Advances in Applied Mathematics 应用数学进展, 2024, 13(3), 1129-1139
Published Online March 2024 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2024.133105

一类分式优化问题的带非单调线搜索的近端梯度次梯度算法研究

张 景

河北工业大学理学院数学系，天津

收稿日期：2024年2月27日；录用日期：2024年3月21日；发布日期：2024年3月28日

摘 要

本文主要研究一类分式优化问题，其中分子是凸非光滑连续函数与非凸光滑函数的和，分母为凸非光滑函数。首先给出了问题的一阶最优性条件，然后给出了求解分式优化问题的新算法，即带非单调线搜索的近端梯度次梯度算法(简称**NL-PGSA**)。此外，基于**Kurdyka-Łojasiewicz**性质，可以保证算法生成的整个序列的全局收敛性，最后，对l_1/l_2稀疏信号恢复问题进行了数值实验，验证了该算法的有效性。

关键词

分式优化，近端梯度次梯度算法，收敛性分析

Research on the Proximal Gradient-Subgradient Algorithm with Nonmonotonic Line Search for a Class of Fractional Optimization Problems

Jing Zhang

Department of Mathematics, School of Sciences, Hebei University of Technology, Tianjin

Received: Feb. 27^{th}, 2024; accepted: Mar. 21^{st}, 2024; published: Mar. 28^{th}, 2024

文章引用: 张景. 一类分式优化问题的带非单调线搜索的近端梯度次梯度算法研究[J]. 应用数学进展, 2024, 13(3): 1129-1139. DOI: 10.12677/aam.2024.133105

Abstract

This paper mainly considers a class of fractional optimization problems where the numerator is a sum of a convex nonsmooth continuous function and a nonconvex smooth function, while the denominator is a convex nonsmooth function. We first give the first-order optimality condition of the problem, and then a new algorithm, called proximal gradient-subgradient algorithm with nonmonotonic line search (NL-PGSA), is proposed for solving the fractional optimization problems. Moreover, the global convergence of the entire sequence generated by NL-PGSA algorithm has been proven based on the Kurdyka-Łojasiewicz property. Finally, some numerical experiments on the l_1/l_2 sparse signal recovery problems are conducted to demonstrate the efficiency of the proposed algorithm.

Keywords

Fractional Optimization, Proximal Gradient-Subgradient Algorithm, Convergence Analysis

1. 引言

本文主要考虑的是一类特殊的分式优化问题，其形式为:

$$\min\left\{\frac{f(x)+h(x)}{g(x)} : x \in \Omega := \{x \in R^n : g(x) \neq 0\}\right\}, \tag{1.1}$$

其中$f, g, h : R^n \to \overline{R} := (-\infty, +\infty]$ 并满足如下假设:
假设条件1.1.

(i) f是恰当凸函数，有下界并且在其有效域上是连续的；

(ii) h是一个具有*Lipschitz*连续梯度的连续可微函数，即存在一个*Lipschitz* 常数L, 满足

$$||\nabla h(x)-\nabla h(y)\|\leq L\|x-y\|,\quad\forall x,y\in R^n.$$

(iii) g是一个实值凸函数；

(iv) $f+h$在R^n上是非负的，在$R^n\setminus\Omega$ 上不等于*0*, $g>0$ 在$\mathrm{dom}(f)\cap\Omega$. 此外，我们还假设问题(1.1)的最优解存在并能够得到。

分式优化问题是最小化或最大化一个或多个函数比值的目标问题，分式优化问题在许多领域都有不同的应用，如人工智能 [1–3]，无线电通信 [4–7]，经济学 [8] 等。因此，针对这一问题许多学者提出了一些有效的方法。其中，参数方法是分式优化问题中常用的思想，他是求解分式优化问题中一种灵活的方法，使用参数方法，我们可以立即证明问题(1.1)的最优解$x^\star\in R^n$ 等价为下列问题的最优解

$$\min\left\{f(x)+h(x)-c_\star g(x):x\in\Omega\right\},\tag{1.2}$$

其中$c_\star$代表问题(1.1)的最优值且$c_\star=\frac{f(x^\star)+h(x^\star)}{g(x^\star)}$, 该方法的优点是他可以求解大规模的分式优化问题，但缺点也很显然，问题(1.1) 的全局最优解不易得到。为了克服这一困难，有学者提出了Dinkelbach's方法 [9], 该方法由当前迭代的函数值c_k取代$c_\star$，具体迭代框架为

$$x^{k+1}\in\operatorname{argmin}\left\{f(x)+h(x)-c_kg(x):x\in\Omega\right\},\tag{1.3}$$

