第 32 卷第 4 期　　重庆工商大学学报(自然科学版)　　2015 年 4 月
Vol. 32　NO.4　　J Chongqing Technol Business Univ. (Nat Sci Ed)　　Apr.2015

doi:10.16055/j.issn.1672-058X.2015.0004.007

解凸优化问题的一类修正线性近似交替方向法

李　慧

(重庆师范大学 数学学院,重庆 401331)

摘　要:在解凸优化问题过程中,对已有文献的线性约束条件推广到非线性约束条件,运用了近似交替分解算法;新提出一类修正线性近似交替方向法,并进行了理论分析和和算例比较.

关键词:近似交替方向法;非线性约束凸规划问题;可分化方法;线性化;增广拉格朗日

中图分类号:O221　　**文献标识码**:A　　**文章编号**:1672-058X(2015)04-0023-05

1　预备知识

介绍以下原始凸规划问题

$$\text{(P)} \ \min \ \theta_1(x) + \theta_2(z) \qquad (1)$$
$$\text{s.t.} \ \boldsymbol{A}x + \boldsymbol{B}z = \boldsymbol{c}, \ x \in X, z \in Z$$

其中 $\theta_1:\mathbf{R}^n \to (-\infty,+\infty]$,$\theta_2:\mathbf{R}^s \to (-\infty,+\infty]$是闭真凸函数,$\boldsymbol{A}$ 是 $m\times n$ 矩阵,$\boldsymbol{B}$ 是 $m\times s$ 矩阵;$X\subseteq\mathbf{R}^n$,$Z\subseteq\mathbf{R}^s$ 是闭凸集;$\boldsymbol{c}$ 是 m 维给定向量.问题(P)的拉格朗日函数 $L:\mathbf{R}^n\times\mathbf{R}^s\times\mathbf{R}^m\to(-\infty,+\infty]$

$$L(x,z,y) = \theta_1(x) + \theta_2(z) - \langle y, G(x,z) \rangle$$

其中$\langle\cdot,\cdot\rangle$表示内积,$\|\cdot\|$表示欧几里得范数,$G(x,z)=\boldsymbol{A}x+\boldsymbol{B}z-\boldsymbol{c}$,$y$ 是约束 $\boldsymbol{A}x+\boldsymbol{B}z=\boldsymbol{c}$ 的相应拉格朗日乘子.问题(P)的增广拉格朗日函数为

$$L_\lambda(x,z,y) = \theta_1(x) + \theta_2(z) - \langle y, G(x,z) \rangle + \frac{\lambda}{2}\|G(x,z)\|^2$$

Peaceman 和 Rochford 在文献[1-3]中首次提出交替方向法(ADM)思想;D.Gabay 和 B.Mercier 在文献[4],R.Glowinsk 和 P.L.Tallec 在文献[5]中分别提出了问题(P)的一个算法,它对增广拉格朗日函数先关于 x 求极小,其次关于 z 求极小,最后更新乘子 y;1991 年,P.Tseng 在文献[6]提出了求解问题(P)的另一种方法,用增广拉格朗日函数 $L_\lambda(x,z,y)$ 代替一般的拉格朗日函数 $L(x,z,y)$ 对 x 极小化;Eckstein 和 Bersekas 在文献[7]中将 Douglas-Rachford 分割方法和邻近点算法建立了联系;1994 年,Chen 和 Teboulle 在文献[8]中提出了一种可分方法,它关于对偶变量进行了两次迭代;Fukushima 在文献[9]中提出把 ADM 应用到凸规划的对偶问题上解决带可分结构的凸规划问题;Paper 在文献[10]中认为还可以通过 ADM 方法有效解决大规模最小二乘法半定规划问题.虽然以上有些算法是全局收敛的,但在实际迭代中,确保线性收敛的前提条件太强且很难满足.为了解决这个问题,文献[11]在交替方向法(Alternating Direction Method,记为 ADM)中的前两步即产生 x^{k+1},z^{k+1} 的过程中分别添加近似二次项 $\frac{\alpha}{2}\|x-x^k\|^2$,$\frac{\beta}{2}\|z-z^k\|^2$,简称为近似交替方向法(Proximal Alternating Direction Method),记为 PADM.

收稿日期:2014-06-03;**修回日期**:2014-09-19.

