一些典型凸优化问题的分裂收缩算法

基于变分不等式和邻近点算法的统一框架

南师大数科院系列讲座 2022 年元月 4 – 9 日

何 炳 生

Homepage: maths.nju.edu.cn/~hebma

二. 从原始对偶混合梯度法(PDHG)到邻近点(PPA)算法

1 引言

本章考虑鞍点问题的求解方法. 鞍点问题可以表述为如下的 $\min-\max$ 问题:

$$\min_x \max_y\{\Phi(x,y)=\theta_1(x)-y^TAx-\theta_2(y)\,|\,x\in\mathcal{X},y\in\mathcal{Y}\}. \tag{1.1}$$

其中 $\theta_1(x):\Re^n\to\Re$, $\theta_2(y):\Re^m\to\Re$ 是凸函数, $\mathcal{X}\subset\Re^n$, $\mathcal{Y}\subset\Re^m$ 是给定的凸集, $A\in\Re^{m\times n}$ 为给定的矩阵. 设 (x^*,y^*) 是问题 (1.1) 的解, 则有

$$(x^*,y^*)\in\mathcal{X}\times\mathcal{Y},\quad \Phi(x^*,y)\le\Phi(x^*,y^*)\le\Phi(x,y^*),\quad \forall(x,y)\in\mathcal{X}\times\mathcal{Y}.$$

也就是说

$$\left\{\begin{array}{ll} x^*\in\mathcal{X}, & \Phi(x,y^*)-\Phi(x^*,y^*)\ge 0,\quad \forall\, x\in\mathcal{X},\\ y^*\in\mathcal{Y}, & \Phi(x^*,y^*)-\Phi(x^*,y)\ge 0,\quad \forall\, y\in\mathcal{Y}.\end{array}\right.$$

利用 $\Phi(x,y)$ 的表达式, 上式就是

$$\left\{\begin{array}{ll} x^*\in\mathcal{X}, & \theta_1(x)-\theta_1(x^*)+(x-x^*)^T(-A^Ty^*)\ge 0,\quad \forall\, x\in\mathcal{X},\\ y^*\in\mathcal{Y}, & \theta_2(y)-\theta_2(y^*)\ +\ (y-y^*)^T(\ Ax^*)\ \ \ge 0,\quad \forall\, y\in\mathcal{Y}.\end{array}\right.$$

这可以表述成紧凑的变分不等式形式:

$$u\in\Omega,\quad \theta(u)-\theta(u^*)+(u-u^*)^TF(u^*)\ge 0,\ \forall\, u\in\Omega, \tag{1.2a}$$

其中

$$u=\begin{pmatrix} x\\ y\end{pmatrix},\quad \theta(u)=\theta_1(x)+\theta_2(y),\quad F(u)=\begin{pmatrix} -A^Ty\\ Ax\end{pmatrix},\quad \Omega=\mathcal{X}\times\mathcal{Y}. \tag{1.2b}$$

因为

$$F(u)=\begin{pmatrix} 0 & -A^T\\ A & 0\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix},\qquad \text{我们有}\quad (u-v)^T(F(u)-F(v))\equiv 0.$$

线性约束的凸优化问题

$$\min\{\theta(x) \mid Ax = b \text{ (or } \geq b), \, x \in \mathcal{X}\} \tag{1.3}$$

对应的拉格朗日函数是定义在 $\mathcal{X} \times \mathcal{Y}$ 上的

$$L(x, y) = \theta(x) - y^T(Ax - b). \tag{1.4}$$

其中

$$\mathcal{Y} = \begin{cases} \Re^m, & \text{if } Ax = b, \\ \Re^m_+, & \text{if } Ax \geq b, \end{cases}$$

这里的 $\Re^m_+$ 表示 $\Re^m$ 中的非负卦限. 如果一对 (x^*, y^*) 满足

$$(x^*, y^*) \in \mathcal{X} \times \mathcal{Y}, \quad L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*), \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y} \tag{1.5}$$

拉格朗日函数 (1.4) 鞍点鞍点中 $u^* = (x^*, y^*)$ 中的 x^* 就是凸优化问题 (1.3) 的解. 求拉格朗日函数 (1.4) 的鞍点就是鞍点问题 (1.1) 的一个特例, 其中

$$\theta_1(x) = \theta(x), \quad \theta_2(y) = -b^T y.$$

2 原始-对偶混合梯度法

求解鞍点问题 (1.1) 的原始-对偶混合梯度法 [15], PDHG 是一个比较自然的想法, 然而它并不能保证一定收敛.

