分裂可行性问题的外推加速线性交替方向乘子法及其全局收敛性

刘洋, 薛中会, 王永全, 曹永胜

http://www.jsjkx.com
DOI:10.11896/jsjkx.230100009

分裂可行性问题的外推加速线性交替方向乘子法及其全局收敛性

刘　洋[1,2]　**薛中会**[3]　**王永全**[1]　**曹永胜**[1]

1 华东政法大学智能科学与信息法学系　上海 201620

2 上海交通大学智慧法院研究院　上海 200030

3 上海出版印刷高等专科学校信息与智能工程系　上海 200093

(lyang@ecupl.edu.cn)

摘　要　针对在图像重建以及语言处理系统等领域有着广泛应用的分裂可行性问题(SFP)的最优化求解，提出了外推加速线性交替方向乘子法。首先将SFP描述为一个具有线性约束的可分离凸极小化问题；然后引进外推线性交替方向乘子法，利用问题的可分离结构，产生了具有闭式解的子问题，并在适当条件下证明了该算法的全局收敛性；最后，通过数值实验验证了该算法的可行性和有效性。

关键词：分裂可行性问题；线性交替方向乘子法；凸极小化问题；外推加速

中图法分类号　TP301.6

Extrapolation Accelerated Linear Alternating Direction Multiplier Method for Split Feasibility Problems and Its Global Convergence

LIU Yang[1,2], XUE Zhonghui[3], WANG Yongquan[1] and CAO Yongsheng[1]

1 Department of Intelligent Science and Information Law, East China University of Political Science and Law, Shanghai 201620, China

2 China Institute for Smart Court, Shanghai Jiao Tong University, Shanghai 200030, China

3 Department of Information and Intelligent Engineering, Shanghai Publishing and Printing College, Shanghai 200093, China

Abstract　This paper deals with a linear alternating direction multiplier method (ADMM) for Split feasibility problems (SFP). More specifically, the SFP has been formulated as a separable convex minimization problem with linear constraints, and then extrapolation accelerated linear ADMM has been proposed, which takes advantage of the separable structure, and then rising to subproblems with closed-form solutions have been given. Furthermore, the global convergence of the proposed method is proved under some suitable conditions. Moreover, the algorithm has been tested by applying to two SFP examples in our theoretical results.

Keywords　Split feasible problem, ADMM, Separable convex problem, Extrapolation accelerated

1　引言

分裂可行性问题(Split Feasibility Problem, SFP)最早由Censer等提出[1]，主要用来解决图像重建和调强放射治疗中的逆问题[2]。设 $\mathfrak{R}^N$ 表示 N 维欧几里德空间，则分裂可行性问题一般可表达为如下形式：

$$x^* \in C, \boldsymbol{A}x^* \in Q \tag{1}$$

其中，C 和 Q 分别是 $\mathfrak{R}^N$ 和 $\mathfrak{R}^M$ 非空闭凸子集，$\boldsymbol{A}:\mathfrak{R}^N \rightarrow \mathfrak{R}^M$ 是给定的有界线性算子。目前，用于解决SFP[3-6]问题的算法有很多，其中包括由Rockafellar提出[7]的增广拉格朗日方法(Augmented Lagrangian Method, ALM)，该方法用于求解线性凸优化问题[8]。

交替方向乘子法(Alternating Direction Multiplier Method, ADMM)作为ALM的一种变形，通常用于解决以下问题：

$$\begin{aligned} &\min_{x,y}\ H(x)+G(y) \\ &\text{s.t.}\ \ \boldsymbol{B}x+\boldsymbol{C}y=b \end{aligned} \tag{2}$$

其中，$H:\mathfrak{R}^N \rightarrow \mathfrak{R}$，$G:\mathfrak{R}^P \rightarrow \mathfrak{R}$ 是凸的、真的和下半连续函数，$\boldsymbol{B}\in\mathfrak{R}^{M\times N}$，$\boldsymbol{C}\in\mathfrak{R}^{M\times P}$，$b\in\mathfrak{R}^M$。图像处理中的许多无约束极小化问题都可以使用式(2)问题的可验证分裂方案进行约束，

到稿日期：2023-01-03　返修日期：2023-04-10

基金项目：国家社科基金重大项目(20&ZD199,21&ZD200)；教育部人文社科青年基金项目(20YJC820030)；国家新闻出版署“智能与绿色柔板印刷”重点实验室招标课题(KLIGFP-02)

This work was supported by the National Social Science Fund Major Project of China (20&ZD199,21&ZD200), Youth Fund for Humanities and Social Sciences of the Ministry of Education (20YJC820030) and Bidding Project of the Key Laboratory of “Intelligent and Green Flexographic Printing” of the State Press and Publication Administration(KLIGFP-02).

