Section 06. Subspace Correction Methods

Projections and Subspace Problems

Definition

Let V be a finite-dimensional Hilbert space with inner product $(\cdot,\cdot)$ and $V_j \subset V$ be a subspace. Define

$$\begin{cases} \text{subspace problem} & \mathcal{A}_j : V_j \mapsto V_j, & (\mathcal{A}_j v_j, w_j) = (\mathcal{A} v_j, w_j), & \forall\, v_j, w_j \in V_j; \\ (\cdot,\cdot)\text{-projection} & \mathcal{Q}_j : V \mapsto V_j, & (\mathcal{Q}_j v, w_j) = (v, w_j), & \forall\, w_j \in V_j; \\ (\cdot,\cdot)_{\mathcal{A}}\text{-projection} & \mathit{\Pi}_j : V \mapsto V_j, & (\mathit{\Pi}_j v, w_j)_{\mathcal{A}} = (v, w_j)_{\mathcal{A}}, & \forall\, w_j \in V_j. \end{cases}$$

Lemma (Relation between projections)

The following equalities hold: $\mathcal{I}_j^T = \mathcal{Q}_j, \quad \mathcal{I}_j^* = \mathit{\Pi}_j, \quad \mathcal{Q}_j \mathcal{A} = \mathcal{A}_j \mathit{\Pi}_j.$

From the definition of $\mathcal{A}_j$, we get

$$\mathcal{A}_j = \mathcal{I}_j^T \mathcal{A} \mathcal{I}_j = \mathcal{Q}_j \mathcal{A} \mathcal{I}_j = \mathcal{Q}_j \mathcal{A} \mathcal{Q}_j^T.$$

We can immediately obtain the error equation on each subspace V_j:

$$\mathcal{A} e = r \quad \Longrightarrow \quad \mathcal{Q}_j \mathcal{A} e = \mathcal{Q}_j r \quad \Longrightarrow \quad \mathcal{A}_j \mathit{\Pi}_j e = \mathcal{Q}_j r \quad \Longrightarrow \quad \mathcal{A}_j e_j = r_j,$$

where $r_j = \mathcal{Q}_j r$ and $e_j = \mathit{\Pi}_j e$. We do not know e, but we can find e_j.

Method of Subspace Corrections

- Problem: Find $u \in V$ such that $a[u, v] = f(v), \ \forall v \in V$.
- Space decomposition: $V = \sum_{i=1}^{n} V_i$, divide and conquer
- Subspace corrections: $e_i \approx \mathcal{A}_i^{-1} \mathcal{Q}_i (f - \mathcal{A}u)$

$$u \leftarrow u + \sum_{i=1}^{n} e_i \qquad \text{(Parallel subspace corrections, Jacobi)}$$

$$\text{for } (i = 1 : n) \ u \leftarrow u + e_i \qquad \text{(Successive subspace corrections, GS)}$$

Algorithm (Method of subspace corrections)

$u^{\text{new}} = \text{MSC}(u^{\text{old}})$

1. Form residual: $r = f - \mathcal{A}u^{\text{old}}$
2. Approximate error equation on V_j: $\mathcal{A}_j e_j = r_j$ by $\hat{e}_j = \mathcal{S}_j r_j$
3. Apply subspace corrections: $u^{\text{new}} = u^{\text{old}} + \hat{e}_j$

Q: How to apply these corrections?

Typical Examples of MSC

Algorithm (Successive subspace corrections)

$u^{\text{new}} = SSC(u^{\text{old}})$

1. $v = u^{\text{old}}$
2. for $(j = 1 : J)$ $v = v + \mathcal{S}_j \mathcal{Q}_j (f - \mathcal{A} v)$
3. $u^{\text{new}} = v$

Algorithm (Parallel subspace corrections)

$u^{\text{new}} = PSC(u^{\text{old}})$

1. $r = f - \mathcal{A} u^{\text{old}}$
2. $\hat{e}_j = \mathcal{S}_j \mathcal{Q}_j r, \quad j = 1, \ldots, J$
3. $u^{\text{new}} = u^{\text{old}} + \sum_{j=1}^{J} \hat{e}_j$

- There are two basic approaches to implement the MSC algorithm, namely SSC and PSC
- Hybrid approaches by combining SSC and PSC can also be introduced easily

Some Notations for Convience

- For convenience, we define an operator

$$\mathcal{T}_j = \mathcal{T}_{\mathcal{S}_j} := \mathcal{S}_j \mathcal{Q}_j \mathcal{A} = \mathcal{S}_j \mathcal{A}_j \Pi_j : V \mapsto V_j .$$

- Apparently, if we restrict the domain to V_j, then we have

$$\mathcal{T}_j = \mathcal{T}_{\mathcal{S}_j} = \mathcal{S}_j \mathcal{A}_j : V_j \mapsto V_j .$$

- We now assume all the subspace solvers (smoothers) $\mathcal{S}_j$ are SPD operators. If $\mathcal{S}_j^T = \mathcal{S}_j$, the operator

$$\mathcal{T}_j = \mathcal{S}_j \mathcal{A}_j : V_j \mapsto V_j$$

  is symmetric and positive definite with respect to $(\cdot,\cdot)_{\mathcal{A}}$.

- If $\mathcal{S}_j = \mathcal{A}_j^{-1}$, i.e., the smoother is the exact solver on each subspace, then we have $\mathcal{T}_j = \Pi_j$.

Operator Form of MSC

- The SSC method:

$$u - u^{\text{new}} = (\mathcal{I} - \mathcal{B}_{\text{SSC}}\mathcal{A})(u - u^{\text{old}}) = (\mathcal{I} - \mathcal{T}_J)\cdots(\mathcal{I} - \mathcal{T}_1)(u - u^{\text{old}}). \tag{34}$$

If $J = N$ and each subspace $V_j = \text{span}\{\phi_j\}$ $(j = 1, \ldots, N)$ and $\mathcal{S}_j = \mathcal{A}_j^{-1}$, then the corresponding SSC method (34) is exactly the G-S method.

- The PSC method:

$$\mathcal{B}_{\text{PSC}} = \sum_{j=1}^{J} \mathcal{S}_j \mathcal{Q}_j = \sum_{j=1}^{J} \mathcal{I}_j \mathcal{S}_j \mathcal{Q}_j \qquad \text{and} \qquad \mathcal{B}_{\text{PSC}}\mathcal{A} = \sum_{j=1}^{J} \mathcal{S}_j \mathcal{Q}_j \mathcal{A} = \sum_{j=1}^{J} \mathcal{T}_j. \tag{35}$$

If $\mathcal{S}_j$'s $(j = 1, \ldots, J)$ are all SPD, then the preconditioner $\mathcal{B}_{\text{PSC}}$ is also SPD. If each subspace $V_j = \text{span}\{\phi_j\}$ $(j = 1, \ldots, N)$, then the resulting PSC methods with $\mathcal{S}_j = \omega(\cdot, \phi_j)\phi_j$ and $\mathcal{S}_j = \mathcal{A}_j^{-1}$ correspond to the Richardson method and the Jacobi method, respectively.

