SUBSPACE CORRECTION METHOD AND AUXILIARY SPACE METHOD

LONG CHEN

In this chapter we shall discuss subspace correction method and auxiliary space method developed by Xu [5, 7, 8] on solving the linear operator equation

$$Au = f, \tag{1}$$

posed on a finite dimensional Hilbert space $\mathbb{V} \cong \mathbb{R}^N$ equipped with an inner product $(\cdot, \cdot)$. Here $A : \mathbb{V} \mapsto \mathbb{V}$ is an *symmetric and positive definite (SPD)* operator, $f \in \mathbb{V}$ is given, and we are looking for $u \in \mathbb{V}$ such that (1) holds. How to solve (1) efficiently remains a basic question in numerical PDEs (and in all scientific computing).

1. Space decomposition and subspace correction methods

In the spirit of dividing and conquering, we decompose the space $\mathbb{V}$ as the summation of subspaces and correspondingly decompose the problem (1) into sub-problems with smaller size which are relatively easy to solve. This method is developed by Xu [5].

Let $\mathbb{V}_i \subset \mathbb{V}$, $i = 1, \ldots, J$, be subspaces of $\mathbb{V}$. If $\mathbb{V} = \sum_{i=1}^J \mathbb{V}_i$, then $\{\mathbb{V}_i\}$ is called a *space decomposition* of $\mathbb{V}$. By the definition, for any $u \in \mathbb{V}$, we can decompose u as

$$u = \sum_{i=1}^J u_i, \quad u_i \in \mathbb{V}_i,\ i = 1, \ldots, J.$$

Since $\sum_{i=1}^J \mathbb{V}_i$ is not necessarily a direct sum, decompositions of $u \in \mathbb{V}$ of the form $u = \sum_{i=1}^J u_i$ are in general not unique. We recall some background on the sum of linear spaces in the appendix.

We introduce the following operators for $i = 1, 1, \ldots, J$:

- $I_i : \mathbb{V}_i \hookrightarrow \mathbb{V}$ the natural inclusion;
- $Q_i : \mathbb{V} \mapsto \mathbb{V}_i$ the projection in the inner product $(\cdot, \cdot)$;
- $P_i : \mathbb{V} \mapsto \mathbb{V}_i$ the projection in the inner product $(\cdot, \cdot)_A$;
- $A_i : \mathbb{V}_i \mapsto \mathbb{V}_i$ the restriction of A on the subspace $\mathbb{V}_i \times \mathbb{V}_i$;
- $R_i : \mathbb{V}_i \mapsto \mathbb{V}_i$ an approximation of A_i^{-1} which is often known as smoothers or local subspace solvers.
- $T_i : \mathbb{V} \to \mathbb{V}_i$ $T_i = R_i Q_i A = R_i A_i P_i$.

We then explore relations between these operators. By definition

$$(Q_i^{\intercal} v_i, v) = (v_i, Q_i v) = (v_i, v) = (I_i v_i, v) \qquad \forall\, v_i \in \mathbb{V}_i,\ v \in \mathbb{V},$$

therefore $Q_i^{\intercal}$ coincides with the natural inclusion I_i or equivalently $I_i^{\intercal} = Q_i$. In the continuous level, I_i is the identity operator and thus skipped in many places. In the implementation, the prolongation matrix is the representation of I_i relative to certain bases and the transpose $I_i^{\intercal}$ is the restriction matrix. The matrix or operator A is understood as the bilinear function on $\mathbb{V} \times \mathbb{V}$. Then the restriction on subspaces is $A_i = I_i^{\intercal} A I_i$.

It follows from the definition that

$$A_i P_i = Q_i A,$$

namely the following diagram is commutative

$$\begin{array}{ccc} \mathbb{V} & \xrightarrow{A} & \mathbb{V} \\ \downarrow{\scriptstyle P_i} & & \downarrow{\scriptstyle Q_i} \\ \mathbb{V}_i & \xrightarrow{A_i} & \mathbb{V}_i \end{array}$$

The consistent notation for the smoother R_i is B_i, the iterator for each local problem. But we reserve the notation B for the iterator of the original problem.

