A Proposal for a Dynamically Adapted Inexact Additive Schwarz Preconditioner

Marcus Sarkis[1] and Daniel B. Szyld[2]

[1] Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, Brazil, and Worcester Polytechnic Institute, Worcester, MA 01609, USA.
msarkis@fluid.impa.br

[2] Department of Mathematics, Temple University, Philadelphia, PA 19122, USA.
szyld@math.temple.edu

1 Introduction

Additive Schwarz is a powerful preconditioner used in conjuction with Krylov subspace methods (e.g., GMRES [7]) for the solution of linear systems of equations of the form $Au = f$, especially those arising from discretizations of differential equations on a domain divided into p (overlapping) subdomains [5], [9], [10]. In this paper we consider right preconditioning, i.e., the equivalent linear system is $AM^{-1}w = f$, with $Mu = w$. The additive Schwarz preconditioner is

$$M^{-1} = \sum_{i=1}^{p} R_i^T A_i^{-1} R_i, \tag{1}$$

where R_i is a restriction operator and $A_i = R_i A R_i^T$ is a restriction of A to a subdomain. The strength of this preconditioner stems in part from having overlap between the subdomains, and in part by the efficiency of local solvers, i.e., solutions of the "local" problems

$$A_i x = R_i v. \tag{2}$$

We also consider a weighted additive Schwarz preconditioner with harmonic extension (WASH), a preconditioner in the family of restricted additive Schwarz (RAS) preconditioners [3] of the form

$$M^{-1} = \sum_{i=1}^{p} R_i^T A_i^{-1} R_i^{\omega}, \tag{3}$$

in which the restriction operator R_i^{ω} is such that all variables corresponding to a point in the overlap are weighted with weights that add up to one, i.e., $\sum_{i=1}^{p} R_i^T R_i^{\omega} = I$ [4].

In this paper we consider the case when the local problems are either too large or too expensive to be solved exactly. Thus, the systems (2) are solved using an iterative method. Usually, one takes a fixed number of (inner) iterations. We are interested instead in prescribing a certain (inner) tolerance so that the iterative method for the solution of (2) stops when the local residual

$$s_{i,k} = A_i x_j - R_i v_k$$

has norm below the inner tolerance ($j = j(i,k)$ being the index of the inner iteration, and we write $x_j = \tilde{A}_{i,k}^{-1} R_i v_k$, where the subscript in $\tilde{A}_{i,k}$ indicates that the inexact local solvers changes also with k). Inexact local solvers have been used extensively (see, e.g., [9]); what is new here is that the inexactness changes as the (outer) iterations proceed. In this case, the (global) preconditioner changes from step to step, i.e.,

$$M_k^{-1} = \sum_{i=1}^{p} R_i^T \tilde{A}_{i,k}^{-1} R_i, \tag{4}$$

and one needs to use a flexible Krylov subspace method, such as FGMRES [6].

Recent results have shown that it is possible to vary how inexact a preconditioner is without degradation of the overall performance of a Krylov method; see [1], [8] and references therein, and in particular we mention [2] where Schur complement methods were studied. More precisely, the preconditioned system has to be solved more exactly at first, while the exactness can be relaxed as the (outer) iterative method progresses. In this paper we propose to apply these new ideas to additive Schwarz preconditioning and its restricted variants, thus providing a way to dynamically choose the inner tolerance for the local solvers in each step k of the (outer) iterative method. Our proposed strategy is illustrated with numerical experiments, which show that there is a great potential in savings while maintaining the performance of the overall process.

2 Dynamic Stopping Criterion for the Local Solvers

The algorithmic setup is as follows, at each step k of the (outer) Krylov subspace method for the solution of $Au = f$ (we use FGMRES here), we apply a preconditioner of the form (4), where the symbol $\tilde{A}_{i,k}$ indicates that the solution of local problem (2) is approximated by a Krylov subspace method (we use GMRES) iterated until $\|s_{i,k}\| \leq \varepsilon_{i,k}$.

