CONVERGENCE ESTIMATES FOR MULTIGRID ALGORITHMS WITH KACZMARZ SMOOTHING

KAB SEOK KANG AND DO YOUNG KWAK*

Abstract. We prove some estimates for the convergence of multigrid algorithms with Kaczmarz smoothing applied to symmetric positive definite problems. To estimate the convergence of multigrid algorithms, we assume some properties of smoothing procedure which are satisfied by Kaczmarz, Gauss-Seidel, and Jacobi smoothing and are weaker than earlier assumptions. We use product formula for multigrid algorithm to show the V-cycle convergence factor is $\delta = 1 - \frac{1}{C(j-1)}$. The theory is presented in an abstract setting which can be applied to finite element multigrid. Also, numerical example is provided.

1. Introduction. Many authors presented various convergence analyses of multigrid methods which are often based on certain assumptions concerning the smoothing process. These assumptions are sometimes verified for specific examples in [2, 5, 6]. In this paper, we provide a weaker assumption under which multigrid algorithms are shown to converge.

Assumptions concerning the smoothing process which given in [6] are satisfied by point, line and block Jacobi and point, line and block Gauss-Seidel smoothing but not by Kaczmarz smoothing, for which the constant C_R the assumption (C.1) in [6] grows with $1/h^2$. Instead, the latter satisfies a weaker condition introduced in this paper.

The outline of the remainder of this paper is as follows. Section 2 describes the basic multigrid algorithm in an abstract setting and gives some of the conditions on the smoothers. In §3, we prove multigrid convergence under the weaker assumption. Kaczmarz smoothing procedures are described and analyzed in §4. Some numerical experiments showing the Kacmarz smoother satisfies our weak assumption are reported there. Finally, in §5, we discuss the finite element multigrid applications.

2. The Multigrid Algorithms. We assume that there is a sequence of nested finite-dimensional inner product spaces $\mathcal{M}_1 \subset \mathcal{M}_2 \subset \ldots \subset \mathcal{M}_j$ with $(\cdot,\cdot)_k$. In addition, we assume that there are symmetric positive definite operators $A_k : \mathcal{M}_k \to \mathcal{M}_k$ for $k = 1, \ldots, j$. We denote $A(\cdot,\cdot) = (A_j\cdot,\cdot)_j$. The multigrid algorithm is an iterative procedures for the solution of the problem on $\mathcal{M}_j$, i.e., given $f \in \mathcal{M}_j$ find $u \in \mathcal{M}_j$ satisfying

$$A_j u = f. \tag{1}$$

We define the projectors $P^0_{k-1} : \mathcal{M}_k \to \mathcal{M}_{k-1}$ and $P_{k-1} : \mathcal{M}_k \to \mathcal{M}_{k-1}$ by, for all $\phi \in \mathcal{M}_{k-1}$, $(P^0_{k-1}v, \phi)_{k-1} = (v, \phi)_k$ and $A(P_{k-1}v, \phi) = A(v, \phi)$, respectively. '$T$' and '$*$' will denote adjoint with respect to $(\cdot,\cdot)_k$ and $A(\cdot,\cdot)$, respectively.

Also, we require a sequence of linear smoothing operators $R_k : \mathcal{M}_k \to \mathcal{M}_k$ for $k = 2, \ldots, j$. We shall always take $R_1 = A_1^{-1}$.

We set

$$R_k^{(l)} = \begin{cases} R_k & \text{if } l \text{ is odd,} \\ R_k^T & \text{if } l \text{ is even} \end{cases}$$

and set $K_k = I - R_k A_k$.

* Department of Mathematicss, Korea Advanced Institute of Science and Technology, Taejon, Korea 305–701, email: dykwak@math.kaist.ac.kr, kks002@math.kaist.ac.kr.

Domain Decomposition Methods in Sciences and Engineering, edited by R. Glowinski *et al.*

We next define a multigrid process for iteratively computing the solution of (1).

