Algebraic Analysis of Schwarz Methods for Singular Systems

Ivo Marek and Daniel B. Szyld

Czech Institute of Technology, School of Civil Engineering

Temple University, Department of Mathematics
(`http://www.math.temple.edu/~szyld/`)

Summary. During the last few years, an algebraic formulation of Schwarz methods was developed. In this paper this algebraic formulation is used to prove new convergence results for multiplicative Schwarz methods when applied to consistent singular systems of linear equations. Coarse grid corrections are also studied. In particular, these results are applied to the numerical solutions of Markov chains.

1 Introduction

We consider the solution of consistent large sparse linear singular systems of the form

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

We study its solution by means of Schwarz methods. Specifically, we analyze the case where the coefficient matrix $A = I - B$, where I is the identity matrix and B is a nonnegative (column) stochastic matrix, i.e., $B^T e = e$, where $e = (1, 1, \dots, 1)^T \in \mathbb{R}^n$. Thus A is a singular M-matrix; see section 2 for definitions. In particular we consider the case of $b = 0$, and thus we look for the nonnegative vector v, normalized so that $v^T e = 1$, satisfying $Av = 0$, i.e., such that $Bv = v$. This is the stationary probability distribution of the Markov chain represented by B.

In our analysis we use the algebraic formulation of Schwarz methods developed in Benzi et al. [2001], Frommer and Szyld [1999], and applied, e.g., in Frommer and Szyld [2001], Nabben [2003], Nabben and Szyld [2003].

There is no separate treatment in the literature of Schwarz methods for singular systems in the p.d.e. context. Nevertheless the implementations derived mostly for the non-singular case can be shown to work in the singular case as well, especially when the null space is known. This is the case, for example, when Neumann boundary conditions are present. The convergence theory developed, e.g., in Dryja et al. [1994], Dryja and Widlund [1994], can be applied to these cases with little or no changes.

We believe that this is the first time that singular systems are analyzed using an algebraic approach to Schwarz methods (with overlap), and that Markov chains problems are studied in this context. One of our goals is to present Schwarz iterations as one more possible tool for the numerical solutions of Markov chains. In fact, multiplicative Schwarz iterations reduce to the block Gauss-Seidel method when the overlap is removed. Having the overlap has proved crucial for the fast convergence of these methods in the nonsingular case; see, e.g., Smith et al. [1996], Dryja and Widlund [1994]. In the singular context, having larger overlap may decrease the convergence rate of the iteration. Comparison theorems may be used to prove such decrease in convergence rate; see Marek and Szyld [2000], Marek and Szyld [2002].

We discuss here one approach, namely that multiplicative Schwarz iterations applied directly to the $n \times n$ system (1) converge. Other approaches are discussed in Marek and Szyld [2004], where "coarse-grid" corrections are also considered.

2 Definitions and auxiliary results

In this section we present some notation, definitions, and preliminaries. Concepts on nonnegative matrices not explicitly defined here can be found in the book by Berman and Plemmons [1979].

An $n \times n$ matrix $C = (c_{jk})$ with $c_{jk} \in \mathbb{R}$, is called nonnegative if $c_{jk} \geq 0,\ j,k = 1,\ldots,n$; this is denoted $C \geq O$. When $c_{jk} > 0,\ j,k = 1,\ldots,n$, we say that the matrix is positive and denote it by $C > O$. The same notation is used for nonnegative and positive vectors. By $\sigma(C)$ we denote the spectrum of C and by $\rho(C)$ its spectral radius. By $\mathcal{R}(C)$ and $\mathcal{N}(C)$ we denote the range and null space of C, respectively.

Let $\lambda \in \sigma(C)$ be a pole of the resolvent operator $R(\mu, C) = (\mu I - C)^{-1}$. The multiplicity of λ as a pole of $R(\mu, C)$ is called the index of C with respect to λ and denoted $ind_\lambda C$. Equivalently, $k = ind_\lambda C$ if it is the smallest integer for which $\mathcal{R}((\lambda I - C)^{k+1}) = \mathcal{R}((\lambda I - C)^k)$. This happens if and only if $\mathcal{R}((\lambda I - C)^k) \oplus \mathcal{N}((\lambda I - C)^k) = \mathbb{R}^n$.

Let A be an $n \times n$ matrix. A is an M-matrix if $A = \beta I - B$, B nonnegative and $\rho(B) \leq \beta$. A pair of matrices (M, N) is called a splitting of A if $A = M - N$ and M^{-1} exists. A splitting of a matrix A is called *of nonnegative type* if the matrix $T = M^{-1}N$ is nonnegative (Marek [1970]). If the matrices M^{-1} and N are nonnegative, the splitting is called *regular* (Varga [1962]).

