

一类求解约束离散不适定问题的积极集 随机迭代方法

郑 宁, 殷俊锋

(同济大学 数学科学学院, 上海 200092)

摘要: 许多科学和工程领域的应用问题都可以归结为线性离散不适定问题的求解。考虑大规模带盒子约束的线性离散不适定问题的求解, 提出一类基于积极集策略的随机内外迭代方法。基于积极集策略的内外迭代法在外层迭代上更新积极集和对应的非积极集, 并采用投影算子, 将不在可行域中的数值解分量投影到可行域边界上, 同时在内层迭代上采用 Krylov 子空间方法求解无约束子问题。提出一类积极集迭代法, 在内层迭代上采用高性能随机算法, 依照概率分布选取子问题系数矩阵的列进行更新, 并利用 Armijo 下降准则对迭代步长进行选择, 这样就可以保证目标函数值随着迭代步数的增加而单调下降。在图像复原问题的数值实验中, 验证所构造算法的高效性。在偏差准则的收敛条件下, 新的积极集内外迭代法所利用的计算量、迭代步数和 CPU 时间都比前人提出的算法更少。

关键词: 线性离散不适定问题; 盒子约束; 随机迭代; 积极集方法; 内外迭代算法

中图分类号: O241.6

文献标志码: A

A Class of Active Set Randomized Iterative Methods for Solving Constrained Discrete Ill-Posed Problems

ZHENG Ning, YIN Junfeng

(School of Mathematical Sciences, Tongji University, Shanghai 200092, China)

Abstract: Linear discrete ill-posed problems arise from many scientific computation and engineering application areas. Considering the solution to large scale ill-posed problems with box constraints, A novel class of randomized internal and external iterative methods are proposed based on the active set strategy, which contain

a two-level iteration that, for the external iteration, updates the active set and free variable set, and orthogonally project the iterate onto the feasible boundary, while for the internal iteration, solves the unconstrained linear system with the Krylov subspace methods. The proposed novel active set algorithm is to utilize the efficient randomized method for the internal iteration and the step size is chosen by the Armijo criterion, so that the objective function value can be monotonically decreased with the increase of the number of iterations. Numerical experiments on the image restoration show the efficiency of the proposed algorithm. Under the condition of the discrepancy principle, the computational complexity, iteration steps, and CPU time of the novel active set randomized iteration methods are less than those of previous methods.

Key words: linear discrete ill-posed problems; box constraints; randomized iteration; active method; internal and external iteration

线性离散不适定问题来源于许多科学计算和工程应用领域^[1]。例如第一类 Fredholm 积分方程数值离散后的问题; 例如图像去模糊复原问题, 已知观测到的模糊带噪声图像和模糊算子, 需要尽可能复原得到清晰图像。也就是说, 在数学上考虑一个不相容的线性方程组

$$Ax = b \quad (1)$$

其中 $A \in \mathbb{R}^{m \times n}$ 是一个大型稀疏的矩阵, 表示离散的模糊算子, 可以是方阵, 也可以是长方形矩阵。右端项 $b \in \mathbb{R}^{m \times 1}$ 表示观测得到的数据, 即带有噪声的模糊图像。式(1)和通常的线性方程组不同之处在于,

收稿日期: 2021-05-10

基金项目: 国家自然科学基金(11971354); 中央高校基本科研业务费专项基金(22120210555)

第一作者: 郑 宁(1990—), 男, 助理教授, 硕士生导师, 理学博士, 主要研究方向为数值代数与金融计算。

E-mail: nzheng@tongji.edu.cn

通信作者: 殷俊锋(1979—), 男, 教授, 博士生导师, 理学博士, 主要研究方向为数值代数与科学计算。

E-mail: yinjf@tongji.edu.cn



论文
拓展
介绍

矩阵 A 是长方形矩阵,且往往行数 m 大于列数 n ,因此式(1)是一个超定的线性系统。从数值角度上分析,矩阵 A 来源于带积分项的模糊算子的离散,往往大型稀疏,且奇异值聚集在原点附近。这样的矩阵往往带有坏条件数,甚至接近奇异矩阵。观测的数据 b 往往带有观测误差。

$$b = \hat{b} + e \quad (2)$$

式中: \hat{b} 为未知的没有噪声的向量,且 $\hat{b} = A\hat{x}$, \hat{x} 为未知的清晰的真实图像;向量 e 表示某种误差或噪声,其中每个分量都独立服从相同的随机分布。因此,线性系统(1)是不相容的,只能求解其对应的最小二乘解。

