Thirteenth International Conference on Domain Decomposition Methods
Editors: N. Debit, M.Garbey, R. Hoppe, J. Périaux, D. Keyes, Y. Kuznetsov

37 An Algebraic Convergence Theory for Restricted Additive and Multiplicative Schwarz Methods

A. Frommer[1], R. Nabben[2], D. B. Szyld[3]

Introduction

In this contribution we use the algebraic representation recently developed for the classical additive and multiplicative Schwarz methods in [FS99, BFNS01] to analyze the *restrictive additive Schwarz* (RAS) and *restrictive multiplicative Schwarz* (RMS) methods; see [CS96, CFS98, CS99, QV99].

RAS was introduced in [CS99] as an efficient alternative to the classical additive Schwarz preconditioner. Practical experiments have proven RAS to be particularly attractive, because it reduces communication time while maintaining the most desirable properties of the classical Schwarz methods [CFS98, CS99]. RAS preconditioners are widely used in practice and are the default preconditioner in the PETSc software package [BGMS97]. Similar savings in communication time can be expected in the case of RMS; see [CS96]. In fact, we announce here that we can prove that RMS is better than RAS, in the sense that the corresponding iteration matrix has a smaller norm, for a certain weighted max norm.

Our results provide the theoretical underpinnings for the behavior of the RAS preconditioners as observed in [CS99]. The theory we develop is not complete in the sense that we do not get quantitative results (like mesh independence in the presence of a coarse grid, for example). However, such results can be obtained indirectly by using some of the comparison results of [FS01] and classical results for the usual Schwarz method.

Our approach is purely algebraic, and therefore our results apply to discretization of differential equations as well as to algebraic additive Schwarz. We believe that the algebraic tools used here and in [FS99, BFNS01] complement the usual analytic tools used for the analysis of Schwarz methods; see, e.g., the books [SBG96, QV99] and the extensive bibliography therein.

One of the reasons why the algebraic approach presented here is a good alternative to the classical approach is that the operators defining RAS and RMS are not orthogonal projections (see [FS01]), and thus the usual theory as described, e.g., in [BM91] does not apply.

This paper is organized as follows. We start by giving algebraic representations of the usual and the restricted additive Schwarz methods and we introduce the splittings associated with each of the methods. Then, our convergence theorem for RAS, as well as results on the effect of overlap on the quality of the preconditioner are presented. Finally, convergence of RMS is shown, together with the comparison between RMS and RAS.

We note that using the same formulation described in this paper, several variants of RAS

[1]Fachbereich Mathematik, Bergische Universität GH Wuppertal, Gauss-Strasse 20, D–42097 Wuppertal, Germany, frommer@math.uni-wuppertal.de

[2]Fakultät für Mathematik, Universität Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany, nabben@mathematik.uni-bielefeld.de

[3]Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, Pennsylvania 19122-6094, USA, szyld@math.temple.edu . Supported by the U.S. National Science Foundation grant DMS-9973219

and RMS preconditioners can be analyzed, including the cases of inexact local solutions and of weighted methods; see [FS01, NS01].

The algebraic representation

The $n \times n$ linear system is given as

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

As in [CS99] we consider p nonoverlapping subspaces $W_{i,0},\ i = 1, \dots, p$ which are spanned by columns of the $n \times n$ identity I and which are then augmented to produce overlap. For a precise definition, let $S = \{1, \dots, n\}$ and let

$$S = \bigcup_{i=1}^{p} S_{i,0}$$

be a partition of S into p disjoint, non-empty subsets. For each of these sets $S_{i,0}$ we consider a nested sequence of larger sets $S_{i,\delta}$ with

$$S_{i,0} \subseteq S_{i,1} \subseteq S_{i,2} \ldots \subseteq S = \{1, \dots, n\}, \tag{2}$$

so that we again have $S = \cup_{i=1}^{p} S_{i,\delta}$ for all values of δ, but for $\delta > 0$ the sets $S_{i,\delta}$ are not necessarily pairwise disjoint, i.e., we have introduced *overlap*. A common way to obtain the sets $S_{i,\delta}$ is to add those indices to $S_{i,0}$ which correspond to nodes lying at distance δ or less from those nodes corresponding to $S_{i,0}$ in the (undirected) graph of A.

