

利用交替方向法求解 H 权重最近相关系数矩阵问题

郭来鹏¹, 顾剑^{*2}, 肖现涛¹

(1. 大连理工大学 数学科学学院, 辽宁 大连 116024;

2. 大连海洋大学 理学院, 辽宁 大连 116023)

摘要: 由于计算 H 权重的半正定矩阵锥投影比较困难, 目前求解带有 H 权重的最近相关系数矩阵问题的方法很少且比较复杂. 考虑用交替方向法求解该问题, 每次迭代只需求解一个有显式解的二次规划问题和一个不带权重的半正定矩阵锥投影, 计算简单, 易于实现. 为提高计算速度, 还考虑了改进的交替方向法. 此外, 通过数值实验对交替方向法与现有方法进行了比较, 说明了交替方向法对解决带有 H 权重的最近相关系数矩阵问题的有效性.

关键词: 交替方向法; 最近相关系数矩阵; 半正定矩阵锥

中图分类号: O221.1

文献标识码: A

doi: 10.7511/dllgxb201505017

0 引言

相关系数矩阵在金融分析和风险管理中起着至关重要的作用, 它是对角线元素全为 1 的对称半正定矩阵. 由历史数据得到的相关系数矩阵往往是不准确的. 因此为了方便金融分析, 就需要计算出它的最邻近的相关系数矩阵. 令 S^n 表示 $n \times n$ 阶实对称矩阵构成的空间, 记 S_+^n 为 S^n 中所有的半正定矩阵构成的锥. 在本文, 一般 $\|\cdot\|$ 表示 Frobenius 范数, 即 $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$, $\langle \cdot, \cdot \rangle$ 表示两个矩阵的内积. 令“ $\circ$ ”表示矩阵的 Hadamard 积, 即 $(A \circ B)_{ij} = A_{ij} B_{ij}$. 在文献[1]中, 作者研究了如下的带有 H 权重的最近相关系数矩阵 (HNCM) 问题:

$$\begin{aligned} \min \quad & \frac{1}{2} \|H \circ (X - G)\|^2 \\ \text{s. t.} \quad & X_{ii} = 1, i = 1, \dots, n \\ & X \in S_+^n \end{aligned}$$

其中矩阵 $H \in S^n$, 为权重矩阵且其元素均为非负的; $G \in S^n$, 是已知矩阵. 在 HNCM 问题中, 当矩阵 H 的元素全为 1 时, 问题就转变成了文献[1]中的不加权重的最近相关系数矩阵 (NCM) 问题:

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - G\|^2 \\ \text{s. t.} \quad & X_{ii} = 1, i = 1, \dots, n \\ & X \in S_+^n \end{aligned}$$

解决 NCM 问题的方法有很多^[1-5]. 在以上方法中, 除了文献[3]中的非精确内点法, 其余方法均需要计算矩阵 $X \in S^n$ 在闭凸锥 S_+^n 上的投影 $\Pi_{S_+^n}(X)$. 这些方法之所以有效, 主要是因为投影 $\Pi_{S_+^n}(X)$ 的计算有显式的公式. 注意到 $\Pi_{S_+^n}(X)$ 是下面优化问题的最优解:

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - G\|^2 \\ \text{s. t.} \quad & X \in S_+^n \end{aligned}$$

对于 HNCM 问题, 由于没有一个显式的公式来计算在 H 权重下 $X \in S^n$ 在 S_+^n 上的投影, 记为 $\Pi_{S_+^n}^H(X)$, 所以用于解决 NCM 问题的方法不能用于解决 HNCM 问题. 其中 $\Pi_{S_+^n}^H(X)$ 是下面问题的最优解:

$$\begin{aligned} \min \quad & \frac{1}{2} \|H \circ (Y - X)\|^2 \\ \text{s. t.} \quad & Y \in S_+^n \end{aligned}$$

