Markov 链平稳分布的迭代解法

郭 广 报
(厦门大学数学科学学院, 福建 厦门 361005)

摘要: 近来, Marek 等第一次将 Schwarz 方法引入了奇异线性方程组的求解问题. 然而, 这种方法对于分裂阵和迭代阵的要求过于严格. 本文在此基础上, 利用 Drazin 逆给出了拟非负分裂的定义. 对 Markov 链分裂阵的要求由非负型分裂推广到拟非负型分裂, 证明了 Markov 链加性 Schwarz 迭代, 诱导分离及其粗网格校正的半收敛性, 扩充了 Schwarz 迭代方法的理论, 使这种方法更具实用性.

关键词: Schwarz 方法; Markov 链; 拟非负型分裂
中图分类号: O 211　　文献标识码: A　　文章编号: 0438-0479(2008)03-0308-04

近些年来, 求解奇异线性方程组的迭代法越来越受到人们的重视, 并为许多计算数学工作者研究, 尤其是针对 Markov 链的平稳概率向量的计算问题[1-4]. 然而, 由于矩阵的奇异性所带来的复杂性, 迭代法的许多问题都未能得到很好解决.

因为易于并行计算等原因, Schwarz 方法在科学计算和工程的很多领域得到了广泛应用[5-7]. 但是, Schwarz 方法仅局限于求解非奇异线性方程组. 近来, Marek 等[2] 第一次将 Schwarz 方法引入了奇异线性方程组的求解问题. 然而, 由于没有引入其它参数和定义等原因, 他们提出的方法对于分裂阵和迭代阵的要求过于严格.

本文在文献[2] 的基础上, 利用 Drazin 逆给出了拟非负分裂的定义, 对分裂阵的要求由非负型分裂推广到拟非负型分裂. 研究了 Markov 链的加性 Schwarz 迭代的半收敛性及其它性质, 使这种方法更具实用性.

1　预备知识

这里研究齐次、不可约 Markov 链的非平凡解, 即 $A = -Q^T$, $B = P^T$ 为列随机矩阵, $A = I - B$, 求奇异线性方程组

$$AX = 0 \tag{1}$$

的非平凡解问题.

下面给出几个相关定义.

分裂矩阵 B 为 n 阶方阵, 且 $B = M - N$, 其中 M 为非奇异矩阵.

(i) 若 $M^{-1} \geqslant 0$, $N \geqslant 0$ 则 B 为正则分裂阵,

(ii) 若 $M^{-1}N \geqslant 0$, 则 B 为非负型分裂阵,

(iii) 若 $M^{-1} \geqslant 0$, $M^{-1}N \geqslant 0$, 则 B 为弱正则分裂阵.

对于 Drazin 逆: n 阶方阵 A 和 X, 若

$$A^{k+1}X = A^kX,\ XAX = X,\ AX = XA,$$

则称 X 为 A 的 Drazin 逆, 即为 A^D.

对于 n 阶方阵 $T = M^{-1}N$, 令 $\sigma(T)$ 表示矩阵 T 的谱, $\rho(T)$ 为矩阵 T 的谱半径,

$$\nu(T) = \max\{|\lambda|, \lambda \in \sigma(T), \lambda \neq 1\}.$$

对于 n 阶方阵 B, M 和 N, 其中 M 为非奇异阵, $E = (I - T)(I - T)^D$. 分裂阵 $B = M - N$, 满足 $\mathrm{index}(I - T) \leqslant 1$, 即 $\mathrm{rank}(I - T)^2 = \mathrm{rank}(I - T)$.

(i) 若 $M^{-1} \geqslant 0$, $NE \geqslant 0$, 则 B 为拟正则分裂,

(ii) 若 $M^{-1}NE \geqslant 0$, 则 B 为拟非负型分裂,

(iii) 若 $M^{-1} \geqslant 0$, $M^{-1}NE \geqslant 0$, 则 B 为拟弱正则分裂.

2　Schwarz 迭代

令 V_i 是 n_i 维 $V = R^n$ 的子空间, $V = \bigcup_{i=1}^{p} V_i$, $i = 1, \cdots, p$, $\sum_{i=1}^{p} n_i > n$. 约束算子 $R_i = [I_i \mid 0]\pi_i$, 其中 I_i 是 R^{n_i} 的恒等阵, π_i 是置换阵, 拓展算子 R_i^T, 对角阵 $E_i = R_i^TR_i$, $i = 1, \cdots, p$. 记 $S_i = \{i_1, i_2, \cdots, i_n\}$ $(i_1 < i_2 < \cdots < i_n)$ 为 R_i 非零元素所在列的指示集, q 为交叉测度, 故 $\sum_{i=1}^{p} E_i \leqslant qI$. 子空间 V_i 对于 A 的约束是 $A_i = R_iAR_i^T$, A_i 是 A 的 $n_i \times n_i$ 主子阵的对称置换阵 $(n_i < n, i = 1, \cdots, p)$, 且 A_i 非奇异.

