Advances in Applied Mathematics 应用数学进展, 2015, 4(3), 271-276
Published Online August 2015 in Hans. http://www.hanspub.org/journal/aam
http://dx.doi.org/10.12677/aam.2015.43033

Hans汉斯

Smoothing Inexact Newton Method for the Second Order Cone Programming

Li Dong[1], Siqi Xu[2], Jingen Yang[1]

[1]College of Mathematics and Information Science, Xinyang Normal University, Xinyang Henan
[2]Department of Civil Engineering, Xinyang Normal University, Xinyang Henan
Email: dongli9261128@163.com

Received: Jul. 30th, 2015; accepted: Aug. 12th, 2015; published: Aug. 18th, 2015

Abstract

A new smoothing inexact Newton method is presented for solving the second-order cone programming. At each iteration, the method uses an inexact Newton method to solve the system of equations, which saves computation work of smoothing Newton methods. Under weak assumptions, our method is proved to have global and local quadratic convergence. Numerical experiments indicate that the proposed method is quite effective.

Keywords

Second-Order Cone Programming, Smoothing Inexact Newton Method, Convergence

二阶锥规划的光滑非精确牛顿法

董　丽[1]，徐思齐[2]，杨金根[1]

[1]信阳师范学院数学与信息科学学院，河南 信阳
[2]信阳师范学院土木工程学院，河南 信阳
Email: dongli9261128@163.com

收稿日期：2015年7月30日；录用日期：2015年8月12日；发布日期：2015年8月18日

摘　要

本文给出了一个新的求解二阶锥规划的光滑非精确牛顿法。在每次迭代时，新方法采用非精确牛顿法去

文章引用: 董丽, 徐思齐, 杨金根. 二阶锥规划的光滑非精确牛顿法[J]. 应用数学进展, 2015, 4(3): 271-276.
http://dx.doi.org/10.12677/aam.2015.43033

求解一个方程组的解，降低了光滑牛顿法的计算量。在较弱条件下，证明了算法具有全局和局部二阶收敛性质。数值试验表明算法是有效的。

关键词

二阶锥规划，光滑非精确牛顿法，收敛性

1. 引言

本文考虑如下二阶锥规划：

$$\text{(P)}\quad \min\left\{c^{\mathrm{T}}x : Ax=b, x\in K\right\}$$

其中 $x\in K$ 是变量， $A\in R^{m\times n}, c\in R^{n}$ 和 $b\in R^{m}$ 是已知的量。 K 是 n 维的二阶锥，其定义为

$$K=\left\{x=(x_1,x_2)\middle| x_1\in R,\ x_2\in R^{n-1},\ x_1\ge\|x_2\|\right\}$$

其中 $\|\cdot\|$ 为向量的欧几里得范数。(P)的对偶规划为

$$\text{(D)}\quad \max\left\{b^{\mathrm{T}}y : A^{\mathrm{T}}y+s=c, s\in K\right\}$$

二阶锥规划作为数学规划领域的一个重要分支，有着非常重要而又广泛的应用背景和实际意义，其研究问题涉及控制、金融、组合优化、工程技术、神经网络、机器学习等诸多领域[1]。许多数学规划问题，如凸二次规划、二次约束的凸二次规划、矩阵分式优化、范数极小化问题、鲁棒最小二乘问题、天线阵列的设计、带损失风险的金融优化问题等均可以转化为二阶锥规划[1]。因此近年来二阶锥规划成为数学规划领域一个值得关注的方向。

光滑算法是近年来求解二阶锥规划的一种新方法[2]-[6]。该方法的基本思想是利用一个光滑函数将二阶锥规划等价转化成一个非线性方程组，然后利用牛顿法求解该方程组，从而得到原问题的最优解。由于光滑算法不但在理论上具有好的收敛性质，而且在具体实施中有很好的实际计算效果，因而近年来得到了迅速的发展。注意到，文献[2]-[6]所研究的光滑算法在每次迭代时，都需要精确求出一个非线性方程组的解。如果要求解的问题规模较大，精确求解此方程组的解往往效率不高，而非精确牛顿法是解决这类问题的一种很好的途径。

