

三个可分离算子凸优化的线性化方法

高雷阜, 潘京乐, 魏 帅
(辽宁工程技术大学理学院, 辽宁 阜新 123000)

摘要: 本文研究了三个可分离算子不含交叉变量的线性约束凸优化问题. 利用定制的邻近点算法, 对其变分不等式子问题进行线性化处理, 并增加一邻近点项, 使其子问题成为易于运算的单调线性变分不等式, 得到了线性化定制的邻近点算法, 并证明了全局收敛性, 推广了文献中的研究结果.

关键词: 可分离算子线性约束问题; 交替方向法; 变分不等式; 全局收敛性

MR(2010) 主题分类号: 90C25

中图分类号: O224

文献标识码: A

文章编号: 0255-7797(2016)02-0365-10

1 引言

对于目标函数不含交叉变量的三个可分离算子线性约束的凸优化最小值问题, 数学表达式

$$\min \left\{ \sum_{i=1}^3 \theta_i(x_i) \mid \sum_{i=1}^3 A_i x_i = b, x_i \in X_i, i = 1, 2, 3 \right\}, \quad (1.1)$$

其中 $A_i \in R^{m \times n_i}$ ($i = 1, 2, 3$), $b \in R^m$, $X_i \subseteq R^{n_i}$ ($i = 1, 2, 3$) 为非空闭凸集, $n_1 + n_2 + n_3 = n$ 及 $\theta_i : X_i \rightarrow R^{n_i}$ ($i = 1, 2, 3$) 是凸函数 (不必光滑).

因映射 θ_i ($i = 1, 2, 3$) 只与变量 x_i 有关, 称这类变分不等式为可分离的变分不等式. 虽然模型 (1.1) 结构特殊, 但它在很多领域都有很重要的作用, 比如在文献 [1] 提到的 fused Lasso 问题, 文献 [2] 涉及的低秩稀疏矩阵的复原问题.

为便于讨论, 假设

- (i) 模型 (1.1) 的解集 Ω^* 非空;
- (ii) 矩阵 A_i ($i = 1, 2, 3$) 均为列满秩的.

对于问题 (1.1), 文献 [4-5] 提到的增广拉格朗日乘子法 (augmented Lagrangian method, ALM) 是可行的, 其拉格朗日函数可表示为

$$L_\beta(x_1, x_2, x_3, \lambda) := \sum_{i=1}^3 \theta_i(x_i) - \lambda^T \left(\sum_{i=1}^3 A_i x_i - b \right) + \frac{\beta}{2} \left\| \sum_{i=1}^3 A_i x_i - b \right\|^2, \quad (1.2)$$

其中 $\lambda \in R^m$ 是增广的拉格朗日乘子, $\beta > 0$ 是事先设定的罚参数. 相应的模型 (1.1) 拉格朗日乘子收缩算法框架.

对给定的 λ^k , 先求

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) \leftarrow \text{Arg min} \{ L_\beta(x_1, x_2, x_3, \lambda^k) \mid x_i \in X_i, i = 1, 2, 3 \}, \quad (1.3)$$

*收稿日期: 2013-06-04

接收日期: 2014-02-10

基金项目: 教育部高校博士学科科研基金资助 (20132121110009).

作者简介: 高雷阜 (1963-), 男, 辽宁阜新, 教授, 主要研究方向: 最优化理论与应用.

然后令

$$\lambda^{k+1} = \lambda^k - \beta \left(\sum_{i=1}^3 A_i x_i - b \right), \quad (1.4)$$

ALM 算法每次迭代过程要求所有的自变量同时最小化, 将其推广用于解决凸优化最小值问题 (1.1) 的交替方向法 (alternating direction method of multipliers, ADM) 的算迭代框架为

$$\begin{cases} x_1^{k+1} \leftarrow \text{Arg min} \{ L_\beta(x_1, x_2^k, x_3^k, \lambda^k) \mid x_1 \in X_1 \}, \\ x_2^{k+1} \leftarrow \text{Arg min} \{ L_\beta(x_1^{k+1}, x_2, x_3^k, \lambda^k) \mid x_2 \in X_2 \}, \\ x_3^{k+1} \leftarrow \text{Arg min} \{ L_\beta(x_1^{k+1}, x_2^{k+1}, x_3, \lambda^k) \mid x_3 \in X_3 \}, \\ \lambda^{k+1} = \lambda^k - \beta \left(\sum_{i=1}^3 A_i x_i^{k+1} - b \right). \end{cases} \quad (1.5)$$

