Estimates of Convergence Rate of Parallel Multisplitting Iterative Methods with Their Applications

Tongxiang Gu

1 Introduction

Consider large-scale systems of linear algebraic equations

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

where $A \in R^{n\times n}$ and $b \in R^n$. O'Leary and White [8] proposed a parallel multisplitting iterative method (the PMI-method) for solving (1) in 1985. Through multisplitting of A

$$A = M_l - N_l; \qquad \text{with} \qquad \det(M_l) \neq 0; \qquad l = 1, 2, \cdots, k, \tag{2}$$

they constructed the iterative procedures

$$y_l^m = M_l^{-1} N_l x^m + M_l^{-1} b; \qquad l = 1, 2, \cdots, k \tag{3}$$

for (1). By introducing weighting matrices E_l for $l = 1, 2, \cdots, k$ with

$$0 \leq E_l \leq I; \qquad \sum_{l=1}^{k} E_l = I, \tag{4}$$

where I is the $n \times n$ identity matrix, they combined (3) and obtained the PMI-method:

$$x^{m+1} = \sum_{l=1}^{k} E_l y_l^m; \qquad m = 0, 1, 2, \cdots \tag{5}$$

[1] Department of Mathematics, Henan Normal University, Xinxiang, Henan, 453002, P.R.China

Domain Decomposition Methods in Sciences and Engineering, edited by R. Glowinski *et al.*

The triple (M_l, N_l, E_l) $l = 1, 2, \cdots, k$ is called a multisplitting of the matrix A and we can rewrite (5) in the equivalent form as:

$$x^{m+1} = Hx^m + Gb \tag{6}$$

where $H = \sum_{l=1}^{k} E_l M_l^{-1} N_l; G = \sum_{l=1}^{k} E_l M_l^{-1}$.

Many authors have presented different schemes based on different multisplittings of the coefficient matrix A, for example, PMI-GS and PMI-SGS [7], PMI-SOR [2], PMI-AOR [10], PMI-GSOR and PMI-GAOR [3]. The convergence of these methods were proved under different conditions. However, the methods used to prove the convergence in [2], [3], [7], [10] were inconvenient and varied. The author knows of new previous estimates of the convergence rate of PMI-method. So it is necessary to simplify and unify the proof of the convergence of PMI-method and to estimate the convergence rate of PMI-methods simply and practically.

2 Estimates of the Convergence Rate of PMI-methods

It is well known that the estimate of asymptotic convergence rate $R(H)$ of iterative method is equivalent to the estimate of the spectral radius $\rho(H)$ of the iterative matrix H, because $R(H) = -\log(\rho(H))$. We will denote M_l, N_l, E_l by (m_{ij}^l), (n_{ij}^l), (e_{ij}^l), respectively, and omit the index $l = 1, 2, \cdots, k$. We denote $\sum_{l=1}^{k}$, $\sum_{j=1}^{n}$, $\sum_{\substack{j=1 \\ j\neq i}}^{n}$, $\max_{1\leq l\leq k}$, $\max_{1\leq i\leq n}$, $\min_{1\leq l\leq k}$, $\min_{1\leq i\leq n}$ by $\sum_l, \sum_j, \sum_{j\neq i}, \max_l, \max_i, \min_l, \min_i$, respectively.

Theorem 1. *Let A be nonsingular and (M_l, N_l, E_l) be a multisplitting of A. If M_l is SDD (strictly diagonally dominant), then*

$$\rho(H) \leq \| H \|_\infty \leq \max_l\{\max_i\{\sum_j \frac{|n_{ij}^l|}{|m_{ii}^l| - \sum_{j\neq i} |m_{ij}^l|}\}\}. \tag{7}$$

Proof. It is well known that $\rho(H) \leq \| H \|_\infty$, so we need only to prove the right inequality.

