

NONSTATIONARY MULTISPLITTINGS WITH GENERAL WEIGHTING MATRICES*

VIOLETA MIGALLÓN[†], JOSÉ PENADÉS[†], AND DANIEL B. SZYLD[‡]

Abstract. In the convergence theory of multisplittings for symmetric positive definite (s.p.d.) matrices it is usually assumed that the weighting matrices are scalar matrices, i.e., multiples of the identity. In this paper, this restrictive condition is eliminated. In its place it is assumed that more than one (inner) iteration is performed in each processor (or block). The theory developed here is applied to nonstationary multisplittings for s.p.d. matrices, as well as to two-stage multisplittings for symmetric positive semidefinite matrices.

Key words. iterative methods, linear systems, symmetric positive definite matrices, block methods, parallel algorithms, multisplitting, two-stage, nonstationary

AMS subject classifications. 65F10, 65F15

PII. S0895479800367038

1. Introduction. Multisplitting methods, first introduced by O’Leary and White [18], developed into an important theoretical tool in the study of parallel block iterative methods for the solution of linear (and nonlinear) systems of equations; see, e.g., [4], [12], [16], and the extensive references therein. In these methods, for the solution of a nonsingular system $Ax = b$, several splittings $A = M_\ell - N_\ell$, $\ell = 1, \dots, p$, (M_ℓ nonsingular) are used, together with a set of diagonal nonnegative (weighting) matrices E_ℓ , such that they add up to the identity; see Algorithm 1.1 below.

Most of the convergence results obtained with the philosophy of multisplittings throughout the literature relate to nonsymmetric matrices. The reason for this is that in these cases, general convergence results were obtained for quite general weighting matrices thus allowing the study of truly parallel methods (with or without overlap), i.e., methods in which each processor computes an approximation to the solution of a problem which is much smaller than the original problem.

In contrast, most results for general symmetric positive definite (s.p.d.), or more generally, Hermitian positive definite, linear systems require the assumption that weighting matrices are multiples of the identity

$$(1.1) \quad E_\ell = \alpha_\ell I, \quad \ell = 1, \dots, p$$

(see, e.g., [7], [8], [15], [18]), thus these results have little applicability for analysis of parallel processing. We note here that in [21], condition (1.1) is not present, but the splittings of A have a very special structure. We also mention [6] where a nonstandard multisplitting is used for s.p.d. matrices.

The main reason for the requirement (1.1) becomes apparent with examples found in the literature, e.g., in [7], [18], showing that relaxing (1.1) may lead to a divergent

*Received by the editors February 4, 2000; accepted for publication (in revised form) by D. O’Leary October 19, 2000; published electronically March 7, 2001.

<http://www.siam.org/journals/simax/22-4/36703.html>

[†]Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Alicante, E-03080 Alicante, Spain (violeta@dccia.ua.es, jpenades@dccia.ua.es). This research was supported by Spanish DGESIC grant PB98-0977.

[‡]Department of Mathematics, Temple University, Philadelphia, PA 19122-6094 (szyld@math.temple.edu). This author was supported by National Science Foundation grants INT-9521226 and DMS-9973219.

multisplitting method. In this note we prove new convergence theorems which, for the first time, show that one can use the multisplitting idea on s.p.d. matrices without the restriction (1.1), i.e., with quite general weighting matrices. As we show later, the price we pay for this generality is a few extra inner iterations.

In the following version of the multisplitting method, which is a special case of the nonstationary two-stage multisplitting method [20] (see also section 3 for another special case), the sequence $s(\ell, k)$ indicates, e.g., the number of local iterations used to approximate the solution of the ℓ th system, at the k th iteration; see also [3]. We call them local iterations, since they correspond to work performed in each processor. In the standard multisplitting method of [18], one has $s(\ell, k) = 1$ for all ℓ and k .

ALGORITHM 1.1 (NONSTATIONARY MULTISPLITTING). *Given an initial vector x_0 and a sequence of numbers of local iterations $s(\ell, k)$, $\ell = 1, \dots, p$, $k = 1, 2, \dots$*

For $k = 1, 2, \dots$, until convergence.

