Overlapping Schwarz Methods

Olof Widlund
Courant Institute of Mathematical Sciences
New York University
http://www.cs.nyu.edu/cs/faculty/widlund/
Based on first three chapters of a new book coauthored with Andrea Toselli.
Introduction

The earliest domain decomposition method, known to us, is the alternating method of Hermann Amandus Schwarz discovered in 1869. Schwarz used this algorithm to establish the existence of harmonic functions with prescribed boundary values on regions with non-smooth boundaries. Regions constructed recursively by forming unions of pairs of regions starting with regions for which existence could be established by elementary means. At the core of this work is a proof that the method converges at a geometric rate in the maximum norm.

For more than two subregions, we can define a step of the algorithm by recursion: i) Solve on the first subregion; ii) Solve on the union of all other subregions, approximately, by recursively invoking the same algorithm.
As pointed out by Pierre-Louis Lions, at DD1, the convergence of this algorithm can be established by two different methods: by a maximum principle and by using Hilbert spaces.

The Hilbert space method more popular; mixes well with classical calculus of variation and finite elements.

The classical Schwarz’s method: Two fractional steps corresponding to two overlapping subregions, $\Omega_1'$ and $\Omega_2'$, with $\Omega = \Omega_1' \cup \Omega_2'$. 
Given an initial guess $u^0$, which vanishes on $\partial \Omega$, the iterate $u^{n+1}$ is determined from the previous iterate $u^n$ in two sequential steps.

\[
\begin{cases}
-\Delta u^{n+1/2} = f & \text{in } \Omega_1', \\
u^{n+1/2} = u^n & \text{on } \partial \Omega_1', \\
u^{n+1/2} = u^n & \text{in } \Omega_2 = \Omega_2 \setminus \overline{\Omega_1'}, \\
-\Delta u^{n+1} = f & \text{in } \Omega_2', \\
u^{n+1} = u^{n+1/2} & \text{on } \partial \Omega_2', \\
u^{n+1} = u^{n+1/2} & \text{in } \Omega_1 = \Omega_1 \setminus \overline{\Omega_2'}.
\end{cases}
\]

We can also write it in terms of projections onto subspaces:

\[
u^{n+1} - u = (I - P_2)(u^{n+1/2} - u) = (I - P_2)(I - P_1)(u^n - u),
\]

where $P_i := R_i^T A_i^{-1} R_i A$. This is the basic multiplicative Schwarz method.
Extends immediately to many subdomains by recursion. We are solving

\[ P_{mu}u := (P_1 + P_2 - P_2 P_1)u = g. \]

We could simplify and just use two terms. We get the basic additive Schwarz method:

\[ P_{add}u := (P_1 + P_2)u = g_{ad}. \]

This is a symmetric operator even for more than two subdomains.

There are a number of other useful Schwarz methods such as

\[ (I - P_1)(I - P_2)(I - P_3)(I - P_2)(I - P_1), \]

for three subdomains.
There are, in fact, at least three ways of analyzing the Schwarz methods. Schwarz used a maximum principle in 1870. We can use an abstract Schwarz theory soon to be briefly discussed.

For two subdomains one can also argue about Schur complements and show,

\[ e_{\Gamma_1}^{n+1} = \left( I - (S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1}(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}) \right) e_{\Gamma_1}^n. \]

We view the iteration in terms of an update of the values on \( \Gamma_1 \). The Schur complements correspond to \( \Omega'_1, \Omega_2, \) and \( \Omega_3 \).
Schwarz Methods

$V$ a finite dimensional space with bilinear form $a(u, v)$ inherited from an elliptic problem. Consider:
Find $u \in V$ such that

$$a(u, v) = f(v) \quad \forall v \in V,$$

and let

$$V = V_0 + V_1 + \cdots + V_N.$$ 

Not necessarily a direct sum of spaces. In many cases, the representation of an element of $V$ in terms of components of the $V_i$ is not unique. The first space $V_0$ often a special coarse subspace; sometimes it is left out.

Let $\tilde{a}_i(u, v), \ i = 0, \cdots, N,$ be symmetric, positive definite bilinear forms
on $V_i \times V_i$. Introduce operators $T_i : V \rightarrow V_i$, by

\[ \tilde{a}_i(T_i u, v) = a(u, v) \quad \forall v \in V_i. \] (2)

One choice is $\tilde{a}_i(u, v) = a(u, v)$. Then, $T_i = P_i$, a projection orthogonal with respect to the inner product $a(\cdot, \cdot)$.

