

一种改进的乘子交替方向法在 ℓ_1 -正则化分裂可行问题中的应用

党亚峥, 唐崇伟

(上海理工大学 管理学院, 上海 200093)

摘要: 提出了一种改进的乘子交替方向法 (ADMM) 算法, 基于松弛技术和预测-校正框架, 将松弛算子引入子问题 x 和对偶变量 λ , 使得每次迭代的步长大于 1, 从而提高了算法的收敛性, 并在变分不等式的框架下证明了该算法的收敛性。此外, 数值实验中通过图像去模糊问题验证了算法的有效性, 并基于多组对照实验, 综合考虑收敛效率和图像质量, 选取适当的收敛准则。

关键词: ℓ_1 范数; 改进的乘子交替方向法; 松弛因子; 图像模糊

中图分类号: O 221 **文献标志码:** A

An improved alternating direction method of multipliers for ℓ_1 -norm regularization splitting feasibility problem

DANG Yazheng, TANG Chongwei

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: we proposes an improved alternating direction method of multipliers (ADMM) algorithm based on the relaxation technique and the prediction-correction framework, which introduces the new parameters in the subproblem x and the dual problem λ , so that the step size of each iteration is greater than 1, thereby improving the convergence of the algorithm. The convergence of the algorithm is proved in the framework of variational inequality. Moreover, the image deblurring problem in numerical experiments verifies that the algorithm is effective. Based on multiple sets of convergence criteria, the appropriate value is selected by comprehensively considering the rate of convergence and the quality of images.

Keywords: ℓ_1 -norm; improved alternating direction method of multipliers; relaxed parameter; image deblurring

1 研究现状

首次提出, 其数学模型实际上是一个可行性问题, 即找一个具有以下属性的点 x^* :

分裂可行性问题 (SFP) 由 Censor 和 Elfving^[1]

$$x^* \in C, \quad Ax^* \in Q \quad (1)$$

其中, C 和 Q 属于非空的闭凸集, 且 $C \in \mathbb{R}^n$, $Q \in \mathbb{R}^m$, $A: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ 为线性运算符。

众所周知, 问题(1)包含现实中各种特殊情况下的逆问题。例如, 当 $Q = b$ 是单元素集合时, 问题(1)简化为经典的凸约束线性逆问题, 即找到一个点 x^* , 使得

$$x^* \in C, \quad Ax^* = b \quad (2)$$

问题(2)在文献 [2-5] 及其参考文献中有着广泛的研究。由于约束集 Q 结构的不确定性, 有关问题(1)的研究比其特殊情况问题(2)要少得多, 同时某些适用于问题(2)的算法并不能直接运用于问题(1)。例如, 目前尚不清楚是否可以将文献 [4] 中针对问题(2)提出的算法扩展应用到问题(1)中。

事实上, 问题(1)等价于如下约束优化问题:

$$\min_{x \in C} f(x) := \frac{1}{2} \|Ax - P_Q Ax\|^2 \quad (3)$$

式中, P_Q 为 Q 上的投影。

受矩阵 A 的影响, 问题(1)和问题(2)时常是不适定的, 因此, 需要对其进行正则化处理, 所以, 在过去的几十年中, 利用正则化技术解决问题(2)已经非常成熟了^[5-8]。于是, 针对问题(1)的正则化方法成为了新的研究话题^[9-12], 由于 ℓ_1 范数会导致解的稀疏性, 故其正则化问题引起了很多学者的关注^[13-17]。因此, 考虑如下 ℓ_1 范数形式, 即

$$\min_{x \in C} \frac{1}{2} \|Ax - P_Q(Ax)\|^2 + \sigma \|x\|_1 \quad (4)$$

其中, $\sigma > 0$ 是正则化参数。显然, lasso 问题、BP 问题、BPDN 问题^[13, 16-19] 皆属于问题(4)的特殊情况, 相较而言, 问题(4)则很少被提及。

在变分不等式(VI)的框架下, 通过乘子交替方向法(ADMM)^[20] 引入一种可实现的分裂算法来解决 ℓ_1 范数正则化的 SFP 问题。为此, 变量拆分后得到

$$\begin{aligned} \min_{x, z} f(x) + \sigma \|z\|_1 \\ \text{s.t. } x = z, \quad x \in C, \quad z \in \mathbb{R}^n \end{aligned} \quad (5)$$

其中, $f(x)$ 为问题(3)中所定义的形式。如今 ADMM 已广泛应用于变分不等式、压缩感知、图像处理及统计学习等领域^[13-14, 21-26]。

He 等^[27] 针对问题(5)的可分离结构, 通过结合线性化思想和邻近点算法, 提出了一种新的可实现 ADMM 算法, 并设计了预测和校正两个步骤。文献 [28] 中引入一组松弛算子, 在取值恰当的情况下具有良好的收敛效果。本文受文献 [27-28] 启发, 提出了一种改进的 ADMM 算法, 通过在预

