DOI:10.19789/j.1004-9398.2022.05.002
文献引用:肖云,温瑞萍. Toeplitz矩阵填充的尾端修正增广拉格朗日乘子算法[J].首都师范大学学报(自然科学版),2022,43(5):8-15. XIAO Y, WEN R P. Tailed-modification augmented Lagrange multiplier algorithm for Toeplitz matrix completion [J]. Journal of Capital Normal University (Natural Science Edition),2022,43(5):8-15.

Toeplitz矩阵填充的尾端修正增广拉格朗日乘子算法*

肖 云,温瑞萍**

(太原师范学院工程科学计算山西省高等学校重点实验室,山西 晋中 030619)

摘要:基于均值的增广拉格朗日乘子(MALM)算法,提出了一种尾端修正的Toeplitz矩阵填充新算法.该算法利用增广拉格朗日乘子(ALM)算法迭代速度较快的优点,对迭代矩阵序列进行结构化与尾端修正.在一定程度上减少了每步均值处理所产生的数据传输量,从而降低了计算代价.同时详细讨论了新算法的收敛性.最后通过数值实验证明了新算法比ℓ步修正的增广垃格朗日乘子(ℓ-MALM)、MALM以及ALM算法在计算时间上有较大程度的减少.

关键词:Toeplitz矩阵;增广拉格朗日乘子算法;均值;矩阵填充
中图分类号:O241.6

Tailed-modification augmented Lagrange multiplier algorithm for Toeplitz matrix completion*

XIAO Yun, WEN Ruiping**

(Key Laboratory for Engineering & Computing Science, Shanxi Provincial Department of Education, Taiyuan Normal University, Jinzhong Shanxi 030619)

Abstract: Based on the mean augmented Lagrange multiplier (MALM) algorithm, a tail modified Toeplitz matrix completion algorithm is proposed. The algorithm uses the advantage of the augmented Lagrange multiplier (ALM) algorithm to structure and modify the iteration matrix sequence. To some extent, the data transmission generated by each step of mean processing is reduced, and thus reduces the computational cost. The convergence of the new algorithm is discussed in detail. Finally, the numerical experiments show that the new algorithm has a greater reduction in computing time than ℓ-MALM, MALM and ALM algorithm.

Keywords: Toeplitz matrix; augmented Lagrange multiplier algorithm; mean value; matrix completion
CLC: O241.6

0 引 言

矩阵填充问题假设当目标矩阵具有低秩或近似低秩的特性时,其能够对含有大量缺失元素的矩阵进行近乎完美的填充.填充一个未知的低秩或近似低秩采样矩阵是一个具有挑战性的问题.目前,

收稿日期:2021-05-17
* 国家自然科学基金项目(11371275);山西省自然科学基金项目(201901D211423)
** 通信作者:wenrp@163.com

矩阵填充在机器学习[1-2]、推荐系统[3]、图像处理[4]、控制理论[5]和计算机视觉[6]等数据科学领域中都有非常重要的应用.

在2009年,Candès和Recht[7]建立了关于矩阵填充问题的凸优化模型:

$$\min_{\boldsymbol{A}\in\mathbf{R}^{m\times n}}\|\boldsymbol{A}\|_*,\ \text{s.t. } P_{\Omega}(\boldsymbol{A})=P_{\Omega}(\boldsymbol{D}),\tag{1}$$

式中:$\boldsymbol{D}\in\mathbf{R}^{m\times n}$是缺失部分元素的已知矩阵,$\boldsymbol{A}\in\mathbf{R}^{m\times n}$是逼近目标矩阵$\boldsymbol{D}$的未知矩阵,核范数$\|\boldsymbol{A}\|_*=\sum_{k=1}^{r}\sigma_k(\boldsymbol{A})$,$\sigma_k(\boldsymbol{A})$表示秩为$r$的矩阵$\boldsymbol{A}\in\mathbf{R}^{m\times n}$的第$k$大奇异值,且$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_k\geqslant\cdots\geqslant\sigma_r>0$.

对于求解问题(1),通常算法中的每次迭代都需要计算一个$m\times n$阶矩阵的奇异值分解(singular value decomposition,SVD).近年来,基于核范数最小化的研究,有不少的成果.如文献[8]提出了一种易于实现的奇异值阈值算法(singular value thresholding,SVT),该算法可以得到具有低秩结构的最优解,并保证了填充矩阵的稀疏性,一定程度上节省了存储空间,可以有效地处理低秩的大规模矩阵填充问题.但当秩较大时,此算法收敛速度缓慢甚至效果不是很理想,在文献[9]给出了加速SVT算法.与此同时,Toh和Yun[10]提出了加速临近梯度算法(accelerated proximal gradient,APG);Lin等[11]提出了比APG算法速度快、精度高的增广拉格朗日乘子算法(augmented lagrange multiplier,ALM),该算法具有很好的操作性和收敛性,并且需要较小的存储空间.

由以上算法结果报告,SVD计算花费占去整个算法花费的80%以上,而Toeplitz矩阵的SVD具有较低的算法复杂度,且一个$n\times n$阶的Toeplitz矩阵$\boldsymbol{T}\in\mathbf{R}^{n\times n}$表达形式为

$$\boldsymbol{T}=\begin{pmatrix}t_0 & t_1 & \cdots & t_{n-2} & t_{n-1}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ t_{-n+1} & t_{-n+2} & \cdots & t_{-1} & t_0\end{pmatrix}\in\mathbf{R}^{n\times n},\tag{2}$$

那是由$2n-1$个元素决定的,即表达式(2)的第一行和第一列元素.近年来,针对Toeplitz矩阵填充问题,方云兰和郑慧娆[12]研究了Toeplitz块矩阵填充的快速正交三角分解算法;于益华和李云翔[13]研究了基于快速小波变换的Toeplitz矩阵填充算法;Wang和Li[14-16]提出了Toeplitz矩阵填充的均值算法、修正的ALM算法、保结构算法[16];刘丽霞和王川龙[17]进一步提出了子空间算法;Wen等[18]、Wen和Li[19]研究了逐步结构化的ALM算法以及ℓ步结构化的ALM算法.

本文主要是基于均值的增广拉格朗日乘子算法(ALM based on mean value,MALM)算法,对中小型规模的低秩矩阵进行填充.为了减少MALM算法中频繁的数据传输量与结构化操作所增加的计算量,采用ALM算法,当迭代矩阵逐渐靠近目标矩阵时,开始对矩阵进行尾部修正并结构化,修正使迭代矩阵从数值上更加靠近目标矩阵,结构化使迭代矩阵与目标矩阵保持同样结构,可利用快速奇异值分解,从而减少迭代步数,缩短奇异值分解时间,节约算法的整体花费.同时给出了算法收敛性分析,并通过数值实验验证了算法的有效性.下面给出一些定义:

定义1[20]**(SVD)**　对于秩为r的矩阵$\boldsymbol{A}\in\mathbf{R}^{m\times n}$,必存在正交矩阵$\boldsymbol{U}\in\mathbf{R}^{m\times r}$和$\boldsymbol{V}\in\mathbf{R}^{n\times r}$,使得

$$\boldsymbol{A}=\boldsymbol{U}\Sigma_r\boldsymbol{V}^{\mathrm{T}},\ \Sigma_r=\operatorname{diag}\left(\sigma_1,\cdots,\sigma_r\right),$$