Last, let us look at $T_i = R_i Q_i A = R_i A_i P_i$. When $R_i = A_i^{-1}$, from the definition, $T_i = P_i = A_i^{-1} Q_i A$. When $T_i|_{\mathbb{V}_i} : \mathbb{V}_i \to \mathbb{V}_i$, the projection P_i is identity and thus $T_i|_{\mathbb{V}_i} = R_i A_i$. With a slight abuse of notation, we use $T_i^{-1} = (T_i|_{\mathbb{V}_i})^{-1}$. The action of T_i and T_i^{-1} is

$$(T_i u_i, u_i)_A = (R_i A_i u_i, A_i u_i), \quad (T_i^{-1} u, u)_A = (R_i^{-1} u, u).$$

Now we describe the method of subspace correction. For a given residual r, let $r_i = Q_i r$ denote the restriction of the residual to the subspace, we shall solve the residual equation in the subspaces

$$A_i e_i^* = r_i \quad \text{by} \quad e_i = R_i r_i.$$

Subspace corrections e_i are assembled to give a correction in the space $\mathbb{V}$ and therefore the method is called subspace correction method.

Basically there are two ways to assemble subspace corrections.

Parallel Subspace Correction (PSC). This method is to do the correction on each subspace in parallel. In operator form, it is

$$u_{k+1} = u_k + B_a(f - Au_k) = (I - B_a A)u_k + B_a f, \tag{2}$$

where

$$B_a = \sum_{i=1}^{J} I_i R_i I_i^{\intercal}. \tag{3}$$

The subspace correction is $\tilde{e}_i = I_i R_i I_i^{\intercal}(f - Au_k)$, and the correction in $\mathbb{V}$ is $\tilde{e} = \sum_{i=1}^{J} \tilde{e}_i$.

Successive Subspace Correction (SSC). This method is to do the correction in a successive way. In operator form, it reads

$$v^1 = u_k, \quad v^{i+1} = v^i + I_i R_i I_i^{\intercal}(f - Av^i), i = 1, \ldots, J, \quad u_{k+1} = v^{J+1}. \tag{4}$$

The iterator B_m is not easy to formulate.

$$\begin{array}{ccccc} & \nearrow & \mathbb{V}_1 & \searrow & \\ r & \longrightarrow & \mathbb{V}_2 & \longrightarrow & e \\ & \searrow & \mathbb{V}_3 & \nearrow & \end{array}$$

(a) PSC

$$r \to \mathbb{V}_1 \to e \to r \to \mathbb{V}_2 \to e \to r \to \mathbb{V}_3 \to e.$$

(b) SSC

FIGURE 1. Illustration of PSC and SSC.

We have the following error formula for PSC and SSC.

- Parallel Subspace Correction (PSC):

$$u - u_{k+1} = \left[I - \sum_{i=1}^{J} T_i \right] (u - u_k);$$

- Successive Subspace Correction (SSC):

$$u - u_{k+1} = \left[\prod_{i=1}^{J} (I - T_i) \right] (u - u_k).$$

Thus PSC is also called additive methods while SSC is called multiplicative method. In the notation $\prod_{i=1}^{J} a_i$, we assume there is a build-in ordering from $i = 1$ to N i.e. $\prod_{i=1}^{J} a_i = a_0 a_1 \ldots a_N$.

We present algorithms of PSC and SSC in the following form to emphasis it is a procedure to solve the residual equation, i.e., given a residual r, return a correction e. One iteration of PSC or SSC can be used as a preconditioner in PCG.

```
function e = PSC(r)
% Solve the residual equation Ae = r by PSC method
for i = 1:J
    ri = Ii'*r;    % restrict the residual to subspace
    ei = Ri*ri;  % solve the residual equation in subspace
    e = e + Ii*ei;  % prolongate the correction to the big space
end
```

```
function e = SSC(r)
% Solve the residual equation Ae = r by SSC method
rd = r;
for i = 1:J
    ri = Ii'*rd;   % restrict the residual to subspace
    ei = Ri*ri;  % solve the residual equation in subspace
    e = e + Ii*ei;  % prolongate the correction to the big space
    rd = r - A*e;  % update residual
end
```

Comparing the above PSC and SSC functions, one can immediately see that in SSC, the residual is updated for each subspace correction while not in PSC. In terms of rate of convergence, SSC is superior. On the other hand, PSC is embarrassingly parallel while SSC is essentially a sequential method.

Example 1.1. Let us consider the matrix equation

$$Au = f,$$

where A is an $N \times N$ SPD matrix. Let us take the trivial decomposition of $\mathbb{R}^N = \sum_{i=1}^{N} \operatorname{span}\{e_i\}$, where $\{e_i, i = 1, \ldots, N\}$ is the canonical basis of $\mathbb{R}^N$. Then

- for $R_i = \omega I$, PSC is Richardson method;
- for $R_i = A_i^{-1}$, PSC is Jacobi method;
- for $R_i = A_i^{-1}$, SSC is the Gauss-Seidal method.

Later we shall also view PSC as a Jacobi method and SSC as a Gauss-Seidel method for a big system formed in a product space formed by subspaces.