令 A^\dagger 为矩阵 A 的摩尔-彭若斯广义逆,则不定问题(1)的最小二乘解为

$$x = A^\dagger b = \hat{x} + A^\dagger e$$

然而,由于不定问题中矩阵 A 病态,式(1)的解对右端项 b 中的误差 e 非常敏感,由噪声引起的分量 $A^\dagger e$ 会将所需要的信息 \hat{x} 掩盖,使得解 $A^\dagger b$ 没有意义。为了解决此问题,需要正则化方法。常用的正则化方法分为2类:当噪声 e 的取值范围已知时,可以利用偏差原理(discrepancy principle)方法;当噪声 e 未知时,可以利用 Tikhonov 正则化方法、广义交叉校验准则和 L 曲线准则等确定正则化参数。本文采用 Tikhonov 正则化方法,即最小化带有惩罚项的目标函数

$$\min f(x) = \|Ax - b\|_2^2 + \mu^2 \|Lx\|_2^2 \quad (3)$$

式中: L 为一个离散微分算子; μ 为正则化参数, $\mu > 0$ 。

在实际应用中,数据往往是在一定的取值范围内选取的。例如,图像的像素值是非负的;灰度图像的像素值则在 $[0, 255]$ 之间。因此,考虑一类带盒子约束的不定问题^[2],即

$$\min f(x) \text{ 使得 } l \leq x \leq u \quad (4)$$

其中, $l \in \mathbb{R}^{n \times 1}$ 和 $u \in \mathbb{R}^{n \times 1}$ 分别是下界向量和上界向量。

针对大规模带约束不定问题(4), Bertero 和 Boccacci^[3] 提出一类投影 Landweber 方法。Hanke 等^[4] 考虑变量无上界的特殊情形,利用显性变量替换将原问题转换成一类非线性方程组,然后采用拟牛顿法进行求解。Nagy 和 Strakos^[5] 将最速下降法与投影技巧结合,在目标函数下降的同时,迭代解满足约束条件。Rojas 和 Steihaug^[6] 从优化算法出发,提出惩罚函数法,使得数值解序列在可行区域内部收敛到最优点。Calvetti 等^[2] 将正则化方法和 Krylow 子空间迭代法相结合,提出一类内外迭代法。

文献[6-7]在此基础上引入积极集策略,外层迭代快速更新积极集,而内层迭代则在子空间通过迭代求解一系列仅在非积极集上有约束的问题。数值实验验证此算法的有效性,并且数值效率高于前人提出的算法,所需的计算复杂度和 CPU 时间都更少。但此方法的缺点在于缺乏理论收敛性,尤其是无法保证目标函数值随着迭代步数逐渐单调下降。

针对内层迭代中需要求解的无约束最小二乘问题,近年来诸多学者提出高性能随机迭代方法。Strohmer 和 Vershynin^[8] 提出随机 Kaczmarz 方法在求解相容系统上的收敛理论。Leventhal 和 Levis^[9] 提出求解不相容系统,即最小二乘问题的随机坐标下降方法,并给出当系数矩阵列满秩时的收敛理论。随着新的有效概率准则的构造,更多高性能的随机迭代方法相继被提出^[9-11],在理论和数值实验中均展示出比传统确定性方法更高效的计算性能。

本文提出一类基于积极集策略的内外迭代法。该算法在内层迭代上利用随机坐标下降方法求解无约束最小二乘问题,并利用 Armijo 下降准则对迭代步长进行选择,这样就可以保证目标函数值随着迭代步数的增加而单调下降;同时在外层迭代上更新积极集和对应的非积极集,并采用投影算子,将不在可行域中的数值解分量投影到可行域边界上。数值实验验证此算法的高效性。在偏差准则的收敛条件下,新的积极集内外迭代法所利用的计算量、迭代步数和 CPU 时间都比前人提出的算法更少。

1 基于积极集策略的内外迭代法

简单介绍求解问题(4)的2种内外迭代法。第1种方法是基于在内迭代中将数值解投影到盒子约束可行域 $\Omega = \{x \in \mathbb{R}^n: l \leq x \leq u\}$ 中。第2种方法是基于积极集策略,通过在每次外迭代中更新积极集,从而在内迭代中只要求解一个规模较小的子问题。

