网络优先数字出版时间:2016-11-21 【DOI】10.13656/j.cnki.gxkx.20161121.007
网络优先数字出版地址:http://www.cnki.net/kcms/detail/45.1206.G3.20161121.1520.014.html

二次半定规划一个原始对偶路径跟踪算法*

A Primal-dual Path-following Algorithm for Quadratic Semi-definite Programming

黎健玲**,王培培
LI Jianling,WANG Peipei

(广西大学数学与信息科学学院,广西南宁 530004)
(College of Mathematics and Information Science,Guangxi University,Nanning,Guangxi,530004,China)

摘要:本文提出求解二次半定规划的一个基于 H..K..M 方向的原始对偶路径跟踪算法.文中首先导出确定 H..K..M 方向的线性方程组,并证明该搜索方向的存在唯一性;然后给出算法的具体步骤,并证明算法产生的迭代点列落在中心路径的某个邻域内.最后采用 Matlab(R2011b)数学软件编程对算法进行数值试验.数值结果表明算法是有效的.

关键词:二次半定规划 原始对偶 算法 路径跟踪 中心路径
中图分类号:C934 **文献标识码**:A **文章编号**:1005-9164(2016)05-0396-08

Abstract:A primal-dual path-following algorithm based on H..K..M direction for quadratic semi-definite programming problems(QSDP) is proposed. Firstly,the system of linear equations yielding the H..K..M direction are derived,and the existence and uniqueness of the search direction are shown;Secondly,the algorithm is described in detail. We show that the iterates generated by the algorithm can fall into some neighborhood of the central path under some mild conditions. Finally,a preliminary numerical experiment is performed for the algorithm by using Matlab (R2011b) mathematical software,and the numerical results show that the proposed algorithm is effective.

Key words:quadratic semi-definite programming,primal-dual,algorithm,path-following,central path

0 引言

【研究意义】(线性)半定规划(简记 SDP)既是定义在半正定矩阵锥上的规划,也是线性规划、凸二次规划、二阶锥规划等的一种推广,在控制论、组合优化等诸多领域有着广泛的应用.而二次半定规划是半定规划的拓展,它在最优控制、证券、金融风险分析等领域中都有着广泛的应用.【前人研究进展】原始对偶内点算法是求解半定规划的有效方法[1-10].由于确定搜索方向的方法不同,因此导出多种形式的原始对偶内点法.目前常用的搜索方向有以下 3 种:(i)AHO 搜索方向[2];(ii)H..K..M 搜索方向[3-5];(iii)Nesterov-Todd 搜索方向(简称 NT 方向)[6-7].基于 SDP 的原始对偶内点算法,文献[11]提出一个求解二次半定规划的势减少(potential reduction)算法;文献[12]结合 Dikin 型方向和牛顿中心步提出一个

收稿日期:2016-08-05
修回日期:2016-09-25
作者简介:黎健玲(1965-),女,博士,教授,主要从事最优化理论与方法研究,E-mail:jianlingli@126.com。
*国家自然科学基金项目(No.11561005)和广西自然科学基金项目(2016GXNSFAA380248,2014GXSFFA118001)资助。
**通信作者。

求解二次半定规划的预估校正算法；文献[13]提出一个求解二次半定规划的投影收缩算法；文献[14]将基于 NT 方向的原始对偶内点算法推广到二次半定规划.**【本研究切入点】**本文考虑如下二次半定规划问题(简记为 QSDP)：

$$\min f(X)=\frac{1}{2}\mathrm{svec}(X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)-a^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)+C\cdot X \tag{0.1}$$

$$\text{s.t.}\quad A_i\cdot X=b_i,i=1,\cdots,m;$$
$$X^{\mathrm{T}}=X\succeq 0,$$

其中 $\mathbb{H}^{\mathrm{T}}=(\mathrm{svec}(H_1),\cdots,\mathrm{svec}(H_l))\in\mathbb{R}^{\frac{1}{2}n(n+1)\times l}$，$H_j(j=1,\cdots,l)$，$A_i(i=1,\cdots,m)$，$C$ 和 X 都是 $n\times n$ 阶实对称矩阵，$a\in\mathbb{R}^l$，$b\in\mathbb{R}^m$，算子 $\mathrm{svec}(X)$ 将 $n\times n$ 阶对称阵 X 的列堆叠(stack) 成一个 $\frac{1}{2}n(n+1)$ 维的列向量，它的逆算子记为 smat. $X\succeq(\succ)0$ 表示 X 是半正定(正定)矩阵. 对任意 $A\in\mathbb{R}^{p\times q}$，$B\in\mathbb{R}^{p\times q}$，矩阵内积 $A\cdot B=\mathrm{Tr}(A^{\mathrm{T}}B)=\sum_{i=1}^{p}\sum_{j=1}^{q}a_{ij}b_{ij}$，其中 $\mathrm{Tr}(\cdot)$ 表示矩阵的迹. 若 A，B 为 n 阶对称阵，则有 $A\cdot B=\mathrm{svec}(A)^{\mathrm{T}}\mathrm{svec}(B)$. 分别记 S^n，S_+^n 和 S_{++}^n 为 n 阶实对称矩阵全体，n 阶实对称半正定矩阵全体和 n 阶实对称正定矩阵的全体. 因为 $\mathbb{H}^{\mathrm{T}}\mathbb{H}$ 是对称半正定矩阵，所以(0.1)式的目标函数是凸的，而约束函数是非光滑但也是凸的，因此 QSDP(0.1)是一个凸优化问题. 据此，本文基于文献[5]的思想，提出一个求解 QSDP(0.1)的原始对偶路径跟踪算法.**【拟解决的关键问题】**首先，利用 Wolfe 对偶理论得到 QSDP(0.1)的对偶规划，从而导出对偶间隙的表达式，并据此定义中心路径. 然后结合最优性条件、矩阵对称化技术以及牛顿法导出确定 H..K..M 方向的线性方程组，并证明该搜索方向的存在唯一性. 随后基于该搜索方向给出短步路径跟踪算法的具体步骤，并证明算法产生的迭代点列落在中心路径的某个邻域内.

