Advances in Applied Mathematics 应用数学进展, **2018, 7(4), 456-465**
Published Online April 2018 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2018.74056

A New Globally Convergent Algorithm for Nonlinear Semidefinite Programming

Hui Zhang, Jianling Li*

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

Received: Apr. 12th, 2018; accepted: Apr. 21st, 2018; published: Apr. 28th, 2018

Abstract

In this paper, we present a sequence quadratic semidefinite programming (SSDP) algorithm for nonlinear semidefinite programming. At each iteration, the search direction is determined by solving two semidefinite programming sub problems; by introducing a distance function, a merit function is constructed for line search. Under some appropriate conditions, any accumulation point of iterative point sequence is either an infeasible stationary point, or a KKT point of the problem.

Keywords

Nonlinear Semidefinite Programming, SSDP Algorithm, KKT Point, Global Convergence

非线性半定规划一个新的全局收敛算法

张　辉，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：2018年4月12日；录用日期：2018年4月21日；发布日期：2018年4月28日

摘　要

本文提出了一个求解非线性半定规划的序列二次半定规划(SSDP)算法。算法每次迭代通过求解两个半定规划子问题确定搜索方向；通过引进距离函数来构造效益函数用于线搜索，从而产生新的迭代点。在适当的假设条件下，算法或收敛到问题的不可行稳定点，或收敛到问题的KKT点。

*通讯作者。

文章引用: 张辉, 黎健玲. 非线性半定规划一个新的全局收敛算法[J]. 应用数学进展, 2018, 7(4): 456-465.
DOI: 10.12677/aam.2018.74056

关键词

非线性半定规划，SSDP算法，KKT点，全局收敛性

1. 引言

本文考虑如下非线性半定规划(NLSDP)：

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \mathcal{A}(x) \preceq 0, \\ & h(x) = 0, \end{array} \tag{1}$$

其中，函数 $f:\mathbb{R}^n \to \mathbb{R}, \mathcal{A}:\mathbb{R}^n \to \mathbb{S}^{m_1}$ 和 $h(x):\mathbb{R}^n \to \mathbb{R}^{m_2}$ 是连续可微函数，$\mathbb{S}^{m_1}$ 表示所有 m_1 阶实对称矩阵构成的集合。“$\mathcal{A}(x) \preceq 0$”表示 $\mathcal{A}(x)$ 为对称半负定矩阵。

非线性半定规划在工程设计、最优控制、金融风险等领域有着广泛的应用，见文[1]-[6]。序列二次规划(SQP)算法是求解传统非线性规划的有效算法之一，一些学者在非线性规划的 SQP 算法思想的基础上，提出了求解非线性半定规划的序列二次半定规划(SSDP)算法。Correa 和 Ramirez [7]提出通过求解一个二次半定规划子问题产生搜索方向并将罚函数作为效益函数用于线搜索的 SSDP 方法。在一定的假设条件下，证明了算法的某些全局收敛性质。但该文没有分析产生搜索方向的子问题的相容性。

本文通过引进距离函数，提出一个新的 SSDP 算法，该算法具有如下特点：

- 产生搜索方向的子问题是相容的；
- 初始点任意，并且算法的线搜索保证效益函数的充分下降；
- 在一定条件下算法或收敛到问题的不可行稳定点，或收敛到问题的 KKT 点。

2. 算法

记 $\mathbb{R}^{m\times n}$ 为所有 $m\times n$ 矩阵构成的集合，$\mathbb{S}_+^m\left(\mathbb{S}_{++}^m\right)$ 为所有对称半正定(正定)矩阵构成的集合，$\mathbb{S}_-^m\left(\mathbb{S}_{--}^m\right)$ 为所有对称半负定(负定)矩阵构成的集合。记 NLSDP (0.1)的可行集为 S。

定义 1.1： 对于任意矩阵 $A=\left(a_{ij}\right)\in\mathbb{R}^{m\times n}$，$B=\left(b_{ij}\right)\in\mathbb{R}^{m\times n}$，定义 A，B 的内积为

$$\langle A,B\rangle := \mathrm{Tr}\left(B^{\mathrm{T}}A\right)=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}b_{ij},$$

其中 $\mathrm{Tr}(\bullet)$ 为矩阵的迹算子。

对于可微函数 $h(x):\mathbb{R}^n\to\mathbb{R}^{m_2}$，我们给出其微分为：

$$Dh(x)=\begin{pmatrix} \dfrac{\partial h_1(x)}{\partial x_1} & \cdots & \dfrac{\partial h_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_{m_2}(x)}{\partial x_1} & \cdots & \dfrac{\partial h_{m_2}(x)}{\partial x_n} \end{pmatrix}. \tag{1.1}$$

