

UCRL-CR-141283
B505008

Algebraic Coarsening Methods for Linear and Nonlinear PDE and Systems

J.C. McWilliams

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

November 6, 2000

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Work performed under the auspices of the U. S. Department of Energy by the University of California Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Algebraic Coarsening Methods for Linear and Nonlinear PDE and Systems

Final Technical Report

PT: James C. McWilliams

Period of Time: 11/09/99 – 6/30/00

Agency Name: Lawrence Livermore National Laboratory

Award Number: #B505008

1 Summary of Accomplishments

The research summarized here was performed mainly by Irad Yavneh and carried out in a long-distance collaboration with Achi Brandt of the Weizmann Institute of Science, whose recent work inspired a large part of the present research.

1.1 Introduction

In [1] Brandt describes a general approach for algebraic coarsening. Given fine-grid equations and a prescribed relaxation method, an approach is presented for defining both the coarse-grid variables and the coarse-grid equations corresponding to these variables. Although, these two tasks are not necessarily related (and, indeed, are often performed independently and with distinct techniques), in the approaches of [1] both revolve around the same underlying observation. To determine whether a given set of coarse-grid variables is appropriate it is suggested that one should employ *compatible relaxation*. This is a generalization of so-called F-relaxation (e.g., [2]). Suppose that the coarse-grid variables are defined as a subset of the fine-grid variables. Then, F-relaxation simply means relaxing only the F-variables (i.e., fine-grid variables that do not correspond to coarse-grid variables), while leaving the remaining fine-grid variables (C-variables) unchanged. The generalization of compatible relaxation is in allowing the coarse-grid variables to be defined differently, say as linear combinations of fine-grid variables, or even nondeterministically (see examples in [1]). For the present summary it suffices to consider the simple case. The central observation regarding the set of coarse-grid variables is the following [1].

I

Observation 1 *A general measure for the quality of the set of coarse-grid variables is the convergence rate of compatible relaxation.*

The conclusion is that a necessary condition for efficient multigrid solution (e.g., with convergence rates independent of problem size) is that the compatible-relaxation convergence be bounded away from 1, independently of the number of variables. This is often a sufficient condition, provided that the coarse-grid equations are sufficiently accurate. Therefore, it is suggested in [1] that the convergence rate of compatible relaxation should be used as a criterion for choosing and

evaluating the set of coarse-grid variables.

Once a coarse grid is chosen for which compatible relaxation converges fast, it follows that the dependence of the coarse-grid variables on each other decays exponentially or faster with the distance between them, measured in mesh-sizes. This implies that highly accurate coarse-grid equations can be constructed locally. A method for doing this by solving local constrained-minimization problems is described in [1]. It is also shown how this approach can be applied to devise prolongation operators, which can be used for Galerkin coarsening in the usual way.

In the present research we studied and developed methods based, in part, on these ideas. We developed and implemented an AMG approach which employs compatible relaxation to define the prolongation operator (but is otherwise similar in its structure to classical AMG); we introduced a novel method for direct (i.e., non-Galerkin) algebraic coarsening, which is in the spirit of the approach originally proposed by Brandt in [1], but is more efficient and well-defined; we investigated an approach for treating systems of equations and other problems where there is no unambiguous correspondence between equations and unknowns.

1.2 AMG with prolongation derived by compatible relaxation

The prolongation operator represents the approximate dependence of the F-variables on C-variables. Various methods of defining the prolongation operator, based on the fine-grid matrix, are of course common and successful. Here we introduce a relatively simple and apparently effective approach, inspired by the ideas discussed in [1] and summarized above. For simplicity of this informal discussion, we consider a 2D discretized PDE, with the coarse-grid points corresponding to fine-grid points both of whose indices are even (standard geometric coarsening). The prolongation operator in the present method is constructed as described below. We denote by $p(I, J, k, C)$ the prolongation coefficient by which coarse-grid variable $U(I, J)$ is multiplied when interpolated to fine-grid point (\mathbf{k}, \mathbf{l}) .

Tentative Prolongation Operator Let Q denote the fine-grid domain.

For each coarse-grid point,

$$(I, J), (i, j) \in C(2)$$

we choose a local fine-grid subdomain, $F \subset Q$, which includes in particular the point (i, j) and all other fine-grid points belonging to the prescribed stencil of coarse-grid point (I, J) . Next we set $u(i, j) = 1$, and $u(k, C) = 0$ for all $(k, C) \notin (i, j)$ in F . We then perform t_4 compatible Gauss-Seidel relaxation sweeps (with zero right-hand side) in F , obtaining $u(i, j)$. Now, for all $(k, C) \in C(2)$ which belong to the prescribed prolongation stencil of point (I, J) , we define a tentative prolongation operator by

$$p^t(I, J, k, C) = u(i, j)$$

For large F , $u(i, j)$ yields the correct contribution of $U(I, J)$ to point (i, j) . However, since we limit severely the size of the prolongation stencil, and hence must ignore the contribution of all but a very small number of nearby coarse-grid points, we need to normalize $p^t(I, J, k, C)$ in order to compensate for the neglected (many, but exponentially small) contributions.

Normalization For all $(i, j) \in C(2)$ we initialize $u(i, j) = 1$, and perform ∞ complete compatible Jacobi

sweeps, obtaining $u^{(glob)}$. The prolongation is given by

$$p(I, J, k, li) = \frac{u^{(Plob)}(k, C)}{Zijpt(I, j, r, s)} p'(I, J, k, L)$$

This normalization is suitable when all the prolongation coefficients are positive. *e.g.*, when the fine-grid problem is an M matrix. If there are both positive and negative tentative prolongation coefficients, they need to be normalized separately, as discussed in [2]. For restriction we use the transpose of the prolongation.

