



Additive Schwarz with variable weights

C Greif, T Rees, DB Szyld

December 2012

©2012 Science and Technology Facilities Council

Enquiries about copyright, reproduction and requests for additional copies of this report should be addressed to:

RAL Library
STFC Rutherford Appleton Laboratory
R61
Harwell Oxford
Didcot
OX11 0QX

Tel: +44(0)1235 445384
Fax: +44(0)1235 446403
email: libraryral@stfc.ac.uk

Science and Technology Facilities Council reports are available online at: <http://epubs.stfc.ac.uk>

ISSN 1358- 6254

Neither the Council nor the Laboratory accept any responsibility for loss or damage arising from the use of information contained in any of their reports or in any communication about their tests or investigations.

Additive Schwarz with Variable Weights

Chen Greif¹, Tyrone Rees², and Daniel B. Szyld³

1 Introduction and Motivation

We consider the numerical solution of nonsymmetric linear systems of equations of the form

$$A\mathbf{u} = \mathbf{f}, \quad (1)$$

which arise, e.g., from the discretization of elliptic partial differential equations (PDEs). In practical problems, the number of mesh points is very large, and thus also the number of unknowns in (1), and the resulting matrix is large and sparse. In these circumstances, iterative methods are often used, due to their ability to deal more effectively with a high degree of sparsity. A popular iterative method is the *Generalized Minimum Residual* iterative scheme, or GMRES [5], [6], [7]. This method is based on minimizing at the k th iterate the residual within the shifted Krylov subspace

$$\mathbf{x}_0 + \mathcal{K}^k(A, \mathbf{r}_0),$$

where x_0 is an initial vector, $\mathbf{r}_0 = b - A\mathbf{x}_0$ is the initial residual, and

$$\mathcal{K}^k(A, \mathbf{r}_0) = \text{span}(\mathbf{r}_0, A\mathbf{r}_0, \dots, A^{k-1}\mathbf{r}_0).$$

The performance of GMRES is often (though not exclusively) determined by the structure of the eigenvalues of the matrix A . Loosely speaking, if the eigenvalues are strongly clustered, then GMRES is expected to converge fast. To accomplish a clustering effect, a *preconditioner* M is typically used: instead of solving (1) we solve, say,

$$AM\tilde{\mathbf{u}} = \mathbf{f},$$

where M is constructed so that AM has a more favorable eigenstructure than A . Upon incorporating a preconditioner, the Krylov subspace changes accordingly: the matrix associated with the subspace will now be AM , and the notion of preconditioned residual arises.

When numerically solving boundary value problems arising from elliptic PDEs, a common way of dealing with the large number of degrees of freedom when the mesh is fine, is to break the problem down into a number of more manageable sub-problems; this amounts to the technique of *domain decomposition*; see, e.g., [8]. We

Department of Computer Science, University of British Columbia, Vancouver, B.C., Canada. greif@cs.ubc.ca · Rutherford Appleton Laboratory, Chilton, Didcot, UK. tyrone.rees@stfc.ac.uk · Department of Mathematics, Temple University, Philadelphia, Pennsylvania, USA. szyld@temple.edu

can then incorporate preconditioners that work on the subdomains into the general iterative framework.

The additive Schwarz preconditioner [8] and its restricted variant (RAS) [2], can be written in the form

$$M = \sum_{i=1}^t \tilde{R}_i A_i^{-1} R_i^T,$$

where t is usually the number of subdomains, \tilde{R}_i is a restriction operator, R_i^T is a prolongation operator, and $A_i = R_i^T A R_i$ is the restriction of A onto the i th subdomain.

We propose in this paper to use a weighted additive or restricted additive Schwarz preconditioner of the form

$$M^{(k)} = \sum_{i=1}^t \alpha_i^{(k)} \tilde{R}_i A_i^{-1} R_i^T, \quad (2)$$

where the weights $\alpha_i^{(k)}$ are chosen at the k th iteration of GMRES so as to minimize the preconditioned residual, cf. [1]. We point out that this is completely different than the approach in [3], where the weights are zeros and ones, and the emphasis is on asynchronous iterations.