For $\ell = 1$ to p

$$\text{padding-left: 120px; } y_{\ell,0} = x_{k-1}$$

For $j = 1$ to $s(\ell, k)$

$$(1.2) \quad \text{padding-left: 120px; } M_{\ell} y_{\ell,j} = N_{\ell} y_{\ell,j-1} + b$$

$$x_k = \sum_{\ell=1}^p E_{\ell} y_{\ell,s(\ell,k)}.$$

The iteration matrix at the k th step of this multisplitting method is

$$(1.3) \quad T_k = \sum_{\ell=1}^p E_{\ell} (M_{\ell}^{-1} N_{\ell})^{s(\ell,k)},$$

i.e., $e_k = T_k e_{k-1}$, where the error at each step is $e_k = x_k - x_{\star}$ and x_{\star} is the solution of $Ax = b$. Convergence of the method is obtained for any initial vector x_0 by showing that $H_k \rightarrow O$ as $k \rightarrow \infty$, where $H_k = T_k T_{k-1} \cdots T_2 T_1$; see, e.g., [4].

The strength of these methods stems from having many zeros in the weighting matrices, indicating that only a small number of variables of $y_{\ell,j}$ in (1.2) need to be computed. This is why multisplittings developed into such a valuable tool for the analysis of block methods, with or without overlap. By overlap we mean that a variable received contributions from more than one processor, i.e., that the corresponding diagonal entry is nonzero in more than one weighting matrix; see, e.g., [5], [9], [10], [13].

In this note we present a convergence theorem for (nonstationary) multisplittings (section 2), where the condition (1.1) is not needed. We apply this general theorem to the case of s.p.d. matrices. Then, in section 3, we use this result to prove convergence of two-stage multisplittings for symmetric positive semidefinite linear systems.

2. Convergence with general weighting matrices. We begin with our general convergence result. To that end, consider *any matrix norm* such that the norm of the identity is equal to one. Thus, from the fact that $\sum_{\ell=1}^p E_{\ell} = I$, we have that

$$(2.1) \quad \sum_{\ell=1}^p \|E_{\ell}\| \geq 1.$$

Furthermore, if each splitting $A = M_{\ell} - N_{\ell}$ is convergent, i.e., if $\rho(M_{\ell}^{-1} N_{\ell}) < 1$, where ρ is the spectral radius, then, since $\lim_{k \rightarrow \infty} \|(M_{\ell}^{-1} N_{\ell})^k\| = 0$, given any positive

number $\eta < 1$ there is an integer $\tilde{s} = \tilde{s}(\eta)$ (which also depends on the chosen norm), so that

$$(2.2) \quad \|(M_\ell^{-1}N_\ell)^s\| \leq \eta \text{ for all } s \geq \tilde{s}, \ell = 1, \dots, p.$$

THEOREM 2.1. *Let A be nonsingular, and let every splitting $A = M_\ell - N_\ell$, $\ell = 1, \dots, p$, be convergent. Given a fixed positive number $\theta < 1$, let $\eta = \theta / (\sum_{\ell=1}^p \|E_\ell\|)$. Let \tilde{s} be such that (2.2) holds. If the sequence of number of local iterations satisfies $s(\ell, k) \geq \tilde{s}$, $\ell = 1, \dots, p$, $k = 1, 2, \dots$, then the nonstationary multisplitting Algorithm 1.1 converges to the solution of $Ax = b$ with convergence factor θ .*

Proof. From (1.3) it follows that

$$\|T_k\| \leq \sum_{\ell=1}^p \|E_\ell\| \|(M_\ell^{-1}N_\ell)^{s(\ell,k)}\| \leq \left(\max_{\ell} \|(M_\ell^{-1}N_\ell)^{s(\ell,k)}\| \right) \sum_{\ell=1}^p \|E_\ell\| \leq \theta < 1.$$

Thus $\|H_k\| = \|T_k T_{k-1} \cdots T_2 T_1\| \leq \theta^k \rightarrow 0$ as $k \rightarrow \infty$. \square

It follows from Theorem 2.1 that even if the standard multisplitting algorithm ($s(\ell, k) = 1$) does not converge, the price to pay for convergence is more local iterations. Furthermore, we can have convergence as fast as desired, i.e., we can prescribe a smaller convergence factor θ , and obtain the desired convergence by performing more local iterations to satisfy the corresponding condition (2.2).