求解鞍点问题 (1.1) 的原始-对偶混合梯度法 设 $r, s > 0$ 是给定的常数
对给定的 (x^k, y^k), PDHG 的第 k-步先由

$$x^{k+1} = \operatorname{argmin}\{\Phi(x, y^k) + \frac{r}{2}\|x - x^k\|^2 \mid x \in \mathcal{X}\}, \tag{2.1a}$$

给出 x^{k+1}, 然后再由

$$y^{k+1} = \operatorname{argmax}\{\Phi(x^{k+1}, y) - \frac{s}{2}\|y - y^k\|^2 \mid y \in \mathcal{Y}\}. \tag{2.1b}$$

产生 y^{k+1}. 完成一次迭代.

利用 $\Phi(x, y)$ 的表达式, 并注意到优化问题的解并不因为变动目标函数中的常数项而改变, PDHG 中的子问题 (2.2) 可以简化成

$$\begin{cases} x^{k+1} = \operatorname{argmin}\{\theta_1(x) - x^T A^T y^k + \frac{r}{2}\|x - x^k\|^2 \mid x \in \mathcal{X}\}, & (2.2a) \\ y^{k+1} = \operatorname{argmin}\{\theta_2(y) + y^T A x^{k+1} + \frac{s}{2}\|y - y^k\|^2 \mid y \in \mathcal{Y}\}. & (2.2b) \end{cases}$$

根据引理 **??**, 子问题 (2.2a) 的最优性条件是

$$\begin{aligned} x^{k+1} \in \mathcal{X}, \;\; & \theta_1(x) - \theta_1(x^{k+1}) \\ & + (x - x^{k+1})^T\{-A^T y^k + r(x^{k+1} - x^k)\} \geq 0, \;\; \forall x \in \mathcal{X}. \end{aligned} \tag{2.3a}$$

类似地, 子问题 (2.2b) 的最优性条件是

$$\begin{aligned} y^{k+1} \in \mathcal{Y}, \;\; & \theta_2(y) - \theta_2(y^{k+1}) \\ & + (y - y^{k+1})^T \{Ax^{k+1} + s(y^{k+1} - y^k)\} \geq 0, \; \forall\, y \in \mathcal{Y}. \end{aligned} \tag{2.3b}$$

将 (2.2a) 和 (2.3b) 加在一起, 我们有

$$\begin{aligned} u^{k+1} \in \Omega, \quad \theta(u) - \theta(u^{k+1}) + \begin{pmatrix} x - x^{k+1} \\ y - y^{k+1} \end{pmatrix}^T \Bigg\{ \begin{pmatrix} -A^T y^{k+1} \\ Ax^{k+1} \end{pmatrix} & \\ + \begin{pmatrix} r(x^{k+1} - x^k) + A^T(y^{k+1} - y^k) \\ s(y^{k+1} - y^k) \end{pmatrix} \Bigg\} & \geq 0, \quad \forall u \in \Omega. \end{aligned}$$

利用 (1.2), 其紧凑的形式是

$$\begin{aligned} u^{k+1} \in \Omega, \;\; & \theta(u) - \theta(u^{k+1}) \\ & + (u - u^{k+1})^T \{F(u^{k+1}) + Q(u^{k+1} - u^k)\} \geq 0, \; \forall u \in \Omega, \end{aligned} \tag{2.4a}$$

其中

$$Q = \begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix} \tag{2.4b}$$

是非对称的. 它跟PPA 形式并不匹配, 我们并不能保证算法收敛.

下面的线性规划的例子也的确说明 PDHG (2.2) 是不能保证收敛的. 对原始-对偶线性规划

$$\text{(Primal)} \quad \begin{array}{ll} \min & c^T x \\ \text{s. t.} & Ax = b, \\ & x \geq 0. \end{array} \qquad \text{(Dual)} \quad \begin{array}{ll} \max & b^T y \\ \text{s. t.} & A^T y \leq c. \end{array} \tag{2.5}$$

我们取如下的一对例子

$$\text{(P)} \quad \begin{array}{ll} \min & x_1 + 2x_2 \\ \text{s. t.} & x_1 + x_2 = 1 \\ & x_1, \; x_2 \geq 0. \end{array} \qquad \text{(D)} \quad \begin{array}{ll} \max & y \\ \text{s. t.} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} y \leq \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \end{array} \tag{2.6}$$