令 T 为一方阵, 平稳迭代的形式为

$$x^{(k+1)} = Tx^{(k)}, k = 0, 1, \cdots \tag{2}$$

收稿日期: 2007-05-16
* 通讯作者: ggb11111111@163.com

假如上面方程解的初始近似值为 $x^{(0)}$, 作为平稳迭代的一水平加性 Schwarz 法的迭代阵为

$$T = I - \sum_{i=1}^{p} P_i = I - \sum_{i=1}^{p} R_i^T A_i^{-1} R_i A \tag{3}$$

这里 $P_i = R_i^T (R_i A R_i^T)^{-1} R_i A$, 其中 R_i 是维数为 $n_i \times n$ 的行满秩阵, $1 \leqslant i \leqslant p$. 在这种交叠的情况下我们有 $\sum_{i=1}^{p} n_i > n$.

相应的同一水平乘性 Schwarz 法的迭代阵为

$$T = \prod_{i=p}^{1} P_i = (I - P_p)(I - P_{p-1}) \cdots (I - P_1) \tag{4}$$

在文献[2] 中, Mark 等在非负性分裂阵的情况下, 对上述方法进行了研究.

最后, 给出阻尼加性 Schwarz 迭代公式

$$x^{(k+1)} = x^{(k)} - \theta \sum_{i=1}^{p} R_i^T A^{-1} R_i A x^{(k)} = (I - \theta \sum_{i=1}^{p} R_i^T A^{-1} R_i A) x^{(k)},$$

$0 < \theta < 1$ 为阻尼因子, 则 $T_\theta = I - \theta \sum_{i=1}^{p} R_i^T A^{-1} R_i A$ 为阻尼加性 Schwarz 方法的迭代阵.

另一种表示方法: 若 $A_{-i} = [0 \mid I_{-i}]\pi_i A_i^T [0 \mid I_{-i}]^T$, 其中 I_{-i} 是 $(n - n_i) \times (n - n_i)$ 单位矩阵, 令 $M_i = \pi_i^T \begin{bmatrix} A_i & 0 \\ 0 & \mathrm{diag}(A_{-i}) \end{bmatrix} \pi_i$, 则

$$E_i M_i^{-1} = R_i^T R_i M_i^{-1} = \pi_i^T \begin{bmatrix} A_i^{-1} & 0 \\ 0 & 0 \end{bmatrix} \pi_i = R_i^T A_i^{-1} R_i,$$

这时 $T_\theta = I - \theta \sum_{i=1}^{p} E_i M_i^{-1} A$.

3　Markov 链的加性 Schwarz 的半收敛性

对于非奇异阵迭代方法收敛的充要条件是迭代阵的谱半径小于 1, 对于奇异阵线性方程组(1) 的迭代收敛称为半收敛. 文献[8] 给出了式(2) 半收敛的充要条件:

(i) $\rho(T) = 1$,

(ii) $\mathrm{rank}(I - T)^2 = \mathrm{rank}(I - T)$ 即 $\mathrm{index}(I - T) = 1$,

(iii) $\nu(T) = \max\{|\mu|, \mu \in \sigma(T), \mu \neq 1\} < 1$.

先给出如下引理:

引理 1[1]　若分裂 $A = M - N$ 满足 $\mathrm{index}(I - T) \leqslant 1$, 则

(i) $(I - TE)^{-1}$ 存在且 $(I - TE)^{-1} = (I - T)^D + I - E$,

(ii) 对于 $i, j \geqslant 0$, $E^i (TE) E^j = TE$.

定理 1　若 $A = M_i - N_i$, $i = 1, \cdots, p$ 为拟非负型分裂阵, 若

$$T_i = M_i^{-1} N_i, E_i = (I - T_i)(I - T_i)^D,$$

则 $T_i E_i = E_i T_i$, $i = 1, \cdots, p$.

