CHAPTER 5

Finite Element Multigrid Methods

The multigrid method provides an optimal complexity algorithm for solving discrete elliptic boundary value problems. The error bounds of the approximate solution obtained from the full multigrid algorithm are comparable to the theoretical error bounds of the finite element solution, while the amount of computational work involved is proportional only to the number of unknowns in the discretized equations.

The multigrid method has two main features: smoothing on the current grid and error correction on the coarse grid. The smoothing step has the effect of damping the oscillatory part of the error. The smooth part of the error can then be corrected on the coarse grid.

5.1. The model problem

Let \( \Omega \subset \mathbb{R}^d \) \((d = 1, 2, 3)\) be a convex polyhedral domain and
\[
a(u, v) = \int_\Omega (\alpha \nabla u \cdot \nabla v + \beta uv) \, dx
\]
(5.1)

where \( \alpha \) and \( \beta \) are smooth functions such that for some \( \alpha_0, \alpha_1, \beta_1 \in \mathbb{R}^+ \) we have \( \alpha_0 \leq \alpha(x) \leq \alpha_1 \) and \( 0 \leq \beta(x) \leq \beta_1 \) for all \( x \in \Omega \). We consider the Dirichlet problem: Find \( u \in V = H^1_0(\Omega) \) such that
\[
a(u, v) = (f, v) \quad \forall \, v \in V,
\]
(5.2)

where \( f \in L^2(\Omega) \) and \((\cdot, \cdot)\) denotes the \( L^2 \) inner product.

Let \( \mathcal{M}_k \) be a sequence of meshes of \( \Omega \) obtained successively by standard uniform refinements. Let \( V_k \) be the \( H^1 \)-conforming linear finite element space over \( \mathcal{M}_k \) whose functions vanish on \( \partial \Omega \). The discrete problem on \( V_k \) is then to find \( u_k \in V_k \) such that
\[
a(u_k, v_k) = (f, v_k) \quad \forall \, v_k \in V_k.
\]
(5.3)

We introduce the \( L^2 \) and \( H^1 \) projection operators
\[
(Q_k \varphi, v_k) = (\varphi, v_k), \quad a(P_k \psi, v_k) = a(\psi, v_k) \quad \forall \, v_k \in V_k,
\]

51
where $\varphi \in L^2(\Omega)$ and $\psi \in H^1_0(\Omega)$. Then by using the Aubin-Nitsche trick (cf. Section 3.3) we have
\[
\|w - P_k w\|_{L^2(\Omega)} \leq C h_k \|w\|_A \quad \forall w \in H^1_0(\Omega),
\]
where $h_k = \max_{K \in \mathcal{M}_k} h_K$ and $\|\cdot\|_A = a(\cdot, \cdot)^{1/2}$. From $v - P_{k-1} v = (I - P_{k-1})(I - P_{k-1}) v$, we then have the following approximation property
\[
\|(I - P_{k-1}) v\|_{L^2(\Omega)} \leq C h_k \|(I - P_{k-1}) v\|_A \quad \forall v \in V_k.
\] (5.4)

5.2. Iterative methods

Let $A_k : V_k \to V_k$ be defined by
\[
(a(w_k, v_k)) = A_k w_k = f_k := Q_k f.
\] (5.5)

Let $\{\phi^i_k : i = 1, \cdots, n_k\}$ denote the nodal basis of $V_k$. Given any $v_k = \sum_{i=1}^{n_k} v_{k,i} \phi^i_k \in V_k$, define $\tilde{v}_k, \tilde{u}_k \in \mathbb{R}^{n_k}$ as follows
\[
(\tilde{v}_k)_i = v_{k,i}, \quad (\tilde{u}_k)_i = (v_k, \phi^i_k), \quad i = 1, \cdots, n_k.
\] (5.6)

Let $\tilde{A}_k = [a(\phi^i_k, \phi^j_k)]_{i,j=1}^{n_k}$ be the stiffness matrix. We have the following matrix representation of (5.5):
\[
\tilde{A}_k \tilde{u}_k = \tilde{f}_k,
\] (5.7)

We want to consider the following linear iterative method for (5.7): Given $\tilde{u}^{(0)} \in \mathbb{R}^{n_k}$
\[
\tilde{u}^{(n+1)} = \tilde{u}^{(n)} + \tilde{R}_k (\tilde{f}_k - \tilde{A}_k \tilde{u}^{(n)}), \quad n = 0, 1, 2, \cdots.
\] (5.8)