Let $n_{i,\delta} = |S_{i,\delta}|$ denote the cardinality of the set $S_{i,\delta}$. For each nested sequence from (2) we can find a permutation π_i on $\{1, \dots, n\}$ with the property that for all $\delta \geq 0$ we have $\pi_i(S_{i,\delta}) = \{1, \dots, n_{i,\delta}\}$.

We now build $n_{i,\delta} \times n$ matrices whose rows are precisely those rows j of the identity for which $j \in S_{i,\delta}$. Formally, such a matrix $R_{i,\delta}$ can be expressed as

$$R_{i,\delta} = [I_{i,\delta} | O]\, \pi_i \tag{3}$$

with $I_{i,\delta}$ the identity on the $n_{i,\delta}$-space. Finally, we define the $n \times n$ weighting matrices

$$E_{i,\delta} = R_{i,\delta}^T R_{i,\delta} \quad \left(= \pi_i^T \begin{bmatrix} I_{i,\delta} & O \\ O & O \end{bmatrix} \pi_i \right)$$

and the subspaces

$$W_{i,\delta} = \text{range}(E_{i,\delta}),\ \ i = 1, \dots, p.$$

Note the inclusion $W_{i,\delta} \supseteq W_{i,\delta'}$ for $\delta \geq \delta'$, and in particular $W_{i,\delta} \supseteq W_{i,0}$ for all $\delta \geq 0$.

We view the matrices $R_{i,\delta}$ as restriction operators and $R_{i,\delta}^T$ as prolongations. We can identify the image of $R_{i,\delta}^T$ with the subspace $W_{i,\delta}$. For each subspace $W_{i,\delta}$ we define a restriction of the operator A on $W_{i,\delta}$ as

$$A_{i,\delta} = R_{i,\delta} A R_{i,\delta}^T.$$

The classical additive Schwarz method consists of using the following preconditioner in a Krylov subspace method for solving (1):

$$M_{AS,\delta}^{-1} = \sum_{i=1}^{p} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta}. \tag{4}$$

In order to describe the *restricted* additive Schwarz method we introduce 'restricted' $n_{i,\delta} \times n$ operators $\widetilde{R}_{i,\delta}$ as

$$\widetilde{R}_{i,\delta} = R_{i,\delta} E_{i,0} \tag{5}$$

The image of $\widetilde{R}_{i,\delta}^T = E_{i,0} R_{i,\delta}^T$ can be identified with $W_{i,0}$, so $\widetilde{R}_{i,\delta}^T$ 'restricts' $R_{i,\delta}^T$ in the sense that the image of the latter, $W_{i,\delta}$, is restricted to its subspace $W_{i,0}$, the space from the non-overlapping decomposition. The restricted additive Schwarz method from [CFS98, CS99] replaces the prolongation operator $R_{i,\delta}^T$ by $\widetilde{R}_{i,\delta}^T$ and thus uses

$$M_{RAS,\delta}^{-1} = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} \tag{6}$$

instead of (4)[4]. For practical parallel implementations, replacing $R_{i,\delta}^T$ by $\widetilde{R}_{i,\delta}^T$ means that the corresponding part of the computation will not require any communication, since the images of the $\widetilde{R}_{i,\delta}^T$ do not overlap. In addition, the numerical results in [CS99] indicate that the restrictive additive Schwarz method is at least as fast (in terms of number of iterations and/or CPU time) as the classical one. Note that we lose symmetry, however, since if A is symmetric, $M_{AS,\delta}^{-1}$ will be symmetric as well, whereas $M_{RAS,\delta}^{-1}$ will usually be nonsymmetric.

For the convergence analysis of these Krylov methods, the relevant matrices are $M_{AS,\delta}^{-1}A$ and $M_{RAS,\delta}^{-1}A$. Alternatively, we can consider the iteration matrices $T_{AS,\delta} = I - M_{AS,\delta}^{-1}A$ and $T_{RAS,\delta} = I - M_{RAS,\delta}^{-1}A$. To analyze these matrices, we write the orthogonal projections

$$P_{i,\delta} = R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A,\ i = 1, \ldots, p$$

and the oblique projections

$$Q_{i,\delta} = \tilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A,\ i = 1, \ldots, p,$$

and thus we have the representation

$$T_{AS,\delta} = I - \sum_{i=1}^{p} P_{i,\delta},\ \ T_{RAS,\delta} = I - \sum_{i=1}^{p} Q_{i,\delta}. \tag{7}$$

