用 AOR 方法求解大型稀疏最小二乘问题的收敛性*

黄德才　杨万年

（重　庆　大　学）

ON CONVERGENCE OF AOR METHODS FOR LARGE SPARSE LEAST-SQUARES PROBLEMS

Huang De-cai　Yang Wan-nian

(Chongqing University, Chongqing)

Abstract

The convergence of SOR methods and SSOR methods for solving a least-squares problem with a large sparse coefficient matrix has already been proved [1—4]. In this paper, the convergence of AOR methods for solving the equation generated by least-squares problems is considered. It is shown that, by properly selecting the parameters, the AOR methods are always convergent.

在许多实际问题中，我们都希望计算以下超定线性方程组

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

的最小二乘解. 其中 A 为一大型疏 $m \times n$ 实矩阵，$m > n$，b 为一给定的 m 维实向量. 这里假定 $\mathrm{Rank}(A) = n$.

我们知道，(1)可叙述成，求唯一向量 $X \in R^n$，使 $\|b - AX\|_2 = \min\|b - Ay\|_2$ 对一切 $y \in R^n$. 由于 $\mathrm{Rank}(A) = n$，上述最小二乘问题等价于求一个 n 维向量 $X \in R^n$ 和一个 m 维向量 $r \in R^m$，使

$$b = AX + r,\ A^T r = 0. \tag{2}$$

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 $(m-n) \times n$ 阶矩阵. 同时把向量 r 和 b 作相应分块

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

其中 V，b_1 为 n 维实向量，W，b_2 为 $m-n$ 维实向量. 这样(2)可等价地写成

* 1990 年 1 月 31 日收到.

$$CZ = d, \tag{3}$$

其中

$$C = \begin{bmatrix} A_1 & 0 & I \\ A_2 & I & 0 \\ 0 & A_2^T & A_1^T \end{bmatrix}, Z = \begin{bmatrix} X \\ W \\ V \end{bmatrix}, d = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}.$$

由于 A_1 为 $n \times n$ 非奇异方阵，容易验证 $(m+n) \times (m+n)$ 矩阵 C 为非奇异的。

在文献[1]—[4]中分别给出了用有一个参数的 SOR 和 SSOR 方法求解(3)的有关收敛定理。本文给出了用两个参数 ω，r 的 AOR 方法求解 (3) 的收敛定理。同时说明，适当选择参数 ω 和 r，AOR 方法求解(3)总是收敛的。特别当取 $\omega = r$ 时，即得(2)中有关 SOR 方法求解(3)的结果。

令 $D = \mathrm{diag}(A_1, I, A_1^T)$，则(3)的 Jacobi 迭代阵为

$$B = I - D^{-1}C = \begin{bmatrix} 0 & 0 & -A_1^T \\ -A_2 & 0 & 0 \\ 0 & -(A_2A_1^{-1})^T & 0 \end{bmatrix} \triangleq \begin{bmatrix} 0 & 0 & B_1 \\ B_2 & 0 & 0 \\ 0 & B_3 & 0 \end{bmatrix} = L + U,$$

其中

$$L = \begin{bmatrix} 0 & 0 & 0 \\ B_2 & 0 & 0 \\ 0 & B_3 & 0 \end{bmatrix}, U = \begin{bmatrix} 0 & 0 & B_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

相应(3)的 SOR 和 AOR 迭代矩阵分别为

$$\mathscr{L}_\gamma = (I - \gamma L)^{-1}\{(1-\gamma)I + \gamma U\}, \quad \gamma \neq 0,$$

$$\mathscr{L}_{\gamma,\omega} = (I - \gamma L)^{-1}\{(1-\omega)I + (\omega-\gamma)L + \omega U\}, \quad \omega \neq 0.$$

若令 $s = \omega/\gamma$，则 AOR 迭代阵可改写为

$$\mathscr{L}_{\gamma,\omega} = s\mathscr{L}_\gamma + (1-s)I. \tag{4}$$

为此，有如下定理：

定理 1. 设 λ 和 μ 分别是(3)相应的 $\mathscr{L}_{\gamma,\omega}$ 和 B 的特征值，则满足

$$[\lambda - (1-\omega)]^3 = [\gamma\lambda - (\gamma - \omega)]^2\omega\mu^3. \tag{5}$$