1 二次半定规划问题的对偶规划

利用凸规划问题的 Wolfe 对偶理论，可得 QSDP(0.1)的对偶规划(简记为 DSDP)为

$$\max D(X,y)=-\frac{1}{2}\mathrm{svec}(X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)+\sum_{i=1}^{m}b_i y_i$$

$$\text{s.t.}\ \sum_{i=1}^{m}y_iA_i+Z=C-\sum_{j=1}^{l}a_jH_j+\sum_{j=1}^{l}H_j(H_j\cdot X),$$

$$X\succeq 0,Z\succeq 0. \tag{1.1}$$

分别记问题(0.1)，(1.1) 的可行集为 $\mathbb{F}_P$，$\mathbb{F}_D$，即

$$\mathbb{F}_P=\{X\mid A_i\cdot X=b_i,X\succeq 0\},$$

$$\mathbb{F}_D=\{(X,y,Z)\mid \sum_{i=1}^{m}y_iA_i+Z=C-\sum_{j=1}^{l}a_jH_j+\sum_{j=1}^{l}H_j(H_j\cdot X),X\succeq 0,Z\succeq 0\}.$$

记 $\mathbb{F}_P$ 和 $\mathbb{F}_D$ 的严格内部为 $\mathbb{F}^{\circ}{}_P$ 和 $\mathbb{F}^{\circ}{}_D$.

易知 DSDP(1.1) 也是一个凸二次半定规划问题. 对于任意的 $(X,y,Z)\in\mathbb{F}_D$，若 $X\in\mathbb{F}_P$，则对偶间隙为

$$\begin{aligned}
f(X)-D(X,y)&=\mathrm{svec}(X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)-a^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)+C\cdot X-\sum_{i=1}^{m}b_iy_i,\\
&=\mathrm{svec}(X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)-a^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)-\sum_{i=1}^{m}b_iy_i+(\sum_{i=1}^{m}y_iA_i+Z+\sum_{j=1}^{l}a_jH_j-\sum_{j=1}^{l}H_j(H_j\cdot X))\cdot X,\\
&=\mathrm{svec}(X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(X)-\sum_{j=1}^{l}(H_j\cdot X)^2+Z\cdot X=(\mathbb{H}\mathrm{svec}(X))^{\mathrm{T}}(\mathbb{H}\mathrm{svec}(X))-\sum_{j=1}^{l}(H_j\cdot X)^2+Z\cdot X\\
&=Z\cdot X.
\end{aligned} \tag{1.2}$$

2 H..K..M 搜索方向

假设：

(A1)设矩阵组 $\{A_i,i=1,\cdots,m\}$ 线性无关.

(A2)Slater 条件成立，即存在 $X\succ 0$，$Z\succ 0$ 和 $y\in\mathbb{R}^m$ 使得 $X\in\mathbb{F}_P$，$(X,y,Z)\in\mathbb{F}_D$.

由(1.2) 式以及文献[13] 的定理 3.5 知，在假设(A2) 下，$X\in\mathbb{F}_P$ 是 QSDP(0.1) 的最优解当且仅当存在 $y\in\mathbb{R}^m$，$Z\in S^n$，使得

$$\sum_{i=1}^{m}y_iA_i+Z=C-\sum_{j=1}^{l}a_jH_j+\sum_{j=1}^{l}H_j(H_j\cdot X), \tag{2.1a}$$

$$XZ=0. \tag{2.1b}$$

点对 (X,Z) 如果满足 $X\in\mathbb{F}_P$，$(X,y,Z)\in\mathbb{F}_D$ 且

$$XZ=\sigma\mu I,$$

其中参数 $\mu>0$，$\mu=\frac{X\cdot Z}{n}$，σ 是中心参数，I 是 n 阶单位阵，则称点对 (X,Z) 在中心路径上，即中心路径上的点对 (X,Z) 满足：

$$A_i\cdot X=b_i,i=1,\cdots,m, \tag{2.2a}$$

$$\sum_{i=1}^{m}y_iA_i+Z=C-\sum_{j=1}^{l}a_jH_j+\sum_{j=1}^{l}H_j(H_j\cdot X), \tag{2.2b}$$

$$XZ=\sigma\mu I. \tag{2.2c}$$

设当前迭代点为(X,y,Z)，且 $X\in \mathbb{F}^{\circ}_P$，$(y,Z)\in \mathbb{F}^{\circ}_D$，用一步 Newton 法求解非线性方程组(2.2)得到的系统如下：

$$A_i\cdot\Delta X=0, i=1,\cdots,m, \tag{2.3a}$$

$$\sum_{i=1}^{m}\Delta y_iA_i+\Delta Z-\sum_{j=1}^{l}H_j(H_j\cdot\Delta X)=0, \tag{2.3b}$$

$$\Delta XZ+X\Delta Z=\sigma\mu I-XZ-\Delta X\Delta Z. \tag{2.3c}$$

忽略(2.3c)式中的 $\Delta X\Delta Z$，得到

$$\Delta XZ+X\Delta Z=\sigma\mu I-XZ. \tag{2.4}$$

注意到 X,Z 不可交换，即 $XZ\neq ZX$，而内点法的一个关键是产生的矩阵 $\Delta X,\Delta Z$ 要满足对称性. 文献[9]给出对称化矩阵 M 的一般通式：

$$H_P(M)=\frac{1}{2}(PMP^{-1}+(PMP^{-1})^{\mathrm{T}}),$$

其中 P 是可逆阵，有多种不同的取法[1-5]. 本文采用文献[1,5]中的方法，取 $P=X^{-\frac{1}{2}}$. 对称化方程(2.4)，得到如下的线性方程组：

$$A_i\cdot\Delta X=0, i=1,\cdots,m, \tag{2.5a}$$

$$\sum_{i=1}^{m}\Delta y_iA_i+\Delta Z-\sum_{j=1}^{l}H_j(H_j\cdot\Delta X)=0, \tag{2.5b}$$

$$X^{-\frac{1}{2}}(X\Delta Z+\Delta XZ)X^{\frac{1}{2}}+X^{\frac{1}{2}}(\Delta ZX+Z\Delta X)X^{-\frac{1}{2}}=2(\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}}). \tag{2.5c}$$

记 $W=\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}}$，求解线性方程组(2.5)得到的解$(\Delta X,\Delta y,\Delta Z)$称为 H..K..M 搜索方向. 记

$$G^{\mathrm{T}}=(\mathrm{svec}(A_1),\cdots,\mathrm{svec}(A_m)), \tag{2.6a}$$

$$r_c=\mathrm{svec}(\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}})=\mathrm{svec}W, \tag{2.6b}$$

$$E=X^{-\frac{1}{2}}\otimes_s X^{\frac{1}{2}}Z=(I\otimes_s X^{\frac{1}{2}}ZX^{\frac{1}{2}})\cdot(X^{-\frac{1}{2}}\otimes_s X^{-\frac{1}{2}}), \tag{2.6c}$$

$$F=X^{\frac{1}{2}}\otimes_s X^{\frac{1}{2}}. \tag{2.6d}$$

其中 $X^{\frac{1}{2}}\otimes_s X^{\frac{1}{2}}$ 表示两个矩阵的对称 Kronecker 积(文献[1]附录). 使用算子 svec，可把线性方程组(2.5)改写成以下形式：

$$\begin{pmatrix}0 & G & 0\\ G^{\mathrm{T}} & -H^{\mathrm{T}}H & I\\ 0 & E & F\end{pmatrix}\begin{pmatrix}\Delta y\\ \mathrm{svec}(\Delta X)\\ \mathrm{svec}(\Delta Z)\end{pmatrix}=\begin{pmatrix}0\\0\\r_c\end{pmatrix}, \tag{2.7}$$

其中 I 是$\frac{n(n+1)}{2}$阶单位阵. 根据假设(A1)知 G 是行满秩的.