Incorporating weights which change from one iteration to the next is significant and we can no longer talk about a standard iterative method with a single preconditioner. Instead, the proposed preconditioner (2) fits into the class of preconditioners we recently described in [4], where one can apply more than one preconditioner simultaneously. This is called multipreconditioned GMRES (MPGMRES). Our main goal in this paper is to show that this methodology is particularly effective in the domain decomposition paradigm, since we can associate each subdomain with a specific, unique preconditioner. We have a rather large number of preconditioners, and domain decomposition naturally lends itself to the approach of multipreconditioned iterations.

An outline of the remainder of this paper follows. In Section 2 we briefly describe Additive and Restricted Additive Schwarz Preconditioning. In Section 3 we describe the MPGMRES algorithm. We address the question of computational cost of the algorithm and characterize the generalized Krylov subspace and its unique features in domain decomposition setting. In Section 4 we provide some details on numerical experiments. Finally, in Section 5 we make some concluding remarks.

2 Additive and Restricted Additive Schwarz Preconditioning

Suppose we divide the domain Ω containing N nodes into t subdomains $\Omega_1, \dots, \Omega_t$, which overlap by bands of width δ nodes. Suppose each subdomain consists of $m_i \ll N$ nodes, which we denote as the entries of the set I_i . We can define an extension matrix $R_{i,\delta}^T \in \mathbb{R}^{N \times m_i}$ which extends vectors $\mathbf{x}^{(i)} \in \mathbb{R}^{m_i}$ to \mathbb{R}^N by

$$(R_{i,\delta}^T \mathbf{x}^{(i)})_k = \begin{cases} (\mathbf{x}^{(i)})_k & \text{if } k \in I_i \\ 0 & \text{otherwise.} \end{cases}$$

The transpose of this matrix defines a restriction operator R_i which restricts vectors in \mathbb{R}^N to the subdomain Ω_i . The restriction of the discretized PDE, A , to the i th subdomain is given by $R_{i,\delta} A R_{i,\delta}^T$.

We can now define the *additive Schwarz* preconditioner as

$$M := \sum_{i=1}^t R_{i,\delta}^T (R_{i,\delta} A R_{i,\delta}^T)^{-1} R_{i,\delta} = \sum_{i=1}^t M_i,$$

where $M_i := R_{i,\delta}^T (R_{i,\delta} A R_{i,\delta}^T)^{-1} R_{i,\delta}$.

Comment: Note that, by the definition of $R_{i,\delta}^T$, there exists some permutation Π_i such that, for all \mathbf{x} ,

$$\Pi_i M_i \mathbf{x} = (\times \cdots \times 0 \cdots \cdots 0)^T.$$

i.e., the vector resulting from multiplication by the M_i (regardless of the permutation) will be sparse.

We can also define a *restricted additive Schwarz* preconditioner [4] by

$$M^{\text{RAS}} := \sum_{i=1}^t R_{i,0}^T (R_{i,\delta} A R_{i,\delta}^T)^{-1} R_{i,\delta} = \sum_{i=1}^t M_i^{\text{RAS}},$$

i.e., the preconditioning operation is as for additive Schwarz, but the final extension operator is replaced by that which would be applied if we didn't have any overlap in the domains. This has been shown to be more efficient than the standard additive Schwarz implementation.

3 The MPMRES Algorithm for Domain Decomposition Problems

MPMRES [4] is a minimal residual algorithm for solving a linear system of equations which allows the user to apply more than one preconditioner simultaneously. The multipreconditioned search directions are all combined into a generalized Krylov subspace, and the minimization procedure requires solving a linear least-squares problem in every iteration. As opposed to standard GMRES, here the subspace grows quickly due to the presence of multiple search directions, and the projection can be expressed in terms of a block upper Hessenberg matrix. It has been shown in [4] that a so-called *selective MPMRES* algorithm which chooses a subset of t search directions, performs quite well on a variety of test problems. See Algorithms 1–3.