其中

$$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_r>0.$$

定义2[9]**(SVT)**　对于任意参数$\tau\geqslant0$,秩为r的矩阵$\boldsymbol{A}\in\mathbf{R}^{m\times n}$,存在奇异值分解$\boldsymbol{A}=\boldsymbol{U}\Sigma_r\boldsymbol{V}^{\mathrm{T}}$,奇异值阈值算子$D_\tau$定义为

$$D_\tau(\boldsymbol{A}):=\boldsymbol{U}D_\tau(\Sigma)\boldsymbol{V}^{\mathrm{T}},D_\tau(\Sigma)=\operatorname{diag}\left(\left\{\sigma_i-\tau\right\}_+\right)$$

其中

$$\left\{\sigma_i-\tau\right\}_+=\begin{cases}\sigma_i-\tau, & \text{当}\sigma_i>\tau\text{时},\\ 0, & \text{当}\sigma_i\leqslant\tau\text{时},\end{cases}$$

定义3[22]**(移位矩阵)**　对$n\times n$阶单位矩阵$\boldsymbol{I}_n=\left(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\right)\in\mathbf{R}^{n\times n}$,称$\boldsymbol{S}_n=\left(\boldsymbol{e}_2,\boldsymbol{e}_3,\cdots,\boldsymbol{e}_n,\boldsymbol{0}\right)\in\mathbf{R}^{n\times n}$为移位矩阵,满足

$$\boldsymbol{S}_n^r=\begin{cases}\begin{pmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{I}_{n-r} & \boldsymbol{O}\end{pmatrix}, & 1<r<n\\ \boldsymbol{O}, & r\geqslant n\end{cases},$$

这里"$\boldsymbol{O}$"表示零矩阵.

定义4[18]**(Toeplitz结构化算子)**　对于任意矩阵$\boldsymbol{A}=\left(a_{ij}\right)\in\mathbf{R}^{n\times n}$,Toeplitz结构化算子$\Gamma$定义为

$$\Gamma(\boldsymbol{A}):=\sum_{\alpha=1}^{n-1}\tilde{a}_{-\alpha}\boldsymbol{S}_n^{\alpha}+\sum_{\alpha=0}^{n-1}\tilde{a}_{\alpha}\left(\boldsymbol{S}_n^{\mathrm{T}}\right)^{\alpha},\tag{3}$$

式中:$\tilde{a}_\alpha=\operatorname{mean}\left(\operatorname{diag}\left(\boldsymbol{A},\alpha\right)\right)$,$\alpha=-n+1,-n+2,\cdots,n-1$.

可见,$\Gamma(\boldsymbol{A})$是由矩阵$\boldsymbol{A}$派生的Toeplitz矩阵.换句话说,$\forall\boldsymbol{A}\in\mathbf{R}^{n\times n}$,都可以通过结构化算子$\Gamma(\cdot)$转化为一个Toeplitz矩阵.

本文拟简单回顾ALM和MALM,并分别比较2种算法所具有的优缺点,介绍关于Toeplitz矩阵填

充的尾端修正ALM算法(Tailed-modification ALM, ℓ-TMALM);运用核范数的性质给出新算法的收敛性分析;通过与ℓ- MALM,MALM以及ALM算法作比较得出新算法的有效性.

符号说明.为方便起见,$\mathbf{R}^{m\times n}$表示的是$m\times n$阶的实数矩阵.矩阵$\boldsymbol{A}$的核范数用$\|\boldsymbol{A}\|_*$表示,而$\|\boldsymbol{A}\|_F$表示矩阵$\boldsymbol{A}$的Frobenius范数.$\boldsymbol{A}^{\mathrm{T}}$表示矩阵$\boldsymbol{A}$的转置.对于一个Toeplitz矩阵$\boldsymbol{A}\in\mathbf{R}^{n\times n}$, $\mathrm{diag}(\boldsymbol{A},\ell)$表示矩阵$\boldsymbol{A}$的第$\ell$条对角线元素,$\ell=-n+1,\cdots,n-1$. $\Omega\subset\{-n+1,\cdots,n-1\}$是采样集,$P_\Omega$是相应矩阵在$\Omega$上的投影算子,满足

$$\mathrm{diag}\left(P_\Omega(\boldsymbol{A}),\ell\right)=\begin{cases}\mathrm{diag}(\boldsymbol{A},\ell), & \ell\in\Omega\\ \boldsymbol{0}, & \ell\notin\Omega\end{cases},$$

其中$\boldsymbol{0}$是一个零向量.

1 相关算法

低秩Toeplitz矩阵的填充问题.为了方便比较,先简单回顾ALM算法,并将其应用到Toeplitz矩阵填充问题中.

1.1 ALM[11]

当问题(1)中的$m=n$时,即矩阵$\boldsymbol{A},\boldsymbol{E},\boldsymbol{D}\in\mathbf{R}^{n\times n}$为Toeplitz矩阵时,求解式(1)的等价形式为

$$\min\|\boldsymbol{A}\|_*,\ \text{s.t.}\ \boldsymbol{A}+\boldsymbol{E}=\boldsymbol{D},P_\Omega(\boldsymbol{E})=\boldsymbol{O}. \tag{4}$$

利用ALM求解式(4),由文献[11]知其拉格朗日函数为

$$\mathcal{L}\left(\boldsymbol{A},\boldsymbol{E},\boldsymbol{Y},\mu\right)=\|\boldsymbol{A}\|_*+\langle\boldsymbol{Y},\boldsymbol{D}-\boldsymbol{A}-\boldsymbol{E}\rangle+\frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{A}-\boldsymbol{E}\|_F^2, \tag{5}$$

式中:$\boldsymbol{Y}\in\mathbf{R}^{n\times n}$为拉格朗日乘数,$\mu$是正常数.算法步骤如下:

第1步:给定下标集合Ω,样本Toeplitz矩阵$\boldsymbol{D}$,参数$\rho>1,\mu_0>0$,以及初始矩阵$\boldsymbol{Y}_0=\boldsymbol{O},\boldsymbol{E}_0=\boldsymbol{O},k:=0$.

第2步:计算矩阵$\left(\boldsymbol{D}-\boldsymbol{E}_k+\mu_k^{-1}\boldsymbol{Y}_k\right)$的奇异值分解

$$[\boldsymbol{U}_k,\Sigma_k,\boldsymbol{V}_k]_{\mu_k^{-1}}=\mathrm{svd}\left(\boldsymbol{D}-\boldsymbol{E}_k+\mu_k^{-1}\boldsymbol{Y}_k\right).$$

第3步:令

$$\boldsymbol{A}_{k+1}=\boldsymbol{U}_k\boldsymbol{D}_{\mu_k^{-1}}\left(\Sigma_k\right)\boldsymbol{V}_k^{\mathrm{T}},$$

$$\boldsymbol{E}_{k+1}=P_{\bar{\Omega}}\left(\boldsymbol{D}-\boldsymbol{A}_{k+1}+\mu_k^{-1}\boldsymbol{Y}_k\right).$$

第4步:通过给定参数,若$\left\|\boldsymbol{D}-\boldsymbol{A}_{k+1}-\boldsymbol{E}_{k+1}\right\|_F\big/\|\boldsymbol{D}\|_F<\varepsilon_1$且$\mu_k\left\|\boldsymbol{E}_{k+1}-\boldsymbol{E}_k\right\|_F\big/\|\boldsymbol{D}\|_F<\varepsilon_2$,停止;否则转第5步.