相当于 $A = [1, 1], \; b = 1, \; c = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. 这对问题拉格朗日函数

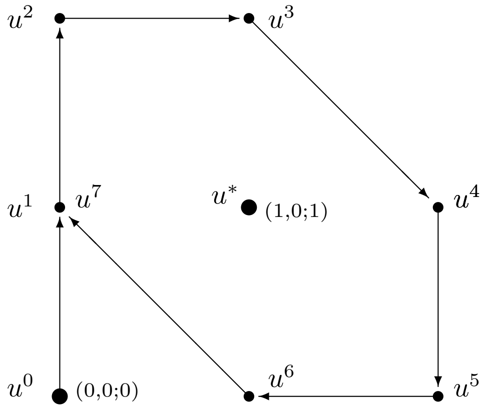
$$L(x, y) = c^T x - y^T (Ax - b) \tag{2.7}$$

定义在 $\Re^2_+ \times \Re$ 上. 问题的唯一最优解(也就是拉格朗日函数的鞍点) 是 $x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $y^* = 1$.

用 PDHG (2.2) 求线性规划 (2.6) 对应的拉格朗日函数 (2.7) 的鞍点, 具体的迭代公式是

$$\begin{cases} x^{k+1} = \max\{(x^k + \frac{1}{r}(A^T y^k - c)), 0\}, \\[2mm] y^{k+1} = y^k - \frac{1}{s}(Ax^{k+1} - b). \end{cases}$$

我们用 $(x_1^0, x_2^0; y^0) = (0, 0; 0)$ 作为初始点, 迭代点列并不收敛.



$$
\begin{aligned}
u^0 &= (0, 0; 0) \\
u^1 &= (0, 0; 1) \\
u^2 &= (0, 0; 2) \\
u^3 &= (1, 0; 2) \\
u^4 &= (2, 0; 1) \\
u^5 &= (2, 0; 0) \\
u^6 &= (1, 0; 0) \\
u^7 &= (0, 0; 1) \\
\boldsymbol{u^{k+6}} &= \boldsymbol{u^k}
\end{aligned}
$$

Fig. 3.1 取 $r = s = 1$ 的 PDHG 迭代序列

是不是增大 r, s 就会收敛呢, 实验结果也不是.

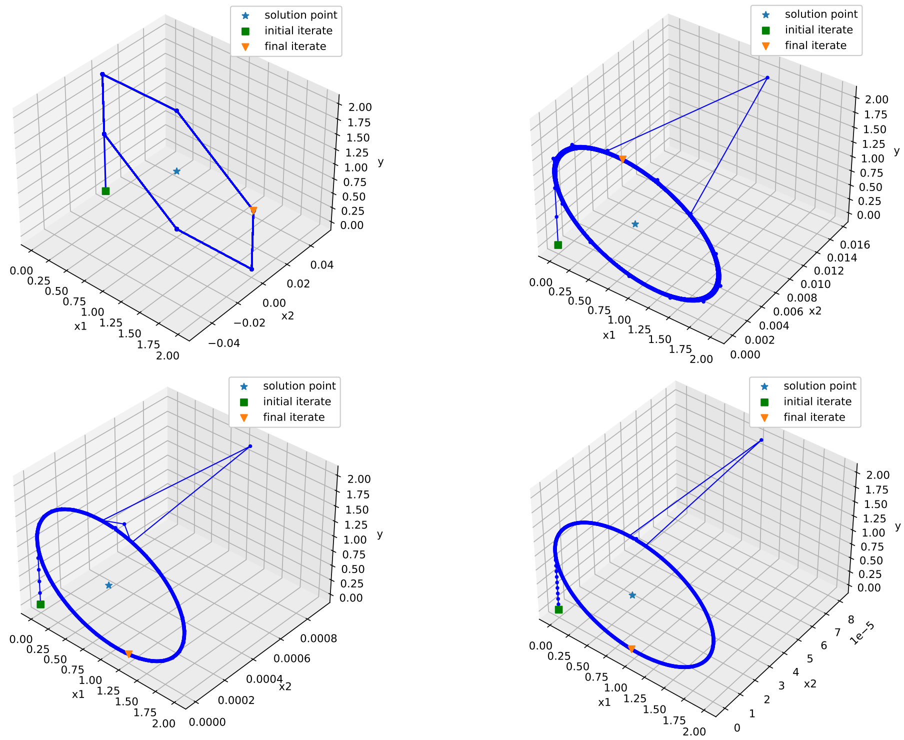


Fig. 3.2 取 $r = s = 1,\ 2,\ 5,\ 10$ 的 PDHG 迭代序列都不收敛

3 求解鞍点问题的邻近点算法 (PPA)

