DOI:10.13718/j.cnki.xsxb.2020.09.003

Douglas-Rachford 分裂法线性收敛性的新证明[①]

陶永凯，　彭建文

重庆师范大学 数学与科学学院，重庆 401331

摘要：结合临近算子和极大单调算子的关系，利用凸优化和压缩算子理论，证明了无约束优化问题的 Douglas-Rachford 分裂法对强凸且光滑的函数具有全局线性收敛性，并给出了相应的收敛率.

关 键 词：Douglas-Rachford 分裂法；临近算子；压缩算子理论；全局收敛性

中图分类号：**O221.2**　　**文献标志码**：**A**　　**文章编号**：1000-5471(2020)09-0013-06

考虑下列无约束优化问题

$$\min_{x\in H} f(\boldsymbol{x})+g(\boldsymbol{x}) \tag{1}$$

其中：H 是实的无穷维希尔伯特空间；f,g：$H\longrightarrow \mathbb{R}\cup\{+\infty\}$ 为闭、真凸函数. 对于优化问题(1)，现在已有多种求解方法[1-5]. 文献[5]第一次提出寻求单调算子和为零的分裂方法(Douglas-Rachford 分裂法)求解如下形式的变分包含问题

$$0\in A\boldsymbol{x}+B\boldsymbol{x} \tag{2}$$

这里 A,B：$H\longrightarrow 2^H$ 是两个极大单调算子. 因为闭的凸函数的次微分是极大单调的[6]，根据一阶最优性条件，求解优化问题(1) 等价于寻找 $\hat{\boldsymbol{x}}\in H$ 使得 $0\in\partial f(\hat{\boldsymbol{x}})+\partial g(\hat{\boldsymbol{x}})$. 近年来，求解优化问题(1) 的 Douglas-Rachford 分裂算法被广泛运用于诸多领域，例如信号处理[7]、压缩感知[8]、图像去噪[9] 和平衡问题[10] 等方面.

对于问题(2) 的求解，常常是通过算子 A,B 相对应的预解式 J_A,J_B 进行分离的，算子 A 预解式 J_A：$H\longrightarrow H$ 定义为

$$J_A:=(A+Id)^{-1} \tag{3}$$

其中 Id 为恒等算子，若 f：$H\longrightarrow\mathbb{R}\cup\{+\infty\}$ 是凸函数，当 $A=\partial f$ 时，有 $J_A=prox_f$，这里 $prox_f$ 为临近算子[11]，其定义如下

$$prox_f(\boldsymbol{z})=\operatorname*{argmin}_{\boldsymbol{x}}\left(f(\boldsymbol{x})+\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{z}\|^2\right),\ \forall\boldsymbol{z}\in H \tag{4}$$

求解优化问题(1) 的 Douglas-Rachford 分裂算法的迭代公式为

$$\boldsymbol{z}^{k+1}=((1-\alpha)Id+\alpha R_g R_f)\boldsymbol{z}^k \tag{5}$$

其中 $\alpha\in(0,2)$ 通常称为松弛因子，R_f 称为反射临近算子，其定义为

$$R_f=2prox_f-Id \tag{6}$$

对于凸优化问题(1)，可以等价转化成变分包含问题(2). 文献[5] 证明了若 A,B 都为极大单调算子，则 Douglas-Rachford 分裂算法产生的序列弱收敛于最优点；文献[12] 证明了若 A,B 中有一个算子既是强单调的又是 Lipschitz 连续的，则 Douglas-Rachford 分裂算法具有收敛性，并给出了其收敛速度为 $O(\sqrt{k})$；

① 收稿日期：2019-03-01
作者简介：陶永凯(1994—)，男，硕士，主要从事优化算法研究.
通信作者：彭建文，教授.

文献[13]利用共轭函数和辅助函数的关系证明了当问题(1)中的一个函数同时具有强凸性和光滑性时，Duglas-Rachford 分裂算法产生的迭代序列具有线性收敛性；紧接着，文献[14]给出了 3 种情形下 Douglas-Rachford 分裂法具有全局线性收敛率.

