数 学 杂 志

J. of Math. (PRC)

Vol. 40 (2020) No. 2

简单约束鞍点优化问题的投影原始 - 对偶算法

郭晓1,陶越2

(1. 桂林电子科技大学材料科学与工程学院,广西 桂林 541004)
 (2. 江苏省广播电视总台,江苏 南京 210000)

摘要: 本文研究了带有简单凸集约束的鞍点优化问题.利用问题的凸凹特性,提出了一个投影 原始 - 对偶梯度方法.算法具有对称结构且每步具有显示解.证明了新算法的收敛性并获得了收敛速 率. 泊松噪音图像恢复问题的数值结果验证了算法的有效性.

 关键词:
 鞍点问题; 原始 – 对偶; 图像恢复

 MR(2010) 主题分类号:
 65K10; 90C25
 中图分类号: O224

 文献标识码:
 A
 文章编号: 0255-7797(2020)02-0245-08

1 引言

考虑如下鞍点优化问题

$$\min_{x \in X} \max_{y \in Y} Q(x, y), \tag{1.1}$$

其中 $X \subseteq R^n, Y \subseteq R^m$ 是两个闭凸集合, Q(x, y) 是光滑的凸 – 凹函数. 该模型在图像处理, 高维统计, 机器学习中有着广泛应用. 注意到鞍点问题实际上可以转化为单调变分不等式 问题, 那么求解单调变分不等式的数值方法 ^[1-3] 都可以用来求解问题 (1.1). 令 $v = (x, y)^T$, $F(v) = (\nabla_x Q(x, y), -\nabla_y Q(x, y))^T$, $K = X \times Y$, 那么寻找极小 – 极大问题 (1.1) 的鞍点 (x^*, y^*) 就可以转换为如下形式的变分不等式问题

$$(v - v^*)^T F(v^*) \ge 0, \quad \forall v \in K.$$

易知, v^* 是上述变分不等式的解当且仅当其满足不定点方程 $v^* = P_K(v^* - \alpha \nabla F(v^*))$, 其中 *P* 是投影算子, $\alpha > 0$ 是步长. 鉴于约束集合 *K* 的可分结构, 不动点方程可以具体写为

$$x^* = P_X(x^* - \alpha \nabla_x Q(x^*, y^*)), y^* = P_Y(y^* + \alpha \nabla_y Q(x^*, y^*)).$$

基于此, Bonettini 和 Ruggiero 在文献 [4] 中提出了如下外梯度方案求解问题 (1.1)

$$\begin{cases} y_{k+\frac{1}{2}} = P_Y(y_k + \alpha_k \nabla_y Q(x_k, y_k)), \\ x_{k+1} = P_X(x_k - \alpha_k \nabla_x Q(x_k, y_{k+\frac{1}{2}})), \\ y_{k+1} = P_Y(y_k + \alpha_k \nabla_y Q(x_{k+1}, y_k)), \end{cases}$$
(1.2)

*收稿日期: 2019-06-26 接收日期: 2019-10-09

基金项目: 广西自然科学基金资助 (2018GXNSFBA281023); 广西自动检测技术与仪器重点实验室基金 资助 (YQ18107).

作者简介:郭晓(1988-),女,陕西渭南,硕士,主要研究方向:最优化方法与应用.

其中 α_k 是自适应步长,由线搜索确定.该方法中,原始变量和对偶变量共用一个步长,且能 自适应确定,这是该方法的优点,不必费大量时间调参.收敛性也能在较弱的条件下证明.该 方法用在全变分图像恢复问题中,数值表现较好.