测步中引入松弛因子对算法进行改进, 提高 ADMM 算法的收敛性。并在一定条件下证明了算法的收敛性, 最后通过图像恢复数值实验验证了算法的有效性。

2 预备知识

现给出一些常用的预备知识。

令 $\mathbb{R}^n$ 为一个 n 维的欧式空间, 给定 $x, y \in \mathbb{R}^n$, $\langle x, y \rangle$ 为内积。此外, 定义 $\|x\|_M = \sqrt{\langle x, Mx \rangle}$ 为 x 的 M 范数, 对于任意的 A , 定义 $\|A\|$ 为 2 范数。

定义 1 由 $\mathbb{R}^n$ 投影到 M 范数下的非空闭凸集上的投影算子

$$P_{Q,M}[x] \sim \operatorname{argmin} \{ \|y - x\|_M \mid y \in Q \}, \quad x \in \mathbb{R}^n \quad (6)$$

正交投影算子可以通过以下关系式来表示:

$$\langle u - P_{Q,M}[u], M(v - P_{Q,M}[u]) \rangle \leq 0, \quad u \in \mathbb{R}^n, v \in Q \quad (7)$$

证明过程参考文献 [29]。

定义 2 若集值算子 $T \in (\mathbb{R}^n, 2^{\mathbb{R}^n})$ 是单调的, 则必有

$$\langle x_1 - x_2, x_1^* - x_2^* \rangle \geq 0, \quad x_1, x_2 \in \mathbb{R}^n, x_1^* \in T(x_1), x_2^* \in T(x_2)$$

如果单调算子的图形没有正确包含在任何其他单调算子的曲线图中, 则称其为最大单调。

定义 3 令 $g: \mathbb{R}^n \rightarrow \mathbb{R}$ 为一个可微的凸函数, 那么, 它的梯度 ∇g 是 Lipschitz 连续的, 常数 $L > 0$, 有

$$\|\nabla g(x_1) - \nabla g(x_2)\| \leq L \|x_1 - x_2\|, \quad x_1, x_2 \in \mathbb{R}^n$$

注意到式(3)中定义的函数 f 的梯度为 $\nabla f = A^T(I - P_Q)A$, 显然 ∇f 是单调且 Lipschitz 连续的, 可知:

$$\|\nabla f(x_1) - \nabla f(x_2)\| \leq \|A\|^2 \|x_1 - x_2\|, \quad x_1, x_2 \in \mathbb{R}^n \quad (8)$$

3 改进的 ADMM 算法

在问题(5)中, 给定第 k 次迭代 $w^k := (x^k, z^k, \lambda^k) \in C \times \mathbb{R}^n \times \mathbb{R}^n$, 则第 $k+1$ 次迭代表示为 $w^{k+1} := (x^{k+1}, z^{k+1}, \lambda^{k+1})$, 由此可以得到以下的子问题:

$$z^{k+1} = \operatorname{argmin}_{z \in \mathbb{R}^n} \left\{ \|z\|_1 + \frac{\gamma}{2} \left\| x^k - z - \frac{\lambda^k}{\gamma} \right\|^2 \right\} \quad (9)$$

$$x^{k+1} = \operatorname{argmin}_{x \in C} \left\{ f(x) + \frac{\gamma}{2} \left\| x - z^{k+1} - \frac{\lambda^k}{\gamma} \right\|^2 \right\} \quad (10)$$

$$\lambda^{k+1} = \lambda^k - \gamma(x^{k+1} - z^{k+1}) \quad (11)$$

式中, λ 为拉格朗日算子, $\gamma > 0$ 为罚因子。

ADMM算法的优势在于充分利用了问题(5)的可分离结构, He等^[27]通过下式生成预测变量 $\tilde{z}^k$

$$\tilde{z}^k = \operatorname{argmin}_{z \in \mathbb{R}^n} \left\{ \|z\|_1 + \frac{\gamma}{2} \left\| x^k - z - \frac{\lambda^k}{\gamma} \right\|^2 + \frac{\rho}{2} \|z - \tilde{z}^k\|^2 \right\} \quad (12)$$

其中, $\rho \geq \|A\|^2$ 为正常数, 通过近似项 $\frac{\rho}{2} \|z - \tilde{z}^k\|^2$ 使子问题 z 正则化。