$\tilde{R}_k$ is called the iterator of $\tilde{A}_k$. Note that (5.8) converges if the spectral radius $\rho(I - \tilde{R}_k \tilde{A}_k) < 1$. If we define a linear operator $R_k : V_k \mapsto V_k$ as
\[
R_k g = \sum_{i,j=1}^{n_k} (\tilde{R}_k)_{ij} (g, \phi^i_k) \phi^j_k,
\] (5.9)

then $\tilde{R}_k g = \tilde{R}_k g$, so that the algorithm (5.8) for the matrix equation (5.7) is equivalent to the following linear iterative algorithm for the operator equation (5.5): Given $u^{(0)} \in V_k$
\[
u^{(n+1)} = u^{(n)} + R_k (f_k - A_k u^{(n)}), \quad n = 0, 1, 2, \cdots.
\]
Here we have used the fact that \( \tilde{A}_k u^{(n)} = \tilde{A}_k \tilde{u}^{(n)} \). It is clear that the error propagation operator is \( I - R_k A_k \).

Noting that \( \tilde{A}_k \) is symmetric and positive definite, we write \( \tilde{A}_k = \tilde{D} - \tilde{L} - \tilde{L}^T \) with \( \tilde{D} \) and \( -\tilde{L} \) being the diagonal and the lower triangular part of \( \tilde{A}_k \) respectively. We recall the following choices of \( \tilde{R}_k \) that result in various different iterative methods:

\[
\tilde{R}_k = \begin{cases} 
\frac{\omega}{\rho(A_k)} \tilde{I} & \text{Richardson;} \\
\omega D^{-1} & \text{Damped Jacobi;} \\
(\tilde{D} - \tilde{L})^{-1} & \text{Gauss-Seidel;} \\
(\tilde{D} - \tilde{L})^{-T} \tilde{D}(\tilde{D} - \tilde{L})^{-1} & \text{Symmetrized Gauss-Seidel.}
\end{cases} \tag{5.10}
\]

**Lemma 5.1.** We have

(i) The Richardson method converges if and only if \( 0 < \omega < 2 \);

(ii) The Damped Jacobi method converges if and only if \( 0 < \omega < \frac{2}{\rho(D^{-1}A_k)} \);


**Lemma 5.2.** The damped Jacobi iterative method for solving (5.7) is equivalent to the following iterative scheme in the space \( V_k \):

\[ u_k^{(n+1)} = u_k^{(n)} + R_k (f_k - A_k u_k^{(n)}) , \quad R_k = \omega \sum_{i=1}^{n_k} P^i_k A_k^{-1}, \]

where \( P^i_k \) is the projection operator to the subspace spanned by \( \{\phi^i_k\} \):

\[ a(P^i_k w_k, \phi^i_k) = a(w_k, \phi^i_k) \quad \forall w_k \in V_k. \tag{5.11} \]

**Proof.** From (5.11) we know that

\[ P^i_k w_k = \frac{a(w_k, \phi^i_k)}{a(\phi^i_k, \phi^i_k)} \phi^i_k , \quad i = 1, 2, \ldots, n_k. \]

Recall that the iterator of the damped Jacobi iterative method is \( \tilde{R}_k = \omega \tilde{D}^{-1} = \text{diag}(\omega/a(\phi^1_k, \phi^1_k), \ldots, \omega/a(\phi^{n_k}_k, \phi^{n_k}_k)) \). It follows from (5.9) that

\[ R_k g = \omega \sum_{i=1}^{n_k} \frac{(g, \phi^i_k)}{a(\phi^i_k, \phi^i_k)} \phi^i_k = \omega \sum_{i=1}^{n_k} \frac{A_k^{-1} g, \phi^i_k}{a(\phi^i_k, \phi^i_k)} \phi^i_k = \omega \sum_{i=1}^{n_k} P^i_k A_k^{-1} g \quad \forall g \in V_k, \]

which completes the proof. \( \square \)
Lemma 5.3. The standard Gauss-Seidel iterative method for solving (5.7) is equivalent to the following iterative scheme in the space $V_k$

$$u_k^{(n+1)} = u_k^{(n)} + R_k(f_k - A_k u_k^{(n)}), \quad R_k = (I - E_k)A_k^{-1},$$

where $E_k = (I - P_k^{n_k}) \cdots (I - P_k^1)$.

Lemma 5.4. The symmetrized Gauss-Seidel iterative method for solving (5.7) is equivalent to the following iterative scheme in the space $V_k$

$$u_k^{(n+1)} = u_k^{(n)} + R_k(f_k - A_k u_k^{(n)}), \quad R_k = (I - E_k^* E_k)A_k^{-1},$$

where $E_1 = (I - P_k^{n_k}) \cdots (I - P_k^1)$ and $E_k^* = (I - P_k^{n_k}) \cdots (I - P_k^1)$ is the conjugate operator of $E_k$ with respect to $a(\cdot, \cdot)$.