Algorithm 1 MPMGMRES

```

calls Algorithm 2 or Algorithm 3
Choose  $\mathbf{x}_0$ ,  $\mathbf{r}_0 = \mathbf{b} - \mathcal{A}\mathbf{x}_0$ 
 $\beta = \|\mathbf{r}_0\|$ ,  $\mathbf{v}_1 = \mathbf{r}_0/\beta$ 
 $Z_1 = \text{completeMPD}(\mathbf{r}_0)$ 
for  $k = 1, \dots$ , until convergence do
   $W = \mathcal{A}Z_k$ 
  for  $j = 1, \dots, k$  do
     $H_{j,k} = (V_j)^T W$ 
     $W = W - V_j H_{j,k}$ 
  end for
   $W = V_{k+1} H_{k+1,k}$  (skinny QR factorization)
   $\mathbf{y}_k = \text{argmin} \|\beta \mathbf{e}_1 - \tilde{H}_k \mathbf{y}\|_2$ 
   $\mathbf{x}_k = \mathbf{x}_0 + [Z_1 \cdots Z_k] \mathbf{y}_k$ 
   $Z_{k+1} = \begin{cases} \text{completeMPD}(V_{k+1}) & \text{for complete MPMGMRES} \\ \text{selectiveMPD}(V_{k+1}) & \text{for selective MPMGMRES} \end{cases}$ 
end for

```

Algorithm 2 Subroutine: complete multipreconditioned directions

```

function  $Z = \text{completeMPD}(V)$ 
 $Z = [\mathcal{P}_1^{-1} V \cdots \mathcal{P}_t^{-1} V]$ 
end function

```

Algorithm 3 Subroutine: selective multipreconditioning directions

```

function  $Z = \text{selectiveMPD}(V)$ 
 $\mathbf{v} = V \mathbf{1}$ 
 $Z = [\mathcal{P}_1^{-1} \mathbf{v} \cdots \mathcal{P}_t^{-1} \mathbf{v}]$ 
end function

```

3.1 Computational Work

In the selective algorithm we need t matrix-vector products and t preconditioner solves per iteration, as opposed to one for both in the standard preconditioned GMRES algorithm. The main other source for work is the inner products. Note that every entry in the Hessenberg matrix H_k is the result of an inner product, and these are the only inner products in the algorithm. MPMGMRES therefore needs $(2k-1)\frac{t^2}{2} + \frac{3}{2}t$ inner products at the k th step [4, Table 4.1].

Significantly, in the domain decomposition setting, due to the nature of the standard Additive Schwarz preconditioner, the preconditioning step is *exactly* the same cost when using both selective MPMGMRES and standard preconditioned GMRES. Moreover, since the vectors we obtain by applying the preconditioners are sparse, the cost of the matrix-vector products will also be of the same order as in the standard GMRES algorithm – the only extra expense coming from the overlapping nodes. Indeed, if we use RAS, then the cost of a matrix-vector product would be identical here too.

The extra cost in the MPGMRES approach therefore lies completely with the inner products. The vectors here are, in general, dense, as we lose sparsity of w in the modified Gram-Schmidt step.

3.2 The subspace in complete MPGMRES

Recall that (complete) MPGMRES minimizes over the multi-Krylov subspace

$$\mathcal{K}_{M_1, \dots, M_t}^k(A, \mathbf{r}_0),$$

where

$$\begin{aligned} \mathcal{K}_{M_1, \dots, M_t}^1(A, \mathbf{r}_0) &= \text{span}\{M_1 \mathbf{A} \mathbf{r}_0, \dots, M_t \mathbf{A} \mathbf{r}_0\}, \\ \mathcal{K}_{M_1, \dots, M_t}^2(A, \mathbf{r}_0) &= \text{span}\{M_1 \mathbf{A} \mathbf{r}_0, \dots, M_t \mathbf{A} \mathbf{r}_0, \\ &\quad M_1 \mathbf{A} M_1 \mathbf{r}_0, \dots, M_1 \mathbf{A} M_t \mathbf{r}_0, \dots \\ &\quad \dots, M_t \mathbf{A} M_1 \mathbf{r}_0, \dots, M_t \mathbf{A} M_t \mathbf{r}_0\}, \end{aligned}$$

etc. Usually the size of this space grows exponentially with each iteration. However, in a domain decomposition context the situation is not quite so dire, as we see below.

