Math 673/4

Multigrid Methods: A Mostly Matrix-Based Approach

Chapter 11: The Axiomatic Method of Subspace Corrections

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

F24/S25

Chapter 11, Part 1 of 2

The Axiomatic Method of Subspace Corrections

The Goal

In this chapter, which is based on the paper Chen, Hu, and Wise (2020), we will build an axiomatic framework for building multiplicative multigrid methods for approximating the solution of nonlinear boundary value problems that arise as the Euler-Lagrange equations of strongly convex functionals.

The method relies on subspace decompositions, as we described in a previous chapter. However, as opposed to the hierarchical basis and BPX preconditioners, these methods are designed to be stand-alone solvers.

Description of the Problem and Basic Axioms

The Problem

Let $\mathcal{V}$ be a Hilbert space with inner product $(\cdot,\cdot)_{\mathcal{V}}$ and norm $\|\cdot\|_{\mathcal{V}}$. Given a functional $E : \mathcal{V} \to \mathbb{R}$, which we shall call an *energy* or *objective*, we consider the following minimization problem: Find $u \in \mathcal{V}$ such that

$$u = \operatorname*{argmin}_{v \in \mathcal{V}} E(v). \tag{1}$$

Often times this problem is framed via the equivalent Euler-Lagrange problem: Find $u \in \mathcal{V}$ such that

$$\langle \mathrm{D}E(u), w \rangle = 0, \quad \forall\, w \in \mathcal{V}. \tag{2}$$

We now make a series of assumptions that guarantee that these problems have a unique solution.

Definition

Suppose that $E : \mathcal{V} \to \mathbb{R}$, where $\mathcal{V}$ is a Hilbert space. We say that **Assumption (E0) holds** iff E is Fréchet differentiable on $\mathcal{V}$ and strongly convex with constant μ. We say that **Assumption (E1) holds** iff Assumption (E0) holds and, additionally, E is locally Lipschitz smooth on $\mathcal{V}$.

In what follows, for a bounded set $\mathcal{B} \subset \mathcal{V}$, we shall denote the (local) Lipschitz smoothness constant by $L_{\mathcal{B}}$.

From the theory in the last chapter, the assumptions above guarantee that problem (1) has a unique solution which, in addition, can be uniquely characterized as the unique solution to the Euler-Lagrange equations (2).

Gradient Descent

As we saw in the last chapter, we can use the gradient descent method to approximate the solution to (1). The convergence, however, will be too slow for practical purposes. For this reason, here we devise other methods based on subspace decompositions.

Definition (Subspace Decomposition)

Let $\mathcal{V}$ be a Hilbert space and $N \in \mathbb{N}$, $N \geq 2$. The collection $\{\mathcal{V}_i\}_{i=1}^N$ is called a **subspace decomposition of $\mathcal{V}$** iff

(i) for $i = 1, \ldots, N$, $\mathcal{V}_i$ is a norm-closed subspace of $\mathcal{V}$, and

(ii) we have

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 + \cdots + \mathcal{V}_N.$$

Definition (Stable Decomposition)

Let $\mathcal{V}$ be a Hilbert space and $\{\mathcal{V}_i\}_{i=1}^N$ be a subspace decomposition of $\mathcal{V}$. We say that the subspace decomposition is **stable**, or, equivalently, **Assumption (NS1) holds** iff there is a constant $C_A > 0$, such that, for every $v \in \mathcal{V}$, there are $v_i \in \mathcal{V}_i$, $i = 1, \dots, N$, such that

$$v = \sum_{i=1}^N v_i, \tag{NS1a}$$

and

$$\sum_{i=1}^N \|v_i\|_{\mathcal{V}}^2 \le C_A^2 \|v\|_{\mathcal{V}}^2. \tag{NS1b}$$

Definition (Strengthened Cauchy-Schwartz Inequality)

