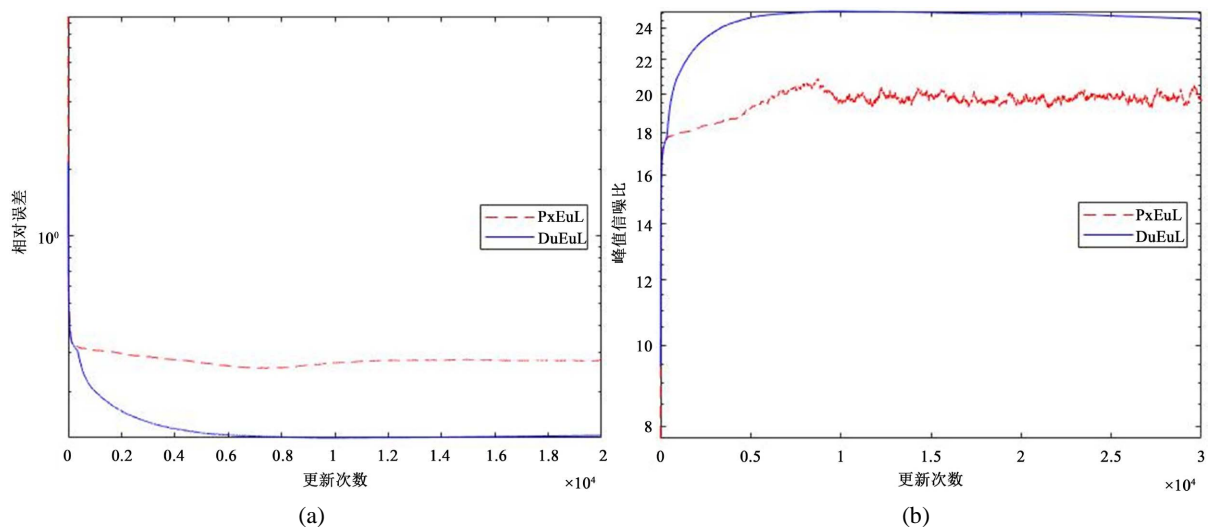
Table 1. Results of two algorithms ($\gamma = \eta = 5$)**表 1.** 两种算法结果($\gamma = \eta = 5$)

算法	输出结果	RelErr	PSNR	SSIM
PxEuL	MAP 估计	0.26672	19.4414	0.99922
	CM 估计	0.21144	21.2131	0.99958
DuEuL	MAP 估计	0.13285	25.2349	0.99985
	CM 估计	0.13109	25.2692	0.99986

Table 2. DuEuL algorithm results with different parameters taken**表 2.** 取不同参数 DuEuL 算法结果

(η, δ)	输出结果	RelErr	PSNR	SSIM
$(1 \times 10^{-3}, 1 \times 10^{-3})$	MAP 估计	0.17088	23.0787	0.99974
	CM 估计	0.16781	23.1574	0.99974
$(1 \times 10^{-2}, 5 \times 10^{-4})$	MAP 估计	0.16851	23.3248	0.99975
	CM 估计	0.16599	23.3482	0.99975
$(1 \times 10^{-1}, 1 \times 10^{-4})$	MAP 估计	0.15941	23.5639	0.99977
	CM 估计	0.15758	23.5726	0.99977
$(5, 1 \times 10^{-5})$	MAP 估计	0.13285	25.2349	0.99985
	CM 估计	0.13109	25.2692	0.99986

图 4 给出了两种算法的相对误差和峰值信噪比随着更新次数的变化趋势。从两种算法比较来看，更新初期，PxEuL 算法和 DuEuL 算法变化速度差不多，而后期 PxEuL 算法存在一定程度的半收敛现象，半收敛现象的程度依赖于参数选取的好坏。另一方面，DuEuL 算法最终达到平稳的相对误差值更低，且变化趋势更光滑，没有半收敛现象。

**Figure 4.** RelErr and PSNR ratio trend of two algorithms ($\gamma = \eta = 5$)**图 4.** 两种算法的相对误差和峰值信噪比趋势($\gamma = \eta = 5$)

