1993年3月
第16卷第2期

重庆大学学报
JOURNAL OF CHONGQING UNIVERSITY

Vol. 16, №. 2
Mar. 1993

20

107-112

2一块 AOR 方法求解大型稀疏最小二乘问题的一个收敛定理*

The Convergence of a 2-Block AOR Method for Large-Scaled Least-Square Problems

O241.6

黄德才
Huang Decai
（系统工程及应用数学系）

摘　要　给出了用2一块 AOR 方法求解大型稀疏最小二乘问题收敛的充分必要条件和若干充分条件。当取 r=w 时，使[3]中的相应结果成为本文的推论。结果表明，适当选择参数，2一块 AOR 方法总是收敛的。

关键词　线性方程组；收敛性；最小二乘问题

中国图书资料分类法分类号　O241.6

ABSTRACT　A necessary and sufficient condition and some sufficient conditions concerning the convergence of a 2-block AOR method for large-scaled least-square problems are given. It is shown that, by appropriately selecting parameters, the 2-block AOR method is always convergent and the theorem in [3] is only a corollary of this paper.

KEY WORDS　system of linear equations; convergence; lease-square problem

0　引　言

在许多实际问题中，我们都希望计算以下超定性方程组：

$$Ax = b \tag{1}$$

的最小范数最小二乘解。其中 A 为一大型稀疏的 $m \times n$ 阶实矩阵，$m > n$ 且 $\mathrm{Rank}(A) = n$，b 为一给定的 m 维实列向量。

以上最小二乘问题等价于：求一个 n 维实向量 $x \in R^n$ 和一个 m 维实向量 $r \in R^m$，使

$$b = Ax + r \qquad A^Tr = 0 \tag{2}$$

为了讨论方便，不妨设 $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$，其中 A_1 为 $n \times n$ 的非奇异矩阵，A_2 为 $(m-n) \times n$ 的矩阵。同时，把向量 r 和 b 也作相应分块：

$$r = \begin{bmatrix} V \\ W \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

* 修改稿收到日期　1992-04-13

其中 V,b_1 均为 n 维实向量,W,b_2 均为 m − n 维实向量。这样,(2) 可等价地写成:

$$cz = d \tag{3}$$

其中
$$c = \begin{bmatrix} A_1 & 0 & I \\ A_2 & I & 0 \\ 0 & A_2^T & A_1^T \end{bmatrix} \qquad z = \begin{bmatrix} X \\ W \\ V \end{bmatrix} \qquad d = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

由于 Rank(A_1) = n,容易验证:矩阵 C 是(m + n) × (m + n) 阶的非奇异方阵。

在文献[2]～[4]分别给出了同带有一个松驰因子 w 的 3 − 块 SOR 和 SSOR 求解(3)的有关收敛定理。黄德才,杨万年*给出了用带有两个松驰因子 w、r 的 3 − 块 AOR 方法解(3)的收敛性条件。[5]研究了用 3 − 块 AGS 方法(参数 $w = 1$ 时的 AOR 方法)解(3)的收敛条件,同时还给出用 2 − 块 AOR 方法求解(3)的一个收敛定理,但没给出证明。本文从不同的角度证明了用 2 − 块方法解(3)收敛的一个充分必要条件和若干充分条件。特别,当取 $r = w$ 时,使[3]中的一个定理成为本文的一个推论。同时还证明:2 − 块 AGS 方法解(3)在收敛性质上比 3 − 块 AGS 方法好。

1 $L_{r,w}$ 和 J 的特征值间的关系

为了给出2−块 AOR 方法求解(3)的有关收敛定理,我们先给出解(3)的2−块 AOR 迭代阵和相应的 Jacobi 迭代阵 J 的特征值间的关系。先将(3)的矩阵 c 作如下分块:

$$c = \left[\begin{array}{cc|c} A_1 & 0 & I \\ A_2 & I & 0 \\ \hline 0 & A_2^T & A_1^T \end{array}\right] \qquad 令 \quad D = \left[\begin{array}{cc|c} A_1 & 0 & 0 \\ A_2 & I & 0 \\ \hline 0 & 0 & A_1^T \end{array}\right]$$

则(3)的 Jacobi 迭代矩阵为:

$$J = I - D^{-1}C = \left[\begin{array}{cc|c} 0 & 0 & -A_1^{-1} \\ 0 & 0 & A_2A_1^{-1} \\ \hline 0 & -(A_2A_1^{-1})^T & 0 \end{array}\right] \triangleq \begin{bmatrix} 0 & B_1 \\ B_2 & 0 \end{bmatrix} \tag{4}$$
$$= L + U = \begin{bmatrix} 0 & 0 \\ B_2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & B_1 \\ 0 & 0 \end{bmatrix}$$