Algorithm Set $B_1^s = A_1^{-1}$. Assume that B_{k-1}^s has been defined and define $B_k^s g$ for $g \in \mathcal{M}_k$ as follows;

1. Set $v^0 = 0$ and $q^0 = 0$.
2. Define v^i for $i = 1, 2, \ldots, m(k)$ by $v^i = v^{i-1} + R_k^{(i+m(k))}(g - A_k v^{i-1})$.
3. Define $w^{m(k)} = v^{m(k)} + q^p$, where q^i, for $i = 1, \ldots, p$, is defined by

$$q^i = q^{i-1} + B_{k-1}^s [P_{k-1}^0 (g - A_k v^{m(k)}) - A_{k-1} q^{i-1}].$$

4. Define w^i for $i = m(k)+1, \ldots, 2m(k)$ by $w^i = w^{i-1} + R_k^{(i+m(k))}(g - A_k w^{i-1})$.
5. Set $B_k^s g = w^{2m(k)}$.

In the above algorithm, by defining $B_k^n g = w^{m(k)}$, we get nonsymmetric multigrid algorithm B_k^n. From the above algorithm, fundamental recurrence relations for the nonsymmetric and the symmetric multigrid algorithm are

$$\begin{aligned} I - B_k^n A_k &= [(I - P_{k-1}) + (I - B_{k-1}^n A_{k-1})^p P_{k-1}] \bar{K}_k^{(m(k))}, \\ I - B_k^s A_k &= (\bar{K}_k^{(m(k))})^* [(I - P_{k-1}) + (I - B_{k-1}^s A_{k-1})^p P_{k-1}] \bar{K}_k^{(m(k))} \end{aligned}$$

on $\mathcal{M}_k$ where

$$\bar{K}_k^{(m(k))} = \begin{cases} (K_k^* K_k)^{m(k)/2} & \text{if } m(k) \text{ is even,} \\ K_k (K_k^* K_k)^{(m(k)-1)/2} & \text{if } m(k) \text{ is odd.} \end{cases}$$

To estimate the convergence of multigrid algorithm, we need some conditions concerning the smoothing operators. The conditions which were often assumed by many authors ([1–7]) are

(C.1) There is a constant C_R which does not depend on k such that the smoothing procedure satisfies

$$\lambda_k^{-1} \|u\|_k^2 \le C_R (\bar{R}_k u, u)_k \qquad \text{for all } u \in \mathcal{M}_k.$$

Here, $\bar{R}_k$ is either $(I - K_k^* K_k) A_k^{-1}$ or $(I - K_k K_k^*) A_k^{-1}$. λ_k is the largest eigenvalue of A_k.

(C.2) Let $T_k = R_k A_k$. There is a constant $\theta < 2$ not depending on k satisfying

$$A(T_k v, T_k v) \le \theta A(T_k v, v).$$

But (C.1) is not satisfied by some smoothers, say Kaczmarz smoothing. Thus we modify (C.1) as follows

(SM.1) There is a constant C_R which does not depend on k such that the smoothing procedure satisfies

$$\lambda_k^{-2} (A_k u, u)_k \le C_R (\bar{R}_k u, u)_k \qquad \text{for all } u \in \mathcal{M}_k.$$

3. Convergence Estimates for Multigrid Algorithms. To estimate the convergence of multigrid algorithm, we need some properties concerning the operator A_k and the subspaces. These are as follows:

(O.1) There exists a sequence of linear operators $Q_k : \mathcal{M}_j \to \mathcal{M}_k$ for $k = 1, \ldots, j$, with $Q_j = I$ satisfying the following properties. There are constants C_1 and C_2 not depending on k for which

$$(2)\quad \begin{aligned} (A_k^{-1}(Q_k - Q_{k-1})u, (Q_k - Q_{k-1})u)_k &\le C_1\lambda_k^{-2}(A_k u, u)_k \quad \text{for } k = 2, \ldots, j, \\ A(Q_k u, Q_k u) &\le C_2 A(u, u) \quad \text{for } k = 1, \ldots, j-1. \end{aligned}$$

THEOREM 1. *Assume that (O.1) hold. Let R_k hold (SM.1) and (C.2). Let B_j^s be defined by symmtric multigrid algorithm with $p = 1$. Then*

$$(3)\qquad A((I - B_j^s)v, v) \le \delta_j A(v, v) \quad \text{for all } v \in \mathcal{M}_j$$

hold with $\delta_j = 1 - \frac{1}{C(j-1)}$ where $C = [(1 + C_2^{1/2})(2\theta/(2-\theta))^{1/2} + (C_R C_1)^{1/2}]^2$.