自从它在文献 [6–7] 中被提出来以后, ADM 算法受到了很大关注, 见文献 [8–12], 并发现它在科学计算的很多领域都有重要的作用, 见文献 [13–16]. 但是, 注意到尽管从经验上其数值实验表现很不错 [2,20], 但对于三个可分离算子的 ADM 算法, 其收敛性依然未从理论上得到证明, 这激励一些建立在 ADM 基础上的算法的诞生, 见文献 [17–19]. 类似于模型 (1.1) 的分离结构, 文献 [17–19, 21] 讨论的算法模型甚至是多个 (大于三个) 可分离算子凸优化的收缩算法.

值得注意的是文献 [3, 24–26] 中的大量数值结果表明, 当用交替方向法求解各类问题时, 罚参数 β 对算法的数值表现影响较大: 较小的 β 值虽然能保证算法的收敛性, 但是会导致收敛速度变慢; 而一个较大的 β 值可能会导致算法不收敛. 因此, 另一改进策略是克服固定罚参数 β 带来的弊端, 即允许 β 在迭代过程中单调变化, 或自适应的调节.

文献 [21] 提出的定制的邻近点算法中对多个 (大于三个) 可分离算子凸优化问题的分析及算法的收敛性证明, 使得多个可分离算子线性约束凸优化问题的研究更深一层. 其中作为特例, 令 $\alpha \in (0, 1)$ 及 $\beta > 0$, 借助于初始迭代向量 $w^0 = (x_1^0, x_2^0, x_3^0, \lambda^0)$, 三个可分离算子定制的邻近点交替方向算法的迭代框架为

(1) 预测阶段, 由给定的 w^0 , 先求

$$\tilde{x}_i^k = \text{Arg min} \left\{ \theta_i(x_i) + \frac{\beta}{2 \times 3} \left\| A_i(3x_i - 2x_i^k) + \sum_{j=1, j \neq i}^3 A_j x_j^k - b - \frac{1}{\beta} \lambda^k \right\|^2 \mid x_i \in X_i \right\}, \quad (1.6)$$

$i = 1, 2, 3,$

$$\tilde{\lambda}^k = \lambda^k - \beta \left(\sum_{j=1}^3 A_j \tilde{x}_j^k - b \right). \quad (1.7)$$

(2) 松弛迭代校正阶段, 由上步得到的 $(\tilde{x}_1^k, \tilde{x}_2^k, \tilde{x}_3^k, \tilde{\lambda}^k)$, 求

$$\begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \lambda^k \end{pmatrix} - \alpha \begin{pmatrix} x_1^k - \tilde{x}_1^k \\ x_2^k - \tilde{x}_2^k \\ x_3^k - \tilde{x}_3^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix}. \quad (1.8)$$

然而, 在实际问题中, 类似于结构型 (二个可分离算子) 优化问题求解过程所遇到的, 精确求解子问题 (1.6) 往往花费很大. 因此, 本文通过子问题的线性化及增加一邻近点项, 将其转化为强单调的线性变分不等式问题, 相对而言, 求解容易很多.

内容安排如下: 第 2 节介绍了分析用到的预备知识; 在第 3 节给出改进的新算法的迭代框架; 第 4 节证明了这种新算法的全局收敛性; 在第 5 节, 作了一些简要总结.

2 预备知识

2.1 基本标记及概念

令 $\|\cdot\|_p$ 表示标准的 p 范数, 令 $\|\cdot\| := \|\cdot\|_2$ 表示欧几里得范数. 对于一个正定对称矩阵 G , 定义 $\|\cdot\|_G$ 为 G -范数, 比如 $\|x\|_G := \sqrt{x^T G x}$.

定义 2.1 令 $f: R^n \rightarrow (-\infty, +\infty)$. 若 $\forall x \in R^n, y \in R^n$, 有

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y),$$

$\forall t \in [0, 1]$ 都成立, 则称 f 是凸函数.