如果我们希望把变分不等式 (2.4) 的形式改造成

$$\begin{aligned} u^{k+1}\in\Omega,\quad & \theta(u)-\theta(u^{k+1}) \\ & +(u-u^{k+1})^T\{F(u^{k+1})+H(u^{k+1}-u^k)\}\geq 0,\ \forall u\in\Omega, \end{aligned} \tag{3.1a}$$

其中

$$H=\begin{pmatrix} rI_n & A^T \\ A & sI_m \end{pmatrix}, \tag{3.1b}$$

当 $rs>\|A^TA$ 时, 矩阵 H 正定, 我们得到第一章中的 PPA 形式 算法就是收敛的. 为此, 要把 (2.4b) 中的非对称的矩阵 Q 改造成 (3.1b) 中的对称矩阵 H:

$$Q=\begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix} \qquad \Rightarrow \qquad H=\begin{pmatrix} rI_n & A^T \\ A & sI_m \end{pmatrix}.$$

为了实现这个目的, 我们只要将 (2.3b) 中的

$$y^{k+1}\in\mathcal{Y},\ \ \theta_2(y)-\theta_2(y^{k+1})+(y-y^{k+1})^T\{Ax^{k+1}+s(y^{k+1}-y^k)\}\geq 0,\ \forall\, y\in\mathcal{Y}$$

改造成

$$y^{k+1}\in\mathcal{Y},\ \ \theta_2(y)-\theta_2(y^{k+1})+(y-y^{k+1})^T\begin{bmatrix} Ax^{k+1}+A(x^{k+1}-x^k) \\ +s(y^{k+1}-y^k) \end{bmatrix}\geq 0,\ \forall\, y\in\mathcal{Y}.$$

也就是说,

$$y^{k+1}\in\mathcal{Y},\ \ \theta_2(y)-\theta_2(y^{k+1})+(y-y^{k+1})^T\{A[2x^{k+1}-x^k]+s(y^{k+1}-y^k)\}\geq 0,\ \forall\, y\in\mathcal{Y}.$$

根据引理 **??**, 上面的形式说明

$$y^{k+1}=\mathrm{argmin}\big\{\theta_2(y)+y^T[A(2x^{k+1}-x^k)]+\frac{s}{2}\|y-y^k\|^2\,\big|\,y\in\mathcal{Y}\big\}.$$

这相当于把 (2.2b) 目标函数中

出现的 $\quad Ax^{k+1} \quad$ 改成 $\quad A(2x^{k+1}-x^k)$.

换句话说,

求解鞍点问题 (1.1) **的 PPA 算法** 设 $r,s>0$ 是给定的满足 $rs>\|A^TA\|$ 的常数
对给定的 (x^k,y^k), PDHG 的第 k-步先由

$$x^{k+1}=\mathrm{argmin}\{\Phi(x,y^k)+\frac{r}{2}\|x-x^k\|^2\,|\,x\in\mathcal{X}\}, \tag{3.2a}$$

给出 x^{k+1}, 然后再由

$$y^{k+1}=\mathrm{argmax}\big\{\Phi\big([2x^{k+1}-x^k],y\big)-\frac{s}{2}\big\|y-y^k\big\|^2\,|\,y\in\mathcal{Y}\big\} \tag{3.2b}$$

产生 y^{k+1}. 得到 (x^{k+1},y^{k+1}) 完成一次迭代.

根据上面的分析, 我们有如下的定理:

定理 3.1 当 $rs > \|A^TA$, 用 *PPA* (3.2) 求解鞍点问题 (1.1) 产生的序列 $\{u^k = (x^k, y^k)\}$ 满足

$$\|u^{k+1} - u^*\|_H^2 \le \|u^k - u^*\|_H^2 - \|u^k - u^{k+1}\|_H^2, \quad \forall u^* \in \Omega^*, \tag{3.3}$$

其中 H 是由 (3.1b) 给出的正定矩阵.