In this setup, at the kth iteration instead of the usual matrix-vector product $AM^{-1}v_k$ we have

$$AM_k^{-1} v_k = A \sum_{i=1}^{p} R_i^T \tilde{A}_{i,k}^{-1} R_i v_k$$

$$= A \sum_{i=1}^{p} R_i^T A_i^{-1} R_i v_k + A \sum_{i=1}^{p} R_i^T (\tilde{A}_{i,k}^{-1} - A_i^{-1}) R_i v_k$$
$$= A M^{-1} v_k + A \sum_{i=1}^{p} R_i^T A_i^{-1} s_{i,k}.$$

Thus, we can write $AM_k^{-1} v_k = (AM^{-1} + E_k) v_k$, where E_k is the inexactness of the preconditioned matrix at the kth step, and $f_k = E_k v_k = A \sum_{i=1}^{p} R_i^T A_i^{-1} s_{i,k}$, so that

$$\|f_k\| = \|E_k v_k\| \leq \sum_{i=1}^{p} \|A R_i^T A_i^{-1}\| \|s_{i,k}\|. \tag{5}$$

In the situation we are describing, namely of inexact preconditioner, the inexact Arnoldi relation that holds is

$$AV_m + [f_1, f_2, \ldots, f_m] = V_{m+1} H_{m+1,m},$$

where the $V_m = [v_1, v_2, \ldots, v_m]$ has orthonormal columns, and $H_{m+1,m}$ is upper Hessenberg. Let $W_m = V_{m+1} H_{m+1,m}$, and r_k be the GMRES (outer) residual at the kth step. It follows from [8, sections 4 and 5] that

$$\|W_m^T r_m\| \leq \kappa(H_{m+1,m}) \sum_{k=1}^{m} \|f_k\| \|r_{k-1}\|, \tag{6}$$
$$\|r_m - \tilde{r}_m\| \leq \frac{1}{\sigma_{min}(H_{m+1,m})} \sum_{k=1}^{m} \|f_k\| \|r_{k-1}\|, \tag{7}$$

where $\kappa(H_{m+1,m}) = \sigma_{max}(H_{m+1,m})/\sigma_{min}(H_{m+1,m})$ is the condition number of $H_{m+1,m}$, and $\tilde{r}_m = r_0 - V_{m+1} H_{m+1,m} y_m$ is the computed residual. In the exact case, i.e., when $\varepsilon_{i,k} = 0$, $i = 1, \ldots, p$, $k = 1, 2, \ldots$, then $W_m^T r_m = 0$. Equation (6) indicates how far from that optimal situation one may be. The residual gap (7) is the norm of the difference between the "true" residual $r_m = f - AV_m y_m$ and the computed one. As $\tilde{r}_m \to 0$, we have that if the right hand side of (7) is of order ε, then $\|r_m\| \to \mathcal{O}(\varepsilon)$; cf. [8, Figure 9.1].

Using (5) we obtain the following result.

Proposition 1. *If the local residuals satisfy* $\|s_{i,k}\| \leq \varepsilon_k$, $i = 1, \ldots, p$, *then the* k*th GMRES (outer) residual satisfies the following two relations:*

$$\|W_m^T r_m\| \leq \kappa(H_{m+1,m}) \sum_{i=1}^{p} \|A R_i^T A_i^{-1}\| \sum_{k=1}^{m} \varepsilon_k \|r_{k-1}\|, \tag{8}$$
$$\|r_m - \tilde{r}_m\| \leq \frac{1}{\sigma_{min}(H_{m+1,m})} \sum_{i=1}^{p} \|A R_i^T A_i^{-1}\| \sum_{k=1}^{m} \varepsilon_k \|r_{k-1}\|. \tag{9}$$

One can then conclude an *a posteriori* result.

Proposition 2. *If ε_k, the bound of the local residual norms, satisfy*

$$\varepsilon_k \leq K_m \frac{1}{\|r_{k-1}\|} \ \varepsilon, \tag{10}$$

with

$$K_m = 1/m\kappa(H_{m+1,m}) \sum_{i=1}^{p} \|AR_i^T A_i^{-1}\|, \tag{11}$$

then $\|W_m^T r_m\| \leq \varepsilon$, and if (10) *holds with*

$$K_m = \sigma_{min}(H_{m+1,m})/m \sum_{i=1}^{p} \|AR_i^T A_i^{-1}\|, \tag{12}$$

then $\|r_m - \tilde{r}_m\| \leq \varepsilon$.

We mention that these results apply to the case of inexact WASH preconditioning as well, where the restriction R_i on the right of each addend in (4) is replaced with R_i^{ω}.

3 Implementation Considerations

The power of Proposition 2 is to point out that one can relax the local residual norms in a way inversely proportional to the norm of the (outer or global) residual from the previous step; cf. [1], [8]. The constants K_m as stated in (11) and (12), which do not depend on k, depend in part on A, i.e., on the problem to be solved, the preconditioner, through the locals problems represented by A_i, as well as on how the inexact strategy is implemented, through $H_{m+1,m}$. Observe that since $m\kappa(H_{m+1,m}) \gg 1$ it is natural from (11) to expect $K_m \leq 1$.

