Advances in Applied Mathematics 应用数学进展, 2020, 9(2), 238-243
Published Online February 2020 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2020.92027

A Feasible SSDP Algorithm for Nonlinear Semidefinite Programming

Meiling He, Jianling Li*

College of Mathematics and Information Science, Guangxi University, Nanning Guangxi
Email: *jianlingli@126.com

Received: Feb. 2nd, 2020; accepted: Feb. 17th, 2020; published: Feb. 24th, 2020

Abstract

This paper proposes a feasible SSDP algorithm for solving nonlinear semidefinite programming. The initial point and iteration points are feasible. The search direction is determined by solving two quadratic semidefinite programming subproblems. The step size is obtained by calculating the line search that satisfies the descent property of the objective function and the feasibility of the constraint function. The global convergence of the algorithm is proved under mild conditions.

Keywords

Nonlinear Semidefinite Programming, Feasible SSDP, Line Search, Global Convergence

非线性半定规划的一个可行SSDP算法

何美玲，黎健玲*

广西大学数学与信息科学学院，广西 南宁
Email: *jianlingli@126.com

收稿日期：2020年2月2日；录用日期：2020年2月17日；发布日期：2020年2月24日

摘 要

本文提出了一个求解非线性半定规划的可行序列半定规划(SSDP)算法。该算法的初始点和迭代点均是可行点，在每次迭代中通过求解两个二次半定规划子问题确定搜索方向，步长由满足目标函数下降性和约束函数可行性的线搜索产生，在某些假设条件下本文证明了算法的全局收敛性。

*通讯作者。

文章引用: 何美玲, 黎健玲. 非线性半定规划的一个可行 SSDP 算法[J]. 应用数学进展, 2020, 9(2): 238-243.
DOI: 10.12677/aam.2020.92027

关键词

非线性半定规划，可行SSDP，线搜索，全局收敛性

1. 引言

本文研究如下仅带半负定矩阵约束的非线性半定规划问题(NLSDP)：

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & G(x) \preceq 0, \end{aligned} \tag{1.1}$$

其中，函数 $f:\mathbb{R}^n \to \mathbb{R}$ 和 $G:\mathbb{R}^n \to \mathbb{S}^m$ 是光滑的函数，$\mathbb{S}^m$ 表示 m 阶实对称矩阵空间，$\preceq$ 表示半负定偏序，即 $A \preceq B$ 表示 $A-B$ 是半负定矩阵。

非线性半定规划问题在最优控制、结构设计、经济金融等领域有广泛应用[1] [2] [3]。非线性半定规划带有半定矩阵约束，结构复杂，不能直接用传统的非线性规划的算法求解，目前求解非线性半定规划问题的主要方法有增广 Lagrange 法[4]，原始－对偶内点法[5] [6]，序列半定规划法[7] [8] [9]，QP-free 法[10]，交替方向乘子法[11]等。

本文基于文[12]的求解非线性规划的可行 SQP 算法的思想，建立了非线性半定规划的一个可行 SSDP 算法。算法产生的迭代点均满足可行性；在每次迭代中算法的搜索方向是通过求解两个二次半定规划子问题产生；目标函数直接作为效益函数用于线搜索，线搜索保证目标函数下降和满足可行性。在适当的假设条件下算法具有全局收敛性。

2. 算法

定义 2.1：对于可微矩阵函数 $G(x):\mathbb{R}^n \to \mathbb{S}^m$，$G(x)$ 在 $x \in \mathbb{R}^n$ 处的微分算子 $DG(x)$ 定义如下：

$$DG(x) := \left(\frac{\partial G(x)}{\partial x_1}, \frac{\partial G(x)}{\partial x_2}, \cdots, \frac{\partial G(x)}{\partial x_n} \right)^{\mathrm{T}},$$

$G(x)$ 在 x 处沿着 $d=(d_1,d_2,\cdots,d_n)^n \in \mathbb{R}^n$ 的方向导数 $DG(x)d$ 为

$$DG(x)d = \sum_{i=1}^{n} d_i \frac{\partial G(x)}{\partial x_i},$$

相对应的伴随算子 $DG(x)^*$ 定义为

$$DG(x)^* Y = \left(\left\langle \frac{\partial G(x)}{\partial x_1}, Y \right\rangle, \left\langle \frac{\partial G(x)}{\partial x_2}, Y \right\rangle, \cdots, \left\langle \frac{\partial G(x)}{\partial x_n}, Y \right\rangle \right)^{\mathrm{T}}, \forall Y \in \mathbb{S}^m,$$

