

非凸不可分离问题的广义交替方向乘子法的收敛性

薛中会¹, 胡惠晴², 党亚峥²

(1. 上海出版印刷高等专科学校, 上海 200093; 2. 上海理工大学 管理学院, 上海 200093)

摘要: 交替方向乘子法 (ADMM) 是求解大规模优化问题和非凸非光滑问题的一种有效的方法, 但当目标函数为非凸非光滑的情况时, 原始 ADMM 算法的收敛性无法保证, 且若目标函数中存在耦合函数, 则算法的收敛性证明将更为复杂。在现实生活中存在的很多问题, 其本质都是非凸的。因此, 本文提出了一种改进的 ADMM 算法。与原始 ADMM 算法相比, 该算法引入了一个松弛因子 α , 构造了一种广义交替方向乘子法 (GADMM) 来求解具有线性约束的非凸不可分离优化问题。在一定的假设条件下, 通过假设增广拉格朗日函数满足 K-L 不等式, 证明了当惩罚参数足够大时, 算法生成的序列收敛到增广拉格朗日函数的稳定点。

关键词: 广义交替方向乘子法; K-L 不等式; 非凸最优化; 不可分离问题; 收敛性分析
中图分类号: O 221.2 **文献标志码:** A

Convergence of the generalized alternating direction method of multipliers for nonseparable nonconvex problem

XUE Zhonghui¹, HU Huiqing², DANG Yazheng²

(1. Shanghai Publishing and Printing College, Shanghai 200093, China; 2. Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The alternating direction method of multipliers (ADMM) is an effective method for solving large-scale optimization and nonconvex non-smooth objective. However, the convergence cannot be guaranteed when the objective function is nonconvex and non-smooth. Moreover, the proof of the convergence is more complex when the objective contains coupled function. Many problems in real life are nonconvex in nature, so the research on nonconvex optimization problems is particularly important. Therefore, an improved ADMM algorithm is proposed in this paper. Compared with the original ADMM algorithm, a relaxation factor α is introduced and a generalized alternating direction multiplier method (GADMM) is constructed to solve the non-convex nonseparable optimization problem with linear constraints. Under certain assumptions, by assuming that the augmented Lagrangian function satisfies the K-L inequality, it is proved that the sequence generated by the algorithm converges to the critical point of the augmented Lagrangian function when the penalty parameter is sufficiently large.

Keywords: generalized alternating direction multiplier method; K-L inequality; nonconvex optimization; nonseparable problem; convergence analysis

收稿日期: 2021-03-18

基金项目: 国家自然科学基金资助项目 (72071130)

第一作者: 薛中会 (1974-), 男, 副教授。研究方向: 最优化理论和数学物理方程。E-mail: hnlxzh@163.com

1 问题的提出

考虑具有非凸不可分离的优化问题

$$\begin{aligned} \min & g(x, y) + f(x) + h(y) \\ \text{s.t. } & \mathbf{A}x + \mathbf{B}y = 0 \end{aligned} \quad (1)$$

式中: $f: \mathbf{R}^n \rightarrow R \cup \{+\infty\}$ 是恰当的下半连续函数; $h: \mathbf{R}^m \rightarrow R$ 是连续可微函数; $g: \mathbf{R}^n \times \mathbf{R}^m \rightarrow R$ 是光滑函数, 且 x 和 y 是不可分离的。

问题(1)的一个特殊形式是没有函数 g , 且目标函数是可分离的, 即

$$\begin{aligned} \min & f(x) + h(y) \\ \text{s.t. } & \mathbf{A}x + \mathbf{B}y = 0 \end{aligned} \quad (2)$$

利用交替方向乘子法 (ADMM) 求解问题(2)的一种有效算法的迭代格式为

$$\begin{cases} x^{k+1} \in \arg \min_x \left\{ f(x) - \langle \lambda, \mathbf{A}x \rangle + \frac{\beta}{2} \|\mathbf{A}x + \mathbf{B}y^k\|^2 \right\} \\ y^{k+1} \in \arg \min_y \left\{ g(y) - \langle \lambda, \mathbf{B}y \rangle + \frac{\beta}{2} \|\mathbf{A}x^{k+1} + \mathbf{B}y\|^2 \right\} \\ \lambda^{k+1} = \lambda^k - \rho (\mathbf{A}x^{k+1} + \mathbf{B}y^{k+1}) \end{cases} \quad (3)$$

式中: λ 为拉格朗日乘子; β 为惩罚参数, $\beta > 0$ 。

ADMM 算法通过引入一个新的辅助变量, 将原问题改写为一个目标函数可分离且辅助变量与原变量是线性约束的形式, 通过交替更新原变量、辅助变量和对偶变量来迭代求解问题的最优解。通过引入合适的辅助变量, 每个迭代步骤可以变成非常简单的子问题, 通常可以收敛到稳定点或者被并行求解。这使得 ADMM 算法适用于求解大规模的优化问题。

对于 f 和 h 都是凸函数的情况, ADMM(式(3))的收敛性得到了很好的证明, 并且对其进行了收敛率分析^[1-3]。在没有凸性的假设下, 更难以证明 ADMM 的收敛性。在这方面的研究取得了一些进展^[4-8]。Guo 等^[4]利用经典 ADMM 算法求解非凸多块可分离最优化问题, 证明了其收敛性, 并且提出了一些充分条件, 保证了算法的超线性和线性收敛速率。Li 等^[5]提出了一种近似 ADMM 算法来解决非凸非光滑优化问题, 证明了当罚参数足够大且生成的序列有聚点时, 算法所产生迭代点列收敛到稳定点。

目前的研究大多数考虑如下的问题:

$$\begin{aligned} \min & g(x, y) + f(x) + h(y) \\ \text{s.t. } & \mathbf{A}x + y = b \end{aligned} \quad (4)$$

然而, 由于函数 g 的存在, 即使目标函数是凸的情况, 对 ADMM(式(4))的收敛性分析的研究

还处于初期, 研究成果很少。并且由于约束条件中矩阵 $\mathbf{B}$ 的存在, 使得收敛性分析变得更加困难。

Gao 等^[9]考虑了函数 g 是光滑函数以及函数 f, h 是凸函数的情况。通过假设 ∇g 是利普希茨连续以及 h 强凸, 证明了由 ADMM 算法生成的序列收敛到问题(4)的最优解。Chen 等^[10]在耦合函数 g 是二次函数的情况下, 分析了 ADMM 解决问题(4)的收敛性。Liu 等^[11]提出了线性化 ADMM 来解决非凸非光滑的目标问题, 通过将目标中的可微项以及增广项线性化, 证明了算法的收敛性, 然后将问题推广到多块并行的 ADMM 算法中。以上研究都是矩阵 $\mathbf{B}$ 为单位阵的情况。在耦合函数缺失的情况下, Wang 等^[12]研究了用 Bregman 距离修饰 ADMM 算法 (BADMM 算法), 分别分析了矩阵 $\mathbf{B}$ 单射和非单射情况下算法的收敛性, 证明了 BADMM 算法生成的序列收敛到稳定点。Guo 等^[13]提出了一种广义的 ADMM 算法来求解问题(4)。