With this notation, the iteration matrix corresponding to the classical multiplicative Schwarz method is

$$T_{MS} = (I - P_{p,\delta})(I - P_{p-1,\delta}) \cdots (I - P_{1,\delta}) \tag{8}$$

[4]We note that the representations (4) and (6) using rectangular matrices $R_{i,\delta}$ and matrices $A_{i,\delta}$ of smaller size is consistent with the standard literature [SBG96, QV99] and different than that of [CS99] where $n \times n$ matrices are used.

and the corresponding iteration matrix for the RMS method is

$$T_{RMS} = (I - Q_{p,\delta})(I - Q_{p-1,\delta}) \cdots (I - Q_{1,\delta}) \tag{9}$$

As in [FS99, BFNS01], the key to our analysis is the use of the nonsingular matrices $M_{i,\delta}$ defined as

$$M_{i,\delta} = \pi_i^T \begin{bmatrix} A_{i,\delta} & O \\ O & D_{\neg i,\delta} \end{bmatrix} \pi_i$$

where $D_{\neg i,\delta}$ is the diagonal part of the principal submatrix of A 'complementary' to $A_{i,\delta}$, i.e.,

$$D_{\neg i,\delta} = \operatorname{diag}\left([O|I_{\neg i,\delta}] \cdot \pi_i \cdot A \cdot \pi_i^T \cdot [O|I_{\neg i,\delta}]^T\right)$$

with $I_{\neg i,\delta}$ the $n - n_{i,\delta} \times n - n_{i,\delta}$ identity. Here, we assume that $A_{i,\delta}$ and $D_{\neg i,\delta}$ are nonsingular. With these matrices we can write

$$P_{i,\delta} = E_{i,\delta} M_{i,\delta}^{-1} A, \tag{10}$$

$$Q_{i,\delta} = E_{i,0} M_{i,\delta}^{-1} A, \tag{11}$$

and this provides a new representation of the matrices (7), (8), and (9); see [FS99, FS01, BFNS01, NS01]. The new representation of the additive Schwarz methods is very much in the spirit of multisplittings; see [OW85], or [BMPS95] and its bibliography.

We note that with the RAS preconditioning the corresponding weighting matrices satisfy

$$\sum_{i=1}^{p} E_{i,0} = I,$$

consistent with the traditional multisplitting theory, while for additive Schwarz we have

$$qI \geq \sum_{i=1}^{p} E_{i,\delta} \geq I,$$

where the inequalities are componentwise and

$$q = \max_{j=1,\dots,n} |\{i : j \in S_{i,\delta}\}|. \tag{12}$$

In the p.d.e. setting, q is the maximum number of subdomains to which each node of the mesh belongs.

Convergence of RAS

We show in this section that for M-matrices the spectral radius $\rho(I - M_{RAS,\delta}^{-1}A)$ of the RAS iteration matrix is less than 1 for all values of $\delta \geq 0$. This implies in particular that the spectrum of the preconditioned system $\sigma(M_{RAS,\delta}^{-1}A)$ is located in the right half plane and contained in a disk of radius less than one around the point 1.

We start by recalling some basic terminology. The natural partial ordering $\leq$ between matrices $A = (a_{ij}),\ B = (b_{ij})$ of the same size is defined component-wise, i.e., $A \leq B$ iff

$a_{ij} \leq b_{ij}$ for all i, j. If $A \geq O$ we call A nonnegative. If all entries of A are positive, we say that A is positive and write $A > O$. This notation and terminology carries over to vectors as well. An $n \times n$ matrix A is called a (nonsingular) M-matrix if it has nonpositive off-diagonal elements and $A^{-1} \geq O$; see [Var62].

Consider the splitting $A = M - N$ with M nonsingular. This splitting is said to be *weak nonnegative of the first type* (also called *weak regular*) if

$$M^{-1} \geq O \text{ and } M^{-1}N \geq O. \tag{13}$$

Theorem 1 [OR70] *Let $A = M - N$ be a weak nonnegative splitting of the first type. Then $\rho(I - M^{-1}A) < 1$ iff A is nonsingular and $A^{-1} \geq 0$.*

We are now able to formulate the central convergence result of this section.

