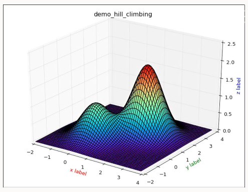
Monotone Operators and Base Splitting Schemes

Lin Wenjie

Nanjing Normal University

版本：1.0

更新：2024 年 6 月 4 日



以下整理了一些算子特别是单调算子的定义与性质，以及一些基本的分裂框架

1 集值算子 (Set-value Operators)

集值算子 (Set-value Operators) 是把一个集合上元素映射为另一个集合的子集的映射，即：

$$\begin{aligned} \mathbb{T} :& \mathbb{E} \to 2^{\mathbb{F}} \\ & x \to \mathbb{T}(x) \subset \mathbb{F} \end{aligned}$$

一般我们考虑的集值算子的形式如下：

$$\begin{aligned} \mathbb{T} :& \mathbb{R}^n \to 2^{\mathbb{R}^n} \\ & x \to \mathbb{T}(x) \subset \mathbb{R}^n \end{aligned}$$

定义 1.1 (定义域) 集值算子 $\mathbb{T}$ 的定义域定义为：

$$\mathrm{dom}\mathbb{T} = \{x | \mathbb{T}x \neq \emptyset\} \subset \mathbb{R}^n$$

定义 1.2 (图) 集值算子 $\mathbb{T}$ 的图定义为：

$$\mathrm{Gra}\mathbb{T} = \{(x, u) | u \in \mathbb{T}x\} \subset \mathbb{R}^n \times \mathbb{R}^n$$

定义 1.3 (L-lipschitz) 称算子 $\mathbb{T}$ 是 L-lipschitz 的，如果：

$$\|u, v\| \leq L\|x - y\|, \forall (x, u), (y, v) \in \mathbb{T}$$

或写成

$$\|\mathbb{T}x - \mathbb{T}y\| \leq L\|x - y\|, \forall x, y \in dom\{\mathbb{T}\}$$

注 可以通过反证得到，L-lipschitzs 算子都是单值算子

定义 1.4 (零点) 如果 $0 \in \mathbb{T}x$，则称 x 为 $\mathbb{T}$ 的一个零点，记为

$$\text{Zer}\mathbb{T} = \{x|0 \in \mathbb{T}x\} = \mathbb{T}^{-1}(0)$$

2 单调算子 (Monotone Operators)

定义 2.1 (单调算子) 称算子 $\mathbb{T}$ 是单调的，如果：

$$\langle u - v, x - y\rangle \geq 0, \forall (x, u), (y, v) \in \mathbb{T}$$

或写成

$$\langle \mathbb{T}x - \mathbb{T}y, x - y\rangle \geq 0, \forall x, y \in \text{dom}\{\mathbb{T}\}$$

定义 2.2 (极大单调算子) 称算子 $\mathbb{T}$ 是极大单调的，如果不存在一个单调算子 $\mathbb{S}$ 使得

$$\text{Gra}\mathbb{T} \subsetneq \text{Gra}\mathbb{S}$$

例 2.1 函数 f 是凸函数，则其次微分算子 ∂f 是单调算子; 又若 f 适当，则则其次微分算子 ∂f 是极大单调算子;

2.1 强单调与余强制

定义 2.3 (强单调) 称算子 $\mathbb{T}$ 是 $\mu-$ 强单调的，如果

$$\langle u - v, x - y\rangle \geq \mu\|x - y\|^2, \forall (x, u), (y, v) \in \mathbb{T}$$

或写成 $\langle \mathbb{T}x - \mathbb{T}y, x - y\rangle \geq \mu\|x - y\|^2, \forall x, y \in \text{dom}\mathbb{T}$

定义 2.4 (余强制) 称算子 $\mathbb{T}$ 是 $\mu-$ 余强制的，如果

$$\langle u - v, x - y\rangle \geq \mu\|u - v\|^2, \forall (x, u), (y, v) \in \mathbb{T}$$

或写成 $\langle \mathbb{T}x - \mathbb{T}y, x - y\rangle \geq \mu\|\mathbb{T}x - \mathbb{T}y\|^2, \forall x, y \in domT$

注 显然，强单调算子和余强制算子都是单调算子。

注 强单调和余强制是一组对偶性质:

$$\mathbb{T}\mu - \text{strongly monotone} \iff \mathbb{T}^{-1}\mu - \text{cocoercive}$$

注 如果 $\mathbb{T}\mu-$ 余强制，即 $\langle\mathbb{T}x-\mathbb{T}y, x-y\rangle \geq \mu\|\mathbb{T}x-\mathbb{T}y\|^2, \forall x, y \in domT$ 则根据 cauchy-swarch 不等式，有：

$$\|\mathbb{T}x-\mathbb{T}y\| \leq \frac{1}{\mu}\|x-y\|, \forall x, y \in domT$$

即余强制算子一定是 $L-lipschitz$ 算子，于是余强制算子也一定是单值算子。

注 同样可以定义极大 $\mu-$ 强单调算子与极大 $\mu-$ 余强制算子，这二者也互为对偶

$$\mathbb{T}\text{maximal}\mu-\text{strongly monotone} \iff \mathbb{T}^{-1}\text{maximal}\mu-\text{cocoercive}$$

注 如果函数 f 是适当闭凸函数，则：

$$f \quad L-smooth \iff \partial f \quad \frac{1}{L}-\text{余强制} \quad \text{文再文老师的书上有证明}$$

$$f\mu-\text{强凸} \iff \partial f\mu-\text{强单调} \iff \partial f^* = (\partial f)^{-1}\mu-\text{余强制} \iff f^*\frac{1}{\mu}-smooth$$

注

$$\begin{aligned}
&\mathbb{T}\mu-\text{强单调}, M \in \mathbb{R}^{n\times m}\text{列满秩} \Rightarrow M^\top\mathbb{T}M\mu\sigma^2_{min}(M)-\text{强单调} \\
pf: &\left\langle M^\top\mathbb{T}Mx - M^\top\mathbb{T}My, x-y\right\rangle = \langle\mathbb{T}Mx-\mathbb{T}My, M(x-y)\rangle \\
&= \langle\mathbb{T}M(x-y), M(x-y)\rangle \\
&\geq \mu\|M(x-y)\|^2 \\
&\geq \mu\sigma^2_{min}(M)\|x-y\|^2
\end{aligned}$$

注

$$\begin{aligned}
&\mathbb{T}\text{L-lipschitz}, M \in \mathbb{R}^{n\times m} \Rightarrow M^\top\mathbb{T}M\mu\sigma^2_{max}(M)-lipschitz \\
pf: &\|M^\top\mathbb{T}Mx - M^\top\mathbb{T}My\| \\
&\leq \sigma_{max}(M)\|\mathbb{T}Mx-\mathbb{T}My\| \\
&\leq \sigma_{max}(M)L\|Mx-My\| \\
&\leq \sigma_{max}(M)L\sigma_{max}(M)\|x-y\| \\
&= \mu\sigma^2_{max}(M)\|x-y\|
\end{aligned}$$

3 非扩张、稳定非扩张与平均算子 (Nonexpansive and averaged operators)

定义 3.1 (非扩张算子、NE) 称一个算子 $\mathbb{T}$ 是非扩张的，如果:

$$\|\mathbb{T}x-\mathbb{T}y\| \leq \|x-y\|, \forall x, y \in dom\mathbb{T}$$

换句话说，非扩张算子就是 $1-\text{lipschitz}$ 算子。如果定义中，当 $x \neq y$ 时不等号严格成立，则称为收缩算子 (contraction)