Let T be a square matrix. T is called *convergent* if $\lim_{k\to\infty} T^k$ exists, and *zero-convergent* if $\lim_{k\to\infty} T^k = O$. Standard stationary iterations of the form

$$x^{k+1} = Tx^k + c, \qquad k = 0, 1, \ldots, \tag{2}$$

converge if and only if either $\lim_{k\to\infty} T^k = O$. or, if $\rho(T) = 1$, T is convergent. A square matrix T with unit spectral radius is convergent if the following two

conditions hold:
(i) if $\lambda \in \sigma(T)$ and $\lambda \neq 1$, then $|\lambda| < 1$.
(ii) $ind_1 T = 1$.
When $T \geq O$, (i) can be replaced with T having positive diagonal entries (Alefeld and Schneider [1982]); see Szyld [1994] for equivalent conditions for (ii).

We state a very useful Lemma; see e.g., Bohl and Marek [1995] for a proof. We note that when $\rho(T) = 1$, this lemma can be used to show condition (ii) above. To prove convergence one needs to show in addition that condition (i) also holds, or equivalently, that the diagonal entries are all positive.

Lemma 1. *Let T be a nonnegative square matrix such that $Tv \leq \alpha v$ with $v > 0$. Then $\rho(T) \leq \alpha$. If furthermore $\rho(T) = \alpha$, then $ind_\alpha T = 1$.*

A square nonnegative matrix B is irreducible if for every pair of indices i, j there is a power $k = k(i, j)$ such that the ij entry of B^k is nonzero.

3 Algebraic formulation of Schwarz methods

Given an initial approximation x^0 to the solution of (1), the (one-level) multiplicative Schwarz method can be written as the stationary iteration (2), where

$$T = (I - P_p)(I - P_{p-1}) \cdots (I - P_1) = \prod_{i=p}^{1} (I - P_i) \tag{3}$$

and c is a certain vector. Here

$$P_i = R_i^T A_i^{-1} R_i A, \tag{4}$$

where $A_i = R_i A R_i^T$, R_i is a matrix of dimension $n_i \times n$ with full row rank, $1 \leq i \leq p$; see, e.g., Smith et al. [1996]. In the case of overlap we have $\sum_{i=1}^{p} n_i > n$. Note that each P_i, and hence each $I - P_i$, is a projection operator; i.e., $(I - P_i)^2 = I - P_i$. Each $I - P_i$ is singular and $\rho(I - P_i) = 1$.

We refer the reader to the papers Benzi et al. [2001], Frommer and Szyld [1999] for details of the algebraic formulation of Schwarz methods. What we will say here is that nonsingular matrices M_i, $i = 1, \ldots, p$, are defined so that $A = M_i - N_i$ are regular splittings (and thus of nonnegative type), and furthermore the following equality holds:

$$E_i M_i^{-1} = R_i^T A_i^{-1} R_i, \; i = 1, \ldots, p, \tag{5}$$

where $E_i = R_i^T R_i$. These diagonal matrices E_i have ones on the diagonal in every row where R_i^T has nonzeros. We can thus rewrite (3) as

$$T = (I - E_p M_p^{-1} A)(I - E_{p-1} M_{p-1}^{-1} A) \cdots (I - E_1 M_1^{-1} A). \tag{6}$$

In the context of discretizations of p.d.e.s, the use of Schwarz methods greatly benefit from the use of coarse grid corrections, and they are needed to guarantee a convergence rate independent of the mesh size; see, e.g., Dryja et al. [1994], Dryja and Widlund [1994], Quarteroni and Valli [1999], Smith et al. [1996]. Coarse grid corrections can be additive or multiplicative. Here we restrict our comments to the multiplicative corrections. To that end consider a new projection P_0 of the form (4) onto the "coarse space", i.e., onto a particular subset of states, usually taken in the overlap between the other set of states. Corresponding to these "coarse" states, there correspond a natural matrix R_0 and $A_0 = R_0 A R_0^T$, so that $E_0 = R_0^T R_0$ and M_0 is similarly defined; see Benzi et al. [2001]. The multiplicative corrected multiplicative Schwarz iteration operator is then

$$T_{\mu c} = (I - P_0)T_\mu = (I - E_0 M_0^{-1} A)T. \tag{7}$$

In Benzi et al. [2001] it was shown that when A is nonsingular, $\rho(T) < 1$, and thus, the method (2) is convergent. The same results hold for $T_{\mu c}$, i.e., with a "coarse grid" correction. In this paper we explore the convergence of (2), using the iterations defined by (6) and (7), when A is singular. Other Schwarz methods for the singular case are studied in Marek and Szyld [2004].

4 Convergence of multiplicative Schwarz

We prove here our main result, namely that when A is irreducible, the multiplicative Schwarz iterations are convergent. As is well known, when $B \geq O$ is irreducible, its Perron eigenvector is strictly positive $v > 0$. If in addition we require that the diagonals of the iteration matrices are positive, we show in the next theorem that the matrix (6) is convergent.