此外, 问题 (1.1) 有着特殊的结构可以利用.因此, 许多高效的数值算法 ^[5-10] 提了出来. 特别地,即 $Q(x,y) = f(x) + \langle Bx, y \rangle - g(y)$.当f(x) 是二次数据拟合项,g(y) 为示性函数, B 是梯度算子时,全变分图像恢复模型可以看作是问题 (1.1) 的一个特殊形式.著名的原始 – 对偶混合梯度算法 (PDHG) 在文献 [5] 中提了出来.该方法虽然数值表现良好,但收敛性 还没被确定.随后,许多学者开始关注此类型算法.He 等学者在文献 [6] 中证明了如果目标 函数中有一个函数是强凸的,那么 PDHG 是收敛的.如果算法中的步长序列满足一定的条 件,Bonettini 等在文献 [7] 中亦证明了 PDHG 的收敛性.Chambolle 和 Pock 在文献 [8] 中 对 PDHG 算法进行了改进,并得到了收敛性结论和次线性收敛速率.利用相关的不动点理 论,Chen 等在文献 [9] 中提出了一个原始 – 对偶不动点算法.当 $X = R^n$ 时,Zhang 等在文 献 [10] 中给出了一个固定步长的原始 – 对偶方法且原始变量的步长与对偶变量的步长不同, 算法描述如下:

$$\begin{cases} y_{k+\frac{1}{2}} = P_Y(y_k + \gamma \nabla_y Q(x_k, y_k)), \\ x_{k+1} = x_k - \beta \nabla_x Q(x_k, y_{k+\frac{1}{2}}), \\ y_{k+1} = P_Y(y_k + \gamma \nabla_y Q(x_{k+1}, y_k)), \end{cases}$$
(1.3)

其中 β, γ > 0 为步长.可以看出该方法每步都有闭示解,子问题无需内迭代求解及矩阵求逆, 适合应用于计算机断层成像和并行磁共振成像等大规模问题上.

可以看到算法 (1.2) 和 (1.3) 区别在于步长的选取不同. 算法 (1.2) 中,每次迭代的步长相同由线搜索确定. 在大规模问题中线搜索可能会增加一些额外的计算时间. 算法 (1.3) 可以取不同的固定步长,但与 (1.2) 相比,它不能有效地处理变量全有约束的鞍点问题. 本文希望结合两个算法的优点,给出一个新算法. 结合投影算子的性质和凸分析的相关知识,给出算法的收敛性分析及收敛速率. 最后,将所设计的算法用于含有泊松噪音的图像恢复问题上,给出相应的数值实验结果.

2 提出的方法

本节给出一个新的外梯度方法,并分析算法的收敛性和收敛速率.结合算法 (1.2) 和 (1.3),一个简单的思路是用固定步长代替算法 (1.2) 中线搜索确定的步长,以期待减少算法的 计算时间,且能处理约束情形下的问题 (1.1).提出的方法形式简单,有如下的迭代步骤

$$\begin{cases} y_{k+\frac{1}{2}} = P_Y(y_k + \gamma \nabla_y Q(x_k, y_k)), \\ x_{k+1} = P_X(x_k - \beta \nabla_x Q(x_k, y_{k+\frac{1}{2}})), \\ y_{k+1} = P_Y(y_k + \gamma \nabla_y Q(x_{k+1}, y_k)), \end{cases}$$
(2.1)

其中 $\beta, \gamma > 0$ 满足一定的条件,这将会在下面给出具体的说明.

接下来的部分,具体分析新方法 (2.1) 产生的迭代序列收敛于问题 (1.1) 的鞍点. 在证明 之前,列举需要用到的假设及投影算子的一些性质,可在相关的文献 [11,12] 中查到. 假设 2.1 存在一个常数 L, 使得下面三个不等式成立

$$\begin{aligned} \|\nabla_x Q(x,y) - \nabla_x Q(\bar{x},y)\| &\leq L \|x - \bar{x}\|, \\ \|\nabla_y Q(x,y) - \nabla_y Q(\bar{x},y)\| &\leq L \|x - \bar{x}\|, \\ \|\nabla_y Q(x,y) - \nabla_y Q(x,\bar{y})\| &\leq L \|y - \bar{y}\|. \end{aligned}$$

引理 2.1 令 Ω 为非空闭凸集, $w, z \in \mathbb{R}^n, u \in \Omega$.