定义 2.2 对于 $VI(\Omega, F)$ 中从 R^n 到 R^n 的映射 F , 即 F 满足

$$(F(u) - F(v))^T(u - v) \geq 0, \forall u, v \in R^n,$$

则称 F 在 R^n 上是单调的, 且变分不等式是单调的; 若存在正常数 η 使得

$$(u - v)^T [F(u) - F(v)] \geq \eta \|u - v\|^2, \forall u, v \in \Omega$$

成立, 则称 F 在 Ω 上是强单调的.

定义 2.3 由算法产生的预测-校正迭代序列 $\{w^k\}$ 满足收缩性质

$$\|w^{k+1} - w^*\|^2 \leq \|w^k - w^*\|^2 - \alpha(2 - \alpha)\|w^k - \tilde{w}^k\|^2,$$

称这样的序列 $\{w^k\}$ 为 F 单调的, 其中 α 为校正过程中的松弛因子.

2.2 变分不等式的等价表述

求解模型 (1.1) 的问题等价于寻找点集 $\Omega^* = (x_1^*, x_2^*, x_3^*, \lambda^*) \in \Omega$, 满足

$$\begin{cases} \theta_1(x_1) - \theta_1(x_1^*) + (x_1 - x_1^*)^T (-A_1^T \lambda^*) \geq 0, \\ \theta_2(x_2) - \theta_2(x_2^*) + (x_2 - x_2^*)^T (-A_2^T \lambda^*) \geq 0, \\ \theta_3(x_3) - \theta_3(x_3^*) + (x_3 - x_3^*)^T (-A_3^T \lambda^*) \geq 0, \quad \forall (x_1, x_2, x_3, \lambda) \in \Omega, \\ (\lambda - \lambda^*)^T \left(\sum_{i=1}^3 A_i x_i^* - b \right) \geq 0, \end{cases} \quad (2.1)$$

则一阶优化条件 (2.1) 可写成紧凑形式

$$w^* \in \Omega, \theta(u) - \theta(u^*) + (w - w^*)^T F(w^*) \geq 0, \forall w \in \Omega. \quad (2.2)$$

注意到 F 是单调的, 且 Ω^* 表示变分不等式 (2.2) 的解集, 其中 $\Omega = X_1 \times X_2 \times X_3 \times R^m$. 记

$$u = (x_1, x_2, x_3)^T, F(w) = \left(-A_1^T \lambda, -A_2^T \lambda, -A_3^T \lambda, \sum_{i=1}^3 A_i x_i - b \right)^T,$$

$$w = (x_1, x_2, x_3, \lambda)^T, \theta(u) = \theta_1(x_1) + \theta_2(x_2) + \theta_3(x_3).$$

通篇假设:

1. 问题 (1.2) 的解集 Ω^* 是非空的;
2. 假设对给定的凸函数 $\theta_i(x_i)$ ($i = 1, 2, 3$) 及 $\forall \gamma > 0, a \in R^n$, 子问题

$$\min \left\{ \theta_i(x_i) + \frac{\gamma}{2} \|A_i x_i - a\|^2 \mid x_i \in X_i \right\}, i = 1, 2, 3$$

的求解是简单的.

3 算法描述

3.1 新方法迭代框架

这一节主要给出了基于三个可分离算子定制邻近点算法 (proximal point algorithm, PPA) 的线性化算法:

线性化 (1.6) 在 x^k 处的二次方项

$$\left\{ \frac{\beta}{2 \times 3} \left\| A_i (3x_i - 2x_i^k) + \sum_{j=1, j \neq i}^3 A_j x_j^k - b - \frac{1}{\beta} \lambda^k \right\|^2 \right\},$$

并在目标函数增加一邻近点项 $\frac{r}{2} \|x_i - x_i^k\|^2$ ($i = 1, 2, 3$), 其中 r 为参数. 线性化后, 由给定的 $(x_1^0, x_2^0, x_3^0, \lambda^0)$ 得到

(1) 预测序列

$$\tilde{x}_i^k = \text{Arg min} \left\{ \theta_i(x_i) + \beta x_i^T A_i^T \left(\sum_{j=1}^3 A_j x_j^k - b - \frac{1}{\beta} \lambda^k \right) + \frac{r}{2} \|x_i - x_i^k\|^2 \mid x_i \in X_i \right\},$$

$$i = 1, 2, 3, \quad (3.1)$$

$$\tilde{\lambda}^k = \lambda^k - \beta \left(\sum_{j=1}^3 A_j \tilde{x}_j^k - b \right). \quad (3.2)$$