The proofs of Lemma 5.3 and Lemma 5.4 are left as Exercise 5.1.

It is well-known that the classical iterative methods listed in (5.10) are inefficient for solving (5.7) when $n_k$ is large. But they have an important “smoothing property” that we discuss now. For example, Richardson iteration for (5.7) reads as

$$\tilde{u}^{(n+1)} = \tilde{u}^{(n)} + \frac{\omega}{\rho(A_k)} (f_k - A_k \tilde{u}^{(n)}), \quad n = 0, 1, 2, \ldots$$

Let $\tilde{A}_k \tilde{\phi}_i = \mu_i \tilde{\phi}_i$ with $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_{n_k}$, $(\tilde{\phi}_i, \tilde{\phi}_j) = \delta_{ij}$ and $\tilde{u}_k - \tilde{u}^0 = \sum_{i=1}^{n_k} \alpha_i \tilde{\phi}_i$, then

$$\tilde{u}_k - \tilde{u}^{(n)} = \sum_i \alpha_i (1 - \omega \mu_i / \mu_{n_k})^n \tilde{\phi}_i.$$

For a fixed $\omega \in (0, 2)$, it is clear that $(1 - \omega \mu_i / \mu_{n_k})^n$ converges to zero very fast as $n \to \infty$ if $\mu_i$ is close to $\mu_{n_k}$. This means that the high frequency modes in the error get damped very quickly.

Let us illustrate the smoothing property of the Gauss-Seidel method by a simple numerical example. Consider the Poisson equation $-\Delta u = 1$ with homogeneous Dirichlet condition on the unit square which is discretized by the uniform triangulation. Figure 1 shows that high frequency errors are well annihilated by Gauss-Seidel iterations.

For the above model problem, Brandt applied the “local mode analysis” to show that: The damped Jacobi method achieve its optimal smoothing property when $\omega = 4/5$; the Gauss-Seidel method is a better smoother than the damped Jacobi method; the Gauss-Seidel method with red-black ordering is a better smoother than the one with lexicographic ordering. We also
5.3. The multigrid V-cycle algorithm

The basic idea in a multigrid strategy is that smoothing on the current grid and error correction on a coarser grid. Let $R_k : V_k \to V_k$ be a linear smoother and $R^t_k$ be the adjoint of $R_k$ with respect to $(\cdot, \cdot)$. The multigrid V-cycle algorithm for solving (5.5) can be written as

$$u_k^{(n+1)} = u_k^{(n)} + B_k(f_k - A_k u_k^{(n)}), \quad n = 0, 1, 2, \cdots, \quad (5.12)$$

where the iterator $B_k$ is defined by the following algorithm.

**Algorithm 5.1.** (V-cycle iterator). For $k = 1$, define $B_1 = A_1^{-1}$. Assume that $B_{k-1} : V_{k-1} \mapsto V_{k-1}$ is defined. For $g \in V_k$, define the iterator $B_k : V_k \mapsto V_k$ through the following steps.

1. **Pre-smoothing:** For $y_0 = 0 \in V_k$ and $j = 1, \cdots, m$,
   $$y_j = y_{j-1} + R_k(g - A_k y_{j-1}).$$
2. **Coarse grid correction:** $y_{m+1} = y_m + B_{k-1} Q_{k-1}(g - A_k y_m), \quad (5.13)$
3. **Post-smoothing:** For $j = m + 2, \cdots, 2m + 1$,
   $$y_j = y_{j-1} + R^t_k(g - A_k y_{j-1}).$$

Define $B_k g = y_{2m+1}$. 

Figure 1. Error after 0, 3, 9 and 200 Gauss-Seidel iterations, respectively, with 2113 unknowns.
In the following, we assume that $R_k$ is symmetric with respect to $(\cdot, \cdot)$ and positive semi-definite. Denote by $y = A_k^{-1} g$, then we have
\[ y_{2m+1} - y = (I - R_k A_k)^m (I - B_{k-1} Q_{k-1} A_k) (I - R_k A_k)^m (y_0 - y). \]
Thus \[ I - B_k A_k = (I - R_k A_k)^m (I - B_{k-1} Q_{k-1} A_k) (I - R_k A_k)^m \]
On the other hand, for any $v_k \in V_k, w_{k-1} \in V_{k-1}$, we have
\[ (Q_{k-1} A_k v_k, w_{k-1}) = (A_k v_k, w_{k-1}) = a(v_k, w_{k-1}) \]
\[ = a(P_{k-1} v_k, w_{k-1}) = (A_{k-1} P_{k-1} v_k, w_{k-1}) \]
that is, $Q_{k-1} A_k = A_{k-1} P_{k-1}$. Therefore we have the following two-level recurrence relation.