本文给出了一个求解二阶锥规划的光滑非精确牛顿法，证明了算法的全局和局部收敛性质。该方法采用非精确牛顿法求解方程组的解，大大提高了光滑算法的效率。数值试验表明算法对求解大规模二阶锥规划是有效的。

2. 预备知识

对任意的 $x=(x_1,x_2)\in R\times R^{n-1}$， $s=(s_1,s_2)\in R\times R^{n-1}$，与二阶锥 K 相伴的约当代数定义为

$$x\circ s=\left(x^{\mathrm{T}}s, x_1s_2+s_1x_2\right)$$

$(R^n,\circ)$ 是一个欧几里得约当代数，其单位元为 $e=(1,0,\cdots,0)^{\mathrm{T}}\in R^n$。

对任意的 $x=(x_1,x_2)\in R\times R^{n-1}$，定义 $L_x=\begin{pmatrix} x_1 & x_2^{\mathrm{T}} \\ x_2 & x_1 I\end{pmatrix}$，其中 I 为 $(n-1)\times(n-1)$ 维单位矩阵。

易知 $x\circ s=L_x s, \forall x,s\in R^n$。

下面给出与二阶锥 K 相伴的 R^n 中向量的谱分解。对任意的 $x=(x_1,x_2)\in R\times R^{n-1}$，它的谱分解为

$x=\lambda_1 u_1+\lambda_2 u_2$，其中谱值 λ_i 和谱向量 u_i 分别为

$$\lambda_i = x_1 + (-1)^i \|x_2\|,\quad u_i = \begin{cases} \frac{1}{2}\left(1,(-1)^i \frac{x_2}{\|x_2\|}\right), \\ \frac{1}{2}\left(1,(-1)^i w\right) \end{cases} , i=1,2,$$

这里 $w\in R^{n-1}$ 是满足 $\|w\|=1$ 的任意向量。利用谱分解，我们可以定义

$$x^2=\lambda_1^2 u_1+\lambda_2^2 u_2,\ \forall x\in R^n,\quad \sqrt{x}=\sqrt{\lambda_1}\,u_1+\sqrt{\lambda_2}\,u_2,\quad \forall x\in K.$$

本文假设原始问题(P)及其对偶问题(D)都严格可行。在此假设条件下，原始问题(P)及其对偶问题(D)都有最优解且最优值相等，并且求解二阶锥规划问题等价于求解其最优性条件：

$$Ax=b,\ A^{\mathrm{T}}y+s=c,\ x\circ s=0,\quad x,s\in K,\ y\in R^m. \tag{1}$$

3. 算法描述

光滑函数在二阶锥规划光滑算法的设计和收敛性分析中起着非常重要的作用。本文采用著名的 Fischer–Burmeister 光滑函数，其定义如下：

$$\phi(\mu,x,s)=x+s-\sqrt{x^2+s^2+2\mu^2 e},\ \forall(\mu,x,s)\in R_+\times R^{2n}. \tag{2}$$

由文献[7]中的命题 4.2 可知

$$\phi(0,x,s)=0 \Leftrightarrow x\in K,\ s\in K,\ x\circ s=0. \tag{3}$$

记 $z=(\mu,x,y)\in R_+\times R^n\times R^m$。利用 $\phi(\mu,x,s)$ 定义函数 $H:R^{1+n+m}\to R^{1+m+n}$ 如下：

$$H(z)=\begin{pmatrix}\mu\\ b-Ax\\ \phi\left(\mu,x,c-A^{\mathrm{T}}y\right)\end{pmatrix}. \tag{4}$$