作者简介:李慧(1990-),女,四川广安人,硕士研究生,从事最优化理论与方法研究.

PADM 算法

初始步:选一个初始点 $M^0=(x^0,z^0,y^0)\in\mathbf{R}^{n+s+m}$,一个正数 $\varepsilon>0,\alpha>0,\beta>0,\gamma\in(0,\frac{\sqrt{5}+1}{2})$,令 $k=0$;

第 1 步:计算

$$x^{k+1}=\arg\min_{x\in X}\{\theta_1(x)-<y^k,G(x,z^k)>+\frac{1}{2}\|G(x,z^k)\|^2+\frac{\alpha}{2}\|x-x^k\|^2\} \tag{2}$$

$$z^{k+1}=\arg\min_{z\in Z}\{\theta_2(z)-<y^k,G(x^{k+1},z)>+\frac{1}{2}\|G(x^{k+1},z)\|^2+\frac{\beta}{2}\|z-z^k\|^2\} \tag{3}$$

$$y^{k+1}=y^k-\gamma G(x^{k+1},z^{k+1}) \tag{4}$$

第 2 步:若 $\|M^{k+1}-M^k\|<\varepsilon$,则停止,$M^{k+1}$ 就是近似解;否则令 $k=k+1$,并返回到第 1 步.

推广模型如下:

$$(\mathrm{P}')\qquad \begin{aligned}&\min\ \theta_1(x)+\theta_2(z)\\ &\text{s.t. } g_1(x)+g_2(z)=b,\ x\in X,z\in Z\end{aligned} \tag{5}$$

其中 $\theta_1:\mathbf{R}^n\to(-\infty,+\infty]$,$\theta_2:\mathbf{R}^s\to(-\infty,+\infty]$ 是闭真凸函数,$g_1:\mathbf{R}^n\to\mathbf{R}$,$g_2:\mathbf{R}^S\to\mathbf{R}$ 是一阶连续可微凸函数,$X\subseteq\mathbf{R}^n$,$Z\subseteq\mathbf{R}^s$ 是闭凸集,b 是给定的常数.

对于这种数学规划问题,提出了一类新的可分算法.在算法中,把 PADM 算法中的二次项依次进行线性化,然后将上一次迭代获得的起始点(x^k,z^k,y^k)和最优点(u^{k+1},v^{k+1},w^{k+1})的凸组合作为下一次迭代的起始点(x^{k+1},z^{k+1},y^{k+1}),称此方法为修正线性近似交替方向法(Modified Linear Proximal Alternating Direction Method),记为 MLPADM.在第 2 部分介绍一类修正线性近似交替方向法;第三部分应用一类修正线性近似交替方向法进行数值实验;最后通过数据比较得出结论.

2 方法介绍

结合以上提出的 PADM 方法,对带有非线性约束的凸规划问题,提出了以下 3 种修正线性近似交替方向法.

第一种修正线性近似交替方向法是把式(2)中的二次项 $\frac{1}{2}\|G(x,z^k)\|^2$ 在 x^k 处线性化,然后将上一次迭代获得的起始点和最优点的凸组合作为下一次迭代的起始点.

因为

$$g_1(x)\approx g_1(x^k)+\nabla g_1(x^k)\cdot(x-x^k)$$

所以

$$\begin{aligned}\|G(x,z^k)\|^2=&\|g_1(x)+g_2(z^k)-b\|^2\approx\\ &\|g_1(x^k)+\nabla g_1(x^k)\cdot(x-x^k)+g_2(z^k)-b\|^2=\\ &\|\nabla g_1(x^k)\cdot(x-x^k)+[g_1(x^k)+g_2(z^k)-b]\|^2=\\ &[\nabla g_1(x^k)\cdot(x-x^k)]^2+2\nabla g_1(x^k)\cdot(x-x^k)\cdot[g_1(x^k)+g_2(z^k)-b]+\cdots\\ &[g_1(x^k)+g_2(z^k)-b]^2\end{aligned}$$