(i) 投影算子是非扩张的, 即 $||P_{\Omega}(w) - P_{\Omega}(z)|| \le ||w - z||$.

(ii) 令 G(z) 为凸可微函数, $w = P_{\Omega}(z - \theta \nabla G(z))$,那么下式成立

$$||w - u||^{2} \le ||z - u||^{2} - ||w - z||^{2} + 2\theta \langle \nabla G(z), u - w \rangle.$$

(iii) $||w - u||^2 \le ||z - u||^2 - ||w - z||^2 + 2\theta \langle \nabla G(w) - \nabla G(z), w - z \rangle + 2(G(u) - G(w)).$ 基于以上假设和引理,下面给出算法 (2.1) 的收敛性证明.

定理 2.1 如果 β, γ 满足 $\frac{1}{2\beta} - L - \gamma L^2 > 0$ 和 $\frac{1}{2\gamma} - L > 0$, 那么新方法 (2.1) 生成的迭代 序列 { (x_k, y_k) } 收敛于优化问题 (1.1) 的鞍点 (x^*, y^*).

证 对式 (2.1) 中的第三个式子应用引理 2.1 中的性质 (ii), 对任意的 $y \in Y$, 可得

$$\|y_{k+1} - y\|^2 \le \|y_k - y\|^2 - \|y_{k+1} - y_k\|^2 + 2\gamma \langle \nabla_y Q(x_{k+1}, y_k), y_{k+1} - y \rangle.$$
(2.2)

同理, 令 $y = y_{k+1}$, 由 (2.1) 中的第一个式子得到

$$\|y_{k+\frac{1}{2}} - y_{k+1}\|^2 \le \|y_k - y_{k+1}\|^2 - \|y_{k+\frac{1}{2}} - y_k\|^2 + 2\gamma \langle \nabla_y Q(x_k, y_k), y_{k+\frac{1}{2}} - y_{k+1} \rangle.$$
(2.3)

在式 (2.3) 中的右式加上和减去 $2\gamma \langle \nabla_y Q(x_{k+1}, y_k), y_{k+\frac{1}{2}} - y_{k+1} \rangle$,整理后有

$$\|y_{k+\frac{1}{2}} - y_{k+1}\|^{2} \leq \|y_{k} - y_{k+1}\|^{2} - \|y_{k+\frac{1}{2}} - y_{k}\|^{2} + 2\gamma \langle \nabla_{y}Q(x_{k+1}, y_{k}), y_{k+\frac{1}{2}} - y_{k+1} \rangle + 2\gamma \langle \nabla_{y}Q(x_{k}, y_{k}) - \nabla_{y}Q(x_{k+1}, y_{k}), y_{k+\frac{1}{2}} - y_{k+1} \rangle.$$

$$(2.4)$$

式 (2.2) 和式 (2.4) 相加后下式成立

$$\|y_{k+1} - y\|^{2} \leq \|y_{k} - y\|^{2} - \|y_{k+\frac{1}{2}} - y_{k}\|^{2} - \|y_{k+\frac{1}{2}} - y_{k+1}\|^{2} + 2\gamma \langle \nabla_{y}Q(x_{k+1}, y_{k}), y_{k+\frac{1}{2}} - y \rangle + 2\gamma \langle \nabla_{y}Q(x_{k}, y_{k}) - \nabla_{y}Q(x_{k+1}, y_{k}), y_{k+\frac{1}{2}} - y_{k+1} \rangle.$$

$$(2.5)$$

又因 Q 关于变量 y 是凹函数, $\langle \nabla_y Q(x_{k+1}, y_k), y - y_k \rangle \ge Q(x_{k+1}, y) - Q(x_{k+1}, y_k)$. 故式 (2.5) 的右边加上和减去 $2\gamma \langle \nabla_y Q(x_{k+1}, y_k), y_k \rangle$ 得

