运筹学与控制论

半定规划的一种 Mehrotra 型预估-校正算法

陈 华 平

（六盘水师范学院 数学系， 贵州 六盘水 553004 ）

摘要：将一种 Mehrotra 型预估-校正算法推广到半定规划。首先给出了半定规划基于 Mehrotra 型预估-校正算法的一些基本理论，尤其是对称化技术；随后通过分析这种算法的迭代复杂性，给出算法的重要思想：在校长步中采用安全策略，给出新算法的最大预估步长的上界，算法过程中对最大预估步长进行削减策略：当最大预估步长大于某个阈值时，对此步长进行削减（可重复），从而得到合适的校正步长下界；最终通过采用以上策略及 NT 搜索方向，得到了该算法的多项式复杂界。

关键词：大步校正算法；Mehrotra 型预估-校正算法；半定规划；多项式复杂性

中图分类号：O221.2

自 1984 年 Karmarkar 为线性规划(LP)提出了一个有效的内点算法以来[1]，线性规划内点算法便成为优化领域研究的热点之一。半定规划(SDO)是线性规划的推广，在许多领域都有广泛的应用[2]。Mehrotra 型预估-校正算法有着良好的计算效果备受大家青睐[3-4]，但是这种算法在校正步可能会出现步长很小甚至为零的不良状况.为此，Salahi 等人通过引入一个“安全策略”(safeguard strategy)[4]，使得这种修正算法最终被证明具有多项式复杂性，数值实验表明它也保持了 Mehrotra 型预估-校正算法的有效性.本文将[4]中的修正算法推广到 SDO 上来.模型的改变导致分析困难，尤其是对称性问题.本文采用备受青睐的 NT 方向作为牛顿搜索方向，通过一些分析技巧，证明了算法的多项式复杂性.

号约定：记 $\mathbf{R}^n$ 为 n 维欧式空间，$\mathbf{R}^{n\times n}$ 为 $n\times n$ 矩阵集，$\boldsymbol{S}^n, \boldsymbol{S}_+^n, \boldsymbol{S}_{++}^n$ 分别表示 $n\times n$ 对称矩阵集、半正定矩阵集、正定矩阵集；任意 $\boldsymbol{M}\in\mathbf{R}^{n\times n}$，$Tr(\boldsymbol{M})=\sum_{i=1}^n \boldsymbol{M}_{ii}$，$\boldsymbol{A}\bullet\boldsymbol{B}=Tr(\boldsymbol{A}^T\boldsymbol{B})$；任意 $\boldsymbol{M}\in\boldsymbol{S}^n$，$\lambda_i(\boldsymbol{M})$ $(i=1,\ldots,n)$、$\lambda_{\max}(\boldsymbol{M})$、$\lambda_{\min}(\boldsymbol{M})$ 分别记 $\boldsymbol{M}$ 的特征值、最大和最小特征 $cond(\boldsymbol{M})=\lambda_{\max}(\boldsymbol{M})/\lambda_{\min}(\boldsymbol{M})$；矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{S}$ 的 Kroneker 积记 $\boldsymbol{X}\otimes\boldsymbol{S}$； $\boldsymbol{X}\in\mathbf{R}^{m\times n}$， $vec(\boldsymbol{X})$ 表示一 mn 维的列向量.

1 SDO 问题的基本理论

1.1 SDO 问题

考虑如下半定规划问题

----

* 收稿日期：2014-07-13 修回日期：2014-09-09
资助项目：贵州省科学技术基金项目(No. 黔科合 J 字 LKLS[2013]14)
作者简介：陈华平，女，讲师，研究方向为最优化理论及其应用，E-mail: chp88ko@aliyun.com

$$\min\{\boldsymbol{C}\bullet\boldsymbol{X} \mid \boldsymbol{A}_i\bullet\boldsymbol{X} = b_i,\ i=1,\ldots,m,\ \boldsymbol{X}\in\boldsymbol{S}_+^n\} \tag{1}$$

其中$\boldsymbol{C}\in\boldsymbol{S}^n$，$b=(b_1,\ldots,b_m)^T\in\mathbf{R}^m$，$\boldsymbol{X}$为原始矩阵变量.问题(1)的对偶问题为:

$$\max\{b^T y \mid \sum_{i=1}^{m} y_i\boldsymbol{A}_i+\boldsymbol{S}=\boldsymbol{C},\quad \boldsymbol{S}\in\boldsymbol{S}_+^n\} \tag{2}$$