引理 2.1 假设(A1)和(A2)成立，$X,Z\in S^n_{++}$，则矩阵 $FH^{\mathrm{T}}H+E$ 可逆.

证明 已知

$$E=X^{-\frac{1}{2}}\otimes_s X^{\frac{1}{2}}Z=(I\otimes_s X^{\frac{1}{2}}ZX^{\frac{1}{2}})\cdot(X^{-\frac{1}{2}}\otimes_s X^{-\frac{1}{2}}),$$

$$F=X^{\frac{1}{2}}\otimes_s X^{\frac{1}{2}},$$

由 X,Z 是对称正定阵知 $X^{\frac{1}{2}},X^{-\frac{1}{2}}$ 也是对称正定阵，进而由对称 Kronecker 积的性质知 E,F 可逆. 而

$$\mathrm{rank}(FH^{\mathrm{T}}H+E)=\mathrm{rank}(F^{-1}(FH^{\mathrm{T}}H+E))=\mathrm{rank}(H^{\mathrm{T}}H+F^{-1}E), \tag{2.8}$$

往证 $F^{-1}E$ 是对称正定矩阵. 由对称 Kronecker 积的性质(文献[1]附录)知

$$F^{-1}E=(X^{\frac{1}{2}}\otimes_s X^{\frac{1}{2}})^{-1}(X^{-\frac{1}{2}}\otimes_s X^{\frac{1}{2}}Z)=(X^{-\frac{1}{2}}\otimes_s X^{-\frac{1}{2}})(X^{-\frac{1}{2}}\otimes_s X^{\frac{1}{2}}Z)=\frac{1}{2}[(X^{-1}\otimes_s Z)+(Z\otimes_s X^{-1})]=X^{-1}\otimes_s Z, \tag{2.9}$$

因 X^{-1} 和 Z 是对称正定阵，故 $X^{-1}\otimes_s Z$，即 $F^{-1}E$ 是对称正定阵. 又因为 $H^{\mathrm{T}}H$ 是对称半正定阵，因此 $H^{\mathrm{T}}H+F^{-1}E$ 是对称正定阵，结合(2.8)式即知矩阵 $FH^{\mathrm{T}}H+E$ 是可逆的.

引理 2.2 假设(A1)和(A2)成立，$X,Z\in S^n_{++}$，则矩阵$(FH^{\mathrm{T}}H+E)^{-1}F\in S^n_{++}$.

证明 因为

$$(FH^{\mathrm{T}}H+E)^{-1}F=(F^{-1}(FH^{\mathrm{T}}H+E))^{-1}=(H^{\mathrm{T}}H+F^{-1}E)^{-1},$$

由引理 2.1 的证明过程知矩阵 $H^{\mathrm{T}}H+F^{-1}E$ 是对称正定阵，所以$(FH^{\mathrm{T}}H+E)^{-1}F$ 也是对称正定阵，即$(FH^{\mathrm{T}}H+E)^{-1}F\in S^n_{++}$.

引理 2.3 假设(A1)和(A2)成立，$X\in \mathbb{F}^{\circ}_P$，$(X,y,Z)\in \mathbb{F}^{\circ}_D$，则线性方程组(2.7)有唯一解.

证明 首先证明线性方程组(2.7)的系数矩阵可逆，即证明方程组(2.7)对应的齐次线性方程组

$$\begin{pmatrix}0 & G & 0\\ G^{\mathrm{T}} & -H^{\mathrm{T}}H & I\\ 0 & E & F\end{pmatrix}\begin{pmatrix}\Delta y\\ \mathrm{svec}(\Delta X)\\ \mathrm{svec}(\Delta Z)\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{2.10}$$

只有平凡解. 记

$$B=\begin{pmatrix}G(FH^{\mathrm{T}}H+E)^{-1}FG^{\mathrm{T}} & 0 & 0\\ G^{\mathrm{T}} & -F^{-1}E-H^{\mathrm{T}}H & 0\\ 0 & E & F\end{pmatrix},$$

利用块高斯消去法对(2.10)式进行化简，得到

$$B\begin{pmatrix}\Delta y\\ \mathrm{svec}(\Delta X)\\ \mathrm{svec}(\Delta Z)\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{2.11}$$

于是有

$$(G(FH^{\mathrm{T}}H+E)^{-1}FG^{\mathrm{T}})\Delta y=0. \tag{2.12}$$

由引理 2.2 知$(FH^{\mathrm{T}}H+E)^{-1}F$ 是对称正定阵，又因

为 G 为行满秩,所以 $G(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}FG^{\mathrm{T}}$ 也是对称正定阵,于是由(2.12)式知 $\Delta y=0$.

将 $\Delta y=0$ 代入(2.11)式,即得

$$\mathrm{svec}(\Delta X)=(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}FG^{\mathrm{T}}\Delta y=0,$$

$$\mathrm{svec}(\Delta Z)=-F^{-1}E\ \mathrm{svec}(\Delta X)=0,$$

所以(2.10)式只有平凡解,从而(2.7)式的解存在并且唯一.