其中$c_k=\frac{f(x^k)+h(x^k)}{g(x^k)}$, 关于该方法的详细论述，可参考文献 [10–12]. 然而，从问题(1.3)可知，该目标函数为非凸函数，要得到全局解也比较困难。

最近，Bot等人 [13]将近端梯度算法(简称PG)应用于当函数g为连续可微时的问题(1.1), 具体迭代框架为

$$x^{k+1}\in\operatorname{argmin}\left\{f(x)+h(x)-c_k\left\langle\nabla g\left(x^k\right),x\right\rangle+\frac{1}{2\eta_k}\left\|x-x^k\right\|^2:x\in\Omega\right\},\tag{1.4}$$

其中$c_k=\frac{f(x^k)+h(x^k)}{g(x^k)}$, $\eta_k>0$. 从(1.4)可知，上述迭代实际上采用了$g(x)$在x^k处的线性展开。然而，从数值性能来看，当问题规模比较大时，近端梯度算法的速度相对较慢。为了提高收敛速度，许多研究人员采用了各种策略，并提出了一些高效的算法。

Zhang等人 [14]提出了一种近端梯度次梯度的算法(简称PGSA)来解决这个问题，算法的迭代框架为

$$x^{k+1}\in\operatorname{argmin}\left\{f(x)+\left\langle\nabla h\left(x^k\right)-c_ky^{k+1},x-x^k\right\rangle+\frac{1}{2\alpha_k}\left\|x-x^k\right\|^2:x\in\Omega\right\},\tag{1.5}$$

其中$y^{k+1}\in\partial g(x^k)$, $0<\alpha_k\leq 1/L$, $c_k=\frac{f(x^k)+h(x^k)}{g(x^k)}$. 我们可以看到，上述迭代采用的是$h(x)-c_kg(x)$的线性展开，这恰好是$\alpha_kf$的近端算子。此外，为了加速算法，作者还提出了一种带线搜索的近端梯度次梯度算法(简称PGSA-L), 为保证收敛性，要求步长$0<\alpha_k<\frac{1}{L}$, 但该算法的弊端是当L较大时，步长会较小，因此收敛速度相对较慢。

本文受到文献 [14]的启发，在f是凸函数的前提下，采用了另一种方式的BB 步长并进一步改进了步长的范围，在函数满足Kurdyka-Łojasiewicz性质的条件下，可以保证算法的收敛性。

2. 预备知识

下面简单介绍与本文相关的一些符号与定义。

定义R^n为n维欧式空间，$\langle\cdot,\cdot\rangle$为标准内积，$\|\cdot\|$, $\|\cdot\|_1$分别为欧式范数与l_1范数，集合S上的示性函数$I_S(x)$定义为

$$I_S(x) := \begin{cases} 0, & \text{if } x \in S, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.1}$$

对于一个扩展的实值函数$\psi : R^n \to (-\infty, +\infty]$, 其有效域为$\operatorname{dom}\psi := \{x \in R^n : \psi(x) < \infty\}$. 若$\forall x \in R^n$, 都满足$\psi > -\infty$且$dom(\psi)$非空，则$\psi$为恰当函数。给定一个非空闭集$S \subseteq R^n$, 设$\operatorname{dist}(\cdot, S) : R^n \to R$表示$x$ 到S的距离函数，即$\operatorname{dist}(x, S) := \inf_{y\in S} \|x - y\|, \forall x \in R^n$.