对于可微矩阵值函数 $\mathcal{A}:\mathbb{R}^n\to\mathbb{R}^{m_1\times m_1}$，引入如下符号表示该函数在点 x 处的微分算子[7]：

$$D\mathcal{A}(x)=\left(\frac{\partial\mathcal{A}(x)}{\partial x_1},\cdots,\frac{\partial\mathcal{A}(x)}{\partial x_n}\right)^{\mathrm{T}},$$

其中 $\frac{\partial\mathcal{A}(x)}{\partial x_p}$ 表示 $\mathcal{A}(x)$ 关于 x_p 的偏导数。对于任意 $d=(d_1,d_2,\cdots,d_n)^{\mathrm{T}}\in\mathbb{R}^n$，$D\mathcal{A}(x)d$ 定义为

$$D\mathcal{A}(x)d:=\sum_{i=1}^{n}d_i\frac{\partial\mathcal{A}(x)}{\partial x_i}. \tag{1.2}$$

若线性算子 $V:\mathbb{R}^n\to\mathbb{S}^{m_1}$ 定义为 $Vy=\sum_{i=1}^{n}y_iV_i,\forall y\in\mathbb{R}^n$，其中 $V_i\in\mathbb{S}^{m_1}(i=1,2,\cdots,n)$，则伴随算子 $V^*:\mathbb{S}^{m_1}\to\mathbb{R}^n$ 为

$$V^*Z=\left(\langle V_1,Z\rangle,\cdots,\langle V_n,Z\rangle\right)^{\mathrm{T}},\forall Z\in\mathbb{S}^{m_1}.$$

NLSDP (0.1)的 Lagrange 函数 $L:\mathbb{R}^n\times\mathbb{S}^{m_1}\times\mathbb{R}^{m_2}\to\mathbb{R}$ 定义为

$$L(x,\wedge,v)=f(x)+\langle\mathcal{A}(x),\wedge\rangle+Dh(x)^{\mathrm{T}}v. \tag{1.3}$$

定义 1.2 [7]： 对于向量 $x\in\mathbb{R}^n$，若存在 $\wedge\in\mathbb{S}^{m_1},v\in\mathbb{R}^{m_2}$，使得

$$\nabla f(x)+D\mathcal{A}(x)^*\wedge+Dh(x)^{\mathrm{T}}v=0, \tag{1.4a}$$

$$\langle\mathcal{A}(x),\wedge\rangle=0,\wedge\succeq0,\mathcal{A}(x)\preceq0,h(x)=0, \tag{1.4b}$$

则称 x 为 NLSDP (0.1)的一个 KKT 点，称 $(\wedge,v)$ 为相应的 Lagrange 乘子。

设 $y\in\mathbb{R}^1,z\in\mathbb{R}^{m_2},K:=\mathbb{R}_-^1\times\{0\}_{\mathbb{R}^{m_2}}$，定义

$$\operatorname{dist}\left[\begin{pmatrix}y\\z\end{pmatrix}\middle|K\right]:=\inf\left\{\left\|\begin{pmatrix}y\\z\end{pmatrix}-\begin{pmatrix}x\\0\end{pmatrix}\right\|:x\le0\right\}$$

为 $\begin{pmatrix}y\\z\end{pmatrix}$ 到集合 K 的距离，其中 $\|\cdot\|$ 表示向量的欧式范数。

本文用于线搜索的效益函数定义如下：

$$\theta_\alpha(x)=f(x)+\alpha P(x), \tag{1.5}$$

其中 $P(x)$ 由下式定义：

$$P(x)=\operatorname{dist}\left[\begin{pmatrix}\lambda_1(\mathcal{A}(x))\\h(x)\end{pmatrix}\middle|K\right], \tag{1.6}$$

显然 $P(x)=0$ 当且仅当 x 为 NLSDP (0.1)的可行点。

下面给出求解 NLSDP (0.1)的 SSDP 算法的具体步骤。

算法 A

参数： $\alpha_0>0,\beta,\sigma\in(0,1)$，$\eta_1>0$，选取初始迭代点 $x^0\in R^n$ 和初始矩阵 $B_0=E_n$ (单位阵)。令 $k:=0$。

步骤 1： 求解如下二次半定规划子问题(简记为 LSDP(x^k))，得最优解 $\{\hat{d}^k,z_1^k,z_2^k\}$

$$
\begin{aligned}
\min \quad & z_1 + \|z_2\| \\
\text{s.t.} \quad & \mathcal{A}(x^k) + D\mathcal{A}(x^k)d \preceq z_1 E_{m_1}, \\
& h(x^k) + Dh(x^k)d = z_2, \\
& z_1 \ge 0, \\
& \|d\| \le 1,
\end{aligned} \tag{1.7}
$$

其中 E_{m_1} 为 m-阶单位阵。记

$$
\Delta(x) := \inf\left\{ \text{dist}\left[\begin{pmatrix} \lambda_1(\mathcal{A}(x) + D\mathcal{A}(x)d) \\ h(x) + Dh(x)d \end{pmatrix} \middle| K \right] \middle| \|d\| \le 1 \right\} - P(x), \tag{1.8}
$$