容易验证,问题(4)中的目标函数等价于

$$f(x) = \left\| \begin{bmatrix} A \\ \mu L \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2^2 = \|\bar{A}x - \bar{b}\|_2^2$$

如果不考虑对变量的盒子约束,极小化 $f(x)$ 在数学上等价于求解一个正规方程组

$$\bar{A}^\top \bar{A}x = \bar{A}^\top \bar{b}$$

令 \hat{x} 为此方程组的解,则通常 $\hat{x} \notin \Omega$ 。定义一个正交投影算子 P_Ω , 则 $x = P_\Omega(\hat{x})$ 为

$$x_j = \begin{cases} l_j, & \text{当 } \hat{x}_j < l_j \\ u_j, & \text{当 } \hat{x}_j > u_j \\ \hat{x}_j, & \text{当 } l_j \leq \hat{x}_j \leq u_j \end{cases}$$

其中 $j=1,2,3,\dots,n$, 这样可以保证 $x \in \Omega$ 。通过反复迭代, 使得数值解对应的目标函数值逐渐下降, 从而得到所需要的解。

投影内外迭代算法^[3]求解带盒子约束的问题(4)如下:

算法1^[3] 投影内外迭代法。①选取一个迭代向量 x^0 。②计算 $r^0 = \bar{b} - \bar{A}x^0$ 。③开始迭代 $k=0, 1, 2, \dots$, 直到收敛。④通过迭代法求得如下问题的数值解 w^k :

$$\min_{w \in R^k} \|\bar{A}w - r^k\|_2 \quad (5)$$

⑤令 $\hat{x}^{k+1} = x^k + w^k$ 并投影到可行区域中 $x^{k+1} = P_\Omega(\hat{x}^{k+1})$ 。⑥计算 $r^{k+1} = \bar{b} - \bar{A}x^{k+1}$ 。⑦结束迭代。其中, 最小二乘问题(5)的求解为内层迭代, k 的更新为外层迭代, 因此算法1为内外迭代法。对于式(5)的求解, 可以利用CGLS算法^[1], 即共轭梯度法求解其正规方程组: $\bar{A}^T \bar{A}w = \bar{A}^T r^k$ 。如果选取内迭代的初始向量为零向量, 则得到的数值解 w^k 在 Krylov 子空间 $K(\bar{A}^T \bar{A}, \bar{A}^T b) = \text{span}\{\bar{A}^T b, \bar{A}^T \bar{A} \bar{A}^T b, (\bar{A}^T \bar{A})^2 \bar{A}^T b, \dots\}$ 中。

Morigi 等^[7]在投影内外迭代算法2的基础上引入积极集策略。令拉格朗日算子 $\lambda^k = -\bar{A}^T r^k$, 其中 x^k 是第 k 步的数值解, $r^k = \bar{b} - \bar{A}x^k$ 是 x^k 对应的残量。根据问题(4)的 Karush-Kuhn-Tucker 条件可知, x^* 是问题(4)的最优解当且仅当其满足

$$x_j^* \geq l_j, \lambda_j \geq 0, \lambda_j(x_j^* - l_j) = 0 \quad (6)$$

$$x_j^* \leq u_j, \lambda_j \leq 0, \lambda_j(u_j - x_j^*) = 0 \quad (7)$$

对于 x^k , 定义积极集为 $B(x^k) = \{j: x_j^k = l_j, \lambda_j^k \geq 0\} \cup \{j: x_j^k = u_j, \lambda_j^k \leq 0\}$, 其对应的补集自由变量为 $F(x^k) = \{1, 2, \dots, n\} \setminus B(x^k)$ 。基于积极集算法的思想是: 如果已知 x^* 的积极集, 则约束最优化问题(4)

可以转化成无约束优化问题 $\min_{x \in R^n} \|\bar{A}_F x - \bar{b}\|_2$ 使得

$l_F \leq x_F \leq u_F$, 其中 $\bar{b} = \bar{b} - \bar{A}_B x_B$, \bar{A}_F 是矩阵 \bar{A} 中列指标属于集合 F 的子矩阵。由于 x^* 未知, 所以先选取初始迭代向量不断更新积极集, 直到收敛为止。