First, note that each of the preconditioners are projection matrices, and so

$$\begin{aligned} M_i \mathbf{A} M_i &= R_{i, \delta}^T (R_{i, \delta} \mathbf{A} R_{i, \delta}^T)^{-1} R_{i, \delta} \mathbf{A} R_{i, \delta}^T (R_{i, \delta} \mathbf{A} R_{i, \delta}^T)^{-1} R_{i, \delta} \\ &= M_i. \end{aligned}$$

Hence applying M_i to $\mathbf{A} M_i$ does nothing to enrich the space.

Next, note that

$$M_i \mathbf{A} M_j = R_{i, \delta}^T (R_{i, \delta} \mathbf{A} R_{i, \delta}^T)^{-1} R_{i, \delta} \mathbf{A} R_{j, \delta}^T (R_{j, \delta} \mathbf{A} R_{j, \delta}^T)^{-1} R_{j, \delta}. \quad (3)$$

In the middle of this expression is the cross-term $R_{i, \delta} \mathbf{A} R_{j, \delta}^T$. Now note that $R_{i, \delta} \mathbf{A} R_{j, \delta}^T = 0 \iff I_i \cap I_j = \emptyset$. Provided the overlap δ isn't big enough to touch two subdomains, this implies that only the contributions from sub-domains that touch each other add anything to the multi-Krylov subspace. This is the number of edges + corners in 2D (a maximum of 8 for a tensor product-based grid), and these plus the number of faces in 3D (a max of 26 for a tensor product-based grid).

Altogether, this means that

$$\dim(\mathcal{K}_{M_1, \dots, M_t}^k(A, \mathbf{r}_0)) = (kc + 1)t,$$

where c is a constant independent of k, t . Therefore, even in the complete MPGMRES case, we only have *linear* growth in the search space.

4 Numerical Experiments

If we split the domain into a small number of subdomains, i.e., we have a high proportion of subdomains lying on an edge, then there may not be much difference between the spaces minimized over by the selective algorithm and the complete algorithm.

For example, consider the special case where we split the domain Ω into two subdomains, Ω_1 and Ω_2 such that $\Omega_1 \cup \Omega_2 = \Omega$. Then it can be shown [4, Section 5.2.1] that, provided the subdomain solves are exact, the space over which we minimize in both selective and complete MPMRES are identical.

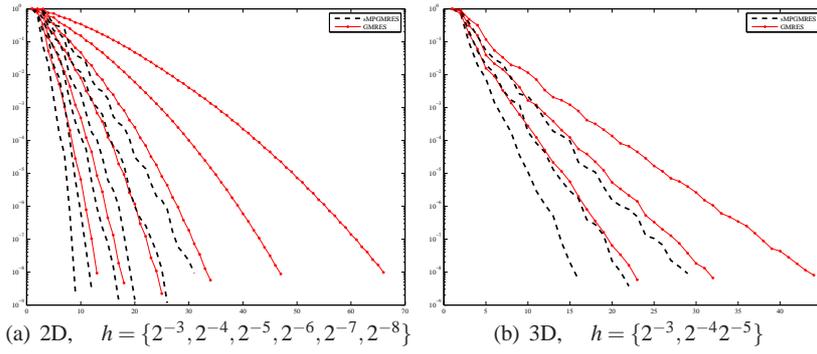


Fig. 1 Convergence curves for two subdomains in 2D and 3D.

For large numbers of subdomains, the work involved in the inner products and vector updates becomes significant, even though the work in actually applying the preconditioners is essentially the same as for the usual AS and RAS methods.