定义2.1. *(Kurdyka-Łojasiewicz性质 [15])* 设函数$\psi : R^n \to \bar{R}$为恰当的下半连续函数，$\hat{x} \in dom(\partial\psi)$, 如果存在$\eta \in (0, +\infty]$, $\hat{x}$的一个领域O及满足下述条件的凹函数$\phi : [0, \eta) \to R_+ := [0, +\infty)$
(i) $\phi(0) = 0$,
(ii) ϕ在$(0, \eta)$上连续可微且$\phi' > 0$,
(iii) $\forall x \in O \cap \{x \in R^n : \psi(\hat{x}) < \psi(x) < \psi(\hat{x}) + \eta\}$, $\phi'(\psi(x) - \psi(\hat{x}))\cdot dist\ (0, \partial\psi(x)) \geq 1$成立.
则称ψ在点$\hat{x}$处满足*Kurdyka-Łojasiewicz*性质(简称*KL*性质)。

3. 主要算法及定理

在本节中，我们将会给出算法的具体迭代框架，并保证算法的收敛性证明。为简便，定义函数F, 具体形式如下：

$$F(x) := \begin{cases} \frac{f(x)+h(x)}{g(x)}, & \text{if } x \in \Omega \cap \operatorname{dom}(f), \\ +\infty, & \text{else,} \end{cases}$$

其中f, g, h为问题(1.1)中的目标函数，因此可以将问题(1.1)写为以下形式：

$$\min \{F(x) : x \in R^n\}. \tag{3.1}$$

在展示算法之前，我们先给出问题(1.1)的一阶最优性条件。下列定义与文献 [?] 中关于稳定点的定义方式相同。

定义3.1. 假设$x^\star \in dom(F)$, $c_\star = F(x^\star)$, 若

$$0 \in \partial f(x^\star) + \nabla h(x^\star) - c_\star \partial g(x^\star). \tag{3.2}$$

则$x^\star$是F的稳定点。

接下来，我们将给出算法NL-PGSA的完整框架。

步骤0 **输入**$x^0 \in \mathrm{dom}(F)$, $a > 0$, $0 < t < 1$, $N \geq 0$. **设**$k = 0$.

步骤1 **计算**$y^{k+1} \in \partial g(x^k)$, $c_k = \frac{f(x^k)+h(x^k)}{g(x^k)}$,

选择$\alpha_{k,0} \in (0, \frac{2}{L})$.

步骤2 **当**$m = 0, 1, \ldots,$**时**

$\alpha_k = \alpha_{k,0} t^m$,

$\tilde{x}^{k+1} \in prox_{\alpha_k f}(x^k - \alpha_k \nabla h(x^k) + \alpha_k c_k y^{k+1})$.

若$\tilde{x}^{k+1} \in \mathrm{dom}(F)$**且**

$$F\left(\tilde{x}^{k+1}\right) \leq \max_{[k-N]_+ \leq i \leq k} c_i - \frac{a}{2} \left\|\tilde{x}^{k+1} - x^k\right\|_2^2, \tag{3.3}$$

令$x^{k+1} = \tilde{x}^{k+1}$**转步骤3.**

步骤3 **令**$k \leftarrow k + 1$**转步骤1.**

根据不等式(3.3)可得，当$N = 0$时，$\left\{F(x^k) : k \in N\right\}$是单调的，当$N > 0$时，$\left\{F(x^k) : k \in N\right\}$是不单调的。为简便，采用记号$\triangle x = x^k - x^{k-1}$, $\triangle h = \nabla h(x^k) - \nabla h(x^{k-1})$, 关于参数$\alpha_{k,0}$ 的生成方式如下：

$$\alpha_{k,0} = \begin{cases} \max\left\{\underline{\alpha}, \min\left\{\bar{\alpha}, \frac{|\langle \Delta x, \Delta h\rangle|}{\|\Delta h\|_2^2}\right\}\right\}, & \text{if } \langle \Delta x, \Delta h\rangle \neq 0, \\ \bar{\alpha}, & \text{else}. \end{cases} \tag{3.4}$$

接下来，我们给出如下引理，该引理可以进一步帮助我们验证算法的收敛性。

引理3.2. 设$\left\{x^k : k \in N\right\}$是*NL-PGSA*为解决问题(1.1) 所生成的迭代点列，则

$$f(x^{k+1}) + h(x^{k+1}) + (\frac{1}{\alpha_k} - \frac{L}{2}) \left|\left|x^{k+1} - x^k\right|\right|^2 \leq c_k g(x^{k+1}), \tag{3.5}$$

且$\forall k \in N$, $\left\{x^k\right\} \subseteq \mathrm{dom}(F)$.