若 z_1, z_2 不全为 0，且 $\Delta(x^k) = 0$，则算法终止，否则转步骤 2；

步骤 2： 令 d^k 为如下二次半定规划子问题(简记为 $\text{QSDP}(x^k)$)的最优解：

$$
\begin{aligned}
\min \quad & \nabla f(x^k)^{\mathrm{T}} d + \frac{1}{2} d^{\mathrm{T}} B_k d \\
\text{s.t.} \quad & \mathcal{A}(x^k) + D\mathcal{A}(x^k)d \preceq z_1^k E_{m_1}, \\
& h(x^k) + Dh(x^k)d = z_2^k,
\end{aligned} \tag{1.9}
$$

若 $d^k = 0$，则算法停止；否则，转步骤 3。

步骤 3： (更新罚参数)罚参数 α_{k+1} 的更新方式如下：

$$
\alpha_{k+1} := \begin{cases} \alpha_k, & \text{若} Q(x^k, \alpha_k) \le -(d^k)^{\mathrm{T}} B_k d^k, \\ \dfrac{\nabla f(x^k)^{\mathrm{T}} d^k + (d^k)^{\mathrm{T}} B_k d^k}{-\Delta(x^k)} + \eta_1, & \text{其他}, \end{cases} \tag{1.10}
$$

其中 $Q(x^k, \alpha_k) := \nabla f(x^k)^{\mathrm{T}} d^k + \alpha_k \Delta(x^k)$。

步骤 4： (计算步长)取 t_k 为 $\{1, \sigma, \sigma^2, \cdots\}$ 中满足如下式子的最大值：

$$
\theta_{\alpha_{k+1}}(x^k + td^k) \le \theta_{\alpha_{k+1}}(x^k) + \beta t Q(x^k, \alpha_{k+1}). \tag{1.11}
$$

步骤 5： 令 $x^{k+1} = x^k + t_k d^k$，利用某种更新方式更新 B_k 为对称正定阵 B_{k+1}，返回步骤 1。

注 2.1： $\text{LSDP}(x^k)(1.7)$ 的最优解中的 $\hat{d}^k$ 可能不唯一，我们将其构成的集合记 $D(x^k)$。

注 2.2： $\text{QSDP}(x^k)(1.9)$ 的 KKT 条件为

$$
\nabla f(x^k)^{\mathrm{T}} d^k + B_k d^k + D\mathcal{A}(x^k)^* \Lambda_k + Dh(x^k)^{\mathrm{T}} v^k = 0, \tag{1.2a}
$$

$$
\mathcal{A}(x^k) + D\mathcal{A}(x^k)d^k \preceq z_1^k E_{m_1}, \tag{1.12b}
$$

$$
h(x^k) + Dh(x^k)d^k = z_2^k, \tag{1.12c}
$$

$$
\Lambda_k \succeq 0,\ \text{Tr}\left(\Lambda_k\left(\mathcal{A}(x^k) + D\mathcal{A}(x^k)d^k - z_1^k E_{m_1}\right)\right) = 0. \tag{1.12d}
$$

注 2.3： 步骤 3 的(1.10)能够保证罚参数的更新是单调不减的。事实上，若 α_k 执行更新，则由 $Q(x^k, \alpha_k) > -(d^k)^{\mathrm{T}} B_k d^k$ 可得 $\alpha_k < \dfrac{\nabla f(x^k)^{\mathrm{T}} d^k + (d^k)^{\mathrm{T}} B_k d^k}{-\Delta(x^k)} < \alpha_{k+1}$。

下面分析算法 A 的适定性，类似非线性半定规划的不可行稳定点的定义[8] [9]，首先给出非线性半定规划 NLSDP (0.1)的不可行稳定点的定义。

定义 1.3： 设 $\tilde{x}$ 为 NLSDP(0.1)的不可行点，若

$$\inf\left\{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}(\tilde{x})+D\mathcal{A}(\tilde{x})d\right)\\ h(\tilde{x})+Dh(\tilde{x})d\end{pmatrix}\middle| K\right]\middle| d\in\mathbb{R}^n\right\}-P(\tilde{x})=0,$$

则称 $\tilde{x}$ 为 NLSDP (0.1)的不可行稳定点。

本文需作如下基本假设：

A1： 函数 $f(x),h(x)$ 和 $\mathcal{A}(x)$ 是连续可微的。

A2： 存在正数 a 和 $\overline{a}$ 使得

$$a\|d\|^2\le d^{\mathrm{T}}B_k d\le\overline{a}\|d\|^2,\forall d\in\mathbb{R}^n.$$