基于上述的原因, 在解决 HNCM 问题时, 需

要避免计算投影 $\Pi_{S_+^n}(\mathbf{X})$. 为此,在文献[6]中,作者研究了求解 HNCM 问题的增广拉格朗日方法. 增广拉格朗日方法每一次的迭代需要求解一个无约束优化问题,作者用非精确的半光滑牛顿法来求解该无约束优化问题. 但是,该方法有两个小的问题:第一,对其他领域的研究人员来说,求解该问题的增广拉格朗日方法比较复杂,实现起来有一定的难度;第二,无论从理论上还是数值实验上都可以看出,文献[6]中的方法具有局部快速收敛的特点,也就是说,当迭代进入最优解的附近,算法会很快地收敛到最优解(通常是两三次迭代). 但是,如果初始点离最优解较远时,想要进入最优解附近需要很多次迭代,该算法的迭代次数会明显地增加. 于是,如何快速地获得一个接近最优解的初始点,是该算法的一个重要问题.

基于上述考虑,本文利用交替方向法(ADM)来求解 HNCM 问题. 首先,交替方向法形式非常简单,本文中求解 $\mathbf{H}$ 权重的最近相关系数矩阵问题的交替方向法的每一次迭代中,要求解一个有显式解的二次规划问题和计算一个半正定矩阵锥投影,只需短短几行程序代码就可实现该算法,这对其他领域的研究人员是非常重要的. 其次,交替方向法作为一阶算法,当迭代进入最优解的附近时,收敛速度会非常缓慢,因此交替方向法要得到一个高精度的近似最优解是非常消耗计算时间的. 但是,交替方向法的优点是,从任意点出发可以较快地收敛到最优解的附近. 因此,可以用交替方向法为增广拉格朗日方法提供一个好的初始点,然后利用文献[6]中的方法快速地得到 HNCM 问题的一个高精度的近似最优解.

1 经典的交替方向法

这一部分将运用经典的交替方向法来求解 HNCM 问题. 令 $I_K(\cdot)$ 表示指示函数,即

$$I_K(x)=\begin{cases} 0; & \text{当 } x \in K \\ +\infty; & \text{否则} \end{cases}$$

HNCM 问题可以等价地写为

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{H}^\circ(\mathbf{X}-\mathbf{G})\|^2 + I_{S_c}(\mathbf{X}) + I_{S_+^n}(\mathbf{Y}) \\ \text{s. t.} \quad & \mathbf{X}-\mathbf{Y}=\mathbf{0} \end{aligned} \tag{1}$$

其中 $S_c := \{\mathbf{X} \in S^n \mid X_{ii}=1, i=1, \dots, n\}$. 问题(1)的增广拉格朗日函数为

$$\begin{aligned} L_c(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) := & \frac{1}{2} \|\mathbf{H}^\circ(\mathbf{X}-\mathbf{G})\|^2 + I_{S_c}(\mathbf{X}) + \\ & I_{S_+^n}(\mathbf{Y}) + \langle \mathbf{Z}, \mathbf{X}-\mathbf{Y} \rangle + \frac{c}{2} \|\mathbf{X}-\mathbf{Y}\|^2 \end{aligned} \tag{2}$$

其中 $c>0$ 是罚参数, $\mathbf{Z} \in S^n$ 是拉格朗日乘子. 运用经典的交替方向法求解问题(1),在第 k 次迭代需要计算

$$\begin{aligned} \mathbf{X}^{k+1} = & \arg \min_{\mathbf{X}} L_{c_k}(\mathbf{X}, \mathbf{Y}^k, \mathbf{Z}^k) \\ \mathbf{Y}^{k+1} = & \arg \min_{\mathbf{Y}} L_{c_k}(\mathbf{X}^{k+1}, \mathbf{Y}, \mathbf{Z}^k) \\ \mathbf{Z}^{k+1} = & \mathbf{Z}^k + c_k(\mathbf{X}^{k+1} - \mathbf{Y}^{k+1}) \end{aligned} \tag{3}$$