Generalized GS Method

Define a weighted GS method $B_\omega = (\omega^{-1}D + L)^{-1}$. We have

$$B_\omega^{-T} + B_\omega^{-1} - A = (\omega^{-1}D + L)^T + (\omega^{-1}D + L) - (D + L + U) = (2\omega^{-1} - 1)D.$$

We assume that there is an invertible smoother or a local relaxation method S for the equation $A\vec{u} = \vec{f}$.

We can define a generalized or modified GS method:

$$B := \left(S^{-1} + L\right)^{-1}. \tag{36}$$

Since $K = B^{-T} + B^{-1} - A$ is symmetric and $\overline{B} = B^T K B$. If B is defined as (36), we have

$$K = (S^{-T} + U) + (S^{-1} + L) - (D + L + U) = S^{-T} + S^{-1} - D.$$

From the definition of K, we notice $B^{-1} = K + A - B^{-T}$. Hence we get an explicit form of $\overline{B}^{-1}$:

$$\overline{B}^{-1} = (K + A - B^{-T})K^{-1}(K + A - B^{-1}) = A + (A - B^{-T})K^{-1}(A - B^{-1}).$$

This identity and the definition of B yield:

$$\left(\overline{B}^{-1}\vec{v}, \vec{v}\right) = (A\vec{v}, \vec{v}) + \left(K^{-1}(D + U - S^{-1})\vec{v},\ (D + U - S^{-1})\vec{v}\right), \quad \forall \vec{v} \in \mathbb{R}^N.$$

Convergence of Generalized GS Method

Corollary (Convergence rate of generalized GS)

If $K = S^{-T} + S^{-1} - D$ is SPD, then the generalized GS method converges and

$$\|I - BA\|_A^2 = \|I - \overline{B}A\|_A = 1 - \frac{1}{1 + c_0}, \quad \text{with } c_0 := \sup_{\|\vec{v}\|_A = 1} \left\| K^{-\frac{1}{2}} \left(D + U - S^{-1} \right) \vec{v} \right\|^2.$$

Example (Solving 1D Poisson's equation using GS)

If $S = D^{-1}$ and $K = D$ in the above generalized GS method and

$$\|I - BA\|_A^2 = 1 - \frac{1}{1 + c_0}, \quad \text{with } c_0 = \sup_{\vec{v} \in \mathbb{R}^N \setminus \{0\}} \frac{(LD^{-1}U\vec{v}, \vec{v})}{\|\vec{v}\|_A^2}.$$

Asymptotically, we have the following estimate

$$c_0 \leq \sup_{\vec{v} \in \mathbb{R}^N \setminus \{0\}} \frac{\frac{1}{2}\|\vec{v}\|^2}{4 \sin^2\left(\frac{\pi}{2(N+1)}\right) \|\vec{v}\|^2} \sim (N+1)^2 = h^{-2}.$$

Expansion of Original System

NCMIS

Next, we will consider an equivalent block matrix form of subspace correction methods.

- Suppose that the finite dimensional vector space V can be decomposed as the summation of linear vector subspaces, $V_1, V_2, \ldots, V_J$, i.e., $V = \sum_{j=1}^J V_j$.
- We define a new vector space
$$\mathbf{V} := V_1 \times V_2 \times \cdots \times V_J.$$
- Define an operator $\mathbf{\Pi} : \mathbf{V} \mapsto V$ such that
$$\mathbf{\Pi u} = \sum_{j=1}^{J} u_j, \quad \text{where } \mathbf{u} := (u_1, \ldots, u_J)^T \in \mathbf{V}$$
with each component $\mathbf{u}_j = u_j \in V_j$. It is easy to see that $\mathbf{\Pi}$ is surjective.
- This operator can be interpreted as
$$\mathbf{\Pi} = (\mathcal{I}_1, \ldots, \mathcal{I}_J),$$
where $\mathcal{I}_j$ is the natural embedding from V_j to V.

Expanded System

- Hence, we obtain

$$\mathbf{\Pi u} = (\mathcal{I}_1, \ldots, \mathcal{I}_J) \begin{pmatrix} u_1 \\ \vdots \\ u_J \end{pmatrix} = \sum_{j=1}^{J} \mathcal{I}_j u_j = \sum_{j=1}^{J} u_j .$$

- So we have

$$\mathbf{\Pi}^T = \begin{pmatrix} \mathcal{I}_1^T \\ \vdots \\ \mathcal{I}_J^T \end{pmatrix} = \begin{pmatrix} \mathcal{Q}_1 \\ \vdots \\ \mathcal{Q}_J \end{pmatrix} .$$

- Note that $\mathbf{\Pi}\,\mathbf{\Pi}^T \neq \mathcal{I}$ in general.
- Define $\mathbf{A} : \mathbf{V} \mapsto \mathbf{V}$ such that $\mathbf{A}_{i,j} = \mathcal{A}_{i,j} := \mathcal{I}_i^T \mathcal{A} \mathcal{I}_j : V_j \mapsto V_i$.
- Find $\mathbf{u} \in \mathbf{V}$, such that $\mathbf{Au} = \mathbf{f}$, where

$$\mathbf{A} := \mathbf{\Pi}^T \mathcal{A}\, \mathbf{\Pi} = \big(\mathbf{A}_{i,j}\big)_{J\times J} = \begin{pmatrix} \mathcal{A}_{1,1} & \cdots & \mathcal{A}_{1,J} \\ \vdots & \ddots & \vdots \\ \mathcal{A}_{J,1} & \cdots & \mathcal{A}_{J,J} \end{pmatrix}, \quad \mathbf{f} := \mathbf{\Pi}^T f = \begin{pmatrix} \mathcal{I}_1^T f \\ \vdots \\ \mathcal{I}_J^T f \end{pmatrix} \in \mathbf{V}.$$

Block Solvers for Expanded System

If $\mathcal{A}$ is SPD, then $\mathbf{A}$ is a symmetric positive semidefinite (SPSD) operator. Note that $\mathbf{A}$ is usually singular due to its nontrivial null space, $\mathrm{null}(\mathbf{\Pi})$. However, its diagonal entries $\mathcal{A}_j$ $(j = 1, 2, \ldots, J)$ are non-singular, where $\mathcal{A}_j := \mathcal{A}_{j,j}$ $(j = 1, \ldots, J)$.