通信作者：薛中会(hnlgxzh@163.com)

因此 ADMM 算法成为图像处理中最常用的算法之一[9-10]。

基于交替技术的 l_1-范数正则分裂可行性问题变分不等式框架下的方向乘子法，He 等提出了一种可实现的分裂算法[11]，数值实验验证了该算法的有效性。受此启发，本文提出了一种可实现的加速 ADMM 算法，即外推加速线性交替方向乘子法(Extrapolation Accelerated Linear Alternating Direction Multiplier Method, EAL-ADMM)，首先将 SFP 表示为式(2)问题所示的一个可分凸线性约束下的极小化问题，然后引入线性 ADMM 来求解可分离结构和具有闭式解的子问题，并证明了该算法的全局收敛性。

2 预备知识

$\forall x, y \in \Re^N$，$\langle x, y\rangle$ 表示标准的内积，$\|x\|_M=\sqrt{\langle x, Mx\rangle}$ 表示 x 的 M-范数。对于任意的矩阵 $\boldsymbol{A}$，$\|\boldsymbol{A}\|$ 表示矩阵 $\boldsymbol{A}$ 的 2-范数。

定义 1 P_Ω 表示在 Ω 上的正交投影，即 $P_\Omega(x)=\arg\min_{y\in\Omega}\|x-y\|$。

对于任意 x, y，正交投影算子 P_Ω 具有以下性质：

引理 1[10] 设 Q 是 $\Re^N$ 上的非空闭凸子集，P_Q 是 x 在 Ω 上的投影，那么对于任何 $x, y\in\Re^N$ 和 $z\in Q$：

(1) $\langle P_\Omega(x)-x, z-P_\Omega(x)\rangle\geqslant 0$；

(2) $\|P_\Omega(x)-P_\Omega(y)\|^2\leqslant\langle P_\Omega(x)-P_\Omega(y), x-y\rangle$；

(3) $\|P_\Omega(x)-z\|^2\leqslant\|x-z\|^2-\|P_\Omega(x)-x\|^2$；

(4) $\|P_\Omega(x)-P_\Omega(y)\|\leqslant\|x-y\|-\|P_\Omega(x)-x+y-P_\Omega(y)\|$。

定义 2 设 $S\subset\Re^N$，T 是 $\Re^N\to\Re^N$ 的映射，如果 $\langle T(x)-T(y), x-y\rangle\geqslant 0, \forall x, y\in S$，那么称 T 在 S 上是单调的。对于单调映射 T，当且仅当 $x-y=0$ 时，$\langle T(x)-T(y), x-y\rangle=0$ 成立，我们称 T 是严格单调映射。

定义 3 如果 $\|T(x)-T(y)\|\leqslant\|x-y\|, \forall x, y\in S$，那么称映射 $T: S\to\Re^N$ 是非扩张的。由上述定义可知，投影算子 T 是单调非扩张映射。

3 外推加速线性交替方向乘子法(EAL-ADMM)

3.1 ADMM 算法

首先将 SFP 重新表述为以下最小化问题：

$$\min F(x)=\frac{1}{2}\|\boldsymbol{A}x-P_Q(\boldsymbol{A}x)\|^2+\frac{1}{2}\|x-P_C(x)\|^2 \tag{3}$$

一种常用的求解式(3)的方法是直接在 $\Re^N$ 上寻找 $F(x)$ 的最小值。显然，利用 $F(x)$ 的梯度构造一个迭代算法，可以在 $\Re^N$ 中生成一个点序列收敛到包含 x^* 的式(3)的解集。但如果矩阵 $\boldsymbol{A}$ 是病态的，则收敛速度较慢。针对此问题，首先引入附加变量 $y\in\Re^M$，在原始变量 $x\in\Re^N$ 中增加约束条件 $\boldsymbol{A}x-y=0$，并在 $\Re^N\times\Re^M$ 中考虑如下约束问题：

$$\begin{aligned}&\min F(x)=\frac{1}{2}\|y-P_Q(y)\|^2+\frac{1}{2}\|x-P_C(x)\|^2\\&\text{s.t.}\ \boldsymbol{A}x-y=0\end{aligned} \tag{4}$$

其中，$(x, y)\in\Re^N\times\Re^M$。显然，问题(4)等价于问题(3)，然后我们用对偶方法消除刚才引入的约束条件。

问题(4)经典的对偶方法是引入一个对应于约束的 Lagrange 乘子 $\lambda\in\Re^N$ 和 Lagrange 函数 L，并且将其定义在 $(\Re^N\times\Re^M)\times\Re^M$ 上，值域为 $(-\infty, +\infty)$，一般表达形式为：

$$\begin{aligned}L(x, y; \lambda)=&\frac{1}{2}\|x-P_C(x)\|^2+\frac{1}{2}\|y-P_Q(y)\|^2+\\&\langle\lambda, \boldsymbol{A}x-y\rangle\end{aligned} \tag{5}$$

因此式(4)的增广拉格朗日函数表达形式为：

$$\begin{aligned}L_\beta(x, y; \lambda)=&\frac{1}{2}\|x-P_C(x)\|^2+\frac{1}{2}\|y-P_Q(y)\|^2+\\&\langle\lambda, \boldsymbol{A}x-y\rangle+\frac{\beta}{2}\|\boldsymbol{A}x-y\|^2\end{aligned} \tag{6}$$

其中 $\lambda\in\Re^N$ 是拉格朗日乘子，$\beta>0$ 是违反线性约束的罚参数。从 (x^k, y^k, λ^k) 到 $(x^{k+1}, y^{k+1}, \lambda^{k+1})$ 的每一步 ALM 迭代方案为：

$$\begin{aligned}&(x^{k+1}, y^{k+1})\leftarrow\arg\min_{x\in\Re^N, y\in\Re^M}L_\beta(x, y; \lambda^k)\\&\lambda^{k+1}\leftarrow\lambda^k-\beta(x^{k+1}-y^{k+1})\end{aligned} \tag{7}$$

