Conference Applications of Mathematics 2013
in honor of the 70th birthday of Karel Segeth.
Jan Brandts, Sergey Korotov, Michal Křížek,
Jakub Šístek, and Tomáš Vejchodský (Eds.),
Institute of Mathematics AS CR, Prague 2013

A SHORT PHILOSOPHICAL NOTE ON THE ORIGIN OF SMOOTHED AGGREGATIONS

Pavla Fraňková, Milan Hanuš, Hana Kopincová, Roman Kužel,
Petr Vaněk, Zbyněk Vastl

University of West Bohemia
Univerzitni 22, 306 14 Pilsen, Czech Republic
frankova@kma.zcu.cz, mhanus@kma.zcu.cz, kopincov@kma.zcu.cz, rkuzel@kma.zcu.cz,
ptrvnk@kma.zcu.cz, zvastl@kma.zcu.cz

Abstract

We derive the smoothed aggregation two-level method from the variational objective to minimize the *final error* after finishing the entire iteration. This contrasts to a standard variational two-level method, where the coarse-grid correction vector is chosen to minimize the error after coarse-grid correction procedure, which represents merely an intermediate stage of computing. Thus, we enforce *the global minimization of the error.* The method with smoothed prolongator is thus interpreted as a qualitatively different, and more optimal, algorithm than the standard multigrid.

1. Introduction

The smoothed aggregation method [13, 14, 15, 12] proved to be a very efficient tool for solving various types of elliptic problems and their singular perturbations. In this short note, we turn to the very roots of smoothed aggregation method and derive its two-level variant on a systematic basis.

The multilevel method consists in combination of a coarse-grid correction and smoothing. The coarse-grid correction of a standard two-level method is derived using the A-orthogonal projection of an error to the range of the prolongator. In other words, the coarse-grid correction vector is chosen to minimize the error *after coarse-grid correction procedure.* This means, the standard two-level method minimizes the error in an intermediate stage of the iteration, while we are, naturally, interested in minimizing *the final error after accomplishing the entire iteration.* In other words, we strive to minimize the error after coarse-grid correction and subsequent smoothing. The two-level smoothed aggregation method is obtained by solving this minimization problem. This, in the opinion of the authors, explains its remarkable robustness.

We derive the two-level smoothed aggregation method from the variational objective to minimize the error after coarse-grid correction and subsequent post-smoothing. Then, by a trivial argument, we extend our result to the two-level method with pre-smoothing, coarse-grid correction and post-smoothing.

The minimization of error after coarse-grid correction and subsequent smoothing leads to a method with smoothed prolongator. We can say that by smoothing the prolongator, we adapt the coarse-space (the range of the prolongator) to the post-smoother so that the resulting iteration is as efficient as possible. Our short explanation applies to any two-level method with smoothed prolongator. The particular case we have in mind is, however, a method with smoothed *tentative* prolongator given by generalized unknowns aggregations [15]. The discrete basis functions of the coarse-space (the columns of the prolongator) given by unknowns aggregations have no overlap; the natural overlap of discrete basis functions (like it is in the case of finite element basis functions) is created by smoothing and, for additive point-wise smoothers, leads to sparse coarse-level matrix.

Our argument is basically trivial. It, however, shows a fundamental property of the method with smoothed prolongator, that is essential. This argument is known to the authors for a long time, but has never been published.

We conclude our paper by a numerical test. Namely, we demonstrate experimentally that smoothed aggregation method with powerful smoother and small coarse-space solves efficiently highly anisotropic problems without the need to perform semi-coarsening (the coarsening that follows only strong connections).

2. Two-level method

We solve a system of linear algebraic equations

$$A\mathbf{x} = \mathbf{f}, \tag{1}$$

where A is a symmetric positive definite matrix of order n and $\mathbf{f} \in \mathbb{R}^n$. We assume that an injective linear *prolongator* $p : \mathbb{R}^m \to \mathbb{R}^n,\ m < n$ is given.