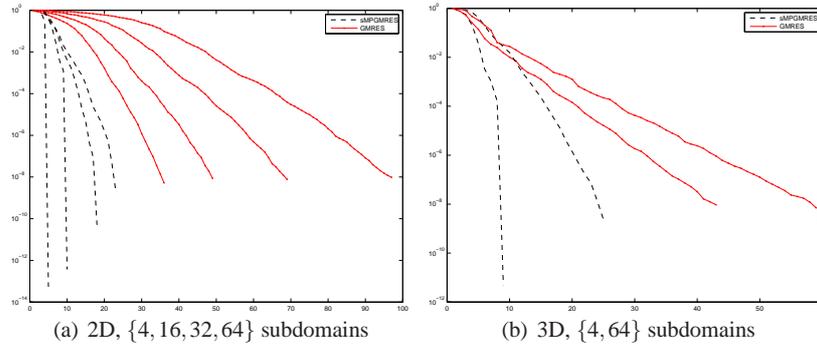


Fig. 2 Convergence curves for multiple subdomains in 2D and 3D.

Although the iteration counts are impressive for a large number of subdomains (with, e.g., 94 iterations for GMRES with an additive Schwarz preconditioner being reduced to 18 iterations with selective MPMGRES), the timings in this case are not yet competitive. This is due to the fact that we are using a proof-of-concept MATLAB code. We are confident, however, that a state-of-the-art implementation will yield great computational savings manifested in a reduced computational running time, due to the highly parallelizable nature of the proposed method. Recall that the only extra work between the methods is in calculating the inner products and the subsequent vector update in the Gram-Schmidt process.

5 Conclusions

We have presented an algorithm that applies Additive Schwarz with Variable Weights. The approach is incorporated as a set of multiple preconditioners into MPMGRES. Domain decomposition has a few unique features that make our approach particularly attractive. First, the preconditioning step entails the same cost when using both selective MPMGRES and standard preconditioned GMRES, and the cost of the matrix-vector products is also of the same order as in the standard GMRES algorithm. Secondly, because there is a very low degree of overlap between nodes in the different subdomains, the growth in the search space for complete MPMGRES is only linear, i.e., very modest. This is in contrast to the situation for other applications, where the search space for complete MPMGRES grows exponentially and one has to settle for a selective algorithm. Domain decomposition is thus unique in that it allows for maintaining a richer subspace where the residual is to be minimized. For this reason, we believe that the approach presented in this paper is promising.

Future work will focus on a parallel implementation with a significant number of subdomains, which are themselves large in size. In such a setting the cost of the inner products and vector-updates would take a small proportion of the total time and this will allow us to further assess and test the method in a large-scale setting.

Acknowledgments. The work of the first author was supported in part by the Natural Sciences and Engineering Research Council (NSERC), and that of the third author in part by the U.S. National Science Foundation under grant DMS-1115520.

References

1. Ayuso de Dios, B., Baker, A.T., Vassilevsky, P.S.: A combined preconditioning strategy for nonsymmetric systems (2012). Arxiv:1208.4544v1
2. Cai, X.C., Sarkis, M.: A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing* **21**, 792 (1999)
3. Frommer, A., Schwandt, H., Szyld, D.B.: Asynchronous weighted additive Schwarz methods. *Electronic Transactions on Numerical Analysis* **5**, 48–61 (1997)

4. Greif, C., Rees, T., Szyld, D.B.: MPGMRES: a generalized minimum residual method with multiple preconditioners. Tech. Rep. 11-12-23, Department of Mathematics, Temple University (2011). Revised September 2012. Also available as Technical Report TR-2011-12, Department of Computer Science, University of British Columbia
5. Saad, Y.: Iterative methods for sparse linear systems. SIAM, Philadelphia (2003)
6. Saad, Y., Schultz, M.H.: GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* **7**, 856–869 (1986)
7. Simoncini, V., Szyld, D.B.: Recent computational developments in Krylov subspace methods for linear systems. *Numerical Linear Algebra with Applications* **14**, 1–59 (2007)
8. Toselli, A., Widlund, O.B.: Domain Decomposition Methods - Algorithms and Theory, *Springer Series in Computational Mathematics*, vol. 34. Springer, Berlin and Heidelberg (2005)