值得注意的是，在 ALM 算法中，式(7)没有考虑它的可分离结构，因此变量 x^{k+1} 和 y^{k+1} 必须同时求解。相反，在 ADMM 算法中，将式(7)中的最小化问题分解为在该方案下连续求解 x^{k+1} 和 y^{k+1} 的两个子问题。

$$x^{k+1}\leftarrow\arg\min_{x\in\Re^N}L_\beta(x, y^k; \lambda^k) \tag{8a}$$

$$y^{k+1}\leftarrow\arg\min_{y\in\Re^M}L_\beta(x^{k+1}, y; \lambda^k) \tag{8b}$$

$$\lambda^{k+1}\leftarrow\lambda^k-\beta(x^{k+1}-y^{k+1}) \tag{8c}$$

鉴于这种分解特性，可以利用式(7)的目标函数中的可分离结构。接下来通过引入线性化思想来考虑式(8)中两个子问题的解。

易证式(8)中的 x-子问题和 y-子问题等价于式(9)和式(10)。

$$x^{k+1}=\arg\min_{x\in\Re^N}\left\{\frac{1}{2}\|x-P_C(x)\|^2+\frac{\beta}{2}\left\|\boldsymbol{A}x-y^k-\frac{\lambda^k}{\beta}\right\|^2\right\} \tag{9}$$

$$y^{k+1}=\arg\min_{y\in\Re^M}\left\{\frac{1}{2}\|y-P_Q(y)\|^2+\frac{\beta}{2}\left\|y-\boldsymbol{A}x^{k+1}+\frac{\lambda^k}{\beta}\right\|^2\right\} \tag{10}$$

接下来，将对式(9)和式(10)进行修正。尽管修正后式(9)和式(10)中的最小化没有被准确求解，但提供了问题(4) (x^{k+1}, y^{k+1}) 的一个近似解。

3.2 ADMM 算法线性化

以下通过线性化式(9)和式(10)的第一个二次项来处理子问题。

引入辅助凸集 $\Omega_1\subset\Re^N$，$\Omega_2\subset\Re^M$，对式(9)和式(10)中的第一个二次项进行线性化，子问题简化为：

$$\begin{aligned}\tilde{x}=&\arg\min_{x\in\Omega_1}\{[\langle x^k-P_C(x^k), x-x^k\rangle+\\&\frac{\tau}{2}\|x-x^k\|^2]+\frac{\beta}{2}\left\|\boldsymbol{A}x-y^k-\frac{\lambda^k}{\beta}\right\|^2\}\end{aligned} \tag{11}$$

$$\begin{aligned}\tilde{y}=&\arg\min_{y\in\Omega_2}\left\{\left[\langle y^k-P_Q(y^k)\rangle+\frac{\tau}{2}\|y-y^k\|^2\right]+\right.\\&\left.\frac{\beta}{2}\left\|y-\boldsymbol{A}\tilde{x}^k+\frac{\lambda^k}{\beta}\right\|^2\right\}\end{aligned} \tag{12}$$

由式(11)和式(12)的最优性条件得出以下方程式：

$$\tilde{x}^k = P_\Omega\{\frac{1}{(\tau+\beta\rho(A^{T}A))}[P_C(x^k)-x^k+\tau x^k+\beta A^{T}y^k+A^{T}\lambda^k]\} \tag{13}$$

$$\tilde{y}^k = P_{\Omega_2}\left\{\frac{1}{(\tau+\beta)}[P_Q(y^k)-y^k+\tau y^k+\beta A\tilde{x}^k-\lambda^k]\right\} \tag{14}$$

3.3　基于线性化 ADMM 方法的快速数值算法(EAL-ADMM)

设

$$M=\begin{bmatrix}\tau I_M & 0 & 0\\ 0 & \tau I_N & 0\\ 0 & 0 & \frac{1}{\beta I_N}\end{bmatrix}$$

$\eta=(\eta_1(x,\tilde{x})\eta_2(y,\tilde{y})0)^{T}$，$\eta_1(x,\tilde{x}):=\nabla f(x)-\nabla f(\tilde{x})$，$\eta_2(y,\tilde{y}):=\nabla g(y)-\nabla g(\tilde{y})$，其中$\nabla f(x)=x-P_C(x)$，$\nabla g(y)=y-P_Q(y)$。基于在第 3.1 节和第 3.2 节中的讨论，现导出两种不同的算法来求解式(4)。