The two-level method consists in the combination of a *coarse-grid correction* and *smoothing*. The smoothing means using point-wise iterative methods at the beginning and at the end of the iteration. The coarse-grid correction is derived by correcting an error $\mathbf{e}$ by a coarse-level vector $\mathbf{v}$ so that the resulting error $\mathbf{e} - p\mathbf{v}$ is minimal in A-norm. In other words, we solve the minimization problem

$$\text{find } \mathbf{v} \in \mathbb{R}^m \text{ so that } \|\mathbf{e} - p\mathbf{v}\|_A \text{ is minimal.} \tag{2}$$

It is well-known that such vector $p\mathbf{v}$ is an A-orthogonal projection of the error $\mathbf{e}$ onto $\mathrm{Range}(p)$, with the projection operator given by

$$Q = p(p^T Ap)^{-1} p^T A.$$

Thus, the error propagation opeartor of the coarse-grid correction is given by $I - Q = I - p(p^T Ap)^{-1} p^T A$ and the error propagation operator of the two-level method by

$$E_{TGM} = S_{post}[I - p(p^T Ap)^{-1} p^T A] S_{pre}, \tag{3}$$

where S_{pre} and S_{post} are error propagation operators of pre- and post- smoothing iterations, respectively.

Clearly, for the error $\mathbf{e}(\mathbf{x}) \equiv \mathbf{x} - A^{-1}\mathbf{f}$ we have $A\mathbf{e}(\mathbf{x}) = A\mathbf{x} - \mathbf{f}$. Hence, the coarse-grid correction can be algorithmized as

$$\mathbf{x} \leftarrow \mathbf{x} - p(p^T A p)^{-1} p^T (A\mathbf{x} - \mathbf{f})$$

and the variational two-level algorithm with post-smoothing step proceeds as follows:

Algorithm 1

1. *Pre-smooth:* $\mathbf{x} \leftarrow \mathcal{S}_{pre}(\mathbf{x}, \mathbf{f})$,
2. *evaluate the residual:* $\mathbf{d} = A\mathbf{x} - \mathbf{f}$,
3. *restrict the residual:* $\mathbf{d}_2 = p^T \mathbf{d}$,
4. *solve a coarse-level problem* $A_2 \mathbf{v} = \mathbf{d}_2, \quad A_2 = p^T A p$,
5. *correct the approximation* $\mathbf{x} = \mathbf{x} - p\mathbf{v}$,
6. *post-smooth* $\mathbf{x} = \mathcal{S}_{post}(\mathbf{x}, \mathbf{f})$.

Here, $\mathcal{S}_{pre}(.,.)$ and $\mathcal{S}_{post}(.,.)$, respectively, represent one or more iterations of point-wise iterative methods for solving (1).

The coarse-grid correction vector $\mathbf{v}$ is chosen to minimize the error after Step 5 of Algorithm 1. Thus, we conclude that in the case of a standard variational multigrid, the coarse-grid correction procedure minimizes the error in an intermediate stage of the iteration, while we are in fact interested in minimizing the final error after accomplishing the entire iteration. This means to minimize the error after coarse-grid correction with subsequent smoothing.

3. The smoothed aggregation two-level method

In the smoothed aggregation method, we construct the coarse-grid correction to minimize the error *after coarse-grid correction with subsequent smoothing*, which means the final error on the exit of the iteration procedure. The minimization of the error after pre-smoothing, coarse-grid correction and post-smoothing then follows immediately by a trivial argument.

Let S be the error propagation operator of the post-smoother $\mathcal{S}(.,.) = \mathcal{S}_{post}(.,.)$. Throughout this section we assume that S is sparse. This is due to the fact that the above minimization problem leads to smoothed prolongator $P = Sp$ and we need a sparse coarse-level matrix $A_2 = P^T A P$. The additive point-wise smoothing methods have, in general, sparse error propagation operator; this is the case of Jacobi method or Richardson's iteration.