Let $\{\mathcal{V}_i\}_{i=1}^N$ be a subspace decomposition of $\mathcal{V}$ and suppose that $E : \mathcal{V} \to \mathbb{R}$ is Fréchet differentiable. We say that the energy functional E satisfies the **Strengthened Cauchy-Schwarz Inequality** with respect to the subspace decomposition, or, equivalently, **Assumption (NS2) holds** iff, given $u_0 \in \mathcal{V}$, there is a constant $C_S = C_S(u_0) > 0$, such that if

$$u_i, v_i, w_i \in \mathcal{V}_i, \quad 1 \le i \le N, \tag{NS2a}$$

with

$$v_{i-1}, v_{i-1} + u_i \in \operatorname{level}_{E(u_0)}(E), \quad 2 \le i \le N, \tag{NS2b}$$

then

$$\left| \sum_{i=1}^N \sum_{j=i+1}^N \langle \mathrm{D}E(v_{j-1} + u_j) - \mathrm{D}E(v_{j-1}), w_i \rangle \right|^2 \le C_S^2 \sum_{i=1}^N \|u_i\|_{\mathcal{V}}^2 \sum_{i=1}^N \|w_i\|_{\mathcal{V}}^2 . \tag{NS2c}$$

Remark (Quadratic Energies)

Notice that, if E is not quadratic, we cannot assume, in general, that the Strengthened Cauchy-Schwarz inequality holds without restriction to the sublevel set $\mathrm{level}_{E(u_0)}(E)$, *as indicated in (NS2).*

Remark (Terminology)

Note that the constant $C_S > 0$ in (NS2) is related to the local Lipschitz constant L. In particular, suppose that E satisfies Assumption (E1) and L is the local Lipschitz constant relative to $\mathrm{level}_{E(u_0)}(E)$*. Since we assume the restriction* (NS2b)*, it follows that*

$$|\langle \mathrm{D}E(v_{j-1}+u_j) - \mathrm{D}E(v_{j-1}), w_i\rangle| \le L\|u_j\|_{\mathcal{V}}\|w_i\|_{\mathcal{V}}.$$

Setting S *equal to the left hand side* (NS2c)*, we have*

$$\begin{aligned}\mathrm{S} &\le L^2\left(\sum_{j=1}^{N}\|u_j\|_{\mathcal{V}}\right)^2\left(\sum_{i=1}^{N}\|w_i\|_{\mathcal{V}}\right)^2\\ &\le L^2N^2\left(\sum_{j=1}^{N}\|u_j\|_{\mathcal{V}}^2\right)\left(\sum_{i=1}^{N}\|w_i\|_{\mathcal{V}}^2\right).\end{aligned}$$

which yields the naïve estimate $C_S = LN$.

Remark (Cont.)

In the case that the dimension of the ambient space, $\mathcal{V}$, is finite, we typically have

$$C_1 \dim(\mathcal{V}) \leq N \leq C_2 \dim(\mathcal{V}),$$

where $0 < C_1 < C_2$ *are constants. Thus, we could take* $C_S = C_2 L \dim(\mathcal{V})$*. Obviously, if* $\dim(\mathcal{V})$ *is large,* C_S *could be undesirably large, if estimated in this crude way.*

On the other hand, if E *is twice continuously differentiable, we may write*

$$\langle \mathrm{D}E(v_{j-1} + u_j) - \mathrm{D}E(v_{j-1}), w_i \rangle = M_j(u_j, w_i),$$

where the bilinear form is

$$M_j = \int_0^1 \mathrm{D}^2 E(v_{j-1} + t u_j)\, \mathrm{d}t \in \mathcal{L}(\mathcal{V}, \mathcal{V}').$$

If this bilinear form is, uniformly, spectrally equivalent to the $\mathcal{V}$*-inner product, a better constant* C_S*, in particular a dimension-independent one, can be obtained. This is the reason we call (NS2) the strengthened Cauchy-Schwarz inequality.*

Successive Subspace Optimization Methods

Definition (SSO)

Let E satisfy (E0) and $u \in \mathcal{V}$ be the unique solution to (1). The **Successive Subspace Optimization (SSO)** method is an algorithm to approximate u that computes a sequence $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ recursively, $u^{k+1} := \mathrm{SSO}\left(u^k\right)$, as follows. Let $u^0 \in \mathcal{V}$ be arbitrary. Then, for $k \geq 0$,

- Set $v_0^k = u^k$.
- For $i = 1$ to N:
  - Define an energy section along $\mathcal{V}_i$:

    $$J_i^k(w) := E(v_{i-1}^k + w), \quad \forall\, w \in \mathcal{V}_i.$$

  - Compute the subspace correction:

    $$e_i^k = \underset{w \in \mathcal{V}_i}{\operatorname{argmin}}\, J_i^k(w). \tag{3}$$

  - Apply the subspace correction:

    $$v_i^k = v_{i-1}^k + e_i^k. \tag{4}$$

- $u^{k+1} = v_N^k$.