而(3)的 SOR 和 AOR 迭代矩阵分别为:

$$L_r = (I - rL)^{-1}[(1-r)I + rU) \qquad r \neq 0$$
$$L_{r,w} = (I - rL)^{-1}[(1-w)I + (w-r)L + wU] \quad r \neq 0$$

当 $w=1$时,AOR 方法即成为 AGS(加速高斯迭代法)方法,其迭代矩阵为:

$$L_{r,1} = (I - rL)^{-1}[(1-r)L + U]$$

若令 $s = w/r$,则 AOR 迭代阵 $L_{r,w}$ 可改写成:

$$L_{r,w} = sL_r + (1-s)I \tag{5}$$

为此,我们有:

定理 1 设入和 μ 分别是 $L_{r,w}$ 和 J 的特征值,则它们满足:

$$[\lambda - (1-w)]^2 = \mu^2 w[\lambda r + (w - r)] \tag{6}$$

---

* AOR 方法求解大稀疏矩阵最小二乘问题的收敛性,即将发表在"数值计算与计算机应用"上

证明：

∵
$$L_r = (I - rL)^{-1}[(1-r)I - rU]$$
$$= (I - rL)^{-1}[(I - rL) - r(I - L - U)]$$
$$= I - rP$$

其中
$$P = (I - rL)^{-1}(I - J)$$

设 τ 和 β 分别是 P 和 L_r 的特征值。则

$$\beta = 1 - r\tau \tag{7}$$

又设 $x \neq 0$ 为 P 关于 τ 的特征向量。

$$\therefore Px = \tau x$$

即：
$$(I - rL)^{-1}(I - J)x = \tau x$$

把上式整理后可得：

$$(J - \tau rL)x = (1 - \tau)x$$

由(4)式可知：

$$J - \tau rL = \begin{pmatrix} 0 & B_1 \\ (1 - \tau r)B_2 & 0 \end{pmatrix}$$

把向量 x 也作相应分块，可得：

$$\begin{cases} B_1x_2 = (1 - \tau)x_1 \\ (1 - \tau r)B_2x_1 = (1 - \tau)x_2 \end{cases}$$

设 $r\tau \neq 1$，则有

$$B_1B_2x_1 = \frac{1 - \tau}{1 - \tau r}B_1x_2 = \frac{(1 - \tau)^2}{1 - \tau r}x_1 \tag{8}$$

又设 μ 为 J 的特征值，$y \neq 0$ 为相应特征向量，则有：$Jy = \mu y$

即：
$$J^2y = \mu^2 y$$

注意到 $J^2 = \begin{bmatrix} B_1B_2 & 0 \\ 0 & B_2B_1 \end{bmatrix}$，并把 y 作相应分块，这样由上式可得：$B_1B_2y_1 = \mu^2 y_1$

把它同(8)式比较可知：$\dfrac{(1 - \tau)^2}{1 - \tau r} = \mu^2$

即：
$$(1 - \tau)^2 = \mu^2(1 - \tau r)$$

由(7)式解出 τ 代入上式，整理后可得：

$$[\beta - (1 - r)]^2 = \mu^2 r^2 \beta \tag{9}$$

又设 λ 为 $L_{r,\omega}$ 的特征值，由(5)式有

$$\lambda = s\beta + (1 - s)$$

从上式中解出 β，并注意 $s = w/r$，我们得：

$$\beta = [\lambda r + (w - r)]/w$$

把 β 代入(9)式并整理就有：

$$[\lambda - (1 - \omega)]^2 = \mu^2\omega[\lambda r + (\omega - r)]$$

2 收敛定理

由 J 的表达式(4)可知：

$$J^2=\begin{bmatrix}0 & A_1^{-1}(A_2A_1^{-1})^T & 0\\ 0 & -(A_2A_1^{-1})(A_2A_1^{-1})^T & 0\\ 0 & 0 & -(A_2A_1^{-1})^T(A_2A_1^{-1})\end{bmatrix}$$