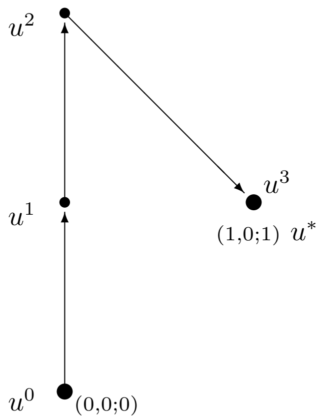
对线性约束为的优化问题 (1.3), 用 PPA 算法 (3.2) 求解的具体格式是

$$\left\{\begin{array}{ll} x^{k+1} = \operatorname{argmin}\big\{\theta(x) + \dfrac{r}{2}\big\|x - \big[x^k + \dfrac{1}{r}A^Ty^k\big]\big\|^2 \,\big|\, x \in \mathcal{X}\big\}, & \text{(3.4a)} \\ y^{k+1} = P_{\mathcal{Y}}\big[y^k - \dfrac{1}{s}\big[A(2x^{k+1} - x^k) - b\big]. & \text{(3.4b)} \end{array}\right.$$

用 PPA (3.2) 求线性规划 (2.6) 对应的拉格朗日函数 (2.7) 的鞍点, 迭代公式就变成

$$\left\{\begin{array}{l} x^{k+1} = \max\{(x^k + \frac{1}{r}(A^Ty^k - c)), 0\}, \\ y^{k+1} = y^k - \frac{1}{s}\big(A(2x^{k+1} - x^k) - b\big). \end{array}\right.$$

我们用 $r = s = 1$, 取 $(x_1^0, x_2^0; y^0) = (0, 0; 0)$ 作为初始点, 迭代序列收敛.



$u^0 = (0, 0; 0)$
$u^1 = (0, 0; 1)$
$u^2 = (0, 0; 2)$
$u^3 = (1, 0; 1)$

$\boldsymbol{u^3 = u^*}$.

Fig. 3.3　取 $r = s = 1$ 的 PPA 迭代序列

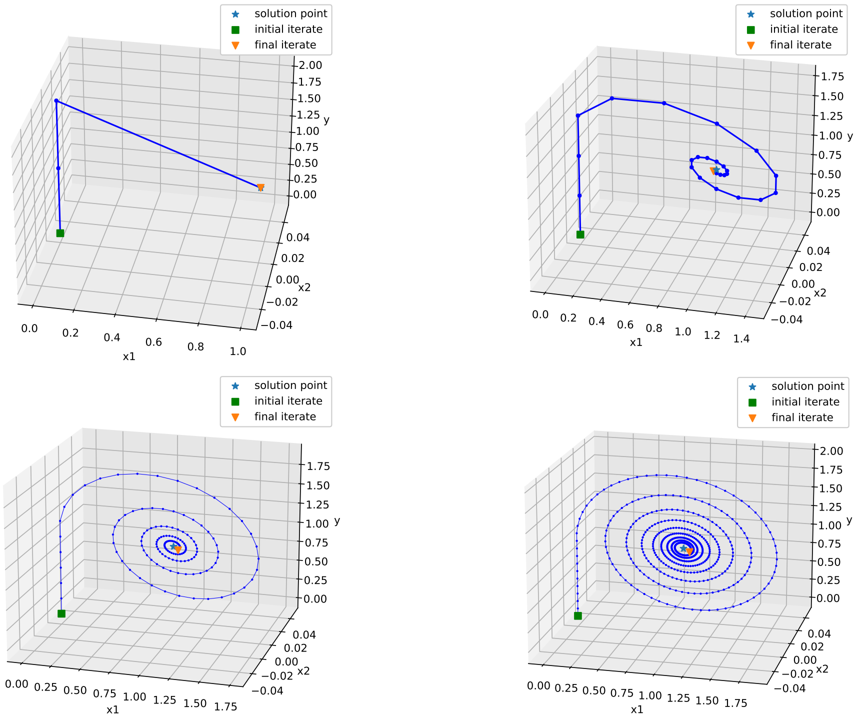


Fig. 3.4. 取 $r = s = 1, 2, 5, 10$, PPA 方法都收敛. 参数越大, 收敛越慢

PPA 算法 (3.2) 用的是 primal-dual 顺序, 我们同样可以采用 dual-primal 顺序的 PPA.