对于子问题 x , 由于函数 f 的非线性, 使得最小化子问题(10)无法获得封闭解。He等^[27]利用线性化思想处理 f 的非线性, 给定 $(x^k, \tilde{z}^k, \lambda^k)$ 生成如下预测变量 $\tilde{x}^k$:

$$\tilde{x}^k = \operatorname{argmin}_{x \in C} \left\{ \langle \nabla f(x^k), x - x^k \rangle + \frac{\gamma}{2} \left\| x - \tilde{z}^k - \frac{\lambda^k}{\gamma} \right\|^2 + \frac{\rho}{2} \|x - x^k\|^2 \right\} = P_C \left[\frac{1}{\rho + \gamma} (\rho x^k + \gamma \tilde{z}^k + \lambda^k - \nabla f(x^k)) \right]$$

本文算法引入 $\omega \tilde{z}^k + (1 - \omega)x^k$ 来对 $\tilde{x}^k$ 子问题中的 $\tilde{z}^k$ 进行改进, 生成新的预测变量 $\tilde{x}^k$, 即

$$\tilde{x}^k = P_C \left[\frac{1}{\rho + \gamma} [\rho \tilde{x}^k + \gamma [\omega \tilde{z}^k + (1 - \omega)x^k] + \lambda^k - \nabla f(x^k)] \right] \quad (13)$$

采用如下形式的预测变量 $\tilde{\lambda}^k$:

$$\tilde{\lambda}^k = \lambda^k - \gamma \left\{ \tilde{x}^k - [\omega \tilde{z}^k + (1 - \omega)x^k] \right\} \quad (14)$$

其中, $\omega \in [1, 2]$ 详见文献[28]。于是, 得到了一组完整的预测序列 $\tilde{w}^k = (\tilde{x}^k, \tilde{z}^k, \tilde{\lambda}^k)$ 。

为方便起见, 定义 $\eta_1(x, \tilde{x}) := \nabla f(x) - \nabla f(\tilde{x})$, 则有

$$\eta(w, \tilde{w}) := \begin{pmatrix} \eta_1(x, \tilde{x}) \\ 0 \\ 0 \end{pmatrix} \quad (15)$$

最后, 对以上给出的预测序列进行校正, 即可得到改进的ADMM(IADMM)算法。

IADMM算法:

a. 令 $\rho \geq \|A\|^2$, $t \in (0, 2)$, 并设 $w^0 := (x^0, y^0, \lambda^0) \in C \times \mathbb{R}^n \times \mathbb{R}^n$;

b. 取 $k = 0, 1, 2, \dots, k$;

c. 根据式(12)~(14)得到预测序列 $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k)$;

d. 根据式(16)迭代 $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$;

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

其中,

$$d(w^k, \tilde{w}^k) = (w^k - \tilde{w}^k) - H^{-1}\eta(w^k, \tilde{w}^k) \quad (17)$$

$$\alpha_k := \frac{\Gamma(w^k, \tilde{w}^k)}{\|d(w^k, \tilde{w}^k)\|_H^2} \quad (18)$$

$$\Gamma(w^k, \tilde{w}^k) = \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle - \langle \tilde{w}^k - w^k, Hd(w^k, \tilde{w}^k) \rangle \quad (19)$$

e. 终止准则

$$\frac{\|x^{k+1} - x^k\|}{\|x^k\|} \leq 10^{-4}$$

f. 结束。

4 收敛性

现证明算法的全局收敛性。

问题(5)的解集可以看成找到对应变分不等式(VI)的解集 $w^* \in \Omega$, 即

$$\langle w - w^*, F(w^*) \rangle \geq 0, w \in \Omega \quad (20)$$

其中,

$$w := \begin{pmatrix} x \\ z \\ \lambda \end{pmatrix}, F(w) := \begin{pmatrix} \nabla f(x) - \lambda \\ \zeta + \lambda \\ x - z \end{pmatrix}, \Omega \in C \times \mathbb{R}^n \times \mathbb{R}^n$$

引理1 令 w^* 为变分不等式(20)的解, 必有

$$\langle w^k - w^*, Hd(w^k, \tilde{w}^k) \rangle \geq \Gamma(w^k, \tilde{w}^k)$$

证明 首先将式(13)改写为对应的变分不等式(VI)形式, 即

$$\langle x - \tilde{x}^k, \nabla f(x^k) + (\rho + \gamma)(\tilde{x}^k - x^k) + \gamma x^k - \gamma [\omega \tilde{z}^k + (1 - \omega)x^k] - \gamma^k \rangle \geq 0, x \in C \quad (21)$$

由式(14)和式(15)对式(21)进行简单的整理, 可得

$$\langle x - \tilde{x}^k, \nabla f(\tilde{x}^k) - \tilde{\lambda}^k + (\rho + \gamma)(\tilde{x}^k - x^k) + \eta_1(x^k, \tilde{x}^k) \rangle \geq \gamma \langle x - \tilde{x}^k, \tilde{x}^k - x^k \rangle, x \in C \quad (22)$$