Let $n_l = (n_1^l, n_2^l, \cdots, n_n^l)^T$ be an n-vector and $M_l^{-1} n_l = x_l := (x_1^l, x_2^l, \cdots, x_n^l)^T$. Thus,

$$\sum_l E_l M_l^{-1} n_l = \sum_l E_l x_l = \left(\sum_l e_{11}^l x_1^l, \sum_l e_{22}^l x_2^l, \cdots, \sum_l e_{nn}^l x_n^l\right)^T.$$

Since $\sum_l e_{ii}^l = 1$ for $i = 1, 2, \cdots, n$, we have

$$\left\| \sum_l E_l M_l^{-1} n_l \right\|_\infty = \max_i\{|\sum_l e_{ii}^l x_i^l|\} \leq \max_l\{\max_i\{|x_i^l|\}\} := |x_{i_0}^{l_0}|.$$

Consider the i_0-th equation of $M_{l_0}x_{l_0} = n_{l_0}$.

$$|m_{i_0i_0}^{l_0}| \cdot |x_{i_0}^{l_0}| = \left|n_{i_0}^{l_0} - \sum_{j\neq i} m_{i_0j}^{l_0} x_{i_0}^{j}\right| \leq |n_{i_0}^{l_0}| + |x_{i_0}^{l_0}| \cdot \sum_{j\neq i} |m_{i_0j}^{l_0}|$$

and this implies

$$|x_{i_0}^{l_0}| \leq \frac{|n_{i_0}^{l_0}|}{|m_{i_0i_0}^{l_0}| - \sum\limits_{j\neq i_0} |m_{i_0j}^{l_0}|} \leq \max_l\{\max_i\{\frac{|n_i^l|}{|m_{ii}^l| - |\sum\limits_{j\neq i} m_{ij}^l|}\}\}.$$

Hence

$$\left\|\sum_l E_l M_l^{-1} n_l\right\|_\infty \leq \max_l\{\max_i\{\frac{|n_{ij}^l|}{|m_{ii}^l| - |\sum\limits_{j\neq i} m_{ij}^l|}\}\}.$$

M_l is SDD. Let $D_l = \text{diag}(M_l)$, $C_l = D_l - M_l$, then $\langle M_l\rangle = |D_l| - |C_l|$, $|D_l^{-1}C_l| \leq |D_l|^{-1}|C_l|$ (in fact, the equality holds). We have

$$\rho(D_l^{-1}C_l) \leq \rho(|D_l|^{-1}|C_l|) < 1.$$

Hence M_l is invertible and

$$|M_l^{-1}| = |(D_l - C_l)^{-1}| = |\sum_{j=0}^{\infty}(D_l^{-1}C_l)^j D_l^{-1}|$$

$$\leq \sum_{j=0}^{\infty}(|D_l|^{-1}|C_l|)^j |D_l|^{-1} = (|D_l| - |C_l|)^{-1} = \langle M_l\rangle^{-1}.$$

Furthermore

$$|M_l^{-1}N_l| \leq \langle M_l\rangle^{-1}|N_l|; \quad |H| \leq \sum_l E_l|M_l^{-1}N_l| \leq \sum_l E_l\langle M_l\rangle^{-1}|N_l| := F.$$

Let $N_l = (N_1^l, N_2^l, \cdots, N_n^l)$, where N_i^l is the i-th column vector of N_l. Then

$$\begin{aligned}
\rho(H) \quad &\leq \| H \|_\infty \leq \| F \|_\infty = \max_i\{\sum_j(\sum_l E_l\langle M_l\rangle^{-1}|N_l|)_{ij}\} \\
&= \max_i\{\sum_j(\sum_l E_l\langle M_l\rangle^{-1}|N_1^l|, \sum_l E_l\langle M_l\rangle^{-1}|N_2^l|, \cdots, \sum_l E_l\langle M_l\rangle^{-1}|N_n^l|)_{ij}\} \\
&= \max_i\{\sum_l(E_l\langle M_l\rangle^{-1}\sum_j |N_j^l|)_i\} = \| \sum_l E_l\langle M_l\rangle^{-1}\sum_j |N_j^l| \|_\infty \\
&\leq \max_l\{\max_i\{\sum_j \frac{(|N_j^l|)_i}{|m_{ii}^l| - \sum\limits_{j\neq i}|m_{ij}^l|}\}\} = \max_l\{\max_i\{\sum_j \frac{|n_{ij}^l|}{|m_{ii}^l| - \sum\limits_{j\neq i}|m_{ij}^l|}\}\}.
\end{aligned}$$

This proves the theorem.