本文针对问题(1)，若 $f: H \longrightarrow \mathbb{R} \cup \{+\infty\}$ 为光滑强凸函数时，给出了 Douglas-Rachford 分裂方法的线性收敛性的新证明，并得到不同于文献[13]的线性收敛结果.

1 预备知识

全文结论建立在实的希尔伯特空间(记作 H)上. 对 H 中任意两点 $\boldsymbol{x}, \boldsymbol{y}$，其内积用 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ 表示，且由内积诱导出的范数定义为 $\| \boldsymbol{x} \| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. 设 X 是 H 的一个子集，集合 X 的子集的全体为 2^X，F 为 X 上的集值映射 $F: X \longrightarrow 2^X$. 函数 $f: H \longrightarrow \mathbb{R} \cup \{+\infty\}$ 在 $\boldsymbol{x} \in H$ 处的次微分用 $\partial f(\boldsymbol{x})$ 表示，定义 $\partial f(\boldsymbol{x}) = \{\boldsymbol{u} \mid f(\boldsymbol{y}) \geqslant f(\boldsymbol{x}) + \langle \boldsymbol{u}, \boldsymbol{y} - \boldsymbol{x} \rangle, \forall \boldsymbol{y} \in H\}$.

定义 1 (强单调算子) 称算子 $A: H \longrightarrow 2^H$ 为 σ - 强单调的，如果 $\exists \sigma > 0$，$\forall \boldsymbol{x}, \boldsymbol{y} \in H$，$\forall \boldsymbol{u} \in A\boldsymbol{x}$，$\boldsymbol{v} \in A\boldsymbol{y}$ 都有

$$\langle \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{x} - \boldsymbol{y} \rangle \geqslant \sigma \| \boldsymbol{x} - \boldsymbol{y} \|^2 \tag{7}$$

注 1 当 $\sigma = 0$ 时，称算子 A 是单调的.

定义 2 (强制算子) 如果 $\exists \beta \geqslant 0$，$\forall \boldsymbol{x}, \boldsymbol{y} \in H$ 都有

$$\langle \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{x} - \boldsymbol{y} \rangle \geqslant \beta \| \boldsymbol{u} - \boldsymbol{v} \|^2, \ \forall \boldsymbol{u} \in A\boldsymbol{x}, \boldsymbol{v} \in A\boldsymbol{y} \tag{8}$$

则称算子 $A: H \longrightarrow 2^H$ 为 β - 强制的.

定义 3 (Lipschitz 连续) 如果存在常数 $L \geqslant 0$，使得

$$\| A\boldsymbol{x} - A\boldsymbol{y} \| \leqslant L \| \boldsymbol{x} - \boldsymbol{y} \| \tag{9}$$

则称算子 $A: H \longrightarrow H$ 为 L-Lipschitz 连续的. 称 L 为 Lipschitz 常数. 当 $L = 1$ 时，称算子 A 为非扩张的；当 $L \in [0, 1)$ 时，称算子 A 为 L- 压缩的.

定义 4 如果 $\exists \sigma > 0$，对于 $\forall \boldsymbol{u} \in \partial f$ 都有

$$f(\boldsymbol{y}) \geqslant f(\boldsymbol{x}) + \langle \boldsymbol{u}, \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{\sigma}{2} \| \boldsymbol{x} - \boldsymbol{y} \|^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in H \tag{10}$$

且当 f 为连续可微函数时有 $\partial f = \{\nabla f(\boldsymbol{x})\}$，其中 $\nabla f(\boldsymbol{x})$ 为 f 在 $\boldsymbol{x}$ 处的梯度，则称函数 $f: H \longrightarrow \mathbb{R} \cup \{+\infty\}$ 是 σ - 强凸函数.

显然，f 是强凸函数等价于

$$\langle \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{x} - \boldsymbol{y} \rangle \geqslant \sigma \| \boldsymbol{x} - \boldsymbol{y} \|^2 \tag{11}$$

其中 $\forall \boldsymbol{u} \in \partial f(\boldsymbol{x})$，$\forall \boldsymbol{v} \in \partial f(\boldsymbol{y})$. 由定义 1 可知，$f$ 是 σ - 强凸函数的充要条件为 ∂f 是 σ - 强单调算子.