证　由 $A = M_i - N_i$, $i = 1, \cdots, p$ 为拟非负型分裂, 故存在非奇异阵 G_i, 使得

$$T_i = G_i \begin{bmatrix} I & 0 \\ 0 & K_i \end{bmatrix} G_i^{-1},$$

$\rho(K_i) = \nu(T_i)$ 且 $1 \notin \sigma(K_i)$.

$$I - T_i = G_i \begin{bmatrix} I & 0 \\ 0 & I - K_i \end{bmatrix} G_i^{-1},$$

$$(I - T_i)^D = G_i \begin{bmatrix} I & 0 \\ 0 & (I - K_i)^{-1} \end{bmatrix} G_i^{-1}.$$

从而

$$E_i = (I - T_i)(I - T_i)^D = G_i \begin{bmatrix} I & 0 \\ 0 & I - K_i \end{bmatrix} G_i^{-1} \circ$$

$$G_i \begin{bmatrix} I & 0 \\ 0 & (I - K_i)^{-1} \end{bmatrix} G_i^{-1} = G_i \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} G_i^{-1},$$

且

$$T_i E_i = G_i \begin{bmatrix} 0 & 0 \\ 0 & K_i \end{bmatrix} G_i^{-1}, E_i T_i = G_i \begin{bmatrix} 0 & 0 \\ 0 & K_i \end{bmatrix} G_i^{-1}.$$

另外, 由引理 1(ii), 令 $i = 0, j = 1$ 也可得结论.

引理 2[8]　对于阵 T, 若 $\rho(T) = 1$, T 收敛的充要条件:

(i) $\lambda \in \sigma(T)$, $|\lambda| = 1$, 则 $\lambda = 1$, 即 $\nu(T) < 1$,

(ii) $\mathrm{rank}(I - T) = \mathrm{rank}(I - T)^2$.

引理 3[9]　当 $T \geqslant 0$ 时, $\lambda \in \sigma(T)$, $|\lambda| = 1$, 则 $\lambda = 1$, 即 $\nu(T) < 1$, 等价于 $\mathrm{diag}(T) \geqslant 0$.

引理 4[10]　令 T 为非负方阵, 使得对 $v > 0$, $Tv \leqslant v$, 则 $\rho(T) \leqslant 1$. 而且 $\rho(T) = 1$, 则 $\mathrm{rank}(I - T) = \mathrm{rank}(I - T)^2$.

引理 5[8]　若分离 $A = M - N$ 满足 $\mathrm{index}(I - T) \leqslant 1$, 则存在非奇异阵 G, 使得 $T = G \begin{bmatrix} I & 0 \\ 0 & K \end{bmatrix} G^{-1}$, 这里 $\rho(K) = \nu(T)$ 和 $1 \notin \sigma(K)$. 而且 T 半收敛的充要条件是 $\rho(K) < 1$.

引理 6[4]　若 $M_i = \pi_i^T \begin{bmatrix} A_i & 0 \\ 0 & \mathrm{diag}(A_{-1}) \end{bmatrix} \pi_i$, $A = M_i - N_i$, 则 $\mathrm{diag}(M_i^{-1} N_i) = 0$, $i = 1, 2, \cdots, p$, 从而 $\mathrm{diag}(\sum_{n=0}^{p} M_i^{-1} N_i) = 0$.

定理 2　若 $A = I - B$, $B \in R^{n \times n}$ 是不可约列随机阵, 令 $p > 1$ 为正整数, $A = M_i - N_i$ 为拟非负型分裂

且$\sum_{i=1}^{p} E_i T_i \geqslant \sum_{i=1}^{p} \bar{E}_i T_i$(或$E_i T_i \geqslant \bar{E}_i T_i$),$i = 1, 2, \cdots, p$.当$\theta < \frac{1}{q}$时,则加性 Schwarz 迭代阵$T_\theta$半收敛.

证 由半收敛充要条件及引理 3 ~ 6,知T_θ半收敛的条件是:

(i) 存在$\nu > 0$,$T_\theta \nu = \nu$,(ii)$T_\theta \geqslant 0$,(iii)$\mathrm{diag}(T_\theta) \geqslant 0$.

首先,$B \in R^{n\times n}$是不可约列随机阵,则有一正的 Perron 根ν,即$\nu^T e = 1$,得$B\nu = \nu$,$A\nu = 0$,所以$T_\theta \nu = \nu - \theta \sum_{i=1}^{p} R_i^T A^{-1} R_i A \nu = \nu$.