(2) 校正序列

$$\begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \lambda^k \end{pmatrix} - \alpha \begin{pmatrix} x_1^k - \tilde{x}_1^k \\ x_2^k - \tilde{x}_2^k \\ x_3^k - \tilde{x}_3^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix}. \quad (3.3)$$

(3) 参数要求 $\alpha \in (0, 2)$, 对给定的 $\beta > 0$, 参数 r 满足 $rI - \beta A^T A \geq 0$, 其中 $A = (A_1, A_2, A_3)$.

注记 此处校正步长 α 是固定不变的, α 的选取对算法收敛及收敛速度都有重要影响, 故允许 α 在迭代过程中自适应地调节是可取的.

3.2 变分不等式的等价表述

公式 (3.1), (3.2) 等价于求解变分不等式

$$\theta(u) - \theta(\tilde{u}^k) + (w - \tilde{w}^k)^T \{F(\tilde{w}^k) + Q(\tilde{w}^k - w^k)\} \geq 0, \quad (3.4)$$

则 $u = (x_1, x_2, x_3)^T$, $w = (x_1, x_2, x_3, \lambda)^T$, $F(w) = \left(-A_1^T \tilde{\lambda}^k, -A_2^T \tilde{\lambda}^k, -A_3^T \tilde{\lambda}^k, \sum_{j=1}^3 A_j \tilde{x}_j^k - b\right)^T$,

$$Q = \begin{pmatrix} rI_{n_1} - \beta A_1^T A_1 & -\beta A_1^T A_2 & -\beta A_1^T A_3 & 0 \\ -\beta A_2^T A_1 & rI_{n_2} - \beta A_2^T A_2 & -\beta A_2^T A_3 & 0 \\ -\beta A_3^T A_1 & -\beta A_3^T A_2 & rI_{n_3} - \beta A_3^T A_3 & 0 \\ 0 & 0 & 0 & \frac{1}{\beta} I_m \end{pmatrix}$$

为半正定矩阵, 且显然 F 为单调的.

分析: 公式 (3.1) 等价于求解变分不等式

$$\theta_i(x_i) - \theta_i(\tilde{x}_i^k) + (x_i - \tilde{x}_i^k) \left\{ \beta A_i^T \left(\sum_{j=1}^3 A_j x_j^k - b - \frac{1}{\beta} \lambda^k \right) + r(x_i - x_i^k) \right\} \geq 0, i = 1, 2, 3,$$

借用公式 (3.2), 上式可化为

$$\theta_i(x_i) - \theta_i(\tilde{x}_i^k) + (x_i - \tilde{x}_i^k) \left\{ \beta A_i^T \left(-\sum_{j=1}^3 A_j (\tilde{x}_j^k - x_j^k) - \frac{1}{\beta} \tilde{\lambda}^k \right) + r(x_i - x_i^k) \right\} \geq 0, i = 1, 2, 3,$$

又 $\forall x_i \in X_i$, 有 $\tilde{x}_i^k \in X_i$, 则上式化为

$$\theta_i(x_i) - \theta_i(\tilde{x}_i^k) + (x_i - \tilde{x}_i^k) \left\{ -A_i^T \tilde{\lambda}^k - \beta A_i^T \sum_{j=1}^3 A_j (\tilde{x}_j^k - x_j^k) + r(\tilde{x}_i^k - x_i^k) \right\} \geq 0, i = 1, 2, 3.$$

上述 $i = 1, 2, 3$ 的情况取加和, 并结合 (3.2) 式的等价变形形式

$$\frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) + \left(\sum_{j=1}^3 A_j \tilde{x}_j^k - b \right) = 0,$$