证明. 我们采用数学归纳法进行证明，首先初始点$x^0 \in dom(F)$, 设$\forall k \in N$, 都有$x^0, x^1, \cdots, x^k \in dom(F)$，由*NL-PGSA*的迭代框架以及近端算子的定义，可得

$$\frac{1}{\alpha_k}\left(x^k - x^{k+1} - \alpha_k \nabla h\left(x^k\right) + \alpha_k c_k y^{k+1}\right) \in \partial f\left(x^{k+1}\right),$$

根据假设1.1，f是凸函数，可以得到

$$f\left(x^{k+1}\right)+\frac{1}{\alpha_k}\left\langle x^k-x^{k+1}-\alpha_k\nabla h\left(x^k\right)+\alpha_k c_k y^{k+1},x^k-x^{k+1}\right\rangle\leq f\left(x^k\right)$$

进一步整理得

$$f\left(x^{k+1}\right)+\left\langle x^{k+1}-x^k,\nabla h\left(x^k\right)-c_k y^{k+1}\right\rangle+\frac{1}{\alpha_k}\left\|x^{k+1}-x^k\right\|_2^2\leq f\left(x^k\right),\tag{3.6}$$

此外，∇h是$Lipschitz$连续函数，可得

$$h(x^{k+1})\leq h(x^k)+\left\langle\nabla h(x^k),x^{k+1}-x^k\right\rangle+\frac{L}{2}\left|\left|x^{k+1}-x^k\right|\right|^2,\tag{3.7}$$

g是凸函数，且$\forall k\in N$，$c_k=\frac{f(x^k)+h(x^k)}{g(x^k)}\geq 0$，因此

$$c_k g(x^k)+\left\langle c_k y^{k+1},x^{k+1}-x^k\right\rangle\leq c_k g(x^{k+1}).\tag{3.8}$$

将(3.6)，(3.7)和(3.8)相加，结合$c_k g(x^k)=\frac{f(x^k)+h(x^k)}{g(x^k)}g(x^k)=f(x^k)+h(x^k)$，故

$$f(x^{k+1})+h(x^{k+1})+(\frac{1}{\alpha_k}-\frac{L}{2})\left|\left|x^{k+1}-x^k\right|\right|^2\leq c_k g(x^{k+1}),$$

设$x^{k+1}\notin dom(F)=\Omega\cap\mathrm{dom}(f)$，因此$g(x^{k+1})=0$，根据假设1.1，$f+h\geq 0$，由算法NL-PGSA可得，$0<\alpha_k<\frac{2}{L}$，结合(3.5)，可得$x^{k+1}=x^k$，这与$x^k\in dom(F)$矛盾，因此$\forall k\in N$，$x^{k+1}\in dom(F)$。引理得证。 □

借助引理3.2，进一步得到以下重要结果。

引理3.3. 设$\left\{x^k:k\in N\right\}$是NL-PGSA生成的序列，则

(i) $\forall k\in N$，$F\left(x^{k+1}\right)+(\frac{1}{\alpha_k}-\frac{L}{2})\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}\leq F\left(x^k\right)$；

(ii) $\lim_{k\to\infty}\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}=0$;

(iii) $\lim_{k\to\infty}F\left(x^k\right)=\lim_{k\to\infty}c_k=c_\star$;

证明. 我们首先证明(i)，根据引理3.2，以及$F(x^{k+1})=c_{k+1}=\frac{f(x^{k+1})+h(x^{k+1})}{g(x^{k+1})}$且$g(x^{k+1})>0$ 可得

$$F\left(x^{k+1}\right)+(\frac{1}{\alpha_k}-\frac{L}{2})\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}\leq F\left(x^k\right),\tag{3.9}$$

接下来，我们证明(ii)，将不等式(3.9)两边的k从$k=0$到K 求和，可得

$$F\left(x^{K+1}\right)+\sum_{k=0}^{K}(\frac{1}{\alpha_k}-\frac{L}{2})\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}\leq F\left(x^0\right).\tag{3.10}$$

基于上式以及$0<\alpha_k<\frac{2}{L}$,可得$\sum_{k=0}^{K}(\frac{1}{\alpha_k}-\frac{L}{2})\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}$有界,令$K\to\infty$,可得$\lim_{k\to\infty}\frac{\left\|x^k-x^{k-1}\right\|^2}{g(x^{k+1})}=0$.