Remark (Fundamental Orthogonality)

Note that, owing to the strong convexity inherited by the energy section J_i^k, the correction e_i^k, computed using (3) is uniquely defined. In fact, the correction satisfies

$$\langle \mathrm{D}E(v_i^k), w\rangle = \langle \mathrm{D}E(v_{i-1}^k + e_i^k), w\rangle = \langle \mathrm{D}J_i^k(e_i^k), w\rangle = 0, \quad \forall\, w \in \mathcal{V}_i.$$

The orthogonality relation satisfied by the corrected approximation v_i^k, specifically,

$$\langle \mathrm{D}E(v_i^k), w\rangle = 0, \quad \forall\, w \in \mathcal{V}_i, \tag{5}$$

is sometimes referred to as the fundamental orthogonality (FO) of the solver. In fact, v_i^k can be understood as the Ritz approximation of u over $\mathcal{V}_i$.

Remark (Relation to Gauss-Seidel)

We point out that, when $\mathcal{V}_i$ is one-dimensional, then the computation of the subspace correction is identical to a nonlinear Gauss-Seidel method. In fact, the SSO method can be considered as a generalization of the nonlinear Gauss-Seidel methodology.

Energy Contraction

In what follows we will show a linear reduction of the energy difference for one iteration of the SSO algorithm. In other words, there is $\rho \in (0,1)$ such that, if u solves (1) and $u^{k+1} = \mathrm{SSO}\left(u^k\right)$, then

$$E(u^{k+1}) - E(u) \leq \rho(E(u^k) - E(u)), \tag{6}$$

Ideally ρ is independent of the size of the problem, i.e., the dimension of $\mathcal{V}$. This is the reason why we merely assume that $\mathcal{V}$ is a Hilbert space, even if we have in mind discretized, read, finite-dimensional, problems.

We mention that the SSO method and its convergence properties were developed in papers by Tai and Xu in 2001 and 2003 in the more general case that $\mathcal{V}$ is a Banach space. Here we present a simplified version for Hilbert spaces.

In what follows we will utilize the following simple result.

Lemma (Monotone Convergence)

Let $\{d_k\}_{k=0}^{\infty} \subset \mathbb{R}_+ \cup \{0\}$*. Assume that there is a constant* $C > 0$ *such that, for every* $k \geq 0$,

$$d_{k+1} \leq C(d_k - d_{k+1}).$$

Then, for $k \geq 0$,

$$d_{k+1} \leq \rho d_k, \qquad \rho = \frac{C}{1+C}. \tag{7}$$

In particular, the sequence $\{d_k\}_{k=0}^{\infty}$ *converges monotonically, and at least linearly, to zero.*

Proof.

Since $C > 0$, inequality (7) follows from a simple algebraic manipulation. In addition, since $\rho \in (0,1)$, (7) implies that $\{d_k\}$ is monotonically decreasing. Thus, the monotone convergence theorem implies that there is $\gamma \geq 0$ for which, as $k \to \infty$, $d_k \to \gamma$.

Proof (Cont.)

Suppose now, for the sake of contradiction, that $\gamma > 0$. Monotone convergence shows that there is $K \in \mathbb{N}$, such that, if $k \geq K$,

$$\gamma \leq d_k < d_K \leq \gamma + \frac{(1-\rho)\gamma}{2\rho}.$$

Therefore, for $k \geq K$,

$$\begin{aligned} d_{k+1} &\leq \rho d_k \\ &\leq \rho\gamma + \frac{1}{2}(1-\rho)\gamma \\ &= \frac{1}{2}(\rho+1)\gamma \\ &< \gamma. \end{aligned}$$

This is a contradiction. The only possibility is $\gamma = 0$. $\square$

The Plan

To analyze the SSO method, we shall assume the following.

- (Lower Bound on Corrections) There is $C_L > 0$ such that, for all $k \geq 0$,

$$E(u^k) - E(u^{k+1}) \geq C_L \sum_{i=1}^{N} \|e_i^k\|_{\mathcal{V}}^2. \tag{8}$$

- (Upper Bound on Corrections) There is $C_U > 0$ such that, for $k \geq 0$,

$$E(u^{k+1}) - E(u) \leq C_U \sum_{i=1}^{N} \|e_i^k\|_{\mathcal{V}}^2. \tag{9}$$

Under these assumptions, as a corollary to the last lemma, we have that (6) holds.