$$\begin{aligned} \|y_{k+1} - y\|^2 &\leq \|y_k - y\|^2 - \|y_{k+\frac{1}{2}} - y_k\|^2 - \|y_{k+\frac{1}{2}} - y_{k+1}\|^2 \\ &+ 2\gamma \langle \nabla_y Q(x_{k+1}, y_k), y_{k+\frac{1}{2}} - y_k \rangle + 2\gamma (Q(x_{k+1}, y_k) - Q(x_{k+1}, y)) \quad (2.6) \\ &+ 2\gamma \langle \nabla_y Q(x_k, y_k) - \nabla_y Q(x_{k+1}, y_k), y_{k+\frac{1}{2}} - y_{k+1} \rangle. \end{aligned}$$

现在考虑式 (2.1) 中的第二个式子, 利用引理 2.1 中的性质 (iii), 对任意的 $x \in X$, 有

$$\begin{aligned} \|x_{k+1} - x\|^2 &\leq \|x_k - x\|^2 - \|x_{k+1} - x_k\|^2 + 2\beta(Q(x, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y_{k+\frac{1}{2}})) \\ &+ 2\beta\langle \nabla_x Q(x_{k+1}, y_{k+\frac{1}{2}}) - \nabla_x Q(x_k, y_{k+\frac{1}{2}}), x_{k+1} - x_k \rangle. \end{aligned}$$
(2.7)

$$\begin{split} & \frac{1}{2\beta} \|x_{k+1} - x\|^2 + \frac{1}{2\gamma} \|y_{k+1} - y\|^2 \\ & \leq \frac{1}{2\beta} \|x_k - x\|^2 + \frac{1}{2\gamma} \|y_k - y\|^2 - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_k\|^2 - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k+1}\|^2 - \frac{1}{2\beta} \|x_{k+1} - x_k\|^2 \\ & + \langle \nabla_y Q(x_{k+1}, y_k), y_{k+\frac{1}{2}} - y_k \rangle + \langle \nabla_y Q(x_k, y_k) - \nabla_y Q(x_{k+1}, y_k), y_{k+\frac{1}{2}} - y_{k+1} \rangle \\ & + \langle \nabla_x Q(x_{k+1}, y_{k+\frac{1}{2}}) - \nabla_x Q(x_k, y_{k+\frac{1}{2}}), x_{k+1} - x_k \rangle \\ & + (Q(x_{k+1}, y_k) - Q(x_{k+1}, y)) + (Q(x, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y_{k+\frac{1}{2}})). \end{split}$$

由 Q 关于变量 y 为凹函数, 可知

$$Q(x_{k+1}, y_k) - \langle \nabla_y Q(x_{k+1}, y_{k+\frac{1}{2}}), y_k - y_{k+\frac{1}{2}} \rangle \le Q(x_{k+1}, y_{k+\frac{1}{2}}).$$

不等式 (2.8) 进一步转化为

$$\frac{1}{2\beta} \|x_{k+1} - x\|^{2} + \frac{1}{2\gamma} \|y_{k+1} - y\|^{2} \\
\leq \frac{1}{2\beta} \|x_{k} - x\|^{2} + \frac{1}{2\gamma} \|y_{k} - y\|^{2} - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k}\|^{2} - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k+1}\|^{2} - \frac{1}{2\beta} \|x_{k+1} - x_{k}\|^{2} \\
+ \langle \nabla_{y} Q(x_{k+1}, y_{k}) - \nabla_{y} Q(x_{k+1}, y_{k+\frac{1}{2}}), y_{k+\frac{1}{2}} - y_{k} \rangle \\
+ \langle \nabla_{y} Q(x_{k}, y_{k}) - \nabla_{y} Q(x_{k+1}, y_{k}), y_{k+\frac{1}{2}} - y_{k+1} \rangle \\
+ \langle \nabla_{x} Q(x_{k+1}, y_{k+\frac{1}{2}}) - \nabla_{x} Q(x_{k}, y_{k+\frac{1}{2}}), x_{k+1} - x_{k} \rangle + Q(x, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y).$$
(2.9)