证明．设 β 是 $\mathscr{L}_\gamma$ 的特征根，则由(4)有

$$\lambda = s\beta + (1-s). \tag{6}$$

[2]中已证明，B 的特征值 μ 和 $\mathscr{L}_\gamma$ 的特征值 β 满足

$$(\beta + \gamma - 1)^3 = \beta^2\gamma^3\mu^3. \tag{7}$$

又由(6)可得 $\beta = (\lambda - 1)/s + 1$ 把 β 之表达式代入(7)可得

$$[(\lambda-1)/s + \gamma]^3 = [(\lambda-1)/s + 1]^2\gamma^3\mu^3.$$

整理上式，并注意到 $s = \omega/\gamma$ 即得(5)式。

为了给出收敛性定理，我们介绍一个引理。

引理[2]. 设 $f(z) = a_0z^n + a_1z^{n-1} + \cdots + a_n$ 为一实系数多项式，且 $a_0 > 0$，则 $f(z) = 0$ 的根全部落在 Z 平面上的左半部分的充分必要条件是

$$\Delta_1 > 0, \Delta_2 > 0, \cdots, \Delta_n > 0,$$

其中

$$\Delta_i = \begin{vmatrix} a_1 & a_3 & a_5 & \cdots\cdots \\ a_0 & a_2 & a_4 & \cdots\cdots \\ 0 & a_1 & a_3 & \cdots\cdots \\ 0 & a_0 & a_2 & \cdots\cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & a_i \end{vmatrix}.$$

当 $K > n$ 时，$a_K = 0$.

由此引理,我们可得:

推论 1. 设 $g(\eta) = \eta^3 + p_1\eta^2 + p_2\eta + p_3$ 为实系数多项式,令

$$a_0 = 1 + p_1 + p_2 + p_3,$$
$$a_1 = 3 + p_1 - p_2 - 3p_3,$$
$$a_2 = 3 - p_1 - p_2 + 3p_3,$$
$$a_3 = 1 - p_1 + p_2 - p_3,$$

则 $g(\eta) = 0$ 的根全部落在单位圆内的充分必要条件是

$$a_0 > 0,\ a_1 > 0,\ a_3 > 0,\ a_1a_2 - a_0a_3 > 0. \tag{8}$$

证明. 作解析变换

$$\eta = (z + 1)/(z - 1),$$

则它把 η 平面上之单位圆内部映射到 Z 平面的左半部分[5]. 从而 $g(\eta) = 0$ 之根全落在单位圆内部之充分必要条件是 $h(z) = g[(z+1)/(z-1)] = 0$ 的根全部落在 Z 平面的左半部分. 而

$$\begin{aligned} h(z) &= (1 + p_1 + p_2 + p_3)z^3 + (3 + p_1 - p_2 - 3p_3)z^2 \\ &\quad + (3 - p_1 - p_2 + 3p_3)z + (1 - p_1 + p_2 - p_3) \\ &= a_0z^3 + a_1z^2 + a_2z + a_3. \end{aligned}$$

由引理可知,$h(z)=0$ 之根全部落在 Z 平面的左半部分的充要条件是 $a_0>0, \Delta_1 = a_1>0$,

$$\Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0,\quad \Delta_3 = \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & 0 \\ 0 & a_1 & a_3 \end{vmatrix} = a_3\Delta_2 > 0,$$

即 $a_0 > 0,\ a_1 > 0,\ a_3 > 0,\ a_1a_2 - a_0a_3 > 0$.

由定理 1 和推论 1 我们可得:

定理 2. 设 $\gamma, \omega \in (0,2)$, 则对于任意 $\rho(B)$ 当 $0 < \omega < 8/[2 + 3\rho(B)]^3$, $\gamma \geq \max\left(3\omega, \frac{1}{6} + \omega\right)$ 时,用分块 AOR 方法求解(3)总是收敛的.