其中 $\langle \bullet,\bullet \rangle$ 表示矩阵的内积，定义如下：

$$\langle A,B \rangle = \mathrm{Tr}\left(B^{\mathrm{T}} A\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ij},\ \forall A = \left[a_{ij}\right],\ B = \left[b_{ij}\right] \in \mathbb{R}^{m\times n}.$$

定义 2.2： 设 $x \in \mathbb{R}^n$ 是 NLSDP(1.1)的一个可行点，若存在矩阵 $M \in \mathbb{S}^m$ 使得以下条件成立：

$$\nabla f(x) + DG(x)^* M = 0,$$

$$G(x) \preceq 0,\ M \succeq 0,\ \mathrm{Tr}\big(MG(x)\big) = 0,$$

则称 x 是 NLSDP(1.1)的一个 KKT 点，M 是相应的拉格朗日乘子，上式称为 NLSDP(1.1)的 KKT 条件。

定义 2.3： NLSDP(1.1)的约束违反度函数定义如下：

$$P(x) = \lambda_1\big(G(x)\big)_+ = \max\big\{0, \lambda_1\big(G(x)\big)\big\},$$

其中，$\lambda_1(A)$表示矩阵 A 的最大特征值。

显然，$P(x) = 0$ 当且仅当 x 为 NLSDP(1.1)的一个可行点。

令 x^k 为当前迭代点，构造如下的二次半定规划子问题(简记为 QSDP)：

$$\begin{aligned} \min \quad & \nabla f\left(x^k\right)^{\mathrm{T}} d + \frac{1}{2} d^{\mathrm{T}} H_k d \\ \text{s.t.} \quad & G\left(x^k\right) + DG\left(x^k\right) d \preceq 0, \end{aligned} \tag{2.1}$$

其中 H_k 为 NLSDP(1.1)的 Lagrange 函数的 Hesse 阵或近似阵。

子问题 QSDP(2.1)的解 d_0^k 为目标函数的下降方向，但不一定是可行方向，因此需对其进行修正，以便得到一个可行下降方向 d^k。

利用 d_0^k 对子问题 QSDP(2.1)的约束不等式的右边进行扰动，得到如下的二次半定规划子问题：

$$\begin{aligned} \min \quad & \nabla f\left(x^k\right)^{\mathrm{T}} d + \frac{1}{2} d^{\mathrm{T}} H_k d \\ \text{s.t.} \quad & G\left(x^k\right) + DG\left(x^k\right) d \preceq -\left\|d_0^k\right\|^{\nu} E_m, \end{aligned} \tag{2.2}$$

其中 E_m 为一个 m 阶的单位矩阵。

记 NLSDP(1.1)的可行集为 $\Omega = \left\{x \in \mathbb{R}^n \mid G(x) \preceq 0\right\}$。本文需作以下基本假设：

假设 1： 集合 Ω 是非空的。

假设 2： 函数 $f(x)$，$G(x)$是连续可微的。

假设 3： 存在正的常数 a 和 b，使得

$$a\|d\|^2 \le d^{\mathrm{T}} H_k d \le b\|d\|^2, \quad \forall d \in R^n.$$

求解 NLSDP(1.1)的可行 SSDP 算法描述如下：

算法 A：

步骤 0： (初始化)选取参数 $\alpha \in \left(0, \frac{1}{2}\right)$，$\beta \in (0,1)$，$\nu > 2$。初始点 $x^0 \in \Omega$，$H_0 \in \mathbb{R}^{n \times n}$ 正定，令 $k = 0$。

步骤 1： (计算搜索方向)

步骤 1.1： 求解子问题 QSDP(2.1)得 d_0^k。如果 $\left\|d_0^k\right\| = 0$，则停止；否则，进入步骤 1.2。

步骤 1.2： 求解子问题 QSDP(2.2)得 d^k，令 $\theta_k = \nabla f\left(x^k\right)^{\mathrm{T}} d^k$。