第5步:令$\boldsymbol{Y}_{k+1}=\boldsymbol{Y}_k+\mu_k\left(\boldsymbol{D}-\boldsymbol{A}_{k+1}-\boldsymbol{E}_{k+1}\right)$.如果$\mu_k\left\|\boldsymbol{E}_{k+1}-\boldsymbol{E}_k\right\|_F\big/\|\boldsymbol{D}\|_F<\varepsilon_2$,令$\mu_{k+1}=\rho\mu_k$;否则转第2步.

数值实验表明,ALM算法在理论和实际应用中均有较好的效果,且该算法迭代快,具有线性全局收敛速度,但奇异值分解时间较长,精度较低,收敛效果不是很理想.

1.2 对偶算法(dual algorithm)[21]

对偶算法是为了通过其对偶性求解问题(4),即首先要解决的问题是

$$\max_{\boldsymbol{Y}}\langle\boldsymbol{D},\boldsymbol{Y}\rangle\ \ \text{s.t.}\ \ J(\boldsymbol{A})\leqslant 1. \tag{6}$$

对于最优增广拉格朗日乘子$\boldsymbol{Y}$,有

$$J(\boldsymbol{Y})=\max\left(\|\boldsymbol{Y}\|_2,\lambda^{-1}\|\boldsymbol{Y}\|_\infty\right). \tag{7}$$

最速上升算法的约束集为$\{\boldsymbol{Y}\}$可以用来解决问题(6),其中被约束的最速上升方向是通过将$\boldsymbol{D}$投影在凸的目标区域$\{\boldsymbol{Y}\}$上而获得.结果表明在寻找被约束的最速上升方向的过程中可以得到原问题(4)的最优解.

1.3 MALM[15]

填充Toeplitz矩阵的MALM,其思想是在ALM算法的基础上,将迭代矩阵的对角元素均值化后重新赋值,使其保持Toeplitz结构.问题表述为

$$\min_{\boldsymbol{A},\boldsymbol{E}\in\mathbf{R}^{m\times n}}\|\boldsymbol{A}\|_*,\text{s.t.}\ \boldsymbol{A}+\boldsymbol{E}=P_\Omega(\boldsymbol{M}),P_\Omega(\boldsymbol{E})=\boldsymbol{O}, \tag{8}$$

式中:$\boldsymbol{A},\boldsymbol{M}\in\mathbf{R}^{n\times n}$均为Toeplitz矩阵,$\Omega\subset\{-n+1,\cdots,n-1\}$.令$\boldsymbol{D}=P_\Omega(\boldsymbol{M})$,增广拉格朗日函数可表示为

$$\mathcal{L}\left(\boldsymbol{A},\boldsymbol{E},\boldsymbol{Y},\mu\right)=\|\boldsymbol{A}\|_*+\langle\boldsymbol{Y},\boldsymbol{D}-\boldsymbol{A}-\boldsymbol{E}\rangle+\frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{A}-\boldsymbol{E}\|_F^2, \tag{9}$$

式中$\boldsymbol{Y}\in\mathbf{R}^{n\times n}$.算法步骤如下:

第1步:给定下标集合Ω,样本Toeplitz矩阵$\boldsymbol{D}$,参数$\rho>1,\mu_0>0$,以及初始矩阵$\boldsymbol{Y}_0=\boldsymbol{O},\boldsymbol{E}_0=\boldsymbol{O},k:=0$.

第2步:利用Lanczos计算矩阵$\left(\boldsymbol{D}-\boldsymbol{E}_k+\mu_k^{-1}\boldsymbol{Y}_k\right)$的奇异值分解

$$[\boldsymbol{U}_k,\Sigma_k,\boldsymbol{V}_k]=\mathrm{lansvd}\left(\boldsymbol{D}-\boldsymbol{E}_k+\mu_k^{-1}\boldsymbol{Y}_k\right).$$

第3步:令

$$\boldsymbol{X}_{k+1}=\boldsymbol{U}_k\boldsymbol{D}_{\mu_k^{-1}}\left(\Sigma_k\right)\boldsymbol{V}_k^{\mathrm{T}},$$

计算 $\tilde{a}_\alpha=\mathrm{mean}\left(\mathrm{diag}\left(\boldsymbol{X}_{k+1},\alpha\right)\right)$, $\alpha=-n+1,-n+2,\cdots,n-1$,

$$\boldsymbol{A}_{k+1}=\Gamma\left(\boldsymbol{X}_{k+1}\right)=\sum_{\alpha=1}^{n-1}\tilde{a}_{-\alpha}\boldsymbol{S}_n^{\alpha}+\sum_{\alpha=0}^{n-1}\tilde{a}_\alpha\left(\boldsymbol{S}_n^{\mathrm{T}}\right)^{\alpha},$$

$$\boldsymbol{E}_{k+1}=P_{\bar{\Omega}}\left(\boldsymbol{D}-\boldsymbol{A}_{k+1}+\mu_k^{-1}\boldsymbol{Y}_k\right).$$

第4步:通过给定参数,若$\left\|\boldsymbol{D}-\boldsymbol{A}_{k+1}-\boldsymbol{E}_{k+1}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_1$且$\mu_k\left\|\boldsymbol{E}_{k+1}-\boldsymbol{E}_k\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_2$,停止;否则转第5步.

第5步:令$\boldsymbol{Y}_{k+1}=\boldsymbol{Y}_k+\mu_k\left(\boldsymbol{D}-\boldsymbol{A}_{k+1}-\boldsymbol{E}_{k+1}\right)$.如果$\mu_k\left\|\boldsymbol{E}_{k+1}-\boldsymbol{E}_k\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_2$,令$\mu_{k+1}=\rho\mu_k$;否则转第2步.

数值实验表明,与ALM和SVT算法相比,MALM算法的精度较高,收敛效果较好,对于在迭代过程中进行结构化处理的MALM算法保持了Toeplitz矩阵的结构,从而可以利用快速奇异值分解,使得在奇异值分解时间上要远短于其他2种算法;但是相比ALM算法,MALM算法在均值步增加了数据的传输.而对于一些计算机来讲,数据传输频繁会造成数据堆积而拖延计算时间,因此,基于减少数据传输量的思考,在迭代的尾端,即接近目标矩阵时再进行结构化处理是增效的选择.

1.4　ℓ-TMALM

对矩阵利用ALM算法[11]的快速迭代ℓ步之后,进行尾部修正及结构化处理.一方面,避免了均值步产生的频繁数据传输;另一方面,尾部修正使得迭代矩阵更加靠近目标矩阵,从而快速达到可行解,减少奇异值分解时间.算法步骤如下:

第1阶段:ALM算法步.

第1步:给定下标集合Ω,样本Toeplitz矩阵$\boldsymbol{D}$,参数$\rho>1,\mu_0>0$,以及初始矩阵$\boldsymbol{Y}_{0,0}=\boldsymbol{O},\boldsymbol{E}_{0,0}=\boldsymbol{O},k=0,q=1,2,\cdots,\ell-1$.