Of course, we do not know a priori how many local iterations are needed for condition (2.2) to hold, and thus, Theorem 2.1 can be seen more as a theoretical result than a computational tool. On the other hand, Theorem 2.1 implies that one can experiment by increasing a value of \tilde{s} , until convergence is achieved. In fact, our discussion before Theorem 2.1 guarantees that such \tilde{s} exists.

We remark that unlike some results in the literature, here we do not need that the sequence $s(\ell, k)$ go to infinity; see [4] and the references therein. On the contrary, all we need is that this sequence be bounded from below by \tilde{s} defined by (2.2). We also mention that similar theorems exist, with the sequence either going to infinity or bounded for two-stage iterative methods; see [11], [17].

In the remainder of this section we apply Theorem 2.1 specifically to the s.p.d. case, using the A -norm, for which (2.1) holds.

If a matrix A is s.p.d., it induces a vector norm $\|x\|_A = (x^T A x)^{1/2}$. A splitting $A = M - N$ of A is called P -regular if $M^T + N$ is positive definite [19]. The following characterization can be found in [2], [12], or [22].

THEOREM 2.2. *Let A be s.p.d. A splitting $A = M - N$ is P -regular if and only if $\|M^{-1}N\|_A < 1$.*

Thus, condition (2.2) can be easily satisfied for P -regular splittings of an s.p.d. matrix.

COROLLARY 2.3. *Let $A = M_\ell - N_\ell$ be P -regular splittings of the symmetric positive definite matrix A , $\ell = 1, \dots, p$. Given a fixed positive number $\theta < 1$, let $\eta = \theta / (\sum_{\ell=1}^p \|E_\ell\|_A)$. Let \tilde{s} be such that $\|(M_\ell^{-1}N_\ell)^s\|_A \leq \eta$, for all $s \geq \tilde{s}$, $\ell = 1, \dots, p$. If the sequence of number of local iterations satisfies $s(\ell, k) \geq \tilde{s}$, $\ell = 1, \dots, p$, $k = 1, 2, \dots$, then the nonstationary multisplitting Algorithm 1.1 converges to the solution of $Ax = b$ with convergence factor θ . Furthermore, we have that for each iteration k , $\|T_k\|_A \leq \theta < 1$.*

We emphasize that in Corollary 2.3 no condition is imposed on the weighting matrices other than adding to the identity, so that (2.1) holds. In other words, we do not have the restriction (1.1).

In Example 2.3 of [7], where there is no convergence, the two splittings are P -regular, and the smallest integer \tilde{s} for which $\|(M_\ell^{-1}N_\ell)^{\tilde{s}}\|_A < 1/\sum_{\ell=1}^p \|E_\ell\|_A$, is $\tilde{s} = 8$. Thus, for $s(\ell, k) \geq 8$, $\|T_k\|_A < 1$. We mention here also that if $s(1, k) = s(2, k) = 4$, $k = 1, 2, \dots$, then $\rho(T_k) < 1$, and this is the smallest integer for which this is true.

3. Two-stage multisplittings for symmetric positive semidefinite matrices. In this section we extend the convergence theory of multisplittings with general weighting matrices to two-stage multisplittings, when the coefficient matrix of the linear system is symmetric positive semidefinite (and in particular our result applies to the s.p.d. case). Here we assume that $s(\ell, k) = s(\ell)$, i.e., that the number of inner iterations may change from one (inner) splitting to another (or from block to block), but it is the same for all (outer) iterations; it is a stationary method.

In the nonsingular case, as in Theorem 2.1, convergence of an algorithm was shown by having the iteration matrix having norm less than one. Here, in the singular case, we consider a consistent linear system $Ax = b$. The iteration matrix T has spectral radius equal to one, and convergence to some solution is achieved when the iteration matrix T is convergent, i.e., when the limit $\lim_{k \rightarrow \infty} T^k$ exists; see, e.g., [2].

Let $A = M - N$ be the outer splitting, and let $M = F_\ell - G_\ell$, $\ell = 1, \dots, p$.