Corollary (Convergence)

Let $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ be obtained by the SSO method. Assume that the lower bound (8) and upper bound (9) hold. Then the sequence of energies $\{E(u^k)\}_{k=0}^{\infty}$ satisfies (6) with

$$\rho = \frac{C_U}{C_L + C_U}.$$

Consequently, as $k \to \infty$, the sequence $\{E(u^k)\}_{k=0}^{\infty}$ converges monotonically, and at least linearly, to $E(u)$. Furthermore, $\{u^k\}_{k=0}^{\infty}$ converges at least linearly to u.

Proof.

We shall use the last lemma with

$$d_k := E(u^k) - E(u).$$

The quantity d_k is the difference between the current value of the energy and its minimum. See the figure on the next slide.

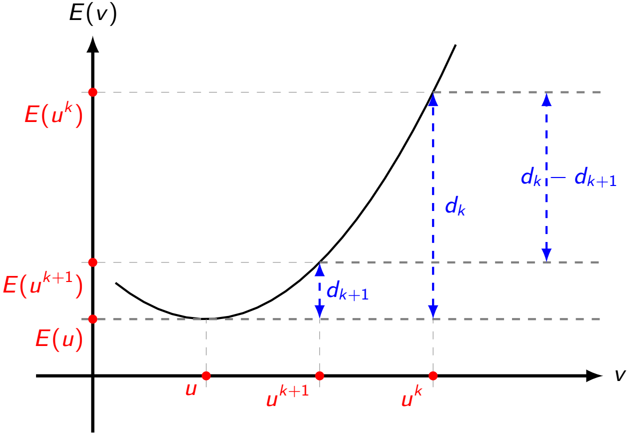

Figure: The sequence $\{d_k\}_{k=0}^{\infty}$ used in the proof of the corollary.

Proof (Cont.)

The difference

$$d_k - d_{k+1} = E(u^k) - E(u^{k+1}),$$

is the energy decrease associated to the $(k+1)^{\mathrm{st}}$ iteration. We have

$$\begin{aligned} d_{k+1} &\overset{(9)}{\leq} C_U \sum_{i=1}^{N} \|e_i^k\|_{\mathcal{V}}^2 \\ &\overset{(8)}{\leq} \frac{C_U}{C_L} \left(d_k - d_{k+1}\right), \end{aligned}$$

so that, by the lemma, we have (7) with

$$\rho = \frac{\frac{C_U}{C_L}}{1 + \frac{C_U}{C_L}} = \frac{C_U}{C_L + C_U}.$$

Thus, the linear convergence of $\{E(u^k)\}_{k=0}^{\infty}$ to $E(u)$ at the rate ρ is guaranteed by the last lemma.

Proof (Cont.)

Finally, using the lower bound from the quadratic energy trap,

$$\frac{\mu}{2}\left\|w-u\right\|_{\mathcal{V}}^2 \le E(w)-E(u) \le \frac{L}{2}\left\|w-u\right\|_{\mathcal{V}}^2, \tag{10}$$

with $w = u^k$, we have

$$\frac{\mu}{2}\left\|u^k-u\right\|_{\mathcal{V}}^2 \le E(u^k)-E(u),$$

which guarantees the linear convergence of u^k to u. $\square$

Next Steps

The next steps then are to verify the lower and upper bounds on corrections (8) and (9), respectively. Verifying the lower bound is relatively easy since E is convex. Solving the convex optimization problem in each subspace will definitely decrease the energy, and this decrease can be quantified in terms of the norms of the corrections. We make essential use of the fundamental orthogonality property.

Theorem (Lower Bound on Corrections)

Assume that the energy E satisfies (E0). Let $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ be obtained with the SSO. Then, condition (8) holds with $C_L := \frac{\mu}{2}$.

Proof.