For a multilevel method with post-smoothing only, the error after coarse-grid correction and subsequent smoothing is given by

$$S(\mathbf{e} - p\mathbf{v}), \tag{4}$$

where $\mathbf{v}$ is a correction vector and $\mathbf{e}$ the error on the entry of the iteration procedure. We choose $\mathbf{v}$ so that the error in (4) is minimal in A-norm, that is, we solve the minimization problem

$$\text{find } \mathbf{v} \in \mathbb{R}^m \text{ such that } \|S(\mathbf{e} - p\mathbf{v})\|_A \text{ is minimal.} \tag{5}$$

Since $\|S(\mathbf{e} - p\mathbf{v})\|_A = \|\mathbf{e} - p\mathbf{v}\|_{S^TAS}$, the minimum is attained for $\mathbf{v}$ satisfying

$$\langle S^T AS(\mathbf{e} - p\mathbf{v}), p\mathbf{w}\rangle = 0 \; \forall \mathbf{w} \in \mathbb{R}^m.$$

We have $\langle S^T AS(\mathbf{e} - p\mathbf{v}), p\mathbf{w}\rangle = \langle p^T S^T AS(\mathbf{e} - p\mathbf{v}), \mathbf{w}\rangle$, hence the above identity is equivalent to $p^T S^T ASp\mathbf{v} = p^T S^T AS\mathbf{e}$ and setting $P = Sp$, it becomes

$$P^T AP\mathbf{v} = P^T AS\mathbf{e}. \tag{6}$$

Here, $\mathbf{e}$ is the error on the entry of the iteration procedure. Assume for now that P is injective. Then by (6), we have $\mathbf{v} = (P^T AP)^{-1} P^T AS\mathbf{e}$ and the error after coarse-grid correction and subsequent smoothing is given by

$$S(\mathbf{e} - p\mathbf{v}) = S\left[\mathbf{e} - p(P^T AP)^{-1} P^T AS\mathbf{e}\right] = \left[I - P(P^T AP)^{-1} P^T A\right] S\mathbf{e}. \tag{7}$$

By comparing the operator

$$E = \left[I - P(P^T AP)^{-1} P^T A\right] S \tag{8}$$

on the right-hand side of (7) with (3), we identify E as the error propagation operator of the variational multigrid with smoothed prolongator $P = Sp$ and pre-smoothing step given by $\mathbf{x} \leftarrow \mathcal{S}(\mathbf{x}, \mathbf{f})$. The algorithm is as follows:

Algorithm 2

1. *Pre-smooth:* $\mathbf{x} \leftarrow \mathcal{S}(\mathbf{x}, \mathbf{f})$*,*
2. *evaluate the residual:* $\mathbf{d} = A\mathbf{x} - \mathbf{f}$*,*
3. *restrict the residual:* $\mathbf{d}_2 = P^T\mathbf{d}$*,*
4. *solve the coarse-level problem:* $A_2\mathbf{v} = \mathbf{d}_2, \quad A_2 = P^T AP$*,*
5. *correct the approximation:* $\mathbf{x} \leftarrow \mathbf{x} - P\mathbf{v}$*.*

Remark 3.1 Note that in the process of the deriving the algorithm in (7), our post-smoother have become a pre-smoother. Nothing was lost in that process; the algorithm minimizes the final error and takes into account the pre-smoother.

Remark 3.2 The smoothed prolongator $P = Sp$ is potentially non-injective, hence the coarse-level matrix $A_2 = P^T AP$ is potentially singular. In this case, we need to replace the inverse of $P^T AP$ in (7) by a pseudo-inverse.

We summarize our considerations in the form of a theorem.

Theorem 3.3 *The error propagation operator E in (8) (the error propagation operator of Algorithm 2) satisfies the identity*

$$\|E\mathbf{e}\|_A = \inf_{\mathbf{v}\in\mathbb{R}^m} \|S(\mathbf{e} - p\mathbf{v})\|_A$$

for all $\mathbf{e} \in \mathbb{R}^n$.

