NUMERICAL LINEAR ALGEBRA

Prof. Paola Antonietti

MOX - Dipartimento di Matematica

Politecnico di Milano

https://antonietti.faculty.polimi.it

TA: Dr. Michele Botti

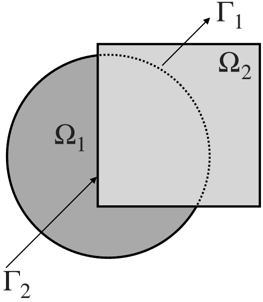
P7: Domani Decomposition Methods

Alternating Schwarz Method

Consider (an elliptic) partial differential equation of the form

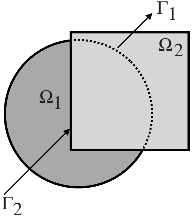
$$Lu = f \qquad \text{in } \Omega = \Omega_1 \cup \Omega_2$$

with boundary condition $u = g$ on $\partial\Omega$



Alternating Schwarz Method

Given $u^{(0)}$



1) On Ω_1 solve

$$\begin{cases} L u_1^{(k+\frac{1}{2})} = f & \text{in } \Omega_1 \\ u_1^{(k+\frac{1}{2})} = g & \text{in } \partial\Omega_1 \setminus \Gamma_1 \\ u_1^{(k+\frac{1}{2})} = u_2^{(k)} & \text{in } \Gamma_1 \end{cases}$$

2) On Ω_2 solve

$$\begin{cases} L u_2^{(k+1)} = f & \text{in } \Omega_2 \\ u_2^{(k+1)} = g & \text{in } \partial\Omega_2 \setminus \Gamma_2 \\ u_2^{(k+1)} = u_1^{(k+\frac{1}{2})} & \text{in } \Gamma_2 \end{cases}$$

3) Define

$$u^{(k+1)} = \begin{cases} u_1^{(k+\frac{1}{2})} & \text{in } \Omega \setminus \Omega_2 \\ u_2^{(k+1)} & \text{in } \Omega_2 \end{cases}$$

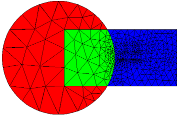
Alternating Schwarz Method

- Alternating iterations continue until convergence to the solution u on the entire domain Ω

- Schwarz proposed this method in 1870 to deal with regions for which analytical solutions are not known

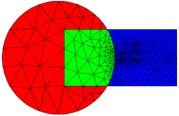
- Today it is of interest, in discretized form, for suggesting one of two major paradigms for solving PDEs numerically by domain decomposition
    - Overlapping subdomains (Schwarz)
    - Non-overlapping subdomains (Schur)

Discretized Schwarz Methods



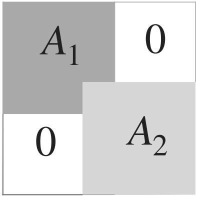
- Discretization yields a $n \times n$ symmetric positive definite linear algebraic system of the form $A\mathbf{x} = \mathbf{b}$
- For $i = 1,2$, let S_i be set of n_i indices of grid points in the interior of Ω_i, where $n_i = |S_i|$.
- Because subdomains overlap, $S_1 \cap S_2 \neq \emptyset$ and $n_1 + n_2 > n$
- For $i = 1,2$, let R_i be $n_i \times n$ the Boolean restriction matrix such that for any vector $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}_i = R_i\mathbf{v} \in \mathbb{R}^{n_i}$ contains precisely those components of $\mathbf{v}$ corresponding to indices in S_i (i.e., those components associated to nodes in Ω_i)

Discretized Schwarz Methods

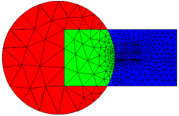


- Conversely, let $R^T \in \mathbb{R}^{n \times n_i}$ be the extension matrix that expands the vector $\mathbf{v}_i \in \mathbb{R}^{n_i}$ into a vector $\mathbf{v} = R_i^T \mathbf{v}_i \in \mathbb{R}^n$, whose components correspond to indices in S_i are same as those of $\mathbf{v}_i$, and whose remaining components are all zero.
- The principal submatrices $A_i \in \mathbb{R}^{n_i \times n_i}, i = 1{,}2,$ of Acorresponding to two subdomains are given by

$$A_1 = R_1 A R_1^T$$
$$A_2 = R_2 A R_2^T$$



Discretized Schwarz Methods



- For discretized problem, the alternating Schwarz iteration takes the following form