2. Auxiliary Space Methods

In this section, we present a variation of the fictitious space method of Nepomnyaschikh [4] and the auxiliary space method of Xu [6]. We follow the presentation in [1].

Let $\tilde{\mathbb{V}}$ and $\mathbb{V}$ be two Hilbert spaces and let $\Pi : \tilde{\mathbb{V}} \to \mathbb{V}$ be a surjective map. Denoted by $\Pi^\intercal : \mathbb{V} \to \tilde{\mathbb{V}}$ the adajoint of Π in the default inner products

$$(\Pi^\intercal u, \tilde{v}) = (u, \Pi\tilde{v}) \quad \text{for all } u \in \mathbb{V}, \tilde{v} \in \tilde{\mathbb{V}}.$$

Here, to save notation, we use $(\cdot,\cdot)$ for inner products in both $\mathbb{V}$ and $\tilde{\mathbb{V}}$. Since Π is surjective, its transpose $\Pi^\intercal$ is injective.

Given an SPD operator $A : \mathbb{V} \to \mathbb{V}$, let $\tilde{A} = \Pi^\intercal A \Pi : \tilde{\mathbb{V}} \to \tilde{\mathbb{V}}$ be the lift of A. To construct a good approximation of A^{-1}, we can project one for $\tilde{A}$. Let $\tilde{B} : \tilde{\mathbb{V}} \to \tilde{\mathbb{V}}$ be SPD, we can define $B := \Pi\tilde{B}\Pi^\intercal : \mathbb{V} \to \mathbb{V}$ and will show it is also SPD. The relation is summarized in the following diagram

$$\begin{tikzcd} \tilde{\mathbb{V}} \arrow[r, shift left, "\tilde{A}"] \arrow[d, "\Pi"] & \tilde{\mathbb{V}} \arrow[l, shift left, "\tilde{B}"] \\ \mathbb{V} \arrow[r, shift left, "A"] & \mathbb{V} \arrow[l, shift left, "B"] \arrow[u, "\Pi^\intercal"'] \end{tikzcd}$$

Theorem 2.1. *Let $\tilde{\mathbb{V}}$ and $\mathbb{V}$ be two Hilbert spaces and let $\Pi : \tilde{\mathbb{V}} \to \mathbb{V}$ be a surjective map. Let $\tilde{B} : \tilde{\mathbb{V}} \to \tilde{\mathbb{V}}$ be a symmetric and positive definite operator. Then $B := \Pi\tilde{B}\Pi^\intercal : \mathbb{V} \to \mathbb{V}$ is also symmetric and positive definite. Furthermore*

$$(B^{-1}v, v) = \inf_{\Pi\tilde{v}=v} (\tilde{B}^{-1}\tilde{v}, \tilde{v}). \tag{5}$$

Proof. We shall adapt the proof given by Xu and Zikatanov [9] (Lemma 2.4).

It is obvious that B is symmetric and positive semi-definite. Since $\tilde{B}$ is SPD and $\Pi^\intercal$ is injective, $(Bv, v) = (\tilde{B}\Pi^\intercal v, \Pi^\intercal v) = 0$ implies $\Pi^\intercal v = 0$ and $v = 0$. Therefore B is positive definite.

Let $\tilde{v}^* = \tilde{B}\Pi^\intercal B^{-1} v$. Then $\Pi\tilde{v}^* = v$. For any $\tilde{w} \in \tilde{\mathbb{V}}$

$$(\tilde{B}^{-1}\tilde{v}^*, \tilde{w}) = (\Pi^\intercal B^{-1}v, \tilde{w}) = (B^{-1}v, \Pi\tilde{w}).$$

In particular

$$(\tilde{B}^{-1}\tilde{v}^*, \tilde{v}^*) = (B^{-1}v, \Pi\tilde{v}^*) = (B^{-1}v, v).$$

For any $\tilde{v} \in \tilde{\mathbb{V}}$, denoted by $v = \Pi\tilde{v}$, we write $\tilde{v} = \tilde{v}^* + \tilde{w}$ with $\Pi\tilde{w} = 0$. Then

$$\begin{aligned} \inf_{\Pi\tilde{v}=v} (\tilde{B}^{-1}\tilde{v}, \tilde{v}) &= \inf_{\Pi\tilde{w}=0} (\tilde{B}^{-1}(\tilde{v}^* + \tilde{w}), \tilde{v}^* + \tilde{w}) \\ &= (B^{-1}v, v) + \inf_{\Pi\tilde{w}=0} \Big(2(\tilde{B}^{-1}\tilde{v}^*, \tilde{w}) + (\tilde{B}^{-1}\tilde{w}, \tilde{w}) \Big) \\ &= (B^{-1}v, v) + \inf_{\Pi\tilde{w}=0} (\tilde{B}^{-1}\tilde{w}, \tilde{w}) \\ &= (B^{-1}v, v). \end{aligned}$$

□

The symmetric positive definite operator B can be used as a preconditioner for solving $Au = f$ using PCG. To estimate the condition number $\kappa(BA)$, we only need to compare B^{-1} and A.