Proof. The proof follows directly from the fact that Algorithm 2 was derived from variational objective (5). □

Remark 3.4 One may also start with the variational objective to minimize the final error after performing the pre-smoothing, the coarse-grid correction and the post-smoothing. Such extension is trivial, the pre-smoother has no influence on the coarse-grid correction operator $I - P(P^T AP)^{-1}P^T A$ and influences only its argument. Indeed, asuming the error propagation operator of the pre-smoother is S^* (the A-adjoint operator), the final error is given by $S(S^*\mathbf{e} - p\mathbf{v})$ and we solve the minimization problem

$$\text{for } \mathbf{e} \in \mathbb{R}^n \text{ find } \mathbf{v} \in \mathbb{R}^m \ : \ \|S(S^*\mathbf{e} - p\mathbf{v})\|_A \text{ is minimal.} \tag{9}$$

Fundamentally, this is the same minimization problem as (5); to derive the corresponding algorithm, it is simply sufficient to follow our manipulations from (5) to (7) with $\mathbf{e} \leftarrow S^*\mathbf{e}$. This way, we end up with a two-level method that has the error propagation operator

$$E = \left[I - P(P^T AP)^{-1}P^T A\right] SS^*, \tag{10}$$

(see (3)) that is, with the algorithm

Algorithm 3

1. *Pre-smooth:* $\mathbf{x} \leftarrow \mathcal{S}_t(\mathbf{x}, \mathbf{f})$*, where* $\mathcal{S}_t$ *is an iterative method with error propagation operator* S^**,*
2. *pre-smooth:* $\mathbf{x} \leftarrow \mathcal{S}(\mathbf{x}, \mathbf{f})$*, where* $\mathcal{S}$ *is an iterative method with error propagation operator* S*,*
3. *evaluate the residual:* $\mathbf{d} = A\mathbf{x} - \mathbf{f}$*,*
4. *restrict the residual:* $\mathbf{d}_2 = P^T\mathbf{d}$*,*
5. *solve the coarse-level problem:* $A_2\mathbf{v} = \mathbf{d}_2, \quad A_2 = P^T AP$*,*
6. *correct the approximation:* $\mathbf{x} \leftarrow \mathbf{x} - P\mathbf{v}$*.*

We summarize the content of Remark 3.4 as a theorem.

Theorem 3.5 *The error propagation operator (10) of Algorithm 3 satisfies the identity*

$$\|E\mathbf{e}\|_A = \inf_{\mathbf{v}\in\mathbb{R}^m} \|S(S^*\mathbf{e} - p\mathbf{v})\|_A$$

for all $\mathbf{e} \in \mathbb{R}^n$.

Proof. The proof follows directly from the fact that Algorithm 3 was derived from variational objective (9). □

Remark 3.6 Our manipulations hold equally for a general pre-smoother with error propagation operator $M \neq S^*$, simply by replacing $S^* \leftarrow M$. The error propagation operator M has no influence on the coarse-space and thus it does not have to be sparse.

4. Numerical example

To demonstrate the robustness of smoothed aggregation method, we consider the algorithm of [6] which is a modification of the method proposed and analyzed in [8] and [10]. Its relationship to Algorithm 2 is obvious. This method uses the smoothing iterative method $\mathcal{S}(\cdot,\cdot)$ which is a sequence of Richardson's iterations with carefully chosen iteration parameters. The error propagation operator S of the smoother $\mathcal{S}(\cdot,\cdot)$ is therefore a polynomial in the matrix A.

In this method, we use massive smoother S and a small coarse-space resulting in sparse coarse-level matrix.

Let $\bar{\lambda} \geq \varrho(A)$ and d be the desired degree of the smoothing polynomial S. We set

$$\alpha_i = \left[\frac{\bar{\lambda}}{2}\left(1 - \cos\frac{2i\pi}{2d+1}\right)\right]^{-1}, \quad i = 1,\dots,d, \tag{11}$$

$$S = (I - \alpha_1 A)\dots(I - \alpha_d A) \tag{12}$$

and

$$P = Sp.$$

Here, p is a *tentative prolongator* given by generalized unknowns aggregation. The simplest aggregation method is described in this section.

