文章编号: 1004-4353(2019)02-0109-05

非线性矩阵方程中心对称解的牛顿-MCG 算法

陈世军
(福建工程学院 应用技术学院, 福建 福州 350001)

摘要: 研究了一类含有高次逆幂非线性矩阵方程中心对称解的数值计算问题. 首先用牛顿算法求等价的线性矩阵方程的中心对称解,然后用修正共轭梯度算法(MCG 算法)求线性矩阵方程的中心对称解或中心对称最小二乘解. 数值算例表明,本文算法有效.

关键词: 含高次逆幂的矩阵方程; 中心对称解; 修正共轭梯度法
中图分类号: O241.6　　**文献标识码:** A

A Newton-MCG algorithm for centrosymmetric solutions of nonlinear matrix equations

CHEN Shijun
(*College of Applied Technology*, *Fujian University of Technology*, *Fuzhou* 350001, *China*)

Abstract: In this paper, we study the numerical computation of central symmetric solution for a class of nonlinear matrix equation with higher order inverse-power. The modified conjugate gradient algorithm (MCG) is used to solve the central symmetric solution or the central symmetric least squares solution of the linear matrix equation derived by Newton's algorithm in each iteration step. The numerical example shows that the algorithm is effective.

Keywords: matrix equation with high order inverse-power; central symmetric solution; modified conjugate gradient method

0 引言

在控制理论、随机滤波和超分辨率图像恢复等领域,会遇到形如

$$\boldsymbol{X}+\boldsymbol{E}_1\boldsymbol{X}^{-1}\boldsymbol{F}_1+\boldsymbol{E}_2\boldsymbol{X}^{-2}\boldsymbol{F}_2+\boldsymbol{E}_3\boldsymbol{X}^{-3}\boldsymbol{F}_3=\boldsymbol{G} \tag{1}$$

的含有未知矩阵高次逆幂的非线性矩阵方程[1-4],其中 $\boldsymbol{E}_i,\boldsymbol{F}_i,\boldsymbol{G},\boldsymbol{X}\in\mathbf{R}^{n\times n}$ $(i=1,2,3)$. 文献[5-8]研究了牛顿算法在求解线性方程和非线性方程中的应用. 程可欣等[8] 研究了方程(1) 的特例 $\boldsymbol{X}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}^{-1}\boldsymbol{A}=\boldsymbol{Q}$ 的牛顿迭代解法,证明了牛顿迭代方法产生的矩阵序列包含在确定的闭球内,且该矩阵序列收敛到闭球内唯一的矩阵方程的解. 张肖肖等[9] 给出了方程(1) 对称解的双迭代算法,其先用牛顿算法求出等价的线性矩阵方程的对称解,然后再采用修正共轭梯度算法(MCG 算法) 求出由牛顿算法每一步迭代计算导出的线性方程的对称解或对称最小二乘解. 目前,对于方程(1) 的中心对称解的研究还未见报道. 为

收稿日期: 2019 - 03 - 26　　**作者简介:** 陈世军(1983—),男,讲师,研究方向为计算数学.
基金项目: 福建工程学院应用技术学院科研项目(YYJS-JB1808);福建省教育厅中青年教育科研项目(JZ180190)

此,本文基于文献[9]的算法原理,将牛顿-MCG 算法推广到求解方程(1)的中心对称解上.

1 求方程(1)中心对称解的牛顿算法

定义 1 划分 n 阶单位矩阵 $\boldsymbol{I}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$,称 $\boldsymbol{S}_n=(\boldsymbol{e}_n,\boldsymbol{e}_{n-1},\cdots,\boldsymbol{e}_1)$ 为 n 阶次单位矩阵. 若 $\boldsymbol{X}\in\mathbf{R}^{n\times n}$ 满足 $\boldsymbol{S}_n\boldsymbol{X}\boldsymbol{S}_n=\boldsymbol{X}$,称 $\boldsymbol{X}$ 为中心对称矩阵,用 $\mathbf{CSR}^{n\times n}$ 表示中心对称矩阵集合.