有 $\theta(u) - \theta(\tilde{u}^k) + (w - \tilde{w}^k)^T \{F(\tilde{w}^k) + Q(\tilde{w}^k - w^k)\} \geq 0$. 其中

$$w = (x_1, x_2, x_3, \lambda)^T,$$

$$F(w) = \left(-A_1^T \tilde{\lambda}^k, -A_2^T \tilde{\lambda}^k, -A_3^T \tilde{\lambda}^k, \sum_{j=1}^3 A_j \tilde{x}_j^k - b\right)^T,$$

$$Q = \begin{pmatrix} rI_{n_1} - \beta A_1^T A_1 & -\beta A_1^T A_2 & -\beta A_1^T A_3 & 0 \\ -\beta A_2^T A_1 & rI_{n_2} - \beta A_2^T A_2 & -\beta A_2^T A_3 & 0 \\ -\beta A_3^T A_1 & -\beta A_3^T A_2 & rI_{n_3} - \beta A_3^T A_3 & 0 \\ 0 & 0 & 0 & \frac{1}{\beta} I_m \end{pmatrix}.$$

下证对称矩阵 Q 为半正定矩阵.

因为参数 $\beta > 0$, 故只需证对称矩阵

$$\tilde{Q} = \begin{pmatrix} rI_{n_1} - \beta A_1^T A_1 & -\beta A_1^T A_2 & -\beta A_1^T A_3 \\ -\beta A_2^T A_1 & rI_{n_2} - \beta A_2^T A_2 & -\beta A_2^T A_3 \\ -\beta A_3^T A_1 & -\beta A_3^T A_2 & rI_{n_3} - \beta A_3^T A_3 \end{pmatrix}$$

半正定即可.

$$\text{又 } \tilde{Q} = rI_n - \beta \begin{pmatrix} A_1^T \\ A_2^T \\ A_3^T \end{pmatrix} (A_1, A_2, A_3), \text{ 其中 } n = n_1 + n_2 + n_3. \text{ 记矩阵 } A = (A_1, A_2, A_3),$$

则矩阵 $\tilde{Q} = rI_n - \beta A^T A$, 由参数的取值范围知矩阵 Q 为半正定阵显然成立.

综上可知本结论得证.

注 1 一般保守估计 $r \geq \beta \lambda_{\max}(A^T A)$.

注 2 在后面的部分, 该性质对算法的收敛性证明很有帮助.

4 收敛性证明

引理 4.1 变分不等式 (2.2) 的解集是凸集, 且可表示为

$$\Omega^* = \bigcap_{w \in \Omega} \left\{ \tilde{w} \in \Omega : \theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T F(w) \geq 0 \right\}. \quad (4.1)$$

证 (1) $(\Rightarrow)$ 若 $\tilde{w} \in \Omega^*$, 有

$$\theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T \geq 0, \forall w \in \Omega,$$

借助 F 在 Ω 上的单调性, 可以得出 $\theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T F(w) \geq 0, \forall w \in \Omega$, 因此 $\tilde{w}$ 属于 (4.1) 式右边的集合.

(2) $(\Leftarrow)$ 设对 $\forall w \in \Omega$, 对所有的 $\alpha \in (0, 1)$, 矢量 $\bar{w} = \alpha \tilde{w} + (1 - \alpha)w \in \Omega$, 因此有

$$\theta(\bar{u}) - \theta(\tilde{u}) + (\bar{w} - \tilde{w})^T F(\bar{w}) \geq 0. \quad (4.2)$$

因为 $\theta(\bullet)$ 是凸的, 故 $\theta(\bar{u}) \leq \alpha \theta(\tilde{u}) + (1 - \alpha)\theta(u)$ 将上式代入 (4.2) 式, 则对所有的 $\alpha \in (0, 1)$, 有

$$(1 - \alpha)[\theta(u) - \theta(\tilde{u})] + (1 - \alpha)(w - \tilde{w})^T F(\bar{w}) \geq 0,$$

则易得

$$\theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T F(\alpha \tilde{w} + (1 - \alpha)w) \geq 0.$$

令 $\alpha \rightarrow 1$ 可得 $\theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T F(\tilde{w}) \geq 0$.

因此 $\tilde{w} \in \Omega^*$. 综合 (1), (2) 知, 变分不等式的解集可表示为

$$\Omega^* = \bigcap_{w \in \Omega} \left\{ \tilde{w} \in \Omega : \theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T F(w) \geq 0 \right\}.$$

(3) 下证 Ω^* 是凸集.

