第15卷 第1期　华侨大学学报（自然科学版）　Vol.15 No.1
1994年1月　Journal of Huaqiao University (Natural Science)　Jan. 1994

2-块AOR迭代解最小二乘问题的最优收敛性*

曾 文 平

（华侨大学管理信息科学系，泉州 362011）

摘要 讨论用2-块AOR迭代法解大型稀疏最小二乘问题的收敛性，给出其收敛的充要条件及其收敛域．进而证明：当$\beta=\bar{\beta}$时，AOR迭代矩阵的谱半径$\rho(L_{\gamma,\omega}^{(2)})=0$，它远比相应的最优2-块AOR迭代矩阵的谱半径好得多．

关键词 最小二乘问题，2-块AOR迭代，收敛性

分类号 O 241.2

诸如大地测量等很多实际问题，最终都导致用最小二乘法求解超定方程组$Ax=b$(式A)，其中，$A\in R^{m\times n}(m>n)$，$b\in R^{n}$，且设Rnak(A)$=n$. 熟知，式(A)的最小二乘解不止1个，但其中，满足欧氏范数最小的最小二乘问题为$\|b-Ax\|_2=\min\limits_{y\in R^n}\|b-Ay\|_2$. 它满足法方程组$A^TAx=A^Tb$. 众所周知，上述最小二乘问题等价于求$x\in R^n$和$\gamma\in R^m$，使得$\gamma+Ax=b$和$A^T\gamma=0$(式B). 1975年，Chen和Gentheman首先建议用3-块SOR迭代解最小二乘问题．最近，Markham, Neumann和Plemmons[1]提出用2-块SOR方法解最小二乘问题．利用Young[2]的结果，给出用2块SOR方法解最小二乘问题时的收敛域，并给出最佳松驰因子及最小谱半径的表达式，同时进一步指出用2-块最优SOR方法求解最小二乘问题，比用3-块最优SOR方法求解最小二乘问题的谱半径小．本文目的在于讨论应用2-块AOR迭代求解最小二乘问题的收敛性，给出其收敛的充要条件(对任意Jacobi阵J_2的谱半径$\rho(J_2)$都可以)．同时，证明当μ接近$\bar{\mu}$时，可以找到使AOR迭代的谱半径$\rho(L_{\gamma,\omega})=0$的$\omega,\gamma$值，此时AOR方法显然优于最优SOR方法．

1 AOR迭代矩阵和Jacobi迭代矩阵特征值之间的关系

因为A为列满秩，总可以通过适当的排列使A,γ和b，具有如下的分块形状$A=\begin{pmatrix}A_1\\A_2\end{pmatrix}$，其中，$A_1$是$n\times n$的非奇异矩阵，$\gamma_1,b_1$为$n$维向量，而$\gamma_2,b_2$为$(m-n)$维向量．$\gamma=\begin{pmatrix}\gamma_1\\\gamma_2\end{pmatrix}$，$b=$

* 本文1993-10-13收到；福建省自然科学基金资助项目

$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. 于是,式(B)可表示为如下等价形式

$$CZ = d, \tag{1}$$

其中

$$C = \begin{bmatrix} A_1 & O & I \\ A_2 & I & O \\ O & A_2^T & A_1^T \end{bmatrix}, \quad Z = \begin{bmatrix} X \\ \gamma_1 \\ \gamma_2 \end{bmatrix}, \quad d = \begin{bmatrix} b_1 \\ b_2 \\ O \end{bmatrix}. \tag{2}$$

因为 A_1 非奇异,故 $(m+n)\times(m+n)$ 矩阵 C 也非奇异. 为了研究用 2-块 AOR 迭代法求解最小二乘问题(1)的解,我们把矩阵 C 划分成如下两块形式,并记

$$C \equiv C_2 \equiv \left[\begin{array}{cc:c} A_1 & O & I \\ A_2 & I & O \\ \hdashline O & A_2^T & A_1^T \end{array}\right], \quad D_2 \equiv \left[\begin{array}{cc:c} A_1 & O & O \\ A_2 & I & O \\ \hdashline O & O & A_1^T \end{array}\right],$$

$$L_2 \equiv \left[\begin{array}{cc:c} O & O & O \\ O & O & O \\ \hdashline O & -A_2^T & O \end{array}\right], \quad U_2 \equiv \left[\begin{array}{cc:c} O & O & -I \\ O & O & O \\ \hdashline O & O & O \end{array}\right]. \tag{3}$$