由(1)和(3)可知 $z^*=(\mu^*,x^*,y^*)\in R\times R^n\times R^m$ 是 $H(z)=0$ 的解当且仅当 $\left(x^*,y^*,c-A^{\mathrm{T}}y^*\right)$ 是(P)和(D)的解。

下面给出我们的算法。

算法 3.1. (光滑非精确牛顿法)

步骤 0: 选取常数 $\delta,\sigma\in(0,1),\mu_0\in R_{++}$。选取初始点 $(x^0,y^0)\in R^n\times R^m$。令 $z^0:=(\mu_0,x^0,y^0)$。选取常数 $\gamma\in(0,1/2)$，使得 $\mu_0\ge\gamma\|H(z^0)\|$。选取常数 $\tau\in(0,1)$，使得 $\gamma+\tau<1$。令 $k:=0$。

步骤 1：如果 $\|H(z^k)\|=0$，则停止迭代。否则，计算

$$\rho_k:=\gamma\|H(z^k)\|\min\{1,\|H(z^k)\|\} \tag{5}$$

步骤 2：求解如下方程组得搜索方向 $(\Delta\mu_k,\Delta x^k,\Delta y^k)\in R\times R^n\times R^m$，

$$H'(z^k)\Delta z^k=-H(z^k)+t_k \tag{6}$$

这里 $t_k:=(\rho_k,r_k)^{\mathrm{T}}\in R^{1+n+m}$，其中 $r_k\in R^{n+m}$ 并且满足

$$\|r_k\|\le\tau\|H(z^k)\|\min\{1,\|H(z^k)\|\} \tag{7}$$

步骤 3：令 l_k 是使得下式成立的最小非负整数

$$\left\|H\left(z^k+\delta^l\Delta z^k\right)\right\| \le \left[1-\sigma\left(1-\gamma-\tau\right)\delta^l\right]\left\|H\left(z^k\right)\right\| \tag{8}$$

令 $\alpha_k := \delta^{l_k}$。

步骤 4：令 $z^{k+1} := z^k + \alpha_k \Delta z^k$。令 $k := k+1$。转步骤 1。

注：令 $h\left(z^k\right) := \begin{pmatrix} b - Ax^k \\ \phi\left(\mu_k, x^k, c - A^{\mathrm{T}} y^k\right) \end{pmatrix}$，则由算法 3.1 的步骤 2 可知

$$h'\left(z^k\right)\Delta z^k = -h\left(z^k\right) + r_k, \quad \left\|r_k\right\| \le \tau \left\|H\left(z^k\right)\right\| \min\left\{1, \left\|H\left(z^k\right)\right\|\right\},$$

即方程组(6)由非精确牛顿法求解。

下面定理给出了 $H(z)$ 的雅可比矩阵及其可逆性，具体证明可参见文献[6]中的定理 3.2。

定理 3.1：设 $H(z)$ 由(4)定义，则有下面结论。

(i) $H(z)$ 在 R^{1+n+m} 上全局 Lipschitz 连续且处处强半光滑，并且在任意的 $(\mu, x, y) \in R_{++} \times R^n \times R^m$ 处连续可微，其雅可比矩阵为

$$H'(z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & A & 0 \\ P(z) & M(z) & -N(z)A^{\mathrm{T}} \end{pmatrix}$$

其中

$$P(z) = -2\mu L_{\overline{w}}^{-1} e, \quad M(z) = I - L_{\overline{w}}^{-1} L_x, \quad N(z) = I - L_{\overline{w}}^{-1} L_s,$$

$$s = c - A^{\mathrm{T}} y, \ \overline{w} := \sqrt{x^2 + \left(c - A^{\mathrm{T}} y\right)^2 + 2\mu^2 e}.$$