$\forall w \in \Omega$, 集合 $\{\tilde{w} \in \Omega : \theta(\tilde{u}) + \tilde{w}^T F(w) \leq \theta(u) + w^T F(w)\}$ 等价于

$$\left\{ \tilde{w} \in \Omega : \theta(u) - \theta(\tilde{u}) + (w - \tilde{w})^T F(w) \geq 0 \right\},$$

因为任意个凸集之交仍然为凸集, 故集合 Ω^* 是凸集得证.

综合 (1)–(3) 知本引理得证.

引理 4.2 由 (3.1) 式产生的序列 $\{w^k\}$ 满足 $\forall w^* \in W^*$,

$$(\tilde{w}^k - w^*)^T Q(w^k - \tilde{w}^k) \geq 0. \quad (4.3)$$

证 记 $w^* \in \Omega^* \subseteq \Omega$, 其中 Ω^* , Ω 分别为不等式的解集及定义域, (3.2) 式中令 $w = w^*$, 有

$$(\tilde{w}^k - w^*)^T Q(w^k - \tilde{w}^k) \geq (\tilde{w}^k - w^*)^T F(\tilde{w}^k) + \theta(\tilde{u}^k) - \theta(u^*) \geq 0, \forall w^* \in \Omega^*.$$

借助 F 的单调性 $(\tilde{w}^k - w^*)^T (F(\tilde{w}^k) - F(w^*)) \geq 0$, 及 $\tilde{w}^k \in \Omega$, $w^* \in \Omega^*$, 有

$$(\tilde{w}^k - w^*)^T F(\tilde{w}^k) + \theta(\tilde{u}^k) - \theta(u^*) \geq (\tilde{w}^k - w^*)^T F(w^*) + \theta(\tilde{u}^k) - \theta(u^*) \geq 0.$$

因此有

$$(\tilde{w}^k - w^*)^T Q(w^k - w^*) \geq 0, \forall w^* \in \Omega^*.$$

因此可得到 (4.3) 式.

引理 4.3 由 (3.1) 式产生的序列 $\{w^k\}$ 满足 $\forall w^* \in W^*$,

$$(w^k - w^*)^T Q(w^k - \tilde{w}^k) \geq \|w^k - \tilde{w}^k\|_Q^2. \quad (4.4)$$

证 由引理 4.2 $(\tilde{w}^k - w^*)^T Q(w^k - \tilde{w}^k) \geq 0$ 知

$$(\tilde{w}^k - w^k + w^k - w^*)^T Q(w^k - \tilde{w}^k) \geq 0,$$

则 $(w^k - w^*)^T Q(w^k - \tilde{w}^k) \geq \|w^k - \tilde{w}^k\|_Q^2$, 即本结论得证.

借助引理 4.3, 可得到一个重要的不等式.

引理 4.4 由 (3.1) 式产生的序列 $\{w^k\}$ 满足

$$\|w^{k+1} - w^*\|_Q^2 \leq \|w^k - w^*\|_Q^2 - \alpha(2 - \alpha) \|w^k - \tilde{w}^k\|_Q^2. \quad (4.5)$$

证 由迭代式 (3.3) 得 $\forall w^* \in W^*$,

$$\begin{aligned} & \|w^k - w^*\|_Q^2 - \|w^{k+1} - w^*\|_Q^2 = \|w^k - w^*\|_Q^2 - \|(w^k - w^*) - \alpha(w^k - \tilde{w}^k)\|_Q^2 \\ & = 2\alpha(w^k - w^*)^T Q(w^k - \tilde{w}^k) - \alpha^2 \|w^k - \tilde{w}^k\|_Q^2. \end{aligned}$$

回顾 (4.4) 式, 则由上述不等式可得

$$\|w^k - w^*\|_Q^2 - \|w^{k+1} - w^*\|_Q^2 \geq \alpha(2 - \alpha) \|w^k - \tilde{w}^k\|_Q^2, \forall w^* \in W^*, \quad (4.6)$$

即引理论断得证.

注 3 由公式 (4.6), 容易知道在迭代步 (3.3) 要求参数 $\alpha \in (0, 1)$, 是为了保证迭代列 $\{w^k\}$ 收敛到 W^* .