引理 2.4 假设(A1)和(A2)成立,设 $X\in\mathbb{F}^{\circ}{}_{P}$,$(X,y,Z)\in\mathbb{F}^{\circ}{}_{D}$,$(\Delta X,\Delta y,\Delta Z)$ 是线性方程组(2.7)对于某个给定的矩阵 $W(\in\mathbb{R}^{n\times n})$ 的解,则以下结论成立:

①$\Delta Z\bullet\Delta X\geqslant 0$;

②$X\bullet\Delta Z+Z\bullet\Delta X=\mathrm{Tr}(W)$;

③ 如果 $W=\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}}$,$\sigma\in\mathbb{R}$,$\mu=\dfrac{X\bullet Z}{n}$,则

$$(X+\alpha\Delta X)\bullet(Z+\alpha\Delta Z)=(1-\alpha+\alpha\sigma)(X\bullet Z)+\alpha^2(\Delta X\bullet\Delta Z).$$

证明 ① 根据(2.5b)式和(2.5a)式,并结合矩阵内积的性质可得

$$\Delta Z\bullet\Delta X=(\sum_{j=1}^{l}H_j(H_j\bullet\Delta X)-\sum_{i=1}^{m}\Delta y_i A_i)\bullet\Delta X=\mathrm{svec}(\Delta X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(\Delta X)-\sum_{i=1}^{m}\Delta y_i(A_i\bullet\Delta X)=\mathrm{svec}(\Delta X)^{\mathrm{T}}\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(\Delta X)\geqslant 0,$$

最后一个不等式成立是因为 $\mathbb{H}^{\mathrm{T}}\mathbb{H}$ 是对称半正定矩阵.

② 由(2.5c)式知 $X^{-\frac{1}{2}}(X\Delta Z+\Delta XZ)X^{\frac{1}{2}}+X^{\frac{1}{2}}(\Delta ZX+Z\Delta X)X^{-\frac{1}{2}}=2W$,则由矩阵迹的性质有

$$2\mathrm{Tr}(W)=\mathrm{Tr}(X^{-\frac{1}{2}}(X\Delta Z+\Delta XZ)X^{\frac{1}{2}})+\mathrm{Tr}(X^{\frac{1}{2}}(\Delta ZX+Z\Delta X)X^{-\frac{1}{2}})=\mathrm{Tr}(X\Delta Z+\Delta XZ)+\mathrm{Tr}(\Delta ZX+Z\Delta X)=2\mathrm{Tr}(X\Delta Z+\Delta XZ)=2(X\bullet\Delta Z+Z\bullet\Delta X),$$

即 $X\bullet\Delta Z+Z\bullet\Delta X=\mathrm{Tr}(W)$.

③ 根据矩阵内积的线性性质以及矩阵迹的性质,再结合结论(2)得到

$$(X+\alpha\Delta X)\bullet(Z+\alpha\Delta Z)=X\bullet Z+\alpha(X\bullet\Delta Z+Z\bullet\Delta X)+\alpha^2(\Delta X\bullet\Delta Z)=X\bullet Z+\alpha\mathrm{Tr}(\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}})+\alpha^2(\Delta X\bullet\Delta Z)=X\bullet Z+\alpha\mathrm{Tr}(\sigma\mu I)-\alpha\mathrm{Tr}(X^{\frac{1}{2}}ZX^{\frac{1}{2}})+\alpha^2(\Delta X\bullet\Delta Z),$$

又因为 $\mathrm{Tr}(\sigma\mu I)=\sigma\mu\mathrm{Tr}(I)=\sigma\mu n$,而 $\mu=\dfrac{X\bullet Z}{n}$,所以有

$$(X+\alpha\Delta X)\bullet(Z+\alpha\Delta Z)=X\bullet Z+\alpha\sigma\mu n-\alpha\mathrm{Tr}(XZ)+\alpha^2(\Delta X\bullet\Delta Z)=X\bullet Z+\alpha\sigma\mu n-\alpha(X\bullet Z)+\alpha^2(\Delta X\bullet\Delta Z)=(1-\alpha+\alpha\sigma)(X\bullet Z)+\alpha^2(\Delta X\bullet\Delta Z),$$

即结论(3)成立.

3 H..K..M 搜索方向的计算

H..K..M 搜索方向 $(\Delta X,\Delta y,\Delta Z)$ 可以通过求解线性方程组(2.7)得到,而方程组(2.7)包含 $m+n(n+1)$ 个线性方程,通过块高斯消去法将其化简为如下的方程

$$(G(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}FG^{\mathrm{T}})\Delta y=-G(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}r_c,\tag{3.1a}$$

$$G^{\mathrm{T}}\Delta y-(F^{-1}E+\mathbb{H}^{\mathrm{T}}\mathbb{H})\mathrm{svec}(\Delta X)=-F^{-1}r_c,\tag{3.1b}$$

$$E\mathrm{svec}(\Delta X)+F\mathrm{svec}(\Delta Z)=r_c,\tag{3.1c}$$

记 $M=G(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}FG^{\mathrm{T}}$,$h=-G(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}r_c$,则(3.1a)式可改写为

$$M\Delta y=h.\tag{3.2}$$

可证 M 是对称正定阵.事实上,因为

$$M=G(F^{-1}(F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E))^{-1}G^{\mathrm{T}}=G(\mathbb{H}^{\mathrm{T}}\mathbb{H}+F^{-1}E)^{-1}G^{\mathrm{T}}=G(\mathbb{H}^{\mathrm{T}}\mathbb{H}+(X^{-1}\otimes_s Z))^{-1}G^{\mathrm{T}},$$

上述最后一个等式成立是由(2.9)式得到.由 X^{-1},Z 是对称正定阵知 $X^{-1}\otimes_s Z$ 也是对称正定阵,而 $\mathbb{H}^{\mathrm{T}}\mathbb{H}$ 是对称半正定阵,因此 $(\mathbb{H}^{\mathrm{T}}\mathbb{H}+(X^{-1}\otimes_s Z))^{-1}$ 是对称正定阵,故 M 是对称正定阵.

于是可利用 Cholesky 分解求解线性方程组(3.2)得到 Δy,进而由(3.1b)式和(3.1c)式可求出

$$\Delta X=\mathrm{smat}((F\mathbb{H}^{\mathrm{T}}\mathbb{H}+E)^{-1}(FG^{\mathrm{T}}\Delta y+r_c)),\tag{3.3a}$$

$$\Delta Z=\mathrm{smat}(\mathbb{H}^{\mathrm{T}}\mathbb{H}\mathrm{svec}(\Delta X)-G^{\mathrm{T}}\Delta y).\tag{3.3b}$$

4 短步原始对偶路径跟踪算法

这一节我们将给出基于 H..K..M 方向的短步原始对偶路径跟踪算法.

算法的具体步骤如下:

算法 A

步骤 0(初始步) 选取 $X^0\in\mathbb{F}^{\circ}{}_{P}$,$(X^0,y^0,Z^0)\in\mathbb{F}^{\circ}{}_{D}$,参数 $\eta>1$,令 $k=0$,$\mu_0=\dfrac{X^0\bullet Z^0}{n}$.