Using these building blocks, we can, e.g. introduce the additive Schwarz operator:

\[ T_{as} = T_0 + T_1 + \cdots + T_N. \]

Replace elliptic finite element problem by transformed, preconditioned problem with the same solution:

\[ T_{as}u = g, \quad g = \sum_{i=0}^{N} g_i, \quad g_i = T_i u. \] (3)
The right hand side $g$ obtained from

$$\tilde{a}_i(g_i, v_i) = a(u, v_i) = f(v_i) \quad \forall v_i \in V_i.$$ 

Use a conjugate gradient method, without further preconditioning, with $a(\cdot, \cdot)$ as the inner product. The extreme eigenvalues of $T_{as}$ used in the standard estimate of the rate of convergence of the conjugate gradient method.

Well known bound for the conjugate gradient error after $k$ steps:

$$2(\sqrt{\kappa} - 1)^k \quad \text{where} \quad \kappa = \frac{\lambda_{\text{max}}(T_{as})}{\lambda_{\text{min}}(T_{as})}.$$ 

This class of algorithms can be parallelized by solving the subproblems simultaneously; this is in contrast to the classical, sequential multiplicative version.
Block Jacobi and matrix form of operators

Turn the finite element variational problem into a linear system of algebraic equations, \( Kx = b \). \( K \) = stiffness matrix, \( b \) = load vector. \( K^T = K > 0 \). A property inherited from the bilinear form if finite elements are conforming.

Consider the block-Jacobi/conjugate gradient method. The stiffness matrix \( K \) preconditioned by a matrix \( K_J \), the direct sum of diagonal blocks of \( K \). Each block corresponds to a set of degrees of freedom, which span a subspace \( V_i \). The space \( V \) is a direct sum of subspaces \( V_i, i = 1, \ldots, N \). The choice of the subspaces is a key issue and so is the choice of basis of \( V \) in, e.g. the spectral element case.

For each subspace \( V_i \), an orthogonal projection \( P_i \):

\[
a(P_iu, v) = a(u, v), \quad \forall v \in V_i, \quad u \in V,
\]
or an approximate projection defined by

\[ \tilde{a}_i(T_iu, v) = a(u, v), \quad \forall v \in V_i, \quad u \in V. \]

\( P_i \) corresponds to the inverse of a block of \( K \), padded by zero blocks, times \( K \). To obtain \( T_i \) replace the relevant block of \( K \) by a local preconditioner.

A matrix form of the operators \( P_i \) and \( T_i \):

Let \( V_i \) correspond to a set of adjacent variables associated with a subregion \( \Omega_i' \); all degrees of freedom outside and on the boundary of \( \Omega_i' \) vanish. After a suitable permutation of variables, \( P_i \) corresponds to

\[ y = P^i x = \begin{pmatrix} (K^{(i)})^{-1} & 0 \\ 0 & 0 \end{pmatrix} K x \]

\( K^{(i)} \) stiffness matrix of Dirichlet problem on \( \Omega_i' \).
Approximate projection $T_i$ obtained by replacing $K^{(i)}$ by a suitable preconditioner for the local Dirichlet problem on $\Omega'_i$.

The relevant spectrum is that of

$$T_{as} = \sum T_i.$$ 

The eigenvalues of $K^{-1}K_J$ are the stationary values of the generalized Rayleigh quotient

$$\sum_{i=1}^{N} \frac{\tilde{a}_i(u_i, u_i)}{a(u, u)}, \quad u = \sum_{i=1}^{N} u_i, \quad u_i \in V_i.$$ 

The upper bound often the most challenging. The lower bound involves upper bounds on $a(u_i, u_i)/\tilde{a}_i(u_i, u_i), \forall u_i \in V_i$. Using these bounds, we obtain the required bound on $\kappa(K_J^{-1}K)$. 
These arguments are also valid for any direct sum decomposition of the space; introduce a basis of the space by merging the bases of $V_1, \cdots, V_N$.

This class of preconditioner can be extended to the case when the finite element space is not a direct sum of subspaces corresponding to the blocks of a block Jacobi splitting. We can, e.g. add a global subspace, $V_0$, which can be a finite element space of degree one on a coarse triangulation of $\Omega$, or, more generally, a subspace with one or just a few degrees of freedom for each coarse element. There are also many other possibilities. The basic formula for the general case, which is easy to prove, is

\[
a(T_{as}^{-1}u, u) = \inf_{u = \sum u_i} \sum \tilde{a}_i(u_i, u_i).
\]

Again, $T_{as} = \sum T_i$ is the operator relevant for the iterative method being considered. There is freedom of choice in the representation of $u$ if the subspaces $V_i$ no longer form a direct sum. Entire spectrum of $T_{as}$ moves
up when subspaces are enlarged. This is often quite advantageous.

We can, e.g., start with a direct sum decomposition of the finite element space and then enrich some or all of the subspaces, increasing their dimensions. This can be related to an increase of the overlap of a decomposition of the given region $\Omega$ into subdomains. We can also strengthen preconditioners by adding additional subspaces, e.g., a coarse global space.