利用 Cauchy-Schwartz 不等式 (2.9) 写为

$$\frac{1}{2\beta} \|x_{k+1} - x\|^{2} + \frac{1}{2\gamma} \|y_{k+1} - y\|^{2}
\leq \frac{1}{2\beta} \|x_{k} - x\|^{2} + \frac{1}{2\gamma} \|y_{k} - y\|^{2} - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k}\|^{2} - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k+1}\|^{2} - \frac{1}{2\beta} \|x_{k+1} - x_{k}\|^{2}
+ \|\nabla_{y}Q(x_{k+1}, y_{k}) - \nabla_{y}Q(x_{k+1}, y_{k+\frac{1}{2}})\|\|y_{k+\frac{1}{2}} - y_{k}\|
+ \|\nabla_{y}Q(x_{k}, y_{k}) - \nabla_{y}Q(x_{k+1}, y_{k})\|\|y_{k+\frac{1}{2}} - y_{k+1}\|
+ \|\nabla_{x}Q(x_{k+1}, y_{k+\frac{1}{2}}) - \nabla_{x}Q(x_{k}, y_{k+\frac{1}{2}})\|\|x_{k+1} - x_{k}\| + Q(x, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y).$$
(2.10)

由引理 2.1 中的性质 (i) 知

$$\|y_{k+1} - y_{k+\frac{1}{2}}\| \le \gamma \|\nabla_y Q(x_{k+1}, y_k) - \nabla_y Q(x_k, y_k)\|.$$
(2.11)

再由假设 2.1 和不等式 (2.11) 得到下式

$$\frac{1}{2\beta} \|x_{k+1} - x\|^2 + \frac{1}{2\gamma} \|y_{k+1} - y\|^2
\leq \frac{1}{2\beta} \|x_k - x\|^2 + \frac{1}{2\gamma} \|y_k - y\|^2 - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_k\|^2 - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k+1}\|^2 - \frac{1}{2\beta} \|x_{k+1} - x_k\|^2
+ L \|y_{k+\frac{1}{2}} - y_k\|^2 + \gamma L^2 \|x_{k+1} - x_k\|^2 + L \|x_{k+1} - x_k\|^2
= \frac{1}{2\beta} \|x_k - x\|^2 + \frac{1}{2\gamma} \|y_k - y\|^2 - \frac{1}{2\gamma} \|y_{k+\frac{1}{2}} - y_{k+1}\|^2
- (\frac{1}{2\beta} - L - \gamma L^2) \|x_{k+1} - x_k\|^2 - (\frac{1}{2\gamma} - L) \|y_{k+\frac{1}{2}} - y_k\|^2 + Q(x, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y).$$
(2.12)

令 $x = x^*, y = y^*,$ 且由鞍点性质 $Q(x^*, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y^*) \le 0$ 和 β, γ 满足 $\frac{1}{2\beta} - L - \gamma L^2 > 0$ 和 $\frac{1}{2\gamma} - L > 0$,可知

$$\frac{1}{2\beta} \|x_{k+1} - x^*\|^2 + \frac{1}{2\gamma} \|y_{k+1} - y^*\|^2 \le \frac{1}{2\beta} \|x_k - x^*\|^2 + \frac{1}{2\gamma} \|y_k - y^*\|^2.$$
(2.13)

(2.8)

由文献 [13] 的结论可证序列 $\{(x_k, y_k)\}$ 收敛于优化问题 (1.1) 的鞍点 (x^*, y^*) .