基于积极集的投影内外迭代算法^[5]求解带盒子约束的问题(4)如下。

算法2^[5] 基于积极集的投影内外迭代法。①选取一个迭代初始向量 x^0 。②计算 $r^0 = \bar{b} - \bar{A}x^0$ 。③开始迭代 $k=0, 1, 2, \dots$, 直到收敛。④计算拉格朗日算子 $\lambda^k = -\bar{A}^T r^k$ 。⑤更新积极集 $B(x^k)$ 和自由变量集 $F(x^k)$ 。⑥通过迭代法求得如下问题的数值解 w^k : $\min_{w \in R^n} \|\bar{A}_F w - r^k\|_2$ 。⑦令 $\hat{x}_F^{k+1} = x_F^k + w^k, \hat{x}_B^{k+1} =$

x_B^k , 并投影到可行区域中, 即 $x^{k+1} = P_\Omega(\hat{x}^{k+1})$ 。⑧计算 $r^{k+1} = \bar{b} - \bar{A}x^{k+1}$ 。⑨结束迭代。

在算法2中, 外层迭代快速更新积极集, 而内层迭代则在子空间通过迭代求解一系列仅在非积极集上有约束的问题。和算法1类似, 内层迭代可以采用CGLS算法快速求解。数值实验^[5]表明, 基于积极集的算法收敛速度更快, 且所需的迭代步数和CPU时间都更少。

算法2的缺点在于理论上无法保证目标函数值随着迭代单调下降。因此, 在第2节中构造新的基于积极集的算法。

2 结合 Armijo 策略的新积极集算法

将 Armijo 充分下降准则和积极集思想相结合, 同时, 在内层无约束最小二乘问题的求解上采用随机坐标下降策略。

算法3 基于积极集和 Armijo 条件的随机投影内外迭代法。①选取一个迭代初始向量 x^0 。②计算 $r^0 = \bar{b} - \bar{A}x^0$ 。③开始迭代 $k=0, 1, 2, \dots$, 直到收敛。④计算拉格朗日算子 $\lambda^k = -\bar{A}^T r^k$ 。⑤更新积极集 $B(x^k)$ 和自由变量 $F(x^k)$ 。⑥通过随机坐标下降迭代法求得如下问题的数值解 w^k :

$$\min_{w \in R^k} \|\bar{A}_F w - r^k\|_2 \quad (8)$$

⑦令 $\hat{x}_F^{k+1} = x_F^k + \beta^m w^k, \hat{x}_B^{k+1} = x_B^k$, 其中 m 是满足如下 Armijo 充分下降准则的最小非负整数:

$$\|\bar{A}x^{k+1} - \bar{b}\|_2^2 \leq \|\bar{A}x^k - \bar{b}\|_2^2 + 2\alpha(\lambda^k)^T(x^{k+1} - x^k) \quad (8)$$

并投影到可行区间中

$$x^{k+1} = P_\Omega(\hat{x}^{k+1})$$

⑧计算 $r^{k+1} = \bar{b} - \bar{A}x^{k+1}$ 。⑨结束迭代。在算法2的步骤⑦中, 参数 $0 \leq \alpha < 1, 0 < \beta < 1$ 。

3 数值实验

用图像去模糊复原的数值实验来验证基于积极集和 Armijo 条件的投影内外迭代法求解带盒子约束的不定问题的可行性和有效性。

在数值实验中, 采用的收敛准则为偏差原理。另外, 选取二阶微分算子为正则化光滑子

$$L = \begin{bmatrix} 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 1 & -2 & 1 \end{bmatrix}$$

在图 1 中分别画出图像复原算例“satellite”的原始图像、模糊带噪声图像、算法 2 复原图像、算法 3 复原图像。图 2、图 3 分别为相对误差和相对残量随着迭代步数变化的曲线。从图 2 可以看出,2 类积极集内外迭代算法随着迭代步数的增加,相对误差都先下降后增加。最优外迭代步数在 5 步附近。从图 3 可以看出,相比于算法 2,算法 3 的相对残量下降更快,振荡更小,这说明随机坐标下降方法结合 Armijo 下降准则可以使计算更快更平稳地收敛,所需要的计算量更少。



图 1 图像复原算例“satellite”

Fig. 1 Restored image of ‘statellite’