其中$(y,\boldsymbol{S})\in\mathbf{R}^m\times\boldsymbol{S}_+^n$为对偶矩阵变量. 原始-对偶可行集及其相对内部分别为:

$$\mathrm{F} = \left\{(\boldsymbol{X},y,\boldsymbol{S})\in\boldsymbol{S}_+^n\times\mathbf{R}^m\times\boldsymbol{S}_+^n \mid \boldsymbol{A}_i\bullet\boldsymbol{X}=b_i, i=1,\ldots,m, \sum_{i=1}^{m} y_i\boldsymbol{A}_i+\boldsymbol{S}=\boldsymbol{C}\right\}$$

$$\mathrm{F}^0 = \left\{(\boldsymbol{X},y,\boldsymbol{S})\in\boldsymbol{S}_{++}^n\times\mathbf{R}^m\times\boldsymbol{S}_{++}^n \mid \boldsymbol{A}_i\bullet\boldsymbol{X}=b_i, i=1,\ldots,m, \sum_{i=1}^{m} y_i\boldsymbol{A}_i+\boldsymbol{S}=\boldsymbol{C}\right\}$$

本文作如下假设: (i) $\mathrm{F}^{\circ}\neq\varnothing$，　(ⅱ)矩阵$\boldsymbol{A}_i$, $i=1,\ldots,m$线性无关.

在以上假设下，问题(1)和(2)的中心路径方程

$$\begin{cases} \boldsymbol{A}_i\bullet\boldsymbol{X}=b_i,\ \boldsymbol{X}\in\boldsymbol{S}_+^n,\ i=1,\ldots,m, \\ \sum_{i=1}^{m} y_i\boldsymbol{A}_i+\boldsymbol{S}=\boldsymbol{C},\ \boldsymbol{S}\in\boldsymbol{S}_+^n, \\ \boldsymbol{X}\boldsymbol{S}=\mu\boldsymbol{I}. \end{cases} \tag{3}$$

有唯一解$(\boldsymbol{X}(\mu),y(\mu),\boldsymbol{S}(\mu))$.随着$\mu$趋于 0，$\mu$-中心$(\boldsymbol{X}(\mu),y(\mu),\boldsymbol{S}(\mu))$ 的极限$(\boldsymbol{X}^*,y^*,\boldsymbol{S}^*)$即为问题(1)和(2)最优解.

1.2 SDO 的搜索方向

当$\boldsymbol{X},\boldsymbol{S}$对称时，$\boldsymbol{XS}$一般不对称，故对方程(3)采用对称化技术[5]: 对任意给定的非奇异矩阵$\boldsymbol{P}\in\mathbf{R}^{n\times n}$，引入对称化算子$H_P:\mathbf{R}^{n\times n}\to\boldsymbol{S}^n$，定义

$$H_{\boldsymbol{P}}(\boldsymbol{M})\equiv\frac{1}{2}[\boldsymbol{PMP}^{-1}+(\boldsymbol{PMP}^{-1})^{\mathrm{T}}]$$

当$\lambda_i(\boldsymbol{M})\in\mathbf{R}$, $i=1,\ldots,n$时，有$H_{\boldsymbol{P}}(\boldsymbol{M})=\mu\boldsymbol{I}\Leftrightarrow\boldsymbol{M}=\mu\boldsymbol{I}$ [5].

本文将非奇异矩阵$\boldsymbol{P}$限制在如下矩阵类上:

$$\mathrm{P}(\boldsymbol{X},\boldsymbol{S})\equiv\{\boldsymbol{P}\in\boldsymbol{S}_{++}^n \mid \boldsymbol{P}^2\boldsymbol{XS}=\boldsymbol{SXP}^2, \boldsymbol{X},\boldsymbol{S}\in\boldsymbol{S}_{++}^n\} \tag{4}$$

则方程(3)的第 3 式替换为$H_{\boldsymbol{P}}(\boldsymbol{M})=\mu\boldsymbol{I}$，运用牛顿法，得如下牛顿方程为

$$\begin{cases} \boldsymbol{A}_i \bullet \Delta \boldsymbol{X} = 0,\ i = 1, \ldots, m, \\ \sum_{i=1}^{m} \Delta y_i \boldsymbol{A}_i + \Delta \boldsymbol{S} = 0, \\ H_{\boldsymbol{P}}(\boldsymbol{X}\Delta \boldsymbol{S} + \Delta \boldsymbol{X}\boldsymbol{S}) = \sigma\mu_g \boldsymbol{I} - H_{\boldsymbol{P}}(\boldsymbol{X}\boldsymbol{S}). \end{cases} \tag{5}$$