ALGORITHM 3.1 (TWO-STAGE MULTISPLITTING). *Given an initial vector x_0 , and a sequence of numbers of inner iterations $s(\ell)$, $\ell = 1, \dots, p$.*

For $k = 1, 2, \dots$, until convergence.

$$\begin{aligned} &\text{For } \ell = 1 \text{ to } p \\ &\quad y_{\ell,0} = x_{k-1} \\ &\quad \text{For } j = 1 \text{ to } s(\ell) \\ &\quad\quad F_\ell y_{\ell,j} = G_\ell y_{\ell,j-1} + N x_{k-1} + b \\ &\quad x_k = \sum_{\ell=1}^p E_\ell y_{\ell,s(\ell)} . \end{aligned}$$

For this two-stage multisplitting algorithm, it follows, e.g., as in [4], that the iteration matrix is

$$(3.1) \quad T = \sum_{\ell=1}^p E_\ell (F_\ell^{-1}G_\ell)^{s(\ell)} + \left(I - \sum_{\ell=1}^p E_\ell (F_\ell^{-1}G_\ell)^{s(\ell)} \right) M^{-1}N.$$

For our convergence proof we will use the following two results. The first can be found, e.g., in [1], and the second, e.g., in [2].

LEMMA 3.2. *Given a nonsingular matrix A and a matrix T such that $(I - T)^{-1}$ exists, there exists a unique pair of matrices P and Q , P nonsingular, such that $A = P - Q$ and $T = P^{-1}Q$. The matrices are $P = A(I - T)^{-1}$ and $Q = P - A$.*

THEOREM 3.3. *Let $A = M - N$ be a P -regular splitting of a symmetric matrix A . Then the matrix $M^{-1}N$ is convergent if and only if A is positive semidefinite.*

THEOREM 3.4. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that M is a s.p.d. matrix and N is a positive semidefinite matrix. Let $M = F_\ell - G_\ell$, $\ell = 1, \dots, p$, be P -regular splittings. Given a fixed positive number $\theta < 1$, let $\eta = \theta / (\sum_{\ell=1}^p \|E_\ell\|_M)$. Let \tilde{s} be such that*

$$(3.2) \quad \|(F_\ell^{-1}G_\ell)^s\|_M \leq \eta \quad \text{for all } s \geq \tilde{s}, \ell = 1, \dots, p.$$

If the numbers of inner iterations satisfies $s(\ell) \geq \tilde{s}$, $\ell = 1, \dots, p$, then the two-stage multisplitting Algorithm 3.1 converges to a solution of the consistent linear system $Ax = b$ for any initial vector x_0 .

Proof. All we need to prove is that T defined in (3.1) is convergent. Observe first that the matrix $S = \sum_{\ell=1}^p E_{\ell}(F_{\ell}^{-1}G_{\ell})^{s(\ell)}$ can be viewed as the iteration matrix of a nonstationary multisplitting method based on the splittings $M = F_{\ell} - G_{\ell}$ and $s(\ell, k) = s(\ell)$, $k = 1, 2, \dots$, $\ell = 1, \dots, p$; cf. (1.3). Furthermore, since $M = F_{\ell} - G_{\ell}$, $\ell = 1, \dots, p$ are P -regular splittings of the s.p.d. matrix M , from Corollary 2.3 it follows that $\|S\|_M < 1$. Then, using Lemma 3.2, this iteration matrix induces a unique splitting $M = P - Q$ such that $P^{-1}Q = S$. Moreover, by Theorem 2.2, this splitting is P -regular. Thus, with these matrices, we have

$$\begin{aligned} T &= P^{-1}Q + (I - P^{-1}Q)M^{-1}N \\ &= P^{-1}(Q + (P - Q)M^{-1}N) = P^{-1}(Q + N). \end{aligned}$$

Thus, the splitting $A = P - (Q + N)$ is a splitting induced by T (this splitting is not unique [1]). Since $P^T + Q$ is positive definite and N is positive semidefinite, $P^T + Q + N$ is positive definite, and thus this splitting is P -regular. Therefore, by Theorem 3.3, T is a convergent matrix and the proof is complete. \square

We mention that the second part of the proof of Theorem 3.4 resembles the proof of Theorem 2.1 of [14], although the context is different.

As in Corollary 2.3, no special condition is imposed on the weighting matrices. Instead we may need to increase the number of inner iterations so that (3.2) holds.