令
$$B=diag(0,-(A_2A_1^{-1})(A_2A_1^{-1})^T,-(A_2A_1^{-1})^T(A_2A_1^{-1}))$$

则由行列式的性质可知：矩阵 J^2 和矩阵 B 有相同的特征值(注：文献[3]～[4]中认为 $J^2=B$，那是不正确的，[5]也没提及这一问题)。又由于 $-(A_2A_1^{-1})(A_2A_1^{-1})^T$，$-(A_2A_1^{-1})^T(A_2A_1^{-1})$ 均是实对称负半定矩阵，所以矩阵 B 的特征值亦即 J^2 的特征值 μ^2 满足：

$$\mu^2\in[-P^2(J),0] \tag{10}$$

为了证明我们的主要定理，先介绍一个熟知的引理，而不给出其证明。

引理 实系数二次方程 $x^2+px+q=0$ 的两根按模均小于1的充分必要条件是：

$$1+p+q>0,\quad 1-p+q>0,\quad 1-q>0.$$

在下面的讨论中，我们总假定松弛因子 r 和 ω 都在开区间(0,2)内。

定理2 2－块AOR方法求解(3)收敛的充分必要条件是参数 r,ω 满足下列条件之一：

① $\omega\geqslant 2r$ 且 $\rho(J)<\sqrt{\dfrac{2-\omega}{\omega-r}}\triangleq f_1(r,\omega)$

② $\omega\leqslant r$ 且 $\rho(J)<\dfrac{2-\omega}{\sqrt{\omega(2r-\omega)}}\triangleq f_2(r,\omega)$

③ $r<\omega<2r$ 且 $\rho(J)<min\{f_1(r,\omega),f_2(r,\omega)\}$

证明：由(10)式可知：μ^2 为负实数。把(6)式右端项移到左端，并仍记 $-\mu^2$ 为 μ^2，整理后有以下等式：

$$\lambda^2+[\omega r\mu^2-2(1-\omega)]\lambda+(1-\omega)^2+\omega\mu^2(\omega-r)=0$$

注意上式中的 $\mu^2\in[0,\rho^2(J)]$，它是 $-J^2$ 的特征值。而 λ 是 $L_{r,\omega}$ 的特征值。因此，以上方程是关于 λ 的实系数二次方程。

2－块AOR方法收敛的充要条件是 $|\lambda|<1$，而由引理可知：$|\lambda|<1$ 的充要条件是：

$$\begin{cases}1+\omega r\mu^2-2(1-\omega)+(1-\omega)^2+\omega\mu^2(\omega-r)>0\\ 1-\omega r\mu^2+2(1-\omega)+(1-\omega)^2+\omega\mu^2(\omega-r)>0\\ 1-(1-\omega)^2-\omega\mu^2(\omega-r)>0\end{cases}$$

以上不等式等价于

$$\begin{cases}(2-\omega)^2+\mu^2\omega(\omega-2r)>0\\ 2-\omega+\mu^2(r-\omega)>0\end{cases} \tag{11}$$

由于 μ^2 是 $-J^2$ 的任意特征值，所以，当 $\mu^2=\rho^2(J)$ 时，(11)式仍然成立。故(11)式等价于

$$\begin{cases}(2-\omega)^2+\rho^2(J)\omega(\omega-2r)>0\\ 2-\omega+\rho^2(J)(r-\omega)>0\end{cases} \tag{12}$$

对(12)式，以下分三种情况来讨论：

(1) 取参数 $\omega\geqslant 2r$，则(12)等价于：

$$(2-\omega)+\rho^2(J)(r-\omega)>0 \tag{13}$$

(2) 取参数 $\omega\leqslant r$，则(12)等价于：

$$(2-\omega)^2+\rho^2(J)\omega(\omega-2r)>0 \tag{14}$$

(3) 取参数 $r<\omega<2r$，则(12)仍是 $|\lambda|<1$ 的充分必要条件。

由以上讨论可知：$|\lambda|<1$ 同以下条件之一等价：

(1) $\omega\geqslant 2r$ 且 $\rho(J)<\sqrt{\dfrac{2-\omega}{\omega-r}}\triangleq f_1(r,\omega)$

(2) $\omega\leqslant r$ 且 $\rho(J)<\dfrac{2-\omega}{\sqrt{\omega(2r-\omega)}}\triangleq f_2(r,\omega)$

(3) $r<\omega<2r$ 且 $\rho(J)<\min\{f_1(r,\omega),f_2(r,\omega)\}$

从定理2可知，无论(3)的 Jacobi 阵 J 的谱半径 $\rho(J)$ 如何，即不管求解(3)的 Jacobi 迭代法是否收敛，总可选择适当的参数 r,ω 使得2—块 AOR 方法解(3)收敛。

推论1　2—块 SOR 方法解(3)收敛的充分必要条件是：$0<\omega<2/(1+\rho(J))$

证明：∵当 $r=\omega$ 时，AOR 方法即是 SOR 方法。所以，在定理2的(2)中令 $r=\omega$，则有

$$\rho(J)<\frac{2-\omega}{\omega}$$

故得

$$0<\omega<2/(1+\rho(J))$$

此推论即是[3]中的定理2。

推论2　当参数 r,ω 满足以下条件之一时，2—块 AOR 方法解(3)收敛。

(1)　$0<2r\leqslant\omega<2/(1+\rho^2(J))\triangleq\alpha_1$

(2)　$0<\omega\leqslant r<2/\{1+\sqrt{2}\rho^2(J)\}\triangleq\alpha_2$

(3)　$0<r<\omega<min\{\alpha_1,\alpha_2\}$

证明：(1) 取 $0<2r\leqslant\omega$，当 $\rho^2(J)<\dfrac{2-\omega}{\omega}$ 时，必有 $\rho(J)<\sqrt{\dfrac{2-\omega}{\omega-r}}$. 由 $\rho^2(J)<\dfrac{2-\omega}{\omega}$ 中解出 ω.