方程(5)的解$(\Delta\boldsymbol{X}, \Delta y, \Delta\boldsymbol{S}) \in \boldsymbol{S}^n \times \mathbf{R}^m \times \boldsymbol{S}^n$ 即为牛顿方向，$\sigma \in [0,1]$为心中参数，$(\boldsymbol{X}, y, \boldsymbol{S})$处的对偶间隙为$\mu_g = \boldsymbol{X} \bullet \boldsymbol{S} / n$. 特别地，当$\boldsymbol{P} = \boldsymbol{W}^{1/2}$[6]且$\boldsymbol{X}, \boldsymbol{S} \in \boldsymbol{S}_+^n$时，有$\boldsymbol{P} \in \mathrm{P}(\boldsymbol{X}, \boldsymbol{S})$，这里

$$\boldsymbol{W} = \boldsymbol{X}^{-1/2}(\boldsymbol{X}^{1/2}\boldsymbol{S}\boldsymbol{X}^{1/2})^{1/2}\boldsymbol{X}^{-1/2} = \boldsymbol{S}^{1/2}(\boldsymbol{S}^{1/2}\boldsymbol{X}\boldsymbol{S}^{1/2})^{-1/2}\boldsymbol{S}^{1/2} \tag{6}$$

此时称$\boldsymbol{P}$为 NT 尺度矩阵，对应于方程(5)的解$(\Delta\boldsymbol{X}, \Delta y, \Delta\boldsymbol{S})$ 为 NT 方向.

本文整个算法在如下宽邻域内

$$N_\infty^-(\gamma) = \{(\boldsymbol{X}, y, \boldsymbol{S}) \in \mathrm{F}^\circ \mid \lambda_{min}(\boldsymbol{X}\boldsymbol{S}) \ge \gamma\mu_g\}$$

内进行，这里$\gamma \in (0, \tfrac{1}{4})$. 设当前点为$(\boldsymbol{X}, y, \boldsymbol{S}) \in N_\infty^-(\gamma)$，则新的迭代点和新的对偶间隙为

$$(\boldsymbol{X}(\alpha), y(\alpha), \boldsymbol{S}(\alpha)) = (\boldsymbol{X}, y, \boldsymbol{S}) + \alpha(\Delta\boldsymbol{X}, \Delta y, \Delta\boldsymbol{S}) \tag{7}$$

$$\mu_g(\alpha) = \frac{\boldsymbol{X}(\alpha) \bullet \boldsymbol{S}(\alpha)}{n} \tag{8}$$

Mehrotra 型预估-校正算法的预估步迭代方向是通过解方程

$$\begin{cases} \boldsymbol{A}_i \bullet \Delta \boldsymbol{X}^a = 0,\ i = 1, \ldots, m, \\ \sum_{i=1}^{m} \boldsymbol{A}_i \Delta y_i^a + \Delta \boldsymbol{S}^a = 0, \\ H_{\boldsymbol{P}}(\boldsymbol{X}\Delta \boldsymbol{S}^a + \Delta \boldsymbol{X}^a\boldsymbol{S}) = -H_{\boldsymbol{P}}(\boldsymbol{X}\boldsymbol{S}). \end{cases} \tag{9}$$

得到，计算此方向上的最大可行步长$\alpha_a \in (0,1]$，使$(\boldsymbol{X}(\alpha_a), \boldsymbol{S}(\alpha_a)) \in \boldsymbol{S}_+^n \times \boldsymbol{S}_+^n$. 但算法并不执行此步而是利用预估步的信息，计算方程组

$$\begin{cases} \boldsymbol{A}_i \bullet \Delta \boldsymbol{X} = 0,\ i = 1, \ldots, m, \\ \sum_{i=1}^{m} \boldsymbol{A}_i \Delta y_i + \Delta \boldsymbol{S} = 0, \\ H_{\boldsymbol{P}}(\boldsymbol{X}\Delta \boldsymbol{S} + \Delta \boldsymbol{X}\boldsymbol{S}) = \sigma\mu_g \boldsymbol{I} - H_{\boldsymbol{P}}(\Delta \boldsymbol{X}^a \boldsymbol{S}^a) - H_{\boldsymbol{P}}(\boldsymbol{X}\boldsymbol{S}). \end{cases} \tag{10}$$

得校正步的迭代方向$(\Delta\boldsymbol{X}, \Delta y, \Delta\boldsymbol{S})$，其中$\sigma = (1-\alpha_a)^3$，计算最大可行步长$\alpha_c$，使新的迭代点$(\boldsymbol{X}(\alpha), y(\alpha), \boldsymbol{S}(\alpha)) \in N_\infty^-(\gamma)$.