第2步:令

$$\boldsymbol{Z}_{k,q}=\boldsymbol{D}-\boldsymbol{A}_{k,q}-\boldsymbol{E}_{k,q},$$
$$\boldsymbol{B}_{k,q}=\boldsymbol{D}-\boldsymbol{E}_{k,q}+\mu_{k,q}^{-1}\boldsymbol{Y}_{k,q}$$

利用Lanczos对矩阵$\boldsymbol{B}_{k,q}$进行奇异值分解

$$[\boldsymbol{U}_{k,q},\Sigma_{k,q},\boldsymbol{V}_{k,q}]=\mathrm{lansvd}\left(\boldsymbol{B}_{k,q}\right).$$

第3步:令

$$\boldsymbol{X}_{k+1,q+1}=\boldsymbol{U}_{k,q}\boldsymbol{D}_{\mu_{k,q}^{-1}}\left(\Sigma_{k,q}\right)\boldsymbol{V}_{k,q}^{\mathrm{T}},$$

$$\boldsymbol{E}_{k+1,q+1}=P_{\bar{\Omega}}\left(\boldsymbol{D}-\boldsymbol{X}_{k+1,q+1}+\mu_{k,q}^{-1}\boldsymbol{Y}_{k,q}\right).$$

第4步:通过给定参数,若$\left\|\boldsymbol{D}-\boldsymbol{X}_{k+1,q+1}-\boldsymbol{E}_{k+1,q+1}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_1$且$\mu_{k,q}\left\|\boldsymbol{E}_{k+1,q+1}-\boldsymbol{E}_{k,q}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_2$,停止;否则转第5步.

第5步:令$\boldsymbol{Y}_{k+1,q}=\boldsymbol{Y}_{k,q}+\mu_{k,q}\left(\boldsymbol{D}-\boldsymbol{X}_{k+1,q+1}-\boldsymbol{E}_{k+1,q+1}\right)$.如果$\mu_{k,q}\left\|\boldsymbol{E}_{k+1,q+1}-\boldsymbol{E}_{k,q}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_2$,令$\mu_{k+1,q+1}=\rho\mu_{k,q}$,继续进行第2阶段尾步修正;否则转第2步.

第2阶段:尾部修正.

第1步:计算矩阵$\boldsymbol{B}_{k,\ell}$的奇异值分解,

$$[\boldsymbol{U}_{k,\ell},\Sigma_{k,\ell},\boldsymbol{V}_{k,\ell}]=\mathrm{lansvd}\left(\boldsymbol{B}_{k,\ell}\right).$$

第2步:令

$$\boldsymbol{X}_{k+1,\ell}=\boldsymbol{U}_{k,\ell}\boldsymbol{D}_{\mu_{k,\ell}^{-1}}\left(\Sigma_{k,\ell}\right)\boldsymbol{V}_{k,\ell}^{\mathrm{T}},$$
$$\boldsymbol{E}_{k+1,\ell}=P_{\bar{\Omega}}\left(\boldsymbol{D}-\boldsymbol{X}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}\right),$$
$$\boldsymbol{Y}_{k+1,\ell}=\boldsymbol{Y}_{k,\ell}+\mu_{k,\ell}\left(\boldsymbol{D}-\boldsymbol{X}_{k+1,\ell}-\boldsymbol{E}_{k+1,\ell}\right),$$
$$\boldsymbol{B}_{k+1,\ell}=\boldsymbol{D}-\boldsymbol{E}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}.$$

第3步:令

$$\boldsymbol{Z}_{k,\ell+1}=\boldsymbol{B}_{k+1,\ell}-\boldsymbol{B}_{k,\ell}+\boldsymbol{Z}_{k,\ell},$$
$$\boldsymbol{E}_{k+1,\ell+1}=P_{\bar{\Omega}}\left(\boldsymbol{D}-\boldsymbol{Z}_{k,\ell+1}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}\right),$$
$$\boldsymbol{Y}_{k+1,\ell+1}=\boldsymbol{Y}_{k+1,\ell}+\mu_{k,\ell}\left(\boldsymbol{D}-\boldsymbol{Z}_{k,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right).$$

第4步:计算均值$\tilde{a}_\alpha=\mathrm{mean}\left(\mathrm{diag}\left(\boldsymbol{Z}_{k,\ell+1},\alpha\right)\right)$,$\alpha=-n+1,-n+2,\cdots,n-1$,令

$$\boldsymbol{A}_{k+1,\ell+1}=\Gamma\left(\boldsymbol{Z}_{k,\ell+1}\right)=\sum_{\alpha=1}^{n-1}\tilde{a}_{-\alpha}\boldsymbol{S}_n^{\alpha}+\sum_{\alpha=0}^{n-1}\tilde{a}_\alpha\left(\boldsymbol{S}_n^{\mathrm{T}}\right)^{\alpha},$$
$$\bar{\boldsymbol{E}}_{k+1,\ell+1}=P_{\bar{\Omega}}\left(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}\right).$$

第5步:若$\left\|\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_1$且$\mu_{k,\ell}\left\|\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k,\ell}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_2$,停止.

第6步:令$\bar{\boldsymbol{Y}}_{k+1,\ell+1}=\boldsymbol{Y}_{k+1,\ell+1}+\mu_{k,\ell}\left(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1}\right)$,若$\mu_{k,\ell}\left\|\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k,\ell}\right\|_F/\|\boldsymbol{D}\|_F<\varepsilon_2$,令$\mu_{k+1,\ell}=\rho\mu_{k,\ell}$;否则转第2阶段第1步.

该算法分为2个阶段执行:第1阶段,使用ALM算法,利用ALM算法本身迭代速度快的性质,当其迭代到一定程度,即迭代矩阵逐渐靠近目标矩阵时;进入第2阶段,对迭代矩阵序列进行尾部修正并结构化,结构化使迭代矩阵与目标矩阵保持同样结构,从而可利用快速奇异值分解,修正使迭代矩阵从数值上更加逼近目标矩阵,以便快速地达到可行解.该算法不仅减少了频繁的数据传输量,而且在

一定程度上节省了奇异值分解时间.

2 收敛性分析

运用Toeplitz矩阵的特殊结构,证明算法1.4的收敛性.在以下叙述中,序列$\{\boldsymbol{Y}_{k,\ell}\}$,$\{\boldsymbol{E}_{k,\ell}\}$,$\{\boldsymbol{A}_{k,\ell}\}$,$\{\bar{\boldsymbol{Y}}_{k+1,\ell+1}\}$,$\{\mu_{k,\ell}\}$的意义均按算法1.4表述的格式产生.并令$(\ddot{\boldsymbol{A}},\ddot{\boldsymbol{E}})$为问题(4)的最优解,序列$\{\ddot{\boldsymbol{Y}}_k\}$的意义按对偶算法1.2表述的格式产生,$\ddot{\boldsymbol{Y}}$为对偶问题(6)的最优解.