因此线性化二次项 $\frac{1}{2}\|G(x,z^k)\|^2$,去掉常数项,就能以近似的式(6)来替代

$$\nabla g_1(x^k)(x-x^k)\cdot[g_1(x^k)+g_2(z^k)-b] \tag{6}$$

MLPADM 算法 1

初始步:选一个初始点 $M^0=(x^0,z^0,y^0)\in\mathbf{R}^{n+s+l}$,一个正数 $\varepsilon>0,\eta\in(0,1),\varphi_1^0\in\nabla g_1(x^0),\alpha>0,\beta>0,\gamma\in\left(0,\frac{\sqrt{5}+1}{2}\right)$,令 $k=0$;

第 1 步:计算

$$u^{k+1}=\arg\min_{x\in X}\left\{\theta_1(x)-<y^k,G(x,z^k)>+\varphi_1^k(x-x^k)(g_1(x^k)+g_2(z^k)-b)+\frac{\alpha}{2}\|x-x^k\|^2\right\}$$

$$v^{k+1}=\arg\min_{z\in Z}\left\{\theta_2(z)-<y^k,G(u^{k+1},z)>+\frac{1}{2}\|G(x^k,z)\|^2+\frac{\beta}{2}\|z-z^k\|^2\right\}$$

$$w^{k+1}=y^k-\gamma G(u^{k+1},v^{k+1})$$

第 2 步:$(x^{k+1},z^{k+1},y^{k+1})=\eta(x^k,z^k,y^k)+(1-\eta)(u^{k+1},v^{k+1},w^{k+1})$;

第 3 步:若$\|M^{k+1}-M^k\|<\varepsilon$,则停止,$M^{k+1}$就是近似解;否则令 $k=k+1$,并返回到第 1 步.

第 2 种修正线性近似交替方向法是把式(3)中的二次项$\frac{1}{2}G(x^{k+1},z)^2$ 在 z^k 处线性化,然后将整个算法由上一次迭代获得的起始点和最优点的凸组合作为下一次迭代的起始点.

因为

$$g_2(z)\approx g_2(z^k)+\nabla g_2(z^k)\cdot(z-z^k)$$

所以

$$\begin{aligned}\|G(x^{k+1},z)\|^2&=\|g_1(x^{k+1})+g_2(z)-b\|^2\approx\\&\|g_1(x^{k+1})+\nabla g_2(z^k)\cdot(z-z^k)+g_2(z^k)-b\|^2=\\&\|\nabla g_2(z^k)\cdot(z-z^k)+[g_1(x^{k+1})+g_2(z^k)-b]\|^2=\\&[\nabla g_2(z^k)\cdot(z-z^k)]^2+2\nabla g_2(z^k)\cdot(z-z^k)\cdot[g_1(x^{k+1})+g_2(z^k)-b]+\cdots\\&[g_1(x^{k+1})+g_2(z^k)-b]^2\end{aligned}$$

因此线性化二次项$\frac{1}{2}\|G(x^{k+1},z)\|^2$,去掉常数项,就可以近似地用式(7)来替代

$$\nabla g_2(z^k)(z-z^k)\cdot[g_1(x^{k+1})+g_2(z^k)-b]\tag{7}$$

MLPADM **算法** 2

初始步:选一个初始点 $M^0=(x^0,z^0,y^0)\in\mathbf{R}^{n+s+l}$,一个正数 $\varepsilon>0,\eta\in(0,1),\varphi_2^0\in\nabla g_2(z^0),\alpha>0,\beta>0,\gamma\in\left(0,\frac{\sqrt{5}+1}{2}\right)$,令 $k=0$;

第 1 步:计算

$$u^{k+1}=\arg\min_{x\in X}\left\{\theta_1(x)-<y^k,G(x,z^k)>+\frac{1}{2}\|G(x,z^k)\|^2+\frac{\alpha}{2}\|x-x^k\|^2\right\}$$

$$v^{k+1}=\arg\min_{z\in Z}\left\{\theta_2(z)-<y^k,G(u^{k+1},z)>+\varphi_2^k(z-z^k)(g_1(u^{k+1})+g_2(z^k)-b)+\frac{\beta}{2}\|z-z^k\|^2\right\}$$

$$w^{k+1}=y^k-\gamma G(u^{k+1},v^{k+1})$$

第 2 步:$(x^{k+1},z^{k+1},y^{k+1})=\eta(x^k,z^k,y^k)+(1-\eta)(u^{k+1},v^{k+1},w^{k+1})$;

第 3 步:若$\|M^{k+1}-M^k\|<\varepsilon$,则停止,$M^{k+1}$就是近似解;否则令 $k=k+1$,并返回到第 1 步.