Theorem 1. *Let $A = I - B$, where B is an $n \times n$ column stochastic matrix such that $Bv = v$ with $v > 0$. Let $p \geq 1$ be a positive integer and $A = M_i - N_i$ be splittings of nonnegative type such that the diagonals of $T_i = M_i^{-1} N_i$, $i = 1, \ldots, p$, are positive. Then (6) is a convergent matrix.*

Proof. Let $v > 0$ be such that $Bv = v$, i.e., $Av = 0$. For each splitting $A = M_i - N_i$, we then have th at $M_i v = N_i v$. This implies that $Tv = v$, and by Lemma 1 we have that $\rho(T) = 1$ and that the index is 1. To show that T is convergent, we show that its diagonal is positive. Each factor in (6) is then

$$I - E_i + E_i(I - M_i^{-1} A) = I - E_i + E_i M_i^{-1} N_i,$$

and since $O \leq E_i \leq I$ and $M_i^{-1} N_i \geq O$, each factor is nonnegative. For a row in which E_i is zero, the diagonal entry in this factor has value one. For a row in which E_i has value one, the diagonal entry in this factor is the positive diagonal entry of $M_i^{-1} N_i$. Thus, we have a product of nonnegative matrices, each having positive diagonals, implying that the product T has positive diagonal entries, and therefore it is convergent. □

Corollary 1. *Theorem 1 applies verbatim to the case of "coarse grid" correction, by considering the additional splitting* $A = M_0 - N_0$*, with* $T_0 = M_0^{-1} N_0$ *having positive diagonals, so that* $T_{\mu c}$ *of (7) is convergent.*

Let $\gamma = \max\{|\lambda|, \lambda \in \sigma(T), \lambda \neq 1\}$. The fact that T is convergent implies that $\gamma < 1$; see, e.g., Berman and Plemmons [1979]. Therefore Theorem 1 and Corollary 1 indicate that for multiplicative Schwarz, $\sigma(M^{-1}A) = \sigma(I - T)$ has zero as an isolated eigenvalue with index 1, and the rest of the spectrum is contained in a ball with center 1 and radius γ. Furthermore, the smaller γ is, the smaller this ball around 1 is. This configuration of the spectrum often gives good convergence properties to Krylov subspace methods preconditioned with multiplicative Schwarz.

5 The reducible case

We consider here the general case, where B might not be irreducible. There is a permutation matrix H such that the symmetric permutation of B is lower block-triangular [Gantmacher, 1959, p.341], and in fact it has the form

$$HBH^T = \begin{bmatrix} G_0 & O & \cdots & O \\ G_1 & C_1 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ G_p & O & \cdots & C_p \end{bmatrix}, \tag{8}$$

where $\lim_{k \to \infty} G_0^k = O$ and C_i is an irreducible and stochastic matrix, $i = 1, \ldots, p$. There are efficient algorithms to compute the permutation matrix H, and thus, the form (8). For example, Tarjan's algorithm has almost linear complexity and good software is available for it; see Duff and Reid [1978].

Solving linear systems with the matrix B reduces then to solving systems with each of the diagonal blocks of (8). This can be accomplished using multiplicative Schwarz iterations, which were shown to converge for irreducible stochastic matrices in section 4.

References

G. Alefeld and N. Schneider. On square roots of M-matrices. *Linear Algebra and its Applications*, 42:119–132, 1982.

M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89:605–639, 2001.

A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, third edition, 1979. Reprinted by SIAM, Philadelphia, 1994.

E. Bohl and I. Marek. A model of amplification. *Journal of Computational and Applied Mathematics*, 63:27–47, 1995.

M. Dryja, B. F. Smith, and O. B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31(6):1662–1694, December 1994.

M. Dryja and O. B. Widlund. Domain decomposition algorithms with small overlap. *SIAM J. Sci.Comput.*, 15(3):604–620, May 1994.

I. S. Duff and J. K. Reid. An implementation of tarjan's algorithm for the block triangularization of a matrix. *ACM Transactions on Mathematical Software*, 4:337–147, 1978.

A. Frommer and D. B. Szyld. Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.

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

F. R. Gantmacher. *Application of the Theory of Matrices*. Interscience, New York, 1959.

I. Marek. Frobenius theory of positive operators. comparison theorems and applications. *SIAM Journal on Applied Mathematics*, 19:608–628, 1970.

I. Marek and D. B. Szyld. Comparison of convergence of general stationary iterative methods for singular matrices. *Linear Algebra and its Applications*, 316:67–87, 2000.

I. Marek and D. B. Szyld. Comparison of convergence of general stationary iterative methods for singular matrices. *SIAM Journal on Matrix Analysis and Applications*, 24:68–77, 2002.

I. Marek and D. B. Szyld. Algebraic Schwarz methods for the numerical solution of Markov chains. *Linear Algebra and its Applications*, 2004. To appear.

R. Nabben. Comparisons between additive and multiplicative Schwarz iterations in domain decomposition methods. *Numerische Mathematik*, 95: 145–162, 2003.

R. Nabben and D. B. Szyld. Convergence theory of restricted multiplicative Schwarz methods. *SIAM Journal on Numerical Analysis*, 40:2318–2336, 2003.

A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, 1999.

B. F. Smith, P. E. Bjørstad, and W. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, 1996.

D. B. Szyld. Equivalence of convergence conditions for iterative methods for singular equations. *Numerical Linear Algebra with Applications*, 1:151–154, 1994.

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