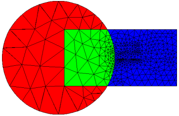
$$\mathbf{x}^{(k+\frac{1}{2})} = \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)})$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k+\frac{1}{2})} + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k+\frac{1}{2})})$$

- This method is analogous to block Gauss-Seidel, but with overlapping blocks
- The method is known as multiplicative Schwarz method
- Overall the error $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$ updated as $\mathbf{e}^{(k+1)} = B_{MS}\mathbf{e}^{(k)}$ where

$$B_{MS} = (I - R_2^T A_2^{-1} R_2 A)(I - R_1^T A_1^{-1} R_1 A)$$

Discretized Schwarz Methods



- We have as yet achieved no parallelism, since two subproblems must be solved sequentially for each iteration, but instead of Gauss-Seidel, we can use the block Jacobi approach

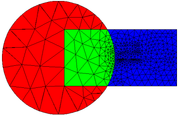
$$\mathbf{x}^{(k+\frac{1}{2})} = \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)})$$
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k+\frac{1}{2})} + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k)})$$

whose subproblems can be solved simultaneously

- The method is known as additive Schwarz method
- Overall the error $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$ updated as $\mathbf{e}^{(k+1)} = B_{AS}\mathbf{e}^{(k)}$ where

$$B_{AS} = (R_2^T A_2^{-1} R_2 + R_1^T A_1^{-1} R_1) A$$

Discretized Schwarz Methods



- With either Gauss-Seidel or Jacobi version, it can be shown that iteration converges at rate independent of mesh size, provided overlap area between subdomains is sufficiently large (and mesh is refined uniformly)

Additive Schwarz Method

$$\mathbf{x}^{(k+\frac{1}{2})} = \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)})$$
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k+\frac{1}{2})} + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k)})$$

Eliminate $\mathbf{x}^{(k+\frac{1}{2})}$ in the Additive Schwarz Methods to obtain

$$\begin{aligned}
\mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)}) + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k)}) \\
&= \mathbf{x}^{(k)} + (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)(\mathbf{b} - A\,\mathbf{x}^{(k)}) \\
&= \mathbf{x}^{(k)} + (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)(\mathbf{r}^{(k)}) \\
&= \mathbf{x}^{(k)} + P_{ad}^{-1}\mathbf{r}^{(k)}
\end{aligned}$$

which is just a Richardson iteration with additive Schwarz preconditioner $P_{ad}^{-1} = (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)$

Additive Schwarz preconditioner

$$\boxed{\begin{aligned}\mathbf{x}^{(k+\frac{1}{2})} &= \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)}) \\ \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k+\frac{1}{2})} + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k)})\end{aligned}}$$

Additive Schwarz Method

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + P_{ad}^{-1}\mathbf{r}^{(k)} \qquad k \geq 0$$

with $P_{ad}^{-1} = (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)$

Remark: Symmetry of preconditioner means that it can be used also in conjunction with PCG, with preconditioner P_{ad} to accelerate convergence

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k P_{ad}^{-1}\mathbf{p}^{(k)} \qquad k \geq 0$$

Symmetrized Multiplicative Schwarz preconditioner

$$\boxed{\begin{aligned}\mathbf{x}^{(k+\frac{1}{2})} &= \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)})\\ \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k+\frac{1}{2})} + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k+\frac{1}{2})})\end{aligned}}$$

The multiplicative Schwarz iteration matrix is not symmetric, but can be made symmetric by additional step with A_1^{-1} each iteration

$$\mathbf{x}^{(k+\frac{1}{3})} = \mathbf{x}^{(k)} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k)})$$
$$\mathbf{x}^{(k+2/3)} = \mathbf{x}^{(k+\frac{1}{3})} + R_2^T A_2^{-1} R_2(\mathbf{b} - A\,\mathbf{x}^{(k+1/3)})$$
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k+\frac{2}{3})} + R_1^T A_1^{-1} R_1(\mathbf{b} - A\,\mathbf{x}^{(k+2/3)})$$

which yields to a symmetric preconditioner that can be used in conjunction with PCG to accelerate convergence

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k P_{mus}^{-1}\mathbf{r}^{(k)} \qquad k \geq 0$$