(ii) 如果矩阵 A 行满秩，则对任意的 $\mu > 0$ 有 $H'(z)$ 可逆。

定理 3.2：设矩阵 A 行满秩，则算法 3.1 有好的定义。

证明：假设对于某个 k 有 $\mu_k > 0$。因为矩阵 A 行满秩，故由定理 3.1 知 $H'\left(z^k\right)$ 可逆，所以在第 k 步迭代步骤 2 有好的定义。令 Δz^k 为方程组(6)的解，则对任意的 $\alpha \in (0,1)$，有

$$\mu_k + \alpha \Delta \mu_k = \mu_k + \alpha\left(-\mu_k + \rho_k\right) = (1-\alpha)\mu_k + \alpha \rho_k > 0. \tag{9}$$

由(9)及定理 3.1 可知 $H(\cdot)$ 在 z^k 附近连续可微。对任意的 $\alpha \in (0,1)$，定义

$$f(\alpha) = H\left(z^k + \alpha \Delta z^k\right) - H\left(z^k\right) - \alpha H'\left(z^k\right)\Delta z^k, \tag{10}$$

则 $f(\alpha) = o(\alpha)$。由(5)和(7)知 $\rho_k \le \gamma \left\|H\left(z_k\right)\right\|$，$r_k \le \tau \left\|H\left(z_k\right)\right\|$，故知

$$\left\|t_k\right\| \le \rho_k + \left\|r_k\right\| \le (\gamma + \tau)\left\|H\left(z_k\right)\right\|.$$

因此，由(6)和(10)可知对任意的 $\alpha \in (0,1)$，有

$$\begin{aligned} \left\|H\left(z^k + \alpha\Delta z^k\right)\right\| &= \left\|H\left(z^k\right) + \alpha H'\left(z^k\right)\Delta z^k + f(\alpha)\right\| = \left\|H\left(z^k\right) + \alpha\left[-H\left(z^k\right) + t_k\right] + f(\alpha)\right\| \\ &\le (1-\alpha)\left\|H\left(z^k\right)\right\| + \alpha(\gamma+\tau)\left\|H\left(z^k\right)\right\| + o(\alpha) = \left[1-(1-\gamma-\tau)\alpha\right]\left\|H\left(z^k\right)\right\| + o(\alpha). \end{aligned}$$

因为 $\gamma + \tau < 1$，所以存在一个常数 $\overline{\alpha} \in (0,1)$ 使得对于任意的 $\alpha \in (0, \overline{\alpha}]$，有

$$\left\|H\left(z^k + \alpha\Delta z^k\right)\right\| \le \left[1 - \sigma(1-\gamma-\tau)\alpha\right]\left\|H\left(z^k\right)\right\|,$$

从而可知步骤 3 在第 k 步迭代有好的定义，即算法步骤 3 可以产生一个步长 $\alpha_k \in (0,1)$。由(9)可知

$\mu_{k+1}=\mu_k+\alpha_k\Delta\mu_k>0$，因此由数学归纳法可知算法 3.1 有好的定义。证毕。

引理 3.1 设 $\{z^k=(\mu_k,x^k,y^k)\}$ 为算法 3.1 产生的迭代点列，则对任意的 $k\ge 0$ 有 $\mu_k\ge\rho_k$。

证明：由算法 3.1 的步骤 3 和步骤 4 可知点列 $\{\|H(z^k)\|\}$ 单调下降，从而由 ρ_k 的定义可知点列 $\{\rho_k\}$ 也是单调下降的。由算法 3.1 的步骤 0 可知 $\mu_0\ge\gamma\|H(z^0)\|\ge\gamma\|H(z^0)\|\min\{1,\|H(z^0)\|\}=\rho_0$。假设对于某个 k 有 $\mu_k\ge\rho_k$，则由(9)可得

$$\mu_{k+1}=(1-\alpha_k)\mu_k+\alpha_k\rho_k\ge(1-\alpha_k)\rho_k+\alpha_k\rho_k=\rho_k\ge\rho_{k+1},$$