基于文献[9]的算法原理,建立求非线性矩阵方程(1)中心对称解的牛顿-MCG算法. 令 $\boldsymbol{X}=\boldsymbol{X}^{-1}$,则方程(1)可以改写为

$$\boldsymbol{X}^{-1}+\boldsymbol{E}_1\boldsymbol{X}\boldsymbol{F}_1+\boldsymbol{E}_2\boldsymbol{X}^2\boldsymbol{F}_2+\boldsymbol{E}_3\boldsymbol{X}^3\boldsymbol{F}_3=\boldsymbol{G}. \tag{2}$$

首先用牛顿算法求方程(2)的中心对称解,然后采用 MCG 算法求由牛顿算法每一步迭代计算导出的线性矩阵方程的中心对称解或中心对称最小二乘解,以此建立求方程(2)中心对称解的牛顿-MCG 算法.

定义 2[10] 方阵 $\boldsymbol{A}$ 的幂级数 $\sum_{k=0}^{\infty}\boldsymbol{A}^k=\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 收敛的充要条件是 $\boldsymbol{A}$ 为收敛矩阵,并且在收敛时,其和为 $(\boldsymbol{I}-\boldsymbol{A})^{-1}$.

根据定义 2 可知,当 $-\boldsymbol{X}^{-1}\boldsymbol{Y}$ 的谱半径小于 1 时,有

$$(\boldsymbol{X}+\boldsymbol{Y})^{-1}=[\boldsymbol{X}(\boldsymbol{I}-(-\boldsymbol{X}^{-1}\boldsymbol{Y}))]^{-1}=[\boldsymbol{I}-(-\boldsymbol{X}^{-1}\boldsymbol{Y})]^{-1}\boldsymbol{X}^{-1}=\sum_{k=0}^{\infty}(-\boldsymbol{X}^{-1}\boldsymbol{Y})^k\boldsymbol{X}^{-1}.$$

引入下列矩阵函数:

$$\psi(\boldsymbol{X})\overset{\text{def}}{=}\boldsymbol{X}^{-1}+\boldsymbol{E}_1\boldsymbol{X}\boldsymbol{F}_1+\boldsymbol{E}_2\boldsymbol{X}^2\boldsymbol{F}_2+\boldsymbol{E}_3\boldsymbol{X}^3\boldsymbol{F}_3-\boldsymbol{G}.$$

因 $\frac{\mathrm{d}(\psi(\boldsymbol{X}+t\boldsymbol{Y}))}{\mathrm{d}t}\Big|_{t=0}=\boldsymbol{E}_1\boldsymbol{Y}\boldsymbol{F}_1+\boldsymbol{E}_2(\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X})\boldsymbol{F}_2+\boldsymbol{E}_3(\boldsymbol{X}\boldsymbol{Y}\boldsymbol{X}+\boldsymbol{X}^2\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X}^2)\boldsymbol{F}_3-\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}$,所以记

$$\phi_{\boldsymbol{X}}(\boldsymbol{Y})\overset{\text{def}}{=}\boldsymbol{E}_1\boldsymbol{Y}\boldsymbol{F}_1+\boldsymbol{E}_2(\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X})\boldsymbol{F}_2+\boldsymbol{E}_3(\boldsymbol{X}\boldsymbol{Y}\boldsymbol{X}+\boldsymbol{X}^2\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X}^2)\boldsymbol{F}_3-\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1},$$

$$\gamma_{\boldsymbol{X}}(\boldsymbol{Y})=\sum_{k=1}^{+\infty}(-\boldsymbol{X}^{-1}\boldsymbol{Y})^k\boldsymbol{X}^{-1}+\boldsymbol{E}_2\boldsymbol{Y}^2\boldsymbol{F}_2+\boldsymbol{E}_3(\boldsymbol{X}\boldsymbol{Y}^2+\boldsymbol{Y}^2\boldsymbol{X}+\boldsymbol{Y}\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}^3)\boldsymbol{F}_3.$$