- The linear stationary iterative methods for the expanded system can be written as

$$\mathbf{u}^{\mathrm{new}} = \mathbf{u}^{\mathrm{old}} + \mathbf{B}(\mathbf{f} - \mathbf{A}\mathbf{u}^{\mathrm{old}}), \tag{37}$$

  where the iterator $\mathbf{B} : \mathbf{V} \mapsto \mathbf{V}$ can be chosen accordingly.

- If $\mathbf{B} = \mathbf{D}^{-1}$, then we have the block Jacobi method.
- If $\mathbf{B} = (\mathbf{D} + \mathbf{L})^{-1}$, then we have the block Gauss–Seidel method.
- Assume there is a non-singular block diagonal smoother (or relaxation operator) $\mathbf{S} : \mathbf{V} \mapsto \mathbf{V}$, i.e.,

$$\mathbf{S} = \mathrm{diag}(\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_J), \quad \text{with } \mathcal{S}_j : V_j \mapsto V_j, \quad j = 1, 2, \ldots, J.$$

- Modified block Jacobi method $\mathbf{B} = \mathbf{S}$ and modified block GS method $\mathbf{B} = (\mathbf{S}^{-1} + \mathbf{L})^{-1}$.

Block Solvers for Expanded System

Theorem (Solution of expanded and original systems)

The linear stationary iteration (37) for the expanded system reduces to an equivalent stationary iteration (11) with the iterator

$$\mathcal{B} = \mathbf{\Pi}\,\mathbf{B}\,\mathbf{\Pi}^T$$

for the original equation. Moreover, these two methods have the same convergence speed, namely,

$$\|\mathcal{I} - \mathcal{B}\mathcal{A}\|_{\mathcal{A}} = |\mathbf{I} - \mathbf{B}\mathbf{A}|_{\mathbf{A}}.$$

- The iterators $\mathcal{B}$ and $\mathbf{B}$ define methods which are equivalent to each other.
- Block solvers (37) for the expanded system $\sim$ MSC for the original system.
- However, $\mathbf{A}$ is oftentimes singular and has multiple solutions.
- It seems useless in practice, except we have special reasons.
- Next, we show a couple of classical examples.

Jacobi, PSC, AS, and Block Jacobi

NCMIS

Example (Block Jacobi method and PSC)

We now apply the block Jacobi method for the expanded system, i.e.,

$$\mathbf{u}^{\text{new}} = \mathbf{u}^{\text{old}} + \mathbf{D}^{-1}(\mathbf{f} - \mathbf{A}\mathbf{u}^{\text{old}}).$$

We notice that

$$\mathbf{D}^{-1}\mathbf{A} = \mathbf{D}^{-1}\mathbf{\Pi}^T \mathcal{A} \mathbf{\Pi},$$

which is spectrally equivalent to $\mathbf{\Pi}\mathbf{D}^{-1}\mathbf{\Pi}^T\mathcal{A}$ because $\sigma(\mathcal{B}\mathcal{A})\backslash\{0\} = \sigma(\mathcal{A}\mathcal{B})\backslash\{0\}$.

In fact, from the above theorem, we can see that the above iterative method is equivalent to

$$u^{\text{new}} = u^{\text{old}} + \mathbf{\Pi}\mathbf{D}^{-1}\mathbf{\Pi}^T(f - \mathcal{A}u^{\text{old}}) = u^{\text{old}} + \sum_{j=1}^{J} \mathcal{I}_j \mathcal{A}_j^{-1} \mathcal{I}_j^T \left(f - \mathcal{A}u^{\text{old}}\right).$$

We immediately recognize that this is the PSC method (or the additive Schwarz method) with exact subspace solvers.

Gauss-Seidel, SSC, MS, and Block GS

NCMIS

Example (Block G-S method and SSC)

Similar to the above example, the block G-S method is just the SSC method (or the multiplicative Schwarz method) for the original problem. We now apply the block G-S method for the expanded system, i.e.,

$$\mathbf{u}^{\text{new}} = \mathbf{u}^{\text{old}} + (\mathbf{D}+\mathbf{L})^{-1}(\mathbf{f} - \mathbf{A}\mathbf{u}^{\text{old}}).$$

We can rewrite this method as

$$(\mathbf{D}+\mathbf{L})\mathbf{u}^{\text{new}} = (\mathbf{D}+\mathbf{L})\mathbf{u}^{\text{old}} + (\mathbf{f} - \mathbf{A}\mathbf{u}^{\text{old}}).$$

Hence we have

$$\mathbf{D}\mathbf{u}^{\text{new}} = \mathbf{D}\mathbf{u}^{\text{old}} + \mathbf{f} - \mathbf{L}\mathbf{u}^{\text{new}} - (\mathbf{D}+\mathbf{U})\mathbf{u}^{\text{old}};$$

in turn, we get

$$\mathbf{u}^{\text{new}} = \mathbf{u}^{\text{old}} + \mathbf{D}^{-1}\Big(\mathbf{f} - \mathbf{L}\mathbf{u}^{\text{new}} - (\mathbf{D}+\mathbf{U})\mathbf{u}^{\text{old}}\Big).$$

We can check that the block G-S method is just the SSC method with exact subspace solvers $\mathcal{S}_j = \mathcal{A}_j^{-1}$ for the original linear equation (10).

Generalized Block Gauss-Seidel Method

Define a general or modified block G-S method:

$$\mathbf{B} := \left(\mathbf{S}^{-1} + \mathbf{L}\right)^{-1}. \tag{38}$$

Let $\mathbf{K} := \mathbf{B}^{-T} + \mathbf{B}^{-1} - \mathbf{A}$ and the symmetrization operator as $\overline{\mathbf{B}} = \mathbf{B}^T\mathbf{K}\mathbf{B}$. Then we get

$$\left(\overline{\mathbf{B}}^{-1}\mathbf{v}, \mathbf{v}\right) = \left(\mathbf{B}^{-1}\mathbf{K}^{-1}\mathbf{B}^{-T}\mathbf{v}, \mathbf{v}\right) = \left((\mathbf{S}^{-1} + \mathbf{L})\mathbf{K}^{-1}(\mathbf{S}^{-T} + \mathbf{U})\mathbf{v},\ \mathbf{v}\right), \quad \forall \mathbf{v} \in \mathbf{V} \tag{39}$$

By the definition of $\mathbf{K}$, it is clear that $\mathbf{K}$ is diagonal and

$$\begin{aligned}\mathbf{K} = (\mathbf{S}^{-T} + \mathbf{U}) + (\mathbf{S}^{-1} + \mathbf{L}) - (\mathbf{D} + \mathbf{L} + \mathbf{U}) \\ = \mathbf{S}^{-T} + \mathbf{S}^{-1} - \mathbf{D} = \mathbf{S}^{-T}\left(\mathbf{S}^T + \mathbf{S} - \mathbf{S}^T\mathbf{D}\mathbf{S}\right)\mathbf{S}^{-1}.\end{aligned}$$