同理, 对式(12)和式(14)作同样的变换, 有

$$\langle z - \tilde{z}^k, \zeta^k + \tilde{\lambda}^k + \rho(\tilde{z}^k - z^k) \rangle \geq \gamma \langle \tilde{z}^k - z, \tilde{x}^k - x^k \rangle, z \in \mathbb{R}^n \quad (23)$$

$$\langle \lambda - \tilde{\lambda}^k, \tilde{x}^k - \tilde{z}^k + \frac{1}{\lambda}(\tilde{\lambda}^k - \lambda^k) \rangle \geq \gamma(1 - \omega) \langle x^k - \tilde{z}^k, \tilde{\lambda}^k - \lambda \rangle, \lambda \in \mathbb{R}^n \quad (24)$$

其中, $\zeta \in \partial(\sigma\|\cdot\|_1)(z)$ 。

综合式(22)~(24), 可得如下紧凑形式:

$$\langle w - \tilde{w}^k, F(\tilde{w}^k) - Hd(w^k, \tilde{w}^k) \rangle \geq \gamma \langle x - \tilde{x}^k + \tilde{z}^k - z, \tilde{x}^k - x^k \rangle + \langle \gamma(1 - \omega)(x^k - \tilde{z}^k), \tilde{\lambda}^k - \lambda \rangle, w \in \Omega$$

$$H \sim \begin{pmatrix} (\rho + \gamma)I_n & 0 & 0 \\ 0 & \rho I_n & 0 \\ 0 & 0 & \frac{1}{\gamma}I_n \end{pmatrix}$$

式中, I_n 为 n 维单位矩阵。

令 $w := w^*$, 显然, $x^* = y^*$, 同时结合式(14), 则有

$$\begin{aligned} & \langle w^* - \tilde{w}^k, F(\tilde{w}^k) - Hd(w^k, \tilde{w}^k) \rangle \geq \\ & \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle + \gamma(1 - \omega) \langle \tilde{z}^k - x^k, \tilde{x}^k - x^k \rangle + \\ & \gamma(1 - \omega) \langle x^k - \tilde{z}^k, \tilde{\lambda}^k - \lambda \rangle \geq \\ & \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle, w \in \Omega \end{aligned}$$

对上式进行恒等变换

$$\langle \tilde{w}^k - w^*, Hd(w^k, \tilde{w}^k) \rangle \geq \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle + \langle \tilde{w}^k - w^*, F(\tilde{w}^k) \rangle$$

由于 $\sigma\|\cdot\|_1$ 和 f 皆属于凸函数, 故易知 F 是单调的, 即

$$\langle \tilde{w}^k - w^*, F(\tilde{w}^k) \rangle \geq \langle \tilde{w}^k - w^*, F(\tilde{w}^*) \rangle \geq 0$$

于是, 有

$$\langle \tilde{w}^k - w^*, Hd(w^k, \tilde{w}^k) \rangle \geq \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle$$

因此,

$$\begin{aligned} & \langle \tilde{w}^k - w^*, Hd(w^k, \tilde{w}^k) \rangle = \\ & \langle \tilde{w}^k - w^k, Hd(w^k, \tilde{w}^k) \rangle + \langle w^k - w^*, Hd(w^k, \tilde{w}^k) \rangle \end{aligned}$$

即

$$\begin{aligned} & \langle w^k - w^*, Hd(w^k, \tilde{w}^k) \rangle \geq \\ & \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle - \langle \tilde{w}^k - w^k, Hd(w^k, \tilde{w}^k) \rangle \quad (25) \end{aligned}$$

令

$$\Gamma(w^k, \tilde{w}^k) = \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle - \langle \tilde{w}^k - w^k, Hd(w^k, \tilde{w}^k) \rangle$$

证毕。

为将生成的预测序列 $\tilde{w}^k$ 进行校正, 定义一组关于步长 α 的新方程,

$$G(\alpha) = \|w^k - w^*\|_H^2 - \|w^{k+1}(\alpha) - w^*\|_H^2 \quad (26)$$

$$w^{k+1}(\alpha) = w^k - \alpha d(w^k, \tilde{w}^k) \quad (27)$$

将式(19), (25), (27)代入式(26)中, 有

$$\begin{aligned} G(\alpha) &= \|w^k - w^*\|_H^2 - \|w^k - \alpha d(w^k, \tilde{w}^k) - w^*\|_H^2 = \\ & 2\alpha \langle w^k - w^*, Hd(w^k, \tilde{w}^k) \rangle - \alpha^2 \|d(w^k, \tilde{w}^k)\|_H^2 \geq \\ & 2\alpha \Gamma(w^k, \tilde{w}^k) - \alpha^2 \|d(w^k, \tilde{w}^k)\|_H^2 = g(\alpha) \end{aligned}$$