The smoother S is chosen to minimize $\varrho(S^2A)$. The reason for this comes from the fact that the convergence of the method of [6] is guided by the constant C in the weak approximation condition

$$\forall \mathbf{e} \in \mathbb{R}^n \; \exists \mathbf{v} \in \mathbb{R}^m \; : \; \|\mathbf{e} - p\mathbf{v}\| \leq \frac{C}{\sqrt{\varrho(S^2A)}}\|\mathbf{e}\|_A. \tag{13}$$

The smaller $\varrho(S^2A)$, the easier it becomes to satisfy (13) with a reasonable (sufficiently small) constant. It holds that ([6])

$$\bar{\lambda}_{S^2A} \equiv \frac{\bar{\lambda}}{(1+2d)^2} \geq \varrho(S^2A). \tag{14}$$

The aggregates $\{\mathcal{A}_j\}$ are sets of fine-level degrees of freedom that form a disjoint covering of the set of all fine-level degrees of freedom. For example, we can choose aggregates to form a decomposition of the set of degrees of freedom induced by a geometrically reasonable partitioning of the computational domain. For standard discretizations of scalar elliptic problems, the tentative prolongator matrix p is the $n \times m$ matrix ($m =$ the number of aggregates)

$$p_{ij} = \begin{cases} 1 & \text{if } i \in \mathcal{A}_j, \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

that is, the j-th column is created by restricting a vector of ones onto the j-th aggregate, with zeroes elsewhere. Thus, the aggregation method can be viewed as a piecewise constant coarsening in a discrete sense. The generalized aggregation method, suitable for non-scalar elliptic problems (like that of linear elasticity), is described in [15].

Algorithm 4 *Given the degree d of the smoothing polynomial $S = pol(A)$, the smoothed prolongator $P = Sp$ where p is the tentative prolongator and the prolongator smoother S is given by (12), the upper bound $\bar{\lambda} \geq \varrho(A)$ and a parameter $\omega \in (0,1)$, one iteration of the two-level algorithm*

$$\mathbf{x} \leftarrow TG(\mathbf{x}, \mathbf{f})$$

proceeds as follows:

1. *perform*
$$\mathbf{x} \leftarrow \mathbf{x} - \frac{\omega}{\bar{\lambda}_{S^2 A}} S^2 (A\mathbf{x} - \mathbf{f}),$$
where $\bar{\lambda}_{S^2 A}$ is given by (14) and S by (12),
2. *perform the iteration with symmetric error propagation operator S given by (12), that is,*
for $i = 1, \ldots, d$ do
$$\mathbf{x} \leftarrow (I - \alpha_i A)\,\mathbf{x} + \alpha_i \mathbf{f},$$
3. *evaluate the residual $\mathbf{d} = A\mathbf{x} - \mathbf{f}$,*
4. *restrict the residual $\mathbf{d}_2 = P^T \mathbf{d}$,*
5. *solve the coarse-level problem $A_2 \mathbf{v} = \mathbf{d}_2$, $A_2 = P^T A P$,*
6. *correct the approximation $\mathbf{x} \leftarrow \mathbf{x} - P\mathbf{v}$,*
7. *for $i = 1, \ldots, d$ do*
$$\mathbf{x} \leftarrow (I - \alpha_i A)\,\mathbf{x} + \alpha_i \mathbf{f},$$
8. *perform*
$$\mathbf{x} \leftarrow \mathbf{x} - \frac{\omega}{\bar{\lambda}_{S^2 A}} S^2 (A\mathbf{x} - \mathbf{f}).$$

512 000 dofs, coarse space 512 dofs, $\deg(S) = 7$, $H/h = 9$.		
ε	rate of conv. q_N	no. iter. N
1000	0.321	19
100	0.241	15
10	0.137	11
1	0.131	11
0.1	0.221	14
0.01	0.317	19
0.001	0.300	18

Table 1: 3D anisotropic problem

Thus, Algorithm 4 is a symmetrized version of Algorithm 2 with added smoothing in steps 1 and 8.

It is generally believed that in order to solve efficiently an anisotropic problem, one has to perform coarsening only by following *strong connections*. This technique is called *semi-coarsening*. In our case, we form aggregates by coarsening by a factor of 10 in all 3 spatial directions, which means, we do not perform semi-coarsening. Despite of this fact, our method gives satisfactory results regardless of the anisotropy coefficient ε. In this experiment, the symmetric Algorithm 4 is used as a conjugate gradient method preconditioner.