当 $\mu_k>2^{-\eta}\mu_0$ 时执行以下步骤:

步骤 1 记 $X=X^k$,$(y,Z)=(y^k,Z^k)$,$\mu=\mu_k$.

步骤 2 选择合适参数 $\sigma=\sigma_k\in[0,1]$,令 $W=\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}}$.

步骤 3 求解线性方程组(2.7)得到 H..K..M 搜索方向 $(\Delta X,\Delta y,\Delta Z)$.

步骤 4 选择 $\alpha\geqslant 0$,使得

$$\hat{X}=X+\alpha\Delta X\in S^n_{++},\ \hat{y}=y+\alpha\Delta y\in\mathbb{R}^m,$$

$\hat{Z}=Z+\alpha\Delta Z\in S^n_{++}$.

步骤 5　令 $X^{k+1}=\hat{X}$,$(y^{k+1},Z^{k+1})=(\hat{y},\hat{Z})$,$\mu=\mu_{k+1}=\frac{X^{k+1}\cdot Z^{k+1}}{n}$. 令 $k=k+1$.

注:算法 A 是一个短步路径跟踪算法[5,8],对所有的 $k\geqslant 0$,取 $\alpha_k=1$,$\sigma_k=1-\delta/\sqrt{n}$,其中 $\delta>0$ 是一个常量.

算法产生的迭代点列将落在中心路径的如下邻域内:

$$\mathbb{N}_F(\gamma)\equiv\{(X,y,Z)\mid X\in\mathbb{F}^\circ_P,(X,y,Z)\in\mathbb{F}^\circ_D,\|X^{\frac{1}{2}}ZX^{\frac{1}{2}}-\mu I\|_F\leqslant\gamma\mu\}=\{(X,y,Z)\mid X\in\mathbb{F}^\circ_P,(X,y,Z)\in\mathbb{F}^\circ_D,(\sum_{i=1}^n(\lambda_i(XZ)-\mu)^2)^{1/2}\leqslant\gamma\mu\},\tag{4.1}$$

其中 $\mu=\frac{X\cdot Z}{n}$,$\gamma\in(0,1)$,$\|\cdot\|_F$ 表示矩阵的 Frobenius 范数.

引理 4.1　假设(A1)和(A2)成立,设 $X\in\mathbb{F}^\circ_P$,$(X,y,Z)\in\mathbb{F}^\circ_D$,$(\Delta X,\Delta Z,\Delta y)$ 是线性方程组(2.7)的解,其中 $W=\sigma\mu I-X^{\frac{1}{2}}ZX^{\frac{1}{2}}$,对任意的 $\alpha\geqslant 0$,令

$$(X(\alpha),Z(\alpha),y(\alpha))\equiv(X,Z,y)+\alpha(\Delta X,\Delta Z,\Delta y),\tag{4.2a}$$

$$\mu(\alpha)\equiv(X(\alpha)\cdot Z(\alpha))/n,\tag{4.2b}$$

$$Q(\alpha)\equiv X^{-\frac{1}{2}}(X(\alpha)Z(\alpha)-\mu(\alpha)I)X^{\frac{1}{2}},\tag{4.2c}$$

则有

$$Q(\alpha)+Q(\alpha)^{\mathrm{T}}=2(1-\alpha)(X^{\frac{1}{2}}ZX^{\frac{1}{2}}-\mu I)-\frac{2\alpha^2}{n}(\Delta X\cdot\Delta Z)I+\alpha^2(X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}+X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}).\tag{4.3}$$

其中 I 为 n 阶单位阵.

证明　假设 $\alpha\geqslant 0$ 是给定的. 根据引理 2.4(3)有

$$X(\alpha)\cdot Z(\alpha)=(X+\alpha\Delta X)\cdot(Z+\alpha\Delta Z)=(1-\alpha+\alpha\sigma)(X\cdot Z)+\alpha^2(\Delta X\cdot\Delta Z),$$

所以由(4.2b)即知

$$\mu(\alpha)=(1-\alpha+\alpha\sigma)\mu+\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z),$$

因此,

$$X(\alpha)Z(\alpha)-\mu(\alpha)I=(X+\alpha\Delta X)(Z+\alpha\Delta Z)-((1-\alpha+\alpha\sigma)\mu+\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z))I=XZ+\alpha X\Delta Z+\alpha\Delta XZ+\alpha^2\Delta X\Delta Z-(1-\alpha)\mu I-\alpha\sigma\mu I-\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z)I=XZ-\alpha XZ-(1-\alpha)\mu I-\alpha\sigma\mu I+\alpha XZ+\alpha(X\Delta Z+\Delta XZ)+\alpha^2\Delta X\Delta Z-\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z)I=(1-\alpha)(XZ-\mu I)+\alpha(XZ-\sigma\mu I)+\alpha(X\Delta Z+\Delta XZ)+\alpha^2\Delta X\Delta Z-\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z)I.\tag{4.4}$$

类似可以得到

$$Z(\alpha)X(\alpha)-\mu(\alpha)I=(1-\alpha)(ZX-\mu I)+\alpha(ZX-\sigma\mu I)+\alpha(\Delta ZX+Z\Delta X)+\alpha^2\Delta Z\Delta X-\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z)I.\tag{4.5}$$

于是,根据(4.2c)式,(4.4)式,(4.5)式得到

$$Q(\alpha)+Q(\alpha)^{\mathrm{T}}=X^{-\frac{1}{2}}(X(\alpha)Z(\alpha)-\mu(\alpha)I)X^{\frac{1}{2}}+X^{\frac{1}{2}}(Z(\alpha)X(\alpha)-\mu(\alpha)I)X^{-\frac{1}{2}}=2(1-\alpha)(X^{\frac{1}{2}}ZX^{\frac{1}{2}}-\mu I)+2\alpha(X^{\frac{1}{2}}ZX^{\frac{1}{2}}-\sigma\mu I)+\alpha(X^{-\frac{1}{2}}(X\Delta Z+\Delta XZ)X^{\frac{1}{2}}+X^{\frac{1}{2}}(\Delta ZX+Z\Delta X)X^{-\frac{1}{2}})+\alpha^2(X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}+X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}})-X^{-\frac{1}{2}}(\frac{\alpha^2}{n}\cdot(\Delta X\cdot\Delta Z)I)X^{\frac{1}{2}}-X^{\frac{1}{2}}(\frac{\alpha^2}{n}(\Delta X\cdot\Delta Z)I)X^{-\frac{1}{2}}=2(1-\alpha)(X^{\frac{1}{2}}ZX^{\frac{1}{2}}-\mu I)+\alpha^2(X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}+X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}})-\frac{2\alpha^2}{n}(\Delta X\cdot\Delta Z)I,$$

最后一个等式成立是根据(2.5c)式.