相应的 Jacobi 迭代矩阵则为

$$J_2 \equiv D_2^{-1}(L_2 + U_2) = \left[\begin{array}{cc:c} O & O & -A_1^{-1} \\ O & O & A_2A_1^{-1} \\ \hdashline O & -(A_2A_1^{-1})^T & O \end{array}\right]. \tag{4}$$

显知,J_2 是 1 个具有相容次序的 2-弱循环阵. 由式(4)可知,由于 $-(A_2A_1^{-1})^T(A_2A_1^{-1})$ 为非负定矩阵,所以

$$J_2^2 = \operatorname{diag}\{0,\ -(A_2A_1^{-1})(A_2A_1^{-1})^T,\ -(A_2A_1^{-1})^T(A_2A_1^{-1})\} \tag{5}$$

相似于 1 个实对称负半定矩阵,故 J_2 的全部特征值为纯虚数 $\mu^2 \leqslant 0$,且 $|\bar{\mu}| = \rho(J_2) = \| A_2A_1^{-1} \|_2$,从而 J_2^2 的全部特征值都落在区间 $[-\rho^2(J_2), 0]$ 内. 这一事实将对收敛性定理的证明起重要作用. 用 2-块 AOR 迭代求解最小二乘问题(1)时,其迭代格式为

$$Z^{(K+1)} = L_{\gamma,\omega}^{(2)} Z^{(K)} + (D_2 - \omega L_2)^{-1} d, \tag{6}$$

其中 AOR 迭代矩阵为

$$L_{\gamma,\omega}^{(2)} = (D_2 - \gamma L_2)^{-1}\{(1-\omega)D_2 + (\omega - \gamma)L_2 + \omega U_2\}, \tag{7}$$

这里 D_2, L_2, U_2 由式(3)所定义. 为了讨论用 2-块 AOR 迭代解(1)的收敛性,先给出解(1)的 AOR 迭代矩阵 $L_{\gamma,\omega}^{(2)}$ 以及相应的 Jacobi 迭代矩阵 J_2 的特征值之间的关系.

引理 1 设 A 为有相容次序的 2-循环阵,其对角线上块子矩阵 $A_{i,i}(i=1,2)$ 非奇异,若 $\omega \neq 0$ 且 $\lambda\gamma + \omega - \gamma \neq 0$,则 AOR 迭代矩阵 $L_{\gamma,\omega}$ 的特征值 λ 及其相伴的 Jacobi 迭代矩阵的特征值 μ 之间,满足如下函数关系式

$$(\lambda + \omega - 1)^2 = (\lambda\gamma + \omega - \gamma)\omega\mu^2. \tag{8}$$

证 注意到文[3]中所指出,AOR 迭代可看作 SOR 迭代的外推,则 $L_{\gamma,\omega}$ 的特征值 λ 与 $L_{\gamma,\gamma}$ 的特征值 υ 之间,有 $\lambda = s\upsilon + (1-s)$ 关系,其中 $s = \omega/\gamma, \gamma \neq 0$. 从中解出 $\upsilon = (\lambda + s - 1)/s$,代入文[4]中定理 4.3 的函数关系式(2.2)(注意此时 $\omega = \gamma, \lambda = \upsilon, p = 2$),即得

$$(\lambda+s\gamma-1)^2=s(\lambda+s-1)\gamma^2\mu^2,$$

再代入 $s=\omega/\gamma$，便得式(8). 证毕.