设 $\Omega_1\subset\Re^N$，$\Omega_2=\Re^M$，是辅助凸集。

取 $\beta>0$，$\tau>0$ 和 $\tau>(\beta^2+1)/\beta$，初始化$(x,y\lambda)=(x^0,y^0,\lambda^0)$，$k=0$。

$$\tilde{x}^k = P_{\Omega_1}\left\{\frac{1}{(\tau+\beta)}[P_C(x^k)-x^k+\tau x^k+\beta A^{T}y^k+A^{T}\lambda^k]\right\}$$

$$\tilde{y}^k = P_{\Omega_2}\left\{\frac{1}{(\tau+\beta)}[P_Q(y^k)-y^k+\tau y^k+\beta A\tilde{x}^k-\lambda^k]\right\}$$

$$\tilde{\lambda}^k=\lambda^k-\beta(A\tilde{x}^k-\tilde{y}^k) \tag{15}$$

通过

$$\omega^{k+1}=\omega^k-t\alpha_k d(\omega^k,\tilde{\omega}^k) \tag{16}$$

更新下一个迭代 $\omega^{k+1}=(x^{k+1},y^{k+1},\lambda^{k+1})$，其中

$$\alpha_k=\frac{\varphi(\omega^k,\tilde{\omega}^k)}{\|d(\omega^k,\tilde{\omega}^k\|_M^2}\quad(\text{外推因子}) \tag{17}$$

$$\phi(\omega^k,\tilde{\omega}^k)=\langle\omega^k-\tilde{\omega}^k,Md(\omega^k,\tilde{\omega}^k)\rangle \tag{18}$$

$$d(\omega^k,\tilde{\omega}^k)=\omega^k-\tilde{\omega}^k-M^{-1}\eta(\omega^k,\tilde{\omega}^k) \tag{19}$$

为了保证收敛性，近端参数 τ 足够大，$\tau>\frac{\beta^2+1}{\beta}$。整个迭代过程不需要计算 $\rho(A^{T}A)$或A^{-1}，有效避免了矩阵 A 的病态情况发生。

4　EAL-ADMM 算法的全局收敛性

下面证明算法 EAL-ADMM 的全局收敛性。

由问题(4)的一阶最优性条件导出一个基本事实，即问题(4)等价于求向量 $\omega^*\in\Omega:=C\times Q\times\Re^M$ 的变分不等式(Va-riable Inequality，VI)，其性质如下：

$$\langle\omega-\omega^*,\tilde{F}(\omega^*)\rangle\geqslant0,\omega\in\Omega \tag{20a}$$

$$\omega:=(x,y,\lambda)^{T}$$
$$\tilde{F}(\omega):=(\nabla f(x)-A^{T}\lambda,\nabla g(y)+\lambda,Ax-y)^{T} \tag{20b}$$

引理 2　若$\forall\omega',\omega\in\Omega$，$(\omega'-\omega)^{T}(\tilde{F}(\omega')-\tilde{F}(\omega))\geqslant0$，则$\tilde{F}(\omega)$单调。

证明：这个证明较为基础，过程省略。

引理 3　假设 ω^* 是式(20)的解。设

$$M=\begin{vmatrix}\tau I_N & 0 & 0\\ 0 & \tau I_M & 0\\ 0 & 0 & \frac{1}{\beta}I_M\end{vmatrix}$$

那么得到

$$\langle\omega^k-\omega^*,Md(\omega^k-\tilde{\omega}^k)\rangle\geqslant\phi(\omega^k,\tilde{\omega}^k) \tag{21}$$

证明：首先，迭代方案式(11)和式(12)可以很容易变化成一个变分不等式，即

$$\langle x-\tilde{x}^k,x^k-P_C(x^k)+\tau(\tilde{x}^k-x^k)+\beta A^{T}A\tilde{x}^k-\beta A^{T}y^k\rangle\geqslant0,x\in\Omega_1 \tag{22}$$

和

$$\langle y-\tilde{y}^k,y^k-P_Q(y^k)+\tau(\tilde{y}^k-y)+\beta\tilde{y}^k-\beta A\tilde{x}^k+\lambda^k\rangle\geqslant0,y\in\Omega_2 \tag{23}$$

使用式(15)并重新排列上述不等式的项，得到

$$\langle x-\tilde{x}^k,\nabla f(\tilde{x}^k)-A^{T}\tilde{\lambda}^k+\tau(\tilde{x}^k-x^k)+\eta_1(x^k,\tilde{x}^k)\rangle\geqslant0 \tag{24}$$

$$\langle y-\tilde{y}^k,\nabla g(\tilde{y}^k)+\tilde{\lambda}^k+\tau(\tilde{y}^k-y^k)+\eta_2(y^k,\tilde{y}^k)\rangle\geqslant0 \tag{25}$$

和

$$\left\langle\lambda-\tilde{\lambda}^k,(A\tilde{x}^k-\tilde{y}^k)-\frac{1}{\beta}(\lambda^k-\tilde{\lambda}^k)\right\rangle=0 \tag{26}$$

在式(7)－式(9)中设置 $\omega=(x,y,\lambda)=\omega^*=(x^*,y^*,\lambda^*)$，得到

$$\langle\omega^*-\tilde{\omega}^k,\tilde{F}(\tilde{\omega}^k)-Md(\omega^k-\tilde{\omega}^k)\rangle\geqslant0 \tag{27}$$

等价于

$$\langle\tilde{\omega}^k-\omega^*,Md(\omega^k-\tilde{\omega}^k)\rangle\geqslant\langle\tilde{\omega}^k-\omega^*,\tilde{F}(\tilde{\omega}^k)\rangle \tag{28}$$

因为 $\tilde{F}$ 是单调的，所以

$$\langle\tilde{\omega}^k-\omega^*,\tilde{F}(\tilde{\omega}^k)\rangle\geqslant\langle\tilde{\omega}^k-\omega^*,F(\omega^*)\rangle\geqslant0$$