不难得到,求 $\mathbf{X}^{k+1}$ 有显式的解,而求 $\mathbf{Y}^{k+1}$ 等价于计算一个半正定矩阵锥的投影. 具体来说,上述公式等价于计算

$$\begin{aligned} X_{ij}^{k+1} = & \begin{cases} \frac{H_{ij}^2 G_{ij} + c_k Y_{ij}^k - Z_{ij}^k}{H_{ij}^2 + c_k}; & i \neq j \\ 1; & i = j \end{cases} \\ & i, j = 1, \dots, n \end{aligned} \tag{4}$$

$$\begin{aligned} \mathbf{Y}^{k+1} = & \Pi_{S_+^n} \left(\mathbf{X}^{k+1} + \frac{1}{c_k} \mathbf{Z}^k \right) \\ \mathbf{Z}^{k+1} = & \mathbf{Z}^k + c_k(\mathbf{X}^{k+1} - \mathbf{Y}^{k+1}) \end{aligned}$$

下面将给出交替方向法的算法步骤和收敛性定理. 算法的终止准则如下:

$$\max\{\|\mathbf{Y}^k - \mathbf{Y}^{k+1}\|_\infty, \|\mathbf{Z}^k - \mathbf{Z}^{k+1}\|_\infty\} \leq \epsilon \tag{5}$$

其中 $\epsilon>0$ 为算法的精度.

算法 1

步骤 1 设 $c_0>0$ 且给定 $\eta>1$. 任取 $\mathbf{X}^0 \in S^n$, 设 $\mathbf{Y}^0 \in S^n$ 和 $\mathbf{Z}^0 \in S_+^n$ 为初始迭代点. 令 $k := 0$.

步骤 2 根据式(4)计算 $\mathbf{X}^{k+1}$ 、 $\mathbf{Y}^{k+1}$ 和 $\mathbf{Z}^{k+1}$. 令 $c_{k+1} = c_k$ 或者 $c_{k+1} = \eta c_k$.

步骤 3 如果式(5)成立,终止迭代;否则,令 $k=k+1$ 并返回步骤 2.

下面的收敛性定理在许多文献中都可以查到.

定理 1 设 $(\mathbf{X}^k, \mathbf{Y}^k, \mathbf{Z}^k)$ 是由算法 1 产生的序列. 则序列 $(\mathbf{X}^k, \mathbf{Y}^k, \mathbf{Z}^k)$ 收敛到问题(1)的 KKT 点. 因此序列 $\mathbf{X}^k$ 收敛到问题 HNCM 的最优解.

证明 参考文献[7]中的附录 A.



2 改进的交替方向法

利用文献[8]中改进的交替方向法来求解 HNCM 问题. 在经典的交替方向法中, 假设第 k 步迭代序列 $(\mathbf{X}^k, \mathbf{Y}^k, \mathbf{Z}^k)$ 已知, 则由式(4)可以得到第 $k+1$ 步的迭代序列. 在这里记该序列为 $(\tilde{\mathbf{X}}^k, \tilde{\mathbf{Y}}^k, \tilde{\mathbf{Z}}^k)$, 然后按照下面的方法获取新的迭代公式:

$$\begin{aligned}\mathbf{X}^{k+1} &= \tilde{\mathbf{X}}^k \\ \mathbf{Y}^{k+1} &= \mathbf{Y}^k - \gamma \alpha_k^* (\mathbf{Y}^k - \tilde{\mathbf{Y}}^k) \\ \mathbf{Z}^{k+1} &= \mathbf{Z}^k - \gamma \alpha_k^* (\mathbf{Z}^k - \tilde{\mathbf{Z}}^k)\end{aligned}\quad (6)$$

其中 $\gamma \in (0, 2)$, 而

$$\alpha_k^* = 1 - \frac{\langle \mathbf{Y}^k - \tilde{\mathbf{Y}}^k, \mathbf{Z}^k - \tilde{\mathbf{Z}}^k \rangle}{c_k \|\mathbf{Y}^k - \tilde{\mathbf{Y}}^k\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \tilde{\mathbf{Z}}^k\|^2} \quad (7)$$

由文献[9]可知改进的交替方向法比经典交替方向法通常快 17%.