又因为

$$\begin{aligned}\psi(\boldsymbol{X}+\boldsymbol{Y})&=(\boldsymbol{X}+\boldsymbol{Y})^{-1}+\boldsymbol{E}_1(\boldsymbol{X}+\boldsymbol{Y})\boldsymbol{F}_1+\boldsymbol{E}_2(\boldsymbol{X}+\boldsymbol{Y})^2\boldsymbol{F}_2+\boldsymbol{E}_3(\boldsymbol{X}+\boldsymbol{Y})^3\boldsymbol{F}_3-\boldsymbol{G}=\\&\sum_{k=1}^{+\infty}(-\boldsymbol{X}^{-1}\boldsymbol{Y})^k\boldsymbol{X}^{-1}+\boldsymbol{X}^{-1}+\boldsymbol{E}_1\boldsymbol{X}\boldsymbol{F}_1+\boldsymbol{E}_1\boldsymbol{Y}\boldsymbol{F}_1+\boldsymbol{E}_2\boldsymbol{X}^2\boldsymbol{F}_2+\boldsymbol{E}_2\boldsymbol{Y}^2\boldsymbol{F}_2+\boldsymbol{E}_2\boldsymbol{X}\boldsymbol{Y}\boldsymbol{F}_2+\\&\boldsymbol{E}_2\boldsymbol{Y}\boldsymbol{X}\boldsymbol{F}_2+\boldsymbol{E}_3\boldsymbol{X}^3\boldsymbol{F}_3+\boldsymbol{E}_3\boldsymbol{X}^2\boldsymbol{Y}\boldsymbol{F}_3+\boldsymbol{E}_3\boldsymbol{X}\boldsymbol{Y}\boldsymbol{X}\boldsymbol{F}_3+\boldsymbol{E}_3\boldsymbol{X}\boldsymbol{Y}^2\boldsymbol{F}_3+\boldsymbol{E}_3\boldsymbol{Y}\boldsymbol{X}^2\boldsymbol{F}_3+\boldsymbol{E}_3\boldsymbol{Y}\boldsymbol{X}\boldsymbol{Y}\boldsymbol{F}_3+\\&\boldsymbol{E}_3\boldsymbol{Y}^2\boldsymbol{X}\boldsymbol{F}_3+\boldsymbol{E}_3\boldsymbol{Y}^3\boldsymbol{F}_3-\boldsymbol{G},\end{aligned}$$

容易导出 $\psi(\boldsymbol{X}+\boldsymbol{Y})=\psi(\boldsymbol{X})+\phi_{\boldsymbol{X}}(\boldsymbol{Y})+\gamma_{\boldsymbol{X}}(\boldsymbol{Y})$. 这里 $\phi_{\boldsymbol{X}}(\boldsymbol{Y})$ 是 $\psi(\boldsymbol{X})$ 在"点 $\boldsymbol{X}$"沿着"方向 $\boldsymbol{Y}$"的 Frechet 导数.

引理 1[9] 设 $\boldsymbol{X}\in\mathbf{CSR}^{n\times n}$ 是方程(1)的近似解,则求方程(1)的解 $\boldsymbol{X}$ 等价于求校正值 $\boldsymbol{Y}\in\mathbf{CSR}^{n\times n}$ 使得 $\psi(\boldsymbol{X}+\boldsymbol{Y})=\boldsymbol{O}$,并可以线性化为求 $\boldsymbol{Y}\in\mathbf{CSR}^{n\times n}$,使得

$$\phi_{\boldsymbol{X}}(\boldsymbol{Y})=-\psi(\boldsymbol{X}). \tag{3}$$

参考文献[9]的算法原理,通过修改某些矩阵的类型,建立求方程(1)中心对称解的牛顿算法,该算法的具体步骤如下:

第 1 步 给定初始矩阵 $\boldsymbol{X}^{(1)}\in\mathbf{CSR}^{n\times n}$,置 $k:=1$.

第 2 步 如果 $\psi(\boldsymbol{X}^{(k)})=\boldsymbol{O}$,停止;否则,求 $\boldsymbol{Y}^{(k)}\in\mathbf{CSR}^{n\times n}$,使得 $\phi_{\boldsymbol{X}^{(k)}}(\boldsymbol{Y}^{(k)})=-\psi(\boldsymbol{X}^{(k)})$.

第 3 步 计算 $\boldsymbol{X}^{(k+1)}=\boldsymbol{X}^{(k)}+\boldsymbol{Y}^{(k)}$,置 $k:=k+1$,转第 2 步.