2　2-块 AOR 迭代收敛的充要条件

引理 2[2]　实系数二次方程 $x^2+px+r=0$ 的两根按模小于1的充要条件为 $|r|<1$，且 $|p|<1+r$.

定理 1　当 2-块 AOR 迭代(6)应用于解最小二乘问题(1)时，且定义参数 $\bar{\beta}=\rho(J_2)=\max\limits_{\mu\in\rho(J_2)} I_m(\mu)=\max|\beta|=\|A_2A_1^{-1}\|_2$ 及 $\underline{\beta}=\min|I_m(\mu)|=\min|\beta|$，则 2-块 AOR 迭代(9)收敛(即 $\rho(L_{\gamma,\omega}^{(2)})<1$)，当且仅当 ω,γ 分别取值 I_ω 及 I_γ 即

(i) $\underline{\beta}\neq0$，　$I_\omega\equiv(-\infty,0)$，　且 $I_\gamma\equiv(M(\bar{\beta}^2),N(\underline{\beta}^2))$，

或　$I_\omega\equiv(0,2)$，　且 $I_\gamma\equiv(N(\bar{\beta}^2),M(\bar{\beta}^2))$，

或　$I_\omega\equiv(2,+\infty)$，　且 $I_\gamma\equiv(N(\underline{\beta}^2),M(\bar{\beta}^2))$；

(ii) $\underline{\beta}=0$，　$I_\omega\equiv(0,2)$，　且 $I_\gamma\equiv(N(\bar{\beta}^2),M(\bar{\beta}^2))$.

其中

$$\begin{cases}M(Z)\equiv\dfrac{1}{\omega Z}\{\dfrac{1}{2}(1+Z)\omega^2+2(1-\omega)\},\\ N(Z)\equiv\dfrac{1}{Z}\{-2+\omega(1+Z)\}.\end{cases}\tag{9}$$

证　因 Jacobi 阵 J_2 的特征值全为纯虚数，故令 $\mu=i\beta$(β 为实数)代入式(8)得 $(\lambda+\omega-1)^2=-\omega\beta^2(\gamma\lambda-\gamma+\omega)$，或展开得 $\lambda^2-[2(1-\omega)-\omega\gamma\beta^2]+[(1-\omega)^2+\omega\beta^2(\omega-\gamma)]=0$. 则由引理2得

$$|\lambda|<1\Leftrightarrow\begin{cases}|(1-\omega)^2+\omega\beta^2(\omega-\gamma)|<1,\\ |2(1-\omega)-\omega\gamma\beta^2|<1+(1-\omega)^2+\omega\beta^2(\omega-\gamma).\end{cases}$$

$$\Leftrightarrow\begin{cases}(a)\,\omega\gamma\beta^2<2(1-\omega)+\omega^2(1+\beta^2),\\ (b)\,\omega\gamma\beta^2>-2\omega+\omega^2(1+\beta^2),\\ (c)\,\omega\gamma\beta^2<2(1-\omega)+\dfrac{1}{2}\omega^2(1+\beta^2),\\ (d)\,\omega^2(1+\beta^2)>0.\end{cases}\tag{10}$$

所以，式(10)的(d)对任意实数 $\omega\neq0$ 及 β 均成立. 而式(10)的(a)，(b)，(c)则可合并为

$$-2\omega+\omega^2(1+\beta^2)<\omega\gamma\beta^2<2(1-\omega)+\frac{1}{2}\omega^2(1+\beta^2).\tag{11}$$