下面给出改进的交替方向法及其收敛性证明, 依然采取式(5)为终止准则.

算法 2

步骤 1 设 $c_0 > 0$ 且给定 $\eta > 1$. 任取 $\mathbf{X}^0 \in \mathbf{S}^n$, 设 $\mathbf{Y}^0 \in \mathbf{S}^n$ 和 $\mathbf{Z}^0 \in \mathbf{S}_+^n$ 为初始迭代点. 设 $k := 0$.

步骤 2 根据式(4)计算 $\tilde{\mathbf{X}}^k, \tilde{\mathbf{Y}}^k$ 和 $\tilde{\mathbf{Z}}^k$.

步骤 3 根据式(6)计算 $\mathbf{X}^{k+1}, \mathbf{Y}^{k+1}$ 和 $\mathbf{Z}^{k+1}$.

令 $c_{k+1} = c_k$ 或者 $c_{k+1} = \eta c_k$.

步骤 4 如果式(5)成立, 终止迭代; 否则, 由 $k+1$ 替代 k 并返回步骤 2.

定理 2 设 $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*)$ 是问题(1)的最优解, 序列 $(\tilde{\mathbf{X}}^k, \tilde{\mathbf{Y}}^k, \tilde{\mathbf{Z}}^k)$ 是由迭代公式(4)产生的, 序列 $(\mathbf{X}^{k+1}, \mathbf{Y}^{k+1}, \mathbf{Z}^{k+1})$ 是由算法 2 产生的. 则有

$$\begin{aligned}& c_k \|\mathbf{Y}^{k+1} - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^{k+1} - \mathbf{Z}^*\|^2 \leq \\ & c_k \|\mathbf{Y}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \mathbf{Z}^*\|^2 - \\ & \frac{\gamma}{2} (2 - \gamma) (c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^k\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^k\|^2) \quad (8)\end{aligned}$$

进而, 可以得到

$$\lim_{k \rightarrow \infty} (c_k \|\mathbf{Y}^k - \tilde{\mathbf{Y}}^k\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \tilde{\mathbf{Z}}^k\|^2) = 0 \quad (9)$$

证明 根据 $(\tilde{\mathbf{X}}^k, \tilde{\mathbf{Y}}^k, \tilde{\mathbf{Z}}^k)$ 的定义, 利用文献

[7]中附录 A 的证明方法可以得到下列结果:

$$\begin{aligned}& c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^*\|^2 \leq \\ & c_k \|\mathbf{Y}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \mathbf{Z}^*\|^2 - \\ & c_k \|\tilde{\mathbf{X}}^k - \tilde{\mathbf{Y}}^k\|^2 - c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^k\|^2\end{aligned}\quad (10)$$

和

$$\langle \mathbf{Y}^k - \tilde{\mathbf{Y}}^k, \mathbf{Z}^k - \tilde{\mathbf{Z}}^k \rangle \geq 0 \quad (11)$$

由式(6)可得

$$\begin{aligned}& c_k \|\mathbf{Y}^{k+1} - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^{k+1} - \mathbf{Z}^*\|^2 = \\ & c_k \|\mathbf{Y}^k - \gamma \alpha_k^* (\mathbf{Y}^k - \tilde{\mathbf{Y}}^k) - \mathbf{Y}^*\|^2 + \\ & \frac{1}{c_k} \|\mathbf{Z}^k - \gamma \alpha_k^* (\mathbf{Z}^k - \tilde{\mathbf{Z}}^k) - \mathbf{Z}^*\|^2 = \\ & (1 - \gamma \alpha_k^*)^2 (c_k \|\mathbf{Y}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \mathbf{Z}^*\|^2) + \\ & \gamma^2 (\alpha_k^*)^2 (c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^*\|^2) + \\ & \gamma \alpha_k^* (1 - \gamma \alpha_k^*) \left[2c_k (\mathbf{Y}^k - \mathbf{Y}^*)^T (\tilde{\mathbf{Y}}^k - \mathbf{Y}^*) + \right. \\ & \left. \frac{2}{c_k} (\mathbf{Z}^k - \mathbf{Z}^*)^T (\tilde{\mathbf{Z}}^k - \mathbf{Z}^*) \right] \quad (12)\end{aligned}$$