结合式(28)，表明$\langle\tilde{\omega}^k-\omega^*,Md(\omega^k-\tilde{\omega}^k)\rangle\geqslant0$。

因此，$\langle\omega^k-\omega^*,Md(\omega^k-\tilde{\omega}^k)\rangle=\langle\omega^k-\tilde{\omega},Md(\omega^k-\tilde{\omega}^k)\rangle+\langle\tilde{\omega}^k-\omega^*,Md(\omega^k-\tilde{\omega}^k)\rangle\geqslant\varphi(\omega^k,\tilde{\omega}^k)$。

引理 4　假设 ω^* 是式(20)的解。定义两个函数

$$\omega^{k+1}(\alpha):=\omega^k-\alpha d(\omega^k-\tilde{\omega}^k) \tag{29}$$

$$\Theta(\alpha):=\|\omega^k-\omega^*\|_M^2-\|\omega^{k+1}(\alpha)-\omega^*\|_M^2 \tag{30}$$

$$q(\alpha)=2\alpha\phi(\omega^k-\tilde{\omega}^k)-\alpha^2\|d(\omega^k-\tilde{\omega}^k)\|_M^2 \tag{31}$$

那么可以得到

$$\Theta(\alpha)\geqslant q(\alpha) \tag{32}$$

证明：由引理 3、式(29)和式(30)得

$$\begin{aligned}\Theta(\alpha)&=\|\omega^k-\omega^*\|_M^2-\|\omega^{k+1}(\alpha)-\omega^*\|_M^2\\&=\|\omega^k-\omega^*\|_M^2-\|\omega^k-\alpha d(\omega^k,\tilde{\omega}^k)-\omega^*\|_M^2\\&=2\alpha\langle\omega^k-\omega^*,Md(\omega^k,\tilde{\omega}^k)\rangle-\alpha^2\|d(\omega^k,\tilde{\omega}^k)\|_M^2\\&\geqslant2\alpha\varphi(\omega^k,\tilde{\omega}^k)-\alpha^2\|d(\omega^k,\tilde{\omega}^k)\|_M^2\end{aligned}$$

由于 $q(\alpha)$是 α 的二次函数，易见 $q(\alpha)$在式(33)处达到最大值，即

$$\alpha_k:=\frac{\phi(\omega^k,\tilde{\omega}^k)}{\|d(\omega^k,\tilde{\omega}^k)\|_M^2} \tag{33}$$

可得对于任何松弛因子 $t>0$，$\Theta(t\alpha)\geqslant2t\alpha_k\varphi(\omega^k,\tilde{\omega}^k)-t^2\alpha_k^2\|d(\omega^k,\tilde{\omega}^k)\|_M^2=t(2-t)\alpha_k\varphi(\omega^k-\tilde{\omega}^k)$。

为了保证每次迭代都能得到改进效果，限制 $t\in(0,2)$，使

上述关系中的右边为正。在实际应用中，可以选择 $t\in[1,2)$ 进行快速收敛。

引理 5 假设 $\tau>\frac{\beta^2+1}{\beta}$，那么式(17)的步长 α_k 从零以下有界，即对于某些正常数 $\alpha_{\min}$，$\inf_{k\geqslant 1}\alpha_k\geqslant\alpha_{\min}>0$。

证明：应用 Cauchy-Schwartz 不等式得到

$$2\langle a,b\rangle\leqslant k\ \|\alpha\|^2+\frac{1}{k}\|b\|^2, a,b\in\Re^N, k>0$$

$$\begin{aligned}\phi(\omega^k,\tilde{\omega}^k)&=\langle\omega^k-\tilde{\omega}^k,\boldsymbol{M}d(\omega^k,\tilde{\omega}^k)\rangle\\&=\langle\omega^k-\tilde{\omega}^k,\boldsymbol{M}(\omega^k-\tilde{\omega}^k)\rangle-\langle\omega^k-\tilde{\omega}^k,\eta(\omega^k,\tilde{\omega}^k)\rangle\tau\\&\quad\|\tilde{x}^k-x^k\|^2+\tau\ \|\tilde{y}^k-y^k\|^2+\\&\quad\frac{1}{\beta}\|\tilde{\lambda}^k-\lambda^k\|^2-\langle x^k-\tilde{x}^k,\eta_1(x^k-\tilde{x}^k)\rangle-\\&\quad\langle y^k-\tilde{y}^k,\eta_1(y^k,\tilde{y}^k)\rangle\end{aligned}$$

利用 Cauchy-Schwartz 不等式和引理 1 中的式(4)，得到

$$-\langle x^k-\tilde{x}^k,\eta_1(x^k-\tilde{x}^k)\rangle\geqslant-\beta\ \|x^k-\tilde{x}^k\|^2-\frac{1}{\beta}\|x^k-\tilde{x}^k\|^2$$

和 $-\langle y^k-\tilde{y}^k,\eta_2(y^k,\tilde{y}^k)\rangle\geqslant-\beta\ \|y^k-\tilde{y}^k\|^2-\frac{1}{\beta}\|y^k-\tilde{y}^k\|^2$。