对于牛顿算法有如下的收敛性结论:假设 $\boldsymbol{X}^*$ 是方程(1)的单根,且初始矩阵 $\boldsymbol{X}^{(1)}$ 充分接近 $\boldsymbol{X}^*$,那么牛顿算法确定的矩阵序列 $\{\boldsymbol{X}^{(k)}\}$ 二次收敛于 $\boldsymbol{X}^*$.

2　求线性矩阵方程(3) 中心对称解的 MCG 算法

记 $\boldsymbol{A}_1=\boldsymbol{E}_1$, $\boldsymbol{B}_1=\boldsymbol{F}_1$, $\boldsymbol{A}_2=\boldsymbol{E}_2\boldsymbol{X}$, $\boldsymbol{B}_2=\boldsymbol{F}_2$, $\boldsymbol{A}_3=\boldsymbol{E}_2$, $\boldsymbol{B}_3=\boldsymbol{X}\boldsymbol{F}_2$, $\boldsymbol{A}_4=\boldsymbol{E}_3\boldsymbol{X}$, $\boldsymbol{B}_4=\boldsymbol{X}\boldsymbol{F}_3$, $\boldsymbol{A}_5=\boldsymbol{E}_3\boldsymbol{X}^2$, $\boldsymbol{B}_5=\boldsymbol{F}_3$, $\boldsymbol{A}_6=\boldsymbol{E}_3$, $\boldsymbol{B}_6=\boldsymbol{X}^2\boldsymbol{F}_3$, $\boldsymbol{A}_7=-\boldsymbol{X}^{-1}$, $\boldsymbol{B}_7=\boldsymbol{X}^{-1}$, $\boldsymbol{F}=-(\boldsymbol{X}^{-1}+\boldsymbol{E}_1\boldsymbol{X}\boldsymbol{F}_1+\boldsymbol{E}_2\boldsymbol{X}^2\boldsymbol{F}_2+\boldsymbol{E}_3\boldsymbol{X}^3\boldsymbol{F}_3-\boldsymbol{G})$. 考虑方程(3) 的一般形式：

$$\sum_{i=1}^{7}\boldsymbol{A}_i\boldsymbol{Y}\boldsymbol{B}_i=\boldsymbol{F},\tag{4}$$

其中 $\boldsymbol{A}_i,\boldsymbol{B}_i,\boldsymbol{C}_i,\boldsymbol{D}_i,\boldsymbol{F}\in\mathbf{R}^{n\times n}(i=1,2,\cdots,7)$. 本文解决下面两个问题：

问题 Ⅰ　设方程(4) 有中心对称解，求 $\boldsymbol{Y}\in\mathbf{CSR}^{n\times n}$，使 $\boldsymbol{Y}$ 满足方程(4).

问题 Ⅱ　设方程(4) 无中心对称解，求 $\boldsymbol{Y}\in\mathbf{CSR}^{n\times n}$，使

$$\left\|\sum_{i=1}^{7}\boldsymbol{A}_i\boldsymbol{Y}\boldsymbol{B}_i-\boldsymbol{F}\right\|=\min.\tag{5}$$

当方程(4) 有中心对称解时，称问题 Ⅰ 相容；否则，称问题 Ⅰ 不相容.

基于 MCG 算法原理，通过修改有关矩阵的类型，建立求解问题 Ⅰ 的 MCG 算法. 引入记号：$\mu(\boldsymbol{Y})\overset{\text{def}}{=}\sum_{i=1}^{7}\boldsymbol{A}_i\boldsymbol{Y}\boldsymbol{B}_i$, $w(\boldsymbol{R})\overset{\text{def}}{=}\sum_{i=1}^{7}\boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{R}\boldsymbol{B}_i^{\mathrm{T}}$. 求解问题 Ⅰ 的 MCG 算法(算法 1) 的具体步骤如下：

第 1 步　给定初始矩阵 $\boldsymbol{Y}^{(k)}\in\mathbf{CSR}^{n\times n}$，置 $k:=1$，计算 $\boldsymbol{R}_k=\boldsymbol{F}-\mu(\boldsymbol{Y}^{(k)})$, $\widetilde{\boldsymbol{R}}_k=w(\boldsymbol{R}_k)$, $\boldsymbol{Q}_k=\frac{1}{2}(\widetilde{\boldsymbol{R}}_k+\boldsymbol{S}_n\widetilde{\boldsymbol{R}}_k\boldsymbol{S}_n)$.