Lemma 2.2. *For two SPD operators A and B, if $c_0(Av, v) \le (B^{-1}v, v) \le c_1(Av, v)$ for all $v \in \mathbb{V}$, then $\kappa(BA) \le c_1/c_0$.*

Proof. Note that BA is symmetric with respect to A. Therefore

$$\lambda_{\min}^{-1}(BA) = \lambda_{\max}((BA)^{-1}) = \sup_{u\in\mathbb{V}\setminus\{0\}} \frac{((BA)^{-1}u,u)_A}{(u,u)_A} = \sup_{u\in\mathbb{V}\setminus\{0\}} \frac{(B^{-1}u,u)}{(Au,u)}.$$

Therefore $(B^{-1}v,v) \le c_1(Av,v)$ implies $\lambda_{\min}(BA) \ge c_1^{-1}$. Similarly $(B^{-1}v,v) \ge c_0(Av,v)$ implies $\lambda_{\max}(BA) \le c_0^{-1}$. The estimate of $\kappa(BA)$ then follows. □

Theorem 2.3. *Let $\tilde{\mathbb{V}}$ and $\mathbb{V}$ be two Hilbert spaces and let $\Pi : \tilde{\mathbb{V}} \to \mathbb{V}$ be a surjective map. Let $\tilde{B} : \tilde{\mathbb{V}} \to \tilde{\mathbb{V}}$ be a symmetric and positive definite operator and $B = \Pi\tilde{B}\Pi^{\intercal}$. If*

$$c_0(Av,v) \le \inf_{\Pi\tilde{v}=v} (\tilde{B}^{-1}\tilde{v},\tilde{v}) \le c_1(Av,v) \quad \textit{for all } v\in\mathbb{V}, \tag{6}$$

then

$$\kappa(BA) \le c_1/c_0.$$

Remark 2.4. In literature, e.g. the fictitious space lemma of [4], the condition (6) is usually decomposed to the following two conditions:

(1) For any $v\in\mathbb{V}$, there exists a $\tilde{v}\in\tilde{V}$, such that $\Pi\tilde{v}=v$ and $\|\tilde{v}\|_{\tilde{B}^{-1}}^2 \le c_1\|v\|_A^2$.
(2) For any $\tilde{v}\in\tilde{\mathbb{V}}$, $\|\Pi\tilde{v}\|_A^2 \le c_0^{-1}\|\tilde{v}\|_{\tilde{B}^{-1}}^2$.

3. An Auxiliary Space of Product Type

Given a space decomposition $\mathbb{V} = \sum_{i=1}^J \mathbb{V}_i$, we construct an auxiliary space of product type $\tilde{\mathbb{V}} = \mathbb{V}_0 \times \mathbb{V}_1 \times ... \times \mathbb{V}_J$, with the standard inner product for product spaces $(\tilde{u},\tilde{v}) := \sum_{i=1}^J (u_i,v_i)$. We define $\Pi : \tilde{\mathbb{V}} \to \mathbb{V}$ as $\Pi\tilde{u} = \sum_{i=1}^J u_i$. In operator form $\Pi = (I_1, I_2, \cdots, I_J)$ if we treat $\tilde{u} = (u_0,\cdots,u_J)^{\intercal}$ as a column vector. Since $\mathbb{V} = \sum_{i=1}^J \mathbb{V}_i$, the operator Π is surjective.

Let $\tilde{A} = \Pi^{\intercal}A\Pi$ and $\tilde{f} = \Pi^{\intercal}f$. If $\tilde{u}$ is a solution of $\tilde{A}\tilde{u} = \tilde{f}$, by multiplying $\Pi^{\intercal}$ both sides, it is straightforward to verify that then $u = \Pi R\tilde{u}$ is the solution of $Au = f$.

We shall derive PSC and SSC by classical iterative methods of solving $\tilde{A}\tilde{u} = \tilde{f}$. To this purpose, let $R_i : \mathbb{V}_i \to \mathbb{V}_i$ be nonsingular operators, often known as smoothers, approximating A_i^{-1}. Define a diagonal matrix of operators $\tilde{R} = \mathrm{diag}(R_0, R_1, \cdots, R_J) : \tilde{\mathbb{V}} \to \tilde{\mathbb{V}}$ which is also non-singular.