进而由数学归纳法可知结论成立。证毕。

4. 算法的收敛性质

定理 4.1：如果矩阵 A 行满秩并且 $\{z^k=(\mu_k,x^k,y^k)\}$ 是由算法 3.1 产生的迭代点列，那么 $\{z^k\}$ 的任意聚点 $z^*:=(\mu^*,x^*,y^*)$ 都是 $H(z)=0$ 的解。

证明：不失一般性，我们假设当 $k\to\infty$ 时，$\{z^k\}$ 收敛到 z^*。因为 $\{\|H(z^k)\|\}$ 单调下降并且有下界，所以由 $H(\cdot)$ 的连续性可知 $\{\|H(z^k)\|\}$ 收敛到一个非负数 $\|H(z^*)\|$。若 $\|H(z^*)\|=0$，则结论成立。假设 $\|H(z^*)\|>0$。因为 $\mu_k\ge\rho_k$ 并且 $\{\rho_k\}$ 单调下降，故

$$\mu^*\ge\rho^*:=\gamma\|H(z^*)\|\min\{1,\|H(z^*)\|\}>0,$$

从而由定理 3.1 可知 $H'(z^*)$ 存在并可逆。因此，存在 z^* 的一个闭邻域 $N(z^*)$，使得对任意的 $z\in N(z^*)$ 有 $\mu\in R_{++}$ 并且 $H'(z)$ 可逆。因为 $\lim_{k\to\infty}z^k=z^*$，故对于充分大的 k 有 $z^k\in N(z^*)$，进而可知 $\mu_k\in R_{++}$ 并且 $H'(z^k)$ 可逆。对于充分大的 k，设 Δz^k 是方程组 $H'(z^k)\Delta z^k=-H(z^k)+t_k$ 唯一的解。类似于定理 3.2 的证明，对于充分大的 k，存在一个非负整数 ν 满足 $\delta^\nu\in(0,\bar\alpha]$，这里 $\bar\alpha\in(0,1)$，使得

$$\|H(z^k+\delta^\nu\Delta z^k)\|\le[1-\sigma(1-\gamma-\tau)\delta^\nu]\|H(z^k)\|$$

对任意的 $\alpha\in(0,\bar\alpha]$ 和 $\sigma\in(0,1)$ 都成立。对于充分大的 k，因为 $\alpha_k=\delta^{l_k}>\delta^\nu$，所以由算法 3.1 的步骤 3 和步骤 4 可得

$$\|H(z^{k+1})\|\le[1-\sigma(1-\gamma-\tau)\delta^{l_k}]\|H(z^k)\|\le[1-\sigma(1-\gamma-\tau)\delta^\nu]\|H(z^k)\|$$

令 $k\to\infty$ 并在上式两边取极限，利用 $\|H(z^*)\|>0$，可得 $1\le 1-\sigma(1-\gamma-\tau)\delta^\nu$。这与 $\gamma+\tau<1$ 矛盾。因此可知 $\|H(z^*)\|=0$。证毕。

定理 4.2：假设矩阵 A 行满秩并且 z^* 是算法 3.1 产生的迭代点列 $\{z^k\}$ 的任意聚点。如果所有的 $V\in\partial H(z^*)$ 都是非奇异的，则 $\{z^k\}$ 二次收敛到 z^*，即 $\|z^{k+1}-z^*\|=O(\|z^k-z^*\|^2)$。

证明：由定理 4.1 可知 $\lim_{k\to\infty}\|H(z^k)\|=0$，故对于充分大的 k，有 $\rho_k=\gamma\|H(z^k)\|^2$ 并且 $\|r_k\|\le\tau\|H(z^k)\|^2$，进而可知 $\|t_k\|\le\rho_k+\|r_k\|\le(\gamma+\tau)\|H(z^k)\|^2=O(\|H(z^k)\|^2)$。