第 3 种修正线性近似交替方向法是同时把式(2)和式(3)中的二次项$\frac{1}{2}\|G(x,z^k)\|^2$,$\frac{1}{2}\|G(x^{k+1},z)\|^2$ 线性化,即分别用$\nabla g_1(x^k)(x-x^k)\cdot[g_1(x^k)+g_2(z^k)-b]$和$\nabla g_2(z^k)(z-z^k)\cdot[g_1(x^{k+1})+g_2(z^k)-b]$来替代,然后将整个算法由上一次迭代获得的起始点和最优点的凸组合为下一次迭代的起始点.具体算法如下:

MLPADM **算法** 3

初始步:选一个初始点 $M^0=(x^0,z^0,y^0)\in\mathbf{R}^{n+s+l}$,一个正数 $\varepsilon>0,\eta\in(0,1),\varphi_1^0\in\partial g_1(x^0),\varphi_2^0\in\nabla g_2(z^0)$,$\alpha>0,\beta>0,\gamma\in\left(0,\frac{\sqrt{5}+1}{2}\right)$,令 $k=0$;

第 1 步:计算

$$u^{k+1}=\arg\min_{x\in X}\left\{\theta_1(x)-<y^k,G(x,z^k)>+\varphi_1^k(x-x^k)(g_1(x^k)+g_2(z^k)-b)+\frac{\alpha}{2}\|x-x^k\|^2\right\}$$

$$v^{k+1}=\arg\min_{z\in Z}\left\{\theta_2(z)-<y^k,G(u^{k+1},z)>+\varphi_2^k(z-z^k)(g_1(u^{k+1})+g_2(z^k)-b)+\frac{\beta}{2}\|z-z^k\|^2\right\}$$

$$w^{k+1}=y^k-\gamma G(u^{k+1},v^{k+1})$$

第 2 步:$(x^{k+1},z^{k+1},y^{k+1})=\eta(x^k,z^k,y^k)+(1-\eta)(u^{k+1},v^{k+1},w^{k+1})$

第 3 步:若$\|M^{k+1}-M^k\|<\varepsilon$,则停止,$M^{k+1}$就是近似解;否则令 $k=k+1$,并返回到第 1 步.

3 数值试验

对于各个算例,分别通过 PADM 和此处提出的 MLPADM 算法 1,算法 2,算法 3 进行比较求解,数值结果记录在表 1 中.其中 $M^0=(x^0,z^0,y^0)$为任意选取的初始点,$M^*=(x^*,z^*,y^*)$代表原始问题的近似最优点,$\theta(x^*,z^*)=\theta_1(x^*)+\theta_2(z^*)$为近似最优值,Iter 代表迭代次数,$t$/s 代表 cpu 运算时间.

例 1 $\min\ 10x^2-x+10z^2-10$

$\text{s.t.}\ x^2+z^2=1$

由文献可知最优点为(1,0),最优值为−1.

表 1 例 1 的 4 种算法比较

	M^0	M^*	$\theta(x^*,z^*)$	Iter	t/s
MPAD	0.7 0.1 9	0.999 998 671 937 88 0.000 006 130 285 08 9.499 998 058 839 79	−1.000 025 232 786 87	13	0.390 000
PLMPAD 算法 1	0.7 0.1 9	0.999 998 169 430 7 0.000 000 842 592 43 9.500 015 348 023 98	−1.000 034 780 731 03	30	0.687 000
PLMPAD 算法 2	0.7 0.1 9	0.999 998 707 355 64 0.000 008 625 005 96 9.499 998 039 851 92	−1.000 024 559 482 25	13	0.328 000
PLMPAD 算法 3	0.7 0.1 9	0.999 998 104 603 17 0.000 000 493 909 40 9.500 015 329 250 08	−1.000 036 012 501 47	30	0.702 000

通过表 1 可以看出,从迭代次数、运算时间以及最优值比较,显然 PMPAD 算法 2 效果好一点.

例 2 $\min\ -12x-7y+y^2$

$\text{s.t.}\ -2x^4+2-y=0,0\leqslant x\leqslant 2,0\leqslant y\leqslant 3$

例 2 最优点为(0.717 51,1.470),最优值为−16.738 89.