为了便于分析，记$\hat{\boldsymbol{X}} \equiv \boldsymbol{P}\boldsymbol{X}\boldsymbol{P},\ \Delta\hat{\boldsymbol{X}} \equiv \boldsymbol{P}\Delta\boldsymbol{X}\boldsymbol{P},\ \hat{\boldsymbol{S}} \equiv \boldsymbol{P}^{-1}\boldsymbol{S}\boldsymbol{P}^{-1},\ \Delta\hat{\boldsymbol{S}} \equiv \boldsymbol{P}^{-1}\Delta\boldsymbol{S}\boldsymbol{P}^{-1}$，$\Delta\hat{\boldsymbol{X}}^a \equiv \boldsymbol{P}\Delta\boldsymbol{X}^a\boldsymbol{P}$，$\Delta\hat{\boldsymbol{S}}^a \equiv \boldsymbol{P}^{-1}\Delta\boldsymbol{S}^a\boldsymbol{P}^{-1}$，则

$$H_{P}(X\mathrm{D}S+\mathrm{D}XS)$$

$$=\frac{1}{2}[(PXPP^{-1}\mathrm{D}SP^{-1}+P\mathrm{D}XPP^{-1}SP^{-1})+(PXPP^{-1}\mathrm{D}SP^{-1}+P\mathrm{D}XPP^{-1}SP^{-1})^{\mathrm{T}}]$$

$$=\frac{1}{2}[(\hat{X}\mathrm{D}\hat{S}+\mathrm{D}\hat{X}\hat{S})+(\hat{X}\mathrm{D}\hat{S}+\mathrm{D}\hat{X}\hat{S})^{\mathrm{T}}]=H(\hat{X}\mathrm{D}\hat{S}+\mathrm{D}\hat{X}\hat{S})$$

这里$H\equiv H_{I}$，同理可得$H_{P}(\mathrm{D}X^{a}S^{a})=H(\mathrm{D}\hat{X}^{a}\mathrm{D}\hat{S}^{a})$，从而 (10)的第三个方程等价为

$$H(\hat{X}\Delta\hat{S}+\Delta\hat{X}\hat{S})=\sigma\mu_{g}I-H(\Delta\hat{X}^{a}\Delta\hat{S}^{a})-H(\hat{X}\hat{S}) \tag{11}$$

引入$\hat{E}\equiv\frac{1}{2}(\hat{S}\otimes I+I\otimes\hat{S})$，$\hat{F}\equiv\frac{1}{2}(\hat{X}\otimes I+I\otimes\hat{X})$，根据 Kroneker 积和$vec(X)$的定义，以及向量相等的条件，可以验证

$$vec(\hat{S}\Delta\hat{X}+\Delta\hat{X}\hat{S})=(\hat{S}\otimes I+I\otimes\hat{S})vec(\Delta\hat{X})$$

$$vec(\hat{X}\Delta\hat{S}+\Delta\hat{S}\hat{X})=(\hat{X}\otimes I+I\otimes\hat{X})vec(\Delta\hat{S})$$

于是

$$vec(H(\hat{X}\Delta\hat{S}+\Delta\hat{X}\hat{S}))=vec(\frac{1}{2}((\hat{X}\Delta\hat{S}+\Delta\hat{X}\hat{S})+(\hat{X}\Delta\hat{S}+\Delta\hat{X}\hat{S})^{\mathrm{T}}))$$

$$=\frac{1}{2}(\hat{S}\otimes I+I\otimes\hat{S})vec(\Delta\hat{X})+\frac{1}{2}(\hat{X}\otimes I+I\otimes\hat{X})vec(\Delta\hat{S})$$

于是矩阵方程(11)等价于如下向量方程

$$\hat{E}vec(\Delta\hat{X})+\hat{F}vec(\Delta\hat{S})=vec(\sigma\mu_{g}I-H(\Delta\hat{X}^{a}\Delta\hat{S}^{a})-H(\hat{X}\hat{S})) \tag{12}$$

易证$\mathrm{P}(X,S)=\{P\in S_{++}^{n}|\ \hat{X}\hat{S}=\hat{S}\hat{X}\}$，即$\hat{X}$和$\hat{S}$可交换，从而$\hat{E}$和$\hat{F}$可交换.

在 LP[4]中，由于 Mehrotra 型预估-校正算法为了保持迭代点始终在中心线的某个邻域内，校正步长可能会很小甚至为零.这种病态可能会导致算法实际计算的低效率和理论上不具备多项式复杂性，通过引入“安全策略”和采用削减最大预估步长的策略，最终可得到算法的多项式复杂性,本文的算法正是鉴于上述思想.

2 SDO 带有安全策略的 Mehrotra 型预估-校正算法

下面的定理给出了最大预估步长的下界,详见文献[7].