现在, 借助引理 4.2, 提出迭代步 (3.3) 对于算法收敛性是有效的. 换句话说, 在 Q -范数意义下, 该步生成的新迭代 w^{k+1} 比 w^k 更靠近集合 W^* , 使得 (3.1) 式产生的序列 $\{w^k\}$ 的收敛性真实可靠.

定理 4.5 令 $\{w^k\}$ 及 $\{\tilde{w}^k\}$ 是定制的邻近点交替方向线性化算法 (3.1) 产生的序列, 则有

- (1) 序列 $\{w^k\}$ 是有界的;
- (2) $\lim_{k \rightarrow \infty} \|w^k - \tilde{w}^k\|^2 = 0$;
- (3) $\{\tilde{w}^k\}$ 的聚点是变分不等式 (2.2) 的解点;
- (4) 聚点唯一.

证 (1) 由引理 4.4 知结论 1 显然成立, 因为有界点列一定有聚点. 故不妨记 w^∞ 为 $\{w^k\}$ 的一个聚点.

(2) 将引理 4.4 中的式 (4.5) 式无穷迭代序列情况 $k = 0, 1, 2, \dots$ 作加和, 可得到

$$\sum_{k=0}^{\infty} \alpha(2-\alpha) \|\mathbf{w}^k - \tilde{\mathbf{w}}^k\|^2 \leq \|\mathbf{w}^0 - \mathbf{w}^*\|^2,$$

而 $\alpha(2-\alpha) > 0$, 故 $\lim_{k \rightarrow \infty} \|\mathbf{w}^k - \tilde{\mathbf{w}}^k\|^2 = 0$, 故 w^∞ 也是 $\{\tilde{w}^k\}$ 的一个聚点. 又因为定义域 Ω 为闭凸集, 则 $w^\infty \in \Omega$.

(3) 由 $\tilde{w}^k \in W$,

$$\lim_{k \rightarrow \infty} \left\{ \theta(u) - \theta(\tilde{u}^k) + (w - \tilde{w}^k)^T F(\tilde{w}^k) \right\} \geq 0, \forall w \in W,$$

显然 $\{\tilde{w}^k\}$ 的聚点是变分不等式 (2.2) 的解点, 即为模型 (1.1) 的解点. 则 w^∞ 为变分不等式的解点.

(4) 依引理 4.4 知 $\{w^k\}$ 是 F 单调的且 $\lim_{k \rightarrow \infty} \|w^k - \tilde{w}^k\|^2 = 0$, 故序列 $\{\tilde{w}^k\}$ 的聚点唯一^[23], 且 $\{\tilde{w}^k\}$ 收敛到 $w^\infty \in \Omega^*$.

故本定理得证.

5 总结

对于三个可分离算子无交叉变量的线性约束凸优化问题, 本文在定制的邻近点交替方向法的基础上提出了线性化算法, 并证明了算法的全局收敛性. 注意到在提出的算法中罚参数 β 及迭代中的校正步长 α 均为固定值, 因此如何使用变化的参数以提高算法的效率, 是下一步研究的问题.

参 考 文 献

- [1] Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused lasso[J]. J. Royal Stat. Soc. Ser. Bist. Meth., 2005, 67: 91-108.