那么

$$\begin{aligned}\varphi(\omega^k,\tilde{\omega}^k)\geqslant&\left(\tau-\frac{\beta^2+1}{\beta}\right)\|\tilde{x}^k-x^k\|^2+\\&\left(\tau-\frac{\beta^2+1}{\beta}\right)\|\tilde{y}^k-y^k\|^2+\\&\frac{1}{\beta}\|\tilde{\lambda}^k-\lambda^k\|^2\geqslant C_{\min}\|\tilde{\omega}^k-\omega^k\|^2\end{aligned}\qquad(34)$$

其中，最后一个不等式引入了一个常数 $C_{\min}:=\min\left\{\tau-\frac{\beta^2+1}{\beta},\frac{1}{\beta}\right\}$。

另一方面，根据 ∇f 和 ∇g 的单调性，得到：

$$\begin{aligned}\|d(\omega^k,\tilde{\omega}^k)\|_{\boldsymbol{M}}^2&=\langle d(\omega^k,\tilde{\omega}^k),\boldsymbol{M}d(\omega^k,\tilde{\omega}^k)\rangle\\&=\|\omega^k-\tilde{\omega}^k\|_{\boldsymbol{M}}^2-2\langle x^k-\tilde{x}^k,\eta_1(x^k,\tilde{x}^k)\rangle-\\&\quad2\langle y^k-\tilde{y}^k,\eta_2(y^k-\tilde{y}^k)\rangle+\\&\quad\|\eta(\omega^k,\tilde{\omega}^k)\|^2\\&\leqslant\|\omega^k-\tilde{\omega}^k\|_{\boldsymbol{M}}^2-2\langle x^k-\tilde{x}^k,\eta_1(x^k,\tilde{x}^k)\rangle-\\&\quad2\langle y^k-\tilde{y}^k,\eta_1(y^k-\tilde{y}^k)\rangle+\\&\quad\|\eta_1(x^k,\tilde{x}^k)\|^2+\|\eta_2(y^k,\tilde{y}^k)\|^2\\&\leqslant(\tau+1)\|x^k-\tilde{x}^k\|^2+(\tau+1)\\&\quad\|y^k-\tilde{y}^k\|^2+\frac{1}{\beta}\|\tilde{\lambda}^k-\lambda^k\|^2\\&\leqslant C_{\max}\|\omega^k-\tilde{\omega}^k\|^2\end{aligned}\qquad(35)$$

其中，$C_{\max}:=\min\{\tau+1,\frac{1}{\beta}\}$。因此 $\alpha_k\geqslant\frac{C_{\min}}{C_{\max}}=:\alpha_{\min}>0$。证毕。

定理 1 设 ω^* 为式(32)的任意解，由 EAL-ADMM 算法生成的序列 $\{\omega^k\}$ 满足以下性质：

$$\|\omega^{k+1}-\omega^*\|_{\boldsymbol{M}}^2\leqslant\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2-t(2-t)\alpha_{\min}C_{\min}\|\omega^k-\tilde{\omega}^k\|^2\qquad(36)$$

证明：由式(16)、式(34)、式(35)以及引理 3 和引理 5，可以得

$$\begin{aligned}\|\omega^{k+1}-\omega^*\|_{\boldsymbol{M}}^2&=\|\omega^k-t\alpha_kd(\omega^k,\tilde{\omega}^k)-\omega^*\|_{\boldsymbol{M}}^2\\&\leqslant\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2-2t\alpha k\langle\omega^k-\omega^*,\\&\quad\boldsymbol{M}d(\omega^k,\tilde{\omega}^k)\rangle+t^2\alpha_k^2\ \|d(\omega^k,\tilde{\omega}^k)\|_{\boldsymbol{M}}^2\\&\leqslant\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2-2t\alpha_k\phi(\omega^k,\tilde{\omega}^k)+\\&\quad t^2\alpha_k^2\ \|d(\omega^k,\tilde{\omega}^k)\|_{\boldsymbol{M}}^2\\&=\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2-t(2-t)\alpha_k\phi(\omega^k,\tilde{\omega}^k)\\&\leqslant\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2-t(2-t)\alpha_{\min}C_{\min}\\&\quad\|\omega^k-\tilde{\omega}^k\|^2\end{aligned}$$

定理 2 EAL-ADMM 算法生成的序列 $\{\omega^k\}$ 收敛于变分不等式式(20)解集中的一点 (x^*,z^*,λ^*)。

证明：设 ω^* 为式(20)的解。由式(36)容易得到

$$\begin{aligned}&\|\omega^{k+1}-\omega^*\|_{\boldsymbol{M}}^2\leqslant\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2\leqslant\|\omega^{k-1}-\omega^*\|_{\boldsymbol{M}}^2\\&\leqslant\cdots\leqslant\|\omega-\omega^*\|_{\boldsymbol{M}}^2\leqslant\|\omega^0-\omega^*\|_{\boldsymbol{M}}^2\end{aligned}\qquad(37)$$

也就是说，序列 $\{\|\omega^k-\omega^*\|_{\boldsymbol{M}}\}$ 是递减序列。特别地，序列 $\{\omega^k\}$ 是有界的并且 $\lim_{k\to\infty}\|\omega^k-\omega^*\|_M$ 存在。由式(36)得到

$$t(2-t)\alpha_{\min}C_{\min}\ \|\omega^k-\tilde{\omega}^k\|^2\leqslant\|\omega^k-\omega^*\|_{\boldsymbol{M}}^2-\|\omega^{k+1}-\omega^*\|_{\boldsymbol{M}}^2$$