引理 1[7] 设$(X,y,S)\in N_{\infty}^{-}(\gamma)$, $(\Delta X^{a},\Delta y^{a},\Delta S^{a})$是方程(9)的解，令$G\equiv\hat{E}^{-1}\hat{F}$，则使

$X(\alpha_{a}),S(\alpha_{a})\in S_{+}^{n}$的最大预估步长$\alpha_{a}$满足

$$\alpha_{a}\geq\frac{\sqrt{\gamma^{2}+2\gamma n\sqrt{cond(G)}}-\gamma}{n\sqrt{cond(G)}}$$

推论 2 设 $\boldsymbol{P}$ 为 NT 尺度矩阵，则 $\alpha_a \ge \frac{1}{2}\sqrt{\frac{\gamma}{n}}$

证明 由于 $\boldsymbol{P}$ 为 NT 尺度矩阵，则 $\hat{\boldsymbol{X}}=\hat{\boldsymbol{S}}$，于是 $\hat{\boldsymbol{E}}=\hat{\boldsymbol{F}}$，由引理 1，知 $\mathrm{cond}(\boldsymbol{G})=1$，故

$\alpha_a \ge \frac{\sqrt{\gamma^2+2\gamma n}-\gamma}{n}$，注意到 $\sqrt{\frac{n}{\gamma}} \ge 1$，则 $\frac{\sqrt{\gamma^2+2\gamma n}-\gamma}{n}=\frac{2}{1+\sqrt{1+\frac{2n}{\gamma}}} \ge \frac{1}{2}\sqrt{\frac{\gamma}{n}}$.　　证毕

引理 3 设 $(\boldsymbol{X},y,\boldsymbol{S}) \in N_\infty^-(\gamma)$，$(\Delta\boldsymbol{X},\Delta y,\Delta\boldsymbol{S})$ 是方程组(10)的解，则

$$H_{\boldsymbol{P}}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha))=(1-\alpha)H_{\boldsymbol{P}}(\boldsymbol{XS})+\alpha\sigma\mu_g I-\alpha H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a)+\alpha^2 H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}) \tag{13}$$

$$\mu_g(\alpha)=(1-\alpha+\alpha\sigma)\mu_g \tag{14}$$

证明 由(7)式可得

$$\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha)=\boldsymbol{XS}+\alpha(\boldsymbol{X}\Delta\boldsymbol{S}+\Delta\boldsymbol{XS})+\alpha^2\Delta\boldsymbol{X}\Delta\boldsymbol{S}$$

由 $H_{\boldsymbol{P}}(\cdot)$ 的线性性及(10)的第 3 式，即可得(13)式. 因为 $Tr(H_{\boldsymbol{P}}(\boldsymbol{M}))=Tr(\boldsymbol{M})$,故

$$\boldsymbol{X}(\alpha)\bullet\boldsymbol{S}(\alpha)=Tr(H_{\boldsymbol{P}}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha)))=$$

$$(1-\alpha)Tr(H_{\boldsymbol{P}}(\boldsymbol{XS}))+\alpha\sigma\mu_g n-\alpha Tr(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a))+\alpha^2 Tr(\boldsymbol{H}_P(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))=$$

$$(1-\alpha)\boldsymbol{X}\bullet\boldsymbol{S}+\alpha\sigma\mu_g n-\alpha\Delta\boldsymbol{X}^a\bullet\Delta\boldsymbol{S}^a+\alpha^2\Delta\boldsymbol{X}\bullet\Delta\boldsymbol{S}$$

由(9)、(10)易证 $\Delta\boldsymbol{X}^a\bullet\Delta\boldsymbol{S}^a=0$，$\Delta\boldsymbol{X}\bullet\Delta\boldsymbol{S}=0$，利用(8)式，即可得(14)式。　　证毕

引理 4[7] 设给定 $(\boldsymbol{X},y,\boldsymbol{S}) \in N_\infty^-(\gamma)$，$(\Delta\boldsymbol{X},\Delta y,\Delta\boldsymbol{S})$ 是方程组(10)的解，则

$$\|H_P(\Delta\boldsymbol{X}\Delta\boldsymbol{S})\|_{\boldsymbol{F}} \le \frac{(cond(\boldsymbol{G}))^{\frac{3}{2}}}{2}\left\{[(1-2\sigma+\frac{\sigma^2}{\gamma})\gamma]^{\frac{1}{2}}+\frac{1}{4}\right\}^2\frac{n^2\mu_g}{\gamma}$$

推论 5 设 $\boldsymbol{P}$ 为 NT 尺度矩阵，若 $\sigma=(1-\alpha_a)^3$，则有

$$\|H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S})\|_{\boldsymbol{F}} \le \frac{n^2\mu_g}{\gamma}$$

证明 设 $\boldsymbol{P}$ 为 NT 尺度矩阵，$\mathrm{cond}(\boldsymbol{G})=1$，利用引理 4 及 $\gamma \in (0,\frac{1}{4})$，即可得结论.　　证毕

引理 6[7] 设 $\boldsymbol{P}$ 为 NT 尺度矩阵，定义 $t:=\max\limits_{\|u\|=1}\left\{\frac{u^T H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a)u}{u^T H_{\boldsymbol{P}}(\boldsymbol{XS})u}\right\}$，则 $t \le \frac{1}{4}$.