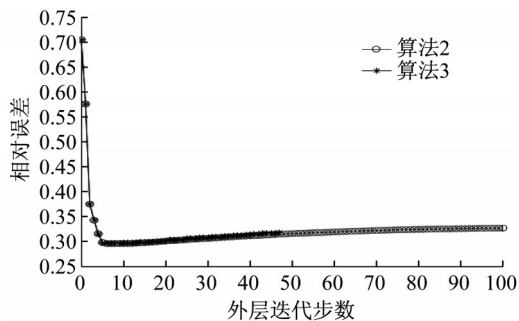


图 2 图像复原算例“satellite”相对误差随迭代步数变化

Fig. 2 Relative error of restored image of ‘satellite’ versus the number of iterative steps

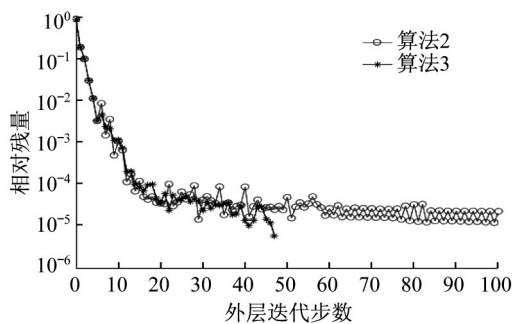


图 3 图像复原算例“satellite”相对残量随迭代步数变化

Fig. 3 Relative residue of restored image ‘satellite’ versus the number of iteration steps

4 结论和展望

考虑带盒子约束的不定问题,构造一类基于积极集策略的内外迭代法。此算法的创新之处在于,在内层迭代上利用随机坐标下降方法,同时采用 Armijo 下降准则对迭代步长进行选择,这样就可以保证目标函数值随着迭代步数的增加而单调下降。

同时在外层迭代上更新积极集和对应的非积极集,并采用投影算子,将不在可行域中的数值解分量投影到可行域边界上。数值实验验证此算法的高效性。在偏差准则的收敛条件下,新的积极集内外迭代法所用的计算量、迭代步数和 CPU 时间都比前人提出的算法更少。在后续工作中,将考虑稀疏约束条件下的最小二乘问题,并构造类似的积极集随机迭代算法进行求解。

作者贡献声明:

郑 宁:负责算法构造、数值实验和分析、论文撰写等。

殷俊锋:负责算法设计和数值案例分析。

参考文献:

- [1] BJORK Å. Numerical methods for least squares problems [M]. Philadelphia :SIAM,1996.
- [2] CALVETTI D, LANDI G, REICHEL L, *et al.* Nonnegativity and iterative methods for ill-posed problems [J]. Inverse Problems, 2004, 20 :1747.
- [3] BERTERO M, BOCCACCI P. Introduction to inverse problems in imaging [M]. Bristol : Institute of Physics Publishing, 1998.
- [4] HANKE M, NAGY J, VOGEL C. Quasi-Newton approach to nonnegative image restorations [J]. Linear Algebra and its Applications, 2000, 316 :223.
- [5] NAGY J, STRAKOS Z. Enforcing non-negativity in image reconstruction algorithms, in mathematical modeling, estimation and imaging [J]. The International Society for Optical Engineering, 2000,4121 :182.
- [6] ROJAS M, STEIHAUG T. An interior point trust-region based method for large-scale non-negative regularization [J]. Inverse Problems, 2002, 18 :1291.
- [7] MORIGI S, REICHEL L, SGALLARI F, *et al.* An iterative method for linear discrete ill-posed problems with box constraints [J]. Journal of Computational and Applied Mathematics, 2007, 198, 505.
- [8] STROHMER T, VERSHYNIN R. A randomized Kaczmarz algorithm with exponential convergence [J]. Journal of Fourier Analysis and Applications, 2009, 15: 262.
- [9] LEVENTHAL D, LEVIS A S. Randomized methods for linear constraints: Convergence rates and conditioning [J]. Mathematics of Operations Research, 2010, 35, 641.
- [10] ZOUZIAS A, FRERIS N M. Randomized extended Kaczmarz for solving least squares[J]. SIAM J Matrix Anal Appl, 2013, 34: 773.
- [11] 杜亦疏,殷俊锋,张科. 求解大型稀疏线性方程组的贪婪距离随机 Kaczmarz 算法[J]. 同济大学学报(自然科学版), 2020, 48(8): 1224.

DU Yishu, YIN Junfeng, ZHANG Ke. Greedy randomized-distance Kaczmarz method for solving large sparse linear systems [J]. Journal of Tongji University (Natural Science), 2020, 48(8): 1224.