显然, 方程 $g(\alpha) = 2\alpha \Gamma(w^k, \tilde{w}^k) - \alpha^2 \|d(w^k, \tilde{w}^k)\|_H^2$

是一个关于 α 的二次函数, 易知 $g(\alpha)$ 取最大值时的 α 的值为

$$\alpha_k := \frac{\Gamma(w^k, \tilde{w}^k)}{\|d(w^k, \tilde{w}^k)\|_H^2} \quad (28)$$

事实证明, 对于任意松弛因子 $t > 0$, 有

$$\begin{aligned} G(t\alpha_k) &\geq 2t\alpha_k \Gamma(w^k, \tilde{w}^k) - t^2 \alpha_k^2 \|d(w^k, \tilde{w}^k)\|_H^2 = \\ & t(2-t)\alpha_k \Gamma(w^k, \tilde{w}^k) \end{aligned}$$

为确保每次迭代都能得到改进, 建议选择 $t \in [1, 2)$ 以便得到更好的收敛结果。

引理 2 假设 $\rho \geq \|A\|^2$, 则有式(28)定义的步长 $\alpha_{\min} > 0$ 。

证明 引用柯西不等式的形式如下:

$$2\langle a, b \rangle \leq \varepsilon \|a\|^2 + \frac{1}{\varepsilon} \|b\|^2, \quad a, b \in \mathbb{R}^n, \quad \varepsilon > 0$$

同理, 可得

$$\langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle \leq \frac{\gamma}{2} \|\tilde{\lambda}^k - \lambda^k\|^2 + \frac{1}{2\gamma} \|\tilde{x}^k - x^k\|^2$$

同样地, 将式(17)代入式(19), 有

$$\begin{aligned} \Gamma(w^k, \tilde{w}^k) &= \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle - \langle \tilde{w}^k - w^k, Hd(w^k, \tilde{w}^k) \rangle = \\ & \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle + \|\tilde{w}^k - w^k\|_H^2 - \langle x^k - \tilde{x}^k, \eta_1(x, \tilde{x}) \rangle = \\ & \langle \tilde{\lambda}^k - \lambda^k, \tilde{x}^k - x^k \rangle + (\rho + \gamma) \|\tilde{x}^k - x^k\|^2 + \rho \|\tilde{z}^k - z^k\|^2 + \\ & \frac{1}{\gamma} \|\tilde{\lambda}^k - \lambda^k\|^2 - \langle x^k - \tilde{x}^k, \eta_1(x, \tilde{x}) \rangle \geq \\ & \left(\rho + \gamma - \frac{1}{2\gamma} \right) \|\tilde{x}^k - x^k\|^2 + \rho \|\tilde{z}^k - z^k\|^2 + \\ & \left(\frac{1}{\gamma} - \frac{\gamma}{2} \right) \|\tilde{\lambda}^k - \lambda^k\|^2 \quad (29) \end{aligned}$$

显然, $\Gamma(w^k, \tilde{w}^k) \geq C_{\min} \|w^k - \tilde{w}^k\|^2$, 其中, $C_{\min} :=$

$$\min \left\{ \rho + \gamma - \frac{1}{2\gamma}, \rho, \frac{1}{\gamma} - \frac{\gamma}{2} \right\}.$$

另一方面, 再次运用柯西不等式

$$\begin{aligned} \|d(w^k, \tilde{w}^k)\|_H^2 &= \|w^k - \tilde{w}^k\|_H^2 - 2\langle x^k - \tilde{x}^k, \eta_1(x, \tilde{x}) \rangle + \\ & \frac{1}{\rho + \gamma} \|\eta_1(x, \tilde{x})\|^2 \leq \left(\rho + \gamma + \frac{\|A\|^2}{\rho + \gamma} \right) \|\tilde{x}^k - x^k\|^2 + \\ & \rho \|\tilde{z}^k - z^k\|^2 + \frac{1}{\gamma} \|\tilde{\lambda}^k - \lambda^k\|^2 \quad (30) \end{aligned}$$

显然, $\|d(w^k, \tilde{w}^k)\|_H^2 \leq C_{\max} \|w^k - \tilde{w}^k\|^2$, 其中,

$$C_{\max} := \max \left\{ \rho + \gamma + \frac{\|A\|^2}{\rho + \gamma}, \rho, \frac{1}{\gamma} \right\}. \text{ 因此, 有 } \alpha_k \geq \frac{C_{\min}}{C_{\max}} =:$$