第 2 步　若 $\boldsymbol{R}_k=\boldsymbol{O}$，或 $\boldsymbol{R}_k\neq\boldsymbol{O}$ 而 $\boldsymbol{Q}_k=\boldsymbol{O}$，停止计算；否则，计算 $\alpha_k=\frac{\|\boldsymbol{R}_k\|^2}{\|\boldsymbol{Q}_k\|^2}$, $\boldsymbol{Y}^{(k+1)}=\boldsymbol{Y}^{(k)}+\alpha_k\boldsymbol{Q}_k$.

第 3 步　计算 $\boldsymbol{R}_{k+1}=\boldsymbol{F}-\mu(\boldsymbol{Y}^{(k+1)})$, $\widetilde{\boldsymbol{R}}_{k+1}=w(\boldsymbol{R}_{k+1})$, $\beta_{k+1}=\frac{\|\boldsymbol{R}_{k+1}\|^2}{\|\boldsymbol{R}_k\|^2}$, $\boldsymbol{Q}_{k+1}=\frac{1}{2}(\widetilde{\boldsymbol{R}}_{k+1}+\boldsymbol{S}_n\widetilde{\boldsymbol{R}}_{k+1}\boldsymbol{S}_n)+\beta_{k+1}\boldsymbol{Q}_k$.

第 4 步　置 $k:=k+1$，转第 2 步.

由以上步骤易见，算法 1 中的矩阵 $\boldsymbol{Y}^{(k)},\boldsymbol{Q}_k\in\mathbf{CSR}^{n\times n}$. 下面给出算法 1 的基本性质，算法 1 在有限步计算之后停止，具体证明可参考文献[11].

引理 2　对任意的 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$, $\boldsymbol{Y}\in\mathbf{CSR}^{n\times n}$ 均有 $\mathrm{tr}((\boldsymbol{A}+\boldsymbol{S}_n\boldsymbol{A}\boldsymbol{S}_n)^{\mathrm{T}}\boldsymbol{Y})=2\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y})$.

性质 1　对于算法 1 中的矩阵 $\boldsymbol{R}_i,\boldsymbol{Q}_i$ 和 $\widetilde{\boldsymbol{R}}_i$，有 $\mathrm{tr}(\boldsymbol{R}_{i+1}^{\mathrm{T}}\boldsymbol{R}_j)=\mathrm{tr}(\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_j)-\frac{\|\boldsymbol{R}_i\|^2}{\|\boldsymbol{Q}_i\|^2}\mathrm{tr}(\boldsymbol{Q}_i^{\mathrm{T}}\widetilde{\boldsymbol{R}}_j)$.

性质 2　设 $k\geqslant 2$，对于算法 1 中的矩阵 $\boldsymbol{R}_i$ 和 $\boldsymbol{Q}_j$，有 $\mathrm{tr}(\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_j)=0$, $\mathrm{tr}(\boldsymbol{Q}_i^{\mathrm{T}}\boldsymbol{Q}_j)=0\ (i\neq j;i,j=1,2,\cdots,k)$.

性质 3　设 $\overline{\boldsymbol{Y}}\in\mathbf{CSR}^{n\times n}$ 是问题 Ⅰ 的任意一组中心对称解，对任意初始矩阵 $\boldsymbol{Y}^{(1)}\in\mathbf{CSR}^{n\times n}$，由算法 1 得到的矩阵 $\boldsymbol{R}_k,\boldsymbol{Y}^{(k)}$ 和 $\boldsymbol{Q}_k$ 满足 $\mathrm{tr}(\boldsymbol{Q}_k^{\mathrm{T}}(\overline{\boldsymbol{Y}}-\boldsymbol{Y}^{(k)}))=\|\boldsymbol{R}_k\|^2(k=1,2,\cdots)$.