**Lemma 5.5.** Let $K_k = I - R_k A_k$. Then
\[ I - B_k A_k = K_m^m ((I - P_{k-1}) + (I - B_{k-1} A_{k-1}) P_{k-1}) K_m^m \] on $V_k$.

The following lemma is left as an Exercise 5.2.

**Lemma 5.6.** We have
\[ a(K_k v, w) = a(v, K_k w) \quad \text{and} \quad (B_k v, w) = (v, B_k w) \quad \forall \ v, w \in V_k. \]

The following abstract estimate plays an important role in the analysis of multigrid method.

**Theorem 5.1.** Assume that $R_k : V_k \to V_k$ is symmetric with respect to $(\cdot, \cdot)$, positive semi-definite, and satisfies
\[ a((I - R_k A_k) v, v) \geq 0 \quad \forall \ v \in V_k. \quad (5.13) \]
Moreover
\[ (R_k^{-1} v, v) \leq \alpha a(v, v) \quad \forall \ v \in (I - P_{k-1}) V_k. \quad (5.14) \]
Then we have
\[ 0 \leq a((I - B_k A_k) v, v) \leq \delta a(v, v) \quad \forall \ v \in V_k, \quad (5.15) \]
where $\delta = \alpha / (\alpha + 2m)$.

**Proof.** We prove by induction, (5.15) is trivial when $k = 1$ since $B_1 = A_1^{-1}$. Let us now assume (5.15) is true for $k - 1$:
\[ 0 \leq a((I - B_{k-1} A_{k-1}) v, v) \leq \delta a(v, v) \quad \forall \ v \in V_{k-1}. \quad (5.16) \]
Thus the eigenvalues of $K$ are all nonnegative.

This completes the proof of the theorem.

For the upper bound, we have

$$a((I - B_k A_k)v,v) \leq a((I - P_{k-1})K^m_k v, K^m_k v) + \delta a(P_{k-1}K^m_k v, P_{k-1}K^m_k v)$$

Now

$$a((I - P_{k-1})K^m_k v, K^m_k v) = ((I - P_{k-1})K^m_k v, A_k K^m_k v)$$

$$= (R_k^{-1}(I - P_{k-1})K^m_k v, R_k A_k K^m_k v)$$

$$\leq (R_k^{-1}(I - P_{k-1})K^m_k v, (I - P_{k-1})K^m_k v)^{1/2}(R_k A_k K^m_k v, A_k K^m_k v)^{1/2}$$

$$\leq \sqrt{\alpha a((I - P_{k-1})K^m_k v, K^m_k v)^{1/2} a((I - K_k)K^m_k v, K^m_k v)^{1/2}}.$$

Thus

$$a((I - P_{k-1})K^m_k v, K^m_k v) \leq \alpha a((I - K_k)K^m_k v, K^m_k v).$$

Since $R_k : V_k \to V_k$ is symmetric and semi-definite, by Lemma 5.6 we know that $K_k$ is symmetric with respect to $a(\cdot, \cdot)$ and $0 \leq a(K_k v, v) \leq a(v,v)$. Thus the eigenvalues of $K_k$ belong to $[0, 1]$. Hence

$$a((I - K_k)K^m_k v, v) \leq a((I - K_k)K^2_k v, v) \quad \forall \ 0 \leq i \leq 2m,$$

and consequently

$$a((I - K_k)K^2_m v, v) \leq \frac{1}{2m} \sum_{i=0}^{2m-1} a((I - K_k)K^i_k v, v) = \frac{1}{2m} a((I - K^2_k) v, v).$$

This yields

$$a((I - B_k A_k)v,v) \leq (1 - \delta) \frac{\alpha}{2m} a(v,v) + \left( \delta - \frac{\alpha}{2m} (1 - \delta) \right) a(K^m_k v, K^m_k v)$$

$$= \frac{\alpha}{\alpha + 2m} a(v,v).$$

This completes the proof of the theorem.
Now we define the smoothers which satisfy the assumptions in Theorem 5.1. Let

\[ V_k = \sum_{i=1}^{K} V_i^k \]

and denote by \( P_i^k : V_k \to V_i^k \) the projection

\[ a(P_i^k v, w) = a(v, w) \quad \forall \ w \in V_i^k, \ \forall \ v \in V_k. \]

We introduce the following additive and multiplicative Schwarz operator

\[ R_a^k = \sum_{i=1}^{K} P_i^k A_k^{-1} \]

and

\[ R_m^k = (I - E_k E_k^*) A_k^{-1}, \]

where \( E_k = (I - P_k^K) \cdots (I - P_k^1) \) and \( E_k^* = (I - P_k^1) \cdots (I - P_k^K) \) is the conjugate operator of \( E_k \) with respect to \( a(\cdot, \cdot) \).