注意到

$$xy = \frac{1}{2} [x^2 + y^2 - (x - y)^2]; \quad \forall x, y \quad (13)$$

因此

$$\begin{aligned}& 2c_k (\mathbf{Y}^k - \mathbf{Y}^*)^T (\tilde{\mathbf{Y}}^k - \mathbf{Y}^*) + \frac{2}{c_k} (\mathbf{Z}^k - \mathbf{Z}^*)^T \times \\ & (\tilde{\mathbf{Z}}^k - \mathbf{Z}^*) = \\ & c_k \|\mathbf{Y}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \mathbf{Z}^*\|^2 + \\ & c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^*\|^2 - \\ & \left(c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^k\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^k\|^2 \right)\end{aligned}$$

把上面的式子代入式(12), 可以得到

$$\begin{aligned}& c_k \|\mathbf{Y}^{k+1} - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^{k+1} - \mathbf{Z}^*\|^2 = \\ & (1 - \gamma \alpha_k^*) \left(c_k \|\mathbf{Y}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\mathbf{Z}^k - \mathbf{Z}^*\|^2 \right) + \\ & \gamma \alpha_k^* \left(c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^*\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^*\|^2 \right) - \\ & \gamma \alpha_k^* (1 - \gamma \alpha_k^*) \left(c_k \|\tilde{\mathbf{Y}}^k - \mathbf{Y}^k\|^2 + \frac{1}{c_k} \|\tilde{\mathbf{Z}}^k - \mathbf{Z}^k\|^2 \right)\end{aligned}$$

利用式(10)和 $\tilde{\mathbf{Z}}^k$ 的定义可得

$$\begin{aligned} &c_k \|Y^{k+1} - Y^*\|^2 + \frac{1}{c_k} \|Z^{k+1} - Z^*\|^2 \leqslant \\ &c_k \|Y^k - Y^*\|^2 + \frac{1}{c_k} \|Z^k - Z^*\|^2 - \\ &\gamma \alpha_k^* (2 - \gamma \alpha_k^*) \left(c_k \|\tilde{Y}^k - Y^k\|^2 + \frac{1}{c_k} \|\tilde{Z}^k - Z^k\|^2 \right) = \\ &c_k \|Y^k - Y^*\|^2 + \frac{1}{c_k} \|Z^k - Z^*\|^2 - \\ &\gamma (2 - \gamma \alpha_k^*) \left(c_k \|\tilde{Y}^k - Y^k\|^2 + \frac{1}{c_k} \|\tilde{Z}^k - Z^k\|^2 - \right. \\ &\left. \langle Y^k - \tilde{Y}^k, Z^k - \tilde{Z}^k \rangle \right) \end{aligned}$$

再次利用式(13)可以得到

$$\begin{aligned} &\|\tilde{Y}^k - Y^k\|^2 + \frac{1}{c_k} \|\tilde{Z}^k - Z^k\|^2 - \langle Y^k - \tilde{Y}^k, Z^k - \tilde{Z}^k \rangle = \\ &\frac{1}{2} \left[c_k \|\tilde{Y}^k - Y^k\|^2 + \frac{1}{c_k} \|\tilde{Z}^k - Z^k\|^2 + \right. \\ &\left. \left\| \sqrt{c_k} (\tilde{Y}^k - Y^k) - \frac{1}{\sqrt{c_k}} (\tilde{Z}^k - Z^k) \right\|^2 \right] \end{aligned} \tag{14}$$