Hence, its inverse matrix is also diagonal and

$$\mathbf{K}^{-1} = \mathbf{S}(\mathbf{S}^T + \mathbf{S} - \mathbf{S}^T\mathbf{D}\mathbf{S})^{-1}\mathbf{S}^T. \tag{40}$$

Since $\mathbf{B}^{-1} = \mathbf{K} + \mathbf{A} - \mathbf{B}^{-T}$, we have a representation of $\overline{\mathbf{B}}^{-1}$ by simple manipulations:

$$\overline{\mathbf{B}}^{-1} = (\mathbf{K} + \mathbf{A} - \mathbf{B}^{-T})\mathbf{K}^{-1}(\mathbf{K} + \mathbf{A} - \mathbf{B}^{-1}) = \mathbf{A} + (\mathbf{A} - \mathbf{B}^{-T})\mathbf{K}^{-1}(\mathbf{A} - \mathbf{B}^{-1}).$$

Convergence of Modified BGS

Suppose $V = \sum_{j=1}^{J} V_j$. It is clear that $\mathbf{\Pi u} = \sum_{j=1}^{J} \mathcal{I}_j \mathbf{u}_j$ and $\mathbf{\Pi} : \mathbf{V} \mapsto V$ is surjective.

Lemma (Technical Lemma)

If the iterator $\mathbf{B}$ in (37) is SPD, then $\mathcal{B} = \mathbf{\Pi}\,\mathbf{B}\,\mathbf{\Pi}^T$ is also SPD and

$$(\mathcal{B}^{-1}v, v) = \inf_{\substack{\mathbf{v}\in\mathbf{V} \\ \mathbf{\Pi v}=v}} (\mathbf{B}^{-1}\mathbf{v}, \mathbf{v}), \quad \forall v \in V.$$

The last equality and (38) immediately yield another important identity:

$$\left(\overline{\mathbf{B}}^{-1}\mathbf{v}, \mathbf{v}\right) = (\mathbf{A}\mathbf{v}, \mathbf{v}) + \left(\mathbf{K}^{-1}(\mathbf{D} + \mathbf{U} - \mathbf{S}^{-1})\mathbf{v},\ (\mathbf{D} + \mathbf{U} - \mathbf{S}^{-1})\mathbf{v}\right), \quad \forall \mathbf{v} \in \mathbf{V}. \tag{41}$$

Theorem (Convergence rate of modified block G-S)

If $\mathbf{K} := \mathbf{S}^{-T} + \mathbf{S}^{-1} - \mathbf{D}$ is SPD, then the modified block G-S method converges and

$$|\mathbf{I} - \mathbf{B}\mathbf{A}|_{\mathbf{A}}^2 = 1 - \frac{1}{1 + c_0}, \quad \text{with } c_0 := \sup_{|\mathbf{v}|_{\mathbf{A}}=1} \left\|\mathbf{K}^{-\frac{1}{2}}(\mathbf{D} + \mathbf{U} - \mathbf{S}^{-1})\mathbf{v}\right\|^2.$$

Proof of the Technical Lemma

(1) It is clear that $(\mathcal{B}v, v) \geq 0$ for any $v \in V$ due to positive definiteness of $\mathbf{B}$. Furthermore, we have

$$0 = (\mathcal{B}v, v) = (\mathbf{B}\mathbf{\Pi}^T v, \mathbf{\Pi}^T v) \quad \Longrightarrow \quad \mathbf{\Pi}^T v = 0 \quad \Longrightarrow \quad v \in \text{null}(\mathbf{\Pi}^T) = \text{range}(\mathbf{\Pi})^{\perp}.$$

Since $\mathbf{\Pi}$ is surjective, we have $v = 0$. This proves the iterator $\mathcal{B}$ is SPD.

(2) Define $\mathbf{v}_* := \mathbf{B}\mathbf{\Pi}^T \mathcal{B}^{-1} v$. It is easy to see that

$$\mathbf{\Pi}\mathbf{v}_* = \mathbf{\Pi}\mathbf{B}\mathbf{\Pi}^T \mathcal{B}^{-1} v = \mathcal{B}\mathcal{B}^{-1} v = v, \quad \forall v \in V,$$

and

$$(\mathbf{B}^{-1}\mathbf{v}_*, \mathbf{w}) = (\mathbf{\Pi}^T \mathcal{B}^{-1} v, \mathbf{w}) = (\mathcal{B}^{-1} v, \mathbf{\Pi}\mathbf{w}).$$

If $\mathbf{w} \in \text{null}(\mathbf{\Pi})$, then $(\mathbf{B}^{-1}\mathbf{v}_*, \mathbf{w}) = 0$. This ensures that, for any vector $\mathbf{v} \in \mathbf{V}$, there exists a $\mathbf{B}^{-1}$-orthogonal decomposition $\mathbf{v} = \mathbf{v}_* + \mathbf{w}$ with $\mathbf{w} \in \text{null}(\mathbf{\Pi})$.

(3) Hence, we get $(\mathbf{B}^{-1}\mathbf{v}, \mathbf{v}) = \big(\mathbf{B}^{-1}(\mathbf{v}_* + \mathbf{w}), \mathbf{v}_* + \mathbf{w}\big) = \big(\mathbf{B}^{-1}\mathbf{v}_*, \mathbf{v}_*\big) + \big(\mathbf{B}^{-1}\mathbf{w}, \mathbf{w}\big)$. Thus

$$\begin{aligned}
\inf_{\substack{\mathbf{v} \in \mathbf{V} \\ \mathbf{\Pi}\mathbf{v} = v}} (\mathbf{B}^{-1}\mathbf{v}, \mathbf{v}) &= \big(\mathbf{B}^{-1}\mathbf{v}_*, \mathbf{v}_*\big) + \inf_{\mathbf{w} \in \text{null}(\mathbf{\Pi})} \big(\mathbf{B}^{-1}\mathbf{w}, \mathbf{w}\big) \\
&= \big(\mathbf{B}^{-1}\mathbf{v}_*, \mathbf{v}_*\big) = \big(\mathbf{\Pi}^T \mathcal{B}^{-1} v, \mathbf{B}\mathbf{\Pi}^T \mathcal{B}^{-1} v\big) = \big(\mathcal{B}^{-1} v, v\big).
\end{aligned}$$

Convergence Results of SSC

Theorem (XZ Identity)

Assume that $\mathcal{B}$ is defined by the SSC Algorithm and, for $j = 1, \ldots, J$,

$$\mathbf{w}_j := \mathcal{A}_j \mathit{\Pi}_j \sum_{i \geq j} \mathbf{v}_i - \mathcal{S}_j^{-1} \mathbf{v}_j .$$