By direct computation, the entry $\tilde{a}_{ij} = Q_iAI_j = A_iP_iI_j$. In particular $\tilde{a}_{ii} = A_i$. The symmetric operator $\tilde{A}$ may be singular with nontrivial kernel $\ker(\Pi)$, but the diagonal of $\tilde{A}$ is always non-singular. Write $\tilde{A} = \tilde{D} + \tilde{L} + \tilde{U}$ where $\tilde{D} = \mathrm{diag}(A_0, A_1, \cdots, A_j)$, $\tilde{L}$ and $\tilde{U}$ are lower and upper triangular matrix of operators, and $\tilde{L}^{\intercal} = \tilde{U}$.

Considering the iteration

$$\tilde{u}_{k+1} = \tilde{u}_k + \tilde{R}(\tilde{f} - \tilde{A}\tilde{u}_k). \tag{7}$$

Let $u_k = \Pi\tilde{u}_k$. Applying Π to (7) and noting that

$$\tilde{f} = \Pi^{\intercal}f, \quad \text{and} \quad \tilde{A}\tilde{u}_k = \Pi^{\intercal}Au_k,$$

we obtain the PSC method

$$u_{k+1} = u_k + \sum_{i=1}^J R_iQ_i(f - Au_k),$$

The multiplicative method is more subtle. Following [3], we shall view the SSC for solving $Au = f$ as a Gauss-Seidel type method for $\tilde{A}\tilde{u} = \tilde{f}$.

Lemma 3.1. *Let $\tilde{A} = \tilde{D} + \tilde{L} + \tilde{U}$ and $\tilde{B} = (\tilde{R}^{-1} + \tilde{L})^{-1}$. Then SSC for $Au = f$ with local solvers R_i is equivalent to the Gauss-Seidel type method for solving $\tilde{A}\tilde{u} = \tilde{f}$:*

$$\tilde{u}_{k+1} = \tilde{u}_k + \tilde{B}(\tilde{f} - \tilde{A}\tilde{u}_k). \tag{8}$$

Proof. By multiplying $\tilde{R}^{-1} + \tilde{L}$ to (8) and rearranging the terms, we have

$$\tilde{R}^{-1}\tilde{u}_{k+1} = \tilde{R}^{-1}\tilde{u}_k + \tilde{f} - \tilde{L}\tilde{u}_{k+1} - (\tilde{D} + \tilde{U})\tilde{u}_k.$$

Multiplying $\tilde{R}$, we obtain

$$\tilde{u}_{k+1} = \tilde{u}_k + \tilde{R}\left(\tilde{f} - \tilde{L}\tilde{u}_{k+1} - (\tilde{D} + \tilde{U})\tilde{u}_k\right),$$

and its component-wise formula, for $i = 1, \cdots, J$

$$\begin{aligned} u_{k+1}^i &= u_k^i + R_i\Big(f_i - \sum_{j=1}^{i-1}\tilde{a}_{ij}u_{k+1}^j - \sum_{j=i}^{J}\tilde{a}_{ij}u_k^j\Big) \\ &= u_k^i + R_iQ_i\Big(f - A\sum_{j=1}^{i-1}u_{k+1}^j - A\sum_{j=i}^{J}u_k^j\Big). \end{aligned}$$

Let the dynamic update

$$v^i = \Pi\,(u_{k+1}^1, \ldots, u_{k+1}^i, u_k^{i+1}, \ldots, u_k^J)^{\intercal} = \sum_{j=1}^{i} u_{k+1}^j + \sum_{j=i+1}^{J} u_k^j.$$

Noting that $v^i - v^{i-1} = u_{k+1}^i - u_k^i$, we then get, for $i = 1, \cdots, J+1$

$$v^i = v^{i-1} + R_iQ_i(f - Av^{i-1}),$$

which is exactly the correction in the subspace $\mathbb{V}_i$; see (4). $\square$

Recall that for SSC, for each subspace problem, we have the operator form $v^{i+1} = v^i + R_i(f - Av^i)$, but it is not easy to write out the iterator for the space $\mathbb{V}$. Let us define B_m to be the error operator so that

$$I - B_mA = (I - R_JQ_JA)(I - R_{J-1}Q_{J-1}A)...(I - R_0Q_0A).$$

We can then derive a formulation of B_m from the auxiliary space method. Let $\tilde{B}_m = (\tilde{R}^{-1} + \tilde{L})^{-1}$ and its symmetrization as

$$\overline{\tilde{B}}_m = \tilde{B}_m^{\intercal} + \tilde{B}_m - \tilde{B}_m^{\intercal}\tilde{A}\tilde{B}_m = \tilde{B}_m^{\intercal}(\tilde{B}_m^{-\intercal} + \tilde{B}_m^{-1} - \tilde{A})\tilde{B}_m. \tag{9}$$