A3： 由算法 A 产生的迭代点列 $\{x^k\}$ 是有界的。

A4： 对任意的 $\overline{x}\in S$，Mangasarian-Fromovitz 约束规格(MFCQ)成立，即

1) $Dh(\overline{x})$ 行满秩；

2) 存在非零向量 $\overline{d}=(\overline{d}_1,\cdots,\overline{d}_n)^{\mathrm{T}}\in\mathbb{R}^n$，使得 $\begin{cases}Dh(\overline{x})\overline{d}=0\\ \mathcal{A}(\overline{x})+\sum\limits_{i=1}^{n}\overline{d}_i\dfrac{\partial\mathcal{A}(\overline{x})}{\partial x_i}\prec 0\end{cases}$

引理 1.1： 若假设 1~3 成立，$\{x^k\}$ 是由算法 A 产生的点列，$d^k\in\mathbb{R}^n$ 为 QSDP$(x^k)(1.9)$ 的最优解，则方向导数 $P'(x^k;d^k)$ 满足如下不等式：

$$P'(x^k;d^k)\le\Delta(x^k)\le 0.$$

进一步的，若 $\Delta(x^k)=0$ 且 x^k 是 NLSDP (0.1)的不可行点，则 x^k 是 NLSDP (0.1)的一个不可行稳定点。

证明： 设 $\overline{d}^k\in D(x^k)$，则

$$\begin{aligned}&\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}(x^k)+D\mathcal{A}(x^k)\overline{d}^k\right)\\ h(x^k)+Dh(x^k)\overline{d}^k\end{pmatrix}\middle| K\right]\\
&\le\operatorname{dist}\left[\begin{matrix}z_1^k\\ z_2^k\end{matrix}\middle| K\right]+\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}(x^k)+D\mathcal{A}(x^k)\overline{d}^k\right)-z_1^k\\ h(x^k)+Dh(x^k)\overline{d}^k-z_2^k\end{pmatrix}\middle| K\right]\\
&\le\left\|\begin{matrix}z_1^k\\ z_2^k\end{matrix}\right\|=\inf\left\{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}(x^k)+D\mathcal{A}(x^k)d\right)\\ h(x^k)+Dh(x^k)d\end{pmatrix}\middle| K\right]\middle|\|d\|\le 1\right\}.\end{aligned}$$

则由方向导数的定义及 Taylor 展式知

$$\begin{aligned}&P'(x^k;d^k)\\
&=\lim_{\rho\to 0}\frac{P(x^k+\rho d^k)-P(x^k)}{\rho}\end{aligned}$$

$$
\begin{aligned}
&= \lim_{\rho\to 0}\frac{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k+\rho d^k\right)\right)\\ h\left(x^k+\rho d^k\right)\end{pmatrix}\middle| K\right]-P\left(x^k\right)}{\rho}\\
&= \lim_{\rho\to 0}\frac{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)+\rho\left(D\mathcal{A}\left(x^k\right)d^k\right)+o(\rho)\right)\\ h\left(x^k\right)+\rho Dh\left(x^k\right)d^k+o(\rho)\end{pmatrix}\middle| K\right]-P\left(x^k\right)}{\rho}\\
&= \lim_{\rho\to 0}\frac{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)+\rho\mathcal{A}\left(x^k\right)-\rho\mathcal{A}\left(x^k\right)+\rho\left(D\mathcal{A}\left(x^k\right)d^k\right)+o(\rho)\right)\\ h\left(x^k\right)+\rho Dh\left(x^k\right)d^k+o(\rho)\end{pmatrix}\middle| K\right]-P\left(x^k\right)}{\rho}\\
&\le \lim_{\rho\to 0}\frac{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)\right)+\rho\lambda_1\left(\mathcal{A}\left(x^k\right)+D\mathcal{A}\left(x^k\right)d^k\right)-\rho\lambda_1\left(\mathcal{A}\left(x^k\right)\right)+o(\rho)\\ h\left(x^k\right)+\rho h\left(x^k\right)-\rho h\left(x^k\right)+\rho Dh\left(x^k\right)d^k+o(\rho)\end{pmatrix}\middle| K\right]-P\left(x^k\right)}{\rho}\\
&\le \inf\left\{\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)+D\mathcal{A}\left(x^k\right)d\right)\\ h\left(x^k\right)+Dh\left(x^k\right)d\end{pmatrix}\middle| K\right]\middle|\,\|d\|\le 1\right\}-\operatorname{dist}\left[\begin{matrix}\lambda_1\left(\mathcal{A}\left(x^k\right)\right)\\ h\left(x^k\right)\end{matrix}\middle| K\right]\\
&\quad+\operatorname{dist}\left[\begin{matrix}\lim\limits_{\rho\to 0}\frac{o(\rho)}{\rho}\\ \lim\limits_{\rho\to 0}\frac{o(\rho)}{\rho}\end{matrix}\middle| K\right]\\
&=\Delta\left(x^k\right),
\end{aligned}
$$