$\alpha_{\min} > 0$ 。

证毕。

定理 1 令 w^* 为式(20)的任意一个可行解, 算法产生的序列 $\{w^k\}$ 必须满足以下属性:

$$\begin{aligned} \|w^{k+1}-w^*\|_H^2 &\leq \|w^k-w^*\|_H^2- \\ &t(2-t)\alpha_{\min}C_{\min}\|w^k-\tilde{w}^k\|^2 \end{aligned} \tag{31}$$

证明 综合式(27)~(29)可得

$$\begin{aligned} \|w^{k+1}-w^*\|_M^2 &= \|w^k-t\alpha_k d(w^k,\tilde{w}^k)-w^*\|_H^2 = \\ &\|w^k-w^*\|_H^2-2t\alpha_k\langle w^k-w^*,Md(w^k,\tilde{w}^k)\rangle + \\ &t^2\alpha_k^2\|d(w^k,\tilde{w}^k)\|_H^2 \leq \|w^k-w^*\|_H^2- \\ &2t\alpha_k\Gamma(w^k,\tilde{w}^k)+t^2\alpha_k^2\|d(w^k,\tilde{w}^k)\|_H^2 = \\ &\|w^k-w^*\|_H^2-t(2-t)\alpha_k\Gamma(w^k,\tilde{w}^k) \leq \\ &\|w^k-w^*\|_H^2-t(2-t)\alpha_{\min}C_{\min}\|w^k-\tilde{w}^k\|^2 \end{aligned}$$

证毕。

定理 2 算法产生的序列必须全局收敛于变分不等式(20)的解。

证明 令 w^* 为式(20)的任意一个可行解, 根据式(31)可以得出结论:

$$\|w^{k+1}-w^*\|_H^2 \leq \|w^k-w^*\|_H^2 \leq \dots \leq \|w^0-w^*\|_H^2 \tag{32}$$

也就是说, $\|w^k-w^*\|_H^2$ 是单调减的有界序列, 即

$$\lim_{k \rightarrow \infty} \|w^k-w^*\|_H^2 \text{ 存在} \tag{33}$$

变换式(31)可得,

$$\begin{aligned} (2-t)\alpha_{\min}C_{\min}\|w^k-\tilde{w}^k\|^2 &\leq \\ \|w^k-w^*\|_H^2 - \|w^{k+1}-w^*\|_H^2 \end{aligned}$$

可得结论

$$\lim_{k \rightarrow \infty} \|w^k-w^*\|_H^2 = 0 \tag{34}$$

这证明了序列 $\{w^k\}$ 全局收敛于 $\{w^\infty\}$ 。

5 数值实验

通过图像恢复问题验证算法的效果, 所有代码均由 Matlab 2017a 编写, 处理器为 Intel Core i5 2.3 GHz, 内存 8 GB, 并且在 macOS 10.15 上运行。

由于图像恢复为问题(4)的特殊情况, 此问题的模型可以表示为

$$\min_x \left\{ \sigma \|Wx\|_1 + \frac{1}{2} \|Ax-b\|^2, x \in C \right\} \tag{35}$$

其中, b 为观测到的图像, 其经过高斯噪声模糊化处理。 $A=BW$ 包含 B 模糊算子和 W 余弦离散变换算子。实例中描述的所有原始图像的所有像素首先被缩放到 0~1 之间。现采用 4 种不同的图片来展示算法。图 1 分别为 house, pepper, camera 和 bridge 的原图与模糊图像。



图 1 原图和模糊图像
Fig.1 Original images and blurred images

随后采用一种常用的信噪比(SNR)概念来观测恢复图像的质量。

$$SNR := 20 \lg \frac{\|x\|}{\|\tilde{x}-x\|} \tag{36}$$

式中: x 表示原本的图像; $\tilde{x}$ 表示恢复后的图像。

SNR 的数值越大, 图像恢复的效果越好。一般使用式

$$\frac{\|x^{k+1}-x^k\|}{\|x^k\|} \leqslant 10^{-4} \quad (37)$$

来终止实验。

实验中设置 $\gamma=0.1$, $\rho=1.05\|A\|^2$, $\sigma=2\times10^{-4}$, $\omega=1.7$, $t=1.9$ 来进行验证, 同时与 ISM^[27](implementable splitting method)进行比较。如图 2 所示, 分别表示在 2 种算法下, 4 组实验的 SNR 和迭代次数的情况。