根据*(i)*和*(ii)*,可知*(iii)*成立。 □

接下来，我们分析NL-PGSA的收敛性，在适当的假设下，我们首先证明对于每个外部循环，其关联的内部循环必须在有限次迭代内终止。我们首先给出下列假设。

假设条件3.1. 对任意的$c_0\in R$, 水平集$\{x\in R^n: F(x)\le c_0\}$ 有界。

假设条件3.2. f在其有效域上是局部$Lipschitz$连续的。

假设条件3.3. g在Ω上是一个连续可微函数且∇g是$Lipschitz$连续的。

引理3.4. 在假设*3.1*成立的条件下，设$M:=sup\{g(x):x\in lev(F(x),c_0)\}$, 则

(i) 在多数T次迭代中，算法*NL-PGSA*的第*3*步会在$\alpha_k\ge\hat{\alpha}$时终止，其中参数$\hat{\alpha}=\frac{2t}{L+\alpha M}$, $T:=\left\lceil\frac{-\log(\hat{\alpha}(aM+L))}{\log t}+1\right\rceil$;

(ii) $\forall k\in N$, 有$x^k\in lev(F,c_0)$且$\left\{x^k:k\in N\right\}$ 是有界的;

(iii) 序列$\left\{F(x^{l(k)}):k\in N\right\}$是非增的。

证明. 首先我们证明*(i)*, 根据引理*3.3*中的(3.10)， 可得$\forall k\in N$, $F(x^k)\le F(x^0)$, 借助假设*3.1*, 可以得到$\left\{x^k:k\in N\right\}$是有界的。此外，因为$g$是凸函数，故$g$是连续函数，因此存在$M\ge 0$, 使得对任意的$k\in N$, 有$g(x^k)\le M$, 根据算法*NL-PGSA*的迭代框架，$\alpha_{k,0}<\frac{2}{L}$, 因此算法在迭代$T$步之后，对任意的$k\in N$, 有$\alpha_k\le\frac{2}{L+\alpha M}=\hat{\alpha}/t$.

接下来，我们对k采用数学归纳法，显然$x^0\in lev(F,c_0)$, 设x^i 是由算法迭代产生的点，并且$x^i\in lev(F,c_0)$, $i=0,1,\cdots,k$, 为了证明*(i)*, 需要说明当$\alpha_k\le\hat{\alpha}/t$ 时，有$\tilde{x}^{k+1}\in\mathrm{dom}(F)$, 且下式成立

$$F\left(\tilde{x}^{k+1}\right)\le c_k-\frac{a}{2}\left\|\tilde{x}^{k+1}-x^k\right\|_2^2, \tag{3.11}$$

由引理*3.3(i)*, 以及$\alpha_k\le\frac{2}{L+\alpha M}<\frac{2}{L}$, 可得$\tilde{x}^{k+1}\in\mathrm{dom}(F)$, 并且

$$F\left(x^{k+1}\right)\le c_k-(\frac{1}{\alpha_k}-\frac{L}{2})\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}\le c_k-\frac{\alpha M}{2}\frac{\left\|x^{k+1}-x^k\right\|^2}{g(x^{k+1})}, \tag{3.12}$$

由上式可知$F\left(\tilde{x}^{k+1}\right)\le c_k\le c_0$, 因此$\tilde{x}^{k+1}\in\mathrm{lev}\left(F,c_0\right)$, 此外，借助$g\left(\tilde{x}^{k+1}\right)\le M$, 可由(3.12)得到(3.11).