现讨论两种情况. (i)如 $\underline{\beta}\neq0$，则令 $\beta^2=Z$，再分两种情况考虑：(a)如 $\omega>0$，则式(11)可化为 $N(Z)\equiv\frac{1}{Z}[-2+\omega(1+Z)]<\gamma<\frac{1}{\omega Z}[\frac{1}{2}(1+Z)\omega^2+2(1-\omega)]\equiv M(Z)$，则当 γ 满足 $\max\limits_Z N(Z)<\gamma<\min\limits_Z M(Z)$ 时式(12)成立；(b)如 $\omega<0$，则 γ 应满足 $\max\limits_Z M(Z)<\gamma<\min\limits_Z N(Z)$，但

$$M'(Z)\equiv\frac{d}{dZ}\{\frac{1}{\omega Z}[\frac{1}{2}(1+Z)\omega^2+2(1-\omega)]\}$$

$$=\frac{-1}{2\omega Z^2}(\omega-2)^2=\begin{cases}<0 & (\text{当 } \omega>0,\omega\neq 2),\\ =0 & (\text{当 } \omega=2),\\ >0 & (\text{当 } \omega<0);\end{cases}$$

而

$$N'(Z)\equiv\frac{d}{dZ}\{\frac{1}{Z}[-2+\omega(1+Z)]\}=\frac{1}{Z^2}(2-\omega)=\begin{cases}<0 & (\text{当 } \omega>2),\\ =0 & (\text{当 } \omega=2),\\ >0 & (\text{当 } \omega<2).\end{cases}$$

于是,可以列出 $N(Z)$ 与 $M(Z)$ 的单调增减性,及当 Z 从 $\underline{\beta}^2=\min|\beta|^2$ 变到 $\bar{\beta}^2=\max|\beta|^2$ 时,γ 的变化区间 I_γ(附表).

附表 $N(Z)$ 与 $M(Z)$ 的单调增减性及 γ 的变化区间 I_γ

I_ω	$N(Z)$	$M(Z)$	I_γ
$(-\infty,0)$	↑	↑	$(M(\bar{\beta}^2),N(\underline{\beta}^2))^*$
$(0,2)$	↑	↓	$(N(\bar{\beta}^2),M(\bar{\beta}^2))$
$(2,+\infty)$	↓	↓	$(N(\bar{\beta}^2),M(\bar{\beta}^2))$

* 当且仅当 $(M(\bar{\beta}^2)<N(\underline{\beta}^2)$ 时成立.

(ii)如果 $\underline{\beta}=\min|\beta|=0$,不等式(11)又可分为两种情况:(a)如果 $\beta=0$,则式(11)给出 $0<\omega<2$;(b)如果 $\beta\neq 0$,,则如情况(i)所分析.因对所有 β 均必须满足不等式(11),从而易得 $I_\omega\equiv(0,2)$,$I_\gamma\equiv(N(\bar{\beta}^2),M(\bar{\beta}^2))$.证毕.

3 最优收敛性

定理 2 当 2-块 AOR 方法应用于解最小二乘问题(1)时,且定义参数 $\bar{\beta}=\rho(J_2)=\max|\beta|=\max\limits_{\mu\in\rho(J_2)} I_m(\mu)=\|A_2A_1^{-1}\|_2$ 及 $\underline{\beta}=\min|\beta|=\min|I_m(\mu)|$.如果相应的 Jacobi 迭代矩阵 J_2 有纯虚数特征值 $\mu=i\beta$,使得 $\beta=\bar{\beta}=\underline{\beta}\neq 0$,则对于 $(\gamma,\omega)=[2(-1+\sqrt{1+\beta^2})/\beta^2,1/\sqrt{1+\beta^2}]$ 或 $[-2(1+\sqrt{1+\beta^2})/\beta^2,-1/\sqrt{1+\beta^2}]$,使得 $\rho(L_{\gamma,\omega}^{(2)})=0$(其中 AOR 迭代矩阵 $L_{\gamma,\omega}^{(2)}$ 由式(7)所定义).