进而有

$$\lim_{k\to\infty}\|\omega^k-\tilde{\omega}^k\|_{\boldsymbol{M}}=0$$

结果表明 $\lim_{k\to\infty}\|x^k-\tilde{x}^k\|=0$，序列 $\{\omega^k\}$ 和 $\{\tilde{\omega}^k\}$ 具有相同聚点。假设 $\{\omega^{k_j}\}$ 是 $\{\omega^k\}$ 的任意子序列，现取式(28)中子序列 $\{k_j\}$，当 $j\to\infty$，利用 $\tilde{F}$ 的连续性，得到 $\langle\omega-\omega^\infty,\tilde{F}(\omega^\infty)\rangle\geqslant0$，$\omega\in\Omega$，即 ω^∞ 为变分不等式(20a)的解，因此 $\{\tilde{\omega}^k\}$ 的子序列 $\{\tilde{\omega}^{k_j}\}$ 也收敛于 ω^∞。

用 ω^∞ 代替式(27)中的 ω^*，利用 $\lim_{k\to\infty}\|\omega^k-\omega^\infty\|_M$ 收敛的事实，得出结论 $\lim_{k\to\infty}\|\omega^k-\omega^\infty\|_M=\lim_{j\to\infty}\|\omega^{k_j}-\omega^\infty\|_M=0$。这表明全序列 $\{\omega^k\}$ 收敛到 ω^∞，是变分不等式式(20)的一个解。

5 数值实验

将本文提出的算法 EAL-ADMM 与经典的 CQ 算法进行比较来评估其性能。所有代码都由 Matlab 2014B 编写，电脑采用 Intel 四核 i7 2.4 GHz CPU，8GB RAM。

例 1 本实验中考虑随机合成数据的一般 SFP(1)，问题构造如下：设 $\boldsymbol{A}=(a_{ij})_{M\times N}$ 是一个随机矩阵，$a_{ij}\in(0,1)$，M 和 N 是正整数。$C=\{x\in\Re^N\mid\sum_{l=1}^{N}x_l^2\leqslant r^2\}$，$Q=\{y\in\Re^M\mid y\leqslant b\}$。为保证问题解的存在性，采用以下方法生成向量 $\boldsymbol{b}$：给定一个随机的 N 维负向量(每个元素都为负)$z\in C$，$r=\|z\|$，取 $\boldsymbol{b}=\boldsymbol{A}z$。$x\in C$ 满足 $\boldsymbol{A}x\in Q$。

我们取 $x^0=e_{01}=(0,0,\cdots,0)\in\Re^N$，以 $y^0=e_{02}=(0,0,\cdots,0)\in\Re^M$ 和 $\lambda^0=(1,1,\cdots,1)\in\Re^M$ 作为该例的初始点。整个过程中设置 $\tau=4$，$\beta=2$，$t=1.5$。设 $Err:=\max\{\|x^k-P_C(x^k)\|,\|\boldsymbol{A}x^k-P_Q(\boldsymbol{A}x^k)\|\}\leqslant10^{-6}$，作为截止规则。

在实验中，我们测试了 3 个情景，维度分别为 $(N,M)=(10,20)$，$(20,40)$，$(50,50)$，$(100,100)$，$(500,500)$。数值实验结果如表 1 所列，其中“Iter.”和“s”分别表示迭代次数和 cpu 时间(以秒为单位)。

表 1　例 1 的数值实验结果

Table 1　Numerical results of example 1

(N,M)	CQ 算法		EAL-ADMM 算法	
(10,20)	$Iter.=13673$	$s=0.6915$	$Iter.=5019$	$s=0.5747$
(20,40)	$Iter.=23694$	$s=0.9182$	$Iter.=2752$	$s=0.3956$
(50,50)	$Iter.=49389$	$s=1.9104$	$Iter.=16794$	$s=1.1770$
(100,100)	$Iter.=96350$	$s=5.1326$	$Iter.=81023$	$s=2.5670$
(500,500)	$Iter.=157895$	$s=21.3421$	$Iter.=12869$	$s=17.3952$

例 2　设 $C=\{x\in R^3 \mid x_1^2+x_2^2\leqslant 40\}$ 且 $Q=\{x\in R^3 \mid x_2-x_3^2\leqslant 1\}$，$A=I$。求 x，其中 $x\in C$ 且 $\boldsymbol{A}x\in Q$。

该实验中，分别取 $x^0=y^0=(1,3,0)^{\mathrm{T}}$，$\lambda^0=(1,1,1)$ 和 $x^0=y^0=(0,-3,1)^{\mathrm{T}}$，$\lambda^0=(1,1,1)^{\mathrm{T}}$ 作为该例的初始点。整个过程中设置的 τ,β,t 和截止规则同上例，得到的数值结果如表 2 所列。