Test problem

- Problem:

$$-\left(\frac{\partial^2}{\partial x^2} + \varepsilon\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) u = f \text{ on } \Omega = (0,1)^3,\ u = 0 \text{ on } \partial\Omega. \tag{16}$$

- Mesh: $82 \times 82 \times 82$ regular square mesh, 512 000 unconstrained degrees of freedom.
- Aggregates: cubic groups of $10 \times 10 \times 10$ unconstrained vertices.
- Coarse-space size: 512 degrees of freedom.
- Degree of smoothing polynomial: 7.
- Stopping criterion: relative residual $< 10^{-9}$.

The results are summed up in Table 1. Note that here, the estimate of the rate of convergence after N iterations is defined as

$$q_N = \left(\|A\mathbf{x}^N - \mathbf{f}\| / \|A\mathbf{x}^0 - \mathbf{f}\|\right)^{\frac{1}{N}}.$$

Here, $\mathbf{x}^i$ denotes the i-th iteration.

Acknowledgements

This work was sponsored by the TAČR (Technologická Agentura České Republiky) grant TA01020352, ITI (Institut Teoretické Informatiky) grant 1R0545, Department of the Navy Grant N62909-11-1-7032.

References

[1] Bramble, J. H., Pasciak, J. E., Wang, J., and Xu, J.: Convergence estimates for multigrid algorithms without regularity assumptions. Math. Comp. **57** (1991).

[2] Xu, J. and Zikatanov, L. T.: The method of alternating projections and the method of subspace corrections in Hilbert space. J. Amer. Math. Soc. **15** (2002).

[3] Vassilevski, P. S.: *Multilevel Block Factorization Preconditioners.* Matrix-Based Analysis and Algorithms for Solving Finite Element Equations, Springer, New York, 2008.

[4] Brandt, A.: Algebraic multigrid theory: the symmetric case. Appl. Math. Comput. **19** (1986).

[5] Ciarlet, P. G.: *The finite element method for elliptic problems.* Series "Studies in Mathematics and its Applications", North-Holland, Amsterdam, 1978.

[6] Brousek, J., Fraňková, P., Kopincová, H., Kužel, R., Tezaur, R., Vaněk, P., and Vastl, Z.: *An overview of multilevel methods with aggressive coarsening and massive polynomial smoothing*, in preparation.

[7] Křížková, J. and Vaněk, P.: Two-level preconditioner with small coarse grid appropriate for unstructured meshes. Numer. Linear Algebra Appl. **3**(4) (1996).

[8] Guillard, H., Kužel, R., Vaněk, P., and Vastl, Z.: An alternative to domain decomposition methods based on polynomial smoothing. Numer. Linear Algebra Appl., submitted.

[9] Vaněk, P.: Smoothed prolongation multigrid with rapid coarsening and massive smoothing. To appear in Appl. Math.

[10] Vaněk, P., Brezina, M., and Tezaur, R.: Two-grid method for linear elasticity on unstructured meshes. SIAM J. Sci Comput. **21** (1999).

[11] Brezina, M., Heberton, C., Mandel, J., and Vaněk, P.: An iterative method with convergence rate chosen a priori. UCD/CCR Report no. 140, 1999.

[12] Vaněk, P., Mandel, J., and Brezina, M.: Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. Computing **56** (1996).

[13] Vaněk, P.: Acceleration of convergence of a two-level algorithm by smoothing transfer operator. Appl. Math. **37** (1992).

[14] Vaněk, P.: Fast multigrid solver. Appl. Math. **40**(1) (1995).

[15] Vaněk, P., Brezina, M., and Mandel, J.: Convergence of algebraic multigrid based on smoothed aggregations. Numer. Math. **88**(3) (2001).

[16] Vaněk, P., Brezina, M: Nearly optimal convergence result for multigrid with aggressive coarsening and polynomial smoothing. Appl. Math, to appear.

[17] Brezina, M., Vaněk, P., Vassilevski, P. S.: An improved convergence analysis of the smoothed aggregation algebraic multigrid. Numer. Linear Algebra Appl., to appear.