注 若算子 $\mathbb{T}$、$\mathbb{S}$ 是非扩张的，则 $\mathbb{TS}$ 是非扩张的且 $\theta\mathbb{T}+(1-\theta)\mathbb{S}$ 是非扩张的,$\forall\theta\in(0,1]$ 又若非扩张算子 $\mathbb{T}$、$\mathbb{S}$ 中至少有一个是收缩算子，则 $\mathbb{TS}$ 是收缩的且 $\theta\mathbb{T}+(1-\theta)\mathbb{S}$ 是收缩算子,$\forall\theta\in(0,1]$

定义 3.2 (稳定非扩张算子、FNE) 称一个算子 $\mathbb{T}$ 是稳定非扩张的，如果:

$$\|\mathbb{T}x-\mathbb{T}y\|^2\leq\|x-y\|^2-\|(\mathbb{I}-\mathbb{T})x-(\mathbb{I}-\mathbb{T})y\|^2,\forall x,y\in dom\mathbb{T}$$

定理 3.1 $\mathbb{T}$ 是一个算子，则以下等价：

$$\begin{aligned}
&(1)\mathbb{T}\text{是稳定非扩张算子}\\
&(2)\mathbb{I}-\mathbb{T}\text{是稳定非扩张算子}\\
&(3)2\mathbb{T}-\mathbb{I}\text{是非扩张算子}\\
&(4)\|\mathbb{T}x-\mathbb{T}y\|^2\leq\langle x-y,\mathbb{T}x-\mathbb{T}y\rangle,\forall x,y\in dom\mathbb{T}\\
&(5)0\leq\langle\mathbb{T}x-\mathbb{T}y,(\mathbb{I}-\mathbb{T})x-(\mathbb{I}-\mathbb{T})y\rangle,\forall x,y\in dom\mathbb{T}
\end{aligned}$$

定义 3.3 (平均算子) 对于 $\theta\in(0,1)$，称算子 $\mathbb{T}$ 是 $\theta-$ 平均算子，如果

$$\exists\text{非扩张算子}\mathbb{S},\text{s.t.}\mathbb{T}=(1-\theta)\mathbb{I}+\theta\mathbb{S}$$

换句话说，平均算子是单位算子和一个非扩张算子的严格凸组合。$\frac{1}{2}-$ 平均算子又被称为稳定非扩张 (firmly nonexpansive)。

注 如果 $\mathbb{T}$ 是 $\frac{1}{2}-$ 平均算子，即 $\mathbb{T}=\frac{1}{2}\mathbb{I}+\frac{1}{2}(2\mathbb{T}-\mathbb{I})$，则 $2\mathbb{T}-\mathbb{I}$ 是非扩张算子，则有定理 3.1 得到 $\mathbb{T}$ 稳定非扩张

定理 3.2

$$\begin{aligned}
&\text{算子}\mathbb{T}\quad\alpha-\text{平均}\\
\Longleftrightarrow\ &(1-\frac{1}{\alpha})\mathbb{I}+\frac{1}{\alpha}\mathbb{T}\text{非扩张}\\
\Longleftrightarrow\ &\|\mathbb{T}x-\mathbb{T}y\|^2\leq\|x-y\|^2-\frac{1-\alpha}{\alpha}\|(\mathbb{I}-\mathbb{T})x-(\mathbb{I}-\mathbb{T})y\|^2,\forall x,y\in dom\mathbb{T}\\
\Longleftrightarrow\ &\|\mathbb{T}x-\mathbb{T}y\|^2+(1-2\alpha)\|x-y\|^2\leq2(1-\alpha)\langle x-y,\mathbb{T}x-\mathbb{T}y\rangle,\forall x,y\in dom\mathbb{T}
\end{aligned}$$

注 从定理 3.2 可以看出，当 $1\leq\frac{1-\alpha}{\alpha}$ 时，即 $0<\alpha\leq\frac{1}{2}$ 时，$\alpha-$ 平均算子 $\mathbb{T}$ 是稳定非扩张算子

注 对于非扩张算子 $\mathbb{T}$,$-1\leq\alpha\leq1$, 则 $\mathbb{A}=\mathbb{I}+\alpha\mathbb{T}$ 是单调算子。

pf:

$$
\begin{aligned}
&\langle x-y, \mathbb{A}x-\mathbb{A}y\rangle \\
=&\langle x-y, x-y+\alpha(\mathbb{T}x-\mathbb{T}y)\rangle \\
=&\|x-y\|^2+\alpha\langle x-y, \mathbb{T}x-\mathbb{T}y\rangle \\
\geq&\|x-y\|^2-|\alpha|\|x-y\|\|\mathbb{T}x-\mathbb{T}y\| \\
\geq&\|x-y\|^2-|\alpha|\|x-y\|^2\geq 0
\end{aligned}
$$

4 不动点迭代 (Fixed-point iteration)

定义 4.1 (不动点) x 为单值算子 $\mathbb{T}$ 的不动点，如果

$$\mathbb{T}x=x$$

记 $\mathrm{Fix}\mathbb{T}=\{x|x=\mathbb{T}x\}=(\mathbb{I}-\mathbb{T})^{-1}(0)$ 为单值算子 $\mathbb{T}$ 的不动点集合。

注 如果单值算子 $\mathbb{T}$ 是非扩张的且 $\mathrm{dom}\mathbb{T}=\mathbb{R}^n$，则 $\mathrm{Fix}\mathbb{T}$ 是闭凸集。

不动点迭代 (FPI) 的格式为：

$$x^{k+1}=\mathbb{T}x^k(k=0,1,2,\dots)$$

如果 $\mathbb{T}$ 是一个收缩算子，则可以利用 *Banach* 不动点定理得到收敛性，且收敛到不动点。但是对于优化问题，收缩的条件不易满足，下面我们来考虑对平均算子使用不动点迭代。

注 注意到，对于任意的非扩张算子 $\mathbb{T}$，则有平均算子 $(1-\theta)\mathbb{I}+\theta\mathbb{T}$ 的不动点集合与 $\mathbb{T}$ 是相同的，于是只要对平均算子 $(1-\theta)\mathbb{I}+\theta\mathbb{T}$ 运用不动点迭代就可以得到非扩张算子 $\mathbb{T}$ 的不动点。

引理 4.1 假设 $\{V^k\},\{S^k\}$ 是非负数列，且 $V^{K+1}\leq V^k-S^k(k=0,1,2,\dots)$，则 $\{V^k\}$ 单调递减，且 $S^K\to 0$
又若 $\{S^k\}$ 单调递减，则 $S^K\leq\frac{V^0}{N+1}$

定理 4.2 (平均算子不动点迭代的收敛性) 有平均算子 $\mathbb{I}=(1-\theta)\mathbb{I}+\theta\mathbb{S}$，其中 $\mathbb{S}$ 为非扩张算子，且 $\mathrm{Fix}\mathbb{T}$，令：

$$x^{k+1}=\mathbb{T}x^k(k=0,1,2,\dots)$$

则 $\{x^k\}$ 收敛于 $\mathrm{Fix}\mathbb{T}$ 中的一点，且 $\{\|x^{k+1}-x^k\|\}$ 单调递减且满足

$$\|x^{k+1}-x^k\|^2\leq\frac{\theta}{(k+1)(1-\theta)}dist^2(x^0,\mathrm{Fix}\mathbb{T})$$