Remark. We can see from the above proof that (7) holds for $M_l \in R^{n\times n}, N_l \in R^{n\times m}$ under the assumption that M_l is SDD. We can make the same remark on the following theorem and corollaries 1 and 2.

Theorem 2. *Under the assumptions of theorem 1 and assumming that M_l is a L-matrix, $N_l \geq 0$, it follows*

$$\min_i\{\sum_j(\sum_l E_l M_l^{-1} N_l)_{ij}\} \geq \min_l\{\min_i\{\sum_j \frac{|n_{ij}^l|}{|m_{ii}^l| - \sum\limits_{j\neq i}|m_{ij}^l|}\}\}. \tag{8}$$

The proof of this theorem is analogous to that of theorem 1.

Corollary 1. *Under the assumptions of theorem 2, we have*

$$\min_l\{\min_i\{\sum_j \frac{n_{ij}^l}{\sum\limits_j m_{ij}^l}\}\} \leq \min_i\{\sum_j(\sum_l E_l M_l^{-1} N_l)_{ij}\} \leq \rho(H) \leq \| H \|_\infty$$

$$= \max_i\{\sum_j(\sum_l E_l M_l^{-1} N_l)_{ij}\} \leq \max_l\{\max_i\{\sum_j \frac{n_{ij}^l}{\sum\limits_j m_{ij}^l}\}\}. \tag{9}$$

In particular, if $\sum\limits_j n_{ij}^l / \sum\limits_j m_{ij}^l \equiv k$(constant), $\rho(H) = \| H \|_\infty = k$. If $\sum\limits_j n_{ij}^l / \sum\limits_j m_{ij}^l \not\equiv$ constant and H is irreducible, the two middle inequalities are strict.

Proof. The first part is obtained from theorem 1 and 2, while the second part can be obtained from theorem 9 in §1.3 of [6].

Corollary 2 ([4, 5]). *Under the assumptions of theorem 2, we have*

$$\min_i\{\sum_j \frac{n_{ij}}{\sum\limits_j m_{ij}}\} \leq \min_i\{\sum_j (M^{-1}N)_{ij}\} \leq \rho(M^{-1}N) \leq \| M^{-1}N \|_\infty$$

$$= \max_i\{\sum_j (M^{-1}N)_{ij}\} \leq \max_i\{\sum_j \frac{n_{ij}}{\sum\limits_j m_{ij}}\}. \tag{10}$$

Proof. We obtain (10) from corollary 1 taking $k = 1$ and $M_1 = M, N_1 = N$.

We can get a further direct result from theorem 1.

Corollary 3. *If M_l, N_l satisfy*

$$|m_{ii}^l| \geq \sum_{j\neq i}|m_{ij}^l| + \sum_j |n_{ij}^l| \qquad i = 1, 2, \cdots, n \tag{11}$$

and

$$|m_{ii}^l| > \sum_{j\neq i}|m_{ij}^l|; \qquad i = 1, 2, \cdots, n, \tag{12}$$

then

$$\rho(H) \leq \| H \|_\infty \leq 1. \tag{13}$$

The second inequality in (13) will be strict when the inequality (11) is strict. The PMI-method is then convergent.

3 Numerical Examples

Consider the systems of linear equations (1), where

$$A = \begin{bmatrix} 5 & -2 & -2 \\ -4 & 10 & -4 \\ -2 & -2 & 5 \end{bmatrix}; \; b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

We split A into several two-splittings as follows

$$(a) \qquad M_1 = \begin{bmatrix} 5 & -1 & -1 \\ -2 & 10 & -2 \\ -1 & -1 & 5 \end{bmatrix}; \; M_2 = \begin{bmatrix} 5 & -2 & 0 \\ -4 & 10 & 0 \\ 0 & -2 & 5 \end{bmatrix};$$

$$N_1 = M_1 - A; \;\; N_2 = M_2 - A; \;\; E_1 = \text{diag}(0,0,1); \;\; E_2 = \text{diag}(1,1,0);$$

(b) Take M_1, M_2, N_1, N_2 as in (a) and take $E_1 = \text{diag}(1,0,0); E_2 = \text{diag}(0,1,1)$;
(c) Take M_1, N_1, E_1, E_2 as in (a) and take $M_2 = \text{diag}(5,10,5); N_2 = M_2 - A$.