接下来的引理给出 scaling 方向 $X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}$,$X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}$ 的界,对任意$(X,y,Z)\in\mathbb{F}^\circ_P\times\mathbb{F}^\circ_D$ 都成立. 证明过程需要使用如下结论,对任意的 $A,B\in\mathbb{R}^{n\times n}$,都有(详见文献[15]5.6 节习题 20).

$$\|A+B\|_F\leqslant\|A\|_F+\|B\|_F,\|AB\|_F\leqslant\|A\|_F\|B\|_F,\|AB\|_F\leqslant\|A\|_F\|B\|.\tag{4.6}$$

引理 4.2　假设(A1)和(A2)成立,设 $X\in\mathbb{F}^\circ_P$,$(X,y,Z)\in\mathbb{F}^\circ_D$,使得 $\|X^{\frac{1}{2}}ZX^{\frac{1}{2}}-\upsilon I\|\leqslant\upsilon\gamma$,对任意的 $\gamma\in[0,1)$,$\upsilon>0$ 成立. $(\Delta X,\Delta y,\Delta Z)$ 是线性方程组(2.7)对于某个确定矩阵 W 的解. 记 $\delta_x=\upsilon\|X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}\|_F$,$\delta_z=\upsilon\|X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}\|_F$,则下列结论成立:

$$\delta_x\delta_z\leqslant\frac{1}{2}(\delta_x^2+\delta_z^2)\leqslant\frac{\|W\|_F^2}{2(1-\gamma)^2}.\tag{4.7}$$

证明　根据(2.5c)式得

$$W=\frac{1}{2}(X^{-\frac{1}{2}}(X\Delta Z+\Delta XZ)X^{\frac{1}{2}}+X^{\frac{1}{2}}(\Delta ZX+Z\Delta X)X^{-\frac{1}{2}})=\frac{1}{2}X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}+\frac{1}{2}X^{-\frac{1}{2}}\Delta XZX^{\frac{1}{2}}+\frac{1}{2}X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}+\frac{1}{2}X^{\frac{1}{2}}Z\Delta XX^{-\frac{1}{2}}=X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}+\upsilon X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}-\frac{1}{2}\upsilon X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}+$$

$\frac{1}{2}X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}X^{\frac{1}{2}}ZX^{\frac{1}{2}} + \frac{1}{2}X^{\frac{1}{2}}ZX^{\frac{1}{2}}X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} - \frac{1}{2}\upsilon X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} = X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}} + \upsilon X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} + \frac{1}{2}X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}(X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \upsilon I) + \frac{1}{2}(X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \upsilon I)X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}$,

由引理2.4(1)知$\Delta Z \cdot \Delta X \geqslant 0$,再根据(4.6)式以及引理4.2的已知条件得到

$\| W \|_F \geqslant \| X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}} + \upsilon X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F - \frac{1}{2}\| X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}(X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \upsilon I) + (X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \upsilon I)X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F \geqslant \| X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}} + \upsilon X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F - \| X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \upsilon I \| \cdot \| X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F \geqslant (\| X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}} \|_F^2 + \upsilon^2 \| X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F^2)^{\frac{1}{2}} - \upsilon\gamma \cdot \frac{\delta_x}{\upsilon} = \sqrt{\delta_x^2 + \delta_z^2} - \gamma \cdot \delta_x \geqslant (1-\gamma)\sqrt{\delta_x^2 + \delta_z^2}$,

所以

$$\frac{\| W \|_F^2}{(1-\gamma)^2} \geqslant \delta_x^2 + \delta_z^2.$$

注意到$\delta_x^2 + \delta_z^2 \geqslant 2\delta_x\delta_z$,于是(4.7)式成立.

以下定理说明本文的短步路径跟踪算法产生的迭代点列将落在中心路径邻域$\mathbb{N}_F(\gamma)$内.

定理4.1 假设(A1)和(A2)成立,设常数$\gamma \in (0,1)$,$\delta \in [0,\sqrt{n})$,并且满足

$$\frac{\gamma^2 + \delta^2}{2(1-\gamma)^2(1-\delta/\sqrt{n})} \leqslant \gamma, \gamma \leqslant \frac{1}{2}. \tag{4.8}$$

又设$(X,y,Z) \in \mathbb{N}_F(\gamma)$,$(\Delta X,\Delta y,\Delta Z)$是线性方程组(2.7)当$W = \sigma\mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}}$时的解,$\sigma = 1 - \delta/\sqrt{n}$.则

$(\hat{X},\hat{y},\hat{Z}) = (X + \Delta X, y + \Delta y, Z + \Delta Z) \in \mathbb{N}_F(\gamma)$.

证明 因为

$(X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \mu I) \cdot I = \mathrm{Tr}((X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \mu I)I) = \mathrm{Tr}(X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \mu I) = \mathrm{Tr}(X^{\frac{1}{2}}ZX^{\frac{1}{2}}) - n\mu = \mathrm{Tr}(XZ) - n\mu = X \cdot Z - n\mu = 0$,

并结合$\sigma = 1 - \delta/\sqrt{n}$,和$\mathbb{N}_F(\gamma)$的定义,则有

$$\| W \|_F^2 = \| \sigma\mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}} \|_F^2 = \| (\sigma-1)\mu I + \mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}} \|_F^2 = \| (\sigma-1)\mu I \|_F^2 + \| \mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}} \|_F^2 \leqslant (1-\sigma)^2\mu^2 \| I \|_F^2 + \gamma^2\mu^2 = (\delta^2 + \gamma^2)\mu^2. \tag{4.9}$$

因为$\| X^{\frac{1}{2}}ZX^{\frac{1}{2}} - \mu I \|_F \leqslant \gamma\mu$,根据引理4.2取$\upsilon = \mu$,$W = \sigma\mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}}$,则有

$$\delta_x^2 \leqslant \delta_x^2 + \delta_z^2 \leqslant \frac{\| W \|_F^2}{(1-\gamma)^2},$$

$$\delta_x\delta_z \leqslant \frac{1}{2}(\delta_x^2 + \delta_z^2) \leqslant \frac{\| W \|_F^2}{2(1-\gamma)^2},$$