引理 7 设 $\boldsymbol{P}$ 为 NT 尺度矩阵，$(\boldsymbol{X},y,\boldsymbol{S}) \in N_\infty^-(\gamma)$，$(\Delta\boldsymbol{X},\Delta y,\Delta\boldsymbol{S})$ 是(10)的解，其中 $\sigma=(1-\alpha_a)^3$，

则当

$\alpha_a \in (0,1]$满足

$$\alpha_a < 1-\left(\frac{\gamma t}{1-\gamma}\right)^{\frac{1}{3}} \tag{15}$$

时，最大校正步长是严格可行的

证明 由题意，只需要在$(0,1]$上找到使$\lambda_{\min}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha)) \ge \gamma\mu_g(\alpha)$的最大校正步长$\alpha_c$，又$\lambda_{\min}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha)) \ge \lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha)))$[7],故只要找到满足

$$\lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha))) \ge \gamma\mu_g(\alpha) \tag{16}$$

的最大α_c即可. 由(13)及$\lambda_{\min}(\cdot)$在对称矩阵空间上为凹函数，可得

$$\begin{aligned}\lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha))) &= \lambda_{\min}((1-\alpha)H_{\boldsymbol{P}}(\boldsymbol{XS})+\alpha\sigma\mu_g\boldsymbol{I}-\alpha H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a)+\alpha^2 H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\\ &\ge \alpha\sigma\mu_g+\lambda_{\min}((1-\alpha)H_{\boldsymbol{P}}(\boldsymbol{XS})-\alpha H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a))+\alpha^2\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\end{aligned}$$

令$\boldsymbol{Q}(\alpha)=(1-\alpha)H_{\boldsymbol{P}}(\boldsymbol{XS})-\alpha H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a)$,由于$\boldsymbol{Q}(\alpha)$为对称矩阵，利用$\lambda_{\min}(\cdot)$的定义[8]可知$\lambda_{\min}(\boldsymbol{Q}(\alpha))=\min\limits_{\|u\|=1}u^T\boldsymbol{Q}(\alpha)u$，即存在$\bar{u}$，且$\|\bar{u}\|=1$，使$\lambda_{\min}(\boldsymbol{Q}(\alpha))=\bar{u}^T\boldsymbol{Q}(\alpha)\bar{u}$，且$\lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{XS})) \le \bar{u}^T(H_{\boldsymbol{P}}(\boldsymbol{XS}))\bar{u}$，又由$H_{\boldsymbol{P}}(\boldsymbol{XS})$是正定的且$t\ge 0$，从而由$t$的定义，可得

$$\bar{u}^T H_{\boldsymbol{P}}(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a)\bar{u} \le t\bar{u}^T H_{\boldsymbol{P}}(\boldsymbol{XS})\bar{u}$$

所以当$0\le\alpha\le\dfrac{4}{5}$（因为$0\le t\le\dfrac{1}{4}$），由上述讨论知

$$\begin{aligned}\lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{X}(\alpha)\boldsymbol{S}(\alpha))) &\ge \alpha\sigma\mu_g+(1-\alpha)\bar{u}^T H_{\boldsymbol{P}}(\boldsymbol{XS})\bar{u}-\alpha t\bar{u}^T H_{\boldsymbol{P}}(\boldsymbol{XS})\bar{u}+\alpha^2\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\\ &= \alpha\sigma\mu_g+(1-\alpha(1+t))\bar{u}^T H_{\boldsymbol{P}}(\boldsymbol{XS})\bar{u}+\alpha^2\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\\ &\ge \alpha\sigma\mu_g+(1-\alpha(1+t))\lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{XS}))+\alpha^2\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\end{aligned}$$

注意到$(\boldsymbol{X},y,\boldsymbol{S})\in N_\infty^-(\gamma)$，$\lambda_{\min}(H_{\boldsymbol{P}}(\boldsymbol{XS}))=\lambda_{\min}(\boldsymbol{XS})$[8]及(14)，故要使(16)成立，只需

$$\alpha\sigma\mu_g+[1-\alpha(1+t)]\gamma\mu_g+\alpha^2\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S})) \ge \gamma(1-\alpha+\alpha\sigma)\mu_g$$

即

$$-t\gamma+\alpha\frac{\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))}{\mu_g}+(1-\gamma)(1-\alpha_a)^3 \ge 0 \tag{17}$$

由于$Tr(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))=Tr(\Delta\boldsymbol{X}\Delta\boldsymbol{S})=\Delta\boldsymbol{X}\bullet\Delta\boldsymbol{S}=0$，故$\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))<0$，所以在最坏情况下，要使

得(17)式成立，则α_a必满足$-t\gamma+(1-\gamma)(1-\alpha_a)^3>0$，故引理结论成立.　　证毕

为使最大校正步长α_c有明确的下界,令

$$\alpha_a \le \alpha_1 := 1-\left(\frac{2\gamma t}{1-\gamma}\right)^{\frac{1}{3}} \tag{18}$$

当α_a不满足(18)时，对α_a进行削减，令$\alpha_a=\alpha_1$，然后继续校正步(可重复此步)；如果最大校正步长小于某个只与n有关的阈值时，我们令$\sigma=\dfrac{\beta}{1-\beta}$，其中$\gamma\le\beta<\dfrac{1}{4}$. 这样的选择，不仅可以保证最大校正步长的下界与$t$无关，还可以保证算法的多项式复杂性.