显然 $d=0$ 时，$\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)+D\mathcal{A}\left(x^k\right)d\right)\\ h\left(x^k\right)+Dh\left(x^k\right)d\end{pmatrix}\middle| K\right]=\operatorname{dist}\left[\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)\right)\\ h\left(x^k\right)\end{pmatrix}\middle| K\right]$，因此结合(1.8)可得 $\Delta\left(x^k\right)\le 0$，因此结合以上分析可知 $P'\left(x^k;d^k\right)\le\Delta\left(x^k\right)\le 0$。

若 $\Delta\left(x^k\right)=0$，此时若 x^k 为 NLSDP (0.1)的不可行点，由不可行稳定点的定义知 x^k 为 NLSDP (0.1)的不可行稳定点。因此综合以上所述引理结论得证。□

引理 1.2：若假设 1~3 成立，算法有限步终止于 x^k，则 x^k 或是 NLSDP (0.1)的不可行稳定点，或是 NLSDP (0.1)的一个 KKT 点。

证明：若算法终止于步骤 1，则结合不可行稳定点的定义知 x^k 是 NLSDP (0.1)的不可行稳定点。若算法终止于步骤 2，则可知 $d^k=0$ 是 $\text{QSDP}\left(x^k\right)(1.9)$ 的最优解，将 $d^k=0$ 带入(1.12)可得

$$\nabla f\left(x^k\right)+D\mathcal{A}\left(x^k\right)^*\Lambda_k+Dh\left(x^k\right)^{\mathrm{T}}v^k=0, \tag{1.13a}$$

$$\mathcal{A}\left(x^k\right)\preceq z_1^kE_m, \tag{1.13b}$$

$$h\left(x^k\right)=z_2^k, \tag{1.13c}$$

$$\Lambda_k \succeq 0,\ \mathrm{Tr}\left(\Lambda_k\left(\mathcal{A}\left(x^k\right)-z_1^k E_m\right)\right)=0. \tag{1.13d}$$

下面证明 $z_1^k=0, z_2^k=0$。(反证)若不成立，则 $\left\|z_1^k\right\|+\left\|z_2^k\right\|>0$，则可知 x^k 为不可行点。显然 $\left(0, \lambda_1\left(\mathcal{A}\left(x^k\right)\right)_+, h\left(x^k\right)\right)$ 为 $\mathrm{LSDP}\left(x^k\right)(1.7)$ 的可行解，因此

$$\left\|\begin{matrix} z_1^k \\ z_2^k \end{matrix}\right\| \le \left\|\begin{matrix} \lambda_1\left(\mathcal{A}\left(x^k\right)\right)_+ \\ h\left(x^k\right) \end{matrix}\right\|,$$

其中 $\lambda_1\left(\mathcal{A}\left(x^k\right)\right)_+ := \max\left\{0, \lambda_1\left(\mathcal{A}\left(x^k\right)\right)\right\}$，又因为算法在步骤 1 不终止，因此 $\left\|\begin{matrix} z_1^k \\ z_2^k \end{matrix}\right\| < \left\|\begin{matrix} \lambda_1\left(\mathcal{A}\left(x^k\right)\right)_+ \\ h\left(x^k\right) \end{matrix}\right\|$，因此可知 $d=0$ 不是 $\mathrm{QSDP}\left(x^k\right)(1.9)$ 的可行解。矛盾。因此 $z_1^k=0, z_2^k=0$，将 $z_1^k=0, z_2^k=0$ 代入(1.13)可知 x^k 是 NLSDP (0.1)的一个 KKT 点。

基于以上引理，若算法 A 在步骤 1~2 未终止，则结合更新公式(1.10)可知 $Q\left(x^k, \alpha_{k+1}\right)<0$，因此，线搜索(1.11)可执行，从而知算法 A 是适定的。□

3. 全局收敛性

基于引理 1.2，本节中不妨假设算法 A 产生无限点列 $\left\{x^k\right\}$。接下来将证明 $\left\{x^k\right\}$ 的聚点或是 NLSDP (0.1)的不可行稳定点，或是 NLSDP (0.1)的一个 KKT 点。

引理 2.1：假设 1~4 成立，$\left\{x^k\right\}$ 是由算法 A 产生的点列，若罚参数 $\alpha_k \to \infty, x^k \xrightarrow{\kappa} x^*$，则 x^* 为 NLSDP(0.1)的一个不可行稳定点。

证明：首先证明 $\left\{d^k\right\}$ 是有界的。由 $\mathrm{LSDP}\left(x^k\right)(1.7)$ 可知 $\left\|\hat{d}^k\right\| \le 1$。显然 $\hat{d}^k$ 是 $\mathrm{QSDP}\left(x^k\right)(1.9)$ 的可行解，因此我们可以得到

$$\nabla f\left(x^k\right)^{\mathrm{T}} d^k + \frac{1}{2}\left(x^k\right)^{\mathrm{T}} B_k d^k \le \nabla f\left(x^k\right)^{\mathrm{T}} \hat{d}^k + \frac{1}{2}\left(\hat{d}^k\right)^{\mathrm{T}} B_k \hat{d}^k,$$