证明.

$$
\begin{aligned}
\|x^{k+1}-x^*\|^2 &= \|(1-\theta)x^k+\theta\mathbb{S}x^k-x^*\|^2 \\
&= \|(1-\theta)x^k+\theta\mathbb{S}x^k-(1-\theta)x^*+\theta\mathbb{S}x^*\|^2 \\
&= \|(1-\theta)(x^k-x^*)+\theta(\mathbb{S}x^k-\mathbb{S}x^*)\|^2 \\
&= (1-\theta)\|x^k-x^*\|^2+\theta\|\mathbb{S}x^k-\mathbb{S}x^*\|^2-\theta(1-\theta)\|\mathbb{S}x^k-x^k\|^2 \\
&\leq (1-\theta)\|x^k-x^*\|^2+\theta\|x^k-x^*\|^2-\theta(1-\theta)\|\mathbb{S}x^k-x^k\|^2 \\
&= \|x^k-x^*\|^2-\theta(1-\theta)\|\mathbb{S}x^k-x^k\|^2 \\
&= \|x^k-x^*\|^2-\frac{(1-\theta)}{\theta}\|x^{k+1}-x^k\|^2
\end{aligned}
$$

则根据引理 4.1 得到 $\{\|x^k-x^*\|\}$ 单调递减，到 $\{\|x^{k+1}-x^k\|\}$ 收敛于 0。
又因为 $\mathbb{T}$ 是非扩张算子，则有 $\{\|x^{k+1}-x^k\|\}$ 单调递减。于是根据引理 4.1，得到：

$$
\|x^{k+1}-x^k\|^2 \leq \frac{\theta}{(k+1)(1-\theta)}\|x^0-x^*\|^2
$$

则：

$$
\|x^{k+1}-x^k\|^2 \leq \frac{\theta}{(k+1)(1-\theta)}\text{dist}^2(x^0,\text{Fix}\mathbb{T})
$$

下面证明点列 $\{x^k\}$ 收敛于 Fix$\mathbb{T}$ 中一点：
由上述证明可知，点列 $\{x^k\}$ 有界，取其中一个聚点 y^*，则有 $\{x^{k_i}\}$ 使得 $x^{k_i}\to y^*(i\to+\infty)$
由于 $\|x^{k+1}-x^k\|\to 0$，则 $(\mathbb{T}-\mathbb{I})(x^k)\to 0$，则 $(\mathbb{T}-\mathbb{I})(x^{k_i})\to 0$，则由海涅定理得到

$$
(\mathbb{T}-\mathbb{I})(y^*)=0
$$

, 则 $y^*\in$ Fix$\mathbb{T}$
将前半部分的 x^* 替换成 y^*，则有 $\{\|x^k-y^*\|\}$ 单调递减有下界，则存在极限，又因为子列 $\{\|x^{k_i}-y^*\|\}$ 有极限 0，则 $\{\|x^k-y^*\|\}$ 有极限 0，于是

$$
x^k\to y^*
$$

，即点列 $\{x^k\}$ 收敛于 Fix$\mathbb{T}$ 中一点。

例 4.1 (Forward Step method)

问题: 找到 $x\in\mathbb{R}^n$ 使得 $0=\mathbb{F}(x)$ 等价于求解 Fix$(\mathbb{I}-\alpha\mathbb{F})$ 则 FPI 为

$$
x^{k+1}=x^k-\alpha\mathbb{F}x^k:=\mathbb{T}(x^k)
$$

如果 $\mathbb{F}$ 是 $\beta-$ 余强制的，即 $\langle \mathbb{F}x-\mathbb{F}y, x-y\rangle \geq \beta\|\mathbb{F}x-\mathbb{F}y\|^2, \forall x,y\in domF$
则

$$\begin{aligned}
\|\mathbb{T}x-\mathbb{T}y\|^2 &= \|(\mathbb{I}-\alpha\mathbb{F})x-(\mathbb{I}-\alpha\mathbb{F})y\|^2 \\
&= \|x-y-\alpha(\mathbb{F}x-\mathbb{F}y)\|^2 \\
&= \|x-y\|^2-2\alpha\langle \mathbb{F}x-\mathbb{F}y, x-y\rangle+\alpha^2\|\mathbb{F}x-\mathbb{F}y\|^2 \\
&\leq \|x-y\|^2-2\alpha\beta\|\mathbb{F}x-\mathbb{F}y\|^2+\alpha^2\|\mathbb{F}x-\mathbb{F}y\|^2
\end{aligned}$$

由此可以得到，当 $\alpha^2-2\alpha\beta\leq 0$(即$0<\alpha\leq 2\beta$) 时，$\mathbb{T}$ 为非扩张算子，则 $\mathbb{T}_0=\mathbb{I}-2\beta\mathbb{F}$ 是非扩张算子。
而 $\mathbb{T}=\mathbb{I}-2\alpha\mathbb{F}=(1-\theta)\mathbb{I}+\theta(\mathbb{I}-2\beta\mathbb{F})=\mathbb{I}-2\theta\alpha\mathbb{F}$，其中 $\theta=\frac{\alpha}{2\beta}$，则 $\mathbb{T}$ 是 $\frac{\alpha}{2\beta}-$ 平均算子，所以 FPI 收敛 (当 $0<\alpha<2\beta$)。

又若 $\mathbb{F}$ 是 $\mu-$ 强单调的，即 $\langle \mathbb{F}x-\mathbb{F}y, x-y\rangle \geq \mu\|x-y\|^2, \forall x,y\in domF$
则

$$\begin{aligned}
\|\mathbb{T}x-\mathbb{T}y\|^2 &= \|(\mathbb{I}-\alpha\mathbb{F})x-(\mathbb{I}-\alpha\mathbb{F})y\|^2 \\
&= \|x-y-\alpha(\mathbb{F}x-\mathbb{F}y)\|^2 \\
&= \|x-y\|^2-2\alpha\langle \mathbb{F}x-\mathbb{F}y, x-y\rangle+\alpha^2\|\mathbb{F}x-\mathbb{F}y\|^2 \\
&\leq \|x-y\|^2-2\alpha\mu\|x-y\|^2+\frac{\alpha^2}{\beta^2}\|x-y\|^2(\beta-\text{余强制}\Rightarrow\frac{1}{\beta}-\text{Lipschitz}) \\
&= (1-2\alpha\mu+\frac{\alpha^2}{\beta^2})\|x-y\|^2
\end{aligned}$$

则当 $0<\alpha<2\mu\beta^2$ 时，$\mathbb{T}$ 是收缩算子，所以收敛。

根据上述分析，当 $\mathbb{F}=\nabla f$ 时，可以推出梯度下降法的收敛性。

5 Resolvents

算子 $\mathbb{A}$ 的 Resolvents 为 $\mathbb{J}_{\mathbb{A}}=(\mathbb{I}+\mathbb{A})^{-1}$

算子 $\mathbb{A}$ 的 Reflected Resolvents 为 $\mathbb{R}_{\mathbb{A}}=2\mathbb{J}_{\mathbb{A}}-\mathbb{I}$

定理 5.1 如果 $\mathbb{A}$ 为极大单调算子，则 $\mathbb{R}_{\mathbb{A}}$ 是非扩张算子，且 $\text{dom}\mathbb{R}_{\mathbb{A}}=\mathbb{R}^n$，$\mathbb{J}_{\mathbb{A}}$ 是 $\frac{1}{2}-$ 平均算子，且 $\text{dom}\mathbb{J}_{\mathbb{A}}=\mathbb{R}^n$