步骤 2： (线搜索)计算序列$\left\{1, \beta, \beta^2, \cdots\right\}$中第一个满足如下两个条件的步长 t_k：

$$f\left(x^k + t d^k\right) \le f\left(x^k\right) + \alpha t \theta_k, \tag{2.3}$$

$$G\left(x^k + t d^k\right) \preceq 0. \tag{2.4}$$

步骤 3：(更新)令 $x^{k+1}=x^k+t_k d^k$，用某种技术产生新的对称正定阵 H_{k+1}，令 $k=k+1$，返回步骤 1。

为分析的简便，现作如下进一步假设：

假设 4：子问题 QSDP(2.2)有解 d^k，且 d^k 满足如下条件：

$$\left\|d^k\right\|\le L,\quad \nabla f\left(x^k\right)^{\mathrm{T}} d^k\le \min\left\{-\left\|d_0^k\right\|^{\sigma},-\left\|d^k\right\|^{\sigma}\right\}.$$

其中 $L>0$，$\sigma>2$。

假设 5：算法 A 生成的迭代点列 $\left\{x^k\right\}$ 有界。

基于假设条件 1~3，易知下面引理成立。

引理 2.1：假设 1~3 成立，则子问题 QSDP(2.1)有唯一解 d_0^k，并且存在矩阵 $M_k\in\mathbb{S}^m$ 满足 QSDP(2.1)的 KKT 条件，即

$$\nabla f\left(x\right)+DG\left(x^k\right)^{*}M_k+H_k d_0^k=0,\tag{2.5a}$$

$$G\left(x^k\right)+DG\left(x^k\right)d_0^k\preceq 0,\tag{2.5b}$$

$$M_k\succeq 0,\tag{2.5c}$$

$$\mathrm{Tr}\left(M_k\left(G\left(x^k\right)+DG\left(x^k\right)d_0^k\right)\right)=0.\tag{2.5d}$$

引理 2.2：假设 1~3 成立，如果 $d_0^k=0$，则当前迭代点 x^k 为 NLSDP(1.1)的一个 KKT 点。

证明：将 $d_0^k=0$ 代入(2.5)，结合定义 2.2 即知当前迭代点 x^k 为 NLSDP(1.1)的一个 KKT 点。

引理 2.3：假设 1~4 成立，算法 A 的线搜索在有限次计算中生成步长 t_k，即算法 A 是适定的。

证明：由 Taylor 展开式知

$$f\left(x^k+td^k\right)=f\left(x^k\right)+t\nabla f\left(x^k\right)^{\mathrm{T}}d^k+o\left(t\right),$$

$$G\left(x^k+td^k\right)=G\left(x^k\right)+tDG\left(x^k\right)d^k+o\left(t\right),$$

由假设 4 知 $\theta_k=\nabla f\left(x^k\right)^{\mathrm{T}}d^k<0$。因为 $\alpha\in\left(0,\dfrac{1}{2}\right)$，所以存在 $\overline{t}_k>0$，使得对任意 $t\in\left(0,\overline{t}_k\right)$，有

$$f\left(x^k+td^k\right)\le f\left(x^k\right)+\alpha t\theta_k,$$

因为 $G\left(x\right)$ 是连续可微函数，由子问题 QSDP(2.2)的约束条件知

$$G\left(x^k\right)+DG\left(x^k\right)d^k\preceq-\left\|d_0^k\right\|^{\nu}E_m\prec 0,$$

从而有

$$\lambda_1\left(G\left(x^k\right)+DG\left(x^k\right)d^k\right)<0,\tag{2.6}$$

而由 $\lambda_1(\cdot)$ 的凸性可得

$$\lambda_1\left(G\left(x^k+td^k\right)\right)\le\left(1-t\right)\lambda_1\left(G\left(x^k\right)\right)+t\lambda_1\left(G\left(x^k\right)+DG\left(x^k\right)d^k\right)+o\left(t\right),$$