本文研究广义 ADMM 算法的收敛性, 通过假设矩阵 $\mathbf{B}$ 列满秩以及增广拉格朗日函数满足 K-L 不等式, 当罚参数足够大时, 证明了算法产生的序列收敛到其稳定点, 从而证明了算法的收敛性。

2 预备知识

现给出理论分析所需要的概念和性质。

对于任意 $x \in \mathbf{R}^n$ 是函数 f 的极小值点的必要条件是 $0 \in \partial f(x)$, 满足这个条件的点称为稳定点, 函数 f 的稳定点集记作 $\text{crit } f$ 。

定义 1 令 $f: \mathbf{R}^n \rightarrow R \cup \{+\infty\}$ 为正常的下半连续函数。

a. Fréchet 次微分。

函数 f 在 $x \in \text{dom } f$ 的 Fréchet 次微分, 定义为满足下列关系 $x^* \in \mathbf{R}^n$ 的集合:

$$\liminf_{\substack{y \neq x \\ y \rightarrow x}} \frac{f(y) - f(x) - \langle x^*, y - x \rangle}{\|y - x\|} \geq 0$$

记作 $\hat{\partial} f(x)$ 。当 $x \notin \text{dom } f$ 时, 记 $\hat{\partial} f(x) := \emptyset$ 。

b. 极限次微分。

函数 f 在 $x \in \text{dom } f$ 的极限次微分, 定义为

$$\partial f(x) := \{x^* \in \mathbf{R}^n : \exists x_n \rightarrow x, f(x_n) \rightarrow f(x), x_n^* \in \hat{\partial} f(x_n),$$

$x_n^* \rightarrow x^*\}, \text{ 记作 } \partial f(x)。$

c. 若在函数 f 的定义域中满足 $0 \in \partial f(x^*)$, 则称 x^* 为 f 的稳定解。

d. 对 $\forall x, y \in \text{dom } f$, 若满足 $\|f(x) - f(y)\| \leq L \|x - y\|$,

则称 f 满足 Lipschitz 连续条件, L 为 Lipschitz 常数。

引理 1 $h: \mathbf{R}^m \rightarrow R$ 是连续可微函数, 且 ∇h 是关于常数 L 利普希茨连续的, 那么, 对于任意的 $x, y \in \mathbf{R}^n$, 有

$$|h(y) - h(x) - \langle \nabla h(x), y - x \rangle| \leq \frac{L}{2} \|y - x\|^2$$

引理 2 (K-L 不等式) 设函数 $f: \mathbf{R}^n \rightarrow R \cup \{+\infty\}$ 是恰当下半连续函数, 对于 $-\infty < \eta_1 < \eta_2 < +\infty$, 令

$$[\eta_1 < f < \eta_2] := \{x \in \mathbf{R}^n : \eta_1 < f(x) < \eta_2\}$$

若存在 $\eta \in (0, +\infty]$, x^* 的领域 U 以及一个连续的凹函数 $\varphi: [0, \eta) \rightarrow R_+$, 满足如下条件:

- a. $\varphi(0) = 0$;
- b. φ 在 $(0, \eta)$ 连续可微且在 0 处连续;
- c. $\varphi'(s) > 0, \forall s \in (0, \eta)$;
- d. 对所有的 $x \in U \cap [f(x^*) < f < f(x^*) + \eta]$, 有

$$\begin{cases} x^{k+1} \in \arg \min_x \left\{ g(x, y^k) + f(x) - \langle \gamma^k, Ax \rangle + \frac{\beta}{2} \|Ax + By^k\|^2 \right\} \\ y^{k+1} \in \arg \min_y \left\{ g(x^{k+1}, y) + h(y) - \langle \gamma^k, By \rangle + \frac{\beta}{2} \|\alpha Ax^{k+1} - (1-\alpha)By^k + By\|^2 \right\} \\ \gamma^{k+1} = \gamma^k - \beta(\alpha Ax^{k+1} + (1-\alpha)y^k + By^{k+1}) \end{cases} \quad (5)$$

其中, $\alpha \in (0, 2)$ 是一个松弛因子, 显然, 当 $\alpha = 1$ 时, GADMM 即退化成经典的 ADMM 算法, 并且当 $g \equiv 0$ 时退化成经典的 GADMM 算法。

在证明收敛性之前, 先给出式(5)的变形, 由最优性条件得到

$$\begin{cases} 0 \in \partial f(x^{k+1}) + \nabla_x g(x^{k+1}, y^k) - A^T \gamma^k + \beta A^T (Ax^{k+1} + By^k) \\ 0 = \nabla h(y^{k+1}) + \nabla_y g(x^{k+1}, y^{k+1}) - B^T \gamma^k + B^T \beta (\alpha Ax^{k+1} - (1-\alpha)By^k + By^{k+1}) \\ \gamma^{k+1} = \gamma^k - \beta (\alpha Ax^{k+1} + (1-\alpha)y^k + By^{k+1}) \end{cases} \quad (6)$$

利用式(6)的最后一个等式, 得到

$$\begin{cases} A^T \gamma^k - \beta A^T (Ax^{k+1} + By^k) - \nabla_x g(x^{k+1}, y^k) \in \partial f(x^{k+1}) \\ \nabla h(y^{k+1}) = B^T \gamma^{k+1} - \nabla_y g(x^{k+1}, y^{k+1}) \\ \gamma^{k+1} = \gamma^k - \beta (\alpha Ax^{k+1} - (1-\alpha)y^k + By^{k+1}) \end{cases} \quad (7)$$

$$\tilde{\beta} := \max \left\{ \frac{\alpha \lambda_{B^T B} (L_3 + \omega) + \sqrt{\alpha^2 (L_3 + \omega)^2 + 16 \alpha \mu \lambda_{B^T B} M^2}}{2 \alpha \mu \lambda_{B^T B}}, \frac{\alpha \lambda_{B^T B} (L_1 + L_2) + \sqrt{\alpha^2 \lambda_{B^T B}^2 (L_1 + L_2)^2 + 16 (2-\alpha) \lambda_{B^T B}^2 (L_1^2 + M^2)}}{2 \lambda_{B^T B}^2 (2-\alpha)} \right\}$$

Kurdyka-Lojasiewicz 不等式

$$\varphi'(f(x) - f(x^*)) d(0, \partial f(x)) \geq 1$$

成立, 则称函数 f 在 $x^* \in \text{dom } \partial f$ 上满足 K-L 性质。

引理 3 (一致 K-L 性质) Ω 是紧集, 设函数 $f: \mathbf{R}^n \rightarrow R \cup \{+\infty\}$ 是恰当下半连续函数。假设 f 在 Ω 上是常数并且在 Ω 上的每个点满足 K-L 性质, 那么, 存在 $\varepsilon > 0, \eta > 0$ 以及 $\varphi \in \Phi_\eta$, 使得对于任意的 $\tilde{x} \in \Omega$ 以及 $x \in \{x \in \mathbf{R}^n : d(x, \Omega) < \varepsilon\} \cap \{f(\tilde{x}) < f < f(\tilde{x}) + \eta\}$, 有

$$\varphi'(f(x) - f(\tilde{x})) d(0, \partial f(x)) \geq 1$$

3 算法及收敛性分析

针对问题(1), 本文提出了一种广义交替方向乘法法(GADMM), 即

假设: 令 $f: \mathbf{R}^n \rightarrow R \cup \{+\infty\}$ 是弱凸函数, 常数 $\omega > 0$; $h: \mathbf{R}^m \rightarrow R$ 是连续可微函数, 它的梯度 ∇h 是利普希茨连续的, 其利普希茨常数 $L_1 > 0$; $g: \mathbf{R}^n \times \mathbf{R}^m \rightarrow R$ 是光滑函数。问题(1)满足下列条件:

- a. 矩阵 B 是列满秩的, 且 $\text{Im}(A) \subset \text{Im}(B)$;
- b. $\|\nabla_y g(x, y) - \nabla_y g(x, \tilde{y})\| \leq L_1(x) \|y - \tilde{y}\|, \forall y, \tilde{y} \in \mathbf{R}^m$;
 $\|\nabla_x g(x, y) - \nabla_x g(\tilde{x}, y)\| \leq L_2(y) \|x - \tilde{x}\|, \forall x, \tilde{x} \in \mathbf{R}^n$;
- c. ∇g 在 $\mathbf{R}^n \times \mathbf{R}^m$ 的有界子集上是利普希茨连续的, 即对于每个有界子集 $B_1 \times B_2 \subseteq \mathbf{R}^n \times \mathbf{R}^m$, 存在 $M > 0$, 使得对于所有的 $(x_i, y_i) \in B_1 \times B_2, i = 1, 2, \|\nabla_x H(x_1, y_1) - \nabla_x H(x_2, y_2), \nabla_y H(x_1, y_1) - \nabla_y H(x_2, y_2)\| \leq M \|(x_1 - y_1, x_2 - y_2)\|$
- d. $A^T A \geq \mu I, \mu > 0$;
- e. 存在 $L_2, L_3 > 0$, 使得 $\sup \{L_2(x^k) : k \in N\} \leq L_2, \sup \{L_3(x^k) : k \in N\} \leq L_3$
- f. $\beta > \tilde{\beta}$,

式中: $\alpha \in (0, 2)$ 为松弛因子。

在收敛分析中, 记 $\omega^k := (x^k, y^k, z^k)$ 及 $v^k := (x^k, y^k)$ 。问题(1)的增广拉格朗日函数为

$$L_\beta(x, y, \gamma) = g(x, y) + f(x) + h(y) - \langle \gamma, Ax + By \rangle + \frac{\beta}{2} \|Ax + By\|^2$$

式中: γ 是与线性约束相关的拉格朗日乘子; β 是惩罚参数, $\beta > 0$ 。

此外, 设

$$\widehat{L}_\beta(x, y, \gamma, \alpha, \vartheta) = g(x, y) + f(x) + h(y) - \langle \gamma, Ax + By \rangle + \frac{\beta}{2} \|\alpha Ax - (1 - \alpha)B\vartheta + By\|^2$$

引理 4 在上述假设条件下, 对于任意的 $l > k$, 有

$$\|\gamma^l - \gamma^k\|^2 \leq \frac{1}{\lambda_{B^T B}} \|B^T \gamma^l - B^T \gamma^k\|^2$$

式中, $\lambda_{B^T B}$ 是 $B^T B$ 的最小的特征值。

证明 由式(4)的第3项, 以及假设的 a, 对

于 2 个整数 $l > k$, 有 $\gamma^l - \gamma^k \in \text{Im}(B)$ 。

由于 $B \in \mathbb{R}^{n \times q}$ 是列满秩的, 因此, 存在 $P \in \mathbb{R}^{q \times q}$, $Q \in \mathbb{R}^{q \times n}$, 且矩阵 P 可逆, $QQ^T = I_{n \times n}$, $B^T = PQ$ 。由 $\text{Im}(B) = \text{Im}(Q^T)$, 可得 $\gamma^l - \gamma^k \in \text{Im}(Q^T)$, $\|\gamma^l - \gamma^k\|^2 = \|Q(\gamma^l - \gamma^k)\|^2$ 。由此得到

$$\begin{aligned} \|B^T(\gamma^l - \gamma^k)\|^2 &= \|PQ(\gamma^l - \gamma^k)\|^2 \geq \\ \lambda_{P^T P} \|Q(\gamma^l - \gamma^k)\|^2 &\geq \lambda_{P^T P} \|\gamma^l - \gamma^k\|^2 \end{aligned}$$

式中, $\lambda_{P^T P}$ 是 $P^T P$ 的最小的特征值。

由矩阵 P 和 Q 的定义, 有 $\lambda_{B^T B} = \lambda_{P^T P}$, 又因为线性代数的常见结论 $\lambda_{P^T P} = \lambda_{P P^T}$, 因此, $\lambda_{P^T P} = \lambda_{B^T B}$ 。

引理 5 令 $\{(x^k, y^k, \gamma^k)\}$ 由算法 GADMM 式(式(5))产生, 则有

$$L_\beta(\omega^k) - L_\beta(\omega^{k+1}) \geq \delta \|v^{k+1} - v^k\|^2 \quad (8)$$

证明 首先将增广拉格朗日函数的差分拆分, 得到

$$\begin{aligned} L_\beta(\omega^k) - L_\beta(\omega^{k+1}) &= L_\beta(x^k, y^k, \gamma^k) - L_\beta(x^{k+1}, y^{k+1}, \gamma^{k+1}) = \widehat{L}_\beta(x^k, y^k, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^{k+1}, \gamma^{k+1}, \alpha, y^k) = \\ &(\widehat{L}_\beta(x^k, y^k, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k)) + (\widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k) - \\ &\widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k)) + (\widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k)) + \\ &(\widehat{L}_\beta(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^{k+1}, \gamma^{k+1}, \alpha, y^k)) \end{aligned} \quad (9)$$

根据定义, 有

$$\begin{aligned} \widehat{L}_\beta(x^k, y^k, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k) &= \\ f(x^k) - f(x^{k+1}) + \langle \gamma^k, Ax^{k+1} - Ax^k \rangle + \\ \beta \langle Ax^{k+1} + By^k, Ax^k - Ax^{k+1} \rangle + \\ \frac{\beta}{2} \|Ax^k - Ax^{k+1}\|^2 + g(x^k, y^k) - g(x^{k+1}, y^k) \end{aligned} \quad (10)$$

因为, f 是关于常数 ω 的弱凸函数, 因此, 由式(7)的第1个关系式得到

$$\begin{aligned} f(x^k) &\geq f(x^{k+1}) + \langle A^T \gamma^k - \beta A^T (Ax^{k+1} + By^k) - \\ &\nabla_x g(x^{k+1}, y^k), x^k - x^{k+1} \rangle - \frac{\omega}{2} \|x^{k+1} - x^k\|^2 \end{aligned}$$