证明. 设 $(x,u),(y,v)\in\mathbb{J}_{\mathbb{A}}$，则 $x\in u+\mathbb{A}u, y\in v+\mathbb{A}v$，则

$$x-u\in\mathbb{A}u, y-v\in\mathbb{A}v$$

$$2u - x \in \mathbb{R}_{\mathbb{A}} x, 2v - y \in \mathbb{R}_{\mathbb{A}} y$$

则由 $\mathbb{A}$ 的单调性得到

$$\langle (x-u) - (y-v), u - v \rangle \geq 0$$

且

$$\begin{aligned} \|(2u-x) - (2v-y)\|^2 &= \|x - y - 2(u-v)\|^2 \\ &= \|x-y\|^2 - 4\langle u-v, x-y \rangle + 4\langle u-v, u-v \rangle \\ &= \|x-y\|^2 - 4\langle u-v, (x-u)-(y-v) \rangle \\ & leq\|x-y\|^2 \end{aligned}$$

于是 $\mathbb{R}_{\mathbb{A}}$ 是非扩张算子，而 $\mathbb{J}_{\mathbb{A}} = \frac{1}{2}\mathbb{I} + \frac{1}{2}\mathbb{R}_{\mathbb{A}}$ 为 $\frac{1}{2}-$ 平均算子。
$\mathrm{dom}\mathbb{R}_{\mathbb{A}} = \mathbb{R}^n, \mathrm{dom}\mathbb{J}_{\mathbb{A}} = \mathbb{R}^n$ 证明略。

定理 5.2 如果 $\mathbb{A}$ 为极大单调算子，则 $\mathrm{Fix}\mathbb{J}_{\mathbb{A}}$ 是闭凸集，且 $\mathrm{Zer}\mathbb{A} = \mathrm{Fix}\mathbb{J}_{\mathbb{A}}$

证明. 由于 $\mathbb{J}_{\mathbb{A}}$ 是非扩展算子，则 $\mathrm{Fix}\mathbb{J}_{\mathbb{A}}$ 是闭凸集。

$$\begin{aligned} x \in \mathrm{Zer}\mathbb{A} &\iff 0 \in \mathbb{A}x \\ &\iff x \in x + \mathbb{A}x \\ &\iff x = \mathbb{J}_{\mathbb{A}} x \\ &\iff x \in \mathrm{Fix}\mathbb{J}_{\mathbb{A}} \end{aligned}$$

$\mathbb{J}$,$\mathbb{R}$ 的基本性质:
假设 $\mathbb{A}$ 为极大单调算子 (根据上面的命题可以得到:$\mathbb{J}_{\mathbb{A}}$ 是单值算子)，$\alpha > 0$

* 如果 $\mathbb{B}(x) = \mathbb{A}x + t$，则 $\mathbb{J}_{\alpha\mathbb{B}}(u) = \mathbb{J}_{\alpha\mathbb{A}}(u - \alpha t)$
* 如果 $\mathbb{B}(x) = \mathbb{A}(x-t)$，则 $\mathbb{J}_{\alpha\mathbb{B}}(u) = \mathbb{J}_{\alpha\mathbb{A}}(u-t) + t$
* 如果 $\mathbb{B}(x) = -\mathbb{A}(t-x)$，则 $\mathbb{J}_{\alpha\mathbb{B}}(u) = t - \mathbb{J}_{\alpha\mathbb{A}}(t-u)$
* $\mathbb{J}_{\alpha^{-1}\mathbb{A}}(x) + \alpha^{-1}\mathbb{J}_{\alpha\mathbb{A}^{-1}}(\alpha x) = x$，特别的有 $\mathbb{J}_{\mathbb{A}} + \mathbb{J}_{\mathbb{A}}^{-1} = \mathbb{I}$
* $\mathbb{R}_{\mathbb{A}}(\mathbb{I} + \alpha\mathbb{A}) = \mathbb{I} - \alpha\mathbb{A}$
* 又若 $\mathbb{A}$ 是单值算子，则 $\mathbb{R}_{\mathbb{A}} = (\mathbb{I} - \alpha\mathbb{A})(\mathbb{I} + \alpha\mathbb{A})^{-1}$

例 5.1 考虑 $min\, f(x)$，其中 f 是适当闭凸函数，则 $\forall \alpha > 0$

$$\begin{aligned} x \in argmin\, f(x) &\iff 0 \in \partial f(x) \\ &\iff 0 \in \alpha \partial f(x) \\ &\iff x \in x + \alpha \partial f(x) \\ &\iff x = (\mathbb{I}) + \alpha \partial f)^{-1}(x) = \mathbb{J}_{\alpha \partial f}(x) \\ &\iff x \in \mathrm{Fix} \mathbb{J}_{\alpha \partial f} \end{aligned}$$

于是其不动点迭代的基本格式为:

$$\begin{aligned} & x^{k+1} = (\mathbb{I}) + \alpha \partial f)^{-1}(x^k) = \mathbb{J}_{\alpha \partial f}(x^k) \\ \iff & 0 \in \partial f(x^{k+1}) + \frac{1}{\alpha}(x^{k+1} - x^k) \\ \iff & x^{k+1} \in \underset{w}{argmin}\{f(w) + \frac{1}{2\alpha}\|w - x^k\|^2\} \\ \iff & x^{k+1} = \mathrm{Prox}_{\alpha f}(x^k) \end{aligned}$$

这就是临近点算法，即 PPA。

例 5.2 考虑带线性约束的优化问题:

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & Ax = b \end{aligned}$$

其中 f 是适当闭凸函数，其对偶问题为:

$$\begin{aligned} & \underset{\lambda}{max}\, \underset{x}{min}\, f(x) + \langle \lambda, Ax - b \rangle \\ = & \underset{\lambda}{max}\, \underset{x}{min}\, f(x) + \left\langle A^\top \lambda, x \right\rangle - \langle \lambda, b \rangle \\ = & \underset{\lambda}{max} - \langle \lambda, b \rangle - \underset{x}{max} \quad \left\langle -A^\top \lambda, x \right\rangle - f(x) \\ = & \underset{\lambda}{max} - \langle \lambda, b \rangle - f^*(-A^\top \lambda) \\ = & - \underset{\lambda}{min}\, f^*(-A^\top \lambda) + \langle \lambda, b \rangle \\ = & - \underset{\lambda}{min}\, g(\lambda) \end{aligned}$$

对对偶问题使用 PPA 算法得到:

$$\begin{aligned} & \lambda^{k+1} = \mathrm{Prox}_{\alpha g}(\lambda^k) = \mathbb{J}_{\alpha \partial g}(\lambda^k) \\ \iff & 0 \in \alpha \partial g(\lambda^{k+1}) + \lambda^{k+1} - \lambda^k \\ \iff & \lambda^{k+1} \in \lambda^k - \alpha \partial g(\lambda^{k+1}) \\ \iff & \lambda^{k+1} \in \lambda^k - \alpha[-A \partial f^*(-A^\top \lambda^{k+1}) + b] \\ \iff & \exists x^k \in \partial f^*(-A^\top \lambda^{k+1}), \lambda^{k+1} = \lambda^k - \alpha(-Ax^k + b) \\ \iff & \exists x^k \in \partial f^*(-A^\top \lambda^{k+1}), \lambda^{k+1} = \lambda^k + \alpha(Ax^k - b) \end{aligned}$$