Lemma 3.2. *For SSC, we have*

$$B_m = \Pi\tilde{B}_m\Pi^{\intercal} \quad \textit{and} \quad \overline{B}_m = \Pi\overline{\tilde{B}}_m\Pi^{\intercal}.$$

Proof. Let $u_k = \Pi\tilde{u}_k$. Applying Π to (8) and noting that

$$\tilde{f} = \Pi^{\intercal}f, \quad \text{and} \quad \tilde{A}\tilde{u}_k = \Pi^{\intercal}Au_k,$$

we then get

$$u_{k+1} = u_k + \Pi\tilde{B}_m\Pi^{\intercal}(f - Au_k).$$

The formulae for $\overline{B}_m$ follows from a similarly computation. $\square$

4. Identities for additive and multiplicative methods

The operator for the additive method is

$$B_a = \Pi \tilde{R} \Pi^{\intercal} = \sum_{i=1}^{J} I_i R_i I_i^{\intercal}. \tag{10}$$

Applying Theorem 2.1, we obtain the following identity for preconditioner B_a.

Theorem 4.1. *If R_i is SPD on $\mathbb{V}_i$ for $i = 1, \ldots, J$, then B_a defined by (10) is SPD on $\mathbb{V}$. Furthermore*

$$(B_a^{-1} v, v) = \inf_{\sum_{i=1}^{J} v_i = v} \sum_{i=1}^{J} (R_i^{-1} v_i, v_i) = \inf_{\sum_{i=1}^{J} v_i = v} \sum_{i=1}^{J} (T_i^{-1} v_i, v_i)_A. \tag{11}$$

To compute $\overline{\tilde{B}}_m$, we define the diagonal matrix of operators $\overline{\tilde{R}} = \operatorname{diag}(\overline{R}_1, \overline{R}_2, \cdots, \overline{R}_J)$, where, for each $R_i, i = 1, \cdots, J$, its symmetrization is

$$\overline{R}_i = R_i^{\intercal}(R_i^{-\intercal} + R_i^{-1} - A_i) R_i.$$

Substituting $\tilde{B}_m^{-1} = \tilde{R}^{-1} + \tilde{L}$, and $\tilde{A} = \tilde{D} + \tilde{L} + \tilde{U}$ into (9), we have

$$\overline{\tilde{B}}_m = (\tilde{R}^{-\intercal} + \tilde{L}^{\intercal})^{-1} (\tilde{R}^{-\intercal} + \tilde{R}^{-1} - \tilde{D}) (\tilde{R}^{-1} + \tilde{L})^{-1} \tag{12}$$

$$= (\tilde{R}^{-\intercal} + \tilde{L}^{\intercal})^{-1} \tilde{R}^{-\intercal} \overline{\tilde{R}} \tilde{R}^{-1} (\tilde{R}^{-1} + \tilde{L})^{-1}. \tag{13}$$

It is obvious that $\overline{\tilde{B}}_m$ is symmetric. To be positive definite, from (13), it suffices to assume $\overline{\tilde{R}}$, i.e. each $\overline{R}_i$, is symmetric and positive definite which is equivalent to the operator $I - \overline{R}_i A_i$ is a contraction and so is $I - R_i A_i$.

(C) $\quad \|I - R_i A_i\|_{A_i} < 1$ for each $i = 1, \cdots, J$.

Theorem 4.2. *Suppose (C) holds. Then $\overline{B}_m = \Pi \overline{\tilde{B}}_m \Pi^{\intercal}$ is SPD, and*

$$(\overline{B}_m^{-1} v, v) = \inf_{\sum_{i=1}^{J} v_i = v} \sum_{i=1}^{J} \|v_i + R_i^{\intercal} A_i P_i \sum_{j>i}^{J} v_j\|_{\overline{R}_i^{-1}}^2. \tag{14}$$

$$(\overline{B}_m^{-1} v, v) = \|v\|_A^2 + \inf_{\sum_{i=1}^{J} v_i = v} \sum_{i=1}^{J} \|R_i^{\intercal}(A_i P_i \sum_{j=i}^{J} v_j - R_i^{-1} v_i)\|_{\overline{R}_i^{-1}}^2. \tag{15}$$

In particular, for $R_i = A_i^{-1}$, we have

$$(\overline{B}_m^{-1} v, v) = \|v\|_A^2 + \inf_{\sum_{i=1}^{J} v_i = v} \sum_{i=1}^{J} \|P_i \sum_{j=i+1}^{J} v_j\|_A^2. \tag{16}$$

Proof. From (13), we have

$$\overline{\tilde{B}}_m^{-1} = (\tilde{I} + \tilde{R}^{\intercal} \tilde{U})^{\intercal} \overline{\tilde{R}}^{-1} (\tilde{I} + \tilde{R}^{\intercal} \tilde{U}).$$

Using component-wise formula of

$$(\tilde{U} \tilde{v})_i = \sum_{j=i+1}^{J} \tilde{a}_{ij} v_j = \sum_{j=i+1}^{J} A_i P_i v_j,$$

and Theorem 2.1, we get (14).