所以由 $G\left(x^k\right)\preceq 0$，并结合(2.6)知存在 $\tilde{t}_k>0$，使得对任意 $t\in\left(0,\tilde{t}_k\right)$，有

$$\lambda_1\left(G\left(x^k+td^k\right)\right)\le 0,$$

从而

$$G\left(x^k+td^k\right)\preceq 0,$$

综上所述知存在充分小的 $t>0$，使得不等式(2.3)和(2.4)成立，故算法 A 适定。

3. 全局收敛性

若算法 A 产生有限的迭代点列，则由引理 2.2 知当前迭代点 x^k 为问题 NLSDP(1.1)的 KKT 点，因此，我们不妨假设算法 A 产生无限的迭代点列 $\left\{x^k\right\}$，本节将证明在适当的假设条件下算法 A 是全局收敛的。

引理 3.1： 假设 1~5 成立，x^* 是 $\left\{x^k\right\}$ 的一个聚点，即存在子列 $K\subseteq\{1,2,\cdots\}$，使得 $x^k\xrightarrow{K}x^*$，则 $d_0^k\xrightarrow{K}0$。

证明： (反证法)假设结论不成立，则存在子列 $\bar{K}\subseteq K$ 和常数 $\underline{d}>0$，使得当 $k\in\bar{K}$ 充分大时，$\left\|d_0^k\right\|\geq\underline{d}$。

下面首先证明存在正数 $\underline{t}>0$，使得

$$t_k\geq\underline{t},\ \forall k\in\bar{K},$$

由假设 4 及子问题 QSDP(2.2)的约束条件知对充分大的 $k\in\bar{K}$，有

$$\theta_k=\nabla f\left(x^k\right)^{\mathrm{T}}d^k\leq-\left\|d_0^k\right\|^{\sigma}\leq-\underline{d}^{\sigma}, \tag{3.1}$$

$$G\left(x^k\right)+DG\left(x^k\right)d^k\preceq-\left\|d_0^k\right\|^{\nu}E_m\preceq-\underline{d}^{\nu}E_m. \tag{3.2}$$

由假设 4 知 $\left\|d^k\right\|\leq L$，即 $\left\{d^k\right\}$ 有界，于是由 Taylor 展开式有

$$f\left(x^k+td^k\right)=f\left(x^k\right)+t\nabla f\left(x^k\right)^{\mathrm{T}}d^k+o(t),$$

对充分大的 $k\in\bar{K}$，结合(3.1)知

$$f\left(x^k+td^k\right)-f\left(x^k\right)-\alpha t\theta_k=t(1-\alpha)\nabla f\left(x^k\right)^{\mathrm{T}}d^k+o(t)\leq-t(1-\alpha)\underline{d}^{\sigma}+o(t),$$

所以对充分大的 $k\in\bar{K}$ 及充分小的 $t>0$，(2.3)成立。

同理，由 Taylor 展开式有

$$G\left(x^k+td^k\right)=G\left(x^k\right)+tDG\left(x^k\right)d^k+o(t),$$

对充分大的 $k\in\bar{K}$，结合(3.2)知

$$\begin{aligned}\lambda_1\left(G\left(x^k+td^k\right)\right)&\leq(1-t)\lambda_1\left(G\left(x^k\right)\right)+t\lambda_1\left(G\left(x^k\right)+DG\left(x^k\right)d^k\right)+o(t)\\&\leq t\lambda_1\left(-\underline{d}^{\nu}E_m\right)+o(t),\end{aligned}$$

所以对充分大的 $k\in\bar{K}$ 及充分小的 $t>0$，(2.4)成立。

综上所述，存在正数 $\underline{t}>0$，使得 $t_k\geq\underline{t}$，$\forall k\in\bar{K}$。

下面基于上述结论导出矛盾。对充分大的 $k\in\bar{K}$，由(2.3)知

$$f\left(x^{k+1}\right)-f\left(x^k\right)\leq\alpha t\theta_k<0. \tag{3.3}$$

因为 $\left\{f\left(x^k\right)\right\}$ 单调下降，且 $x^k\xrightarrow{\bar{K}}x^*$，所以 $\lim\limits_{k\to\infty}f\left(x^k\right)=f\left(x^*\right)$。在(3.3)式中令 $k\xrightarrow{\bar{K}}\infty$，得出矛盾，于是引理成立。