引理2.1[7] 设任意矩阵$\boldsymbol{A}\in\mathbf{R}^{m\times n}$和其奇异值分解为$\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$,则$\boldsymbol{A}$的核范数的次梯度集为

$$\partial\|\boldsymbol{A}\|_*=\left\{\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}+\boldsymbol{W}:\boldsymbol{W}\in\mathbf{R}^{m\times n},\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}=\boldsymbol{O},\boldsymbol{W}\boldsymbol{V}=\boldsymbol{O},\|\boldsymbol{W}\|_2\leqslant 1\right\}. \tag{10}$$

引理2.2[11] 如果μ_k是不减序列,则下列级数的每一项都是非负的,且

$$\sum_{k=1}^{+\infty}\mu_k^{-1}\left(\left\langle\boldsymbol{Y}_{k+1}-\boldsymbol{Y}_k,\boldsymbol{E}_{k+1}-\boldsymbol{E}_k\right\rangle+\left\langle\boldsymbol{A}_{k+1}-\ddot{\boldsymbol{A}},\hat{\boldsymbol{Y}}_{k+1}-\ddot{\boldsymbol{Y}}\right\rangle+\left\langle\boldsymbol{E}_{k+1}-\ddot{\boldsymbol{E}},\boldsymbol{Y}_{k+1}-\ddot{\boldsymbol{Y}}\right\rangle\right)<+\infty, \tag{11}$$

即级数是收敛的.

引理2.3[11] 序列$\{\ddot{\boldsymbol{Y}}_k\}$,$\{\boldsymbol{Y}_k\}$和$\{\hat{\boldsymbol{Y}}_k\}$均是有界的,且$\hat{\boldsymbol{Y}}_k=\boldsymbol{Y}_{k-1}+\mu_{k-1}(\boldsymbol{D}-\boldsymbol{A}_k-\boldsymbol{E}_k)$.

引理2.4 由算法1.4生成的矩阵序列$\{\bar{\boldsymbol{Y}}_{k+1,\ell+1}\}$是有界的.

证明令$\boldsymbol{B}=\mu_{k,\ell}(\boldsymbol{D}-\boldsymbol{E}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}-\boldsymbol{X}_{k+1,\ell})$,$\Gamma(\boldsymbol{B})=\sum_{\alpha=1}^{n-1}\tilde{b}_{-\alpha}\boldsymbol{S}_n^{\alpha}+\sum_{\alpha=0}^{n-1}\tilde{b}_{\alpha}(\boldsymbol{S}_n^{\mathrm{T}})^{\alpha}$,定义见式(3).

证明 $\boldsymbol{Y}_{k,\ell},\boldsymbol{E}_{k,\ell},k=1,2,\cdots$是Toeplitz矩阵.由算法1.4可知,$\boldsymbol{Y}_{0,0}=\boldsymbol{O},\boldsymbol{E}_{0,0}=\boldsymbol{O}$为Toeplitz矩阵.假定$\boldsymbol{Y}_{k,\ell},\boldsymbol{E}_{k,\ell}$均为Toeplitz矩阵,可得$\bar{\boldsymbol{E}}_{k+1,\ell+1}=P_{\bar{\Omega}}(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell})$也为Toeplitz矩阵.因此,算法1.4第6步中的$\bar{\boldsymbol{Y}}_{k+1,\ell+1}$也是一个Toeplitz矩阵.

$$\bar{\boldsymbol{Y}}_{k+1,\ell+1}=\boldsymbol{Y}_{k+1,\ell+1}+\mu_{k,\ell}(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1})=\boldsymbol{Y}_{k+1,\ell+1}+\mu_{k,\ell}(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1})+\mu_{k,\ell}(\boldsymbol{E}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1}).$$

且

$$\mu_{k,\ell}(\boldsymbol{E}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1})=\mu_{k,\ell}P_{\bar{\Omega}}(\boldsymbol{D}-\boldsymbol{Z}_{k,\ell+1}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}-\bar{\boldsymbol{E}}_{k+1,\ell+1})=$$

$$\mu_{k,\ell}P_{\bar{\Omega}}(\boldsymbol{D}-\boldsymbol{D}+\boldsymbol{E}_{k+1,\ell}-\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}+\boldsymbol{A}_{k,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}-\bar{\boldsymbol{E}}_{k+1,\ell+1})=$$

$$\mu_{k,\ell}P_{\bar{\Omega}}(\boldsymbol{D}-\boldsymbol{X}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}-\boldsymbol{D}+\boldsymbol{A}_{k+1,\ell+1}-\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}+\boldsymbol{A}_{k,\ell})=$$

$$\mu_{k,\ell}P_{\bar{\Omega}}(-\boldsymbol{D}+\boldsymbol{E}_{k+1,\ell}-\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{B}+\Gamma(\boldsymbol{Z}_{k,\ell+1})+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k,\ell}+\boldsymbol{A}_{k,\ell})=$$

$$\mu_{k,\ell}P_{\bar{\Omega}}\Gamma(\mu_{k,\ell}^{-1}\boldsymbol{B})=P_{\bar{\Omega}}\Gamma(\boldsymbol{B}).$$

由算法1.4中的第1步和第2步可知,

$$\boldsymbol{D}-\boldsymbol{E}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}=\check{\boldsymbol{U}}_{k,\ell}\check{\Sigma}_{k,\ell}\check{\boldsymbol{V}}_{k,\ell}+\tilde{\boldsymbol{U}}_{k,\ell}\tilde{\Sigma}_{k,\ell}\tilde{\boldsymbol{V}}_{k,\ell},$$

式中:$\check{\boldsymbol{U}}_{k,\ell},\check{\boldsymbol{V}}_{k,\ell}$是奇异值$>\mu_{k,\ell}^{-1}$的奇异向量且$\tilde{\boldsymbol{U}}_{k,\ell},\tilde{\boldsymbol{V}}_{k,\ell}$奇异值$<\mu_{k,\ell}^{-1}$的奇异向量,对角矩阵$\check{\Sigma}_{k,\ell}$的元素$>\mu_{k,\ell}^{-1}$,而对角矩阵$\tilde{\Sigma}_{k,\ell}$的元素$\leqslant\mu_{k,\ell}^{-1}$.因此,$\boldsymbol{X}_{k+1,\ell}=\check{\boldsymbol{U}}_{k,\ell}(\check{\Sigma}_{k,\ell}-\mu_{k,\ell}^{-1}\boldsymbol{I})\check{\boldsymbol{V}}_{k,\ell}^{\mathrm{T}}$以及

$$\|\boldsymbol{B}\|_F=\left\|\mu_{k,\ell}(\boldsymbol{D}-\boldsymbol{E}_{k+1,\ell}+\mu_{k,\ell}^{-1}\boldsymbol{Y}_{k+1,\ell}-\boldsymbol{X}_{k+1,\ell})\right\|_F=\left\|\mu_{k,\ell}(\mu_{k,\ell}^{-1}\check{\boldsymbol{U}}_{k,\ell}\check{\boldsymbol{V}}_{k,\ell}^{\mathrm{T}}+\tilde{\boldsymbol{U}}_{k,\ell}\tilde{\Sigma}_{k,\ell}\tilde{\boldsymbol{V}}_{k,\ell}^{\mathrm{T}})\right\|_F=\left\|\check{\boldsymbol{U}}_{k,\ell}\check{\boldsymbol{V}}_{k,\ell}^{\mathrm{T}}+\mu_{k,\ell}\tilde{\boldsymbol{U}}_{k,\ell}\tilde{\Sigma}_{k,\ell}\tilde{\boldsymbol{V}}_{k,\ell}^{\mathrm{T}}\right\|_F\leqslant\sqrt{n}.$$