此外，由上式结合假设 2，有

$$\nabla f\left(x^k\right)^{\mathrm{T}} d^k + \frac{1}{2}\left(d^k\right)^{\mathrm{T}} B_k d^k \le \left\|\nabla f\left(x^k\right)\right\|\left\|\hat{d}^k\right\| + \frac{1}{2}\left\|\hat{d}^k\right\|^2 \bar{a} \le M_1 + \frac{1}{2}\bar{a}.$$

另一方面我们根据假设 2，有

$$\begin{aligned} \nabla f\left(x^k\right)^{\mathrm{T}} d^k + \frac{1}{2}\left(d^k\right)^{\mathrm{T}} B_k d^k &\ge -\left\|\nabla f\left(x^k\right)\right\|\left\|d^k\right\| + a\left\|d^k\right\|^2 \\ &\ge -M_1\left\|d^k\right\| + a\left\|d^k\right\|^2. \end{aligned}$$

以上两式可以得到 $\left\{d^k\right\}$ 是有界的。因 $\alpha_k \to +\infty$，则由罚参数更新公式(1.10)可知 $\xi_k := \dfrac{\nabla f\left(x^k\right)^{\mathrm{T}} d^k + \left(d^k\right)^{\mathrm{T}} B_k d^k}{-\Delta\left(x^k\right)}$ 趋于 $+\infty$。而又结合 $\left\{f\left(x^k\right)\right\}$ 和 $\left\{d^k\right\}$ 的有界性可知：$\Delta\left(x^k\right)$ 趋于 0，若此式成立，则结合 κ 的定义及罚参数的更新方式可知 $\Delta\left(x^k\right) \to \Delta\left(x^*\right)=0$，且 x^k 为不可行点。下面证明 x^* 为不可行点。

由(1.12a)可知

$$\nabla f\left(x^k\right)^{\mathrm{T}} d^k=-\left(D\mathcal{A}\left(x^k\right)^* \Lambda_k\right)^{\mathrm{T}} d^k-\left(Dh\left(x^k\right)v^k\right)^{\mathrm{T}} d^k-\left(d^k\right)^{\mathrm{T}} B_k d^k$$
$$=-\sum_{i=1}^n d_i^k \operatorname{Tr}\left(\Lambda_k \frac{\partial \mathcal{A}\left(x^k\right)}{\partial x_i}\right)-\left(v^k\right)^{\mathrm{T}} Dh\left(x^k\right)d^k-\left(d^k\right)^{\mathrm{T}} B_k d^k. \tag{2.1}$$

根据(1.12b)可知

$$-\operatorname{Tr}\left(\Lambda_k\left(\sum_{i=1}^n d_i^k \frac{\partial \mathcal{A}\left(x^k\right)}{\partial x_i}\right)\right)=-\operatorname{Tr}\left(\Lambda_k\left(D\mathcal{A}\left(x^k\right)d^k\right)\right)=\operatorname{Tr}\left(\Lambda_k\left(\mathcal{A}\left(x^k\right)-z_1^k E_m\right)\right). \tag{2.2}$$

由(2.1)，(2.2)，结合(1.13c)可知

$$\begin{aligned}\xi_k&=\frac{\operatorname{Tr}\left(\Lambda_k\left(\mathcal{A}\left(x^k\right)-z_k E_m\right)\right)-\left(Dh\left(x^k\right)d^k\right)^{\mathrm{T}} v^k}{-\Delta(x^k)}\\&\le\frac{\lambda_1\left(\mathcal{A}\left(x^k\right)-z_1^k E_m\right)\operatorname{Tr}\left(\Lambda_k\right)+\left\|h\left(x^k\right)-z_2^k\right\|\left\|v^k\right\|}{P\left(x^k\right)-\left\|\begin{matrix}z_1^k\\ z_2^k\end{matrix}\right\|}\\&\le\frac{1}{\left\|\begin{matrix}\lambda_1\left(\mathcal{A}\left(x^k\right)-z_1^k E_{m_1}\right)\\ h(x^k)-z_2^k\end{matrix}\right\|}\left\|\begin{matrix}\lambda_1\left(\mathcal{A}\left(x^k\right)-z_1^k E_{m_1}\right)\\ h\left(x^k\right)-z_2^k\end{matrix}\right\|\left\|\begin{matrix}\operatorname{Tr}\left(\Lambda_k\right)\\ v^k\end{matrix}\right\|\\&=\left\|\begin{matrix}\operatorname{Tr}\left(\Lambda_k\right)\\ v^k\end{matrix}\right\|.\end{aligned}$$