It follos from Corollary 2 that

$$\begin{array}{ll} \rho(H) = 2/3 & \text{for case (a) and case (b),} \\ 2/3 < \rho(H) < 4/5 & \text{for case (c).} \end{array}$$

In fact, we get from practical computing that

$$\begin{array}{lll} (a) & \det(\lambda I - H) = \lambda(3\lambda - 2)(9\lambda + 5)/27 & \text{and } \rho(H) = 2/3, \\ (b) & \det(\lambda I - H) = (3\lambda - 2)(6\lambda + 1)(18\lambda + 5)/972 & \text{and } \rho(H) = 2/3, \\ (c) & \det(\lambda I - H) = (5\lambda + 2)(15\lambda^2 - 6\lambda - 4)/75 & \text{and} \\ & 2/3 < \rho(H) = (6 + \sqrt{276})/30 \approx 22.61/30 < 4/5 & . \end{array}$$

4 Applications

In this section, we give some convergence and divergence theorems of relaxed PMI-methods by using the estimates established in §2. In order to do this, the concepts of optimally scaled matrix and its several properties introduced in [5] are useful.

Lemma 1 **([5]).** *Let $A = (a_{ij})$ bea irreducible matrix with nonzero diagonal entries, $D = diag(A), B = D - A$. Then there exists a diagonal matrix $Q = diag(q_1, q_2, \cdots, q_n)$ with positive diagonal entries such that $\tilde{A} = (\tilde{a}_{ij}) = AQ$,*

$$\sum_{j \neq i} |\, \tilde{a}_{ij} \,| / |\, \tilde{a}_{ii} \,| = \rho(|D|^{-1}|B|) \qquad i = 1, 2, \cdots, n \tag{14}$$

here $\tilde{A}$ is unique except for a constant factor. Furthermore, for an arbitrary $\bar{A} = (\bar{a}_{ij}) = \tilde{A} \;\; P$ where $P = diag(p_1, p_2, \cdots, p_n)$ with $0 < p_i \not\equiv constant$ for $i = 1, 2, \cdots, n$, we have

$$\min_i \{\sum_{j \neq i} |\, \bar{a}_{ij} \,| / |\, \bar{a}_{ii} \,|\} \leq \rho(|D|^{-1}|B|) \leq \max_i \{\sum_{j \neq i} |\, \bar{a}_{ij} \,| / |\, \bar{a}_{ii} \,|\}. \tag{15}$$

We call $\widetilde{A}$ the optimally scaled matrix of A. Several properties can be established for it.

Property 1 ([5]). Under the assumptions of lemma 1, we have

$$\rho(|\widetilde{D}|^{-1}|\widetilde{B}|) = \| |\widetilde{D}|^{-1}|\widetilde{B}| \|_\infty = \rho(|D|^{-1}|B|)$$

where $\widetilde{D} = \text{diag}(\widetilde{A})$, $\widetilde{B} = \widetilde{D} - \widetilde{A}$.

Property 2 ([5]). Under the assumptions of lemma 1, the following four properties are equivalent.

(*a*) $\rho(|D|^{-1}|B|) < 1$, (*b*) $\rho(|\widetilde{D}|^{-1}|\widetilde{B}|) < 1$,
(*c*) A is H-matrix, i. e., $\langle A\rangle$ is M-matrix, (*d*) $\widetilde{A}$ is H-matrix, i.e.,$\langle\widetilde{A}\rangle$ is M-matrix .

And if one of the above holds, $\widetilde{A}$ is SDD, i.e., $|\widetilde{a}_{ii}| > \sum_{j\neq i} |\widetilde{a}_{ij}|$.