表 2 例 2 的 4 种算法比较

	M^0	M^*	$\theta(x^*,z^*)$	Iter	t/s
MPAD	1 1.5 5	0.717 536 170 914 14 1.469 839 467 769 92 4.060 318 441 096 99	−16.738 882 264 345 00	19	0.437 000
PLMPAD 算法 1	1 1.5 5	0.717 536 164 916 65 1.469 839 510 981 15 4.060 318 435 837 67	−16.738 882 367 826 50	18	0.390 000
PLMPAD 算法 2	1 1.5 5	0.717 535 097 674 65 1.469 839 332 234 74 4.060 329 907 321 35	−16.738 888 752 547 9	18	0.404 000

续表

	M^0	M^*	$\theta(x^*, z^*)$	Iter	t/s
PLMPAD 算法 3	1	0.717 536 271 010 99	−16.738 888 750 576 58	19	0.405 000
	1.5	1.469 844 898 756 41			
	5	4.060 312 298 545 41			

通过表 2 观察分析,显然这 4 种算法中 PMPAD 算法 2 效果好一点.

4　结　论

从上面的数值结果可知,对于同一个问题,MPAD 算法与 PLMPAD 算法 1,PLMPAD 算法 2,PLMPAD 算法 3 比较,PMPAD 算法 2 效果好一点.从以上计算结果来看,算法对求解带非线性约束的规划问题是有效的.在计算过程中发现,只有松弛因子 γ 在 $\left(0,\dfrac{\sqrt{5}+1}{2}\right)$ 中才有效,并且大于等于 1 的效果更好.

参考文献:

[1] PEACEMAN D W,RACJFORD H H.The Numerical Solution of Parabolic and Elliptic Differential Equations[J].J Soc Ind Appl Math ,1955(3):28-41

[2] DOUGLAS J,RACHFORD H H.On the Numerlcal Solution of the Heat Conduction Problem in 2 and 3 Space Variables Trans[J].Am Math Soc,1956(82):421-439

[3] DOUGLAS J.GUNN J E.A General Formulation of Alternating Direction Methods,Part I Parabolic and Hyperbolic Problems[J].Numer Math,1964(6):428-453

[4] GABAY D,MERCIAR B.A Dual Algorithm for the Solution of Nonlinear Variation Problems via Finite-element Approximations[J].Computers and Mathematics with Application,1976(2):17-40

[5] GLOWINSKI R,LETALLEC P.Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics[M].SIAM Studies in Applied Mathematics,SIAM,Philadelphia,PA, 1989

[6] TSENG P.Application of a Splitting Algorithm to Decomposition in Convex Programming and Cariational Inequalities [J].SIAM J Control and Opti,1991,29(1):119-138

[7] ECKSTEIN J,BERTSEKAS D P.On the Douglas-Rachford Splitting Menthod and Proximal Point Algorithm for Maximal Monotone Operators[J].Math Program,1992(55): 293-318

[8] GONG C,TEBOULLE M. A Proximal-based Decomposition Method for Convex Minimization Problems[J].1994,64(1-3):81-101

[9] 徐以汎,吴芳.A Decomposition Method for Convex Minization Problems and its Application[J].Acta Mathematice Application sinca,2001(17):20-28

[10] HE B S,XU M H,YUAN X M.Solving Large-scale Least Squares Semidefinite Programming by Alternating Direction Method[J].SIAM J Matrix Anal Appl,2011,32(1):136-152

[11] XU M H,WU T.A Class of Linearized Proximal Alternating Direction Methods[J].Optim Theory Appl,2011,151(2):321-337

A Class of Modified Linear Proximal Alternating Direction Methods for Convex Optimization

LI Hui

(College of Mathematics,Chongqing Normal University, Chongqing 401331,China)

Abstract: In this paper, the liner constraints of the original literature is extended to nonlinear constraints. Approximate Solution alternating decomposition algotithm is used, and a new type of nwdified linear proximal alternating direction method is proposed with theoretical analyses and algorithm comparison.

Key words: Proximal alternating direction method; Nonlinear constrained Convex programming; Decomposition methods; Linearized; Augmented Lagrangian