PROOF. We observe that

$$I - B_j^s A_j = (I - B_j^n A_j)^*(I - B_j^n A_j)$$

and for $T_k = (I - (\bar{K}_k^{(m(k))})^*)P_k$, that

$$(I - B_j^n A_j)^* = (I - T_j)(I - T_{j-1}) \cdots (I - T_1).$$

To use a product analysis, we set $E_0 = I$ and $E_k = (I - T_k)E_{k-1}$. We get

$$\begin{aligned} A(u, u) - A(E_j u, E_j u) &= \textstyle\sum_{k=1}^{j} [A(E_{k-1}u, E_{k-1}u) - A(E_k u, E_k u)] \\ &= \textstyle\sum_{k=1}^{j} A((2I - T_k)E_{k-1}u, T_k E_{k-1}u). \end{aligned}$$

Note that $I - B_j^n A_j = E_j^*$ and hence the inequality (3) will follow if we show that

$$(4)\qquad A(u, u) \le C(j-1) \sum_{k=1}^{j} A((2I - T_k)E_{k-1}u, T_k E_{k-1}u).$$

From the fact that $Q_j = I$, we get

$$(5)\qquad \begin{aligned} A(u, u) &= \sum_{k=2}^{j} A(E_{k-1}u, (Q_k - Q_{k-1})u) + A(u, Q_1 u) \\ &+ \sum_{k=2}^{j} A((I - E_{k-1})u, (Q_k - Q_{k-1})u). \end{aligned}$$

For the first sum on the right-hand side (5), from (O.1) and (SM.1), we see that

$$\begin{aligned} &\sum_{k=2}^{j} A(E_{k-1}u, (Q_k - Q_{k-1})u) = \sum_{k=2}^{j} (A_k^{-1} A_k^2 P_k E_{k-1}u, (Q_k - Q_{k-1})u)_k \\ &\le \sum_{k=2}^{j} A(A_k P_k E_{k-1}u, A_k P_k E_{k-1}u)^{1/2} \cdot (A_k^{-1}(Q_k - Q_{k-1})u, (Q_k - Q_{k-1})u)_k^{1/2} \\ &\le (C_R C_1)^{1/2} A^{1/2}(u, u) \sum_{k=2}^{j} A^{1/2}((I - K_k^* K_k)P_k E_{k-1}u, P_k E_{k-1}u) \\ &\le (C_R C_1 (j-1))^{1/2} A^{1/2}(u, u) \left(\sum_{k=2}^{j} A((I - K_k^* K_k)P_k E_{k-1}u, P_k E_{k-1}u) \right)^{1/2} \end{aligned}$$

The remainder of proof is the same with the proof of Theorem 4.3 in [6]. □

4. Smoothing Procedures in Multigrid Algorithms. We define the multicative smoother(Gauss-Seidel smoother) by the following algorithm.

Algorithm 4.1. Let $f \in \mathcal{M}_k$, we define $R_k f \in \mathcal{M}_k$ as follows;

1. Set $v_0 = 0$.
2. Define v_i for $i = 1, \ldots, l$ by $v_i = v_{i-1} + A_{k,i}^{-1} Q_k^i (f - A_k v_{i-1})$ where $A_{k,i}$ is an i-th diagonal element of A_k and Q_k^i is a projection onto span$\{e_i\}$ with respect to $(\cdot,\cdot)_k$.
3. Set $R_k f = v_l$.

Let P_k^i is the projection onto span$\{e_i\}$ with respet to $(A_k\cdot,\cdot)_k$. It immediately follows from the identity $A_{k,i} P_k^i = Q_k^i A_k$ that $K_k = (I - P_k^l) \cdots (I - P_k^1)$. It was shown that Gauss-Seidel smoother satisfies (C.1)([6]).

Theorem 2 Let R_k be a smoother which satisfy (C.1). Then R_k satisfy (SM.1).

PROOF. Let $u = A_k w$, then (C.1) becomes

$$\lambda_k^{-1}(A_k w, A_k w)_k \le C_R[(A_k w, w)_k - (A_k K_k w, K_k w)_k].$$

From the fact, for all $w \in \mathcal{M}_k$, $(A_k w, w)_k \le \lambda_k (w, w)_k$, (SM.1) is satisfied. □

Kaczmarz smoother is defined by the following algorithm.