If $\mathcal{S}_j^{-T} + \mathcal{S}_j^{-1} - \mathcal{A}_j$ are SPD's for $j = 1, \ldots, J$, then

$$\|\mathcal{I} - \mathcal{B}\mathcal{A}\|_{\mathcal{A}}^2 = 1 - \frac{1}{1 + c_0} = 1 - \frac{1}{c_1}, \tag{42}$$

where

$$c_0 := \sup_{\|v\|_{\mathcal{A}}=1} \inf_{\sum_j \mathbf{v}_j = v} \sum_{j=1}^{J} \|\mathcal{S}_j^T \mathbf{w}_j\|_{\overline{\mathcal{S}}_j^{-1}}^2 \tag{43}$$

and

$$c_1 := \sup_{\|v\|_{\mathcal{A}}=1} \inf_{\sum_j \mathbf{v}_j = v} \sum_{j=1}^{J} \left\|\overline{\mathcal{S}}_j \mathcal{S}_j^{-1} \mathbf{v}_j + \mathcal{S}_j^T \mathbf{w}_j\right\|_{\overline{\mathcal{S}}_j^{-1}}^2 . \tag{44}$$

Some Remarks on XZ Identity

We have introduced operators $\mathcal{T}_j = \mathcal{T}_{\mathcal{S}_j} := \mathcal{S}_j\mathcal{A}_j : V_j \mapsto V_j$ earlier. Hence

$$\mathcal{T}_{\overline{\mathcal{S}}_j} := \overline{\mathcal{S}}_j\mathcal{A}_j = \mathcal{T}_j + \mathcal{T}_j^* - \mathcal{T}_j^*\mathcal{T}_j.$$

Furthermore,

$$\mathcal{S}_j^{-T}\mathbf{v}_j + \sum_{i>j}\mathcal{Q}_j\mathcal{A}\mathcal{I}_i\mathbf{v}_i = \mathcal{A}_j\left(\mathcal{S}_j^T\mathcal{A}_j\right)^{-1}\mathbf{v}_j + \mathcal{A}_j\mathit{\Pi}_j\sum_{i>j}\mathbf{v}_i = \mathcal{A}_j\left[\left(\mathcal{T}_j^*\right)^{-1}\mathbf{v}_j + \mathit{\Pi}_j\sum_{i>j}\mathbf{v}_i\right]$$

$$\Longrightarrow \quad \left(\mathcal{S}_j^{-1} + \mathcal{S}_j^{-T} - \mathcal{A}_j\right)^{-1}\mathcal{A}_j = \left(\mathcal{T}_j^{-1} + (\mathcal{T}_j^*)^{-1} - \mathcal{I}_j\right)^{-1} = \mathcal{T}_j\mathcal{T}_{\overline{\mathcal{S}}_j}^{-1}\mathcal{T}_j^*.$$

Theorem (Another form of XZ identity)

We can rewrite the above estimate (44) in another form:

$$c_1 = \sup_{\|v\|_{\mathcal{A}}=1}\inf_{\sum_j \mathbf{v}_j = v}\sum_{j=1}^{J}\left\|\mathcal{T}_{\overline{\mathcal{S}}_j}^{-\frac{1}{2}}\left(\mathbf{v}_j + \mathcal{T}_j^*\mathit{\Pi}_j\sum_{i>j}\mathbf{v}_i\right)\right\|_{\mathcal{A}}^2. \tag{45}$$

Employing exact subspace solvers $\quad\Longrightarrow\quad c_1 = \sup_{\|v\|_{\mathcal{A}}=1}\inf_{\sum_j \mathbf{v}_j = v}\sum_{j=1}^{J}\left\|\mathit{\Pi}_j\sum_{i\geq j}\mathbf{v}_i\right\|_{\mathcal{A}}^2$

Application of XZ Identity: Linear Stationary Method

Example (Linear stationary iterative method)

One-level linear stationary iterative method

$$u^{\text{new}} = u^{\text{old}} + \overline{\mathcal{S}}(f - \mathcal{A}u^{\text{old}}),$$

can be viewed as a special subspace correction method with only one subspace V. Hence, using (45), we immediately have

$$c_1 = \sup_{\|v\|_{\mathcal{A}}=1} \left\| \mathcal{T}_{\overline{\mathcal{S}}}^{-\frac{1}{2}} v \right\|_{\mathcal{A}}^2 = \sup_{\|v\|_{\mathcal{A}}=1} \left((\overline{\mathcal{S}}\mathcal{A})^{-1} v, v \right)_{\mathcal{A}} = \sup_{\|v\|_{\mathcal{A}}=1} \left(\overline{\mathcal{S}}^{-1} v, v \right),$$

which is exactly the convergence rate given in Theorem 20.

Application of XZ Identity: TG Method

Example (Two-grid method)

Theorem 33 can be viewed as a special case of the XZ identity in the case of space decomposition with two subspaces, i.e., $V = V_c + V$. Suppose we use $\mathcal{A}_c^{-1}$ and $\overline{\mathcal{S}}$ as subspace solvers, respectively. According to (45), we get

$$c_1 = \sup_{\|w\|_{\mathcal{A}}=1} \inf_{\substack{w=v_c+v \\ v_c\in V_c, v\in V}} \|v_c + \Pi_c v\|_{\mathcal{A}}^2 + \|(\overline{\mathcal{S}}\mathcal{A})^{-\frac{1}{2}} v\|_{\mathcal{A}}^2.$$

We can prove that

$$c_1 = \sup_{\|v\|_{\mathcal{A}}=1} \left\| \mathcal{T}_{\overline{\mathcal{S}}}^{-\frac{1}{2}} (\mathcal{I} - \mathcal{Q}_{\overline{\mathcal{S}}^{-1}}) v \right\|_{\mathcal{A}}^2,$$

which coincides with (32) in Theorem 33.

Remark: For a complete proof of this result, we refer to Zikatanov [Zikatanov 2008].