接下来分析算法 (2.1) 的次线性收敛速率. 为了确定算法的复杂性, 需要一些额外的记号. 令 $N \ge 1$ 为正整数, $x^N = \frac{1}{N+1} \sum_{k=0}^{N} x_{k+1}, y^N = \frac{1}{N+1} \sum_{k=0}^{N} y_{k+\frac{1}{2}}.$ 定理 2.2 如果 β, γ 满足 $\frac{1}{2\beta} - L - \gamma L^2 > 0$ 和 $\frac{1}{2\gamma} - L > 0$, 那么

$$Q(x^N, y) - Q(x, y^N) \le \frac{1}{N+1} (\frac{1}{2\beta} ||x - x_0||^2 + \frac{1}{2\gamma} ||y - y_0||^2).$$

证 由式 (2.12) 可得到

$$\frac{1}{2\beta} \|x_{k+1} - x\|^2 + \frac{1}{2\gamma} \|y_{k+1} - y\|^2 \le \frac{1}{2\beta} \|x_k - x\|^2 + \frac{1}{2\gamma} \|y_k - y\|^2 + Q(x, y_{k+\frac{1}{2}}) - Q(x_{k+1}, y).$$
(2.14)

对式 (2.14) 从 $k = 0, 1, \dots, N$ 相加, 那么

$$\frac{1}{2\beta} \|x_{N+1} - x\|^2 + \frac{1}{2\gamma} \|y_{N+1} - y\|^2 + \sum_{k=0}^{N} (Q(x_{k+1}, y) - Q(x, y_{k+\frac{1}{2}})) \\
\leq \frac{1}{2\beta} \|x_0 - x\|^2 + \frac{1}{2\gamma} \|y_0 - y\|^2.$$
(2.15)

再由函数 Q(x,y) 的凸凹性和 Jensen 不等式可知

$$Q(x^{N}, y) - Q(x, y^{N}) \le \frac{1}{N+1} \sum_{k=0}^{N} (Q(x_{k+1}, y) - Q(x, y_{k+\frac{1}{2}})).$$
(2.16)

结合式 (2.15) 和 (2.16), 易知要证的结论成立.

3 数值实验

本节给出泊松噪音下图像去模糊的数值表现. 令 $g \in R^m$ 是观测数据, 每个观测值 g_i 由含有泊松随机变量的 $(Hx + b)_i$ 期望值确定, 其中 x 是原始图像, H 可认为是采集系统造成的失真模糊矩阵, $b \in R^m$ 是一个正的常数背景项. 在泊松噪音的假设下, Kullback–Leibler 函数经常用来作为数据拟合项. 由于问题的病态性, 选取能保持图像边缘的全变分^[14] 作为正则项. 令外, 也可以增加一些物理约束, 如像素的非负性. 综上, 图像去模糊问题转化为如下的最优化问题

$$\min_{x \ge 0} \sum_{i} \{g_i \ln \frac{g_i}{(Hx+b)_i} + (Hx+b)_i - g_i\} + \alpha \sum_{i}^{N} \|(\nabla x)_i\|,$$
(3.1)

其中 $(\nabla x)_i$ 是 x 在像素 i 处的梯度. 如果 $g_i = 0$, 那么 $g_i \ln g_i = 0$.

模型 (3.1) 是非光滑凸优化问题, 直接求解有一定的难度. 利用全变分函数的对偶公式, 可转化为光滑的鞍点优化问题

$$\min_{x \ge 0} \max_{|y| \le 1} \sum_{i} \{g_i \ln \frac{g_i}{(Hx+b)_i} + (Hx+b)_i - g_i\} + \alpha y^T \nabla x.$$
(3.2)

易知模型 (3.2) 是问题 (1.1) 的特殊形式, 且关于原始变量 x 为凸的, 关于对偶变量 y 是 凹的. 当 Q(x,y) 由式 (3.2) 定义, 其梯度表达式为 $\nabla_x Q(x,y) = e_n - H^T Z(x)^{-1}g - \alpha \nabla$. $y, \nabla_y Q(x,y) = \alpha \nabla x, 其中 e_n = (1,1,\dots,1)^T \in R^n, \nabla \cdot$ 是散度算子; Z(x) 是对角矩阵, 对 角元素由 $(Hx + b)_i$ 给定. 当 $b \neq 0$ 时, 由文献 [12] 中引理 4.6 知 $\nabla_x Q(x,y)$ 李布希兹连续, 可知该问题满足假设 2.1. 故所提算法可以应用于问题 (3.2).