Algorithm 4.2 Let $f \in \mathcal{M}_k$. We define $R_k f \in \mathcal{M}_k$ as follows:

1. Set $v_0 = 0$.
2. Define v_i for $i = 1, \ldots, l$ by $v_i = v_i - \frac{a_i}{a_i^T a_i}(a_i^T v_{i-1} - f_i)$ where $a_i^T = i^{\text{th}}$ row of A_k.
3. Set $R_k f = v_l$.

From the above algorithm, we know that

$$K_k = I - R_k A_k = (I - B_k^l) \cdots (I - B_k^1) \text{ where } B_k^i = \frac{a_i a_i^T}{a_i^T a_i} = \frac{A_k^T Q_k^i A_k}{(A_k A_k^T)_{ii}}. \tag{6}$$

Let S_k^i be the projection with respect to $((A_k A_k^T)\cdot,\cdot)_k$. Then we get $(A_k A_k^T)_{ii} S_k^i = Q_k^i (A_k A_k^T)$. From $B_k^i = A_k^T S_k^i A_k^{-T}$, (6) becomes $K_k = A_k^T (I - S_k^l) \cdots (I - S_k^1) A_k^{-T}$. Above presentation reflects that Kaczmarz iteration can be regarded as a Gauss-Seidel iteration applied to $A_k A_k^T u = f$. In fact, from Gauss-Seidel iteration, we know that $(I - S_k^l) \cdots (I - S_k^1) = I - (D + L)^{-1}(A_k A_k^T)$ where $A_k A_k^T = D + L + L^T$. Therefore $K_k = I - A_k^T (D + L)^{-1} A_k$.

Numerical verification of (SM.1) and (C.2). We consider an elliptic partial differential equation of the form

$$\begin{aligned} -\textstyle\sum_{i,j=1}^2 \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial v}{\partial x_j}(x)\right) &= f(x) \quad \text{in } \Omega, \\ v(x) &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{7}$$

where Ω is a unit square and $(a_{ij})_{i,j=1}^2$ is a symmetric positive matrix.

To obtain A_k, we discretize Ω by quasi-uniform triangular element with $h_k = 2^{-k}$ and define $\mathcal{M}_k$ to be the set of piecewise linear functions which vanish on $\partial\Omega$.

First, for $(a_{ij}) = I$, we calculate C_R's in (C.1) and (SM.1) of damped Jacobi with $\omega = 0.8$, Gauss-Seidel, and Kaczmarz smoothing.

	Jacobi		Gauss-Seidel		Kaczmarz	
h	(C.1)	(SM.1)	(C.1)	(SM.1)	(C.1)	(SM.1)
1/8	1.409707	1.409707	1.118052	1.118052	2.281606	1.349255
1/16	1.519271	1.519271	1.123504	1.123504	8.349859	1.371804
1/32	1.551332	1.551332	1.124665	1.124665	32.667680	1.376693
1/64	1.559570	1.559776	1.124900	1.124900	129.937300	1.377789

Table I

Table I show that Kaczmarz smoother does not satisfy (C.1).

Next, we calculate C_R in (SM.1) and θ in (C.2) of Kaczmarz smoother for other (a_{ij})'s.

	I		$p(x)I$		$q(x)$	
h	C_R	θ	C_R	θ	C_R	θ
1/8	1.349255	1.496815	1.321122	1.455358	1.322647	1.460811
1/16	1.371804	1.516820	1.342990	1.479987	1.341607	1.483321
1/32	1.376693	1.522054	1.356463	1.495506	1.353792	1.497318
1/64	1.377789	1.523371	1.364844	1.505585	1.361640	1.506282

Table II

Here $p(x) = e^{0.2x+0.7y}$ and $q(x) = \text{diag}(e^{0.2x+0.3y}, x + 0.5)$.