其次,由定理 1 得

$$T_\theta = I - \theta \sum_{i=1}^{p} R_i^T A^{-1} R_i A = I - \theta \sum_{i=1}^{p} E_i M_i^{-1} A \geqslant \frac{1}{q} \sum_{i=1}^{p} E_i - \theta \sum_{i=1}^{p} E_i M_i^{-1} A \geqslant \theta \sum_{i=1}^{p} E_i (I - M_i^{-1} A) \geqslant \theta \sum_{i=1}^{p} E_i T_i \geqslant \theta \sum_{i=1}^{p} E_i T_i \geqslant \theta \sum_{i=1}^{p} T_i E_i \geqslant 0.$$

最后,取$\theta < \frac{1}{q}$,有$\mathrm{diag}(I) \geqslant \mathrm{diag}(\theta \sum_{i=1}^{p} E_i)$.

因而

$$\mathrm{diag}(T_\theta) = \mathrm{diag}(I - \theta \sum_{i=1}^{p} E_i M_i^{-1} A) \geqslant \mathrm{diag}(\theta \sum_{i=1}^{p} E_i - \theta \sum_{i=1}^{p} E_i M_i^{-1} A) \geqslant \mathrm{diag}(\sum_{i=1}^{p} E_i M_i^{-1} N_i) = 0.$$

4 诱导分离和粗网格校正

4.1 诱导分离

诱导分离的存在和半收敛性的研究是非常必要的,下面给出了在拟非负型分裂情况下的半收敛性.

定理 3 若$A = I - B$,$B \in R^{n\times n}$是不可约列随机阵,令$p > 1$是正整数,令$G = \sum_{i=0}^{p} E_i M_i^{-1}$为非奇异矩阵,则存在拟非负型分裂$A = M - N$使迭代阵为$T_\theta = M^{-1} N$,这里$M^{-1} = \theta G$,且对于迭代阵$T_\theta$半收敛.

证 令矩阵$M = (\frac{1}{\theta}) G^{-1}$,取矩阵$N = A - M$,则$T_\theta = I - M^{-1} A$.

因为$T_\theta E_\theta \geqslant 0$($E_\theta = (I - T_\theta)(I - T_\theta)^D$),由 Perron-Frobenius 定理[8],存在向量$x \geqslant 0$,且$x \neq 0$,使

$$T_\theta E_\theta x = \rho(T_\theta E_\theta) x, \rho(T_\theta E_\theta) \neq 0.$$

显然$E_\theta x \neq 0$.故

$$T_\theta (E_\theta x) = T_\theta E_\theta^2 x = E_\theta T_\theta E_\theta x = \rho(T_\theta E_\theta)(E_\theta x),$$

又$\rho(T_\theta E_\theta) \in \nu(T_\theta)$,得$\rho(T_\theta E_\theta) < 1$.

由定理 1 和引理 5 知迭代阵T_θ半收敛.

4.2 粗网格校正

我们可以选择一个附加子空间V_0,约束算子$R_0 \geqslant 0$,$A_0 = R_0 A R_0^T$是A的一个主子式.$A_0^{-1} \geqslant 0$,拟非负型分裂$A = M_0 - N_0$,对角阵$0 \leqslant E_0 \leqslant I$,有$R_0^T A_0^{-1} R_0 = E_0 M_0^{-1}$,令$\bar{T}_\theta = T_\theta - \theta R_0^T A_0^{-1} R_0 A = I - \theta \sum_{i=1}^{p} R_i^T A^{-1} R_i A$即为全解迭代阵.$\bar{T}_\theta$和$T_\theta$收敛速度至少是一样的.

定理 4 若$A = I - B$,$B \in R^{n\times n}$是不可约列随机阵,令$p > 1$是正整数,$A = M_i - N_i$是拟非负型分裂,$i = 1, 2, \cdots, p$,$\theta < \frac{1}{q}$.则如果$\theta < \frac{1}{q+1}$时,加性 Schwarz 迭代阵$\bar{T}_\theta$半收敛.

证 由引理 2 ~ 6,知$\bar{T}_\theta$半收敛的条件是

(i) 对$\nu > 0$,$\bar{T}_\theta \nu = \nu$,(ii)$\theta \geqslant 0$,(iii)$\mathrm{diag}(\bar{T}_\theta) \geqslant 0$.

首先,$B \in R^{n\times n}$是不可约列随机阵,则有一正的 Perron 根ν,即$\nu^T e = 1$,得$B\nu = \nu$,$A\nu = 0$,所以$\bar{T}_\theta \nu = \nu - \theta \sum_{i=0}^{p} R_i^T A^{-1} R_i A \nu = \nu$.