证明. 注意由文献[2]中已证明, B^3 的特征值 $\mu^3 \in [-\rho^3(B), 0]$, 此处 $\rho(B)$ 为 B 的谱半径.

把(5)式展开可得

$$\begin{aligned} &\lambda^3 - 3(1-\omega)\lambda^2 + 3(1-\omega)^2\lambda - (1-\omega)^3 \\ &\quad = [\lambda^2\gamma^2 - 2\gamma(\gamma - \omega)\lambda + (\gamma - \omega)^2]\omega\mu^3. \end{aligned} \tag{9}$$

对(9)进行移项,合并同类项,并仍记$-\mu^3$ 为 μ^3, 则可得以下等式

$$\lambda^3+[\omega\gamma^2\mu^3-3(1-\omega)]\lambda^2+[3(1-\omega)^2-2\gamma(\gamma-\omega)\omega\mu^3]\lambda$$
$$+(\gamma-\omega)^2\omega\mu^3-(1-\omega)^3=0. \tag{10}$$

注意，这时(10)中的 $\mu^3\in[0,\rho^3(B)]$.

令
$$p_1=-3(1-\omega)+\omega\gamma^2\mu^3,$$
$$p_2=3(1-\omega)^2-2\gamma(\gamma-\omega)\omega\mu^3,$$
$$p_3=-(1-\omega)^3+\omega(\gamma-\omega)^2\mu^3,$$

则得

$$a_0=1+p_1+p_2+p_3=\omega^3(1+\mu^3),$$
$$a_1=3+p_1-p_2-3p_3=3(2-\omega)\omega^2+\omega^2\mu^3(4\gamma-3\omega),$$
$$a_2=3-p_1-p_2+3p_3=3\omega(2-\omega)^2+\omega\mu^3(2\gamma-\omega)\ (2\gamma-3\omega),$$
$$a_3=1-p_1+p_2-p_3=(2-\omega)^3-\omega\mu^3(2\gamma-\omega)^2.$$

由推论 1 知，$|\lambda|<1$ 的充分必要条件是

$$a_0=\omega^3(1+\mu^3)>0, \tag{11}$$
$$a_1=3(2-\omega)\omega^2+\omega^2\mu^3(4\gamma-3\omega)>0, \tag{12}$$
$$a_3=(2-\omega)^3-\omega\mu^3(2\gamma-\omega)^2>0, \tag{13}$$
$$\begin{aligned}a_1a_2-a_0a_3&=8(2-\omega)^3+8(-2\omega^2+6\omega^2+3\gamma^2+4\omega^2\gamma-\gamma^2\omega\\&\quad-12\gamma\omega-3\omega+6\gamma-1)\mu^3+8\mu^6(2\gamma-\omega)(\omega-\gamma)^2\\&=8(2-\omega)^2+8\{(\omega-1)^3+(\omega-\gamma)[(3\omega-\gamma)\\&\quad\times(3-\omega)-6]\}\mu^3+8\mu^6(2\gamma-\omega)(\omega-\gamma)^2>0.\end{aligned} \tag{14}$$

因为 $\omega>0$，$\mu^3\geqslant 0$ 所以(11)式自然成立．若取

$$4\gamma\geqslant 3\omega, \tag{15}$$

则(12)式也成立．

由(13)可得

$$\mu^3<\frac{(2-\omega)^3}{\omega(2\gamma-\omega)^2}, \tag{16}$$

所以只要

$$\rho^3(B)<\frac{(2-\omega)^3}{\omega(2\gamma-\omega)^2} \tag{17}$$

成立，必有(13)成立．又由(15)式并注意到

$$\frac{(2-\omega)^3}{\omega(2\gamma-\omega)^2}>\frac{(2-\omega)^3}{4^2\omega},\ (\omega,\gamma\in(0,2)),$$