Many Overlapping Subdomains

- To achieve higher degree of parallelism with Schwarz method, we can apply a two-domain algorithm recursively or use many subdomains
- If there are N overlapping subdomains, then define matrices R_i and A_i as before, $i = 1,...,N$
- The Additive Schwarz preconditioner then takes form

$$P_{ad}^{-1} = \sum_{i=1,\dots,N} R_i^T A_i^{-1} R_i$$

Many Overlapping Subdomains

- Resulting generalisation of block-Jacobi iteration is highly parallel, but not algorithmically scalable because the convergence rate degrades as N grows
- The convergence rate can be restored by using a coarse grid correction to provide global coupling
- The Additive Schwarz preconditioner then takes the form

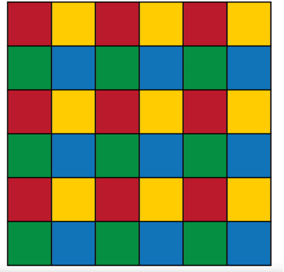
$$P_{ad} = \sum_{i=0,\dots,N} R_i^T A_i^{-1} R_i$$

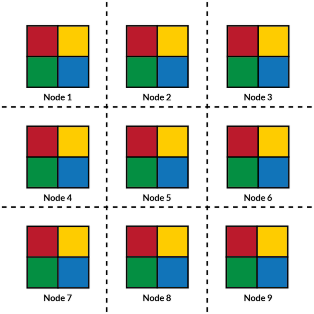
Many Overlapping Subdomains

- Multiplicative Schwarz iteration for p domains is defined analogously
- As with classical Gauss-Seidel vs. Jacobi, multiplicative Schwarz has a faster convergence rate than corresponding additive Schwarz (though it still requires coarse grid correction to remain scalable)
- Unfortunately, multiplicative Schwarz appears to provide no parallelism, as p subproblems per iteration must be solved sequentially
- As with classical Gauss-Seidel, parallelism can be introduced by coloring subdomains to identify independent subproblems that can be solved simultaneously

Colouring techniques

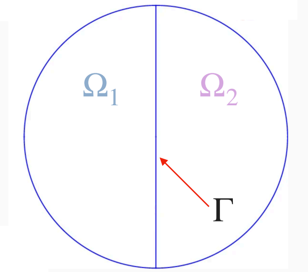
- The multiplicative Schwarz preconditioner is inherently serial.
- We must use a subdomain colouring mechanism in order to identify a set of subdomains that can be processed concurrently.
- This may limit the degree of parallelism if there is a low number of subdomains per color.
- In general, the Multiplicative Schwarz method converges faster than the Additive Schwarz method, while the latter can result in better parallel speedup.

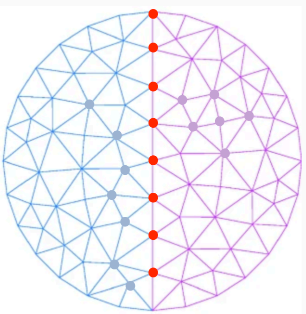




Non Overlapping Subdomains

- We now consider adjacent subdomains whose only points in common are along their mutual boundary Γ
- We partition indices of unknowns in the corresponding discrete linear system into three sets:
  - S_1 corresponding to interior nodes in Ω_1
  - S_2 corresponding to interior nodes in Ω_2
  - S_Γ corresponding to interface nodes in Γ



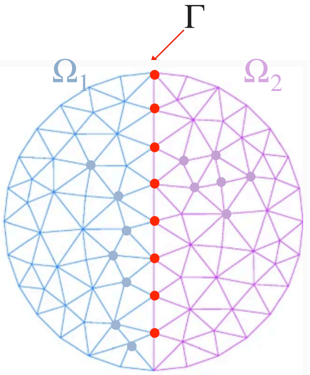


Non Overlapping Subdomains

- Partitioning matrix and right-hand-side vector accordingly, we obtain a symmetric block linear system

$$\begin{pmatrix} A_{11} & 0 & A_{1\Gamma} \\ 0 & A_{22} & A_{2\Gamma} \\ A_{1\Gamma}^T & A_{2\Gamma}^T & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_\Gamma \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_\Gamma \end{pmatrix}$$

- Zero blocks result from assumption that nodes in Ω_1 are not directly connected to nodes in Ω_2, but only through interface nodes in Γ
-