则结合罚参数 α_k 无界易得 $\left\|\begin{matrix}\operatorname{Tr}\left(\Lambda_k\right)\\ v^k\end{matrix}\right\|$ 无界，因此 x^* 为不可行点(否则若 x^* 为 NLSDP (0.1)的可行点，则由假设 4，可知 MFCQ 在 x^* 处成立。类似于文献[10]中定理 5.1 的证明，可知 QSDP$\left(x^*,B_*\right)$(1.9) 的 KKT 乘子集合 Ω 是非空且有界的。注意到 $\Lambda_k \xrightarrow{\kappa} \Lambda_*, v^k \xrightarrow{\kappa} v^*$ 且 $\Lambda_*, v^* \in \Omega$，因此 $\left\{\Lambda_k\right\},\left\{v^k\right\}$ 是有界的)。再结合 $\Delta\left(x^*\right)=0$，由定义即知 x^* 为 NLSDP (0.1)的不可行稳定点。□

基于引理 2.1，在下面的分析和讨论中不妨假设 $\alpha_k<+\infty$，结合罚参数的更新公式(1.10)，可得如下结论成立：

引理 2.2： 若假设 1~4 成立。则存在自然数 k_0，使得对任意的 $k \ge k_0$，都有 $\alpha_k \equiv \alpha_{k_0} \triangleq \alpha>0$。基于引理 2.2，在下面的分析中，不失一般性，假设 $\alpha_k \equiv \alpha, k=1,2,\cdots$

定理 2.1： 假设 1~4 成立，x^* 为算法 A 产生的无限序列 $\left\{x^k\right\}$ 的任意一个聚点，即 $x^k \xrightarrow{\kappa} x^*$，则 x^* 或是 NLSDP(0.1)的一个不可行稳定点，或是 NLSDP (0.1)的一个 KKT 点。

证明： 不失一般性，假设 x^* 不是 NLSDP (0.1)的不可行稳定点，下面将证明 x^* 是 NLSDP (0.1)的一个 KKT 点。首先证明若 x^* 不是 NLSDP (0.1)的不可行稳定点，则

$$d^k \xrightarrow{K} 0.$$

反证法，假设 $d^k \stackrel{K}{\nrightarrow} 0$，则存在常数 $b>0$ 以及一个指标集 $K' \subseteq K$，使得 $d^k \xrightarrow{K'} d^*$，且

$$\left\|d^*\right\| \ge b>0. \tag{2.3}$$

下面将证明分成两部分。

a) 证明存在 $\underline{t}>0$，使得 $\underline{t}:=\inf\{t_k,k\in K'\}>0$。

$$\begin{aligned}\theta_\alpha\left(x^k+td^k\right)&=f\left(x^k+td^k\right)+\alpha\operatorname{dist}\left[\left.\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k+td^k\right)\right)\\ h\left(x^k+td^k\right)\end{pmatrix}\right|K\right]\\ &\le f\left(x^k\right)+\alpha\operatorname{dist}\left[\left.\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)\right)\\ h\left(x^k\right)\end{pmatrix}\right|K\right]+t\nabla f\left(x^k\right)^{\mathrm{T}}d^k\\ &\quad+\alpha t\left\{\inf\left\{\operatorname{dist}\left[\left.\begin{pmatrix}\lambda_1\left(\mathcal{A}\left(x^k\right)+D\mathcal{A}\left(x^k\right)d\right)\\ h\left(x^k\right)+Dh\left(x^k\right)d\end{pmatrix}\right|K\right]\middle|\|d\|\le 1\right\}-P\left(x^k\right)\right\}+o(t)\\ &=\theta\left(x^k\right)+t\nabla f\left(x^k\right)^{\mathrm{T}}d^k+\alpha t\Delta\left(x^k\right)+o(t)\\ &=\theta\left(x^k\right)+tQ\left(x^k,\alpha\right)+o(t),\end{aligned}$$

于是有

$$\theta_\alpha\left(x^k+td^k\right)-\theta\left(x^k\right)-\beta tQ\left(x^k,\alpha\right)\le(1-\beta)tQ\left(x^k,\alpha\right)+o(t).$$

结合 $Q\left(x^k,\alpha\right)\le-\left(d^k\right)^{\mathrm{T}}B_kd^k$ 以及(2.3)可知 $Q\left(x^*,\alpha\right)\le-0.5\left(d^*\right)^{\mathrm{T}}B_*d^*<0$，因此当 $k\in K'$ 充分大时有

$$\theta_\alpha\left(x^k+td^k\right)-\theta\left(x^k\right)-\beta tQ\left(x^k,\alpha\right)\le-0.5(1-\beta)t\left(d^*\right)^{\mathrm{T}}B_*d^*+o(t)<0. \tag{2.4}$$