表 2　例 2 的数值实验结果

Table 2　Numerical results of example 2

初始点	CQ 算法		EAL-ADMM 算法	
$x^0=y^0=(1,3,0)^{\mathrm{T}}$，$\lambda^0=(1,1,1)^{\mathrm{T}}$	$Iter.=89$	$s=0.1015$	$Iter.=43$	$s=0.0841$
$x^0=y^0=(0,-3,1)^{\mathrm{T}}$，$\lambda^0=(1,1,1)^{\mathrm{T}}$	$Iter.=42$	$s=0.0982$	$Iter.=22$	$s=0.0518$

例 3　设 C 为 R^{100} 上的子集，Q 为 R^{300} 上的子集且是闭凸的，设子集 C 和 Q 的定义为：

$$C:=\{x\in R^{100}:\langle \boldsymbol{b},x\rangle\leqslant c\}$$

$$Q:=\{y\in R^{300}:\langle \boldsymbol{q},y\rangle\leqslant p\}$$

其中，向量 $\boldsymbol{b}$ 和 $\boldsymbol{q}$ 的坐标和实数 c,p 分别在闭区间[2,6]和[0,1]中随机生成。

设 $\boldsymbol{A}:R^{100}\rightarrow R^{300}$，是有界的线性算子，并且矩阵中的元素在闭区间[−5,5]中随机生成，容易得到 $S=C\cap Q\neq\varnothing$，因为它包含了 R^{100} 的零点。考虑如下两种情况：

情形 1　$x^0\in R^{100}$ 和 $y^0\in R^{300}$ 作为初始点在闭区间[50,100]上随机产生，$\lambda^0=(1,1,\cdots,1)^{\mathrm{T}}$。

情形 2　$x^0\in R^{100}$ 和 $y^0\in R^{300}$ 作为初始点，要求他们的各分量元素都为 10，$\lambda^0=(1,1,\cdots,1)^{\mathrm{T}}$。

整个过程中设置 τ,β,t 和终止准则同上例，得到的数值结果如表 3 所列。

表 3　例 3 的数值结果

Table 3　Numerical results of example 3

算法	$\varepsilon=10^{-6}$	
	$Iter.$	Time/s
情形 1 EAL-ADMM	60	3.34809×10^{-2}
情形 2 EAL-ADMM	56	3.27447×10^{-2}
情形 1 CQ 算法	324	3.89850×10^{-2}
情形 2 CQ 算法	284	3.55002×10^{-2}

从表 1—表 3 可以看出，算法 EAL-ADMM 的迭代次数和计算时间都比 CQ 算法要少得多。数值实验结果表明，该算法对求解 SFP(1)是有效的且性能良好。数值实验结果表明了加速战略的有效性。

结束语　解决 SFP 问题的 ADMM 算法是 CQ 算法的扩展延伸，本文在此基础上研究了将 SFP 的一般形式构造为目标函数的可分形式(无交叉变量的单个函数之和形式)。数值实验结果表明，本文所提出的 EAL-ADMM 算法是一种简单而有效的迭代框架，比 CQ 算法具有更快的收敛速度。该算法在分析和设计其他加速技术及应用方面，具有潜在的理论价值和应用价值。

参 考 文 献

[1] CENSOR Y, ELFVING T. A multiprojection algorithm using bregman projections in a product space [J]. Numerical Algorithms, 1994, 8: 221-239.

[2] BYRNE C. A unified treatment of some iterative algorithm algorithms in signal processing and image reconstruction [J]. Inverse Problems, 2004, 20: 103-120.

[3] DANG Y Z, GAO Y. The strong convergence of a KM-CQ-like algorithm for a split feasibility problem [J]. Inverse Problems, 2011, 27(1): 015007.

[4] QU B, XIU N. A note on the CQ algorithm for the split feasibility problem [J]. Inverse Problem, 2005, 21: 1655-1665.

[5] XU H. A variable Krasnosel'skii-Mann algorithm and the multiple-set split feasibility problem [J]. Inverse Problems, 2006, 22: 2021-2034.

[6] ZHAO J, YANG Q. Several solution methods for the split feasibility problem [J]. Inverse Problems, 2005, 21: 1791-1799.

[7] ROCKAFELLAR R. Augmented lagrangians and applications of the proximal point algorithm in convex programming [J]. Mathmetics of Operations Research, 1976, 1(2): 97-116.

[8] XUE Z H, YIN Q W, DANG Y Z. Inertia approximate relaxation alternating direction multiplier method for separable convex optimization problems[J]. Journal of University of Shanghai for Science and Technology, 2022, 44(2): 204-212.

[9] YANG J, ZHANG Y. Alternating direction algorithms for l1-problems in compressive sensing [J]. SIAM Journal on Scientific Computing, 2011, 33(1): 250-278.

[10] YANG J, ZHANG Y, YIN W. A fast alternating direction me-thod for TV l1Cl2 signal reconstruction from partial fourier data [J]. IEEE Journal of Selected Topics in Signal Processing, 2010, 4(2): 288-297.

[11] HE H J, LING C, XU H. An implementable splitting algorithm for the l1norm regularized split feasibility problem [J]. Journal of Scientific Computing, 2016, 67(1): 281-298.

LIU Yang, born in 1983, Ph.D, lecturer. Her main research interests include algorithm optimization and artificial intelligence.

XUE Zhonghui, born in 1974, Ph.D, associate professor. His main research interests include operational research and optimization theory, image reconstruction and restoration, artificial intelligence and brain science.

(责任编辑：何杨)