而

$$
\begin{aligned}
& x^k \in \partial f^*(-A^\top \lambda^{k+1}) \\
\Longleftrightarrow & -A^\top \lambda^{k+1} \in \partial f(x^k) \\
\Longleftrightarrow & 0 \in \partial f(x^k) + A^\top \lambda^{k+1} \\
\Longleftrightarrow & 0 \in \partial f(x^k) + A^\top \left(\lambda^k + \alpha(Ax^k - b)\right) \lambda^{k+1} \\
\Longleftrightarrow & x^k \in \underset{w}{argmin}\{f(w) + \langle \lambda^k, Aw - b\rangle + \frac{\alpha}{2}\|Aw - b\|^2\}
\end{aligned}
$$

综合一下得到迭代过程为：

$$
\begin{aligned}
x^k &\in \underset{w}{argmin}\{f(w) + \langle \lambda^k, Aw - b\rangle + \frac{\alpha}{2}\|Aw - b\|^2\} \\
\lambda^{k+1} &= \lambda^k + \alpha(Ax^k - b)
\end{aligned}
$$

这就是增广拉格朗日方法，即 ALM，他是对对偶问题进行 PPA 算法。

6 算子分裂 (Operator Splitting)

6.1 Forward-backward splitting(FBS) and Backward-forward splitting(BFS)

问题:

$$
\underset{x\in\mathbb{R}^n}{find} \quad 0 \in (\mathbb{A} + \mathbb{B})x
$$

其中 $\mathbb{A}, \mathbb{B}$ 为极大单调算子，且 $\mathbb{A}$ 是 $\beta-$ 余强制的单值算子。

分析,$\forall \alpha > 0$：

$$
\begin{aligned}
& 0 \in (\mathbb{A} + \mathbb{B})x \\
\Longleftrightarrow \ & 0 \in (\mathbb{I} + \alpha\mathbb{B})x - (\mathbb{I} - \alpha\mathbb{A})x \\
\Longleftrightarrow \ & (\mathbb{I} - \alpha\mathbb{A})x \in (\mathbb{I} + \alpha\mathbb{B})x \\
\Longleftrightarrow \ & x = (\mathbb{I} + \alpha\mathbb{B})^{-1}(\mathbb{I} - \alpha\mathbb{A})x \\
\Longleftrightarrow \ & x = \mathbb{J}_{\alpha\mathbb{B}}(\mathbb{I} - \alpha\mathbb{A})x \\
\Longleftrightarrow \ & x \in \text{Fix}\mathbb{J}_{\alpha\mathbb{B}}(\mathbb{I} - \alpha\mathbb{A})
\end{aligned}
$$

由此，他的不动点迭代为 FBS:

$$
x^{k+1} = \mathbb{J}_{\alpha\mathbb{B}}(\mathbb{I} - \alpha\mathbb{A})x^k
$$

且当 $0 < \alpha < 2\beta$ 时，该 FBS 收敛。

在上面的分析中，可以改变一下分析思路，可以得到：

$$
\begin{aligned}
& 0 \in (\mathbb{A}+\mathbb{B})x \\
\Longleftrightarrow\ & 0 \in (\mathbb{I}+\alpha\mathbb{B})x - (\mathbb{I}-\alpha\mathbb{A})x \\
\Longleftrightarrow\ & (\mathbb{I}-\alpha\mathbb{A})x \in (\mathbb{I}+\alpha\mathbb{B})x \\
\Longleftrightarrow\ & \exists z \in (\mathbb{I}+\alpha\mathbb{B})x, z = (\mathbb{I}-\alpha\mathbb{A})x \\
\Longleftrightarrow\ & \exists z : x = (\mathbb{I}+\alpha\mathbb{B})^{-1}z, z = (\mathbb{I}-\alpha\mathbb{A})x \\
\Longleftrightarrow\ & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, z = (\mathbb{I}-\alpha\mathbb{A})\mathbb{J}_{\alpha\mathbb{B}}z \\
\Longleftrightarrow\ & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, z \in \mathrm{Fix}(\mathbb{I}-\alpha\mathbb{A})\mathbb{J}_{\alpha\mathbb{B}}
\end{aligned}
$$

由此，他的不动点迭代为 BFS:

$$
\begin{cases}
x^{k+1} = \mathbb{J}_{\alpha\mathbb{B}}z^k \\
z^{k+1} = (\mathbb{I}-\alpha\mathbb{A})z^k
\end{cases}
$$

且当 $0<\alpha<2\beta$ 时，该 FBS 收敛。

例 6.1 考虑问题：

$$
\min_{x\in\mathbb{R}^n} \quad f(x)+g(x)
$$

其中，f,g 都是适当闭凸函数，且 f 是梯度 $L-$ 利普西茨连续函数。对其利用 FBS 得到：

$$
\begin{aligned}
& x^{k+1} = \mathbb{J}_{\alpha\partial g}(\mathbb{I}-\alpha\nabla f)x^k \\
\Longleftrightarrow\ & x^{k+1} = \mathrm{Prox}_{\alpha g}((x^k - \alpha\nabla f(x^k)) \\
\Longleftrightarrow\ & x^{k+1} = \underset{w}{argmin}\{g(w) + \frac{1}{2\alpha}\|w-(x^k-\alpha\nabla f(x^k))\|^2\}
\end{aligned}
$$

即临近点梯度法 PG。

6.2 Peaceman-Rachford splitting(PRS) and Douglas-Rachford splitting(DRS)

问题:

$$
\underset{x\in\mathbb{R}^n}{find} \quad 0 \in (\mathbb{A}+\mathbb{B})x
$$

其中 $\mathbb{A},\mathbb{B}$ 为极大单调算子。

分析,$\forall \alpha > 0$：

$$
\begin{aligned}
& 0 \in (\mathbb{A}+\mathbb{B})x \\
\iff & 0 \in (\mathbb{I}+\alpha\mathbb{A})x - (\mathbb{I}-\alpha\mathbb{B})x \\
\iff & 0 \in (\mathbb{I}+\alpha\mathbb{A})x - \mathbb{R}_{\alpha\mathbb{A}}(\mathbb{I}+\alpha\mathbb{A})x \\
\iff & \exists z \in (\mathbb{I}+\alpha\mathbb{B})x, 0 \in (\mathbb{I}+\alpha\mathbb{A})x - \mathbb{R}_{\alpha\mathbb{B}}z \\
\iff & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, 0 \in (\mathbb{I}+\alpha\mathbb{B})\mathbb{J}_{\alpha\mathbb{B}}z - \mathbb{R}_{\alpha\mathbb{B}}z \\
\iff & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, \mathbb{R}_{\alpha\mathbb{B}}z \in (\mathbb{I}+\alpha\mathbb{B})\mathbb{J}_{\alpha\mathbb{B}}z \\
\iff & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, \mathbb{J}_{\alpha\mathbb{B}}z = \mathbb{J}_{\alpha\mathbb{B}}\mathbb{R}_{\alpha\mathbb{B}}z \\
\iff & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, 2\mathbb{J}_{\alpha\mathbb{B}}z - \mathbb{R}_{\alpha\mathbb{B}}z = 2\mathbb{J}_{\alpha\mathbb{B}}\mathbb{R}_{\alpha\mathbb{B}}z - \mathbb{R}_{\alpha\mathbb{B}}z \\
\iff & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, z = \mathbb{R}_{\alpha\mathbb{B}}\mathbb{R}_{\alpha\mathbb{B}}z \qquad (PRS) \\
\iff & \exists z : x = \mathbb{J}_{\alpha\mathbb{B}}z, z = \left(\frac{1}{2}\mathbb{I} + \frac{1}{2}\mathbb{R}_{\alpha\mathbb{B}}\mathbb{R}_{\alpha\mathbb{B}}\right) z \qquad (DRS)
\end{aligned}
$$