又因为 $\nabla_x g(\cdot, y^k)$ 是关于常数 $L_3(y^k)$ 利普希茨连续的, 由引理 1 可得

$$\begin{aligned} g(x^k, y^k) - g(x^{k+1}, y^k) &\geq \langle \nabla_x g(x^{k+1}, y^k), x^k - x^{k+1} \rangle - \\ \frac{L_3(y^k)}{2} \|x^{k+1} - x^k\|^2 \end{aligned}$$

根据假设的 d, 有

$$\|Ax^{k+1} - Ax^k\|^2 \geq \mu \|x^{k+1} - x^k\|^2$$

将上述 3 个不等式代入式(10), 可得

$$\begin{aligned} \widehat{L}_\beta(x^k, y^k, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k) &\geq \\ \frac{\beta\mu - L_3(y^k) - \omega}{2} \|x^{k+1} - x^k\|^2 \end{aligned} \quad (11)$$

类似地,

$$\begin{aligned} \widehat{L}_\beta(x^{k+1}, y^k, \gamma^k, \alpha, y^k) - \widehat{L}_\beta(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k) &= \\ h(y^k) - h(y^{k+1}) + g(x^{k+1}, y^k) - g(x^{k+1}, y^{k+1}) + \\ \langle \gamma^k, B(y^{k+1} - y^k) \rangle - \alpha\beta \langle Ax^{k+1} + By^k, B(y^{k+1} - y^k) \rangle - \\ \frac{\beta}{2} \|By^{k+1} - By^k\|^2 &= h(y^k) - h(y^{k+1}) + g(x^{k+1}, y^k) - \\ g(x^{k+1}, y^{k+1}) + \langle \gamma^{k+1}, B(y^{k+1} - y^k) \rangle + \\ \frac{\beta}{2} \|By^{k+1} - By^k\|^2 \end{aligned} \quad (12)$$

其中,最后1个等式由

$$\|B y^{k+1}-B y^k\|^2 \geq \lambda_{B^T B}\left\|y^{k+1}-y^k\right\|^2$$

得到.通过 ∇h 的利普希茨连续性,由引理1及式(7)的第2个不等式,可得

$$h\left(y^k\right)-h\left(y^{k+1}\right) \geq\left\langle B^T \gamma^{k+1}-\nabla_y g\left(x^{k+1}, y^{k+1}\right), y^k-y^{k+1}\right\rangle-\frac{L_1}{2}\left\|y^{k+1}-y^k\right\|^2$$

又因为 $\nabla_y g\left(x^{k+1}, \cdot\right)$ 是关于常数 $L_2\left(x^{k+1}\right)$ 利普希茨连续的,由引理1可得

$$g\left(x^{k+1}, y^k\right)-g\left(x^{k+1}, y^{k+1}\right) \geq\left\langle \nabla_y g\left(x^{k+1}, y^{k+1}\right), y^k-y^{k+1}\right\rangle-\frac{L_2\left(x^{k+1}\right)}{2}\left\|y^{k+1}-y^k\right\|^2$$

又由于 B 是列满秩的,可得

$$\left\|B y^{k+1}-B y^k\right\|^2 \geq \lambda_{B^T B}\left\|y^{k+1}-y^k\right\|^2$$

将上述3个等式相加并代入式(12),可得

$$\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k\right) \geq \frac{\beta \lambda_{B^T B}-L_2\left(x^{k+1}\right)-L_1}{2}\left\|y^{k+1}-y^k\right\|^2 \quad (13)$$

现计算式(9)剩余的项。

$$\begin{aligned} & \left(\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)\right)+ \\ & \left(\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k\right)-\right. \\ & \left.\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^{k+1}, \alpha, y^k\right)\right) \end{aligned}$$

事实上,

$$\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)=\frac{\beta}{2}\left\|A x^{k+1}+B y^k\right\|^2-\frac{\beta}{2}\left\|\alpha\left(A x^{k+1}+B y^k\right)\right\|^2 \quad (14)$$

且

$$\begin{aligned} & \widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^{k+1}, \alpha, y^k\right)= \\ & \left\langle \gamma^{k+1}-\gamma^k, A x^{k+1}+B y^{k+1}\right\rangle+\frac{\beta}{2}\left\|\alpha\left(A x^{k+1}+B y^k\right)+\right. \\ & \left.B\left(y^{k+1}-y^k\right)\right\|^2-\frac{\beta}{2}\left\|A x^{k+1}+B y^{k+1}\right\|^2= \\ & \left\langle \gamma^{k+1}-\gamma^k, A x^{k+1}+B y^{k+1}\right\rangle+\frac{\beta}{2}\left\|\alpha\left(A x^{k+1}+B y^k\right)\right\|^2+ \\ & \alpha \beta\left\langle A x^{k+1}+B y^k, B\left(y^{k+1}-y^k\right)\right\rangle-\frac{\beta}{2}\left\|A x^{k+1}+B y^k\right\|^2- \\ & \beta\left\langle A x^{k+1}+B y^k, B\left(y^{k+1}-y^k\right)\right\rangle \end{aligned} \quad (15)$$

将式(14)和式(15)相加,可得

$$\begin{aligned} & \left(\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)\right)+ \\ & \left(\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^{k+1}, \alpha, y^k\right)\right)= \\ & \left\langle \gamma^{k+1}-\gamma^k, A x^{k+1}+B y^{k+1}\right\rangle+\alpha \beta\left\langle A x^{k+1}+B y^k, B\left(y^{k+1}-y^k\right)\right\rangle= \\ & \left\langle \gamma^{k+1}-\gamma^k,-\frac{1}{\alpha \beta}\left(\gamma^{k+1}-\gamma^k\right)-\frac{1-\alpha}{\alpha} B\left(y^{k+1}-y^k\right)\right\rangle+ \\ & \frac{1-\alpha}{\alpha}\left\langle \gamma^{k+1}-\gamma^k+\beta\left(B y^{k+1}-B y^k\right), B\left(y^{k+1}-y^k\right)\right\rangle= \\ & -\frac{1}{\alpha \beta}\left\|\gamma^{k+1}-\gamma^k\right\|^2+\frac{(1-\alpha) \beta}{\alpha}\left\|B\left(y^{k+1}-y^k\right)\right\|^2 \quad (16) \end{aligned}$$

根据引理4可得

$$\begin{aligned} & \left\|\gamma^{k+1}-\gamma^k\right\|^2 \leq \frac{1}{\lambda_{B^T B}}\left\|B^T \gamma^{k+1}-B^T \gamma^k\right\|^2 \\ & \left\|B^T \gamma^{k+1}-B^T \gamma^k\right\|^2=\left\|\nabla h\left(y^{k+1}\right)+\nabla_y g\left(x^{k+1}, y^{k+1}\right)-\right. \\ & \left.\nabla h\left(y^k\right)-\nabla_y g\left(x^k, y^k\right)\right\|^2 \leq 2\left\|\nabla h\left(y^{k+1}\right)-\right. \\ & \left.\nabla h\left(y^k\right)\right\|^2+2\left\|\nabla_y g\left(x^{k+1}, y^{k+1}\right)-\nabla_y g\left(x^k, y^k\right)\right\|^2 \leq \\ & 2 L_1^2\left\|y^{k+1}-y^k\right\|^2+2 M^2\left\|x^{k+1}-x^k\right\|^2+2 M^2\left\|y^{k+1}-y^k\right\|^2= \\ & \left(2 L_1^2+2 M^2\right)\left\|y^{k+1}-y^k\right\|^2+2 M^2\left\|x^{k+1}-x^k\right\|^2 \end{aligned}$$