求解鞍点问题 (1.1) **的 PPA 算法** (in dual-primal order)
设 $r, s > 0$ 是给定的满足 $rs > \|A^T A\|$ 的常数, 第 k-步迭代从给定的 $u^k = (x^k, y^k)$ 开始,

$$\begin{cases} y^{k+1} = \operatorname{argmax}\{\Phi(x^k, y) - \frac{s}{2}\|y - y^k\|^2 \mid y \in \mathcal{Y}\}, & (3.5a) \\ x^{k+1} = \operatorname{argmin}\{\Phi(x, (2y^{k+1} - y^k)) + \frac{r}{2}\|x - x^k\|^2 \mid x \in \mathcal{X}\}. & (3.5b) \end{cases}$$

由 (3.5) 产生的 $u^{k+1} \in \Omega$ 满足

$$\begin{aligned} u^{k+1} \in \Omega, \;\; & \theta(u) - \theta(u^{k+1}) \\ & + (u - u^{k+1})^T \{F(u^{k+1}) + Q(u^{k+1} - u^k)\} \geq 0, \;\; \forall u \in \Omega, \end{aligned} \tag{3.6a}$$

其中

$$H = \begin{pmatrix} rI_n & -A^T \\ -A & sI_m \end{pmatrix} \tag{3.6b}$$

$$u^{k+1} \in \Omega, \;\; \theta(u) - \theta(u^{k+1}) + (u - u^{k+1})^T \{F(u^{k+1}) + H(u^{k+1} - u^k)\} \geq 0, \;\; \forall u \in \Omega,$$

其中

$$H = \begin{pmatrix} rI_n & -A^T \\ -A & sI_m \end{pmatrix}.$$

对线性约束为的优化问题 (1.3), 用 PPA 算法 (3.5) 求解的具体格式是

$$\begin{cases} y^{k+1} = P_{\mathcal{Y}}\big[y^k - \frac{1}{s}(Ax^k - b)\big] & (3.7a) \\ x^{k+1} = \mathrm{argmin}\big\{\theta(x) + \frac{r}{2}\big\|x - \big[x^k + \frac{1}{r}A^T(2y^{k+1} - y^k)\big]\big\|^2 \,\big|\, x \in \mathcal{X}\big\}. & (3.7b) \end{cases}$$

4 Relationship to Chambolle-Pock Method

Chambolle and Pock [2] have proposed a method for solving the convex-concave $\min-\max$ problem, in short, C-P method. Applied C-P method to the problem (1.1), it is also required $rs > \|A^TA\|$.

CP method. For given (x^k, y^k), C-P method obtains x^{k+1} via

$$x^{k+1} = \arg\min\{\Phi(x, y^k) + \frac{r}{2}\|x - x^k\|^2 \mid x \in \mathcal{X}\}. \tag{4.1a}$$

Then, y^{k+1} is given by

$$y^{k+1} = \arg\max\{\Phi([x^{k+1} + \tau(x^{k+1} - x^k)], y) - \frac{s}{2}\|y - y^k\|^2 \mid y \in \mathcal{Y}\} \tag{4.1b}$$

where $\tau \in [0, 1]$.

- 原始-对偶混合梯度法(PDHG) (2.2) 和按需定制的邻近点算法(C-PPA) (3.2) 都是 Chambolle-Pock 方法 [2] 分别取 $\tau = 0$ 和 $\tau = 1$ 的特例.
- 对 $\tau = 0$ 的 PDHG 方法 (2.2), §2 中已经说明不能保证收敛. 对 $\tau = 1$ 的 CPPA 方法 (3.2), 其收敛性在 §3 中有了结论.
- 根据我们的知识, 对于 $\tau \in (0, 1)$ 的 CP 方法 (4.1), 收敛性还没有定论.

附上我以前写的: CP 方法十年记 2020 年9 月

- Chambolle 和 Pock 在 2010 年提出的求解 $\min-\max$ 问题的原始-对偶方法, 在图像处理领域有着广泛的应用和很大的影响, 被称为 CP 方法。
- Chambolle 和Pock 方法的第一个版本公布于2010 年6 月. 他们的方法中有个 $[0, 1]$ 之间的参数, 但在文章中, 只对参数为1 的方法给了证明. 读了他们的这篇文章以后, 我们对这类方法的收敛性进行了研究.
- 由于我们多年研究单调变分不等式的求解方法, 很快发现, 参数为1 的 CP 方法, 可以解释为变分不等式H-模(H为对称正定矩阵) 的邻近点算法 (PPA), 因此收敛性证明特别简单. 五个月后的 2010 年 11 月 4 日, 我们把相关证明的第一稿, OO-2790, 公布在 Optimization Online 上. 同时, 对参数为 0 的 CP 方法, 我们找到了不收敛的例子
- 参数在 $(0, 1)$ 间的CP 方法, 能不能保证收敛, 这个问题至今没有解决.