Property 3 ([5]). Under the assumptions of lemma 1, if A and $\widetilde{A}$ have matrix splittings $A = M - N; \widetilde{A} = \widetilde{M} - \widetilde{N}$ where $\widetilde{M} = MQ, \widetilde{N} = NQ, M^{-1}$ exists. Here Q is given in lemma 1. Then

$$\rho(M^{-1}N) = \rho(\widetilde{M}^{-1}\widetilde{N}). \tag{16}$$

Let $(D - L_l, U_l, E_l)$ be a multisplitting of A, where $D = \text{diag}(A)$, L_l is strictly lower triangular. Let $R = \text{diag}(r_1, r_2, \cdots, r_n)$, $\Omega = \text{diag}(\omega_1, \omega_2, \cdots, \omega_n)$ be relatation matrices, where $r_i \geq 0, \omega_i > 0$ for $i = 1, 2, \cdots, n$ and let $\omega > 0, r \geq 0$. We can then write the iterative matrices of PMI-SOR [2], PMI-GSOR [3], PMI-AOR [10] and PMI-GAOR [3] as follows:

$$\pounds_\omega(A) = \sum_l E_l(D - \omega L_l)^{-1}[(1 - \omega)D + \omega U_l],$$
$$\pounds_\Omega(A) = \sum_l E_l(D - \Omega L_l)^{-1}[(I - \Omega)D + \Omega U_l],$$
$$\pounds_{r,\omega}(A) = \sum_l E_l(D - rL_l)^{-1}[(1 - \omega)D + (\omega - r)L_l + \omega U_l],$$
$$\pounds_{R,\Omega}(A) = \sum_l E_l(D - RL_l)^{-1}[(I - \Omega)D + (\Omega - R)L_l + \Omega U_l].$$

If we let $D = \text{diag}(A), B = D - A$, then the Jacobi iterative matrix is $J(A) = D^{-1}B$. We denote the intervals $\left(0, \frac{2}{1+\rho(J(A))}\right)$ and $\left[0, \frac{2}{1+\rho(J(A))}\right)$ as I_A and $\widetilde{I}_A$, respectively. We then have the following theorem.

Theorem 3. *Let A be an H-matrix with $\langle A\rangle = D - |L_l| - |U_l|$. Then*
(*a*) $\rho(\pounds_\omega(A)) < 1$, $\forall\omega \in I_A$.
(*b*) $\rho(\pounds_\Omega(A)) < 1$, $\forall\omega_i \in I_A(\forall i)$.
(*c*) $\rho(\pounds_{r,\omega}(A)) < 1$, $\forall r \leq \omega, r \in \widetilde{I}_A, \omega \in I_A$.
(*d*) $\rho(\pounds_{R,\Omega}(A)) < 1$, $\forall r_i \leq \omega_i, r_i \in \widetilde{I}_A, \omega_i \in I_A(\forall i)$.

Proof. Since $(a), (b), (c)$ are special cases of (d), we only give a proof of (d). We first assume that A is irreducible. Hence, by theorem 1, we have

$$\begin{aligned}\rho(\pounds_{R,\Omega}(A)) &= \rho(\pounds_{R,\Omega}(\widetilde{A})) \leq \| \pounds_{R,\Omega}(\widetilde{A}) \|_\infty \\ &\leq \max_l\{\max_i\{\frac{|1-\omega_i||\widetilde{a}_{ii}|+(\omega_i-r_i)\sum_{j<i}|l^l_{ij}|++\omega_i\sum_j|u^l_{ij}|}{\widetilde{a}_{ii}-r_i\sum_{j<i}|l^l_{ij}|}\}\} \\ &= \max_l\{\max_i\{\frac{|1-\omega_i|+\omega_i\rho(|J(A)|)-r_i\sum_{j<i}|l^l_{ij}|/|\widetilde{a}_{ii}|}{1-r_i\sum_{j<i}|l^l_{ij}|/|\widetilde{a}_{ii}|}\}\}.\end{aligned} \tag{17}$$

When $\omega_i \in I_A$ for $i = 1, 2, \cdots, n$, we obtain $|1-\omega_i| + \omega_i\rho(|J(A)|) < 1$. When $a, b, c > 0, a < b$, we have $(a-c)/(b-c) < a/b$. Hence, when $r_i \leq \omega_i, r_i \in \widetilde{I}_A, \omega_i \in I_A(\forall i)$, we get from (17) that

$$\rho(\pounds_{R,\Omega}(\widetilde{A})) = \rho(\pounds_{R,\Omega}(A)) < |1-\omega_i| + \omega_i\rho(|J(A)|) < 1.$$

If A is reducible, we can change some zero entries of A into sufficient a small positive number $\epsilon > 0$ such that A change into A_ϵ and A_ϵ is irreducible. We can work with A_ϵ as above. Finally, we can show that theorem 3 holds when A is reducible by taking $\epsilon \to 0$ and using the continuity of the spectral radius of the entries of the matrix.