其中,第2个不等式是由 ∇h 的利普希茨连续性和假设的c得到.因此,

$$\left\|\gamma^{k+1}-\gamma^k\right\|^2 \leq \frac{1}{\lambda_{B^T B}}\left(\left(2 L_1^2+2 M^2\right)\left\|y^{k+1}-y^k\right\|^2+2 M^2\left\|x^{k+1}-x^k\right\|^2\right) \quad (17)$$

将式(17)代入式(16),得到

$$\begin{aligned} & \left(\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^k, \gamma^k, \alpha, y^k\right)\right)+ \\ & \left(\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^k, \alpha, y^k\right)-\widehat{L}_\beta\left(x^{k+1}, y^{k+1}, \gamma^{k+1}, \alpha, y^k\right)\right) \geq \\ & -\frac{1}{\alpha \beta \lambda_{B^T B}}\left(\left(2 L_1^2+2 M^2\right)\left\|y^{k+1}-y^k\right\|^2+\right. \\ & \left.2 M^2\left\|x^{k+1}-x^k\right\|^2\right)+\frac{(1-\alpha) \beta \lambda_{B^T B}}{\alpha}\left\|y^{k+1}-y^k\right\|^2= \\ & \left(\frac{(1-\alpha) \beta \lambda_{B^T B}}{\alpha}-\frac{2 L_1^2+2 M^2}{\alpha \beta \lambda_{B^T B}}\right)\left\|y^{k+1}-y^k\right\|^2- \\ & \frac{2 M^2}{\alpha \beta \lambda_{B^T B}}\left\|x^{k+1}-x^k\right\|^2 \end{aligned} \quad (18)$$

因此,将式(11),(13),(18)代入式(9),得到

$$\begin{aligned}
L_\beta(\omega^k) - L_\beta(\omega^{k+1}) &\geq \left(\frac{\beta\mu - L_3(y^k) - \omega}{2} - \frac{2M^2}{\alpha\beta\lambda_{B^T B}} \right) \times \\
&\quad \left\| x^{k+1} - x^k \right\|^2 + \left(\frac{\beta\lambda_{B^T B} - L_2(x^{k+1}) - L_1}{2} + \right. \\
&\quad \left. \frac{(1-\alpha)\beta\lambda_{B^T B}}{\alpha} - \frac{2L_1^2 + 2M^2}{\alpha\beta\lambda_{B^T B}} \right) \left\| y^{k+1} - y^k \right\|^2 \geq \\
&\quad \left(\frac{\beta\mu - L_3 - \omega}{2} - \frac{2M^2}{\alpha\beta\lambda_{B^T B}} \right) \left\| x^{k+1} - x^k \right\|^2 + \\
&\quad \left(\frac{\beta\lambda_{B^T B} - L_2 - L_1}{2} + \frac{(1-\alpha)\beta\lambda_{B^T B}}{\alpha} - \right. \\
&\quad \left. \frac{2L_1^2 + 2M^2}{\alpha\beta\lambda_{B^T B}} \right) \left\| y^{k+1} - y^k \right\|^2 \geq \delta \left\| v^{k+1} - v^k \right\|^2
\end{aligned}$$

其中, 第2个不等式由假设的 e 得到, 在最后1个等式中, 令

$$\delta := \min \left\{ \frac{\beta\mu_1 - L_3 - \omega}{2} - \frac{2M^2}{\alpha\beta\lambda_{B^T B}}, \frac{\beta\lambda_{B^T B} - L_2 - L_1}{2} + \frac{(1-\alpha)\beta\lambda_{B^T B}}{\alpha} - \frac{2L_1^2 + 2M^2}{\alpha\beta\lambda_{B^T B}} \right\}$$

由假设的 c 可知, $\delta > 0$ 。

引理 6 令序列 $\{\omega^k\}_{k \in \mathbb{N}}$ 由算法 GADMM 生成, 且假设有界, 因此, 有

$$\sum_{k=0}^{+\infty} \left\| \omega^{k+1} - \omega^k \right\|^2 < +\infty \quad (19)$$

证明 由于序列 $\{\omega^k\}_{k \in \mathbb{N}}$ 是有界的, 则存在收敛子列 $\{\omega^{k_j}\}_{j \in \mathbb{N}}$ 且 $\omega^{k_j} \rightarrow \omega^*$ 。由 h 和 g 的连续性和 f 的下半连续性可知 $L_\beta(\cdot)$ 是下半连续的, 从而可得

$$L_\beta(\omega^*) \leq \liminf_{j \rightarrow \infty} L_\beta(\omega^{k_j})$$

因此, 序列 $\{L_\beta(\omega^{k_j})\}_{j \in \mathbb{N}}$ 有下界, 又由式 (17) 可知 $\{L_\beta(\omega^{k_j})\}_{j \in \mathbb{N}}$ 单调递减, 所以, $\{L_\beta(\omega^{k_j})\}_{j \in \mathbb{N}}$ 收敛, 此外, $\{L_\beta(\omega^k)\}_{k \in \mathbb{N}}$ 也是收敛的, 且 $L_\beta(\omega^k) \geq L_\beta(\omega^*)$ 。由引理 4 可得

$$\begin{aligned}
\sum_{k=0}^m \delta \left\| v^{k+1} - v^k \right\|^2 &\leq L_\beta(\omega^0) - L_\beta(\omega^{m+1}) \leq \\
L_\beta(\omega^0) - L_\beta(\omega^*) &< +\infty
\end{aligned}$$

由 $\delta > 0$ 可得

$$\sum_{k=0}^{+\infty} \left\| v^{k+1} - v^k \right\|^2 < +\infty$$

因此,

$$\sum_{k=0}^{+\infty} \left\| x^{k+1} - x^k \right\|^2 < +\infty, \quad \sum_{k=0}^{+\infty} \left\| y^{k+1} - y^k \right\|^2 < +\infty \quad (20)$$

此外, 由式 (17) 和式 (20) 可得

$$\sum_{k=0}^{+\infty} \left\| \gamma^{k+1} - \gamma^k \right\|^2 < +\infty$$

因此,

$$\sum_{k=0}^{+\infty} \left\| \omega^{k+1} - \omega^k \right\|^2 < +\infty$$

引理 7 存在 $\xi > 0$, 使得

$$d(0, \partial L_\beta(\omega^{k+1})) \leq \xi \left\| v^{k+1} - v^k \right\| \quad (21)$$