根据式(7)和(11),易得

$$\frac{1}{2} \leqslant \alpha_k^* \leqslant 1 \tag{15}$$

于是

$$\begin{aligned} &c_k \|Y^{k+1} - Y^*\|^2 + \frac{1}{c_k} \|Z^{k+1} - Z^*\|^2 \leqslant \\ &c_k \|Y^k - Y^*\|^2 + \frac{1}{c_k} \|Z^k - Z^*\|^2 - \\ &\frac{\gamma}{2} (2 - \gamma) \left[c_k \|\tilde{Y}^k - Y^k\|^2 + \frac{1}{c_k} \|\tilde{Z}^k - Z^k\|^2 + \right. \\ &\left. \left\| \sqrt{c_k} (\tilde{Y}^k - Y^k) - \frac{1}{\sqrt{c_k}} (\tilde{Z}^k - Z^k) \right\|^2 \right] \leqslant \\ &c_k \|Y^k - Y^*\|^2 + \frac{1}{c_k} \|Z^k - Z^*\|^2 - \\ &\frac{\gamma}{2} (2 - \gamma) \left(c_k \|\tilde{Y}^k - Y^k\|^2 + \frac{1}{c_k} \|\tilde{Z}^k - Z^k\|^2 \right) \end{aligned}$$

这样式(8)就得到了证明,下面证明式(9).由式(8)可知

$$\begin{aligned} &c_k \|Y^{k+1} - Y^*\|^2 + \frac{1}{c_k} \|Z^{k+1} - Z^*\|^2 \leqslant \\ &c_k \|Y^0 - Y^*\|^2 + \frac{1}{c_k} \|Z^0 - Z^*\|^2 - \\ &\sum_{i=0}^k \frac{\gamma}{2} (2 - \gamma) \left(c_k \|\tilde{Y}^i - Y^i\|^2 + \frac{1}{c_k} \|\tilde{Z}^i - Z^i\|^2 \right) \end{aligned}$$

因此

$$\sum_{i=0}^k \frac{\gamma}{2} (2 - \gamma) (c_k \|\tilde{Y}^i - Y^i\|^2 + \frac{1}{c_k} \|\tilde{Z}^i - Z^i\|^2) \leqslant$$

$$c_k \|Y^0 - Y^*\|^2 + \frac{1}{c_k} \|Z^0 - Z^*\|^2$$

所以对任意的 k ,可得

$$\sum_{i=0}^k \frac{\gamma}{2} (2 - \gamma) (c_k \|\tilde{Y}^i - Y^i\|^2 + \frac{1}{c_k} \|\tilde{Z}^i - Z^i\|^2)$$

是有界的.由此可得

$$\sum_{i=0}^k \frac{\gamma}{2} (2 - \gamma) \left(c_k \|\tilde{Y}^i - Y^i\|^2 + \frac{1}{c_k} \|\tilde{Z}^i - Z^i\|^2 \right) \leqslant \infty$$

所以

$$\lim_{k \rightarrow \infty} \left(c_k \|Y^k - \tilde{Y}^k\|^2 + \frac{1}{c_k} \|Z^k - \tilde{Z}^k\|^2 \right) = 0$$

□

基于上面的定理,可以得到下面的收敛定理.

定理 3 设 $(X^{k+1}, Y^{k+1}, Z^{k+1})$ 由算法 2 产生,其中 $\gamma \in (0, 2)$,且 α_k^* 由式(7)给出.则 $(X^{k+1}, Y^{k+1}, Z^{k+1})$ 收敛到式(1)的 KKT 点.因此, X^k 收敛到 HNCM 的最优解.

证明 参照文献[8]中定理 4.1 的证明.

□

3 数值实验

本文用 MATLAB R2012a 编写算法程序.在数值实验中,采用相同的初始点 $(X^0 = I_n, Y^0 = I_n, Z^0 = I_n)$ 和相同的参数 $c=15, \gamma=1.5$,并且都采用式(5)作为算法程序的终止准则.表 1、2 中的 n 表示矩阵的阶数,即问题规模; t 表示计算时间, K 表示迭代次数, T 表示算法精度. ADM 表示经典的交替方向法;EADM 表示改进的交替方向法;EADM-MEX 表示结合了特征值分解加速的交替方向法;ALDA 表示增广拉格朗日方法;Hybrid 表示 EADM 与 ALDA 两种算法的结合,即首先由交替方向法得到一个好的初始点,然后由增广拉格朗日方法从该初始点出发得到一个高精度的最优解.