接下来，我们证明$x^{k+1}\in\mathrm{lev}\left(F,c_0\right)$, 为简单起见，定义

$$l(k):=\max\left\{j:j\in\arg\max\left\{F\left(x^i\right):[k-N]_+\le i\le k\right\}\right\}, \tag{3.13}$$

借助上述定义以及(3.3)，可得$\forall i \leq k$, $F\left(x^{i+1}\right) \leq F\left(x^{l(i)}\right)$, 因此，

$$\begin{aligned} F\left(x^{l(i+1)}\right) &= \max_{[i+1-N]_+ \leq j \leq i+1} F\left(x^j\right) \\ &= \max\left\{F\left(x^{i+1}\right), \max_{[i+1-N]_+ \leq j \leq i} F\left(x^j\right)\right\} \\ &\leq \max\left\{F\left(x^{i+1}\right), F\left(x^{l(i)}\right)\right\} \\ &= F\left(x^{l(i)}\right). \end{aligned} \tag{3.14}$$

得$F\left(x^{k+1}\right) \leq F\left(x^{l(k)}\right) \leq F\left(x^{l(0)}\right) = c_0$. 引理得证。 □

下述两个定理即文献 [14]中定理5.2与定理5.3, 定理3.5表明NL-PGSA的子序列是收敛的。定理3.6表明NL-PGSA是全局收敛的。

定理3.5. *在假设1.1, 假设3.1成立的前提下,设$\left\{x^k : k \in N\right\}$是NL-PGSA生成的序列, 则$\left\{x^k : k \in N\right\}$的任意聚点都是$F$的临界点。*

定理3.6. *在假设1.1, 3.1, 3.2以及假设3.3成立的前提下，设F在其有效域中的任意一点都满足KL性质，$\left\{x^k : k \in N\right\}$是NL-PGSA生成的序列，则$\sum_{k=1}^{\infty}\left\|x^{k+1} - x^k\right\|_2 < +\infty$ 且$\left\{x^k : k \in N\right\}$收敛到$F$的临界点。*

4. 数值实验

在数值测试中，我们主要考虑以下带约束的l_1/l_2稀疏信号恢复问题：

$$\min\left\{\frac{\|x\|_1}{\|x\|} : Ax = b, \underline{x} \leq x \leq \bar{x}, x \in R^n\right\}, \tag{4.1}$$

其中$\underline{x}$, $\bar{x} \in R^n$分别代表基础信号的下限与上限，通过引入罚参数$\lambda > 0$, 上述带约束的优化问题与下列问题等价：

$$\min\left\{\frac{\lambda\|x\|_1 + \frac{1}{2}\|Ax - b\|^2}{\|x\|} : \underline{x} \leq x \leq \bar{x}, x \in R^n\right\}. \tag{4.2}$$

显然，问题(4.2)是问题(1.1)的特例，具体来说，f是函数$\lambda\|x\|_1$ 与示性函数在$\{\underline{x} \leq x \leq \bar{x}, x \in R^n\}$上的和，$h = \frac{1}{2}\|Ax - b\|^2$, $g = \|x\|$, $\Omega = \{x \in R^n : g(x) \neq 0\}$, 这里$A \in R^{m \times n}$, $b \in R^m$. 由于目标函数是一类分式优化问题，因为要确保分母的函数值非零，结合g函数的具体形式，此时初值点不能设为零向量。在这里，我们采用真实信号$\tilde{x}$来生成初值x_0. 接下来，我们构造了一种稀疏度大小为K的真实信号$\tilde{x} \in R^n$, 我们随机选择一个大小为K的支持子集，其最小分离度至少为$2D$, 并在该集合上生成一个支持向量$\tilde{v} \in R^n$, sgn代表标准符号函数，设$\tilde{x} = sgn(\tilde{v})$. 最后，设$b \in R^m$, $b = A\tilde{x}$, $\underline{x} = -2 \times 1_n$, $\bar{x} = 2 \times 1_n$, 这里1_n代表所有分量都等于1的n维向量。

接下来，我们需要指出$\tilde{x}$满足问题(4.2)的一阶必要性条件。首先$\tilde{x}$ 在问题(4.2)处的函数值为$\tilde{c} = \lambda\sqrt{K}$, 因为$\widetilde{x} \in \partial\|\cdot\|_1(\widetilde{x})$且$\nabla\left(\|\cdot\|_2\right)(\widetilde{x}) = \widetilde{x}/\sqrt{K}$, 则$0 \in \partial\left(\lambda\|\cdot\|_1\right)(\widetilde{x}) - \widetilde{c}\nabla\left(\|\cdot\|\right)(\widetilde{x})$. 此外，