基于引理 3.1 以及子问题 QSDP(2.1)的 KKT 条件(2.5)，可得算法 A 的全局收敛性，即：

定理 3.1： 假设 1~5 成立，$\left\{x^k\right\}$ 是算法 A 产生的一个迭代点列，则 $\left\{x^k\right\}$ 的每一个聚点都是 NLSDP(1.1)的一个 KKT 点。

4. 结束语

本文将非线性规划可行 SQP 算法进行了推广，提出了一个求解非线性半定规划的可行 SSDP 算法。该算法的搜索方向是通过求解两个二次半定规划子问题产生，目标函数直接作为效益函数，线搜索保证目标函数下降和满足可行性。在适当的假设条件下，证明了算法的适定性和全局收敛性。该算法可以应用于对可行性有严格要求的实际问题。

基金项目

获国家自然科学基金(No. 11561005)，广西自然科学基金(No. 2016GXNSFAA380248)资助。

参考文献

[1] Kanno, Y and Takewaki, I. (2006) Sequential Semidefinite Program for Maximum Robustness Design of Structures under Load Uncertainty. *Journal of Optimization Theory and Applications*, **130**, 265-287. https://doi.org/10.1007/s10957-006-9102-z

[2] Roche, J.R., Herskovits, J., Bazán, E. and Zúñiga, A. (2017) A Feasible Direction Algorithm for General Nonlinear Semidefinite Programming. *Structural and Multidisciplinary Optimization*, **55**, 1261-1279. https://doi.org/10.1007/s00158-016-1564-5

[3] Yang, L. and Yu, B. (2013) A Homotopy Method for Nonlinear Semidefinite Programming. *Computational Optimization and Applications*, **56**, 81-96. https://doi.org/10.1007/s10589-013-9545-8

[4] Sun, D.F., Sun, J. and Zhang, L.W. (2008) The Rate of Convergence of the Augmented Lagrangian Method for Nonlinear Semidefinite Programming. *Mathematical Programming*, **114**, 349-391. https://doi.org/10.1007/s10107-007-0105-9

[5] Yamashita, H., Yabe, H. and Harada, K. (2012) A Primal-Dual Interior Point Method for Nonlinear Semidefinite Programming. *Mathematical Programming*, **135**, 89-121. https://doi.org/10.1007/s10107-011-0449-z

[6] Yamashita, H. and Yabe, H. (2012) Local and Superlinear Convergence of a Primal-Dual Interior Point Method for Nonlinear Semidefinite Programming. *Mathematical Programming*, **132**, 1-30. https://doi.org/10.1007/s10107-010-0354-x

[7] Correa, R. and Ramirez, H.C. (2004) A Global Algorithm for Nonlinear Semidefinite Programming. *SIAM Journal on Optimization*, **15**, 303-318. https://doi.org/10.1137/S1052623402417298

[8] Zhao, Q. and Chen, Z. (2016) On the Superlinear Local Convergence of a Penalty-Free Method for Nonlinear Semidefinite Programming. *Journal of Computational and Applied Mathematics*, **308**, 1-19. https://doi.org/10.1016/j.cam.2016.05.007

[9] Zhao, Q. and Chen, Z. (2018) An SQP-Type Method with Superlinear Convergence for Nonlinear Semidefinite Programming. *Asia-Pacific Journal of Operational Research*, **35**, 1-25. https://doi.org/10.1142/S0217595918500094

[10] Li, J.L., Yang, Z.P. and Jian, J.B. (2017) A Globally Convergent QP-Free Algorithm for Nonlinear Semidefinite Programming. *Journal of Inequalities and Applications*, **2017**, Article No. 145. https://doi.org/10.1186/s13660-017-1415-y

[11] Sun, J. and Zhang, S. (2010) A Modified Alternating Direction Method for Convex Quadratically Constrained Quadratic Semidefinite Programs. *European Journal of Operational Research*, **207**, 1210-1220. https://doi.org/10.1016/j.ejor.2010.07.020

[12] Panier, E.R. and Tits, A.L. (1987) A Superlinearly Convergent Feasible Method for the Solution of Inequality Constrained Optimization Problems. *SIAM Journal on Control & Optimization*, **25**, 934-950. https://doi.org/10.1137/0325051