证 由文[4]定理 4.3 可知,在定理 1 的条件下,$L_{\gamma,\gamma}$ 矩阵特征值 υ 与 Jacobi 迭代矩阵特征值 μ 之间,满足如下函数关系式 $(\upsilon+\gamma-1)^2=\gamma^2\mu^2\upsilon$,令 $\mu=i\beta$,得

$$(\upsilon-1+\gamma)^2=-\beta^2\gamma^2\upsilon, \tag{12}$$

展开得 $\upsilon^2-[2(1-\gamma)-\beta^2\gamma^2]\upsilon+(1-\gamma)^2=0$.其中,$\mu$ 为 Jacobi 阵 J_2 的特征值,υ 为 $L_{\gamma,\gamma}$ 的特征值.由于假定 β^2 有且只有一个固定值,于是令

$$\Delta=b^2-4ac=[2(1-\gamma)-\beta^2\gamma^2]^2-4(1-\gamma)^2=\beta^2\gamma^2[\beta^2\gamma^2-4(1-\gamma)]=0. \tag{13}$$

若 $\beta\neq 0,\gamma\neq 0$,则方程(13)成为 $\beta^2\gamma^2+4\gamma-4=0$,它有 2 个实根

$$\gamma_1=\frac{2(-1+\sqrt{1+\beta^2})}{\beta^2} \text{ 及 } \gamma_2=\frac{-2(1+\sqrt{1+\beta^2})}{\beta^2}.$$

这时方程(12)有二重根，即 $v=[2(1-\gamma)-\beta^2\gamma^2]/2=1-\gamma-\frac{1}{2}\beta^2\gamma^2$. 由文[3]知 AOR 迭代可看作 SOR 迭代的外推，且其特征值之间满足关系式 $\lambda=sv+(1-s)$，若令 $\lambda=s(v-1)+1=0$，易得 $s=1/(1-v)$. 再将 $s=\omega/\gamma$ 代入上式，得 $\omega=\gamma s=\gamma/(1-v)=1/[1+(1/2)\beta^2\gamma]$，从而对应于 γ_1,γ_2 分别有 $\omega_1=1/\sqrt{1+\beta^2}$，　$\omega_2=-1/\sqrt{1+\beta^2}$. 因此，当取 (γ_1,ω_1) 或 (γ_2,ω_2) 时，$\rho(L_{\gamma,\omega}^{(2)})=0$；所以，当且仅当 β^2 有一固定值(即 $\underline{\mu}=\bar{\mu}$)时，$\rho(L_{\gamma,\omega}^{(2)})=0$. 显然，这比 SOR 迭代的最优值(最小谱半径 $\rho(L_{\omega_b}^{(2)})=[\bar{\beta}/(1+\sqrt{1+\bar{\beta}^2})^2]$[1]好得多．由于连续性，因此至少在 $\underline{\mu}$ 很接近 $\bar{\mu}$ 的情况下，我们可以找到最优 AOR 方法，此时最优 AOR 方法比对应的最优 SOR 方法好．

顺便指出，文[5]也讨论用 2-块 AOR 方法求解最小二乘问题(1)的收敛性，但其所用的 AOR 迭代矩阵特征值与相伴的 Jacobi 迭代阵特征值之间的函数关系式是错误的，因而导致错误的结论．

参　考　文　献

1 Markham T L, Neumann M, Plemmons R J. Comvergence of a direct-iterative method for large-scale least squares problems. LAA, 1985, 69: 155～167

2 Young D M. Iterative solution of large linear systems. New York: Academic Pr., 1971. 145～163

3 Hadjidimos A. Accelerated overrelaxation method. Mathematics of comput., 1978, 32(141): 149～157

4 Varga R S. Matrix iterative analysis. New York: Prentice-Hall, 1962. 48～64

5 沈光星．用 2-块 AOR 方法求解最小二乘问题的收敛域．数学研究与评论，1990，10(3)：429～432

Optimal Convergence of Two-Block AOR Iteration for Solving Least Square Problems

Zeng Wenping

(Dept. of Man. Info. Sci., Huaqiao Univ., 362011, Quanzhou)

Abstract　For solving large-scale sparse least square problems, the author discusses the convergence of two-block AOR iterative method; and gives the necessary and sufficient conditions and the domain of its convergence; and further demonstrates that the spectral radius $\rho(L_{\gamma,\omega}^{(2)})$ of two-block AOR optimal iterative matrix.

Keywords　least square problem, two-block AOR iteration, convergence