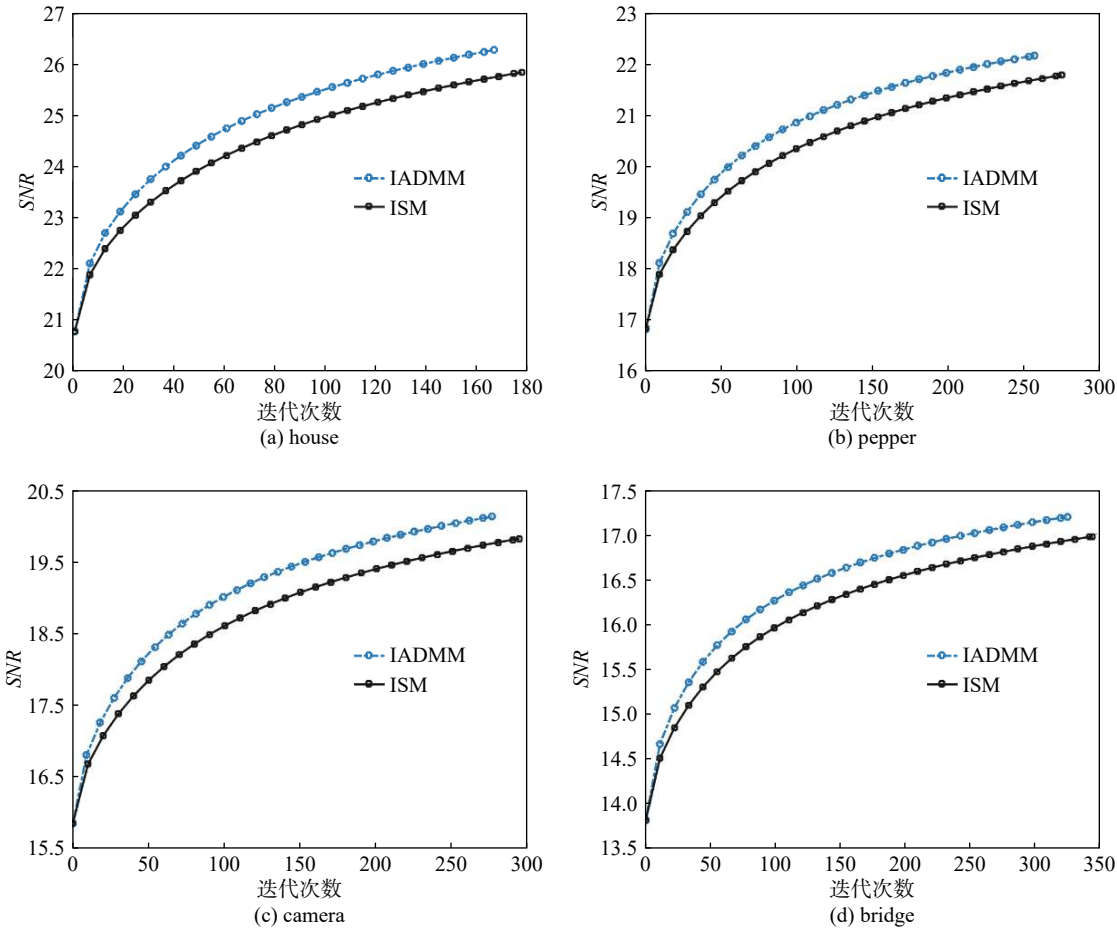


图 2 ISM 与 IADMM 算法的试验比较

Fig.2 Experimental comparison of ISM and IADMM algorithms

经过实验比较可知, 改进后的 IADMM 算法无论是在信噪比 (SNR) 上, 还是迭代次数上都有明显的优势。如图 3 所示, 经 IADMM 算法运算后恢复的图像, 相较于被噪声损坏的模糊图像,

算法处理过的图像已剔除了绝大部分噪声对图像清晰度的影响, 模糊的细节变得清晰可见。由此可见, 本算法在去模糊化的过程中得到了良好的体现。

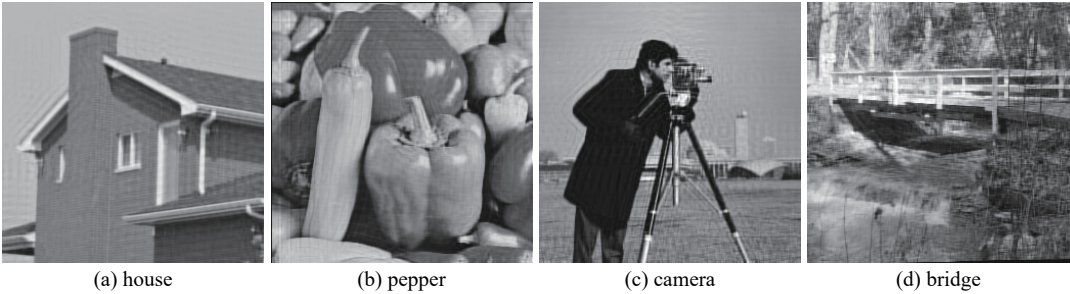


图 3 恢复后的图像

Fig.3 Restored images

同时，在 $t = 1.9$ 的条件下对不同的终止准则分别进行了实验，针对 IADMM 算法设置了 4 个不同的公差数值，分别为 $tol = \{10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times$