其次,

$$\bar{T}_\theta = I - \theta \sum_{i=0}^{p} R_i^T A^{-1} R_i A = I - \theta \sum_{i=0}^{p} E_i M_i^{-1} A \geqslant \frac{1}{q+1} \sum_{i=0}^{p} E_i - \theta \sum_{i=0}^{p} E_i M_i^{-1} A \geqslant \theta \sum_{i=0}^{p} E_i (I - M_i^{-1} A) \geqslant \theta \sum_{i=0}^{p} E_i T_i \geqslant \theta \sum_{i=0}^{p} E_i T_i \geqslant \theta \sum_{i=0}^{p} T_i E_i \geqslant 0.$$

最后,取$\theta < \frac{1}{q+1}$,有$\mathrm{diag}(I) \geqslant \mathrm{diag}(\theta \sum_{i=0}^{p} E_i)$,因而

$$\mathrm{diag}(\bar{T}_\theta) = \mathrm{diag}(I - \theta \sum_{i=0}^{p} E_i M_i^{-1} A) \geqslant \mathrm{diag}(\theta \sum_{i=0}^{p} E_i - \theta \sum_{i=0}^{p} E_i M_i^{-1} A) \geqslant \mathrm{diag}(\sum_{i=0}^{p} E_i M_i^{-1} N_i) = 0.$$

定理 5[比较定理] 如果有$\omega > 0$,使$A\omega = e > 0$,则$\bar{T}_\theta \omega \leqslant T_\theta \omega$.

证 因为

$$\bar{M}_\theta^{-1} = \theta \sum_{i=0}^{p} R_i^T A^{-1} R_i = \theta R_0^T A_0^{-1} R_0 + M_\theta^{-1} \geqslant$$

$$M_\theta^{-1} \geqslant 0.$$

故

$$T_\theta\omega = M_\theta^{-1}N_\theta\omega \leqslant \omega - M_\theta^{-1}A\omega = \omega - M_\theta^{-1}e \leqslant \omega - M_\theta^{-1}A\omega = T_\theta\omega.$$

参考文献:

[1] Song Yongzhong. Semiconvergence of nonnegative splittings for singular matrices[J]. Numer Math, 2000, 85: 109—127.

[2] Marek I, Szyld D B. Algebraic Schwarz methods for the numerical solution of Markov chains[J]. Linear Algebra Appl, 2004, 386: 67—81.

[3] Bru R, Pedroche F, Szyld D B. Iterations for Markov chains, society for industrial and applied mathematics[J]. Society for Industrial and Applied Mathematics, 2005, 27(2): 445—458.

[4] Philippe B, Saad Y, Stewart W J. Numerical methods in Markov chain modelling[J]. Operat Research, 1992, 40: 1156—1179.

[5] Frommer A, Szyld D B. Weighted max norms, splittings, and overlapping additive Schwarz iterations[J]. Numerische Mathematik, 1999, 83: 259—278.

[6] Benzi M, Frommer A, Nabben R, et al. Algebraic theory of multiplicative Schwarz methods[J]. Numerische Mathematik, 2001, 89: 605—639.

[7] Nabben R. Comparisons between additive and multiplicative Schwarz iterations in domain decomposition methods[J]. Numer Math, 2003, 95: 145—162.

[8] Berman A, Plemmons R J. Nonnegative matrices in the mathematical sciences[M]. New York: Academic Press, 1979.

[9] Chan T F, Mathew T P. Domain decomposition algorithms[J]. Acta Numer, 1994, 3: 61—143.

[10] Szyld D B. Equivalence of convergence conditions for iterative methods for singular equations[J]. Numer Linear Algebra Appl, 1994, 1: 151—154.

Iterative Methods for Solving the Stationary Distribution of Markov Chains

GUO Guang-bao

(School of Mathematical Sciences, Xiamen University, Xiamen 361005, China)

Abstract: Up to now, singular systems are analyzed using Schwarz methods by Marek, and it is the first time that Markov chains problems are studied in that context. In this paper, the author gives the definition of quasi-nonnegative splittings. Splitting matrixs vary from nonnegative splittings to quasi-nonnegative splittings. It is shown that the semiconvergence of the additive Schwarz method, induced splitting and corse grid corrections in the case.

Key words: Schwarz method; Markov chains; quasi-nonnegative splittings