Theorem 2 *Let A be a nonsingular M-matrix. Then for each value of $\delta \geq 0$, the splitting $A = M_{RAS,\delta} - N_{RAS,\delta}$, corresponding to the RAS method, is weak nonnegative of the first type. In particular, the iteration matrix $M_{RAS,\delta}^{-1} N_{RAS,\delta} = I - M_{RAS,\delta}^{-1} A$ satisfies*

$$\rho(I - M_{RAS,\delta}^{-1} A) < 1. \tag{14}$$

The proof consists of showing that $M_{RAS,\delta}^{-1} \geq O$, and that $I - M_{RAS,\delta}^{-1} A \geq O$, as per (13), and then apply Theorem 1; see [FS01].

We point out that in general a convergence result such as (14) does not hold for the classical additive Schwarz preconditioner (4). To guarantee convergence, a damping (or relaxation) parameter $\theta > 0$ is introduced. It can be shown that if $\theta \leq 1/q$, then $\rho(I - \theta M_{AS,\delta}^{-1} A) < 1$, where q is defined in (12); see [FS99, BFNS01]. Thus, one of the attractive features of the RAS preconditioner is that no damping parameter is needed for convergence.

Using an appropriate norm, we study the effect of varying the overlap. More precisely, we prove comparison results on the spectral radii and/or on certain weighted max norms for the corresponding iteration matrices $T_{RAS,\delta}$ as defined in (7) for different values of $\delta \geq 0$.

We want to compare one RAS splitting, defined through the sets $S_{i,\delta'}$ with another one with more overlap defined through sets $S_{i,\delta}$ where $S_{i,\delta'} \subseteq S_{i,\delta}, i = 1, \ldots, p$. We show that the larger the overlap ($\delta \geq \delta'$), the faster RAS method converges as measured in certain weighted max norms. This is consistent with the experiments in Tables 1 and 2 of [CS99], where an increase of the overlap is associated with fewer iterations.

For a positive vector w we denote $\|x\|_w$ the weighted max norm in n-space given by

$$\|x\|_w = \max_{i=1,\ldots,n} |x_i|/w_i.$$

The resulting operator norm in $n \times n$-space is denoted similarly.

The following theorem from [FS01] is very similar to [FP95, Theorem 2.1].

Theorem 3 *Let A be a nonsingular M-matrix and let $w > 0$ be any positive vector such that $Aw > 0$, e.g., $w = A^{-1}v$ with $v > 0$. Then, if $\delta \geq \delta'$,*

$$\|T_{RAS,\delta}\|_w \leq \|T_{RAS,\delta'}\|_w. \tag{15}$$

Moreover, if the Perron vector $w_{\delta'}$ of $T_{RAS,\delta'}$ satisfies $w_{\delta'} > 0$ and $Aw_{\delta'} \geq 0$, then we also have

$$\rho(T_{RAS,\delta}) \leq \rho(T_{RAS,\delta'}). \tag{16}$$

In the case that (16) holds, Theorem 3 indicates that the spectrum of the preconditioned matrix is included in a possibly smaller disk when the overlap is increased.

We remark that (15) (as well as most results using the weighted max norms in the paper) holds for *any* positive vector w such that Aw is positive, so that one has a lot of freedom in choosing the norm. For example, if all row-sums of A are positive we can choose as w the vector of all ones, and thus the weighted max norm is simply the max norm. A commonly chosen vector w is the row-sums of A^{-1}, which is always positive.

For $\delta' = 0$, i.e., for the block Jacobi preconditioner we can always provide the comparison of the spectral radii (16), in addition to the comparison (15). The following theorem is in fact [FP95, Theorem 2.2].

Theorem 4 *Let A be a nonsingular M-matrix. Then, for any value of $\delta \geq 0$,*

$$\rho(T_{RAS,\delta}) \leq \rho(T_{RAS,0}) \ .$$

Convergence of RMS

Using the new algebraic representation (10), it was shown in [BFNS01] that for any $w = A^{-1}e > 0$ with $e > 0$, we have $\rho(T_{MS}) \leq \|T_{MS}\|_w < 1$. In a similar way, using the representation (11), we can prove the following result; see [NS01].