Acknowledgments. We thank Michele Benzi for helpful comments on an early manuscript. We also thank the referees for their suggestions. These comments and suggestions helped improve our presentation.

REFERENCES

- [1] M. BENZI AND D. B. SZYLD, *Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods*, Numer. Math., 76 (1997), pp. 309–321.
- [2] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, 3rd ed., Academic Press, New York, 1979. Reprinted by SIAM, Philadelphia, 1994.
- [3] R. BRU, L. ELSNER, AND M. NEUMANN, *Models of parallel chaotic iteration methods*, Linear Algebra Appl., 103 (1988), pp. 175–192.
- [4] R. BRU, V. MIGALLÓN, J. PENADÉS, AND D. B. SZYLD, *Parallel, synchronous and asynchronous two-stage multisplitting methods*, Electron. Trans. Numer. Anal., 3 (1995), pp. 24–38.
- [5] Z.-H. CAO, *Nonstationary two-stage multisplitting methods with overlapping blocks*, Linear Algebra Appl., 285 (1998), pp. 153–163.
- [6] Z.-H. CAO AND Z.-Y. LIU, *Symmetric multisplitting of a symmetric positive definite matrix*, Linear Algebra Appl., 285 (1998), pp. 309–319.
- [7] M. J. CASTEL, V. MIGALLÓN, AND J. PENADÉS, *Convergence of non-stationary multisplitting methods for Hermitian positive definite matrices*, Math. Comp., 67 (1998), pp. 209–220.
- [8] J.-J. CLIMENT AND C. PEREA, *Convergence and comparison theorems for multisplittings*, Numer. Linear Algebra Appl., 6 (1999), pp. 93–107.
- [9] A. FROMMER AND B. POHL, *A comparison result for multisplittings and waveform relaxation methods*, Numer. Linear Algebra Appl., 2 (1995), pp. 335–346.
- [10] A. FROMMER, H. SCHWANDT, AND D. B. SZYLD, *Asynchronous weighted additive Schwarz methods*, Electron. Trans. Numer. Anal., 5 (1997), pp. 48–61.
- [11] A. FROMMER AND D. B. SZYLD, *Asynchronous two-stage iterative methods*, Numer. Math., 69 (1994), pp. 141–153.
- [12] A. FROMMER AND D. B. SZYLD, *Weighted max norms, splittings, and overlapping additive Schwarz iterations*, Numer. Math., 83 (1999), pp. 259–278.
- [13] M. T. JONES AND D. B. SZYLD, *Two-stage multisplitting methods with overlapping blocks*, Numer. Linear Algebra Appl., 3 (1996), pp. 113–124.
- [14] V. MIGALLÓN AND J. PENADÉS, *Convergence of two-stage iterative methods for Hermitian positive definite matrices*, Appl. Math. Lett., 10 (1997), pp. 79–83.
- [15] R. NABBEN, *A note on comparison theorems of splittings and multisplittings of Hermitian positive definite matrices*, Linear Algebra Appl., 233 (1996), pp. 67–80.

- [16] M. NEUMANN AND R. J. PLEMMONS, *Convergence of parallel multisplitting iterative methods for M -matrices*, *Linear Algebra Appl.*, 88/89 (1987), pp. 559–573.
- [17] N. K. NICHOLS, *On the convergence of two-stage iterative processes for solving linear equations*, *SIAM J. Numer. Anal.*, 10 (1973), pp. 460–469.
- [18] D. P. O'LEARY AND R. E. WHITE, *Multisplittings of matrices and parallel solution of linear systems*, *SIAM J. Algebraic Discrete Methods*, 6 (1985), pp. 630–640.
- [19] J. M. ORTEGA, *Numerical Analysis, A Second Course*, Academic Press, New York, 1972. Reprinted by SIAM, Philadelphia, 1990.
- [20] D. B. SZYLD AND M. T. JONES, *Two-stage and multisplitting methods for the parallel solution of linear systems*, *SIAM J. Matrix Anal. Appl.*, 13 (1992), pp. 671–679.
- [21] R. E. WHITE, *Multisplitting of a symmetric positive definite matrix*, *SIAM J. Matrix Anal. Appl.*, 11 (1990), pp. 69–82.
- [22] D. M. YOUNG, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.