- Chambolle 和 Pock 很快发现了我们的工作, 一个多月后的 2010 年 12 月 21 日, 他们的文章在 J. MIV online 正式发表. 我们高兴地看到, Chambolle 和 Pock 这么快就注意到并引用了我们的文章, 也提到了我们的证明. 我们的文章正式发表以后, CP 后来就不再提参数在 $[0,1)$ 间的方法了.
- 特别感谢CP 方法的原创者认可我们给出的简单证明. 他们在2011年的IEEE ICCV 会议论文中, 称赞我们的工作极大地简化了收敛性分析 (which greatly simplifies the convergence analysis).
- 后来CP 方法的作者又有多篇相关的文章发表(后面的文章他们都只讨论参数为 1 的方法). 他们于2016 年在Math. Progr. 发表的文章中, 继续利用我们的 PPA 解释, 文章的引言中就开诚布公(In particular, exploiting a proximal-point interpretation due to [16], we are able to give a very elementary proof). 这里的[16] 是我们 2010 年的预印本 OO-2790, 2012 年春发表在 SIAM Imaging Science.

参考文献

[1] X.J. Cai, G.Y. Gu, B.S. He and X.M. Yuan, A proximal point algorithms revisit on the alternating direction method of multipliers, Science China Mathematics, 56 (2013), 2179-2186.

[2] A. Chambolle, T. Pock, A first-order primal-dual algorithms for convex problem with applications to imaging, J. Math. Imaging Vison, 40, 120-145, 2011.

[3] Chambolle A and Pock T. On the ergodic convergence rates of a first-order primal-dual algorithm, *Mathematical Programming*, 2016, 159: 253–287.

[4] C. H. Chen, B. S. He, Y. Y. Ye and X. M. Yuan, *The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent*, to appear in *Mathematical Programming, Series A*.

[5] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[6] Gu G Y, He B S and Yuan X M. Customized proximal point algorithms for linearly constrained convex minimization and saddle-point problems: a unified approach. Comput. Optim. Appl., 2014, 59: 135-161.

[7] B. S. He, PPA-like contraction methods for convex optimization: a framework using variational inequality approach. J. Oper. Res. Soc. China **3** (2015) 391 – 420.

[8] B. S. He, Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities, *Computational Optimization and Applications* **42**(2009), 195–212.

[9] B. S. He, M. Tao and X.M. Yuan, Alternating direction method with Gaussian back substitution for separable convex programming, *SIAM Journal on Optimization* **22**(2012), 313-340.

[10] B.S. He, M. Tao and X.M. Yuan, A splitting method for separable convex programming, *IMA Journal of Numerical Analysis*, **31**, 394-426, 2015.

[11] B. S. He and X. M. Yuan, On the $O(1/t)$ convergence rate of the alternating direction method, *SIAM J. Numerical Analysis* **50**(2012), 700-709.

[12] B.S. He and X. M. Yuan, Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective, SIAM Journal on Imaging Science, **5**, 119-149, 2012.

[13] B. Martinet, Regularisation, d'inéquations variationelles par approximations succesives, *Rev. Francaise d'Inform. Recherche Oper.*, **4**, 154-159, 1970.

[14] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Cont. Optim.*, **14**, 877-898, 1976.

[15] M. Zhu and T. F. Chan, An efficient primal-dual hybrid gradient algorithm for total variation image restoration, CAM Reports 08-34, UCLA, 2008.

[16] 何炳生. 从变分不等式的统一收缩算法到凸优化的分裂收缩算法, 高等学校计算数学学报, 2016, 38: 74–96.

[17] 何炳生. 凸优化的一阶分裂算法–变分不等式为工具的统一框架, 见 http://maths.nju.edu. cn/~hebma 中的《My Talk》.

[18] 何炳生. 凸优化和单调变分不等式的收缩算法, 见 http://maths.nju.edu.cn/~hebma 中的系列讲义.