Application of XZ Identity: Alternating Projection Method

- Define $\Theta_j := \mathcal{I} - \mathit{\Pi}_j : V \mapsto V_j^{\perp} =: U_j$. Now we can define a projection

$$\Theta_0 : V \mapsto U_0, \quad \text{where } U_0 := \bigcap_{j=1}^{J} U_j.$$

- We notice that $\Theta_j \Theta_0 = \Theta_0$. From the XZ identity with exact subspace solvers, we have

$$\left\| \Pi_{j=J:-1:1} \Theta_j \right\|_{\mathcal{A}}^2 = \left\| \Pi_{j=J:-1:1} (\mathcal{I} - \mathit{\Pi}_j) \right\|_{\mathcal{A}}^2 = 1 - \frac{1}{1+c_0}$$

$$\Longrightarrow \quad \left\| \Pi_{j=J:-1:1} \Theta_j (\mathcal{I} - \Theta_0) v \right\|_{\mathcal{A}}^2 \le \frac{c_0}{1+c_0} \left\| (\mathcal{I} - \Theta_0) v \right\|_{\mathcal{A}}^2.$$

- Hence, we have

$$\left\| \left(\Pi_{j=J:-1:1} \Theta_j - \Theta_0 \right) v \right\|_{\mathcal{A}}^2 \le \frac{c_0}{1+c_0} \|v\|_{\mathcal{A}}^2.$$

- Besides, we have $\left(\Pi_{j=J:-1:1} \Theta_j - \Theta_0 \right)^k = \left(\Pi_{j=J:-1:1} \Theta_j \right)^k - \Theta_0$ and

$$\lim_{k \to \infty} \left(\Pi_{j=J:-1:1} \Theta_j \right)^k = \Theta_0.$$

Proof of XZ Identity

(1) From (41), we have, for any $\mathbf{v} \in \mathbf{V}$, that

$$\left(\overline{\mathbf{B}}^{-1}\mathbf{v}, \mathbf{v}\right) = (\mathbf{A}\mathbf{v}, \mathbf{v}) + \left(\mathbf{K}^{-1}(\mathbf{D} + \mathbf{U} - \mathbf{S}^{-1})\mathbf{v},\ (\mathbf{D} + \mathbf{U} - \mathbf{S}^{-1})\mathbf{v}\right).$$

By simple calculations, we get

$$\begin{aligned}
\left(\mathbf{D} + \mathbf{U}\right)\mathbf{v} &= \left(\sum_{j\geq 1} \mathcal{Q}_1 \mathcal{A} \mathcal{Q}_j^T \mathbf{v}_j,\ \sum_{j\geq 2} \mathcal{Q}_2 \mathcal{A} \mathcal{Q}_j^T \mathbf{v}_j,\ \cdots\right)^T \\
&= \left(\sum_{j\geq 1} \mathcal{A}_1 \mathit{\Pi}_1 \mathcal{I}_j \mathbf{v}_j,\ \sum_{j\geq 2} \mathcal{A}_2 \mathit{\Pi}_2 \mathcal{I}_j \mathbf{v}_j,\ \cdots\right)^T \\
&= \left(\mathcal{A}_1 \mathit{\Pi}_1 \sum_{j\geq 1} \mathbf{v}_j,\ \mathcal{A}_2 \mathit{\Pi}_2 \sum_{j\geq 2} \mathbf{v}_j,\ \cdots\right)^T.
\end{aligned}$$

Hence we can denote

$$\left(\mathbf{D} + \mathbf{U} - \mathbf{S}^{-1}\right)\mathbf{v} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_J)^T, \quad \text{with } \mathbf{w}_j := \mathcal{A}_j \mathit{\Pi}_j \sum_{i\geq j} \mathbf{v}_i - \mathcal{S}_j^{-1} \mathbf{v}_j.$$

Proof of XZ Identity, Continued

Due to (40) and the fact that $\mathbf{K}$ is diagonal, we have

$$\left(\mathbf{K}^{-1}(\mathbf{D}+\mathbf{U}-\mathbf{S}^{-1})\mathbf{v},\ (\mathbf{D}+\mathbf{U}-\mathbf{S}^{-1})\mathbf{v}\right) = \sum_{j=1}^{J}\left(\mathcal{S}_j\overline{\mathcal{S}}_j^{-1}\mathcal{S}_j^T\mathbf{w}_j, \mathbf{w}_j\right) = \sum_{j=1}^{J}\left\|\mathcal{S}_j^T\mathbf{w}_j\right\|_{\overline{\mathcal{S}}_j^{-1}}^2,$$

where $\overline{\mathcal{S}}_j := \mathcal{S}_j^T + \mathcal{S}_j - \mathcal{S}_j^T\mathcal{A}_j\mathcal{S}_j$. For any $v \in V$, that

$$\sup_{\|v\|_{\mathcal{A}}=1}\inf_{\mathbf{\Pi v}=v}\left(\overline{\mathbf{B}}^{-1}\mathbf{v},\mathbf{v}\right) = 1 + \sup_{\|v\|_{\mathcal{A}}=1}\inf_{\mathbf{\Pi v}=v}\sum_{j=1}^{J}\left\|\mathcal{S}_j^T\mathbf{w}_j\right\|_{\overline{\mathcal{S}}_j^{-1}}^2.$$

By applying Theorem 20 and Lemma 45, we know

$$\|\mathcal{I}-\mathcal{BA}\|_{\mathcal{A}}^2 = 1-\left(\sup_{\|v\|_{\mathcal{A}}=1}\left(\overline{\mathcal{B}}^{-1}v,v\right)\right)^{-1} = 1-\left(\sup_{\|v\|_{\mathcal{A}}=1}\inf_{\mathbf{\Pi v}=v}\left(\overline{\mathbf{B}}^{-1}\mathbf{v},\mathbf{v}\right)\right)^{-1}. \tag{46}$$

This gives the desired estimate for the constant c_0.

Proof of XZ Identity, Continued

(2) On the other hand, from (39), we have

$$\begin{aligned}\left(\overline{\mathbf{B}}^{-1}\mathbf{v},\mathbf{v}\right) &= \left(\mathbf{K}^{-1}(\mathbf{S}^{-T}+\mathbf{U})\mathbf{v},\ (\mathbf{S}^{-T}+\mathbf{U})\mathbf{v}\right) \\ &= \sum_{j=1}^{J}\left\|(\mathcal{S}_j^{-1}+\mathcal{S}_j^{-T}-\mathcal{A}_j)^{-\frac{1}{2}}(\mathcal{S}_j^{-T}\mathbf{v}_j+\sum_{i>j}\mathcal{Q}_j\mathcal{A}\mathcal{I}_i\mathbf{v}_i)\right\|^2. \end{aligned} \tag{47}$$

We notice that

$$\begin{aligned}\mathcal{S}_j^{-T}\mathbf{v}_j+\sum_{i>j}\mathcal{Q}_j\mathcal{A}\mathcal{I}_i\mathbf{v}_i = \mathcal{S}_j^{-T}\mathbf{v}_j+\mathcal{A}_j\Pi_j\sum_{i>j}\mathbf{v}_i &= (\mathcal{S}_j^{-T}+\mathcal{S}_j^{-1}-\mathcal{A}_j)\mathbf{v}_j+\mathbf{w}_j \\ &= \mathcal{S}_j^{-T}\overline{\mathcal{S}}_j\mathcal{S}_j^{-1}\mathbf{v}_j+\mathbf{w}_j = \mathcal{S}_j^{-T}\left(\overline{\mathcal{S}}_j\mathcal{S}_j^{-1}\mathbf{v}_j+\mathcal{S}_j^{T}\mathbf{w}_j\right).\end{aligned}$$