证明 根据增广拉格朗日函数的定义, 可得

$$\begin{cases} \partial_x L_\beta(\omega^{k+1}) = \partial f(x^{k+1}) + \nabla_x g(x^{k+1}, y^{k+1}) - \\ \quad A^T \gamma^{k+1} + \beta A^T (Ax^{k+1} + By^{k+1}) \\ \partial_y L_\beta(\omega^{k+1}) = \nabla h(y^{k+1}) + \nabla_y g(x^{k+1}, y^{k+1}) - \\ \quad B^T \gamma^{k+1} + \beta B^T (Ax^{k+1} + By^{k+1}) \\ \partial_\gamma L_\beta(\omega^{k+1}) = -(Ax^{k+1} + By^{k+1}) \end{cases} \quad (22)$$

进一步结合式 (7), 可得

$$\begin{cases} A^T(\gamma^k - \gamma^{k+1}) + \beta A^T B(y^{k+1} - y^k) + \\ \quad \nabla_x g(x^{k+1}, y^{k+1}) - \nabla_x g(x^{k+1}, y^k) \in \partial_x L_\beta(\omega^{k+1}) \\ \frac{1}{\alpha} B^T(\gamma^k - \gamma^{k+1}) + \frac{(1-\alpha)\beta}{\alpha} B(y^k - y^{k+1}) \in \\ \quad \partial_y L_\beta(\omega^{k+1}) \\ \frac{1}{\alpha\beta}(\gamma^{k+1} - \gamma^k) + \frac{1-\alpha}{\alpha} B(y^{k+1} - y^k) \in \\ \quad \partial_\gamma L_\beta(\omega^{k+1}) \end{cases} \quad (23)$$

令

$$\begin{aligned}
x_{k+1}^* &= A^T(\gamma^k - \gamma^{k+1}) + \beta A^T B(y^{k+1} - y^k) + \\ &\quad \nabla_x g(x^{k+1}, y^{k+1}) - \nabla_x g(x^{k+1}, y^k) \\
y_{k+1}^* &= \frac{1}{\alpha} B^T(\gamma^k - \gamma^{k+1}) + \frac{(1-\alpha)\beta}{\alpha} B(y^k - y^{k+1}) \\
\gamma_{k+1}^* &= \frac{1}{\alpha\beta}(\gamma^{k+1} - \gamma^k) + \frac{1-\alpha}{\alpha} B(y^{k+1} - y^k)
\end{aligned}$$

则 $(x_{k+1}^*, y_{k+1}^*, \gamma_{k+1}^*) \in L_\beta(\omega^{k+1})$, 且存在 $\xi_1, \xi_2, \xi_3 > 0$, 使得

$$\begin{aligned}
\left\| x_{k+1}^*, y_{k+1}^*, \gamma_{k+1}^* \right\| &\leq \xi_1 \left\| y^{k+1} - y^k \right\| + \xi_2 \left\| \gamma^{k+1} - \gamma^k \right\| + \\ &\quad \xi_3 \left\| \nabla_x g(x^{k+1}, y^{k+1}) - \nabla_x g(x^{k+1}, y^k) \right\| \quad (24)
\end{aligned}$$

又由假设的 c 可得

$$\left\|\nabla_x g\left(x^{k+1}, y^{k+1}\right)-\nabla_x g\left(x^{k+1}, y^k\right)\right\| \leq M\left\|y^{k+1}-y^k\right\| \quad(25)$$

根据式(17)可得

$$\left\|\gamma^{k+1}-\gamma^k\right\| \leq \frac{1}{\sqrt{\lambda_{B^T B}}}\left(\sqrt{\left(2 L_1^2+2 M^2\right)}\left\|y^{k+1}-y^k\right\|+\sqrt{2} M\left\|x^{k+1}-x^k\right\|\right) \quad(26)$$

因此, 由式(24)~(26)可得

$$d\left(0, \partial L_{\beta}\left(\omega^{k+1}\right)\right) \leq\left\|x_{k+1}^*, y_{k+1}^*, \gamma_{k+1}^*\right\| \leq\left(\xi_1+\sqrt{\frac{\left(2 L_1^2+2 M^2\right)}{\lambda_{B^T B}}+M \xi_3}\right)\left\|y^{k+1}-y^k\right\|+\sqrt{\frac{2}{\lambda_{B^T B}}} M\left\|x^{k+1}-x^k\right\|$$

令

$$\xi=\sqrt{\left(\xi_1+\sqrt{\frac{\left(2 L_1^2+2 M^2\right)}{\lambda_{B^T B}}+M \xi_3}\right)^2+\left(\sqrt{\frac{2}{\lambda_{B^T B}}} M\right)^2}$$

由上式可以得到式(21)。

引理 8 假设序列 $\left\{\omega^k\right\}_{k \in N}$ 为算法生成的有界序列, 其所有极限点记为 $S\left(\omega^0\right)$, 则

a. $S\left(\omega^0\right)$ 是一个非空紧集, 并且 $d\left(\omega^k, S\left(\omega^0\right)\right) \rightarrow 0, k \rightarrow+\infty$;

b. $S\left(\omega^0\right) \subset \operatorname{crit} L_{\beta}$;

c. $L_{\beta}(\cdot)$ 在 $S\left(\omega^0\right)$ 上取有限值且为常数, 且

$$\inf _{k \in N} L_{\beta}\left(\omega^k\right)=\lim _{k \rightarrow+\infty} L_{\beta}\left(\omega^k\right)$$

证明 a. 由 $S\left(\omega^0\right)$ 的定义可直接得到。

b. 对于任意点 $\left(x^*, y^*, \gamma^*\right) \in S\left(\omega^0\right)$, 存在子列

$$\left\{\left(x^{k_j}, y^{k_j}, \gamma^{k_j}\right)\right\} \rightarrow\left(x^*, y^*, \gamma^*\right), k_j \rightarrow+\infty$$

根据增广拉格朗日函数的定义, 式(5)的 x 子问题等价于

$$x^{k+1} \in \arg \min _x\left\{L_{\beta}\left(x, y^k, \gamma^k\right)\right\}$$

即 x^{k+1} 是 $L_{\beta}\left(x, y^k, \gamma^k\right)$ 关于变量 x 的全局最小点, 由此可得

$$L_{\beta}\left(x^{k+1}, y^k, \gamma^k\right) \leq L_{\beta}\left(x^*, y^k, \gamma^k\right)$$

又因为 $L_{\beta}(\cdot)$ 关于 y 和 γ 连续, 则有

$$\limsup _{j \rightarrow+\infty} L_{\beta}\left(x^{k_j+1}, y^{k_j}, \gamma^{k_j}\right)=\limsup _{j \rightarrow+\infty} L_{\beta}\left(x^{k_j+1}, y^{k_j+1}, \gamma^{k_j+1}\right) \leq L_{\beta}\left(x^*, y^*, \gamma^*\right) \quad(27)$$

由引理 6 可得, $\sum_{k=0}^{+\infty}\left\|\omega^{k+1}-\omega^k\right\|^2<+\infty$, 从而 $\left\{\left(x^{k_j+1}, y^{k_j+1}, \gamma^{k_j+1}\right)\right\} \rightarrow\left(x^*, y^*, \gamma^*\right), k_j \rightarrow+\infty$