If \( K = n_k \) and \( V_i^k = \text{span}\{ \phi_i^k \} \), then, from Lemma 5.2 and Lemma 5.4, \( R_a^k \) and \( R_m^k \) are the iterators of the Jacobi method and the symmetrized Gauss-Seidel method, respectively.

**Theorem 5.2.** Assume there exist constants \( \beta, \gamma > 0 \) such that

(i) \[ \sum_{i=1}^{K} \sum_{j=1}^{K} a(v_i^j, w_i^j) \leq \beta \left( \sum_{i=1}^{K} a(v_i^i, v_i^i) \right)^{1/2} \left( \sum_{j=1}^{K} a(w_i^j, w_i^j) \right)^{1/2}, \]

\[ \forall v_i^j, w_i^j \in V_i^k; \]

(ii) \[ \inf_{v_k^1 \in V_k^1} \sum_{i=1}^{K} a(v_i^1, v_i^1) \leq \gamma a(v, v) \quad \forall v \in (I - P_{k-1})V_k. \]

Then

1°) For \( \omega \leq 1/\beta, R_k = \omega R_a^k \) satisfies the assumptions in Theorem 5.1 with \( \alpha = \gamma/\omega. \)

2°) \( R_k = R_m^k \) satisfies the assumptions in Theorem 5.1 with \( \alpha = \beta^2 \gamma. \)
5.3. THE MULTIGRID V-CYCLE ALGORITHM

Proof. 1°) For any \( v \in V_k \),

\[
a(R_k^a A_k v, v) = a \left( \sum_{i=1}^{K} P_i^k v, v \right) \leq a(v, v)^{1/2} a \left( \sum_{i=1}^{K} P_i^k v, \sum_{i=1}^{K} P_i^k v \right)^{1/2} \\
\leq \beta^{1/2} a(v, v)^{1/2} \left( \sum_{i=1}^{K} a(p_i^k v, v) \right)^{1/2} \\
= \beta^{1/2} a(v, v)^{1/2} a(R_k^a A_k v, v)^{1/2}.
\]

Thus

\[
a((I - \omega R_k^a A_k) v, v) \geq 0 \quad \text{if} \quad \omega \leq 1/\beta.
\]

Next we prove (5.14) by showing that

\[
((R_k^a)^{-1} v, v) = \inf_{\sum_{i=1}^{K} v_i^k = v} \sum_{i=1}^{K} a(v_i^k, v_i^k) \quad \text{for any} \quad v \in (I - P_{k-1})V_k. \quad (5.17)
\]

Denote by \( \Theta = R_k^a \), and \( v = \sum_{i=1}^{K} v_i^k \cdot v_i^k \in V_k^i \). Then

\[
(\Theta^{-1} v, v) = \sum_{i=1}^{K} (\Theta^{-1} v, v_i^k) = \sum_{i=1}^{K} a(A_k^{-1} \Theta^{-1} v, v_i^k) = \sum_{i=1}^{K} a(P_i^k A_k^{-1} \Theta^{-1} v, v_i^k) \\
\leq \left( \sum_{i=1}^{K} a(P_i^k A_k^{-1} \Theta^{-1} v, P_i^k A_k^{-1} \Theta^{-1} v) \right)^{1/2} \left( \sum_{i=1}^{K} a(v_i^k, v_i^k) \right)^{1/2} \\
= \left( \sum_{i=1}^{K} (P_i^k A_k^{-1} \Theta^{-1} v, \Theta^{-1} v) \right)^{1/2} \left( \sum_{i=1}^{K} a(v_i^k, v_i^k) \right)^{1/2}, \quad (5.18)
\]

that is

\[
(\Theta^{-1} v, v) \leq (v, \Theta^{-1} v)^{1/2} \left( \sum_{i=1}^{K} a(v_i^k, v_i^k) \right)^{1/2}.
\]

Thus

\[
(\Theta^{-1} v, v) \leq \sum_{i=1}^{K} a(v_i^k, v_i^k) \quad \forall v = \sum_{i=1}^{K} v_i^k.
\]

To show the equality in (5.17) we only need to take \( v_i^k = P_i^k A_k^{-1} \Theta^{-1} v \). This proves the assertion for \( R_k^a \).