从而

$$\| X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F = \frac{\delta_x}{\upsilon} \leqslant \frac{\| W \|_F}{(1-\gamma)\upsilon} = \frac{\| \sigma\mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}} \|_F}{(1-\gamma)\mu}, \tag{4.10}$$

$$\| X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}} \|_F \| X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}} \|_F = \frac{\delta_x \cdot \delta_z}{\upsilon} \leqslant \frac{\| W \|_F^2}{2\upsilon(1-\gamma)^2} = \frac{\| \sigma\mu\ I - X^{\frac{1}{2}}ZX^{\frac{1}{2}} \|_F^2}{2\mu(1-\gamma)^2}. \tag{4.11}$$

根据(4.2b)式,(4.2c)式和(4.3)式得

$$\hat{Q} := Q(1) = X^{-\frac{1}{2}}(\hat{X}\hat{Z} - \hat{\mu}I)X^{\frac{1}{2}}, \hat{\mu} = \frac{\hat{X} \cdot \hat{Z}}{n},$$

$$Q(1) + Q(1)^{\mathrm{T}} = (X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} + X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}) - \frac{2}{n}(\Delta X \cdot \Delta Z)I.$$

再根据(4.6)式,有

$$\frac{1}{2}\| \hat{Q} + \hat{Q}^{\mathrm{T}} \|_F = \frac{1}{2}\| X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} - \frac{1}{n}(\Delta X \cdot \Delta Z)I + X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}} - \frac{1}{n}(\Delta X \cdot \Delta Z)I \|_F \leqslant \| X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} - \frac{1}{n}(\Delta X \cdot \Delta Z)I \|_F, \tag{4.12}$$

由矩阵Frobenious范数的定义知

$\| X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} - \frac{1}{n}(\Delta X \cdot \Delta Z)I \|_F^2 = \mathrm{Tr}(X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}) - \frac{2}{n}(\Delta X \cdot \Delta Z)\mathrm{Tr}(X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}) + \frac{1}{n^2}(\Delta X \cdot \Delta Z)^2\mathrm{Tr}(I) = \mathrm{Tr}(X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}) - \frac{2}{n}(\Delta X \cdot \Delta Z)(\Delta X \cdot \Delta Z) + \frac{1}{n^2}(\Delta X \cdot \Delta Z)^2 n = \mathrm{Tr}(X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}) - \frac{1}{n}(\Delta X \cdot \Delta Z)^2 \leqslant \mathrm{Tr}(X^{\frac{1}{2}}\Delta Z\Delta XX^{-\frac{1}{2}}X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}) = (X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}) \cdot (X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}) = \| X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} \|_F^2$,

即

$$\| X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} - \frac{1}{n}(\Delta X \cdot \Delta Z)I \|_F \leqslant \| X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}} \|_F.$$

将上式代入(4.12)式,再结合(4.11)式,(4.9)式和(4.8)式,得

$$\frac{1}{2}\|\hat{Q}+\hat{Q}^{\mathrm{T}}\|_F \leqslant \|X^{-\frac{1}{2}}\Delta X\Delta ZX^{\frac{1}{2}}\|_F = \|X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}\|_F \leqslant \|X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}\|_F\|X^{\frac{1}{2}}\Delta ZX^{\frac{1}{2}}\|_F \leqslant \frac{\|\sigma\mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}}\|_F^2}{2\mu(1-\gamma)^2} \leqslant \frac{(\delta^2+\gamma^2)\mu}{2(1-\gamma)^2} \leqslant \gamma(1-\delta/\sqrt{n})\mu \leqslant \gamma\hat{\mu}, \tag{4.13}$$

其中 $\hat{\mu}=\frac{\hat{X}\bullet\hat{Z}}{n}\geqslant(1-\delta/\sqrt{n})\frac{X\bullet Z}{n}=(1-\delta/\sqrt{n})\mu$.

根据(4.10)式,(4.9)式和定理4.1的条件,得

$$\|X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}\|_F \leqslant \frac{\|\sigma\mu I - X^{\frac{1}{2}}ZX^{\frac{1}{2}}\|_F}{(1-\gamma)\mu} \leqslant \frac{(\delta^2+\gamma^2)^{\frac{1}{2}}\mu}{(1-\gamma)\mu} \leqslant (2\gamma(1-\delta/\sqrt{n}))^{\frac{1}{2}} < 1,$$

从而知 $I+X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}}$ 是正定矩阵. 注意到 $X^{\frac{1}{2}}=(X^{\frac{1}{2}})^{\mathrm{T}}$,于是 $\hat{X}=X+\Delta X=X^{\frac{1}{2}}(I+X^{-\frac{1}{2}}\Delta XX^{-\frac{1}{2}})X^{\frac{1}{2}}$ 也是正定矩阵. 而

$$\hat{Q}=X^{-\frac{1}{2}}(\hat{X}\hat{Z}-\hat{\mu}I)X^{\frac{1}{2}}=(X^{-\frac{1}{2}}\hat{X}^{\frac{1}{2}})(\hat{X}^{\frac{1}{2}}\hat{Z}\hat{X}^{\frac{1}{2}}-\hat{\mu}I)(X^{-\frac{1}{2}}\hat{X}^{\frac{1}{2}})^{-1},$$

记 $D=X^{-\frac{1}{2}}\hat{X}^{\frac{1}{2}}$,$G=\hat{X}^{\frac{1}{2}}\hat{Z}\hat{X}^{\frac{1}{2}}-\hat{\mu}I$,则 $\hat{Q}=DGD^{-1}$. 由文献[5]的引理3.3以及(4.13)式可得

$$\|\hat{X}^{\frac{1}{2}}\hat{Z}\hat{X}^{\frac{1}{2}}-\hat{\mu}I\|_F \leqslant \frac{1}{2}\|\hat{Q}+\hat{Q}^{\mathrm{T}}\|_F \leqslant \gamma\hat{\mu}, \tag{4.14}$$

于是 $\lambda_{\min}(\hat{X}^{\frac{1}{2}}\hat{Z}\hat{X}^{\frac{1}{2}}) \geqslant (1-\gamma)\hat{\mu}>0$,所以 $\hat{X}^{\frac{1}{2}}\hat{Z}\hat{X}^{\frac{1}{2}}$ 是正定矩阵. 注意到 $\hat{X}^{\frac{1}{2}}=(\hat{X}^{\frac{1}{2}})^{\mathrm{T}}$,从而知 $\hat{Z}$ 也是正定矩阵. 根据(2.3a)式,(2.3b)式以及 $(X,y,Z)\in\mathbb{N}_F(\gamma)$ 可得到 $(\hat{X},\hat{y},\hat{Z})\in\mathbb{F}^{\circ}{}_D$ 且 $\hat{X}\in\mathbb{F}^{\circ}{}_P$,再结合(4.14)式,即知 $(\hat{X},\hat{y},\hat{Z})\in\mathbb{N}_F(\gamma)$.