由此根据倒数第二个式子得到的不动点迭代为 PRS，但是由于 $\mathbb{R}_{\alpha\mathbb{B}}\mathbb{R}_{\alpha\mathbb{B}}$ 仅仅时非扩张算子，则该迭代不一定收敛，而倒数第一个式子得到的不动点迭代 DRS 是收敛的。

6.3 Davis-Yin splitting(DYS)

问题:

$$
\underset{x \in \mathbb{R}^n}{find} \quad 0 \in (\mathbb{A}+\mathbb{B}+\mathbb{C})x
$$

其中 $\mathbb{A},\mathbb{B},\mathbb{C}$ 为极大单调算子，$\mathbb{C}$ 是 $\beta-$ 余强制的单值算子。
分析,$\forall\alpha>0$：

$$
\begin{aligned}
& 0\in(\mathbb{A}+\mathbb{B})x \\
\Longleftrightarrow\ & 0\in(\mathbb{I}+\alpha\mathbb{A})x-(\mathbb{I}-\alpha\mathbb{B})x+\alpha\mathbb{C}x \\
\Longleftrightarrow\ & 0\in(\mathbb{I}+\alpha\mathbb{A})x-\mathbb{R}_{\alpha\mathbb{A}}(\mathbb{I}+\alpha\mathbb{B})x+\alpha\mathbb{C}x \\
\Longleftrightarrow\ & \exists z\in(\mathbb{I}+\alpha\mathbb{B})x, 0\in(\mathbb{I}+\alpha\mathbb{A})x-\mathbb{R}_{\alpha\mathbb{B}}z+\alpha\mathbb{C}x \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, 0\in(\mathbb{I}+\alpha\mathbb{A})\mathbb{J}_{\alpha\mathbb{B}}z-\mathbb{R}_{\alpha\mathbb{B}}z+\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}z \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, \left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z\in(\mathbb{I}+\alpha\mathbb{A})\mathbb{J}_{\alpha\mathbb{B}}z \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, \mathbb{J}_{\alpha\mathbb{B}}z=\mathbb{J}_{\alpha\mathbb{A}}\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, 2\mathbb{J}_{\alpha\mathbb{B}}z-\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z=2\mathbb{J}_{\alpha\mathbb{A}}\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z-\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, z+\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}z=\mathbb{R}_{\alpha\mathbb{A}}\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, z=\left(\mathbb{R}_{\alpha\mathbb{A}}\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)z \\
\Longleftrightarrow\ & \exists z: x=\mathbb{J}_{\alpha\mathbb{B}}z, z=\left(\frac{1}{2}\mathbb{I}+\frac{1}{2}\mathbb{T}\right)z, \mathbb{T}=\left(\mathbb{R}_{\alpha\mathbb{A}}\left(\mathbb{R}_{\alpha\mathbb{B}}-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)-\alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}\right)
\end{aligned}
$$

由此得到的不动点迭代即为 DYS:

$$
\begin{cases}
x^{k+\frac{1}{2}}=\mathbb{J}_{\alpha\mathbb{B}}z^k \\
x^{k+1}=\mathbb{J}_{\alpha\mathbb{A}}(2x^{k+\frac{1}{2}}-z^k-\alpha\mathbb{C}x^{k+\frac{1}{2}}) \\
z^{k+1}=z^k+x^k-x^{k+\frac{1}{2}}
\end{cases}
$$

可以证明当 $0<\alpha<2\beta$ 时，算法收敛。

7 矩阵变换方法 (Variable metric methods)

7.1 带有矩阵变换的 PPA(Variable metric proximal point method)

对于极大单调算子 $\mathbb{A}$ 以及正定矩阵 M，由第一节的知识得到算子 $M^{-\frac{1}{2}}\mathbb{A}M^{-\frac{1}{2}}$ 也是极大单调算子，对这个算子利用 PPA 得到:

$$
y^{k+1}=(\mathbb{I}+\alpha M^{-\frac{1}{2}}\mathbb{A}M^{-\frac{1}{2}})^{-1}y^k
$$

令 $x^k=M^{-\frac{1}{2}}y^k$，则:

$$
\begin{aligned}
& y^k\in(\mathbb{I}+\alpha M^{-\frac{1}{2}}\mathbb{A}M^{-\frac{1}{2}})y^{k+1} \\
\Longleftrightarrow\ & x^k\in(\mathbb{I}+\alpha M^{-1}\mathbb{A})x^{k+1} \\
\Longleftrightarrow\ & x^{k+1}=\mathbb{J}_{\alpha M^{-1}\mathbb{A}}(x^k)
\end{aligned}
$$

对此, 当 $\mathbb{A}=\partial f$, 其中 f 是一个适当闭凸函数, 则

$$\begin{aligned}
& x^k \in (\mathbb{I}+\alpha M^{-1}\partial f)x^{k+1} \\
\Longleftrightarrow\ & 0 \in \partial f(x^{k+1}) + \frac{1}{\alpha}M(x^{k+1}-x^k) \\
\Longleftrightarrow\ & x^{k+1} \in \underset{x}{argmin}\{f(x)+\frac{1}{2\alpha}\|x-x^k\|_M^2\}
\end{aligned}$$

7.2 带有矩阵变换的 FBS(Variable metric forward-backward splitting)

考虑问题:

$$\underset{x\in\mathbb{R}^n}{find} \quad 0 \in (\mathbb{A}+\mathbb{B})x$$

其中 $\mathbb{A},\mathbb{B}$ 为极大单调算子，且 $\mathbb{A}$ 是 $\beta-$ 余强制的单值算子。

取正定矩阵 M,$M^{-\frac{1}{2}}\mathbb{A}M^{-\frac{1}{2}}$,$M^{-\frac{1}{2}}\mathbb{B}M^{-\frac{1}{2}}$ 也是极大单调算子。则原问题等价于:

$$\underset{x\in\mathbb{R}^n}{find} \quad 0 \in (M^{-\frac{1}{2}}\mathbb{A}M^{-\frac{1}{2}}+M^{-\frac{1}{2}}\mathbb{B}M^{-\frac{1}{2}})x$$

对此使用 FBS 得到迭代格式为:

$$y^{k+1} = (\mathbb{I}+\alpha M^{-\frac{1}{2}}\mathbb{B}M^{-\frac{1}{2}})^{-1}(\mathbb{I}-\alpha M^{-\frac{1}{2}}\mathbb{A}M^{-\frac{1}{2}})y^k$$

设 $x^k = M^{-\frac{1}{2}}y^k$, 则得到:

$$\begin{aligned}
x^{k+1} &= (M+\alpha\mathbb{B})^{-1}(M-\alpha\mathbb{A})x^k \\
&= (\mathbb{I}+\alpha M^{-1}\mathbb{B})^{-1}(\mathbb{I}-\alpha M^{-1}\mathbb{A})x^k \\
&= \mathbb{J}_{\alpha M^{-1}\mathbb{B}}(\mathbb{I}-\alpha M^{-1}\mathbb{A})x^k
\end{aligned}$$