Remark. Theorem 3 shows that PMI-SOR, PMI-GSOR, PMI-AOR and PMI-GAOR are convergent if the parameters are in the intervals given in theorem 3. This is in keeping with results given in [2], [3], [10]. We unify the proof of convergence of these relaxed PMI-methods. Using theorem 1, we can get more general results than those given in [2], [3], [10] and theorem 3.

Theorem 4. *If A is SDD, $\langle A\rangle = |D| - |L_l| - |U_l|$ and $\omega > 0, \omega_i > 0, 0 \leq r \leq \omega, 0 \leq r_i \leq \omega_i$ for $i = 1, 2, \cdots, n$ are all smaller than $2|a_{ii}|/\sum_j |a_{ij}|$, then all relaxed PMI-methods in theorem 3 are convergent.*

Theorem 5. *If A is an irreducible L-matrix but not a M-matrix, then*

(a) $\rho(\pounds_\omega(A)) \geq 1$ *for sufficiently small* ω .
(b) $\rho(\pounds_\Omega(A)) \geq 1$ *for sufficiently small* $\omega_i(\forall i)$.
(c) $\rho(\pounds_{r,\omega}(A)) \geq 1$ *for sufficiently small* r, ω *with* $0 \leq r \leq \omega$.
(d) $\rho(\pounds_{R,\Omega}(A)) \geq 1$ *for suitable small* r_i, ω_i *with* $0 \leq r_i \leq \omega_i(\forall i)$.

Proof. We also only show that (*d*) holds. First, we choose r_i sufficiently small such that $D - RL_l$ is SDD. Then we have from theorem 2 and property 3, that

$$\begin{aligned}\rho(\pounds_{R,\Omega}(A)) &= \rho(\pounds_{R,\Omega}(\widetilde{A})) \\ &\geq \min_l\{\min_i\{\frac{[1-\omega_i] + \omega_i\rho(|J(A)|) - r_i\sum_{j<i}|l^l_{ij}|/|\widetilde{a}_{ii}|}{1 - r_i\sum_{j<i}|l^l_{ij}|/|\widetilde{a}_{ii}|}\}\}\end{aligned} \tag{18}$$

and from property 2, we get that the right hand of inequality (18) is not smaller than unit for any l and i. So (*d*) holds.

REFERENCES

[1] Berman A. and Plemmons R. J. (1979) *Nonnegative Matrices in the Mathematical and Science*, Academic Press, New York.

[2] Frommer F. and Mayer G. (1989) Convergence of relaxed parallel multisplitting methods. *Lin. Alg. Appl.* 119:141-152.

[3] Gu Tongxiang and Wang Nengchao (1992) A class of multisplitting iterative methods. *Proceedings of Third National Conference on Parallel Algorithm of China,* Huazhong Uni. Sci. Tech. Press, Wuhan, China, 186-190.

[4] Hu Jiagan (1982) Estimates of $\| B^{-1}A \|$ and their applications. *Numerica Mathematica Sinica,* 4(3):272-282.

[5] Hu Jiagan (1983) Scaling transformation and convergence of splitting of matrix,*Numerica Mathematica Sinica,* 5(1):72-78.

[6] Hu Jiagan (1991) *Iterative Methods for Solving Linear Systems of Algebraic Equations*, Science Press, Beijing, China.

[7] Neumann M. and Plemmons R. J. (1987) Convergence of parallel multisplitting iterative methods for M-matrices.*Lin. Alg. Appl.* 88/89:559-573.

[8] O'Leary D. P. and White R. E. (1985) Multi-splitting of matrices and parallel solution of linear systems. *SIAM J. Alg. Disc. Meth.* 6(4):630-640.

[9] Varga R. S. (1962) *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N. J.

[10] Wang Deren (1991) On the convergence of parallel multisplitting AOR algorithm. *Lin. Alg. Appl.* 154/156:473-486.

[11] Young D. M. (1971) *Iterative Solution of Large Linear Systems*, Academic Press, New York .