定理 1　设问题 Ⅰ 相容，则对任意的初始矩阵 $\boldsymbol{Y}^{(1)}\in\mathbf{CSR}^{n\times n}$，算法 1 可在有限步计算后求得问题 Ⅰ 的一组解. 问题 Ⅰ 不相容的充要条件是存在正整数 k，使得由算法 1 得到的 $\boldsymbol{R}_k\neq\boldsymbol{O}$ 而 $\boldsymbol{Q}_k=\boldsymbol{O}$.

定理 2　设问题 Ⅰ 相容，选取初始矩阵 $\boldsymbol{Y}^{(1)}\in\mathbf{CSR}^{n\times n}$ 满足

$$\boldsymbol{Y}^{(1)}=\frac{1}{2}(\mu(\boldsymbol{H})+\boldsymbol{S}_n\mu(\boldsymbol{H})\boldsymbol{S}_n)\ \ (\forall\,\boldsymbol{H}\in\mathbf{R}^{n\times n}),$$

则算法 1 可在有限步计算后得到问题 Ⅰ 的唯一极小范数解，即方程(4) 的唯一极小范数解.

3 求解问题 Ⅱ 的 MCG 算法

根据定理 1,在算法 1 中当 $\boldsymbol{R}_k \neq \boldsymbol{O}$ 而 $\boldsymbol{Q}_k = \boldsymbol{O}$ 时,算法 1 中断. 这表明线性方程(4) 没有中心对称解,因此需要求解问题 Ⅱ,即求线性方程(4) 的中心对称最小二乘解. 下面通过构造等价的线性矩阵方程组,将求矩阵方程(4) 中心对称最小二乘解问题转化为求等价的正规方程中心对称解问题;然后参照算法 1,建立求矩阵方程(4) 中心对称最小二乘解的迭代算法.

记 $\boldsymbol{y} = \overline{\mathrm{vec}}(\boldsymbol{Y})$, $\boldsymbol{f} = \overline{\mathrm{vec}}(\boldsymbol{F})$, $\boldsymbol{T}_{m,n}$ 表示满足 $\overline{\mathrm{vec}}(\boldsymbol{A}^{\mathrm{T}}) = \boldsymbol{T}_{m,n}\,\overline{\mathrm{vec}}(\boldsymbol{A}_{m\times n})$ 的列交换矩阵.

$$f(\boldsymbol{Y}) = w(\mu(\boldsymbol{Y})) + \boldsymbol{S}_n w(\mu(\boldsymbol{S}_n\boldsymbol{Y}\boldsymbol{S}_n))\boldsymbol{S}_n,\ \boldsymbol{N} = w(\boldsymbol{F}) + \boldsymbol{S}_n w(\boldsymbol{F})\boldsymbol{S}_n,\ \boldsymbol{M} = \begin{pmatrix} \sum\limits_{i=1}^{7} \boldsymbol{A}_i \otimes \boldsymbol{B}_i^{\mathrm{T}} \\ \sum\limits_{i=1}^{7} \boldsymbol{A}_i\boldsymbol{S}_n \otimes \boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{S}_n \end{pmatrix}.$$

定理 3 求解问题 Ⅱ 等价于求线性矩阵方程

$$f(\boldsymbol{Y}) = \boldsymbol{N} \tag{6}$$

的中心对称解,并且方程(6) 一定有中心对称解.

证明 求解问题 Ⅱ 等价于求 $\boldsymbol{Y} \in \mathbf{CSR}^{n\times n}$,使得

$$\left\| \begin{pmatrix} \sum\limits_{i=1}^{7} \boldsymbol{A}_i\boldsymbol{Y}\boldsymbol{B}_i \\ \sum\limits_{i=1}^{7} \boldsymbol{A}_i\boldsymbol{S}_n\boldsymbol{Y}\boldsymbol{S}_n\boldsymbol{B}_i \end{pmatrix} - \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{F} \end{pmatrix} \right\| = \min. \tag{7}$$