Before we prove the identity (15), we first prove the special case (16) to present the main idea. Obviously (**C**) holds for the exact local solver $R_i = A_i^{-1}$. When $R_i = A_i^{-1}$, $\tilde{B}_m = (\tilde{D}+\tilde{L})^{-1}$ and, by direct computation,

$$\overline{\tilde{B}}_m^{-1} = \tilde{A} + \tilde{L}\tilde{D}^{-1}\tilde{U}. \tag{17}$$

Therefore

$$(\overline{B}_m^{-1}v, v) = \inf_{\Pi\tilde{v}=v}(\overline{\tilde{B}}_m^{-1}\tilde{v},\tilde{v}) = (\tilde{A}\tilde{v},\tilde{v}) + \inf_{\Pi\tilde{v}=v}(\tilde{D}^{-1}\tilde{U}\tilde{v},\tilde{U}\tilde{v}).$$

For any $\tilde{v}\in\tilde{\mathbb{V}}$, denoted by $v = \Pi\tilde{v}$, we have

$$(\tilde{A}\tilde{v},\tilde{v}) = (\Pi^{\intercal}A\Pi\tilde{v},\tilde{v}) = (A\Pi\tilde{v},\Pi\tilde{v}) = \|v\|_A^2,$$

and

$$(\tilde{D}^{-1}\tilde{U}\tilde{v},\tilde{U}\tilde{v}) \quad = \quad \sum_{i=1}^J (A_i^{-1}\sum_{j=i+1}^J A_iP_iv_j, \sum_{j=i+1}^J A_iP_iv_j) = \sum_{i=1}^J \|P_i\sum_{j=i+1}^J v_j\|_A^2.$$

The identity (16) then follows.

In general let

$$\mathscr{M} = \tilde{R}^{-\intercal} + \tilde{R}^{-1} - \tilde{D} = \tilde{R}^{-\intercal}\overline{\tilde{R}}\tilde{R}^{-1},\quad \mathscr{U} = \tilde{D}+\tilde{U}-\tilde{R}^{-1},\quad \mathscr{L} = \mathscr{U}^{\intercal}.$$

then $\tilde{R}^{-1}+\tilde{L} = \mathscr{M}+\mathscr{L}$ and $\tilde{A} = \mathscr{M}+\mathscr{L}+\mathscr{U}$. We then compute, from (12), that

$$\begin{aligned}\overline{\tilde{B}}_m^{-1} &= (\tilde{R}^{-1}+\tilde{L})(\tilde{R}^{-\intercal}+\tilde{R}^{-1}-\tilde{D})^{-1}(\tilde{R}^{-\intercal}+\tilde{L}^{\intercal})\\ &= (\mathscr{M}+\mathscr{L})\mathscr{M}^{-1}(\mathscr{M}+\mathscr{U}),\\ &= \tilde{A} + \mathscr{L}\mathscr{M}^{-1}\mathscr{U}\\ &= \tilde{A} + \Big[\tilde{R}^{\intercal}(\tilde{D}+\tilde{U}-\tilde{R}^{-1})\Big]^{\intercal}\overline{\tilde{R}}^{-1}\Big[\tilde{R}^{\intercal}(\tilde{D}+\tilde{U}-\tilde{R}^{-1})\Big].\end{aligned}$$

The identity (15) then follows from the component-wise formula. □

If we use the operator $T_i = R_iA_iP_i : \mathbb{V}\to\mathbb{V}_i$, then $T_i^* = R_i^{\intercal}A_iP_i$, $\overline{T}_i := T_i + T_i^* - T_i^*T_i = \overline{R}_iA_iP_i$, and $(\overline{R}_i^{-1}u_i,u_i) = (\overline{T}_i^{-1}u_i,u_i)_A$. Here $\overline{T}_i^{-1} := (\overline{T}_i|_{\mathbb{V}_i})^{-1} : \mathbb{V}_i\to\mathbb{V}_i$ is well defined due to the assumption (**C**). The identity (15) can be written as the original formulation in [9]