推论 8 设$\boldsymbol{P}$为 NT 尺度矩阵，$0<\gamma\le\beta<\dfrac{1}{4}$. 若$\sigma=\dfrac{\beta}{1-\beta}$，则$\| H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S})\|_{\boldsymbol{F}} \le \dfrac{n^2\mu_g}{3\gamma}$

证明 cond(G)=1 且$\sigma=\dfrac{\beta}{1-\beta}\le\dfrac{1}{3}$，利用引理 4 及$\gamma\in(0,\dfrac{1}{4})$，即可得结论.　　证毕

定理 9 设$0<\gamma\le\beta<\dfrac{1}{4}$，$(\boldsymbol{X},y,\boldsymbol{S})\in N_\infty^-(\gamma)$，$(\Delta\boldsymbol{X},\Delta y,\Delta\boldsymbol{S})$是(10)的解，其中$\sigma=\dfrac{\beta}{1-\beta}$，则$\alpha_c\ge\dfrac{9\gamma^2}{4n^2}$.

证明 由[8]知$\lambda_{\min}^2(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\le\| H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S})\|_{\boldsymbol{F}}^2$，故$\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\ge-\| H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S})\|_{\boldsymbol{F}}$,由推论 8 可得$\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))\ge-\dfrac{n^2\mu_g}{3\gamma}$，由(17)知，只要

$$-t\gamma+\alpha\frac{\lambda_{\min}(H_{\boldsymbol{P}}(\Delta\boldsymbol{X}\Delta\boldsymbol{S}))}{\mu_g}+(1-\gamma)\frac{\beta}{1-\beta}>0 \tag{19}$$

即可保证$(\boldsymbol{X}(\alpha),y(\alpha),\boldsymbol{S}(\alpha))\in N_\infty^-(\gamma)$，又因$0\le t\le\dfrac{1}{4}$，故只要$\alpha$满足$-\dfrac{1}{4}\gamma-\alpha\dfrac{n^2}{3\gamma}+(1-\gamma)\dfrac{\beta}{1-\beta}>0$

即$\alpha\le\dfrac{9\gamma^2}{4n^2}$时，(19)成立，从而最大可行步长$\alpha_c\ge\dfrac{9\gamma^2}{4n^2}$，从而定理得证.　　证毕

下面给出 SDO 带有安全策略的 Mehrotra 型预估-校正算法的步骤：

设邻近参数$\gamma\in(0,\dfrac{1}{4})$，安全参数$\beta\in[\gamma,\dfrac{1}{4})$，精度$\varepsilon>0$，初始点$(\boldsymbol{X}^0,y^0,\boldsymbol{S}^0)\in N_\infty^-(\gamma)$

步骤 1：若$\boldsymbol{X}\bullet\boldsymbol{S}<\grave{o}$，则算法终止；否则转步骤 2；

步骤 2：计算 NT 尺度矩阵 $\boldsymbol{P}=(\boldsymbol{X}^{-1/2}(\boldsymbol{X}^{1/2}\boldsymbol{S}\boldsymbol{X}^{1/2})^{1/2}\boldsymbol{X}^{-1/2})^{1/2}$；

步骤 3：（预估步）解方程(9)，计算最大预估步长 α_a，使 $(\boldsymbol{X}(\alpha_a),\boldsymbol{S}(\alpha_a))\in\boldsymbol{S}_+^n\times\boldsymbol{S}_+^n$；

步骤 4：（校正步）若 $\alpha_a>\alpha_1$，则令 $\alpha_a=\alpha_1$(削减策略)；解方程(10)，其中 $\sigma=(1-\alpha_a)^3$,计算最大可行步长 α_c，使得 $(\boldsymbol{X}(\alpha_c),y(\alpha_c),\boldsymbol{S}(\alpha_c))\in N_\infty^-(\gamma)$. 若 $\alpha_c<\dfrac{9\gamma^2}{4n^2}$，解方程(10)，其中 $\sigma=\dfrac{\beta}{1-\beta}$ 计算最大可行步长 α_c，使得 $(\boldsymbol{X}(\alpha_c),y(\alpha_c),\boldsymbol{S}(\alpha_c))\in N_\infty^-(\gamma)$；

步骤 5：令 $(\boldsymbol{X},y,\boldsymbol{S})=(\boldsymbol{X}(\alpha_c),y(\alpha_c),\boldsymbol{S}(\alpha_c))$，转步骤 1。

下面的定理给出了上述算法的多项式迭代复杂界.