所以只要

$$\rho^3(B)<\frac{(2-\omega)^3}{4^2\omega}, \tag{18}$$

则有(17)成立，也就有(13)成立．由(18)可得

$$\omega+\omega^{\frac{1}{3}}4^{\frac{2}{3}}\rho(B)<2,$$

即

$$\omega^{\frac{1}{3}}[\omega^{\frac{2}{3}}+4^{\frac{2}{3}}\rho(B)]<2.$$

因为 $\omega\in(0,2)$ 有 $[\omega^{\frac{2}{3}}+4^{\frac{2}{3}}\rho(B)]<[2+3\rho(B)]$，所以只要

$$\omega^{\frac{1}{3}}[2+3\rho(B)]<2,$$

也就是

$$0<\omega<\left[\frac{2}{2+3\rho(B)}\right]^3, \tag{19}$$

必有(13)式成立.

又注意到(15)式和 $\omega\in(0,2)$ 可知,只要取 γ, ω 使

$$(\omega-1)^3+(\omega-\gamma)\{(3\omega-\gamma)(3-\omega)-6\}\geqslant 0 \tag{20}$$

成立,必有(14)成立.

注意到(19)式可知 $-1<(\omega-1)^3\leqslant 0$. 因此只要

$$(\omega-\gamma)\{(3\omega-\gamma)(3-\omega)-6\}\geqslant 1, \tag{21}$$

就有(20)成立.

只要取

$$\gamma\geqslant\max\left\{3\omega,\ \frac{1}{6}+\omega\right\}, \tag{22}$$

必有(21)成立,也就有(14)成立.

综合(15),(19),(22)可知,只要选

$$0<\omega<\left[\frac{2}{2+3\rho(B)}\right]^3,$$

$$\gamma\geqslant\max\left\{3\omega,\ \frac{1}{6}+\omega\right\},$$

则有(11)—(14)式同时成立,即有 $|\lambda|<1$.

由于 λ 是 AOR 方法求解(3)的迭代阵 $\mathscr{L}_{\gamma,\omega}$ 的特征根，所以，用 AOR 方法求解(3)时,当选择参数 γ, ω 满足条件时,总是收敛的.

由(4)式可知,当 $\gamma=\omega$ 时，AOR 方法即是 SOR 方法. 因此,我们有:

推论 2. 用分块 SOR 方法求解(3)时,当 $0\leqslant\rho(B)\leqslant 2$, 选 $0<\omega<\dfrac{2}{1+\rho(B)}$. 当 $2<\rho(B)<3$ 选 $(\rho(B)-2)/(\rho(B)-1)<\omega<2/(1+\rho(B))$, 迭代法收敛. 若 $\rho(B)\geqslant 3$, 则发散.

证明. 当 $\omega=\gamma$ 时,条件(11)—(14)变为

$$a_0=\omega^3(1+\mu^3)>0, \tag{11'}$$

$$a_1=3(2-\omega)\omega^2+\omega^3\mu^3>0, \tag{12'}$$

$$a_3=(2-\omega)^3-\omega^3\mu^3>0, \tag{13'}$$

$$a_1a_2-a_0a_3=8(2-\omega)^3+8(\omega-1)^3\mu^3>0. \tag{14'}$$

其余证明同[2]中完全一样.

此推论 2 即是[2]中关于 SOR 方法的结果. 从结果可知，用分块 SOR 方法求解(3)时,当 $\rho(B)\geqslant 3$, SOR 永远发散. 定理 2 则告诉我们,若用 AOR 方法求解(3),对任何 $\rho(B)$, 我们总可以适当选择参数 ω, γ 使得 AOR 方法为收敛的. 这说明,用分

块 AOR 方法求解(3)，在收敛性质上比用 SOR 方法要好.

下面，我们用一个简单的数值例子说明. 当相应于(3)的 Jacobi 迭代阵 B 有 $\rho(B)>3$ 时，可用定理 2 选择适当的 γ,ω 用 AOR 方法求解(3)时是收敛的.