例 1^[6,10] 矩阵 G^* 是来自于网站 RiskMetric.com 中的一个 387 阶的相关系数矩阵.对它进行扰动,设 $G = (1 - \alpha)G^* + \alpha E$,其中 $\alpha \in (0, 1)$. E 按照如下方法生成: $E = 2.0 * rand(n, n) - ones(n, n)$, $E = triu(E) + triu(E, 1)'$, $E = (E + E')/2$;置 $G_{ii} = 1, i = 1, 2, \dots, n$,权重矩阵 H 按照文献[6]中的方法生成.

例 2^[6] 设矩阵 G 在 MATLAB 中按照如下方法生成： $X=10. [-4; 4/(n-1):0]$ ， $G=gallery('randcorr', n * X/sum(X))$ ， $G=(1-\alpha)\times G+\alpha E$ ，其中 $\alpha\in(0,1)$ 。 E 按照如下方法生成： $E=2. 0 * rand(n, n) - ones(n, n)$ ， $E=triu(E)+triu(E,1)'$ ， $E=(E+E')/2$ ；置 $G_{ii}=1$ ， $i=1,2,\cdots,n$ 。权重矩阵 H 按照文献[6]中的方法生成。

例 3 矩阵 G 来自于股票市场的真实数据，并且已知数据 $G(1,6)=G(6,1)$ 是不准确的，该矩阵的特征值为 $-0.02,0.04,0.05,0.11,1.96,3.86$ 。

$$G=\begin{pmatrix} 1.00 & 0.98 & 0.94 & 0.92 & -0.04 & -0.10 \\ 0.98 & 1.00 & 0.95 & 0.92 & -0.07 & -0.06 \\ 0.94 & 0.95 & 1.00 & 0.95 & -0.06 & -0.05 \\ 0.92 & 0.92 & 0.95 & 1.00 & -0.13 & -0.12 \\ -0.04 & -0.07 & -0.06 & -0.13 & 1.00 & 0.98 \\ -0.10 & -0.06 & -0.05 & -0.12 & 0.98 & 1.00 \end{pmatrix}$$
$$H=\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0.001 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0.001 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

对上述 3 个算例进行数值实验，数值结果汇集在表 1~3 中。下面对数值结果进行一些简要的说明。

表 1 为算例 1 的数值实验结果，比较了 ADM 和 EADM 的计算效果。ADM 的程序非常简单，只需要十几行程序代码就可实现。从数值实验结果可以看出，两种方法都可以较快地得到一个中等精度的近似最优解，而 EADM 在计算速度上相对于经典的 ADM 有一定的提升。

表 1 例 1 的数值实验结果($T=10^{-4}$)

Tab. 1 Numerical experimental results of Example 1 ($T=10^{-4}$)

n	ADM		EADM	
	K	t/s	K	t/s
387	225	45.991	155	33.109

表 2 中列出的是两种方法分别求解算例 2 的数值结果，算例 2 是一个随机产生的数值算例。表中两种方法 EADM 和 EADM-MEX 都是改进的交替方向法(因此两种方法的迭代次数都是一样的)，唯一的区别是在计算半正定矩阵锥投影时用的特征值分解不同，EADM 采用 MATLAB 中自带的 eig 函数，而 EADM-MEX 中特征值分解采用的是分治算法(divide-and-conquer)，速度更快。如前文所述，求解 HNCM 问题的交替方向法的主要计算量都消耗在矩阵的特征值分解上面。从表 2 中可以看出，当采用更快的特征值分解方法时，算法的计算时间会显著地减少。

表 2 例 2 的数值实验结果($T=10^{-4}$)

Tab. 2 Numerical experimental results of Example 2 ($T=10^{-4}$)

n	α	EADM		EADM-MEX	
		K	t/s	K	t/s
100	0.05	104	1.134	104	0.516
300	0.05	88	8.470	88	3.935
500	0.05	56	20.627	56	6.544