5. 结论

本文基于非光滑约束下的混合先验,提出了对偶 Langevin 算法,并与近端 Langevin 算法对比,通过 MAP 估计和 CM 估计解相应的指标比较算法的优劣,并基于 CT 成像问题进行了数值模拟。结果表明 DuEuL 算法显著优于现有的近端 Langevin 算法,可以给出更好的重构解,更能够利用 TV 约束刻画反演解的特征。

基金项目

国家自然科学基金(批准号: 12071184)资助项目。

参考文献

- [1] 程晋, 刘继军, 张波. 偏微分方程反问题: 模型、算法和应用[J]. 中国科学: 数学, 2019, 49(4): 643-666.
- [2] Engl, H., Hanke, M. and Neubauer, A. (1996) Regularization of Inverse Problems. Springer, New York.
- [3] Burger, M. and Osher, S. (2004) Convergence Rates of Convex Variational Regularization. *Inverse Problems*, **20**, 1411-1421. <https://doi.org/10.1088/0266-5611/20/5/005>
- [4] Yao, Z., Hu, Z. and Li, J. (2016) A TV-Gaussian Prior for Infinite-Dimensional Bayesian Inverse Problems and Its Numerical Implementations. *Inverse Problems*, **32**, 075006. <https://doi.org/10.1088/0266-5611/32/7/075006>
- [5] Cotter, S.L., Roberts, G.O., Stuart, A.M. and White, D. (2023) MCMC Methods for Functions: Modifying Old Algorithms to Make Them Faster. *Statistical Science*, **28**, 424-446. <https://doi.org/10.1214/13-STS421>
- [6] Lassas, M. and Siltanen, S. (2004) Can One Use Total Variation Prior for Edge-Preserving Bayesian Inversion? *Inverse Problems*, **20**, 1537. <https://doi.org/10.1088/0266-5611/20/5/013>
- [7] Durmus, A., Majewski, S. and Miasojedow, B. (2019) Analysis of Langevin Monte Carlo via Convex Optimization. *The Journal of Machine Learning Research*, **20**, 2666-2711.
- [8] Durmus, A., Moulines, E. and Pereyra, M. (2018) Efficient Bayesian Computation by Proximal Markov Chain Monte Carlo: When Langevin Meets Moreau. *SIAM Journal on Imaging Sciences*, **11**, 473-506. <https://doi.org/10.1137/16M1108340>
- [9] Dalalyan, A.S. (2017) Theoretical Guarantees for Approximate Sampling from Smooth and Log-Concave Densities. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **79**, 651-676. <https://doi.org/10.1111/rssb.12183>
- [10] Bonettini, S. and Ruggiero, V. (2012) On the Convergence of Primal-Dual Hybrid Gradient Algorithms for Total Variation Image Restoration. *Journal of Mathematical Imaging and Vision*, **44**, 236-253. <https://doi.org/10.1007/s10851-011-0324-9>
- [11] Jin, Q. and Wang, W. (2013) Landweber Iteration of Kaczmarz Type with General Non-Smooth Convex Penalty Functionals. *Inverse Problems*, **29**, 085011. <https://doi.org/10.1088/0266-5611/29/8/085011>
- [12] Zhong, M., Wang, W. and Jin, Q. (2019) Regularization of Inverse Problems by Two-Point Gradient Methods in Banach Spaces. *Numerische Mathematik*, **143**, 713-747. <https://doi.org/10.1007/s00211-019-01068-0>
- [13] 金其年, 王薇. Banach 空间中求解线性反问题的对偶梯度流方法[J]. 中国科学: 数学, 2023, 53(10): 1377-1396.
- [14] Zhong, M., Wang, W. and Zhu, K. (2022) On the Asymptotical Regularization with Convex Constraints for Nonlinear Ill-Posed Problems. *Applied Mathematics Letters*, **133**, 108247. <https://doi.org/10.1016/j.aml.2022.108247>
- [15] Hansen, P.C. and Saxild-Hansen, M. (2012) AIR Tools—A MATLAB Package of Algebraic Iterative Reconstruction Methods. *Journal of Computational and Applied Mathematics*, **236**, 2167-2178. <https://doi.org/10.1016/j.cam.2011.09.039>
- [16] Zhu, M. and Chan, T. (2008) An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration. *Ucla Cam Report*, **34**, 8-34.