Fix $k \in \mathbb{N}_0$. Owing to a previous lemma, we have that, for every $i = 1, \ldots, N$, the energy section J_i^k is strictly convex over $\mathcal{V}_i$ and is Fréchet differentiable, as it inherits the structure of E. It follows that

$$\langle \mathrm{D}J_i^k(e_i^k), w \rangle = 0, \quad \forall\, w \in \mathcal{V}_i.$$

But, using the definition of J_i^k, we have, for any $w \in \mathcal{V}_i$,

$$\langle \mathrm{D}J_i^k(e_i^k), w \rangle = \langle \mathrm{D}E(v_{i-1}^k + e_i^k), w \rangle = \langle \mathrm{D}E(v_i^k), w \rangle.$$

Therefore, the fundamental orthogonality, $\mathrm{D}E(v_i^k) = 0$ in $\mathcal{V}_i'$, holds. As $e_i^k = v_i^k - v_{i-1}^k \in \mathcal{V}_i$, in view of a result from the last chapter, we have

$$E(v_{i-1}^k) - E(v_i^k) = J_i^k(0) - J_i^k(e_i^k) \geq \frac{\mu}{2} \left\| e_i^k \right\|_{\mathcal{V}}^2. \tag{11}$$

Proof (Cont.)

Summing (11) over $i = 1, \ldots, N$, we see that the left hand side telescopes, and we obtain

$$E(u^k) - E(u^{k+1}) = \sum_{i=1}^{N} \left(E(v_{i-1}) - E(v_i) \right) \geq \frac{\mu}{2} \sum_{i=1}^{N} \|e_i^k\|_{\mathcal{V}}^2. \qquad \square$$

Here is the result from the last chapter of which we made careful use.

Lemma (Convexity of Sections)

Let $\mathcal{V}$ be a Hilbert space and $E : \mathcal{V} \to \mathbb{R}$ be strongly convex and locally Lipschitz smooth. Fix $\alpha \in \mathbb{R}$ and set $\mathcal{B} = \mathrm{level}_\alpha(E)$. Let $\xi \in \mathcal{B}$ be arbitrary, and $\mathcal{W} \subseteq \mathcal{V}$ be a subspace. Define the section

$$J_\xi(w) := E(\xi + w), \quad \forall\, w \in \mathcal{W}.$$

Then, $J_\xi : \mathcal{W} \to \mathbb{R}$ is differentiable, strongly convex, and locally Lipschitz smoother. Furthermore, there exists a unique element $\eta \in \mathcal{W}$ such that $\xi + \eta \in \mathcal{B}$, η is the unique global minimizer of J_ξ, and

$$\langle \mathrm{D}E(\xi + \eta), w\rangle = \langle \mathrm{D}J_\xi(\eta), w\rangle = 0, \quad \forall\, w \in \mathcal{W}.$$

As a consequence, for all $w \in \mathcal{W}$ with $w + \xi \in \mathcal{B}$,

$$\frac{\mu}{2}\left\|w - \eta\right\|_{\mathcal{V}}^2 \le J(w) - J(\eta) = E(\xi + w) - E(\xi + \eta) \le \frac{L}{2}\left\|w - \eta\right\|_{\mathcal{V}}^2.$$

The lower bound holds for any $w \in \mathcal{W}$, without restriction.

We now focus our attention on the upper bound on correction (9). The proof of this bound is more elaborate, as it relies on the assumptions about the decomposition of spaces. This is the content of the following result.

Theorem (Upper Bound on Corrections)

Assume that the energy E satisfies (E0). Given $u^0 \in \mathcal{V}$, let $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ be obtained with the SSO method. If the space decomposition satisfies (NS1) and (NS2), then we have (9) *with $C_U := \frac{C_S^2 C_A^2}{2\mu}$, where $C_S = C_S(u^0)$.*

Proof.

Fix $k \geq 0$. Owing to a prior lemma, with the choice $v = u^{k+1}$, we have

$$E(u^{k+1}) - E(u) \leq \frac{1}{2\mu} \|\mathrm{D}E(u^{k+1})\|_{\mathcal{V}'}^2.$$

Proof (Cont.)

Observe next that, in (3), we solve the minimization problem on each subspace exactly. Thus, the energy decreases. This guarantees that, if

$$\mathcal{B} = \mathrm{level}_{E(u^0)}(E),$$

then, for all $k \geq 0$ and $j = 1, \ldots, N$, we have $v_j^k \in \mathcal{B}$. Moreover, $\mathrm{D}E(v_j^k) = 0$ in $\mathcal{V}_j'$. Let now $w \in \mathcal{V}$ and, by invoking (NS1), we may choose a stable decomposition $w = \sum_{i=1}^N w_i$.

Proof (Cont.)