5 数值试验

采用Matlab(R2011b)数学软件编程对算法A进行数值试验,程序运行环境为Windows7(64bite),Intel(R) Core(TM) i3-2330M CPU @ 2.20GHZ,RAM:4G.

考虑如下两类测试问题:(1)随机二次半定规划问题;矩阵 A_i,H_j,C 都是随机生成的对阵矩阵,初始点 (X,y,Z) 的选取需在 $\mathbb{F}_P$,$\mathbb{F}_D$ 内. 记该类测试问题为RP.(2)最近相关矩阵 Nearest correlation matrix(简记为NCM)问题[16-17].

$$\min \frac{1}{2}\|X-G\|^2$$

$$\text{s.t.}\quad \mathrm{diag}(X)=e,$$

$$X\in S_+^n,$$

其中 G 是对称正定阵,对角线上的元素为1,对角线外的元素在 -1 到1之间. e 是所有元素都为1的 n 维向量.

NCM问题具有广泛的应用,例如在市场营销以及经济学方面. 但是遗憾的是,由于缺乏数据等信息,不能得到一个完整的矩阵 G,即矩阵 G 的一些元素是未知的. 于是通过求解NCM问题得到一个有效的、且与 G 最近的相关矩阵. 在数值测试中,取 G 是随机矩阵且 $G:=(1-\alpha)B+\alpha E$[18],其中 $\alpha\in(0,1)$,E 是元素属于 $[-1,1]$ 的随机对称阵,B 是具有指定特征值的测试矩阵,在NCM问题的数值测试中取 $n=10$,20,30.

利用对称Kronecker积的矩阵形式

$$A\otimes_s B=\frac{1}{2}Q(A\otimes B+B\otimes A)Q^{\mathrm{T}}.$$

将矩阵的对称Kronecker积转化为矩阵的Kronecker积计算,其中 Q 的取法详见文献[1]附录.

在数值测试中,取算法A中的参数 $\delta=0.3$,$\gamma=0.3$,终止准则是 $Z^k\bullet X^k\leqslant 10^{-6}$.

测试的数值结果见表1(所有结果都是每个维数测试10次,然后取平均值得到的),其中表中符号含义如下:prob,测试问题的编号;n,矩阵 X 的维数;m,不等式约束的个数;Itr,算法迭代次数;Nf,目标函数值的计算次数;time(s),算法迭代所需CPU时间(s).

表1 数值结果

Table 1 Numerical results

prob	n	m	Itr	Nf	time(s)
RP	5	0	126	126	1.009955e+00
	6	0	146	146	2.871151e+00
NCM	10	10	163	163	1.842785e+01
	20	20	244	244	1.767764e+02
	30	30	307	307	1.529803e+03

上述数值结果表明本文提出的基于H..K..M方向的路径跟踪算法是有效的,能在合理的时间内求解中等规模的NCM问题. 为进一步提高本文算法的数值效果,我们将在后续的研究中深入探讨步长 α 的更有效的确定方法,即线搜索技术.

参考文献:

[1] TODD M J,TOH K C,TÜTÜNCÜ R H. On the nesterov-todd direction in SDP[J]. SIAM Journal on Optimization,1998,8(3):769-796.

[2] ALIZADEH F,HAEBERLY J P,OVERTON M L. Primal-dual interior-point methods for semidefinite pro-

gramming: Convergence rates, stability and numerical results[J]. SIAM Journal on Optimization, 1998, 8: 746-768.

[3] KOJIMA M, SHINDOH S, HARA S. Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices[J]. SIAM Journal on Optimization, 1997, 7(1): 86-125.

[4] HELMBERG C, RENDL F, VANDERBEI R J, et al. An interior-point methods for semidefinite programming [J]. SIAM Journal on Optimization, 1996, 6(2): 342-361.

[5] MONTEIRO R D C. Primal-dual path-following algorithms for semidefinite programming[J]. SIAM Journal on Optimization, 1997, 7(3): 663-678.

[6] NESTEROV Y E, TODD M J. Primal-dual interiorpoint methods for self-scaled cones[J]. SIAM Journal on Optimization, 1998, 8(2): 324-364.

[7] NESTEROV Y E, TODD M J. Self-scaled barriers and interior-point methods for convex programming[J]. Mathematics of Operations Research, 1997, 22(1): 1-42.

[8] STURM J F, ZHANG S Z. Symmetric primal-dual path-following algorithms for semidefinite programming[J]. Applied Numerical Mathematics, 1999, 29(3): 301-315.

[9] ZHANG Y. On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming[J]. SIAM Journal on Optimization, 1998, 8(2): 365-386.

[10] NESTEROV Y, NEMIROVSKII A S. Interior-Point Polynomial Methods in Convex Programming[M]. Philadelphia PA: SIAM, 1994.

[11] NIE J W, YUAN Y X. A potential reduction algorithm for an extended SDP problem[J]. Science in China Series A Mathematics, 2000, 43(1): 35-46.

[12] NIE J W, YUAN Y X. A Predictor-corrector algorithm for QSDP combining dikin-type and newton centering steps[J]. Annals of Operations Research, 2001, 103 (1/2/3/4): 115-133.

[13] 关秀翠，刁在筠. 二次半定规划问题及其投影收缩算法[J]. 全国高等学校计算数学学报，2002，24(2): 97-108.
GUAN X C, SIAO Z Y. The quadratic semi-definite programming problem and its projection and contraction algorithm[J]. Numerical Mathematics a Journal of Chinese Universities, 2002, 24(2): 97-108.

[14] 徐凤敏，徐成贤. 求解二次半定规划的原对偶内点算法[J]. 工程数学学报，2006，23(4): 590-598.
XU F M, XU C X. Primal-dual algorithm for quadratic semi-definite programming[J]. Chinese Journal of Engineering Mathematics, 2006, 23(4): 590-598.

[15] HORN R A, JOHNSON C R. Matrix Analysis[M]. New York: Cambridge University Press, 1985.

[16] HIGHAM N J. Computing the nearest correlation matrixa problem from finance[J]. IMA Journal of Numerical Analysis, 2002, 22(3): 329-343.

[17] TOH K C. An inexact primal-dual path following algorithm for convex quadratic SDP[J]. Mathematical Programming, 2008, 112(1): 221-254.

[18] ZHAO X Y. A Semismooth Newton-CG Augmented Lagrangian Method for Large Scale Linear and Convex Quadratic SDPs[D]. Singapore: National University of Singapore, 2009.

（责任编辑：米慧芝）