综合引理2.2和引理2.3的结论可知,

$$\boldsymbol{Y}_{k+1,\ell+1}+\mu_{k,\ell}(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1})\in\partial\left\|\boldsymbol{A}_{k+1,\ell+1}\right\|_*.$$

再由引理2.1,$\boldsymbol{A}_{k+1,\ell+1}=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$,

$$\partial\left\|\boldsymbol{A}_{k+1,\ell+1}\right\|_*=\left\{\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}+\boldsymbol{W}:\boldsymbol{W}\in\mathbf{R}^{n\times n},\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}=\boldsymbol{O},\boldsymbol{W}\boldsymbol{V}=\boldsymbol{O},\|\boldsymbol{W}\|_2\leqslant 1\right\},$$

且

$$\left\|\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}+\boldsymbol{W}\right\|_F^2=\mathrm{tr}\left((\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}+\boldsymbol{W})^{\mathrm{T}}(\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}+\boldsymbol{W})\right)=\mathrm{tr}\left(\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}+\boldsymbol{W}^{T}\boldsymbol{W}\right)\leqslant n.$$

因此,有下列不等式成立

$$\left\|\boldsymbol{Y}_{k+1,\ell+1}+\mu_{k,\ell}(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1})\right\|_F\leqslant\sqrt{n},$$

以及

$$\left\|P_{\bar{\Omega}}(\Gamma(\boldsymbol{B}))\right\|_F\leqslant\left\|\Gamma(\boldsymbol{B})\right\|_F\leqslant\|\boldsymbol{B}\|_F\leqslant\sqrt{n}.$$

所以矩阵序列$\{\bar{\boldsymbol{Y}}_{k+1,\ell+1}\}$是有界的.

定理2.1 如果当$\mu_{k,\ell}\to\infty$,$\sum_{k=1}^{+\infty}\mu_{k,\ell}^{-1}=+\infty$,$\left\langle\boldsymbol{A}_{k+1,\ell+1}-\boldsymbol{A}_{k,\ell},\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right\rangle\geqslant 0$,那么由算法1.4生成的矩阵序列$\{\boldsymbol{A}_{k,\ell}\}$收敛于矩阵填充问题(4)的最优解$\ddot{\boldsymbol{A}}$.

证明　由于 $\mu_{k,\ell}^{-1}\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right)=\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1}$,通过引理2.4可知 $\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right)$ 是有界的,所以

$$\lim_{k\to\infty}\left(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\bar{\boldsymbol{E}}_{k+1,\ell+1}\right)=$$
$$\lim_{k\to\infty}\mu_{k,\ell}^{-1}\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right)=0.$$

又 $\left(\ddot{\boldsymbol{A}},\ddot{\boldsymbol{E}}\right)$ 表示问题(4)的最优解,$\ddot{\boldsymbol{Y}}$ 为对偶问题(6)的最优解.由于 $\boldsymbol{A}_{k+1,\ell+1}$, $\bar{\boldsymbol{Y}}_{k+1,\ell+1}$, $\bar{\boldsymbol{E}}_{k+1,\ell+1}$, $k=1,2,\cdots$ 均为Toeplitz矩阵,且 $\ddot{\boldsymbol{A}}+\ddot{\boldsymbol{E}}=\boldsymbol{D}$,首先证明

$$\left\|\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\|_F^2+\mu_{k,\ell}^{-2}\left\|\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\|_F^2=$$
$$\left\|\boldsymbol{E}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\|_F^2-\left\|\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right\|_F^2+$$
$$\mu_{k,\ell}^{-2}\left\|\boldsymbol{Y}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\|_F^2-\mu_{k,\ell}^{-2}\left\|\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right\|_F^2-$$
$$2\mu_{k,\ell}^{-1}\left(\left\langle\boldsymbol{A}_{k+1,\ell+1}-\ddot{\boldsymbol{A}},\hat{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\rangle+\right.$$
$$\left\langle\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}},\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\rangle+$$
$$\left.\left\langle\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1},\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right\rangle\right),\quad(12)$$

式中 $\hat{\boldsymbol{Y}}_{k+1,\ell+1}=\boldsymbol{Y}_{k+1,\ell+1}+\mu_{k,\ell}\left(\boldsymbol{D}-\boldsymbol{A}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right)$.

$$\left\|\boldsymbol{E}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\|_F^2=\left\|P_{\bar{\Omega}}\left(\boldsymbol{E}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right)\right\|_F^2=$$
$$\left\|P_{\bar{\Omega}}\left(\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}-\bar{\boldsymbol{E}}_{k+1,\ell+1}+\boldsymbol{E}_{k+1,\ell+1}\right)\right\|_F^2=$$
$$\left\|P_{\bar{\Omega}}\left(\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right)\right\|_F^2+\left\|P_{\bar{\Omega}}\left(\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right)\right\|_F^2-$$
$$2\left\langle P_{\bar{\Omega}}\left(\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right),P_{\bar{\Omega}}\left(\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right)\right\rangle=$$
$$\left\|\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\|_F^2+\left\|\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1}\right\|_F^2+$$
$$2\mu_{k,\ell}^{-1}\left(\left\langle\bar{\boldsymbol{E}}_{k+1,\ell+1}-\boldsymbol{E}_{k+1,\ell+1},\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right\rangle\right).$$

同理,可得:

$$\mu_{k,\ell}^{-2}\left\|\boldsymbol{Y}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\|_F^2=\mu_{k,\ell}^{-2}\left\|P_{\bar{\Omega}}\left(\boldsymbol{Y}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right)\right\|_F^2=$$
$$\mu_{k,\ell}^{-2}\left\|P_{\bar{\Omega}}\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right)\right\|_F^2+$$
$$\mu_{k,\ell}^{-2}\left\|P_{\bar{\Omega}}\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right)\right\|_F^2-$$
$$2\mu_{k,\ell}^{-1}\left\langle P_{\bar{\Omega}}\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right),P_{\bar{\Omega}}\left(\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right)\right\rangle=$$
$$\mu_{k,\ell}^{-2}\left\|\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\|_F^2+\mu_{k,\ell}^{-2}\left\|\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\boldsymbol{Y}_{k+1,\ell+1}\right\|_F^2+$$
$$2\mu_{k,\ell}^{-1}\left(\left\langle\boldsymbol{A}_{k+1,\ell+1}-\ddot{\boldsymbol{A}},\hat{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\rangle+\right.$$
$$\left.\left\langle\bar{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}},\bar{\boldsymbol{E}}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\rangle\right).$$

所以式(12)成立.

再由 $\hat{\boldsymbol{Y}}_{k+1,\ell+1}\in\partial\|\boldsymbol{A}\|_*$, $\left\langle\boldsymbol{A}_{k+1,\ell+1}-\ddot{\boldsymbol{A}},\hat{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\rangle\geqslant 0$ 可知,$\sum_{k=1}^{\infty}\mu_{k,\ell}^{-1}\left\langle\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}},\hat{\boldsymbol{Y}}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\rangle$ 是有界的,且 $\left\|\boldsymbol{E}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\|^2+\mu_{k,\ell}^{-2}\left\|\boldsymbol{Y}_{k+1,\ell+1}-\ddot{\boldsymbol{Y}}\right\|^2$ 是递减的.另外,根据算法1.4有

$\left\langle\boldsymbol{Y}_{k+1,\ell+1},\boldsymbol{E}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\rangle=0$ 和 $\left\langle\ddot{\boldsymbol{Y}},\boldsymbol{E}_{k+1,\ell+1}-\ddot{\boldsymbol{E}}\right\rangle=0$.