Then,

$$
\begin{aligned}
\langle \mathrm{D}E(u^{k+1}), w\rangle &= \sum_{i=1}^{N} \langle \mathrm{D}E(u^{k+1}), w_i\rangle \\
&\overset{(5)}{=} \sum_{i=1}^{N} \langle \mathrm{D}E(u^{k+1}) - \mathrm{D}E(v_i^k), w_i\rangle \\
&= \sum_{i=1}^{N} \sum_{j=i+1}^{N} \langle \mathrm{D}E(v_j^k) - \mathrm{D}E(v_{j-1}^k), w_i\rangle \\
&\overset{(4)}{=} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \langle \mathrm{D}E(v_{j-1}^k + e_j^k) - \mathrm{D}E(v_{j-1}^k), w_i\rangle \\
&\overset{(NS2)}{=} C_S \left(\sum_{i=1}^{N} \| e_i^k \|_{\mathcal{V}}^2 \right)^{1/2} \left(\sum_{i=1}^{N} \| w_i \|_{\mathcal{V}}^2 \right)^{1/2} \\
&\overset{(NS1)}{\leq} C_S C_A \left(\sum_{i=1}^{N} \| e_i \|_{\mathcal{V}}^2 \right)^{1/2} \| w \|_{\mathcal{V}}.
\end{aligned}
$$

Proof (Cont.)

We may then gather the two previous estimates and use the definition of the norm of $\mathcal{V}'$ to obtain

$$
\begin{aligned}
E(u^{k+1}) - E(u) &\leq \frac{1}{2\mu}\|\mathrm{D}E(u^{k+1})\|_{\mathcal{V}'}^2 \\
&= \frac{1}{2\mu}\left(\sup_{w\in\mathcal{V}\setminus\{0\}} \frac{\langle \mathrm{D}E(u^{k+1}), w\rangle}{\|w\|_{\mathcal{V}}}\right)^2 \\
&\leq \frac{C_S^2 C_A^2}{2\mu}\sum_{i=1}^N \|e_i\|_{\mathcal{V}}^2,
\end{aligned}
$$

which finishes the proof. □

Corollary (Convergence of SSO)

Assume that the energy E satisfies (E0). Given $u^0 \in \mathcal{V}$, let $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ be obtained with the SSO method. If the space decomposition satisfies (NS1) and (NS2), then the SSO method converges linearly. In other words, set

$$\rho = \frac{C_S^2 C_A^2}{C_S^2 C_A^2 + \mu^2}.$$

Then, for every $k \geq 0$, we have

$$E(u^{k+1}) - E(u) \leq \rho(E(u^k) - E(u)).$$

Proof.

We have

$$C_L = \frac{\mu}{2} \quad \text{and} \quad C_U = \frac{C_S^2 C_A^2}{2\mu}.$$

Thus,

$$\rho = \frac{C_U}{C_L + C_U} = \frac{\frac{C_S^2 C_A^2}{2\mu}}{\frac{\mu}{2} + \frac{C_S^2 C_A^2}{2\mu}} = \frac{C_S^2 C_A^2}{C_S^2 C_A^2 + \mu^2}. \qquad \square$$

FAS Method Overview

In this section, we will take a very brief detour to describe a standard two-level method for solving nonlinear elliptic equations known as the Full Approximation Storage (FAS) method (Brandt, 1977). To put things into the context of the current chapter, suppose that $\mathcal{V} = V_h$, where the latter is a finite element space. We wish to solve the following finite element problem: Find $u_h \in V_h$ such that

$$\langle \mathrm{D}E(u_h), w_h \rangle = 0, \quad \forall\, w_h \in V_h, \tag{12}$$

where E is the appropriate energy for the underlying partial differential equation. We will give an example of this in the next chapter.

Two-Grid Revisted

First, let us assume that the derivative of the energy can be represented as a linear operator of the form

$$\mathrm{D}E(u_h) = \mathcal{L}_h(u_h) - f_h.$$

In other words, the Euler equation that we need to solve can be expressed in operator form as follows: Find $u_h \in V_h$, such that

$$\mathcal{L}_h(u_h) = f_h.$$

Next, let us suppose that $V_H \subset V_h$ is a finite element subspace of V_h based on a coarsening of the grid. In the classical two-grid algorithm, we wish to use corrections from the coarse space $V_H \subset V_h$.