例. 设

$$\begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix} x = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \tag{1$'$}$$

则(3)的相应方程为

$$CZ = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 4 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 0 & 4 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \omega \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 0 \end{bmatrix}. \tag{3$'$}$$

这时，(3)′的 Jacobi 迭代阵为

$$B = I - D^{-1}C = \begin{bmatrix} 0 & 0 & 0 & -1 \\ -4 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 0 & -4 & -4 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 0 & -4 & -4 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = L + U.$$

因为

$$|\lambda I - B| = \begin{vmatrix} \lambda & 0 & 0 & 1 \\ 4 & \lambda & 0 & 0 \\ 4 & 0 & \lambda & 0 \\ 0 & 4 & 4 & \lambda \end{vmatrix} = \lambda^4 + 32\lambda,$$

所以

$$\rho(B) = (32)^{\frac{1}{3}} > 3.$$

由[2]知，当用 Jacobi 和 SOR 方法求解(3)′时，永远发散. 但由定理 2，我们只要选择

$$0 < \omega < \left[\frac{2}{2+3\rho(B)}\right]^3 \doteq 5.2267809 * 10^{-3},$$

$$\gamma \geqslant \max\left\{3\omega, \frac{1}{6} + \omega\right\},$$

即可使用 AOR 方法解(3)′为收敛的.

这里我们选 $\omega = 0.005$，则可选 $\gamma = 1$. 这时，相应于(3)′的 AOR 迭代阵

$$\begin{aligned} \mathscr{L}_{\gamma,\omega} &= (I - \gamma L)^{-1}((1-\omega)I + (\omega - \gamma)L + \omega U) \\ &= (I - L)^{-1}[(1 - 0.005)I + (0.005 - 1)L + 0.005U] \end{aligned}$$

$$
=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 0 & 4 & 4 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 0.995 & 0 & 0 & -0.005 \\ 4*0.995 & 0.995 & 0 & 0 \\ 4*0.995 & 0 & 0.995 & 0 \\ 0 & 4*0.995 & 4*0.995 & 0.995 \end{bmatrix}
$$

$$
=\begin{bmatrix} 1 & 0 & 0 & 0 \\ -4 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ 32 & -4 & -4 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 0.995 & 0 & 0 & -0.005 \\ 4*0.995 & 0.995 & 0 & 0 \\ 4*0.995 & 0 & 0.995 & 0 \\ 0 & 4*0.995 & 4*0.995 & 0.995 \end{bmatrix}
$$

$$
=\begin{bmatrix} 0.995 & 0 & 0 & -0.005 \\ 0 & 0.995 & 0 & 0.02 \\ 0 & 0 & 0.995 & 0.02 \\ 0 & 0 & 0 & 0.835 \end{bmatrix}, \ (\omega \neq 0),
$$

即选 $\omega = 0.005$, $\gamma = 1$ 时, $\rho(L_{\gamma,\omega}) = 0.995 < 1$, 故 $\rho(B) > 3$ 时,可选择 γ, ω 使得用 AOR 方法解(3)′是收敛的.

注意到定理 2 中关于 ω 和 γ 选择范围,只是收敛的充分条件. 因此,定理 2 中选出的参数 ω 和 γ 并不是使得 AOR 收敛得最快的. 事实上,在本例中我们可以选择 $\omega = 0.05$ $\gamma = 1$ 而使得 $\rho(\mathscr{L}_{\gamma,\omega}) = 0.95 < 0.995$, 即这组参数使得收敛速度更快. 在实际应用中,可就定理 2 给出的参数范围适当放宽,以选出使收敛较快的参数.

参 考 文 献

[1] W. Niethammer, J. DE. Pillis, R. S. Varga, Linear Algebra and its Applications, 58, (1984) 327—341.

[2] 蔡大用,计算数学,7: 3(1985)295—301.

[3] T. L. Markham, M. Neumann, R. J. Plemmons, Linear Algebra and its Applications, 69 (1985) 155—167.

[4] 汤健康,高等学校计算数学学报,10: 2(1988),189－191.

[5] 钟玉泉,复变函数,高等教育出版社出版,1988 年 5 月.

[6] 蔡大用,数值代数,清华大学出版社出版, 1987 年 9 月.