$\underline{x} < \tilde{x} < \bar{x}$, $A\tilde{x} = b$, 可得

$$0 \in \partial\left(\lambda\|\cdot\|_1 + \iota_{\{x\in\mathbb{R}^n:\underline{x}\le x\le\bar{x}\}}\right)(\widetilde{x}) + A^T(A\widetilde{x} - b) - \widetilde{c}\nabla\left(\|\cdot\|\right)(\widetilde{x}).$$

这意味着$\tilde{x}$为问题(4.2)的最优解。我们将算法NL-PGSA与算法PGSA-L [14], 进行比较，接下来，我们将给出算法的具体参数设置。

· NL-PGSA. $L = \|A\|^2$, 初始步长$\alpha_{0,0} = 1/L$, $\alpha = 10^{-3}$, $t = 0.5$, $N = 4$.

· PGSA-L. 我们选择的步长参数与文献 [14]一致。

在接下来的实验中，取维数分别为(m,n)=(512,8192)和(m,n)=(640,10240), 选择不同的稀疏度$D = \{1, 5, 10, 15\}$, $K = \{12, 16\}$ 来测试l_1/l_2稀疏信号恢复问题。在每个(D, K) 设置中，我们如上述所述随机生成100次数据，并根据100次数据的平均结果比较上述两种算法。两种算法每次选取的初值为$x_0 = \tilde{x} + 0.4\xi$, 这里$\xi \in R^n$其分量满足$[-1, 1]$ 的均匀分布，当迭代次数达到1000或$\|x^k - x^{k-1}\|/\max\{1, \|x^k\|\} \le 10^{-5}$时迭代终止。如表 1和表 2所示，我们给出两种算法在迭代次数，CPU时间(以秒为单位)和迭代终止时的函数值。从表中我们得出结论，我们的方法表现较好。

Table 1. Solving problem (4.2) when $(m, n) = (512, 8192)$, $\lambda = 5 \times 10^{-4}$

表 1. 解决问题(4.2), $(m, n) = (512, 8192)$, $\lambda = 5 \times 10^{-4}$

	iter		time		F_val	
	l_1/l_2_PGSA-L	l_1/l_2_NL-PGSA	l_1/l_2_PGSA-L	l_1/l_2_NL-PGSA	l_1/l_2_PGSA-L	l_1/l_2_NL-PGSA
D=1,K=12	482	402	2.23297	2.21622	0.098759	0.0907651
D=1,K=16	484	406	2.24221	2.21645	0.098803	0.0906339
D=5,K=12	483	403	2.23481	2.21423	0.098282	0.0907574
D=5,K=16	482	403	2.24031	2.22685	0.098458	0.0908579
D=10,K=12	482	401	2.23166	2.21356	0.098117	0.0907067
D=10,K=16	483	402	2.23102	2.21256	0.098973	0.0905531
D=15,K=12	486	402	2.23308	2.21204	0.098665	0.0909149
D=15,K=16	483	402	2.23208	2.21072	0.098606	0.0902784

Table 2. Solving problem (4.2) when $(m, n) = (640, 10240)$, $\lambda = 5 \times 10^{-4}$

表 2. 解决问题(4.2), $(m, n) = (640, 10240)$, $\lambda = 5 \times 10^{-4}$

	iter		time		F_val	
	l_1/l_2_PGSA-L	l_1/l_2_NL-PGSA	l_1/l_2_PGSA-L	l_1/l_2_NL-PGSA	l_1/l_2_PGSA-L	l_1/l_2_NL-PGSA
D=1,K=12	483	404	3.44432	2.99739	0.192345	0.1917892
D=1,K=16	487	402	3.43452	2.99816	0.193024	0.1915356
D=5,K=12	482	403	3.44371	2.99515	0.192964	0.1912004
D=5,K=16	483	404	3.44735	2.99225	0.192718	0.1911423
D=10,K=12	485	405	3.44345	2.99341	0.192343	0.1911445
D=10,K=16	482	403	3.44462	2.99235	0.192274	0.1911234
D=15,K=12	482	403	3.44351	2.99532	0.192494	0.1912451
D=15,K=16	483	405	3.66209	2.53988	0.254709	0.0563043