Plugging this into the previous identity, we get

$$\left(\overline{\mathbf{B}}^{-1}\mathbf{v},\mathbf{v}\right)=\sum_{j=1}^{J}\left\|\overline{\mathcal{S}}_j\mathcal{S}_j^{-1}\mathbf{v}_j+\mathcal{S}_j^{T}\mathbf{w}_j\right\|^2_{\overline{\mathcal{S}}_j^{-1}}.$$

Relation Between PSC and SSC

NCMIS

Theorem (PSC and SSC)

If $\mathcal{S}_j = \mathcal{A}_j^{-1}$ for all j and V_j are subspaces of V, then there exists a constant c_* depends only on topology of the overlaps between the subspaces such that

$$\frac{1}{4}\big(\mathcal{B}_{\text{PSC}}^{-1} v, v\big) \le \big(\overline{\mathcal{B}}_{\text{SSC}}^{-1} v, v\big) \le c_* \big(\mathcal{B}_{\text{PSC}}^{-1} v, v\big), \quad \forall\, v \in V.$$

Sketch of proof: Given $v = \sum_{j=1}^J v_j$ with $v_j \in V_j$. It follows that

$$\|v\|_{\mathcal{A}}^2 = \sum_{k,j=1}^J (v_k, v_j)_{\mathcal{A}} = \sum_{k=1}^J \Big((v_k, v_k)_{\mathcal{A}} + 2 \sum_{j>k}^J (v_k, v_j)_{\mathcal{A}} \Big) = 2 \sum_{k=1}^J \sum_{j \ge k}^J (v_k, v_j)_{\mathcal{A}} - \sum_{k=1}^J (v_k, v_k)_{\mathcal{A}}.$$

Since Π_k is an $\mathcal{A}$-projection, it follows that

$$\sum_{k=1}^J \|v_k\|_{\mathcal{A}}^2 \le 2 \sum_{k=1}^J \Big(v_k, \sum_{j=k}^J v_j \Big)_{\mathcal{A}} = 2 \sum_{k=1}^J \Big(v_k, \Pi_k \sum_{j=k}^J v_j \Big)_{\mathcal{A}} \le 2 \Big(\sum_{k=1}^J \|v_k\|_{\mathcal{A}}^2 \Big)^{\frac{1}{2}} \Big(\sum_{k=1}^J \Big\| \Pi_k \sum_{j=k}^J v_j \Big\|_{\mathcal{A}}^2 \Big)^{\frac{1}{2}}.$$

In turn, it gives that $\sum_{k=1}^J \|v_k\|_{\mathcal{A}}^2 \le 4 \sum_{k=1}^J \big\| \Pi_k \sum_{j=k}^J v_j \big\|_{\mathcal{A}}^2$. □

Convergence Analysis of PSC

NCMIS

Assumptions:

1. For any $v \in V$, there exists a decomposition $v = \sum_{j=1}^{J} v_j$ with $v_j \in V_j$ such that

$$\sum_{j=1}^{J} \left(\mathcal{S}_j^{-1} v_j, v_j\right) \leq K_1 \left(\mathcal{A}v, v\right);$$

2. For any $u, v \in V$,

$$\sum_{(i,j)} \left(\mathcal{T}_i u, \mathcal{T}_j v\right)_{\mathcal{A}} \leq K_2 \left(\sum_{i=1}^{J} (\mathcal{T}_i u, u)_{\mathcal{A}}\right)^{\frac{1}{2}} \left(\sum_{j=1}^{J} (\mathcal{T}_j v, v)_{\mathcal{A}}\right)^{\frac{1}{2}}.$$

By checking the above two assumptions, we can show the convergence performance of PSC method $\mathcal{B}$:

Theorem (Condition number of PSC)

If the above assumptions hold, the PSC preconditioner (35) satisfies $\kappa(\mathcal{B}\mathcal{A}) \leq K_1 K_2$.

Proof of Condition Number Estimate

(1) Lower Bound:

For any $v \in V$, suppose that $v = \sum_{j=1}^{J} v_j$ is a decomposition that satisfies the first assumption. It is easy to see that

$$
\begin{aligned}
(v,v)_{\mathcal{A}} &= \sum_{j=1}^{J}(v_j, v)_{\mathcal{A}} = \sum_{j=1}^{J}(v_j, \Pi_j v)_{\mathcal{A}} = \sum_{j=1}^{J}(v_j, \mathcal{A}_j \Pi_j v) = \sum_{j=1}^{J}\left(\mathcal{S}_j^{-\frac{1}{2}} v_j, \mathcal{S}_j^{\frac{1}{2}} \mathcal{A}_j \Pi_j v\right) \\
&\le \sum_{j=1}^{J}\left(\mathcal{S}_j^{-1} v_j, v_j\right)^{\frac{1}{2}} \left(\mathcal{S}_j \mathcal{A}_j \Pi_j v, \mathcal{A}_j \Pi_j v\right)^{\frac{1}{2}} = \sum_{j=1}^{J}\left(\mathcal{S}_j^{-1} v_j, v_j\right)^{\frac{1}{2}} \left(\mathcal{S}_j \mathcal{A}_j \Pi_j v, v\right)_{\mathcal{A}}^{\frac{1}{2}} \\
&\le \left(\sum_{j=1}^{J}\left(\mathcal{S}_j^{-1} v_j, v_j\right)\right)^{\frac{1}{2}} \left(\sum_{j=1}^{J}\left(\mathcal{T}_j v, v\right)_{\mathcal{A}}\right)^{\frac{1}{2}} \le \sqrt{K_1}\, \|v\|_{\mathcal{A}} \left(\mathcal{B}\mathcal{A}v, v\right)_{\mathcal{A}}^{\frac{1}{2}}.
\end{aligned}
$$

Consequently, we have the lower bound

$$
\frac{1}{K_1}(v,v)_{\mathcal{A}} \le (\mathcal{B}\mathcal{A}v, v)_{\mathcal{A}}, \quad \forall\, v \in V.
$$

Proof of Condition Number Estimate, Continued

(2) Upper Bound:

From the second assumption, we have

$$\|\mathcal{B}\mathcal{A}v\|_{\mathcal{A}}^2 = \sum_{i,j=1}^{J} \big(\mathcal{T}_i v, \mathcal{T}_j v\big)_{\mathcal{A}} \le K_2(\mathcal{B}\mathcal{A}v, v)_{\mathcal{A}} \le K_2\|\mathcal{B}\mathcal{A}v\|_{\mathcal{A}}\|v\|_{\mathcal{A}}.$$

So with some calculation, we can obtain the upper bound

$$(\mathcal{B}\mathcal{A}v, v)_{\mathcal{A}} \le K_2(v, v)_{\mathcal{A}}, \quad \forall\, v \in V.$$