定义 5 如果对于 $\forall \boldsymbol{x}, \boldsymbol{y} \in H$，$\beta > 0$ 都有

$$f(\boldsymbol{y}) \leqslant f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{x} - \boldsymbol{y} \rangle + \frac{\beta}{2} \| \boldsymbol{x} - \boldsymbol{y} \|^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in H \tag{12}$$

其中 $\nabla f(\boldsymbol{x})$ 为 f 在 $\boldsymbol{x}$ 处的梯度，则称函数 f 是 β - 光滑凸函数.

对于不等式(12)，交换 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的位置，将两个不等式相加得到 f 是 β - 光滑函数的另一种表达形式

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \leqslant \beta \| \boldsymbol{x} - \boldsymbol{y} \|^2 \tag{13}$$

引理 1[15] 若 $f, g: H \longrightarrow \mathbb{R} \cup \{+\infty\}$ 是连续函数，f 是 σ - 强凸函数当且仅当 $f - \frac{\sigma}{2} \| \cdot \|^2$ 是凸函数；g 是 β - 光滑函数当且仅当 $\frac{\beta}{2} \| \cdot \|^2 - g$ 是凸函数.

引理 2[16] 设 f 是 H 中的实、闭、真凸函数，下列命题等价：

1) f 是 β - 光滑函数；

2) ∇f 是 β -Lipschitz 连续的；

3) ∇f 是 $\frac{1}{\beta}$ 强制的.

引理 3[16]　若算子 $A: H \longrightarrow 2^H$ 是极大单调的，那么算子 $J_A, Id - J_A: H \longrightarrow H$ 是严格非扩张极大单调的；反射预解式 $R_A: \boldsymbol{x} \longrightarrow 2J_A\boldsymbol{x} - \boldsymbol{x}$ 是非扩张的.

引理 4[16]　若算子 $A: H \longrightarrow 2^H$ 是 σ - 强单调的，则预解式 J_A 是 $(1+\sigma)$ - 强制的.

2　Douglas-Rachford 分裂法

求解优化问题(1) 等价寻找最优点 $\hat{\boldsymbol{x}}$ 使得 $0 \in \partial f(\hat{\boldsymbol{x}}) + \partial g(\hat{\boldsymbol{x}})$，因此求解优化问题(1) 转化为求解变分包含问题(2). 利用Douglas-Rachford分裂法能解决问题(2)，任意给定一个初始点 $\boldsymbol{z}^0$，Douglas-Rachford分裂算法如下

$$\begin{cases} \boldsymbol{x}_f^k = prox_f(\boldsymbol{z}^k) \\ \boldsymbol{x}_g^k = prox_g(2\boldsymbol{x}_f^k - \boldsymbol{z}^k) \\ \boldsymbol{z}^{k+1} = \boldsymbol{z}^k + 2\alpha(\boldsymbol{x}_g^k - \boldsymbol{x}_f^k) \end{cases} \tag{14}$$

其中 $prox_f$ 为临近算子，其定义见(4) 式. 将(14) 式的 3 式相加，利用 R_f 的定义得到如下的迭代结果

$$\boldsymbol{z}^{k+1} = (1-\alpha)\boldsymbol{z}^k + \alpha R_g R_f \boldsymbol{z}^k \tag{15}$$

其中 $\alpha \in (0, 2)$，当 $\alpha = 1$ 时，即有 $\boldsymbol{z}^{k+1} = R_g R_f \boldsymbol{z}^k$，此时产的算法称之为 Peaceman-Rachford 分裂法[2]. 由引理 3 知，R_g, R_f 是非扩张的，因此它们的复合 $R_g R_f$ 也是非扩张的，故在一般情形下 Peaceman-Rachford 分裂算法产生的点列不具有收敛性. 当 $\alpha = \dfrac{1}{2}$ 时，式(15) 退化为 $\boldsymbol{z}^{k+1} = \dfrac{1}{2}(Id + R_g R_f)\boldsymbol{z}^k$，这就是文献[17]中提及的 Douglas-Rachford 分裂算法. 在本文中引入松弛因子 α，相比较于 Peaceman-Rachford 分裂法和特定的 Douglas-Rachford 分裂法，具有良好的收敛性.