Depending on the problem, on could obtain an *a priori* bound for K_m which would not depend on the specifics of the inexact strategy, for example by setting $\kappa(H_{m+1,m}) \approx \gamma\kappa(AM^{-1})$, for some fixed number γ, or similarly $\sigma_{min}(H_{m+1,m}) \approx \gamma\sigma_{min}(AM^{-1})$. While this may appear as an oversimplification, one is justified in part because the bounds (8) and (9) are very far from being tight.

In many problems though, the value of K_m may not be known in advance, or it may be hard to estimate, and one can just try some number, say 1, and decrease it until a good convergence behavior is achieved. One could also use the information from a first run, to estimate some value of K_m. In our preliminary experiments, reported in the next section, we have used the value of $K_m = 1$.

4 Numerical Experiments

We present numerical experiments on finite difference discretizations of two partial differential equations with Dirichlet boundary conditions on the two-dimensional unit square: the Laplacian $-\Delta u = f$, and a convection diffusion equation $-\Delta u + b.\nabla u = f$, with $b^T = [10, 20]$, where upwind differences are used, and f is random, uniformly distributed between 0 and 1. We use an uniform discretization in each direction of 128 points, so the matrices are of order 16129, i.e., 16129 nodes in the grid. We partition the grid into 8×8 subdomains. In Table 1 we report experiments with varying degree of overlap: no overlap (0), one or two lines of overlap (1,2). Our (global) tolerance is $\varepsilon = 10^{-6}$. We compare the performance of using a fixed inner tolerance in each local solve, $\varepsilon_k = 10^{-4}$ for $k = 1, \ldots$, with the dynamic choice (10) using $K = K_m = 1$. We remark that both of these strategies correspond to varying the degree of inexactness and are expressed by the preconditioner (4). We run our experiments with the Additive Schwarz preconditioner (4) (ASM) and with weighted additive Schwarz preconditioner with harmonic extension (WASH). We used a minimum of five (inner) iterations in each of the local solvers. We report the average number of inner iterations, which in this case well reflects the total work in each case, and in parenthesis the number of outer FGMRES iterations needed for convergence.

Table 1. Average number of inner iterations (and number of outer iterations). Fixed or dynamic inner tolerance ($K = 1$).

	problem	Laplacian			Convection Diffusion		
	overlap	0	1	2	0	1	2
ASM	Fixed 10^{-4}	1923(64)	1536(46)	1388(38)	1825(60)	1458(43)	1295(35)
	Dynamic	1557(73)	1316(60)	1201(53)	1762(66)	1434(51)	1288(44)
WASH	Fixed 10^{-4}	1692(56)	1317(40)	1100(31)	1601(53)	1220(37)	1020(29)
	Dynamic	1387(61)	1089(45)	948(38)	1570(56)	1216(40)	1060(35)

As it can be appreciated from Table 1, the proposed dynamic strategy for the inexact local solvers can reach the same (outer) tolerance using up to 20% less work. We point out that we have used the same value of $K = 1$ for all overlaps, although the preconditioners certainly change. A better estimate of K as a function of the overlap is expected to produce better results. We also mention that both the fixed inner tolerance and the dynamically chosen one usually require less storage than the exact local solvers (1) and (3).

Acknowledgments

The first author was supported in part by CNPQ (Brazil) under grant 305539/2003-8 and by the U.S. National Science Foundation under grant CGR 9984404. The second author was supported in part by the U.S. National Science Foundation under grant DMS-0207525.

References

1. A. Bouras and V. Frayseé. Inexact matrix-vector products in Krylov methods for solving linear systems: a relaxation strategy. *SIAM J. Matrix Anal. Appl.*, 26:660–678, 2005.
2. A. Bouras, V. Frayssé, and L. Giraud. A relaxation strategy for inner–outer linear solvers in domain decomposition methods. Technical Report TR/PA/00/17, CERFACS, Toulouse, France, 2000.
3. X.-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM J. Scientific Comput.*, 21:792–797, 1999.
4. A. Frommer and D. B. Szyld. Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numer. Math.*, 83:259–278, 1999.
5. A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, Clarendon Press, Oxford, 1999.
6. Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM J. Scientific Comput.*, 14:461–469, 1993.
7. Y. Saad and M. H. Schultz. GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM J. Scientific Stat. Comput.*, 7:856–869, 1986.
8. V. Simoncini and D. B. Szyld. Theory of Inexact Krylov Subspace Methods and Applications to Scientific Computing. *SIAM J. Scientific Comput.*, 25:454–477, 2003.
9. B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain decomposition: Parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press, Cambridge, 1996.
10. A. Toselli and O. B. Widlund. *Domain Decomposition: Algorithms and Theory*. Springer Series in Computational Mathematics v. 34. Springer, Heidelberg, Berlin, New York, 2004.