定理 10 算法至多经过 $O\left(n^{\frac{5}{2}}\log\dfrac{\boldsymbol{X}^0\bullet\boldsymbol{S}^0}{\varepsilon}\right)$ 次迭代，便可得到满足精度 $\boldsymbol{X}\bullet\boldsymbol{S}<\varepsilon$ 的解.

证明 若 $\alpha_a\ge\alpha_1$ 且 $\alpha_c\ge\dfrac{9\gamma^2}{4n^2}$，由(14)可

$$\mu_g(\alpha)=(1-\alpha+\alpha\sigma)\mu_g=\left(1-\alpha+\alpha\frac{2\gamma t}{1-\gamma}\right)\mu_g\le\left[1-\frac{9\gamma^2(2-3r)}{4(2-2r)n^2}\right]\mu_g$$

若 $\alpha_a\ge\alpha_1$ 且 $\alpha_c<\dfrac{9\gamma^2}{4n^2}$， 有

$$\mu_g(\alpha)=(1-\alpha+\alpha\sigma)\mu_g=\left(1-\alpha+\alpha\frac{\beta}{1-\beta}\right)\mu_g\le\left[1-\frac{9\gamma^2(1-2\beta)}{4(1-\beta)n^2}\right]\mu_g$$

若 $\alpha_a<\alpha_1$ 且 $\alpha_c\ge\dfrac{9\gamma^2}{4n^2}$ 时，由(14)及推论 2，可得

$$\mu_g(\alpha)=(1-\alpha+\alpha\sigma)\mu_g=\left(1-\alpha+\alpha(1-\alpha_a)^3\right)\mu_g\le\left[1-\frac{9\gamma^{\frac{5}{2}}}{8n^{\frac{5}{2}}}\right]\mu_g$$

由文献[8]的引理 6, 可得定理结论.　　　　证毕

参考文献:

[1] Karmarkar N K. A new polynomial-time algorithm for linear programming[J]. Combinatorica,1984，4(4): 373-395.

[2] Boyd S, Haoui L EI, Feron G E. Linear Matrix Inequalities in System and Control Theory[M]. SIAM, Philadelphia, PA,1994

[3] Mehrotra S. On the implementation of a (primal-dual) interior point method[J]. SIAM Journal on Optimization 1992，2(2): 575-601

[4] Salahi M, Terlaky T. Mehrotra-type predictor corrector algorithm revisited[J]. Optimization Methods and Software, 2008, 23(2): 259-273

[5] Zhang Y. On extending some primal-dual interior point from linear programming to semidefinite programming[J]. SIAM Journal on Optimization, 1998，8(2): 365-383.

[6] Nesterov Y E, Todd M J. Self-scaled barriers and interior-point method for convex programming[J]. Mathematics of Operations Research, 1997, 22(1): 1-42

[7] Koulaei M H, Terlaky T. On the extension of a Mehrotra-type algorithm for semi-definite optimization[M]. Ontario,Canada.Advanced optimization Lab.，Department of Computing and Software, Memaster University, Hamilton, Technical Report, 2007

[8] Monteiro R D C, Zhang Y. A unified analysis for a class of long step primal-dual path following interior-point algorithms for semidefinite programming[J].Mathematical Programming , 1998,81 (3): 281–299

Operations Research and Cybernetics

A Mehrotra-Type Predictor-Corrector Algorithm for Semidefinite Optimization

CHEN Huaping

(Department of Mathematics, Liupanshui Normal University, Liupanshui Guizhou 553004, China)

Abstract: The extension of Mehrotra-type predictor-corrector algorithm which was proposed by Salahi for linear optimization is obtained for semidefinite optimization. Firstly ,the basic theories of algorithm based on Mehrotra-type predictor-corrector algorithm for Semi-definite are introduced, especially the selection of symmetrical technique. Then followed by giving the iteration complexity analysis of this algorithm, the important thought of the algorithm is explained: the safeguard strategy is used in the algorithm, the above bound on the maximum feasible step in the predictor step is given. A cut strategy for the maximum feasible step in the predictor step is provided：when the maximum prediction step size is greater than a certain threshold, this step is cut (repeated), so the lower bound in the corrector step is obtained. Based on these strategy and the NT search direction, the polynomial complexity of new algorithm is obtained.

Key words: large-update method；mehrotra-type predictor-corrector algorithm；semidefinite optimization；polynomial complexity