因此，类似于文献[2]中的定理 4.3 可证得结论成立。在此省略。证毕。

Table 1. Numerical results of Algorithm 3.1
表 1. 算法 3.1 的数值结果

n	m	AIT	ACPU	MHK
100	50	5.7	0.07	3.056(−7)
200	100	5.6	0.41	9.813(−7)
300	150	5.8	1.12	6.251(−7)
400	200	6.0	2.77	1.019(−7)
500	250	6.2	5.11	8.702(−7)
600	300	6.1	8.98	2.313(−7)
700	350	6.5	15.43	2.691(−7)
800	400	6.5	21.73	4.101(−7)

5. 数值试验

为检验算法 3.1 的实算效果，我们用 Matlab7.0.1 编程在 Windows XP 操作系统的电脑上做数值试验。

在所有的试验中，参数选择为 $\mu_0 = 0.01$，$\sigma = 0.2$，$\delta = 0.8$，$\tau = 0.1$，$\gamma = 0.01\big/\left(1+\left\|H\left(z^0\right)\right\|\right)$。我们选择 $x_0 = e$，$y_0 = 0$ 作为初始点。终止准则为 $\left\|H\left(z^k\right)\right\| \le 10^{-6}$。

测试问题是随机生成的二阶锥规划问题。具体步骤为：首先生成随机的行满秩矩阵 A 和随机向量 $x \in \operatorname{int} K$，$s \in \operatorname{int} K$，$y \in R^m$，这里 $\operatorname{int} K = \left\{x = \left(x_1, x_2\right) \middle| x_1 \in R, x_2 \in R^{n-1}, x_1 > \left\|x_2\right\|\right\}$，然后令 $b = Ax$，$c = A^{\mathrm{T}} y + s$，则得到的二阶锥规划的原问题和对偶问题都存在最优解且最优值相等。每个算例测试 10 次，计算结果列于表 1，其中 $a(-b)$ 表示 $a \times 10^{-b}$，AIT 表示算法所需的迭代次数的平均值，ACPU 表示算法所需的 CPU 时间(单位：秒)的平均值，MHK 表示算法终止时 10 个算例中 $\left\|H\left(z^k\right)\right\|$ 的最大值。

由表 1 可以看出，算法 3.1 是有效的并且能够求解大规模二阶锥规划问题，它只需要很少的迭代次数和 CPU 时间就可以得到满足终止条件的解。

基金项目

河南省自然科学基金(142300410437)，河南省高等学校重点科研项目(15A110039)。

参考文献 (References)

[1] Alizadeh, F. and Goldfarb, D. (2003) Second-order cone optimization. *Mathematical Programming*, **95**, 3-51. http://dx.doi.org/10.1007/s10107-002-0339-5

[2] Chi, X.N. and Liu, S.Y. (2009) A non-interior continuation method for second-order cone optimization. *Optimization*, **58**, 965-979. http://dx.doi.org/10.1080/02331930701763421

[3] Chi, X.N. and Liu, S.Y. (2009) A one-step smoothing Newton method for second-order cone programming. *Journal of Computational and Applied Mathematics*, **223**, 114-123. http://dx.doi.org/10.1016/j.cam.2007.12.023

[4] Fang, L., He, G.P. and Hu, Y.H. (2009) A new smoothing Newton-type method for second-order cone programming problems. *Applied Mathematics and Computation*, **215**, 1020-1029. http://dx.doi.org/10.1016/j.amc.2009.06.029

[5] Tang, J.Y., He, G.P., Dong, L. and Fang, L. (2011) A smoothing Newton method for second-order cone optimization-based on a new smoothing function. *Applied Mathematics and Computation*, **218**, 1317-1329. http://dx.doi.org/10.1016/j.amc.2011.06.015

[6] 汤京永，贺国平 (2012) 一个新的求解二阶锥规划的非内部连续化算法. *应用数学*, **1**, 26-31.

[7] Fukushima, M., Luo, Z.Q. and Tseng, P. (2002) Smoothing functions for second-order-cone complementarity problems. *SIAM Journal on Optimization*, **12**, 436-460. http://dx.doi.org/10.1137/S1052623400380365