表 3 通过求解一个小规模的数值算例比较了增广拉格朗日方法与结合 EADM 算法的增广拉格朗日方法的数值结果。表 3 中的数据 82+6 表示首先由交替方向法经过 82 次迭代得到一个好的初始点，然后再由增广拉格朗日方法经过 6 次迭代得到一个好的近似最优解。从数值结果可以看出，与交替方向法结合后的增广拉格朗日方法在速度上有明显的优势。

表 3 例 3 的数值实验结果($T=10^{-8}$)

Tab. 3 Numerical experimental results of Example 3 ($T=10^{-8}$)

算法	K	$R/10^{-9}$	t/s
Hybrid	82+6	9.45	0.015
ALDA	10	7.00	0.031

4 结 语

本文利用交替方向法求解带有 H 权重的最近相关系数矩阵(HNCM)问题。对于优化领域外的研究人员，该方法简单易懂，容易实现。另一方

面,可以利用该方法得到一个好的初始点,然后结合二阶方法得到 HNCM 问题的一个高精度的近似最优解.

参考文献:

- [1] Higham N J. Computing the nearest correlation matrix - a problem from finance [J]. **IMA Journal of Numerical Analysis**, 2002, **22**(3):329-343.
- [2] Borsdorf R, Higham N J. A preconditioned Newton algorithm for the nearest correlation matrix [J]. **IMA Journal of Numerical Analysis**, 2010, **30**(1): 94-107.
- [3] QI Hou-duo, SUN De-feng. A quadratically convergent Newton method for computing the nearest correlation matrix [J]. **SIAM Journal on Matrix Analysis and Applications**, 2006, **28**(2):360-385.
- [4] Toh K C. An inexact primal-dual path following algorithm for convex quadratic SDP [J]. **Mathematical Programming**, 2008, **112**(1): 221-254.
- [5] Malick J. A dual approach to semidefinite least-squares problems [J]. **SIAM Journal on Matrix**

Analysis and Applications, 2004, **26**(1):272-284.

- [6] QI Hou-duo, SUN De-feng. An augmented Lagrangian dual approach for the H -weighted nearest correlation matrix problem [J]. **IMA Journal of Numerical Analysis**, 2011, **31**(2):491-511.
- [7] Boyd S, Parikh N, Chu E, *et al.* Distributed optimization and statistical learning via the alternating direction method of multipliers [J]. **Foundations and Trends in Machine Learning**, 2010, **3**(1):1-122.
- [8] YE Cai-hong, YUAN Xiao-ming. A descent method for structured monotone variational inequalities [J]. **Optimization Methods and Software**, 2007, **22**(2): 329-338.
- [9] He B, Xu M, Yuan X. Solving large-scale least squares semidefinite programming by alternating direction methods [J]. **SIAM Journal on Matrix Analysis and Applications**, 2011, **32**(1):136-152.
- [10] LI Qing-na, QI Hou-duo. A sequential semismooth Newton method for the nearest low-rank correlation matrix problem [J]. **SIAM Journal on Optimization**, 2011, **21**(4):1641-1666.

ADM for H -weighted nearest correlation matrix problem

GUO Lai-peng¹, GU Jian^{*2}, XIAO Xian-tao¹

(1. School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China;

2. School of Sciences, Dalian Ocean University, Dalian 116023, China)

Abstract: There are few methods for solving H -weighted nearest correlation matrix (HNCM) problem in the literatures. The main reason is that it is difficult to obtain a computable formula for the projection of a matrix onto the positive semidefinite matrix cone under the H -weight. The alternating direction method (ADM) is applied for solving HNCM problems. At each iteration, a quadratic program with closed-form solution and a projection on positive semidefinite matrix cone without weight are calculated. The method is very simple and easy to be implemented. To speed up the calculation, a modified ADM is also studied. The numerical results show that compared with the existing methods ADM is very efficient for solving HNCM problem.

Key words: alternating direction method; the nearest correlation matrix; positive semidefinite matrix cone