3　主要结果

定理 1　假设 f 是 σ - 强凸函数、β - 光滑函数，且满足 $\beta + \dfrac{1}{\sigma} \geqslant 2$，则算法(15) 迭代产生的 R_f 是 $\sqrt{\dfrac{\beta\sigma - 2\sigma + 1}{\beta\sigma + 2\sigma + 1}}$ - 压缩的.

证　因 f 是 σ - 强凸函数，由式(3) 和引理 4 知 $prox_f$ 是 $(1+\sigma)$ - 强制的，令 $T = prox_f$，由定义 2 可知，对 $\forall\, \boldsymbol{x}, \boldsymbol{y} \in H$

$$\langle T\boldsymbol{x} - T\boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle \geqslant (\sigma + 1)\, \| T\boldsymbol{x} - T\boldsymbol{y} \|^2 \tag{16}$$

对式(16) 两边同乘以 2 并减去 $\| \boldsymbol{x} - \boldsymbol{y} \|^2$ 得

$$\langle 2T\boldsymbol{x} - \boldsymbol{x} - (2T\boldsymbol{y} - \boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y}\rangle \geqslant \frac{(\sigma+1)}{2}\, \| 2T\boldsymbol{x} - \boldsymbol{x} - (2T\boldsymbol{y} - \boldsymbol{y}) + \boldsymbol{x} - \boldsymbol{y} \|^2 - \| \boldsymbol{x} - \boldsymbol{y} \|^2 \tag{17}$$

由式(6) 知 $R_f = 2prox_f - Id = 2T - Id$，整理式(17) 可得与之等价的表达

$$\langle R_f\boldsymbol{x} - R_f\boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle \geqslant \frac{\sigma+1}{2}\, \| R_f\boldsymbol{x} - R_f\boldsymbol{y} + \boldsymbol{x} - \boldsymbol{y} \|^2 - \| \boldsymbol{x} - \boldsymbol{y} \|^2 \tag{18}$$

对式(18) 两边同乘以 $\dfrac{2}{\sigma+1}$，并将不等式右侧第一项展开，整理得

$$-\frac{2\sigma}{\sigma+1}\langle R_f\boldsymbol{x} - R_f\boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle \geqslant \| R_f\boldsymbol{x} - R_f\boldsymbol{y} \|^2 + \frac{\sigma-1}{\sigma+1}\, \| \boldsymbol{x} - \boldsymbol{y} \|^2 \tag{19}$$

又因 f 是 β - 光滑函数，由引理 2 可知 ∇f 是 $\dfrac{1}{\beta}$ - 强制的. 则对 $\boldsymbol{u}, \boldsymbol{v} \in H$ 有

$$\langle \nabla f(\boldsymbol{u}) - \nabla f(\boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v}\rangle \geqslant \frac{1}{\beta}\, \| \nabla f(\boldsymbol{u}) - \nabla f(\boldsymbol{v}) \|^2 \tag{20}$$

对式(20) 两边同时加 $\| \boldsymbol{u} - \boldsymbol{v} \|^2$，整理得

$$\langle \nabla f(\boldsymbol{u}) + \boldsymbol{u} - (\nabla f(\boldsymbol{v}) + \boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v}\rangle \geqslant \frac{1}{\beta}\, \| \nabla f(\boldsymbol{u}) + \boldsymbol{u} - (\nabla f(\boldsymbol{v}) + \boldsymbol{v}) - (\boldsymbol{u} - \boldsymbol{v}) \|^2 + \| \boldsymbol{u} - \boldsymbol{v} \|^2 \tag{21}$$

对式(21)，令 $\boldsymbol{x} = (\nabla f + Id)\boldsymbol{u}$，$\boldsymbol{y} = (\nabla f + Id)\boldsymbol{v}$，由式(3)，式(4) 可知：$\boldsymbol{x} = (\nabla f + Id)\boldsymbol{u}$ 当且仅当 $\boldsymbol{u} =$