2°) Since \( R_k^a = (I - E_k E_k) A_k^{-1} \), we have

\[
a((I - R_k^a A_k) v, v) = a(E_k v, E_k v) \geq 0.
\]
Note that (5.18) holds for any invertible operator on $V_k$. By letting $\Theta = R_k^m$ in (5.18) we have
\[
((R_k^m)^{-1}v, v) \leq (R_k^m(R_k^m)^{-1}v, (R_k^m)^{-1}v)^{1/2} \left( \sum_{i=1}^{K} a(v_k^i, v_k^i) \right)^{1/2}.
\]
It follows from (ii) that
\[
((R_k^m)^{-1}v, v) \leq \gamma^{1/2}(R_k^m(R_k^m)^{-1}v, (R_k^m)^{-1}v)^{1/2} a(v, v)^{1/2}.
\] (5.19)

Now we show
\[
(R_k^m v, v) \leq \beta^2(R_k^m v, v), \quad \forall v \in V_k.
\] (5.20)

Denote by $y = A_k^{-1}v$, then
\[
(R_k^m v, v) = ((I - E^*E)A_k^{-1}v, v) = a((I - E^*E)y, y) = a(y, y) - a(E_k y, E_k y)
\]
Let $E_k^0 = I$ and $E_k^i = (I - P_k^i)\cdots(I - P_k^1), \ i = 1, \ldots, K$. Then
\[
E_k^i = (I - P_k^i)E_k^{i-1} \quad \text{and} \quad E_k^K = E_k.
\]
Therefore
\[
a(E_k^i y, E_k^i y) = a((I - P_k^i)E_k^{i-1}y, E_k^{i-1}y)
\]
\[
= a(E_k^{i-1}y, E_k^{i-1}y) - a(P_k^iE_k^{i-1}y, E_k^{i-1}y),
\]
which yields
\[
a(E_k y, E_k y) = a(y, y) - \sum_{i=1}^{K} a(P_k^iE_k^{i-1}y, E_k^{i-1}y).
\]
Consequently,
\[
(R_k^m v, v) = \sum_{i=1}^{K} a(P_k^iE_k^{i-1}y, E_k^{i-1}y).
\]

On the other hand,
\[
(R_k^a v, v) = \sum_{i=1}^{K} (P_k^iA_k^{-1}v, v) = \sum_{i=1}^{K} a(P_k^i y, y) = \sum_{i=1}^{K} a(P_k^i y, P_k^i y).
\]
We deduce from $E_k^j = E_k^{j-1} - P_k^j E_k^{j-1}$ that
\[
E_k^j y = y - \sum_{j=1}^{i} P_k^j E_k^{j-1} y
\]
and thus
\[
P_k^i y = P_k^i E_k^i y + P_k^i \sum_{j=1}^{i} P_k^j E_k^{j-1} y = P_k^i \sum_{j=1}^{i} P_k^j E_k^{j-1} y.
\]
Now

\[
(R_k^a v, v) = \sum_{i=1}^{K} a(P_k^i y, \sum_{j=1}^{i} P_j^i E_{k}^{j-1} y) = \sum_{i=1}^{K} \sum_{j=1}^{i} a(P_k^i y, P_j^i E_{k}^{j-1} y)
\]

\[
\leq \beta \left( \sum_{i=1}^{K} a(P_k^i y, P_k^i y) \right)^{1/2} \left( \sum_{j=1}^{K} a(P_k^i E_{k}^{j-1} y, P_k^i E_{k}^{j-1} y) \right)^{1/2}
\]

\[
= \beta (R_k^a v, v)^{1/2} (R_k^m v, v)^{1/2}.
\]

This proves (5.20). Finally, we deduce from (5.19) that

\[
((P_k^m)^{-1} v, v) \leq \beta \gamma^{1/2} ((P_k^m)^{-1} v, v)^{1/2} a(v, v).
\]

This completes the proof of the theorem.

5.4. The finite element multigrid V-cycle algorithm

Now we apply the abstract result in last section to solve the discrete elliptic problem (5.3). Let \( K = n_k \) and \( V_k^i = \text{span} \{ \phi_k^i \} \), the subspace spanned by nodal basis function \( \phi_k^i \), \( i = 1, 2, \cdots, n_k \). Then the condition (i) in Theorem 5.2 is easily satisfied by the local property of finite element nodal basis functions. For any \( v \in (I - P_{k-1}) V_k \), it remains to find the decomposition \( v = \sum_{i=1}^{n_k} v_k^i \), \( v_k^i \in V_k^i \), so that (ii) of Theorem 5.2 is satisfied. To do so, we take the canonical decomposition \( v = \sum_{i=1}^{n_k} v(x_k^i) \phi_k^i \) with \( v_k^i = v(x_k^i) \phi_k^i \in V_k^i \). It is easy to see that

\[
\sum_{i=1}^{n_k} \| v_k^i \|^2_{L^2(\Omega)} \leq C \sum_{i=1}^{n_k} h_k^d v(x_k^i)^2 \leq C \| v \|^2_{L^2(\Omega)}
\]

by the scaling argument. Thus by the inverse estimate and (5.4) we get

\[
\sum_{i=1}^{n_k} a(v_k^i, v_k^i) \leq Ch_k^{-2} \sum_{i=1}^{n_k} \| v_k^i \|^2_{L^2(\Omega)} \leq Ch_k^{-2} \| v \|^2_{L^2(\Omega)} \leq C a(v, v),
\]

for any \( v \in (I - P_{k-1}) V_k \).