记所提算法为算法1,对比算法为文献[4]中的交替外梯度方法 (AEM),可在 www.unife.it /prin/software 下载,以及算法 (1.3),按照文献 [10] 建议记为 LPD. 计算环境如下, Matlab 版本为 R2011a, 内存为 2GB, 处理器为 i3 的个人电脑. 在两个实验中,图像分别为 128 × 128 的 micro 数据和 256 × 256 的 circles 数据,详细可见文献 [4] 的说明. 背景数据项均为 b = 3. 根据文献 [12] 知,李布希兹常数为 $\frac{\|g\|_{\infty} \|H\|^2}{b^2}$. 由此可计算出 *L* 在两个问题中的值分别为 0.14 和 0.1. 据此,测试了满足定理 2.1 条件下的不同参数,选取了恢复效果比较好的参数,具体如 下:在第一个实验中,算法 1 参数设置为 $\alpha = 0.09$, $\beta = 1.2$, $\gamma = 3.5$;第二个实验中,参数设置为 $\alpha = 0.25$, $\beta = 1.3$, $\gamma = 1.6$.其它两个算法按其文章里的建议设置. 三个算法的停止准则 为

$$\left\| \left(\begin{array}{c} x_{k+1} - x_k \\ y_{k+1} - y_k \end{array} \right) \right\| / \left\| \left(\begin{array}{c} x_{k+1} \\ y_{k+1} \end{array} \right) \right\| \le 0.0001.$$

为了评价算法恢复图像的质量,采用相对误差 (RE) 作为指标,定义为 $RE = \frac{\|x_k - x\|}{\|x\|}$,其中 x_k 为算法恢复的图像, x为原始图像. 它能衡量算法恢复的图像与真实图像的接近程度. 显然,相对误差的值越小,图像恢复的质量越好.

衣 I				
	算法	迭代次数	时间 (秒)	相对误差
micro	AEM	369	3.3076	0.808
	LPD	432	3.5316	0.1099
	算法1	388	3.1702	0.807
circles	AEM	156	7.3290	0.0441
	LPD	245	9.9959	0.0410
	算法1	141	5.7923	0.0439

表1 数值结果



图 1 micro 图像恢复效果对比图:真实图像,模糊噪音图像,AEM 恢复,LPD 恢复,算法 1 恢复



图 2 circles 图像恢复效果对比图: 真实图像, 模糊噪音图像, AEM 恢复, LPD 恢复, 算法 1 恢复



图 3 目标函数随时间变化曲线图.

表1给出了算法在两个含有泊松噪音数据上的去模糊表现.在micro问题中,虽然算法1的迭代步数多于AEM,但发现计算时间稍微少于AEM.这说明由线搜索确定的步长在一定程度上会消耗一点时间.LPD 算法无论是迭代步数,计算时间和相对误差均不占优势.在 circles 问题中,无论迭代步数和计算时间都少于AEM.LPD 方法虽然相对误差较小,但计算步数较多,计算时间较长.图1和图2显示了两个算法恢复的效果图,可以看到算法基本上都能很好的恢复模糊的图像.这也可以从表1的评价标准相对误差中得到,因为相对误差的值相差不大.为更清楚地说明算法的表现,在图3中画出了目标函数值随计算时间的曲线图.从曲线下降趋势看,算法1在迭代的前几步或十几步函数值下降快于AEM,随后两个算法函数值基本保持不变.表明算法1可以更快的收敛于最优解.虽然算法LPD 的目标函数值在 micro问题中下降较快,但它是在目标函数(3.2)在 *X* = *R*ⁿ 的情况下.由于无相关物理条件下的约束,故其相对误差较大.这也说明了如果有先验知识的加入,如非负约束,则可能得到较好的恢复结果.最后,和相关算法对比的数值结果表明了新方法是有效的.