Thus Lemmas 10 and 11 yield the desired estimate. □

Remark (Similar estimate for SSC)

With the same assumptions, we can also show that the SSC method also converges with

$$\|\mathcal{I} - \mathcal{B}\mathcal{A}\|_{\mathcal{A}}^2 \le 1 - \frac{2-\omega_1}{K_1(1+K_2)^2} \quad \text{and} \quad \omega_1 := \max_j \rho(\mathcal{S}_j\mathcal{A}_j) = \max_j \rho(\mathcal{T}_j). \tag{48}$$

Estimates of K_1

Lemma (Estimates of K_1)

Assume that, for any $v \in V$, there exists a decomposition $v = \sum_{j=1}^{J} v_j$ with $v_j \in V_j$:

(i) If the decomposition satisfies that

$$\sum_{j=1}^{J} (v_j, v_j)_{\mathcal{A}} \le C_1 (v, v)_{\mathcal{A}},$$

then we have

$$K_1 \le C_1/\omega_0, \quad \text{where} \ \ \omega_0 := \min_{j=1,\ldots,J} \big\{ \lambda_{\min}(\mathcal{S}_j \mathcal{A}_j) \big\};$$

(ii) If $\rho_j := \rho(\mathcal{A}_j)$ and

$$\sum_{j=1}^{J} \rho_j (v_j, v_j) \le \hat{C}_1 (v, v)_{\mathcal{A}},$$

then we have

$$K_1 \le \hat{C}_1/\hat{\omega}_0, \quad \text{where} \ \ \hat{\omega}_0 := \min_{j=1,\ldots,J} \big\{ \rho_j \lambda_{\min}(\mathcal{S}_j) \big\}.$$

Estimates of K_2

We introduce a nonnegative symmetric matrix

$$\Sigma = \left(\sigma_{i,j}\right) \in \mathbb{R}^{J\times J}, \tag{49}$$

where each entry $\sigma_{i,j}$ is the smallest constant such that

$$\left(\mathcal{T}_i u, \mathcal{T}_j v\right)_{\mathcal{A}} \le \omega_1 \sigma_{i,j} \left(\mathcal{T}_i u, u\right)_{\mathcal{A}}^{\frac{1}{2}} \left(\mathcal{T}_j v, v\right)_{\mathcal{A}}^{\frac{1}{2}}, \quad \forall\, u, v \in V. \tag{50}$$

Here ω_1 has been defined in (48). It is clear that

- $0 \le \sigma_{i,j} \le 1$.
- $\sigma_{i,j} = 0$, if $\mathit{\Pi}_i \mathit{\Pi}_j = 0$ and exact subspace solvers are used.

Lemma (Estimate of K_2)

The constant $K_2 \le \omega_1 \rho(\Sigma)$. Furthermore, if $\sigma_{i,j} \lesssim \gamma^{|i-j|}$ holds for some parameter $0 < \gamma < 1$, then

$$\rho(\Sigma) \lesssim (1-\gamma)^{-1};$$

in this case, the second assumption is the well-known strengthened Cauchy–Schwarz inequality.

Auxiliary Space Preconditioning

Sometimes, we cannot apply subspace correction methods directly due to difficulties in obtaining an appropriate space decomposition.

We introduce an auxiliary space $\tilde{V}$. Suppose $\Pi : \tilde{V} \mapsto V$ is surjective and satisfies:

- Firstly, Π is stable
$$\|\Pi\tilde{v}\|_{\mathcal{A}} \le C\|\tilde{v}\|_{\tilde{\mathcal{A}}}, \qquad \forall\, \tilde{v} \in \tilde{V}.$$
- Secondly, for any $v \in V$, there exists $\tilde{v} \in \tilde{V}$ such that $\Pi\tilde{v} = v$ and
$$c\|\tilde{v}\|_{\tilde{\mathcal{A}}} \le \|v\|_{\mathcal{A}}.$$

Under the above assumptions, if $\tilde{\mathcal{B}}$ is a SPD preconditioner for $\tilde{\mathcal{A}}$, then $\mathcal{B} = \Pi\,\tilde{\mathcal{B}}\,\Pi^T$ is SPD and

$$\kappa(\mathcal{B}\mathcal{A}) \le \left(\frac{C}{c}\right)^2 \kappa(\tilde{\mathcal{B}}\tilde{\mathcal{A}}).$$

Remark: This result is known as the Fictitious Space Lemma or the Fictitious Domain Lemma.

Construction of Efficient Preconditioners

NCMIS

How to obtain a preconditioner for $\mathcal{A}$? $\quad \|v\|_0^2 \lesssim (\mathcal{A}v, v) \lesssim h^{-2}\|v\|_0^2, \quad \forall v \in V_h.$

MSC $\approx$ Block solvers for the expanded system

- Convergence rate of stationary methods: $c_1 = \sup_{\|v\|_{\mathcal{A}}=1} (\overline{\mathcal{B}}^{-1} v, v)$
- XZ identity for SSC: $c_1 = \sup_{\|v\|_{\mathcal{A}}=1} \inf_{\sum_j \mathbf{v}_j = v} \sum_{j=1}^{J} \left\| \Pi_j \sum_{i \geq j} \mathbf{v}_i \right\|_{\mathcal{A}}^2$
- Convergent iterative method as a preconditioner: $\kappa(\mathcal{B}\mathcal{A}) \leq \dfrac{1+\rho}{1-\rho}$
- Stable decomposition and strengthened Cauchy–Schwarz inequality

Multilevel MSC:

- Introduce a multilevel space decomposition $\Longrightarrow$ Multilevel method of subspace corrections
- Subspace solvers $\Longrightarrow$ Smoothers (local relaxations)
- Recursive calls to two-grid methods $\Longrightarrow$ Apply CGC to deal with smooth error components

Setup Multilevel Methods

As mentioned before, we can apply a general SETUP step for constructing multilevel hierarchy.

Algorithm (Setup step for multigrid methods)

For a given sparse matrix $A \in \mathbb{R}^{N\times N}$, we apply the following steps:

1. Obtain a suitable matrix for coarsening $A_f \in \mathbb{R}^{N_f\times N_f}$ (for example, $A_f = A_{\rm sym}$);
2. Define a coarse space with N_c variables (C/F splitting or aggregation);
3. Construct a prolongation (usually an interpolation) $P \in \mathbb{R}^{N_f\times N_c}$:
   - 3.1. Give a sparsity pattern for the interpolation P;
   - 3.2. Determine weights of the interpolation P;
4. Construct a restriction $R \in \mathbb{R}^{N_c\times N_f}$ (for example, $R = P^T$);
5. Form a coarse-level coefficient matrix (for example, $A_c = RA_fP$);
6. Give a sparse approximation of A_c whenever necessary.