因此,类似于文献[11]中定理2的证明,可以得到矩阵序列 $\left\{\boldsymbol{A}_{k,\ell}\right\}$ 收敛于(4)的最优解 $\ddot{\boldsymbol{A}}$.

定理2.2　在算法1.4中,由矩阵 $\boldsymbol{X}_{k,\ell}$ 得到Toeplitz矩阵 $\boldsymbol{A}_{k,\ell}$,则有下式成立

$$\left\|\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F<\left\|\boldsymbol{X}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F,$$

式中 $\ddot{\boldsymbol{A}}$ 是问题(4)的解.

证明　由于

$$\left\|\boldsymbol{X}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F^2=\left\|\boldsymbol{X}_{k,\ell}-\boldsymbol{A}_{k,\ell}+\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F^2=$$
$$\left\langle\boldsymbol{X}_{k,\ell}-\boldsymbol{A}_{k,\ell},\boldsymbol{X}_{k,\ell}-\boldsymbol{A}_{k,\ell}\right\rangle+2\left\langle\boldsymbol{X}_{k,\ell}-\boldsymbol{A}_{k,\ell},\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}}\right\rangle+$$
$$\left\langle\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}},\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}}\right\rangle=$$
$$\left\|\boldsymbol{X}_{k,\ell}-\boldsymbol{A}_{k,\ell}\right\|_F^2+\left\|\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F^2.$$

因此,有 $\left\|\boldsymbol{A}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F<\left\|\boldsymbol{X}_{k,\ell}-\ddot{\boldsymbol{A}}\right\|_F$.

由该定理可得,经Toeplitz结构化后的矩阵序列 $\left\{\boldsymbol{A}_{k,\ell}\right\}$ 仍收敛于问题的最优解 $\ddot{\boldsymbol{A}}$.

3　数值实验

通过数值实验比较了 ℓ- TMALM、ℓ- MALM[19]、MALM[15]以及ALM算法[11],所有的数值实验均在相同工作环境中完成.实验中,通过对迭代步数(IT)、计算时间(time,单位为s)、收敛误差1和2(E_1,E_2)的分析和比较,表明 ℓ- TMALM算法无论在迭代步数上还是在奇异值的分解时间上,明显优于其他3种算法.E_1 和 E_2 的计算式分别为

$$E_1=\frac{\left\|\boldsymbol{A}+\boldsymbol{E}-\boldsymbol{D}\right\|_F}{\left\|\boldsymbol{D}\right\|_F},\ E_2=\frac{\left\|\boldsymbol{A}-\boldsymbol{M}\right\|_F}{\left\|\boldsymbol{M}\right\|_F}.$$

在数值实验中,$\boldsymbol{M}\in\mathbf{R}^{n\times n}$ 是Toeplitz矩阵($\boldsymbol{M}$ 是原始矩阵).在 ℓ- TMALM算法中,选取参数 $\mu=\frac{1}{\|\boldsymbol{D}\|_2}$, $\rho=1.217\,2+\frac{1.858\,8m}{n^2}$, $\varepsilon_1=10^{-9}$, $\varepsilon_2=2\times10^{-6}$, Ω 是目标集,其他3种算法参数的选取与 ℓ- TMALM算法一致.

选取的样本矩阵阶数为 $m=n\in\{800,1\,500,2\,000,3\,000,4\,000\}$,经过多次数据实验得出当第1阶段迭代步数 $\ell=6$ 时,计算出的总迭代时间以及误差均

可达到最小,所以第1阶段ALM算法步中的迭代步数ℓ取值为6,采样率为$p\in\{0.6,0.5,0.4\}$,样本矩阵的秩为10.

4种算法的数值实验结果列于表1,4种算法均可成功地计算出满足规定停止准则的迭代矩阵.且$\ell-$ TMALM算法在保证迭代步数和奇异值分解时间上均优于$\ell-$ MALM,MALM以及ALM算法的情况下,其迭代矩阵与原始矩阵的逼近程度最好,即E_2最小,收敛效果最好.同一采样率下,4种算法的时间比较如图1所示.表明同一矩阵阶数下,$\ell-$ TMALM算法所用计算时间最短.

表1　不同采样率时4种算法比较

算法	矩阵阶数/10^2	p=0.6			p=0.5			p=0.4		
		迭代步数	计算时间/s	收敛误差2	迭代步数	计算时间/s	收敛误差2	迭代步数	计算时间/s	收敛误差2
ALM	8	70	20.202 5	0.012 3	74	20.351 8	0.017 0	71	18.128 3	0.014 4
	15	475	249.851 2	0.002 6	455	390.588 1	0.003 1	464	300.049 1	0.006 2
	20	102	58.552 5	$7.960\,6\times10^{-4}$	110	65.376 2	0.001 3	153	130.840 4	0.001 5
	30	328	564.296 7	0.002 0	320	617.982 9	0.003 4	350	1 024.500 0	0.003 5
	40	140	292.150 6	$6.842\,8\times10^{-4}$	184	403.136 3	0.003 1	165	358.829 0	0.002 2
MALM	8	73	9.520 0	$2.311\,0\times10^{-7}$	125	20.361 7	$1.481\,2\times10^{-5}$	92	13.595 3	$1.587\,5\times10^{-7}$
	15	321	115.256 3	$5.616\,3\times10^{-7}$	294	129.428 3	$2.068\,4\times10^{-7}$	478	223.391 3	$1.247\,5\times10^{-6}$
	20	95	60.790 0	$2.766\,2\times10^{-8}$	169	99.179 7	$9.901\,5\times10^{-7}$	243	166.270 5	$1.731\,0\times10^{-7}$
	30	177	217.453 6	$8.106\,5\times10^{-8}$	195	264.675 7	$2.004\,4\times10^{-7}$	240	325.554 3	$5.256\,5\times10^{-8}$
	40	162	336.013 0	$7.015\,5\times10^{-8}$	260	567.684 3	$4.896\,1\times10^{-7}$	187	415.635 1	$9.410\,1\times10^{-8}$
$\ell-$ MALM	8	67	8.266 0	$2.698\,2\times10^{-7}$	127	19.869 0	$1.511\,4\times10^{-5}$	90	13.008 3	$1.561\,4\times10^{-7}$
	15	276	99.529 8	$5.353\,4\times10^{-7}$	273	122.954 7	$1.964\,1\times10^{-7}$	461	288.672 6	$1.230\,4\times10^{-6}$
	20	94	58.835 1	$2.604\,5\times10^{-8}$	166	94.522 3	$9.805\,3\times10^{-7}$	242	161.345 5	$1.722\,4\times10^{-7}$
	30	147	177.028 9	$6.849\,4\times10^{-8}$	170	226.104 0	$1.410\,8\times10^{-7}$	174	233.566 5	$5.326\,1\times10^{-8}$
	40	162	332.037 9	$5.036\,9\times10^{-8}$	257	557.230 7	$4.843\,3\times10^{-7}$	183	394.201 6	$1.102\,8\times10^{-7}$
ℓ-TMALM	8	55	7.889 3	$2.313\,1\times10^{-8}$	67	13.091 6	$7.691\,7\times10^{-6}$	57	10.549 0	$7.287\,3\times10^{-8}$
	15	175	87.423 8	$1.544\,3\times10^{-8}$	116	65.377 1	$7.996\,6\times10^{-9}$	312	207.456 6	$3.255\,0\times10^{-7}$
	20	70	51.829 1	$1.898\,3\times10^{-9}$	83	65.100 7	$4.011\,0\times10^{-7}$	101	91.781 6	$6.180\,5\times10^{-8}$
	30	86	138.467 0	$2.658\,2\times10^{-8}$	102	180.624 9	$1.704\,7\times10^{-8}$	97	181.079 7	$2.764\,1\times10^{-9}$
	40	85	277.835 4	$2.573\,1\times10^{-9}$	108	360.397 7	$3.515\,7\times10^{-8}$	95	322.609 0	$6.042\,1\times10^{-9}$

注:ALM为增广拉格朗日乘子;MALM为基于均值的增广拉格朗日乘子;ℓ-MALM为ℓ步修正的增广拉格朗日乘子;ℓ-TMALM为Toeplitz矩阵填充的尾端修正的增广拉格朗日乘子.