Note that the general theory shows that the upper bound on the Rayleigh quotient decreases if two subspaces are merged or if a subspace is increased, e.g. by increasing the overlap. This often leads to a considerably much faster algorithm.

Introductions to this material can be found in books by Smith, Bjørstad, and Gropp, Cambridge University Press, 1996 and by Toselli and Widlund, Springer 2004.
Abstract Schwarz theory

We assume that $V$ admits the following decomposition

$$V = R_0^T V_0 + \sum_{i=1}^{N} R_i^T V_i. \tag{4}$$

Note that the $V_i$ do not need to be subspaces of $V$, but that, as is customary, we refer to them as ‘subspaces’ or ‘local spaces’.

Bilinear forms on the subspaces,

$$\tilde{a}_i(\cdot, \cdot) : V_i \times V_i \longrightarrow \mathbb{R}, \quad i = 0, \ldots, N,$$

and the local stiffness matrices associated with them,

$$\tilde{A}_i : V_i \longrightarrow V_i.$$
\[ P_i = R^T_i \tilde{A}^{-1}_i R_i A, \quad 0 \leq i \leq N. \]  \hspace{1cm} (5)

In order to prove bounds, it is enough to make three assumptions.

**Assumption 1. [Stable Decomposition]** There exists a constant \( C_0 \), such that every \( u \in V \) admits a decomposition

\[ u = \sum_{i=0}^{N} R^T_i u_i, \quad \{u_i \in V_i, \ 0 \leq i \leq N\} \]

that satisfies

\[ \sum_{i=0}^{N} \tilde{a}_i(u_i, u_i) \leq C_0^2 a(u, u). \]
Assumption 2. [Strengthened Cauchy-Schwarz Inequalities] There exist constants $0 \leq \epsilon_{ij} \leq 1$, $1 \leq i, j \leq N$, such that

$$|a(R_i^T u_i, R_j^T u_j)| \leq \epsilon_{ij} a(R_i^T u_i, R_i^T u_i)^{1/2} a(R_j^T u_j, R_j^T u_j)^{1/2},$$

for $u_i \in V_i$ and $u_j \in V_j$. We will denote the spectral radius of $E = \{\epsilon_{ij}\}$ by $\rho(E)$.

We note that Assumption 2 does not involve the space $V_0$. 
Assumption 3. [Local Stability] There exists $\omega > 0$, such that

$$a(R_i^T u_i, R_i^T u_i) \leq \omega \tilde{a}_i(u_i, u_i), \quad u_i \in \text{range}(\tilde{P}_i) \subset V_i, \quad 0 \leq i \leq N.$$ 

Assumption 3 ensures that the local bilinear forms are coercive and gives a one-sided measure of their approximation properties.

**Theorem 1.** Let Assumptions 1, 2, and 3 be satisfied. Then the condition number of the additive Schwarz operator satisfies

$$\kappa(P_{ad}) \leq C_0^2 \omega (\rho(\mathcal{E}) + 1).$$
A two-level additive Schwarz method with small (or generous) overlap.

Consider Poisson’s equation on a bounded Lipschitz region $\Omega$ in two or three dimensions.

Two triangulations, a coarse, and a fine, (which might be a refinement of the coarse.) The overall space is $V^h$, the space of continuous, piecewise linear finite element functions on the fine triangulation. There is also a covering of the region by overlapping subregions $\Omega'_i$. Let $\delta/H$ measure the relative overlap between adjacent subregions, each of which is a union of elements. Assume shape regularity, but not necessarily quasi-uniformity of the triangles.

Spaces chosen for the Schwarz methods:

$$V_0 = V^H \quad \text{based on coarse triangulation},$$
\[ V_i = V^h \cap H^1_0(\Omega'_i), \quad i > 0. \]

**Theorem 1.** The condition number of the additive Schwarz method satisfies

\[ \kappa(T_{as}) \leq C(1 + H/\delta). \]

*The constant $C$ is independent of the parameters $H$, $h$ and $\delta$.*

Result cannot be improved: Sue Brenner.

There are quite similar results for multiplicative algorithms as well.

Without a coarse space component, $H/\delta$ must be replaced by $1/(H\delta)$.

There are also many alternative coarse spaces; cf. Toselli and W., Chapter 3.
To prove this bound, we must come up with a recipe of how to decompose any function in $V$.

We choose

$$u_0 = \tilde{I}^H u \in V_0,$$  \hspace{1cm} (6)

where we use averages over neighborhoods of nodes of coarse triangles and interpolation into the coarse space. Let

$$w = u - R^T_0 u_0 = u - I^h u_0.$$  \hspace{1cm} (7)

The local components are defined by

$$u_i = R_i(I^h(\theta_i w)) \in V_i, \quad 1 \leq i \leq N.$$  \hspace{1cm} (8)

Here $\{\theta_i\}$ is a piecewise linear partition of unity associated with the overlapping partition.