**Theorem 5.3.** Let \( B_k \) be the standard multigrid V-cycle with symmetric Gauss-Seidel relaxation as the smoothing operator. Then the exists a constant \( C \) independent of \( M_k \) and \( m \geq 1 \) such that

\[
\| I - B_k A_k \|_A \leq \frac{C}{C + m}.
\]
where
\[ \|I - B_k A_k\|_A = \sup_{0 \neq v \in V_k} \frac{a((I - B_k A_k)v, v)}{\|v\|_A}. \]

**Example 5.4.** Consider the Poisson equation \(-\Delta u = 1\) with homogeneous Dirichlet condition on unit square discretized with uniform triangulations. We solve the problem by the V-cycle algorithm (5.12) with zero initial value, Gauss-Seidel smoother \((m = 2)\), and stopping rule
\[ \|\tilde{f}_k - \tilde{A}_k \tilde{u}_k^{(n)}\|_\infty / \|\tilde{f}_k - \tilde{A}_k \tilde{u}_k^{(0)}\|_\infty < 10^{-6}. \]

The initial mesh consists of 4 triangles. Table 1 shows the number of multigrid iterations after 1–10 uniform refinements by the “newest vertex bisection” algorithm. The final mesh consists of 4194304 triangles and 2095105 interior nodes. For an implementation of the V-cycle algorithm we refer to Section 10.4.

<table>
<thead>
<tr>
<th>(N)</th>
<th>5</th>
<th>25</th>
<th>113</th>
<th>481</th>
<th>1985</th>
<th>8065</th>
<th>32513</th>
<th>130561</th>
<th>523265</th>
<th>2095105</th>
</tr>
</thead>
<tbody>
<tr>
<td>(l)</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 1. Number of multigrid iterations \((l)\) versus number of degrees of freedom \((N)\) for Example 5.4.

**5.5. The full multigrid and work estimate**

We shall now describe a more efficient multigrid technique, called the *full multigrid cycle*. Recall that \(u_k\) is the \(k^{th}\) level finite element solution. By the convergence theory of finite element methods in Chapter 3 we have the following error estimates:
\[ \|u - u_k\|_A \leq c_1 h_k, \quad k \geq 1, \tag{5.21} \]
where \(c_1 > 0\) is a constant independent of \(k\). The full multigrid method \((FMG)\) is based on the following two observations:

1. \(u_{k-1} \in V_{k-1} \subset V_k\) is closed to \(u_k \in V_k\) and hence can be used as an initial guess for an iterative scheme for solving \(u_k\);
2. Each \(u_k\) can be solved within its truncation error by a multigrid iterative scheme.

**Algorithm 5.2.** \((FMG)\).

For \(k = 1\), \(\hat{u}_1 = A_1^{-1}f_1\);
For $k \geq 2$, let $\hat{u}_k = \hat{u}_{k-1}$, and iterate $\hat{u}_k \leftarrow \hat{u}_k + B_k(f_k - A_k\hat{u}_k)$ for $l$ times.

Denote by $\tilde{h}_k = \max_{K \in M_k} |K|^{1/d}$. It is clear that there exists a positive number $p > 1$ such that $\tilde{h}_k = \tilde{h}_{k-1}/p$ and that $\tilde{h}_k$ is equivalent to $h_k$, that is, there exist positive constants $c_2$ and $c_3$ depending only on the minimum angle of the meshes such that $c_2 \tilde{h}_k \leq h_k \leq c_3 \tilde{h}_k$. The following theorem says that the above full multigrid algorithm can produce results with errors comparable to the errors of the finite element solutions.

**Theorem 5.5.** Assume that Theorem 5.1 holds and that $\delta_l < 1/p$. Then

$$\|u_k - \hat{u}_k\|_A \leq \frac{c_3 \rho \delta^l}{c_2 (1 - p \delta^l)} c_1 h_k, \quad k \geq 1.$$  

**Proof.** By Theorem 5.1 we have

$$\|u_k - \hat{u}_k\|_A \leq \delta^l \|u_k - \hat{u}_{k-1}\|_A \leq \delta^l (\|u_k - u_{k-1}\|_A + \|u_{k-1} - \hat{u}_{k-1}\|_A).$$