The linear two-grid method,

$$u_h^{k+1} = \mathrm{TG}(u_h^k),$$

is summarized by the following pseudo-code:

1. $v_0 = u_h^k$.
2. Smooth (e.g, using linear Gauss-Seidel) on the fine grid: $v_1 = S(v_0)$.
3. Compute the coarse-grid residual:

$$r_H := R_h(f_h - \mathcal{L}_h(v_1)) \in V_H, \tag{13}$$

   where the R_h is a restriction operator

$$R_h : V_h \searrow V_H.$$

4. Solve the coarse-grid correction problem: Find $s_H \in V_H$, such that

$$\mathcal{L}_H(s_H) = r_H \longrightarrow s_H \in V_H, \tag{14}$$

   where $\mathcal{L}_H$ is a version of the linear operator when realized on the coarse finite element space V_H. Typically, $\mathcal{L}_H$ will be determined via some Galerkin condition.

5. Apply the coarse-grid correction:

$$v_2 := v_1 + P_H s_H, \tag{15}$$

where P_H is the prolongation operator

$$P_H : V_H \nearrow V_h,$$

which, as usual, is related to the restriction operator through transposition.

6. Smooth (e.g., using linear Gauss-Seidel) on the fine grid: $v_3 = S(v_2)$.

7. $u_h^{k+1} = v_3$.

The Error-Residual Equation

This design of the linear two-grid algorithm is predicated upon on few facts. Paramount among these is the fact that the error equation is equivalent to the original equation. For example, if $w_h \in V_h$ is an approximation to the true solution u_h. Then,

$$u_h = w_h + e_h,$$

where $e_h \in V_h$ is the error function, which satisfies

$$\mathcal{L}_h(e_h) = r_h := f - \mathcal{L}_h(w_h).$$

Here $r_h \in V_h$ is the finite element residual. We can think about $P_H s_H$ as an approximation for the error function, in particular.

The Nonlinear Setting

Now, suppose that the derivative of the energy can be represented as a nonlinear operator of the form

$$\mathrm{D}E(u_h) = \mathcal{N}_h(u_h) - f_h.$$

The Euler equation that we need to solve can be expressed in operator form as follows: Find $u_h \in V_h$, such that

$$\mathcal{N}_h(u_h) = f_h.$$

Since the equation is nonlinear, we can no longer use the linear two-grid process. This is due, essentially, to the fact that the error-residual equation is no longer equivalent to the original equation.

The Full Approximation Storage (FAS) method, originating in (Brandt, 1977), was designed for nonlinear equations. Specifically, it avoids a dependence on the equivalence with the error-residual equation, and like the two-grid process, it is an iterative scheme.

The nonlinear FAS method

$$u_h^{k+1} = \mathrm{FAS}(u_h^k),$$

is summarized by the following pseudo-code:

1. $v_0 = u_h^k$.
2. Smooth (e.g., using nonlinear Gauss-Seidel) on the fine grid: $v_1 = S(v_0)$.
3. Compute the coarse-grid τ-*perturbation*:

$$\tau_H := \mathcal{N}_H(Q_h v_1) + R_h(f_h - \mathcal{N}_h(v_1)) \in V_H, \tag{16}$$

   where R_h is the restriction operator mentioned earlier, and Q_h is a projection operator,

$$Q_h : V_h \searrow V_H.$$

   Unlike R_h it may not be related to the prolongation operator directly through transposition.
4. Solve the coarse-grid correction problem: Find $\eta_H \in V_H$, such that

$$\mathcal{N}_H(\eta_H) = \tau_H \longrightarrow s_H := \eta_H - Q_h v_1 \in V_H. \tag{17}$$

5. Apply the coarse-grid correction:

$$v_2 := v_1 + P_H s_H. \tag{18}$$

6. Smooth (e.g., using nonlinear Gauss-Seidel) on the fine grid: $v_3 = S(v_2)$.
7. $u_h^{k+1} = v_3$.

Problems with FAS

The FAS method is analyzed in (Hackbusch, 1985) and elsewhere. However, the analysis is not very general, and the initial guess u_h^0 needs to be very close to u_h for the convergence to be proven. In the next section, we want to generalize this method so that it always converges, regardless of how u_h^0 is chosen, as long as $E : \mathcal{V} \to \mathbb{R}$ strongly convex and the subspace decomposition is sufficiently nice.