对此, 当 $\mathbb{A}=\nabla f$,$\mathbb{B}=\partial g$ 其中 f,g 是适当闭凸函数, 且 f 是梯度 $L-$ 利普西茨连续函数, 则:

$$\begin{aligned}
x^{k+1} &= \mathbb{J}_{\alpha M^{-1}\partial g}(\mathbb{I}-\alpha M^{-1}\nabla f)x^k \\
&= \underset{x}{argmin}\{g(x)+\langle\nabla f(x^k),x\rangle\ \frac{1}{2\alpha}\|x-x^k\|_M^2\}
\end{aligned}$$

7.3 M-范数意义下的平均算子 (Averagedness with respect to M-Norm)

取一个对称正定矩阵 M, 称算子 $\mathbb{T}$ 是在 $M-$ 范数下非扩张的, 如果:

$$\|\mathbb{T}x-\mathbb{T}y\|_M \leq \|x-y\|_M, \forall x,y \in dom\mathbb{T}$$

对于 $\theta\in(0,1)$, 称算子 $\mathbb{S}$ 是在 $M-$ 范数下的 $\theta-$ 平均算子, 如果 $\mathbb{S}=(1-\theta)\mathbb{I}+\theta\mathbb{T}$, 其中 $\mathbb{T}$ 是一个在 $M-$ 范数下非扩张算子

定理 7.1 M 为一个对称正定矩阵,$\mathbb{T}$ 为一个算子, 则算子 $M^{-\frac{1}{2}}\mathbb{T}M^{-\frac{1}{2}}$ 在 2-范数意义下是非扩张的等价于 $M^{-1}\mathbb{T}$ 在 $M-$ 范数意义下是非扩张的.

8 分裂方法中的对偶

8.1 Attouch-Thera duality

考虑问题:

$$\underset{x}{find} \quad 0 \in (\mathbb{A}+\mathbb{B})x$$

其中 $\mathbb{A},\mathbb{B}$ 是极大单调算子，则:

$$\begin{aligned}
& x \in \mathrm{Zer}(\mathbb{A}+\mathbb{B}) \\
\iff & \exists u : -u \in \mathbb{A}x, u \in \mathbb{B}x \\
\iff & \exists u : -x \in -\mathbb{A}^{-1}(-u), x \in \mathbb{B}^{-1}(u) \\
\Rightarrow & \exists u : 0 \in (-\mathbb{A}^{-1}(-\mathbb{I}) + \mathbb{B}^{-1})u \\
\iff & \exists u \in \mathrm{Zer}(-\mathbb{A}^{-1}(-\mathbb{I}) + \mathbb{B}^{-1})
\end{aligned}$$

记算子 $-\mathbb{A}^{-1}(-\mathbb{I}) = A^{-ⓥ}$, 则记以下问题为原问题的对偶问题:

$$\underset{u}{find} \quad 0 \in (A^{-ⓥ} + \mathbb{B}^{-1})u$$

可以证明:

$$\mathrm{Zer}(\mathbb{A}+\mathbb{B}) \neq \emptyset \iff \mathrm{Zer}(A^{-ⓥ} + \mathbb{B}^{-1}) \neq \emptyset$$

证明.

$$\begin{aligned}
& \mathrm{Zer}(\mathbb{A}+\mathbb{B}) \neq \emptyset \\
\iff & \exists x : 0 \in (\mathbb{A}+\mathbb{B})x \\
\iff & \exists x, u : -u \in \mathbb{A}x, u \in \mathbb{B}x \\
\iff & \exists x, u : -x \in -\mathbb{A}^{-1}(-u) = A^{-ⓥ}u, x \in \mathbb{B}^{-1}u \\
\iff & \exists u : 0 \in (A^{-ⓥ} + \mathbb{B}^{-1})u \\
\iff & \mathrm{Zer}(A^{-ⓥ} + \mathbb{B}^{-1}) \neq \emptyset
\end{aligned}$$

例 8.1 考虑优化问题:

$$\underset{x}{min} \quad f(x)+g(x)$$

其中 f,g 是适当闭凸函数，其等价于问题:

$$\underset{x}{find} \quad 0\in(\partial f+\partial g)x$$

其中 $\partial f,\partial g$ 是极大单调算子，其对偶问题为:

$$\begin{aligned}&\underset{u}{find} \quad 0\in((\partial f)^{-Ⓥ}+(\partial g)^{-1})u\\ &\Longleftrightarrow 0\in-(\partial f)^{-1}(-u)+(\partial g)^{-1}(u)\\ &\Longleftrightarrow 0\in-\partial f^*(-u)+\partial g^*(u)\end{aligned}$$

这相当于优化问题:

$$\underset{u}{max} \quad f^*(-u)+g^*(u)$$

即原优化问题的对偶问题。

8.2 具体的例子

在分裂算法中加入对偶变量的迭代，可以验证原始问题的正确性

例 8.2

$$\underset{x}{find} \quad 0\in(\mathbb{A}+\mathbb{B})x$$

其中 $\mathbb{A},\mathbb{B}$ 是极大单调算子，且 $\mathbb{A}$ 是单值算子，则 FBS 的迭代格式为:

$$\begin{cases}x^{k+\frac{1}{2}}=(\mathbb{I}-\alpha\mathbb{A})x^k\\ x^{k+1}=\mathbb{J}_{\alpha\mathbb{B}}x^{k+\frac{1}{2}}\end{cases}$$

加入了对偶迭代则变成:

$$\begin{cases}x^{k+\frac{1}{2}}=(\mathbb{I}-\alpha\mathbb{A})x^k\\ u^{k+\frac{1}{2}}=-\mathbb{A}x^k \quad (x^k\in A^{-Ⓥ}u^{k+\frac{1}{2}})\\ x^{k+1}=\mathbb{J}_{\alpha\mathbb{B}}x^{k+\frac{1}{2}}\\ u^{k+1}=\dfrac{1}{\alpha}(x^{k+\frac{1}{2}}-x^{k+1})\in\mathbb{B}x^{k+1} \quad (x^{k+1}\in\mathbb{B}^{-1}u^{k+1})\end{cases}$$

当 $x\to x^*\in\mathrm{Zer}(\mathbb{A}+\mathbb{B})$，设 $u^{k+\frac{1}{2}}\to-\mathbb{A}x^*=u^*$，则:

$$\begin{aligned}x^{k+\frac{1}{2}}&\to x^*-\alpha\mathbb{A}x^*=x^*+\alpha u^*\\ u^{k+1}&\to\frac{1}{\alpha}(x^*+\alpha u^*-x^*)=u^*\in\mathbb{B}x^*\end{aligned}$$

于是，$u^*\in\mathrm{Zer}(A^{-Ⓥ}+\mathbb{B}^{-1})$

例 8.3

$$\underset{x}{find} \quad 0 \in (\mathbb{A}+\mathbb{B})x$$

其中 $\mathbb{A},\mathbb{B}$ 是极大单调算子，则 DRS 的迭代格式为：

$$\begin{cases} x^{k+\frac{1}{2}} = \mathbb{J}_{\alpha\mathbb{B}} z^k \\ x^{k+1} = \mathbb{J}_{\alpha\mathbb{A}}(2x^{k+\frac{1}{2}} - z^k) \\ z^{k+1} = z^k + x^{k+1} - x^{k+\frac{1}{2}} \end{cases}$$