The Shur complement

Block LU factorization of matrix A yields

$$\begin{pmatrix} A_{11} & 0 & A_{1\Gamma} \\ 0 & A_{22} & A_{2\Gamma} \\ A_{1\Gamma}^T & A_{2\Gamma}^T & A_{\Gamma\Gamma} \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ A_{1\Gamma}^T & A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 & A_{1\Gamma} \\ 0 & A_{22} & A_{2\Gamma} \\ 0 & 0 & S \end{pmatrix}$$

where S is the Shur complement

$$S = A_{\Gamma\Gamma} - A_{1\Gamma}^T A_{11}^{-1} A_{1\Gamma} - A_{2\Gamma}^T A_{22}^{-1} A_{2\Gamma}$$

The Shur complement system

We can now determine interface unknowns $\mathbf{u}_\Gamma$ by solving system

$$S\,\mathbf{x}_\Gamma = \widetilde{\mathbf{b}}_\Gamma$$

where

$$\widetilde{\mathbf{b}}_\Gamma = \mathbf{b}_\Gamma - A_{1\Gamma}^T A_{11}^{-1}\mathbf{b}_1 - A_{2\Gamma}^T A_{22}^{-1}\mathbf{b}_2$$

The remaining unknowns (which can be computed simultaneously) are then given by

$$\mathbf{x}_1 = A_{11}^{-1}(\mathbf{b}_1 - A_{1\Gamma}\mathbf{x}_\Gamma)$$

$$\mathbf{x}_2 = A_{22}^{-1}(\mathbf{b}_2 - A_{2\Gamma}\mathbf{x}_\Gamma)$$

The Shur complement system: remarks

- Schur complement matrix S is expensive to compute and is generally dense even if A is sparse.
- If Schur complement system $S\mathbf{x}_\Gamma = \widetilde{\mathbf{b}}_\Gamma$ is solved iteratively, then S need not be formed explicitly
- Matrix-vector multiplication by S requires the solution in each subdomain, implicitly involving A_{11}^{-1} and A_{22}^{-1}, which can be done in parallel
- The condition number of S is generally better than that of A, typically $O(h^{-1})$instead of $O(h^{-2})$for mesh size h
- In practice, suitable interface preconditioners are still needed to accelerate convergence

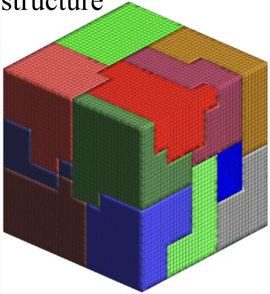
Many Non-Overlapping Subdomains

- To improve parallelism with the Schur method, we can use many subdomains
- If there are N non-overlapping subdomains, let I be set of indices of interior nodes of subdomains and, as before, let Γ be set of indices of interface nodes
- Then the discrete linear system has the following block form

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ A_{I\Gamma}^T & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} \mathbf{x}_I \\ \mathbf{x}_\Gamma \end{pmatrix} = \begin{pmatrix} \mathbf{b}_I \\ \mathbf{b}_\Gamma \end{pmatrix}$$

- The matrix A_{II} is block diagonal and has the following structure

$$A_{II} = \begin{pmatrix} A_{11} & 0 & \dots & 0 \\ 0 & A_{22} & \dots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & \dots & 0 & A_{NN} \end{pmatrix}$$



Many Non-Overlapping Subdomains

- As before, block LU factorization of matrix A yields a system

$$S\,\mathbf{x}_\Gamma = \widetilde{\mathbf{b}}_\Gamma$$

where the Schur complement matrix S is given by

$$S = A_{\Gamma\Gamma} - A_{I\Gamma}^T A_{II}^{-1} A_{I\Gamma}$$

and $\widetilde{\mathbf{b}}_\Gamma = \mathbf{b}_\Gamma - A_{I\Gamma}^T A_{II}^{-1}\mathbf{b}_I$

- This system can be solved again iteratively without forming S explicitly
- Suitable interface preconditioners can be used to accelerate convergence
- Interior unknowns are then given by

$$\mathbf{x}_I = A_{II}^{-1}(\mathbf{b}_I - A_{I\Gamma}\mathbf{x}_\Gamma)$$

- All the occurrences of A_{II}^{-1} can be performed on all subdomains in parallel because A_{II} is block diagonal

How to evaluate the efficiency of a domain decomposition?

- Weak scalability: How the solution time varies with the number of processors for a fixed problem size per processor
- It is not achieved with the one-level method

Number of subdomains	16	32	64
ASM-one level	35	66	128
ASM-two levels	27	28	27

Without the coarse correction, the iteration counts increases linearly with the number of subdomains

Some numerics