$10^{-5}\}$ ，并在表 1 中列出了相应的结果。分析表 1 可知，综合考虑迭代时间和重建质量，当 tol 取值为 10^{-4} 时，为本算法最优终止准则。

表 1 不同 tol 值下 IADMM 算法的实验结果
Tab.1 Experimental results of the IADMM algorithm under different tol

图片	$tol=10^{-3}$			$tol=5 \times 10^{-4}$			$tol=10^{-4}$			$tol=5 \times 10^{-5}$		
	迭代次数	迭代时间	SNR	迭代次数	迭代时间	SNR	迭代次数	迭代时间	SNR	迭代次数	迭代时间	SNR
house	18	0.93	23.05	39	1.55	24.06	167	6.27	26.26	293	10.70	27.06
pepper	28	1.12	19.10	63	2.35	20.18	257	9.86	22.16	437	15.60	22.79
camera	30	1.26	17.66	68	2.99	18.55	277	10.60	20.12	478	17.15	20.64
bridge	36	1.42	15.39	79	3.59	16.06	325	12.01	17.19	567	19.85	17.56

将 2 个算法进行比较，可以看出，IADMM 算法比原始的 ISM 算法花费更少的迭代次数的同时获得更高的信噪比，这验证了本文提出的改进算法对迭代方面性能的提高。本文的实验验证了 IADMM 算法的有效性和可靠性。

6 结 论

由于 ℓ_1 范数正则化问题不可微的性质，很多传统的迭代方法无法直接应用于 ℓ_1 正则化的模型。本文对子问题 x 和对偶算子 λ 引入了新的参数，在取值恰当的情况下，提高了算法的收敛性，并减少了迭代次数。数值实验的结果也验证了算法的有效性和可靠性。

参考文献:

[1] CENSOR Y, ELFVING T. A multiprojection algorithm using Bregman projections in a product space[J]. *Numerical Algorithms*, 1994, 8(2): 221–239.

[2] EICKE B. Iteration methods for convexly constrained ill-posed problems in Hilbert space[J]. *Numerical Functional Analysis and Optimization*, 1992, 13(5/6): 413–429.

[3] SABHARWAL A, POTTER L C. Convexly constrained linear inverse problems: iterative least-squares and regularization[J]. *IEEE Transactions on Signal Processing*, 1998, 46(9): 2345–2352.

[4] POTTER L C, ARUN K S. A dual approach to linear inverse problems with convex constraints[J]. *SIAM Journal on Control and Optimization*, 1993, 31(4): 1080–1092.

[5] ENGL H W, HANKE M, NEUBAUER A. Regularization of inverse problems[M]. Dordrecht: Kluwer, 1996.

[6] FRICK K, GRASMAIR M. Regularization of linear ill-posed problems by the augmented Lagrangian method and variational inequalities[J]. *Inverse Problems*, 2012, 28(10): 104005.

[7] HOCHSTENBACH M E, REICHEL L. Fractional Tikhonov regularization for linear discrete ill-posed problems[J]. *BIT Numerical Mathematics*, 2011, 51(1): 197–215.

[8] SCHÖPFER F, LOUIS A K, SCHUSTER T. Nonlinear iterative methods for linear ill-posed problems in Banach spaces[J]. *Inverse Problems*, 2006, 22(1): 311–329.

[9] CENG L C, ANSARI Q H, YAO J C. Relaxed extragradient methods for finding minimum-norm solutions of the split feasibility problem[J]. *Nonlinear Analysis: Theory, Methods & Applications*, 2012, 75(4): 2116–2125.

[10] CENSOR Y, GIBALI A, REICH S. Algorithms for the split variational inequality problem[J]. *Numerical Algorithms*, 2012, 59(2): 301–323.

[11] XU H K. Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces[J]. *Inverse Problems*, 2010, 26(10): 105018.

[12] YAO Y H, WU J G, LIOU Y C. Regularized methods for the split feasibility problem[J]. *Abstract and Applied Analysis*, 2012, 2012: 140679.

[13] BOYD S, PARIKH N, CHU E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. *Foundations and Trends in Machine Learning*, 2011, 3(1): 1–122.

[14] YANG J F, ZHANG Y. Alternating direction algorithms for ℓ_1 -problems in compressive sensing[J]. *SIAM Journal on Scientific Computing*, 2011, 33(1): 250–278.

[15] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit[J]. *SIAM Review*, 2001, 43(1): 129–159.

[16] DONOHO D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289–1306.

(下转第 503 页)