下面证明求解极小值问题(7) 等价于求方程(6) 的中心对称解. 约定:矩阵的乘积运算优于矩阵的 Kronecker 积运算. 将矩阵方程组 $\begin{cases} \sum\limits_{i=1}^{7} \boldsymbol{A}_i\boldsymbol{Y}\boldsymbol{B}_i = \boldsymbol{F} \\ \sum\limits_{i=1}^{7} \boldsymbol{A}_i\boldsymbol{S}_n\boldsymbol{Y}\boldsymbol{S}_n\boldsymbol{B}_i = \boldsymbol{F} \end{cases}$ 按行拉直可得线性方程 $\boldsymbol{M}\boldsymbol{y} = \boldsymbol{f}$. 该线性方程的正规方程组为 $\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{f}$,将其还原为矩阵形式即为矩阵方程(6). 因此求方程(6) 的最小二乘解,就是求它的正规方程组 $\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{f}$ 的解. 因为正规方程组 $\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{f}$ 一定有解,所以方程(6) 也一定有解. 设 $\widetilde{\boldsymbol{Y}}$ 是它的一组解(未必是中心对称解),则 $f(\widetilde{\boldsymbol{Y}}) = \boldsymbol{N}$. 令 $\widetilde{\boldsymbol{Y}}^{*} = \frac{1}{2}(\widetilde{\boldsymbol{Y}} + \boldsymbol{S}_n\widetilde{\boldsymbol{Y}}\boldsymbol{S}_n)$,显然 $\widetilde{\boldsymbol{Y}}^{*} \in \mathbf{CSR}^{n\times n}$,且有 $f(\widetilde{\boldsymbol{Y}}^{*}) = \boldsymbol{N}$,故方程(6) 一定有中心对称解.

参照算法 1 以及 MCG 算法的基本原理,建立求矩阵方程(6) 中心对称解的算法,即求解问题 Ⅱ 的 MCG 算法. 求解问题 Ⅱ 的 MCG 算法(算法 2) 的具体步骤如下:

第 1 步 给定初始矩阵 $\boldsymbol{Y}^{(1)} \in \mathbf{CSR}^{n\times n}$,置 $k := 1$,计算 $\boldsymbol{R}_k = \boldsymbol{N} - f(\boldsymbol{Y}^{(k)})$, $\widetilde{\boldsymbol{R}}_k = f(\boldsymbol{R}_k)$, $\boldsymbol{Q}_k = \widetilde{\boldsymbol{R}}_k$.

第 2 步 如果 $\boldsymbol{R}_k = \boldsymbol{O}$,停止计算;否则,计算 $\alpha_k = \dfrac{\|\boldsymbol{R}_k\|^2}{\|\boldsymbol{Q}_k\|^2}$, $\boldsymbol{Y}^{(k+1)} = \boldsymbol{Y}^{(k)} + \alpha_k\boldsymbol{Q}_k$.

第 3 步 计算 $\boldsymbol{R}_{k+1} = \boldsymbol{N} - f(\boldsymbol{Y}^{(k+1)})$, $\widetilde{\boldsymbol{R}}_{k+1} = f(\boldsymbol{R}_{k+1})$, $\beta_{k+1} = \dfrac{\|\boldsymbol{R}_{k+1}\|^2}{\|\boldsymbol{R}_k\|^2}$, $\boldsymbol{Q}_{k+1} = \widetilde{\boldsymbol{R}}_{k+1} + \beta_{k+1}\boldsymbol{Q}_k$.

第 4 步 置 $k := k + 1$,转第 2 步.

由以上步骤易见,算法 2 中的矩阵 $\boldsymbol{Y}^{(k)}$, $\boldsymbol{Q}_k \in \mathbf{CSR}^{n\times n}$. 对于算法 2 有类似于算法 1 的收敛定理.

定理 4 对任意的初始矩阵 $\boldsymbol{Y}^{(1)} \in \mathbf{CSR}^{n\times n}$,算法 2 可在有限步计算后求得问题 Ⅱ 的一组解,即矩阵方程(6) 的中心对称最小二乘解. 若取初始矩阵 $\boldsymbol{Y}^{(1)}$ 满足 $\boldsymbol{Y}^{(1)} = f(\boldsymbol{H})$ ($\forall\, \boldsymbol{H} \in \mathbf{CSR}^{n\times n}$),则算法 2 可在有

限步计算后得到矩阵方程(6) 的唯一极小范数中心对称解.