$$(\overline{B}_m^{-1}v,v) = \|v\|_A^2 + \inf_{\sum_{i=1}^J v_i = v}\sum_{i=1}^J(\overline{T}_i^{-1}T_i^*w_i, T_i^*w_i)_A,$$

with $w_i = \sum_{j=i}^J v_j - T_i^{-1}v_i$. The identity (14) becomes the formula in [2]

$$(\overline{B}_m^{-1}v,v) = \inf_{\sum_{i=1}^J v_i=v}\sum_{i=1}^J(\overline{T}_i^{-1}(v_i + T_i^*\sum_{k=i+1}^J v_k), v_i + T_i^*\sum_{k=i+1}^J v_k)_A.$$

Combining Lemma 3.2 and Theorem 4.2, we obtain the X-Z identity for multiplicative methods.

Theorem 4.3 (X-Z identity)**.** *Suppose assumption (C) holds. Then*

$$\|I - B_mA\|_A^2 = \|I - \overline{B}_mA\|_A = 1 - \frac{1}{K}, \tag{18}$$

where

$$K = \sup_{\|v\|_A=1} \inf_{\sum_{i=1}^J v_i = v} \sum_{i=1}^J \|v_i + R_i^{\intercal} A_i P_i \sum_{j>i}^J v_j\|_{\overline{R}_i^{-1}}^2 .$$

Or

$$\|I - B_m A\|_A^2 = \|I - \overline{B}_m A\|_A = 1 - \frac{1}{1+c_0}, \tag{19}$$

where

$$c_0 = \sup_{\|v\|_A=1} \inf_{\sum_{i=1}^J v_i = v} \sum_{i=1}^J \|R_i^{\intercal}(A_i P_i \sum_{j=i}^J v_j - R_i^{-1} v_i)\|_{\overline{R}_i^{-1}}^2 .$$

In particular, for $R_i = A_i^{-1}$,

$$\|(I - P_J)(I - P_{J-1}) \cdots (I - P_0)\|_A^2 = 1 - \frac{1}{1+c_0}, \tag{20}$$

where

$$c_0 = \sup_{\|v\|_A=1} \inf_{\sum_{i=1}^J v_i = v} \sum_{i=1}^J \|P_i \sum_{j=i+1}^J v_j\|_A^2 .$$

Appendix: Sum and Product of Two Linear Spaces

Given two linear spaces V_1, V_2 and assume they are subspaces of a larger linear space V. We have the following operations of these two spaces

- $V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$;
- $V_1 \oplus V_2 = V_1 + V_2$ and $V_1 \cap V_2 = \{0\}$;
- $V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$;
- $V_1 \otimes V_2 = \{v_1 \otimes v_2 : v_1 \in V_1, v_2 \in V_2\}$.

The tensor product $v_1 \otimes v_2$ is a bilinear mapping on the dual space $V_1' \times V_2'$. A natural product topology can be defined for $V_1 \times V_2$ component-wise.

The relation of dimensions are

- $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2) \le \dim(V_1) + \dim(V_2)$;
- $\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$;
- $\dim(V_1 \times V_2) = \dim(V_1) + \dim(V_2)$;
- $\dim(V_1 \otimes V_2) = \dim(V_1) \times \dim(V_2)$.

We emphasize the sum $V_1 + V_2$ may not enlarge the space. For example, when $V_1 \subset V_2$ (a line on a plane), $V_1 + V_2 = V_2$.

References

[1] L. Chen. Deriving the X-Z Identity from Auxiliary Space Method. In Y. Huang, R. Kornhuber, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering XIX*, pages 309–316. Springer Berlin Heidelberg, 2010. 4

[2] D. Cho, J. Xu, and L. Zikatanov. New estimates for the rate of convergence of the method of subspace corrections. *Numer. Math. Theor. Meth. Appl.*, 1:44–56, 2008. 8

[3] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz methods. *Numer. Math.*, 70:163–180, 1995. 5

[4] S. V. Nepomnyaschikh. Decomposition and fictitious domains methods for elliptic boundary value problems. In *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations (Norfolk, VA, 1991)*, pages 62–72. SIAM, Philadelphia, PA, 1992. 4, 5

[5] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992. 1

[6] J. Xu. The auxiliary space method and optimal multigrid preconditioning techniques for unstructured meshes. *Computing*, 56:215–235, 1996. 4

[7] J. Xu. An introduction to multilevel methods. 1997. published by Oxford University Press, New York. 1

[8] J. Xu. *Multilevel Finite Element Methods*. Lecutre Notes, 2004. 1

[9] J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Amer. Math. Soc.*, 15:573–597, 2002. 4, 8