∴　$0<\omega<2/(1+\rho^2(J))$ 时

必有定理2的条件(1)成立，由定理2的充分性知：此时2—块 AOR 方法解(3)收敛。

(2) 取 $0<\omega\leqslant r$，当 $\rho^2(J)<\dfrac{(2-r)^2}{r\cdot 2r}$ 时，必有 $\rho(J)<\dfrac{2-\omega}{\sqrt{\omega(2r-\omega)}}$ 成立。由 $\rho^2(J)<\dfrac{(2-r)^2}{2r^2}$ 解出 r，

∴　$0<\omega\leqslant r<2/(1-2\rho(J))$ 时，

必有定理2的条件(2)成立，故知，此时2—块 AOR 方法解(3)收敛。

(3)由(1)(2)及定理2的(3)即得。

定理3　若松驰因子 r 满足：

$$\max\{0,\rho^2(J)-1)/\rho^2(J)\}<r<min\{(2,(1+\rho^2(J))/2\rho^2(J)\}$$

则：用2—块 AGS 方法求解(3)收敛。

证明：∵AGS 方法是参数 $\omega=1$ 时的 AOR 方法，

∴　当 $\omega=1$ 时，定理2的证明中(12)式以前的推导仍然成立。在(12)式中令 $\omega=1$，可得：

2—块 AGS 方法求解(3)收敛的充分必要条件是：

$$\begin{cases}1+\rho^2(J)(1-2r)>0\\1+\rho^2(J)(1-r)>0\end{cases}$$

以上不等式等价于：

$$(\rho^2(J)-1)/\rho^2(J)<r<(1+\rho^2(J))/2\rho^2(J)$$

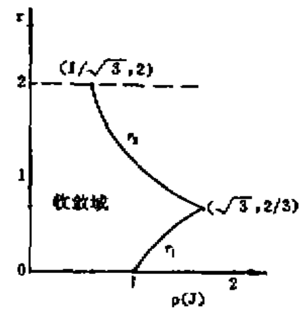
注意到松弛因子 $r \in (0,2)$

故，当 r 满足定理条件时，2 — 块 AGS 方法求解(3) 是收敛的。

令
$$r_1(\rho(J)) = (\rho^2(J) - 1)/\rho^2(J)$$
$$r_2(\rho(J)) = (1 + \rho^2(J))/2\rho^2(J)$$

则 $r_1(\rho(J)) = r_2(\rho(J))$ 的交点是 $\rho(J) = \sqrt{3}$，此时
$$r_1(\rho(J)) = r_2(\rho(J)) = \frac{2}{3}$$

所以，定理3所描述的收敛区域如下图所示：



2-块 AGS 方法的收敛区域(边界除外)

从图上不难看出，当分块 Jacobi 迭代解(3) 发散($1 \leqslant \rho(J) < \sqrt{3}$) 时，2 — 块 *AGS* 方法解(3)仍然是收敛的。为此，我们有

推论3　若 $0 \leqslant \rho(J) < \sqrt{3}$，则可选择适当松弛因子 r，使2—块 *AGS* 方法解(3)是收敛的。

文献[5]中证明仅当 $0 \leqslant \rho(J) < 4^{\frac{1}{3}}$ 时，才可选择适当松弛因子 r，使3—块 *AGS* 法解(3)收敛。本推论表明：用2—块 *AGS* 法解(3)在收敛性质上比用3—块 *AGS* 法好。

作者衷心感谢我系杨万年教授，他在百忙之中仔细阅读了全文，并对本文的修改提出了宝贵建议。

参　考　文　献

1　W. Niethammer, etc. Linear Algebra and its Applications. 1984, 58: 327～341

2　蔡大用. 计算数学. 1985, 3: 295～301

3　T. L. Markham, ect. Linear Algebra and its Applications, 1985, 69: 155～167

4　汤健康. 高等学校计算数学学报, 1988, 2: 189～191

5　E. P. Papadopoulour, etc. SIAM. J. Numer. Anal. 1989, 26(3): 637～660