由此可知存在 $\underline{t}>0$，使得 $\underline{t}:=\inf\{t_k,k\in K'\}>0$。

b) 基于 $\underline{t}:=\inf\{t_k,k\in K'\}>0$，将导出矛盾。

由线搜索(1.11)可知 $\left\{\theta_\alpha\left(x^k\right)\right\}$ 为非增序列。由 $\theta_\alpha\left(x^k\right)=f\left(x^k\right)+\alpha P\left(x^k\right)\ge f\left(x^k\right)$ 以及 $\left\{f\left(x^k\right)\right\}$ 的有界性可知 $\left\{\theta_\alpha\left(x^k\right)\right\}$ 有下界。因此 $\left\{\theta_\alpha\left(x^k\right)\right\}$ 收敛。令 $k(\in K')\to\infty$，对(2.4)取极限得

$$-0.5ab^2\beta\underline{t}>0,$$

显然矛盾。因此 $\lim\limits_{K}d^k=0$。进而知 $d^*=0$。将 $d^*=0$ 带入 $\mathrm{QSDP}\left(x^*,B_*\right)$(1.9) 的 KKT 条件，得

$$\nabla f\left(x^*\right)+D\mathcal{A}\left(x^*\right)^*\Lambda_k+Dh\left(x^*\right)^{\mathrm{T}}v^*=0, \tag{2.5a}$$

$$\mathcal{A}\left(x^*\right)\preceq z_1^*E_{m_1}, \tag{2.5b}$$

$$h\left(x^*\right)=z_2^*, \tag{2.5c}$$

$$\Lambda_*\succeq 0, \tag{2.5d}$$

$$\operatorname{Tr}\left(\Lambda_*\left(\mathcal{A}\left(x^*\right)-z_1^*E_{m_1}\right)\right)=0. \tag{2.5e}$$

下证 z_1^* 和 z_2^* 均为 0。反证法，若不成立，易知 x^* 为不可行点，则 $\left\|\begin{matrix}z_1^*\\ z_2^*\end{matrix}\right\|>0$，结合 d^* 为 $\mathrm{LSDP}\left(x^k\right)$ 的一个可行解知 $\Delta\left(x^*\right)=0$，故 x^* 为不可行稳定点，矛盾。因此 z_1^* 和 z_2^* 均为 0。将 $z_1^k=0,z_2^k=0$ 代入(2.5)即知 x^* 为 NLSDP (0.1)的 KKT 点。□

4. 结束语

本文提出了带有等式和半负定矩阵约束的非线性半定规划一个全局收敛的 SSDP 算法。基于非线性规划修正的序列二次规划算法思想构建两个半定规划子问题，通过求解子问题产生搜索方向，线搜索保证效益函数充分下降。在 MFCQ 等温和条件下，论证了算法的适定性和全局收敛性。

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] Ben-Tal, A., Jarre, F., Kocvara, M., Nemirovski, A. and Zowe, J. (2000) Optimal Design of Trusses under a Nonconvex Global Buckling Constraint. *Optimization and Engineering*, **1**, 189-213. https://doi.org/10.1023/A:1010091831812

[2] Fares, B., Noll, D. and Apkarian, P. (2002) Robust Control via Sequential Semidefinite Programming. *SIAM Journal on Control and Optimization*, **40**, 1791-1820. https://doi.org/10.1137/S0363012900373483

[3] Kocvara, M. and Stingl, M. (2004) Solving Nonconvex SDP Problems of Structural Optimization with Stability Control. *Optimization Methods and Software*, **19**, 595-609. https://doi.org/10.1080/10556780410001682844

[4] Jarre, F. (2000) An Interior Method for Nonconvex Semidefinite Programs. *Optimization and Engineering*, **1**, 347-372. https://doi.org/10.1023/A:1011562523132

[5] Fares, B., Apkarian, P. and Noll, D. (2001) An Augmented Lagrangian Method for a Class of LMI-Constrained Problems in Robust Control Theory. *International Journal of Control*, **74**, 3702-3706. https://doi.org/10.1080/00207170010010605

[6] Nishimura, R., Hayashi, S. and Fukushima, M. (2012) Semidefinite Complement Arit Reformulation for Robust Nash Equilibrium Problems with Euclidean Uncertainty Sets. *Journal of Global Optimization*, **53**, 107-120.

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

[8] Burke, J.V. and More, J.J. (1988) On the Identification of Active Constraints. *SIAM Journal on Numerical Analysis*, **25**, 1197-1211. https://doi.org/10.1137/0725068

[9] Dunn, J.C. (1987) On the Convergence of Projected Gradient Processes to Singular Critical Points. *Journal of Optimization Theory and Applications*, **55**, 203-216. https://doi.org/10.1007/BF00939081

[10] Burke, J.V. and Han, S.P. (1989) A Robust Sequential Quadratic Programming Method. *Mathematical Programming*, **43**, 277-303. https://doi.org/10.1007/BF01582294