4　数值算例

用上述建立的算法 1 和算法 2 求矩阵方程(1) 的一组中心对称解和中心对称最小二乘解. 取方程(1) 的系数矩阵均为块对角矩阵，且均为 $5n$ 阶方阵. 系数矩阵、初始矩阵和终止准则如下：

$\boldsymbol{A}$ 为三对角矩阵，其中 $a_{i,i}=3$，$a_{i,i-1}=2$，$a_{i-1,i}=2\,(i=1,2,\cdots,5n)$，其余元素均为 0；

$\boldsymbol{E}_1=-\boldsymbol{A}^{\mathrm{T}}$，$\boldsymbol{F}_1=\boldsymbol{A}$，$\boldsymbol{E}_2=\boldsymbol{O}$，$\boldsymbol{F}_2=\boldsymbol{O}$，$\boldsymbol{E}_3=\boldsymbol{O}$，

$\boldsymbol{F}_3=\boldsymbol{O}$，$\boldsymbol{G}=\boldsymbol{I}$，$\boldsymbol{X}_1=\boldsymbol{I}$，$\varepsilon=10^{-9}$(终止准则).

选取初始矩阵 $\boldsymbol{Y}_1=2\boldsymbol{I}$，按照算法 1 和算法 2 可求得矩阵方程(1) 的一组中心对称解，结果见表 1. 表 1 中，$5n$ 为系数矩阵的阶数，t 为计算时间，k_1 为算法 1 的迭代次数，k_{12} 为算法 1 的中断次数，k_2 为算法 2 的迭代次数，error 为实际误差的范数.

表 1　矩阵方程(1) 的计算结果

$5n$	t	k_1	k_{12}	k_2	error
5	0.895 1	63	0	0	6.186 3e－08
10	0.081 9	358	0	0	9.896 8e－09
15	0.241 3	995	0	0	8.342 0e－09
20	0.287 0	2 027	0	0	8.439 5e－09

从表 1 可以看出，运用算法 1 和算法 2 求解方程(1) 是有效的. 当算法 1 中断时，可以通过算法 2 计算方程(1) 的中心对称矩阵最小二乘解. 若修改算法 1 中矩阵 $\boldsymbol{Q}_k$ 的类型，则可以建立求其他特殊矩阵解的迭代算法.

参考文献：

[1]　张冰尘，戴博伟. 一种基于随机滤波的神经动作电位信号压缩感知采样方法[J]. 电子与信息学报，2013，35(9)：2283-2286.

[2]　郑岩，李健，李智. 一种基于随机滤波器的测量矩阵优化算法[J]. 小型微型计算机系统，2016，37(3)：632-636.

[3]　唐智礼. 约束最优控制理论及其在气动优化中的应用[J]. 力学学报，2007，39(2)：273-277.

[4]　于秀萍，官英双. 基于 H_∞ 控制理论的 BTT 导弹自动驾驶仪设计[J]. 系统工程与电子技术，2008，30(5)：905-908.

[5]　吕巍，魏良亭，冯恩民. 一类求解非线性奇异方程组的牛顿改进算法[J]. 控制与决策，2017，32(12)：2240-2246.

[6]　徐莹莹，张惠. 一类改进的拟牛顿算法[J]. 井冈山大学学报(自然科学版)，2018，39(1)：21-23.

[7]　万中，冯冬冬. 无约束优化问题的精细修正牛顿算法[J]. 高校应用数学学报，2011，26(2)：179-186.

[8]　程可欣，彭振赟，杜丹丹. 矩阵方程 $\boldsymbol{X}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}^{-1}\boldsymbol{A}=\boldsymbol{Q}$ 的牛顿迭代解法[J]. 工程数学学报，2016，33(1)：63-72.

[9]　张肖肖，张凯院，宋卫红. 含高次逆幂的矩阵方程对称解的双迭代算法[J]. 数学杂志，2016，36(2)：437-444.

[10]　程云彭. 矩阵论[M]. 西安：西北工业大学出版社，2004：147.

[11]　武见，张凯院. 多变量矩阵方程异类约束解的修正共轭梯度算法[J]. 工程数学学报，2012，29(1)：112-116.