5. Finite Elements Applications . We shall consider the problem of approximating the solution v of (7). The form A corresponding to the above operator is given by

$$A(v,w) = \sum_{i,j=1}^{n} \int_\Omega a_{ij} \frac{\partial v}{\partial x_i} \frac{\partial w}{\partial x_j} dx \qquad \text{for all } v, w \in \mathcal{M}_k.$$

Clearly, $U \in H_0^1(\Omega)$ is the solution of

$$A(U,\theta) = (F,\theta) \qquad \text{for all } \theta \in H_0^1(\Omega).$$

By positive definiteness of $\{a_{ij}\}$, $\|\cdot\|_A = A^{1/2}(\cdot,\cdot)$ is a norm on $H_0^1(\Omega)$ and this norm is equivalence to $\|\cdot\|_1$ which denote $H^1(\Omega)$-norm.

We let τ_k^i be a sequence of quasi-uniform triangulations of size h_k for $k = 1, \dots, j$. We define $\mathcal{M}_k$ as in the previous section. Since Ω is polygonal, the subspaces are nested.

Let $\{y_k^i\}$ be the collection of nodes corresponding to the triangulation for $\mathcal{M}_k$. Let $(u,v)_k = h_k^2 \sum_i u(y_k^i) v(y_k^i)$. Note that the quasi-uniformity of the triangulations implies that the norm $\|\cdot\|_k$ is equivalent to the L^2 norm on the subspace $\mathcal{M}_k$. The operator A_k, $k = 1, \dots, j$, are then defined by, for all $\phi \in \mathcal{M}_k$, $(A_k v, \phi)_k = A(v, \phi)$.

Let Q_k denote the $L^2(\Omega)$ projection onto $\mathcal{M}_k$. We know that, since the triangulations are quasi-uniform and inverse property, for all $v \in H_0^1(\Omega)$,

$$\|(I - Q_k)v\| \le c h_k \|v\|_1 \qquad \text{and} \qquad \|Q_k v\|_1 \le C \|v\|_1. \tag{8}$$

From (8) and definition, we get

$$\begin{aligned}
&(A_k^{-1}(Q_k - Q_{k-1})v, (Q_k - Q_{k-1})v)_k \\
&= \sup_{u \in \mathcal{M}_k} \frac{((Q_k - Q_{k-1})v, u)^2}{(A_k u, u)_k} = \sup_{u \in \mathcal{M}_k} \frac{((Q_k - Q_{k-1})v, (Q_k - Q_{k-1})u)^2}{(A_k u, u)_k} \\
&\le \sup_{u \in \mathcal{M}_k} \frac{\|(Q_k - Q_{k-1})v\|^2 \|(Q_k - Q_{k-1})u\|^2}{(A_k u, u)_k} \le \frac{c h_k^2 \|v\|_1^2 c h_k^2 \|u\|_1^2}{\|u\|_1^2} \\
&\le C h_k^4 (A_k v, v)_k \le C \lambda_k^{-2} (A_k v, v)_k
\end{aligned} \tag{9}$$

since $\|(Q_k - Q_{k-1})v\| \le \|(I - Q_k)v\| + \|(I - Q_{k-1})v\|$ and $\lambda_k = O(h_k^{-2})$. Combining (8) and (9) shows that (O.1).

REFERENCES

[1] R.E. Bank and C.C. Douglas, *Sharp estimates for multigrid rates of convergence with general smoothing and acceleration*, SIAM J. Numer. Anal. **22** (1985), 617–633.

[2] R.E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.

[3] D. Braess and W. Hackbush, *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal. **20** (1983), 967–975.

[4] J.H. Bramble, D.Y. Kwak, and J.E. Pasciak, *Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems*, SIAM J. Numer. Anal. **31** (1994).

[5] J.H. Bramble and J. E. Pasciak, *New convergence estimates for multigrid algorithms* Math. Comp. **49** (1987), 311–329.

[6] J.H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms* Math. Comp. **58** (1992), 467–488.

[7] J.H. Bramble and J.E. Pasciak, *New estimates for multigrid algorithms including the V-cycle*, Math. Comp. **60** (1994), 447–471.

[8] J.H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions* Math. Comp. **56** (1991), 23–45.

[9] W. Hackbush, *Multi-grid methods and applications*, Springer-Verlag, NewYork, 1985.

[10] S. McCormick (Ed.) *Multigrid methods*, SIAM, Philadelphia, PA, 1987.