由函数 $L_{\beta}(\cdot)$ 的下半连续性可得

$$\liminf _{j \rightarrow+\infty} L_{\beta}\left(x^{k_j+1}, y^{k_j+1}, \gamma^{k_j+1}\right) \geq L_{\beta}\left(x^*, y^*, \gamma^*\right) \quad(28)$$

因此, 结合式(27)和式(28), 可得

$$\lim _{j \rightarrow+\infty} L_{\beta}\left(x^{k_j+1}, y^{k_j+1}, \gamma^{k_j+1}\right)=L_{\beta}\left(x^*, y^*, \gamma^*\right)$$

由此可以推断

$$\lim _{j \rightarrow+\infty} f\left(x^{k_j+1}\right)=f\left(x^*\right) \quad(29)$$

结合 $\nabla h, \nabla g_x(\cdot, \cdot), \nabla g_y(\cdot, \cdot)$ 的连续性, 在式(6)中对序列 $\left\{\left(x^{k_j}, y^{k_j}, \gamma^{k_j}\right)\right\}$ 取极限, 可得

$$\left\{\begin{array}{l} A^T \gamma^k-\nabla_x g\left(x^{k+1}, y^k\right) \in \partial f\left(x^*\right) \\ B^T \gamma^{k+1}-\nabla_y g\left(x^{k+1}, y^{k+1}\right)=\nabla h\left(y^*\right) \\ A x^*+B y^*=0 \end{array}\right.$$

即 $\left(x^*, y^*, \gamma^*\right) \in S\left(\omega^0\right)$ 。

c. 对于任意点 $\left(x^*, y^*, \gamma^*\right) \in S\left(\omega^0\right)$, 存在子列

$$\left\{\left(x^{k_j}, y^{k_j}, \gamma^{k_j}\right)\right\} \rightarrow\left(x^*, y^*, \gamma^*\right), k_j \rightarrow+\infty$$

由于 $L_{\beta}\left(\omega^k\right)$ 单调递减, 结合式(27)和式(28), 可得

$$\lim _{j \rightarrow+\infty} L_{\beta}\left(x^k, y^k, \gamma^k\right)=L_{\beta}\left(x^*, y^*, \gamma^*\right)$$

因此, $L_{\beta}(\cdot)$ 在 $S\left(\omega^0\right)$ 是常数, 且

$$\inf _{k \in N} L_{\beta}\left(\omega^k\right)=\lim _{k \rightarrow+\infty} L_{\beta}\left(\omega^k\right)$$

定理 1 若 $L_{\beta}(\cdot)$ 满足 K-L 性质, 则 $\sum_{k=0}^{+\infty}\left\|\omega^{k+1}-\omega^k\right\|<+\infty$, 且序列 $\left\{\omega^k\right\}_{k \in N}$ 收敛到 $L_{\beta}(\cdot)$ 的一个稳定点。

证明 由引理 8 可知, $\lim _{k \rightarrow+\infty} L_{\beta}\left(\omega^k\right)=L_{\beta}\left(\omega^*\right), \forall \omega^* \in S\left(\omega^0\right)$, 因此, 考虑以下 2 种情况:

a. 存在整数 k_0 使得 $L_{\beta}\left(\omega^{k_0}\right)=L_{\beta}\left(\omega^*\right)$, 由引理 5 可知,

$$\delta\left\|v^{k+1}-v^k\right\|^2 \leq L_{\beta}\left(\omega^k\right)-L_{\beta}\left(\omega^{k+1}\right) \leq L_{\beta}\left(\omega^{k_0}\right)-L_{\beta}\left(\omega^*\right)=0$$

因此, 对任意的 $k>k_0$, 有 $v^{k+1}=v^k$, 结合式(16)可知, 对于任意的 $k>k_0+1$, 有 $\omega^{k+1}=\omega^k$ 成立。

b. 假设对于任意的 k 都有 $L_{\beta}\left(\omega^k\right)>L_{\beta}\left(\omega^*\right)$ 成立, 存在 $\tilde{k}>0$ 使得对于所有的 $k>\tilde{k}$, 有

$$\delta\left\|v^{k+1}-v^k\right\|^2 \leq \xi\left\|v^k-v^{k-1}\right\| \Delta_{k, k+1} \quad(30)$$

其中,

$$A_{p,q} = \varphi(L_\beta(\omega^p) - L_\beta(\omega^*)) - \varphi(L_\beta(\omega^q) - L_\beta(\omega^*))$$

由于 $d(\omega^k, S(\omega^0)) \rightarrow 0$ 且 $L_\beta(\omega^k) \rightarrow L_\beta(\omega^*)$, 那么, 对于任意的 $\varepsilon, \kappa > 0$, 存在 $\tilde{k} > 0$ 使得对于所有的 $k > \tilde{k}$, 有

$$d(\omega^k, S(\omega^0)) < \varepsilon, L_\beta(\omega^*) < L_\beta(\omega^k) < L_\beta(\omega^*) + \kappa$$

由于 $S(\omega^0)$ 是非空紧集, 且 $L_\beta(\cdot)$ 在 $S(\omega^0)$ 是常数, 由引理 3 可得, 对所有的 $k > \tilde{k}$, 有

$$\varphi'(L_\beta(\omega^k) - L_\beta(\omega^*))d(0, \partial L_\beta(\omega^k)) \geq 1$$

由于

$$L_\beta(\omega^k) - L_\beta(\omega^{k+1}) = (L_\beta(\omega^k) - L_\beta(\omega^*)) - (L_\beta(\omega^{k+1}) - L_\beta(\omega^*))$$

利用函数 φ 的凹性, 可得

$$\begin{aligned} & \varphi(L_\beta(\omega^k) - L_\beta(\omega^*)) - \varphi(L_\beta(\omega^{k+1}) - L_\beta(\omega^*)) \geq \\ & \varphi'(L_\beta(\omega^k) - L_\beta(\omega^*)) (L_\beta(\omega^k) - L_\beta(\omega^{k+1})) \end{aligned}$$

因此, 结合引理 7 以及 $\varphi'(L_\beta(\omega^k) - L_\beta(\omega^*)) > 0$, 可得

$$\begin{aligned} & L_\beta(\omega^k) - L_\beta(\omega^{k+1}) \leq \\ & \frac{\varphi(L_\beta(\omega^k) - L_\beta(\omega^*)) - \varphi(L_\beta(\omega^{k+1}) - L_\beta(\omega^*))}{\varphi'(L_\beta(\omega^k) - L_\beta(\omega^*))} \leq \\ & \xi \|v^k - v^{k-1}\| \|\varphi(L_\beta(\omega^k) - L_\beta(\omega^*)) - \\ & \varphi(L_\beta(\omega^{k+1}) - L_\beta(\omega^*))\| = \xi \|v^k - v^{k-1}\| \Delta_{k,k+1} \end{aligned}$$