Theorem 5 *Let A be a nonsingular M-matrix. For any $w = A^{-1}e > 0$ with $e > 0$, we have $\rho(T_{RMS}) \leq \|T_{RMS}\|_w < 1$. Furthermore, there exists a unique splitting $A = B - C$ such that $T_{MRS} = B^{-1}C$, and this splitting is weak nonnegative of the first type.*

It is well known that bounds for the convergence using the standard multiplicative Schwarz preconditioner are better than those obtained for the standard additive Schwarz; see, e.g. [SBG96, QV99]. For the restrictive preconditioners we can actually show that the weighted max norm of the RMS iteration matrix is smaller than that of RAS.

Theorem 6 *Let A be a nonsingular M-matrix and let $w > 0$ be any positive vector such that $Aw > 0$, e.g., $w = A^{-1}e$ with $e > 0$. Then,*

$$\|T_{RMS,\delta}\|_w \leq \|T_{RAS,\delta}\|_w \ .$$

Moreover, if the Perron vector w_δ of $T_{RAS,\delta}$ satisfies $w_\delta > 0$ and $Aw_\delta \geq 0$, then we also have

$$\rho(T_{RMS,\delta}) \leq \rho(T_{RAS,\delta}) \ .$$

The proof consists of showing that $M_{RMS,\delta}^{-1} \geq M_{RAS,\delta}^{-1}$, where $M_{RMS,\delta} = (I - T_{RMS,\delta})^{-1}A$. This inequality together with theorems 2 and 5, and Theorem 4.1 of [FS99] provides the needed norm and spectral radii inequalities; see [NS01].

As is the case for RAS, one can also show that by increasing the overlap, the weighted max norm of the iteration matrix decreases, i.e., that if $\delta \geq \delta'$,

$$\|T_{RMS,\delta}\|_w \leq \|T_{RMS,\delta'}\|_w < 1$$

for any $w > 0$ such that $Aw > 0$. Furthermore, it can be shown that overlap is always better than no overlap, i.e., for any value of $\delta \geq 0$,

$$\rho(T_{RMS,\delta}) \leq \rho(T_{RMS,0}) \ .$$

References

[BFNS01]Michele Benzi, Andreas Frommer, Reinhard Nabben, and Daniel B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89:605–639, 2001.

[BGMS97]Satish Balay, William D. Gropp, Louis C. McInnes, and Barry F. Smith. *PETSc 2.0 User Manual*. Argonne National Laboratory, http://www.mcs.anl.gov/petsc/, 1997.

[BM91]Petter E. Bjørstad and Jan Mandel. On the spectra of sums of orthogonal projections with applications to parallel computing. *BIT*, 31:76–88, 1991.

[BMPS95]Rafael Bru, Violeta Migallón, José Penadés, and Daniel B. Szyld. Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Transactions on Numerical Analysis*, 3:24–38, 1995.

[CFS98]X.-C. Cai, C. Farhat, and M. Sarkis. A minimum overlap restricted additive Schwarz preconditioner and applications to 3D flow simulations. *Contemporary Mathematics*, 218:479–485, 1998.

[CS96]Xiao-Chuan Cai and Youcef Saad. Overlapping domain decomposition algorithms for general sparse matrices. *Numerical Linear Algebra with Applications*, 3:221–237, 1996.

[CS99]Xiao-Chuan Cai and Marcus Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21:239–247, 1999.

[FP95]Andreas Frommer and Bert Pohl. A comparison result for multisplittings and waveform relaxation methods. *Numerical Linear Algebra with Applications*, 2:335–346, 1995.

[FS99]Andreas Frommer and Daniel B. Szyld. Weighted max norms, splittings, and overlapping additive schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.

[FS01]Andreas Frommer and Daniel B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. *SIAM Journal on Numerical Analysis*, 39:463–479, 2001.

[NS01]Reinhard Nabben and Daniel B. Szyld. Convergence theory of restrictive multiplicative Schwarz methods. Technical Report 01-05-17, Department of Mathematics, Temple University, Philadelphia, Pa., May 2001.

[OR70]James M. Ortega and Werner C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York and London, 1970.

[OW85]D. P. O'Leary and R.E. White. Multi-splittings of matrices and parallel solution of linear systems. *SIAM J. Alg. Disc. Meth.*, 6:630–640, 1985.

[QV99]Alfio Quarteroni and Alberto Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, 1999.

[SBG96]Barry F. Smith, Petter E. Bjørstad, and William Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, 1996.

[Var62]Richard S. Varga. *Matrix Iterative Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1962. Second Edition, Springer, Berlin, 2000.