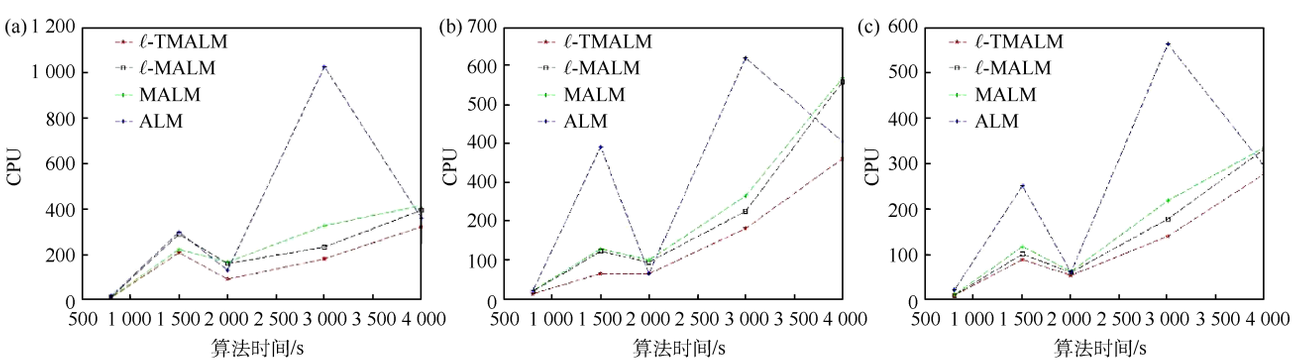


图1　不同采样率时4种算法时间比较

(a)p=0.4;(b)p=0.5;(c)p=0.6

4　结束语

对于Toeplitz矩阵填充问题,本文基于MALM提出了一种ℓ- TMALM算法.算法通过尾部修正再结构化,加速迭代矩阵与目标矩阵的逼近过程,以便快速地达到可行解,较大程度地缩短迭代步数;结构化使迭代矩阵与目标矩阵保持原有的Toeplitz结构,从而可利用快速奇异值分解缩短算法运行时间,降低计算代价.因此,新算法不仅克服了原始ALM算法较长的奇异值分解时间,也避免了MALM算法额外的数据传输量;此外,新算法的收敛效果要优于ℓ- MALM,MALM以及ALM算法.基于新算法的合理性和有效性,建立了与之相对应的收敛性理论.理论分析和数值结果表明,ℓ- TMALM算法是求解矩阵填充问题的一种有效算法.另外,未来拟将研究推广到高维的张量填充问题,比如将其修正的思想进一步推广到张量填充问题,从而求解出对于具有Toeplitz结构矩阵的张量填充算法.

参 考 文 献

[1] 史劲.人工智能领域的机器学习算法研究[J].中国新通信,2021,23(24):48-49.

[2] ARGYRIOU A, EVGENIOU T, PONTIL M. Multi-task feature learning [J]. Advances in Neural Information Processing Systems,2007,19:41-48.

[3] 潘伟.矩阵填充在推荐系统中的应用研究[D].赣州:江西理工大学,2021.

[4] BERTALMIO M, SAPIRO G, CASELLES V, et al. Multi-task feature learing, image inpainting [J]. Computer Graphics,2000,34:417-424.

[5] MESBAHI M, PAPAVASSILOPOULOS G P. On the rank minimization problem over a positive semidefinite linear matrix inequality [J]. IEEE Transactions on Automatic Control,1997,42(2):239-243.

[6] TOMASI C,KANADE T. Shape and motion from image streams under orthography: a factorization method [J]. International Journal of Computer Vision, 1992, 9(2): 137-154.

[7] CANDÈS E J, RECHT B. Exact matrix completion via convex optimization [J]. Foundations of Computational Mathematics,2009,9(6):717-772.

[8] CAI J F, CANDÈS E J, SHEN Z W. A singular value thresholding method for matrix completion [J]. Society for Industrial and Applied Mathematics Journal of Optimization,2010,20(4):1956-1982.

[9] LI W, HU J, CHEN C. On accelerated singular value thresholding algorithm for matrix completion [J]. Applied Mathematics,2014,5(21):3445-3451.

[10] TOH K C, YUN S. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems [J]. Pacific Journal of Optimization, 2010,6(3):615-640.

[11] LIN Z, CHEN M, MA Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices[R/OL]. (2010-09-26)[2021-03-02]. http://arxiv.org/abs/1009.5055.

[12] 方云兰,郑慧娆.TOEPLITZ-块矩阵的快速QR分解算法及其并行实现[J].武汉大学学报(自然科学版),1995,41(3):287-293.

[13] 于益华,李云翔.Toeplitz矩阵的快速小波变换与性能分析[J].湖南城市学院学报(自然科学版),2004,13(2):37-40.

[14] WANG C L, LI C. A mean value algorithm for Toeplitz matrix completion [J]. Applied Mathematics Letters, 2015,41:35-40

[15] WANG C L, LI C, WANG J. A modified augmented larange multiplier algorithm for toeplitz matrix completion[J]. Advances in Computational Mathematics, 2016,42:1209-1224.

[16] 王川龙,李超.Toeplitz矩阵填充的保结构算法[J].中国科学(数学),2016,46:1-16.

[17] 刘丽霞,王川龙.基于均值的Toeplitz矩阵填充的子空间算法[J].计算数学,2017,39(2):179-188.

[18] WEN R P, LI S Z, ZHOU F. Toeplitz matrix completion via smoothing augmented Lagrange multiplier algorithm [J]. Applied Mathematics and Computation, 2019, 355: 299-310.

[19] WEN R P, LI S Z. An ℓ-step modified augmented lagrange multiplier algorithm for completing a Toeplitz matrix [J]. Applied Mathematics and Computation, 2019,32(4):887-899.

[20] GOLUB G H, VAN LOAN C F. Matrix computations [M]. 4th ed. Baltimore: the Johns Hopkins University Press,1996,47:392-396.

[21] CHEN M M, GANESH A, LIN Z C, et al. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix [J]. Marine Biological Association of the UK,2009,56(3):707-722.

[22] 温瑞萍,王俊霞.应用矩阵分析[M].太原:山西人民出版社,2018:177-179.

(责任编辑:马田田)