结合引理 5, 可得式 (30) 成立, 且由式 (30) 可得

$$\|v^{k+1} - v^k\| \leq \sqrt{\frac{\xi}{\delta} \Delta_{k,k+1}} \cdot \|v^k - v^{k-1}\|^{\frac{1}{2}}$$

利用不等式 $2\sqrt{\alpha\beta} \leq \alpha + \beta (\forall \alpha, \beta > 0)$, 可得

$$2\|v^{k+1} - v^k\| \leq \|v^k - v^{k-1}\| + \frac{\xi}{\delta} \Delta_{k,k+1}$$

进一步, 将上式从 $\tilde{k}+1$ 到 m 求和, 得到

$$2 \sum_{k=\tilde{k}+1}^m \|v^{k+1} - v^k\| \leq \sum_{k=\tilde{k}+1}^m \|v^k - v^{k-1}\| + \frac{\xi}{\delta} \Delta_{\tilde{k}+1,m}$$

由于 $\varphi(L_\beta(\omega^{m+1}) - L_\beta(\omega^*)) > 0$, 令 $m \rightarrow +\infty$, 上式变为

$$\begin{aligned} & \sum_{k=\tilde{k}+1}^{+\infty} \|v^{k+1} - v^k\| \leq \|v^k - v^{k-1}\| + \\ & \frac{\xi}{\delta} \varphi(L_\beta(\omega^{\tilde{k}+1}) - L_\beta(\omega^*)) \quad (31) \end{aligned}$$

即 $\sum_{k=0}^{+\infty} \|v^{k+1} - v^k\| < +\infty$, 因此,

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\| < +\infty, \sum_{k=0}^{+\infty} \|y^{k+1} - y^k\| < +\infty$$

进一步由式 (17) 可得

$$\sum_{k=0}^{+\infty} \|\gamma^{k+1} - \gamma^k\| < +\infty$$

此外, 由于

$$\begin{aligned} \|\omega^{k+1} - \omega^k\| &= \left(\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2 + \right. \\ & \left. \|\gamma^{k+1} - \gamma^k\|^2 \right)^{\frac{1}{2}} \leq \|x^{k+1} - x^k\| + \\ & \|y^{k+1} - y^k\| + \|\gamma^{k+1} - \gamma^k\| \end{aligned}$$

因此,

$$\sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\| < +\infty$$

且 $\{\omega^k\}_{k \in N}$ 是 Cauchy 序列, 因此, $\{\omega^k\}_{k \in N}$ 收敛。

4 结 论

研究了非凸不可分离的线性约束优化问题, 以及求解该问题的广义交替方向乘子法 (GADMM), 在矩阵 B 不是单位阵的情况下, 要求其列满秩, 通过假设增广拉格朗日函数满足 K-L 不等式, 证明了当惩罚参数大于一定的阈值时, 算法生成的序列收敛到增广拉格朗日函数的稳定点。

参考文献:

- [1] BOLEY D. Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs[J]. *SIAM Journal on Optimization*, 2013, 23(4): 2183–2207.
- [2] HAN D R, YUAN X M. Local linear convergence of the alternating direction method of multipliers for quadratic programs[J]. *SIAM Journal on Numerical Analysis*, 2013, 51(6): 3446–3457.
- [3] HE B S, YUAN X M. On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method[J]. *SIAM Journal on Numerical Analysis*, 2012, 50(2): 700–709.
- [4] GUO K, HAN D R, WANG D Z W, et al. Convergence of ADMM for multi-block nonconvex separable optimization models[J]. *Frontiers of Mathematics in China*, 2017, 12(5): 1139–1162.

- [5] LI G Y, PONG T K. Global convergence of splitting methods for nonconvex composite optimization[J]. *SIAM Journal on Optimization*, 2015, 25(4): 2434–2460.
 - [6] GUO K, HAN D R, WU T T. Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints[J]. *International Journal of Computer Mathematics*, 2017, 94(8): 1653–1669.
 - [7] 蒋峰, 党亚峥. 求解凸优化问题的改进对称交替方向乘法[J]. *上海理工大学学报*, 2020, 42(3): 269–274.
 - [8] 王欣, 郭科. 一类非凸优化问题广义交替方向法的收敛性[J]. *应用数学和力学*, 2018, 39(12): 1410–1425.
 - [9] GAO X, ZHANG S Z. First-order algorithms for convex optimization with nonseparable objective and coupled constraints[J]. *Journal of the Operations Research Society of China*, 2017, 5(2): 131–159.
 - [10] CHEN C H, LI M, LIU X, et al. Extended ADMM and BCD for nonseparable convex minimization models with quadratic coupling terms: convergence analysis and insights[J]. *Mathematical Programming*, 2019, 173(1/2): 37–77.
 - [11] LIU Q H, SHEN X Y, GU Y T. Linearized ADMM for nonconvex nonsmooth optimization with convergence analysis[J]. *IEEE Access*, 2019, 7: 76131–76144.
 - [12] WANG F H, XU Z B, XU H K. Convergence of Bregman alternating direction method with multipliers for nonconvex composite problems[J]. *Eprint Arxiv*, 2014, 2014: 1–17.
 - [13] GUO K, WANG X. Convergence of generalized alternating direction method of multipliers for nonseparable nonconvex objective with linear constraints[J]. *Journal of Mathematical Research with Applications*, 2018, 38(5): 523–540.
- (编辑: 石 瑛)

[上接第 579 页]

- [14] WU X, LI H L, CHENG K, et al. Modified graphene/polyimide composite films with strongly enhanced thermal conductivity[J]. *Nanoscale*, 2019, 11(17): 8219–8225.
 - [15] QIU H X, HAN X B, QIU F L, et al. Facile route to covalently-jointed graphene/polyaniline composite and it's enhanced electrochemical performances for supercapacitors[J]. *Applied Surface Science*, 2016, 376: 261–268.
 - [16] HUMMERS W S JR, OFFEMAN R E. Preparation of graphitic oxide[J]. *Journal of the American Chemical Society*, 1958, 80(6): 1339.
 - [17] 江陆洋, 李昊亮, 吴限, 等. 石墨烯量子点用于修复石墨烯结构缺陷及其薄膜导热性能研究 [J]. 上海理工大学学报, 2020, 42(2): 163–171.
 - [18] MIAO J, LI H L, QIU H X, et al. Graphene/PANI hybrid
 - [19] ZOU R, LIU F, HU N, et al. Carbonized polydopamine nanoparticle reinforced graphene films with superior thermal conductivity[J]. *Carbon*, 2019, 149: 173–180.
 - [20] ZOU R, LIU F, HU N, et al. 1-Pyrenemethanol derived nanocrystal reinforced graphene films with high thermal conductivity and flexibility[J]. *Nanotechnology*, 2020, 31(6): 065602.
 - [21] GHOSH T, BISWAS C, OH J, et al. Solution-processed graphite membrane from reassembled graphene oxide[J]. *Chemistry of Materials*, 2012, 24(3): 594–599.
 - [22] ZENGIN H, ZHOU W S, JIN J Y, et al. Carbon nanotube doped polyaniline[J]. *Advanced Materials*, 2002, 14(20): 1480–1483.
- (编辑: 丁红艺)