Noting that $\|u_1 - \hat{u}_1\|_A = 0$, we conclude that

$$\|u_k - \hat{u}_k\|_A \leq \sum_{n=1}^{k-1} (\delta^l)^n \|u_{k-n+1} - u_{k-n}\|_A \leq \sum_{n=1}^{k-1} (\delta^l)^n \|u - u_{k-n}\|_A$$

$$\leq c_1 \sum_{n=1}^{k-1} (\delta^l)^n \tilde{h}_{k-n} \leq c_1 c_3 \sum_{n=1}^{k-1} (\delta^l)^n \tilde{h}_{k-n}$$

$$\leq c_1 c_3 \tilde{h}_k \sum_{n=1}^{k-1} (p \delta^l)^n \leq \frac{c_1 c_3}{c_2} \frac{p \delta^l}{1 - p \delta^l} h_k.$$  

This completes the proof. \qed

We now turn our attention to the work estimate. It is clear that

$$n_k = \dim V_k \sim \frac{1}{h_k^d} \sim \frac{1}{\tilde{h}_k^d} \sim (p^d)^k.$$  \hspace{1cm} (5.22)

**Theorem 5.6.** The work involved in the FMG is $O(n_k)$.

**Proof.** Let $W_k$ denote the work in the $k^{th}$ level V-cycle iteration. Together, the smoothing and correction steps yield

$$W_k \leq C m n_k + W_{k-1}.$$  

Hence

$$W_k \leq C m (n_1 + n_2 + \cdots + n_k) \leq C n_k.$$
Let \( \hat{W}_k \) denote the work involved in obtaining \( \hat{u}_k \) in the FMG. Then
\[
\hat{W}_k \leq \hat{W}_{k-1} + \hat{W}_k \leq \hat{W}_{k-1} + Cn_k.
\]
Thus we have
\[
\hat{W}_k \leq C(n_1 + \cdots + n_k) \leq Cn_k.
\]
This completes the proof. \( \square \)

This theorem shows that the FMG has an optimal computational complexity \( O(n_k) \) to compute the solution within truncation error. In contrast, the computational complexity of the \( k \)th level V-cycle iteration is not optimal, because its number of operations required to compute the solution within truncation error is \( O(n_k \log \frac{1}{h_k}) = O(n_k \log n_k) \).

### 5.6. The adaptive multigrid method

The distinct feature of applying multigrid methods on adaptively refined finite element meshes is that the number of nodes of \( M_k \) may not grow exponentially with respect to the number of mesh refinements \( k \). In practice, local relaxation schemes are used in applying multigrid methods on adaptively refined finite element meshes.

Let \( \tilde{N}_k \) be the set of nodes on which local Gauss-Seidel relaxation are carried out
\[
\tilde{N}_k = \{ z \in \tilde{N}_k : z \text{ is a new node or } z \in N_{k-1} \text{ but } \phi^z_k \neq \phi^{z}_{k-1} \},
\]
where \( \phi^z_k \) is the nodal basis function at the node \( z \) in \( V_k \). For convenience we denote \( \tilde{N}_k = \{ x^j_k : j = 1, \ldots, \tilde{n}_k \} \). The local Gauss-Seidel iterative operator is given by
\[
R_k = (I - (I - P_{n_k}^{1}) \cdots (I - P_{1}^{k}))A^{-1}_k.
\]

The following theorem is proved by Wu and Chen [52].

**Theorem 5.7.** Let the meshes \( M_k, 0 \leq k \leq J \), be obtained by the “newest vertex bisection” algorithm in Section 10.4. Let each element \( K \in M_k \) is obtained by refining some element \( K' \in M_{k-1} \) finite number of times so that \( h_{K'} \leq Ch_K \). Then the standard multigrid V-cycle with local Gauss-Seidel relaxation satisfies
\[
\| I - E_k A_k \|_A < \delta
\]
for some constant \( \delta < 1 \) independent of \( k \) and \( M_k \).
Bibliographic notes. There is a rich literature on the mathematical theory of multigrid methods. We refer to Brandt [14], the book Bramble [12], and the review paper Xu [53] for further mathematical results. Our development in Section 5.3 follows Arnold et al [3]. The full multigrid method is introduced in Brandt [15]. The convergence of the adaptive multigrid finite element method is considered in Wu and Chen [52].

5.7. Exercises

Exercise 5.1. Prove Lemma 5.3 and Lemma 5.4.

Exercise 5.2. Prove Lemma 5.6.

Exercise 5.3. Let $R_k$ be symmetric with respect to $(\cdot, \cdot)$ and let $K_k = I - R_k A_k$. Then $R_k$ is semi-definite and satisfies

$$a(K_k v, v) \geq 0 \quad \forall \; v \in V_k$$

is equivalent to

$$\|K_k\|_A \leq 1 \quad \text{and} \quad \|I - K_k\|_A \leq 1.$$