加入对偶迭代得到：

$$\begin{cases} x^{k+\frac{1}{2}} = \mathbb{J}_{\alpha\mathbb{B}} z^k \\ u^{k+\frac{1}{k}} = \dfrac{1}{\alpha}(z^k - x^{k+\frac{1}{2}}) \in \mathbb{B}x^{k+\frac{1}{2}} \\ x^{k+1} = \mathbb{J}_{\alpha\mathbb{A}}(2x^{k+\frac{1}{2}} - z^k) \\ u^{k+1} = \dfrac{1}{\alpha}(x^k - x^{k+\frac{1}{2}} + \alpha u^{k+\frac{1}{2}}) \in -\mathbb{A}x^{k+1} \\ z^{k+1} = z^k + x^{k+1} - x^{k+\frac{1}{2}} \end{cases}$$

则可以得到：

$$\begin{aligned} &x^k, x^{k+\frac{1}{2}} \to x^* \in \mathsf{Zer}(\mathbb{A}+\mathbb{B}) \\ &z^k \to z^* = x^* + \alpha u^* \in \mathsf{Fix}(\mathbb{R}_{\alpha\mathbb{A}}\mathbb{R}_{\alpha\mathbb{B}}) \\ &u^k, u^{k+\frac{1}{2}} \to u^* \in \mathsf{Zer}(A^{-ⓥ} + \mathbb{B}^{-1}) \end{aligned}$$

可以证明得到：

$$\begin{aligned} \mathbb{J}_{\frac{1}{\alpha}\mathbb{A}^{-ⓥ}} &= \mathbb{I} + \frac{1}{\alpha}\mathbb{J}_{\alpha\mathbb{A}}(-\alpha\mathbb{I}) \\ \mathbb{J}_{\frac{1}{\alpha}\mathbb{B}^{-1}} &= \mathbb{I} - \frac{1}{\alpha}\mathbb{J}_{\alpha\mathbb{B}}(\alpha\mathbb{I}) \end{aligned}$$

则

$$\begin{aligned} &\mathbb{R}_{\frac{1}{\alpha}\mathbb{A}^{-ⓥ}}\mathbb{R}_{\frac{1}{\alpha}\mathbb{B}^{-1}} \\ =&\left(2\mathbb{J}_{\frac{1}{\alpha}\mathbb{A}^{-ⓥ}} - \mathbb{I}\right)\left(2\mathbb{J}_{\frac{1}{\alpha}\mathbb{B}^{-1}} - \mathbb{I}\right) \\ =&\left(\frac{2}{\alpha}\mathbb{J}_{\alpha\mathbb{A}}(-\alpha\mathbb{I}) + \mathbb{I}\right)\left(\mathbb{I} - \frac{2}{\alpha}\mathbb{J}_{\alpha\mathbb{B}}(\alpha\mathbb{I})\right) \\ =&\left(\frac{2}{\alpha}\mathbb{J}_{\alpha\mathbb{A}}(-\alpha\mathbb{I}) + \mathbb{I}\right)\left(-\frac{1}{\alpha}\mathbb{I}\right)(-\alpha\mathbb{I})\left(\mathbb{I} - \frac{2}{\alpha}\mathbb{J}_{\alpha\mathbb{B}}(\alpha\mathbb{I})\right) \\ =&\frac{1}{\alpha}\mathbb{R}_{\alpha\mathbb{A}}\mathbb{R}_{\alpha\mathbb{B}}(\alpha\mathbb{I}) \\ \Longleftrightarrow\ &\mathbb{R}_{\alpha\mathbb{A}}\mathbb{R}_{\alpha\mathbb{B}} = \alpha\mathbb{R}_{\frac{1}{\alpha}\mathbb{A}^{-ⓥ}}\mathbb{R}_{\frac{1}{\alpha}\mathbb{B}^{-1}}(\alpha\mathbb{I}) \end{aligned}$$

则加入对偶迭代的 DRS 为：

$$\begin{cases} x^{k+\frac{1}{2}} = \mathbb{J}_{\alpha\mathbb{B}} z^k \\ u^{k+\frac{1}{k}} = \mathbb{J}_{\frac{1}{\alpha}\mathbb{B}^{-1}} \left(\frac{1}{\alpha}\mathbb{I}\right)(z^k) = \mathbb{I} - \frac{1}{\alpha}\mathbb{J}_{\alpha\mathbb{B}}(\alpha\mathbb{I})\left(\frac{1}{\alpha}z^k\right) = \frac{1}{\alpha}\left(z^k - \mathbb{J}_{\alpha\mathbb{B}}(z^k)\right) = \frac{1}{\alpha}(z^k - x^{k+\frac{1}{2}}) \\ x^{k+1} = \mathbb{J}_{\alpha\mathbb{A}}(2x^{k+\frac{1}{2}} - z^k) \\ u^{k+1} = \mathbb{J}_{\frac{1}{\alpha}\mathbb{A}^-\text{ⓥ}} \left(2u^{k+\frac{1}{2}} - \frac{1}{\alpha}z^k\right) = \left(\mathbb{I} + \frac{1}{\alpha}\mathbb{J}_{\alpha\mathbb{A}}(-\alpha\mathbb{I})\right)\left(2u^{k+\frac{1}{2}} - \frac{1}{\alpha}z^k\right) = \frac{1}{\alpha}(x^k - x^{k+\frac{1}{2}} + \alpha u^{k+\frac{1}{2}}) \\ z^{k+1} = z^k + x^{k+1} - x^{k+\frac{1}{2}} = z^k + \frac{1}{\alpha}\left(u^{k+1} - u^{k+\frac{1}{2}}\right) \end{cases}$$

这是一种自对偶性质。

例 8.4

$$\underset{x}{find} \quad 0 \in (\mathbb{A} + \mathbb{B})x$$

其中 $\mathbb{A}, \mathbb{B}, \mathbb{C}$ 是极大单调算子，且 $\mathbb{C}$ 是单值算子，则 DYS 的迭代格式为：

$$\begin{cases} x^{k+\frac{1}{2}} = \mathbb{J}_{\alpha\mathbb{B}} z^k \\ x^{k+1} = \mathbb{J}_{\alpha\mathbb{A}}(2x^{k+\frac{1}{2}} - z^k - \alpha\mathbb{C}x^{k+\frac{1}{2}}) \\ z^{k+1} = z^k + x^{k+1} - x^{k+\frac{1}{2}} \end{cases}$$

加入对偶迭代得到：

$$\begin{cases} x^{k+\frac{1}{2}} = \mathbb{J}_{\alpha\mathbb{B}} z^k \\ u^{k+\frac{1}{k}} = \frac{1}{\alpha}(z^k - x^{k+\frac{1}{2}}) \in \mathbb{B}x^{k+\frac{1}{2}} \\ x^{k+1} = \mathbb{J}_{\alpha\mathbb{A}}(2x^{k+\frac{1}{2}} - z^k - \alpha\mathbb{C}x^{k+\frac{1}{2}}) \\ u^{k+1} = \frac{1}{\alpha}(x^k - x^{k+\frac{1}{2}} + \alpha u^{k+\frac{1}{2}}) \in -\left(\mathbb{A}x^{k+1} + \mathbb{C}x^{k+\frac{1}{2}}\right) \\ z^{k+1} = z^k + x^{k+1} - x^{k+\frac{1}{2}} \end{cases}$$