4 结语

对带有凸集约束的光滑鞍点优化问题,提出了一个简单的原始 – 对偶梯度方法. 算法每步具有显示解,不需要内迭代求解子问题. 新方法应用在全变分正则化泊松噪音图像恢复问题上,结果表明在计算时间上具有一定的优势. 值得注意的是该方法也可以拓展求解其它噪音类型的图像恢复问题.

参考文献

- [1] Harker P T, Pang J S. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications[J]. Math. Program, 1990, 48: 161–220.
- [2] Xiu N, Zhang J. Some recent advances in projection-type methods for variational inequalities[J]. J. Comput. Appl. Math., 2003, 152(1): 559–585.
- [3] Patriksson M. Nonlinear programming and variational inequality problems: a unified approach[M]. New York: Springer Science Business Media, 2013.
- [4] Bonettini S, Ruggiero V. An alternating extragradient method for total variation-based image restoration from Poisson data[J]. Inverse. Probl., 2011, 27(9): 095001.

- [5] Zhu M, Chan T. An efficient primal-dual hybrid gradient algorithm for total variation image restoration[R]. UCLA CAM Report, 2008, 34.
- [6] He B, You Y, Yuan X. On the convergence of primal-dual hybrid gradient algorithm[J]. Siam J. Imag. Sci., 2014, 7(4): 2526–2537.
- [7] Bonettini S, Ruggiero V. On the convergence of primal-dual hybrid gradient algorithms for total variation image restoration[J]. J. Math. Imaging. Vis., 2012, 44(3): 236–253.
- [8] Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging[J]. J. Math. Imaging. Vis., 2011, 40(1): 120–145.
- [9] Chen P, Huang J, Zhang X. A primal-dual fixed point algorithm for convex separable minimization with applications to image restoration[J]. Inverse. Probl., 2013, 29(2): 025011.
- [10] Zhang B, Zhu Z, Wang S. A simple primal-dual method for total variation image restoration[J]. J. Vis. Comm. Image Repre., 2016, 38: 814–823.
- [11] Calamai P H, MoréJ J. Projected gradient methods for linearly constrained problems[J]. Math. Program, 1987, 39(1): 93–116.
- [12] Krol A, Li S, Shen L. Preconditioned alternating projection algorithms for maximum a posteriori ECT reconstruction[J]. Inverse. Probl., 2012, 28(11): 115005.
- [13] Rockafellar R T. Monotone operators and the proximal point algorithm[J]. SIAM J. Control Optim., 1976, 14(5): 877–898.
- [14] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. Physica D., 1992, 60: 259–268.

PROJECTION PRIMAL-DUAL METHOD FOR SADDLE POINT OPTIMIZATION PROBLEMS WITH SIMPLE CONSTRAINED

GUO Xiao¹ , TAO Yue²

(1.School of Material Science and Engineering, Guilin University of Electronic Technology, Guilin 541004, China)

(2. Jiangsu Broadcasting Corporation, Nanjing 210000, China)

Abstract: In this paper, we study the saddle point optimization problems with simple convex set constraints. Using the property of convex-concave, a projection primal-dual gradient method is proposed. The algorithm has a symmetric structure and each step has a closed-form solution. The convergence of the new algorithm is proved and the convergence rate is also obtained. The numerical results of Poisson noise image restoration problem verify the effectiveness of the new method.

Keywords: saddle point problem; primal-dual; image restoration **2010 MR Subject Classification:** 65K10; 90C25