$(\nabla f+Id)^{-1}\boldsymbol{x}=prox_f\boldsymbol{x}$；$\boldsymbol{y}=(\nabla f+Id)\boldsymbol{v}$ 当且仅当 $\boldsymbol{v}=(\nabla f+Id)^{-1}\boldsymbol{y}=prox_f\boldsymbol{y}$. 则式(21)化为

$$(\beta+2)\langle T\boldsymbol{x}-T\boldsymbol{y},\boldsymbol{x}-\boldsymbol{y}\rangle \geqslant \|\boldsymbol{x}-\boldsymbol{y}\|^2+(\beta+1)\|T\boldsymbol{x}-T\boldsymbol{y}\|^2 \tag{22}$$

对式(22)两侧同乘以2再同时减去 $\|\boldsymbol{x}-\boldsymbol{y}\|^2$，整理得

$$(\beta+2)\langle R_f\boldsymbol{x}-R_f\boldsymbol{y},\boldsymbol{x}-\boldsymbol{y}\rangle \geqslant \frac{\beta+1}{2}\|R_f\boldsymbol{x}-R_f\boldsymbol{y}+(\boldsymbol{x}-\boldsymbol{y})\|^2-\beta\|\boldsymbol{x}-\boldsymbol{y}\|^2$$

将不等式右侧第一项展开，整理得

$$2\langle R_f\boldsymbol{x}-R_f\boldsymbol{y},\boldsymbol{x}-\boldsymbol{y}\rangle \geqslant (1+\beta)\|R_f\boldsymbol{x}-R_f\boldsymbol{y}\|^2+(1-\beta)\|\boldsymbol{x}-\boldsymbol{y}\|^2 \tag{23}$$

将式(19)加上式(23)再乘以 $\frac{\sigma}{\sigma+1}$ 得

$$\|R_f\boldsymbol{x}-R_f\boldsymbol{y}\|^2 \leqslant \frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}\|\boldsymbol{x}-\boldsymbol{y}\|^2$$

当 $\beta\sigma-2\sigma+1\geqslant 0$，即 $\beta+\frac{1}{\sigma}\geqslant 2$ 时，有 $0\leqslant\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}<1$，则有

$$\|R_f\boldsymbol{x}-R_f\boldsymbol{y}\| \leqslant \sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\|\boldsymbol{x}-\boldsymbol{y}\|$$

即 R_f 是压缩的.

定理 2　假设 f 是 σ-强凸函数、β-光滑函数，$\alpha\in\left(0,1+\frac{\beta\sigma+1}{2\sigma}-\sqrt{\left(\frac{\beta\sigma+1}{2\sigma}\right)^2-1}\right)$，则 Douglas-Rachford 分裂算法(15)线性收敛到不动点 $\hat{\boldsymbol{z}}$，且收敛率至少为 $|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}$. 即有

$$\|\boldsymbol{z}^{k+1}-\hat{\boldsymbol{z}}\| \leqslant \left(|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\right)\|\boldsymbol{z}^k-\hat{\boldsymbol{z}}\| \tag{24}$$

证　由引理3和定理1可知 R_g 是非扩张的，R_f 是 $\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}$ 压缩的，则 R_gR_f 也是 $\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}$ 压缩的. 因为对 $\forall\boldsymbol{z}_1,\boldsymbol{z}_2\in H$，$\|R_gR_f\boldsymbol{z}_1-R_gR_f\boldsymbol{z}_2\|\leqslant\|R_f\boldsymbol{z}_1-R_f\boldsymbol{z}_2\|\leqslant\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\|\boldsymbol{z}_1-\boldsymbol{z}_2\|$. 令 $T=(1-\alpha)Id+\alpha R_gR_f$，则式(15)可化为 $\boldsymbol{z}^{k+1}=T\boldsymbol{z}^k$，因为 $\hat{\boldsymbol{z}}$ 是 Douglas-Rachford 分裂算法(15)的不动点，故有 $\hat{\boldsymbol{z}}=T\hat{\boldsymbol{z}}$

$$\begin{aligned}
&\|\boldsymbol{z}^{k+1}-\hat{\boldsymbol{z}}\| = \|T\boldsymbol{z}^k-T\hat{\boldsymbol{z}}\| = \\
&\|(1-\alpha)(\boldsymbol{z}^k-\hat{\boldsymbol{z}})+\alpha(R_gR_f\boldsymbol{z}^k-R_gR_f\hat{\boldsymbol{z}})\| \leqslant \\
&\|(1-\alpha)(\boldsymbol{z}^k-\hat{\boldsymbol{z}})\| + \|\alpha(R_gR_f\boldsymbol{z}^k-R_gR_f\hat{\boldsymbol{z}})\| = \\
&|1-\alpha|\,\|\boldsymbol{z}^k-\hat{\boldsymbol{z}}\| + \alpha\|R_gR_f\boldsymbol{z}^k-R_gR_f\hat{\boldsymbol{z}}\| \leqslant \\
&\left(|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\right)\|\boldsymbol{z}^k-\hat{\boldsymbol{z}}\|
\end{aligned}$$