- [2] Tao Min, Yuan XiaoMing. Recovering low-rank and sparse components of matrices from incomplete and noisy observations[J]. *Siam J. Contr. Opti.*, 2011, 21: 57–81.
- [3] Fukushima M. Application of the alternating directions method of multipliers to separable convex programming problems[J]. *Comp. Opti. Appl.*, 1992, 2: 93–111.
- [4] Hestenes M. Multiplier and gradient methods[J]. *J. Opti. The. Appl.*, 1969, 4: 303–320.
- [5] Powell M. A method for nonlinear constraints in minimization problems[M]. New York: UKAEA, 1967.
- [6] Gabay D, Mercier B. A dual algorithm for the solution of nonlinear variational problems via finite element approximations[J]. *Comp. Math. Appl.*, 1976, 2: 17–40.
- [7] Glowinski R, Marrocco A. Approximation par éléments finis d'ordre un et résolution par pénalisation-dualité d'une classe de problèmes non linéaires[J]. *RAIRO. Anal. Numér.*, 1975, 9(2): 41–76.
- [8] Chen G, Teboulle M. A proximal-based decomposition method for convex minimization problems[J]. *Math. Prog.*, 1994, 64: 81–101.
- [9] Fukushima M. Application of the alternating directions method of multipliers to separable convex programming problems[J]. *Comp. Opti. Appl.*, 1992, 2: 93–111.
- [10] Glowinski R, Le Tallec P. Augmented lagrangian and operator-splitting methods in nonlinear mechanics[M]. Philadelphia: SIAM Stud. Appl. Math., 1989.
- [11] He Bingsheng, Liao Lizhi, Han Deren, Yang Hai. A new inexact alternating direction method for monotone variational inequalities[J]. *Math. Prog.*, 2002, 92: 103–118.
- [12] Kontogiorgis S, Meyer R. A variable-penalty alternating directions method for convex optimization[J]. *Math. Prog.*, 1998, 83: 29–53.
- [13] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. *Found. Trends Mach. Learn.*, 2010, 3: 1–122.
- [14] Chen Caihua, He Bingsheng, Yuan Xingming. Matrix completion via alternating direction method[J]. *SIAM J. Numer. Anal.*, 2012, 32: 227–245.
- [15] Esser E. Applications of lagrangian-based alternating direction methods and connections to split bregman[J]. *UCLA CAM Report*, 2009, 9–31.
- [16] He Bingsheng, Yuan Xiaoming. Solving large-scale least squares covariance matrix problems by alternating direction methods[J]. *SIAM J. Matrix Anal. Appl.*, 2011, 32: 136–152.
- [17] He Bingsheng, Tao Min, Yuan Xiaoming. Alternating direction method with Gaussian back substitution for separable convex programming[J]. *SIAM J. Opti.*, 2012, 22: 313–340.
- [18] He Bingsheng, Tao Min, Yuan Xingming. An alternating direction-based contraction method for linearly constrained separable convex programming problems[J]. *Opti.*, 2013, 62(4): 573–596.
- [19] He Bingsheng, Tao Min, Yuan Xiaomng. A splitting method for separable convex programming[J]. *IMA J. Numer. Anal.*, 2014: dit060.
- [20] Peng Y, Ganesh A, Wright J. Robust alignment by sparse and low-rank decomposition for linearly correlated images[J]. *Pattern Anal. Machine Intel., IEEE Trans.*, 2012, 34(11): 2233–2246.
- [21] Gu Guoyong, He Bingsheng, Yuan Xiaoming. Customized proximal point algorithms for linearly constrained convex minimization and saddle-point problems: a uniform approach[J]. *Comp. Opti. Appl.*, 2011: 1–27.
- [22] Blum E, Oettli W. Mathematische optimierung, grundlagen und verfahren, Ökonometrie und Unternehmensforschung[M]. Berlin-Heidelberg-New York: Springer-Verlag, 1975.
- [23] Bauschke H H, Combettes P L. A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces[J]. *Math Oper. Res.*, 2001, 26(2): 248–264.

- [24] Han Deren. Inexact operator splitting methods with self-adaptive strategy for variational inequality problems[J]. J. Opti. The. Appl., 2007, 132: 227–243.
- [25] Han Deren, Xu Wei, Yang Hai. An operator splitting method for variational inequalities with partially unknown mappings[J]. Numer. Math., 2008, 111: 207–237.
- [26] He Bingsheng, Liao L Z, Wang S L. Self-adaptive operator splitting methods for monotone variational inequalities[J]. Numer. Math., 2003, 94: 715–737.

A LINEARIZED METHOD FOR THE SEPARABLE CONVEX PROGRAMMING WITH OBJECTIVE FUNCTION REPRESENTED AS THREE FUNCTIONS

GAO Lei-fu, PAN Jing-le, WEI Shuai

(College of Science, Liaoning Technical University, Fuxin 123000, China)

Abstract: In this paper, we study the convex minimization problem with linear constraints and a block-separable objective function which is represented as the sum of three functions without coupled variables. Based on the customized proximal point algorithm, linearizing its variable inequality subproblem and adding an approximate item to the subproblem, the new linearized method is offered and its global convergence is proved finally. The results have extended the corresponding studies documented.

Keywords: linear constraints problems with separable operators; alternating direction method of multipliers; variational inequality; global convergence.

2010 MR Subject Classification: 90C25