5. 结论

本文针对一类特殊的分式优化问题，提出了一种新的非单调线搜索的近端梯度次梯度算法。首先，在保证收敛性的前提下，我们扩大了步长的选择范围，这有利于算法更快地移动到目标函数的更优解附近。另一方面，借助数值实验，我们进一步验证了算法的有效性。

参考文献

[1] Baldacci, R., Lim, A., Traversi, E. and Calvo, R.W. (2020) Optimal Solution of Vehicle Routing Problems with Fractional Objective Function. *Transportation Science*, **54**, 434-452. https://doi.org/10.1287/trsc.2019.0929

[2] Rahimi, Y., Wang, C., Dong, H. and Lou, Y. (2019) A Scale-Invariant Approach for Sparse Signal Recovery. *SIAM Journal on Scientific Computing*, **41**, A3649-A3672. https://doi.org/10.1137/18M123147X

[3] Studer, C. and Baraniuk, R.G. (2014) Stable Restoration and Separation of Approximately Sparse Signals. *Applied and Computational Harmonic Analysis*, **37**, 12-35. https://doi.org/10.1016/j.acha.2013.08.006

[4] Candès, E., Romberg, J.K. and Tao, T. (2006) Stable Signal Recovery from Incomplete and Inaccurate Measurements. *Communications on Pure and Applied Mathematics*, **59**, 1207-1223. https://doi.org/10.1002/cpa.20124

[5] Hoyer, P.O. (2004) Non-Negative Matrix Factorization with Sparseness Constraints. *Journal of Machine Learning Research*, **5**, 1457-1469.

[6] Shen, K. and Yu, W. (2018) Fractional Programming for Communication Systems—Part I: Power Control and Beamforming. *IEEE Transactions on Signal Processing*, **66**, 2616-2630. https://doi.org/10.1109/TSP.2018.2812733

[7] Zappone, A., Sanguinetti, L. and Debbah, M. (2017) Energy-Delay Efficient Power Control in Wireless Networks. *IEEE Transactions on Communications*, **66**, 418-431. https://doi.org/10.1109/TCOMM.2017.2755644

[8] Konno, H. and Inori, M. (1989) Bond Portfolio Ptimization by Bilinear Fractional Programming. *Journal of the Operations Research Society of Japan*, **32**, 143-158. https://doi.org/10.15807/jorsj.32.143

[9] Clemmensen, L., Hastie, T., Witten, D. and Ersbøll, B. (2100) Sparse Discriminant Analysis. *Technometrics*, **53**, 406-413. https://doi.org/10.1198/TECH.2011.08118

[10] Ibaraki, T. (1983) Prametric Approaches to Fractional Programs. *Mathematical Programming*, **26**, 345-362. https://doi.org/10.1007/BF02591871

[11] Pang, J.S. (1980) A Parametric Linear Complementarity Technique for Optimal Portfolio Selection with a Risk-Free Asset. *Operation Research*, **28**, 927-941. https://doi.org/10.1287/opre.28.4.927

[12] Schaible, S. (1976) Fractional Programming. II, on Dinkelbach's Algorithm. *Management Science*, **22**, 868-873. https://doi.org/10.1287/mnsc.22.8.868

[13] Bot, R.I. and Csetnek, E.R. (2017) Proximal-Gradient Algorithms for Fractional Programming. *Optimization*, **66**, 1383-1396. https://doi.org/10.1080/02331934.2017.1294592

[14] Zhang, N. and Li, Q. (2022) First-Order Algorithms for a Class of Fractional Optimization Problems. *SIAM Journal on Optimization*, **32**, 100-129. https://doi.org/10.1137/20M1325381

[15] Attouch, H., Bolte, J., Redont, P. and Soubeyran, A. (2010) Proximal Alternating Minimization and Projection Methods for Nonconvex Problems: An Approach Based on the Kurdyka-Łojasiewicz Inequality. *Mathematics of Operations Research*, **35**, 438-457. https://doi.org/10.1287/moor.1100.0449