其中第一个小于等于关系式是利用三角不等式，第二个小于等于关系式是利用 R_gR_f 的压缩性. 令 $|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}<1$，当 $\alpha\leqslant 1$ 时，显然 $|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}<1$ 恒成立；当 $\alpha>1$ 时，解得 $\alpha<1+\frac{\beta\sigma+1}{2\sigma}-\sqrt{\left(\frac{\beta\sigma+1}{2\sigma}\right)^2-1}$. 综上，当 $\alpha\in\left(0,1+\frac{\beta\sigma+1}{2\sigma}-\sqrt{\left(\frac{\beta\sigma+1}{2\sigma}\right)^2-1}\right)$ 时，Douglas-Rachford 分裂算法式(17)线性收敛率至少为 $|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}$.

注 2　文献[13]的定理1说明当 f 是 σ-强凸函数、β-光滑函数，$\gamma\in(0,+\infty)$ 时，算法(15)迭代产生的 R_f 是 $\max\left(\frac{\gamma\beta-1}{\gamma\beta+1},\frac{1-\gamma\sigma}{\gamma\sigma+1}\right)$ 压缩的，而本文的定理1说明了当 f 是 σ-强凸函数、β-光滑函数，且 $\beta+$

$\frac{1}{\sigma}\geqslant 2$ 时，算法(15) 迭代产生的 R_f 是 $\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}$- 压缩的.

注 3　文献[13] 的定理 2 与本文的定理 2 都说明了在适当条件下，Douglas-Rachford 分裂算法(15) 产生的序列$\{\boldsymbol{z}^k\}$ 是线性收敛的.

推论 1　若 f 是 σ- 强凸函数、β- 光滑函数，设 $\hat{\boldsymbol{x}}$ 是优化问题(1) 的解，则利用 Douglas-Rachford 分裂算法式(16) 产生的迭代序列$\{\boldsymbol{x}^k\}$ 满足

$$\|\boldsymbol{x}^{k+1}-\hat{\boldsymbol{x}}\|\leqslant\frac{1}{\sigma+1}\left(|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\right)^{k+1}\|\boldsymbol{z}^0-\hat{\boldsymbol{z}}\| \tag{25}$$

证　f 是 σ - 强凸函数，则 ∂f 是 σ - 强单调算子，可知 $prox_f$ 是$(1+\sigma)$- 强制的，由定义 2 及 Cauchy-Schwarz 不等式可得 $prox_f$ 是$\frac{1}{\sigma+1}$-Lipschitz 连续. 因 $\hat{\boldsymbol{x}}$ 是优化问题(1) 的解，那么有 $\hat{\boldsymbol{x}}=prox_f(\hat{\boldsymbol{z}})$，其中 $\hat{\boldsymbol{z}}$ 是 Douglas-Rachford 分裂算法(15) 的不动点. 因此，

$$\begin{aligned}\|\boldsymbol{x}^{k+1}-\hat{\boldsymbol{x}}\| &= \|prox_f(\boldsymbol{z}^{k+1})-prox_f(\hat{\boldsymbol{z}})\|\leqslant\\ &\frac{1}{\sigma+1}\|\boldsymbol{z}^{k+1}-\hat{\boldsymbol{z}}\|\leqslant\\ &\frac{1}{\sigma+1}\left(|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\right)\|\boldsymbol{z}^k-\hat{\boldsymbol{z}}\|\leqslant\\ &\frac{1}{\sigma+1}\left(|1-\alpha|+\alpha\sqrt{\frac{\beta\sigma-2\sigma+1}{\beta\sigma+2\sigma+1}}\right)^{k+1}\|\boldsymbol{z}^0-\hat{\boldsymbol{z}}\|\end{aligned}$$

第一个不等式是利用 $prox_f$ 是$\frac{1}{1+\sigma}$-Lipschitz 连续的，第二、三个不等式利用定理 2 的结论，则推论得证.

参考文献：

[1] COMBETTES P L，WAJS V R. Signal Recovery by Proximal Forward-Backward Splitting [J]. Multiscale Modeling Simulation，2005，4(4)：1168-1200.

[2] PEACEMAN D W，RACHFORD H H. The Numerical Solution of Parabolic and Elliptic Differential Equations [J]. Journal of the Society for Industrial and Applied Mathematics，1955，3(1)：28-41.

[3] COMBETTES P L，PESQUET J C. Proximal Splitting Methods in Signal Processing [M]. Berlin：Springer，2011：185-212.

[4] BOYD S. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers [J]. Foundations and Trends in Machine Learning，2010，3(1)：1-122.

[5] LIONS P L，MERCIER B. Splitting Algorithms for the Sum of Two Nonlinear Operators [J]. SIAM Journal on Numerical Analysis，1979，16(6)：964-979.

[6] ROCKAFELLAR R T. Monotone Operators and the Proximal Point Algorithm [J]. SIAM Journal on Control and Optimization，1976，14(5)：877-898.

[7] COMBETTES P L，PESQUET J C. A Douglas-Rachford Splitting Approach to Nonsmooth Convex Variational Signal Recovery [J]. IEEE Journal of Selected Topics in Signal Processing，2007，1(4)：564-574.

[8] LI S J，QI H R. A Douglas-Rachford Splitting Approach to Compressed Sensing Image Recovery Using Low-Rank Regularization [J]. IEEE Transactions on Image Processing，2015，24(11)：4240-4249.

[9] STEIDL G，TEUBER T. Removing Multiplicative Noise by Douglas-Rachford Splitting Methods [J]. Journal of Mathematical Imaging and Vision，2010，36(2)：168-184.

[10] BRICEÑO-ARIAS L M. A Douglas-Rachford Splitting Method for Solving Equilibrium Problems [J]. Nonlinear Analysis：Theory，Methods & Applications，2012，75(16)：6053-6059.

[11] PARIKH N. Proximal Algorithms [J]. Foundations and Trends® in Optimization，2014，1(3)：127-239.

[12] DAVIS D，YIN W T. Convergence Rate Analysis of Several Splitting Schemes [M]//Splitting Methods in Communication，Imaging，Science，and Engineering. Berlin：Springer，2016：115-163.

[13] GISELSSON P, BOYD S. Linear Convergence and Metric Selection for Douglas-Rachford Splitting and ADMM [J]. IEEE Transactions on Automatic Control, 2017, 62(2): 532-544.

[14] GISELSSON P. Tight Global Linear Convergence Rate Bounds for Douglas-Rachford Splitting [J]. Journal of Fixed Point Theory and Applications, 2017, 19(4): 2241-2270.

[15] GISELSSON P, BOYD S. Diagonal Scaling in Douglas-Rachford Splitting and ADMM [C]//53rd IEEE Conference on Decision and Control. Los Angeles: IEEE Press, 2014: 5033-5039.

[16] BAUSCHKE H H, COMBETTES P L. Convex Analysis and Monotone Operator Theory in Hilbert Spaces [M]. Berlin: Springer, 2011.

[17] ECKSTEIN J, BERTSEKAS D. On the Douglas—Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators [J]. Mathematical Programming, 1992, 55(1-3): 293-318.

A New Proof of Linear Convergence of the Douglas-Rachford Splitting Method

TAO Yong-kai, PENG Jian-wen

School of Mathematical Science, Chongqing Normal University, Chongqing 401331, China

Abstract: In this paper, the relationship between the proximity operator and the maximal monotone Operator has been combined, and the convex optimization and compression operator theory used to prove that the Douglas-Rachford splitting method with unconstrained optimization problems has global linear convergence under strong convex and smooth functions, and give the corresponding convergence rate.

Key words: Douglas-Rachford splitting method; proximity operator; compression operator theory; global convergence

责任编辑 张 枸