Remarks on Computational Cost In numerical tests we found that quite small values (usually 2) for v_g and v_r were required to obtain essentially optimal efficiency. We conclude that the method seems comparable to classical AMG in its cost of implementation.

Numerical Tests Several numerical tests were performed with this solver. First, we compared convergence rates of V(1,1) cycles, employing pointwise Red-Black relaxation, against the classical Black-Box multigrid method of Dendy (1982). “Geometric” coarsening was used in this comparison. The problem solved was a diffusion equation with random coefficients between zero and one on a domain with randomly distributed “holes”, where Dirichlet boundary conditions were prescribed. The present solver proved to be robust throughout the parameter range tested, even where the Black-Box solver stagnated. We also implemented the solver in a purely algebraic setting, suitable for unstructured grids, employing AMG coarsening. Again, the solver seemed robust, performing similarly to AMG, but the tests have so far been limited to fairly standard problems which standard AMG also handles easily.

Conclusions: The new AMG algorithm seems comparable to standard AMG for “standard” problems, though a detailed comparison of computational cost has not been carried out. The new approach offers the potential of generalization to problems where the coarse-grid variables are not (or cannot be) defined as subsets of the fine-grid variables. This has not yet been implemented. Also, here, unlike “standard” AMG, we do not require a direct dependence of every F-variable on some C-variable, and the basic algorithm applies when this is not the case. It would be interesting to combine this algorithm with a coarsening based also on compatible relaxation, as suggested in [1]. This has been implemented only for some very basic cases, and we do not yet know the optimal implementation of this idea.

1.3 Directly derived coarse-grid equations

In [1] both direct and Galerkin coarsening are considered. We sketch here briefly our variant of the direct approach, which was developed as part of this research. It is described in greater generality in Section 4.1 of [1].

To determine the coarse-grid equation for coarse-grid variable I , first choose some local fine-grid subdomain, F . Then take a linear combination of the fine-grid equations corresponding to variables belonging to F . This linear combination is itself obviously an exact fine-grid equation. The coarse-grid equation is obtained by ignoring all coefficients corresponding to F -variables. The idea is to choose the coefficients of the linear combination such that the neglected coefficients will be minimal (in some norm), and that the coefficient of the C -variable corresponding to I will be relatively large

so that there will remain an unambiguous correspondence between equations and variables on the coarse grid.

The coefficients of the linear combination define a restriction operator. The neglected coefficients of the resulting stencil represent the discrete relative truncation error. We minimize the weighted L_2 norm of the latter, subject to certain constraints. For example, we normally impose that the sum of the neglected coefficients equal zero, which ensures that the coarse-grid approximation of constant functions is exact (since the relative truncation error vanishes in this case). Alternative constraints are also possible, for example, imposing that all the neglected coefficients be of the same sign, resulting in a coarse-grid operator which satisfies some fine-grid maximum principle. This can be used to obtain monotonic convergence, for example.

In [1] some examples are given using this approach for direct determination of the coarse-grid equations. We can also use this approach to define a prolongation, which can be used in the framework of Galerkin coarsening. For each F-variable, i , we take a local linear combination of the fine-grid equations (with zero right-hand sides), and then neglect all coefficients that correspond to F-variables, except variable i . We minimize the L_2 norm of the neglected coefficients subject to the zero-sum constraint and also the constraint that the coefficient of equation i is 1. This gives us an equation for variable i in terms of its neighboring C-variables (which are included in the prescribed prolongation stencil.) We implemented this for the problem of Section 1.2. The convergence rate was satisfactory, but this approach is significantly more expensive than that of Section 1.2. It remains to be seen whether it is advantageous in more general problems.

1.4 AMG for PDE systems

The success of algebraic multigrid methods for systems (more generally, to problems where there is no obvious correspondence between variables and equations) has thus far been rather limited. This is clearly an important goal. Some "grey" approaches are discussed in [1], and some very significant advances are alluded to in [2]. Here we describe an approach that we have pursued in this research. This is work in progress, and useful results have only been achieved so far for very simple cases.

The classical AMG approach assumes the existence of an unambiguous correlation between equations and unknowns. The coarse-grid definition process relies on this relation. Also, ordinary Gauss Seidel relaxation is often a useful smoother in this situation. But in many cases, in particular in problems resulting from the discretization of systems of PDE, no such relation exists. Some common examples are the Cauchy-Reimann equations and the Stokes system.

We propose a pre-processing step whose purpose is to transform the system into one for which some fairly standard AMG would be efficient. Specifically, given some set of equations,

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

we compute some "distribution matrix", M , define new variables.

$$u=Mv, \tag{2}$$

and solve the problems in terms of the new variables.

$$L\tilde{u}=\tilde{f} \tag{3}$$

where $L = Li^{-1}$. Of course, M needs to be chosen such that L can be handled efficiently by AMG. (Using the AMG approaches discussed here, we need not necessarily compute L explicitly, but can instead use M as in the classical distributive relaxation, and the distributive coarsening described in

[1].) Alternatively, we may compute some preconditioning or “weighting matrix” W and solve the problem

$$LWu = Wf, \quad (4)$$

Again, L must be such that (4) can be handled efficiently by AMG.

For simplicity of the discussion, consider the second case, where $L = WL$. This can be thought of as replacing each equation, i , by a linear combination of equations, with weights given by the elements of W_i —the i th row of W . Clearly, for this approach to work L must have a substantial diagonal element. A natural approach then is to choose H so as to obtain L that is “as diagonally dominant as possible”. For example, we can minimize $\sum w_i |L_{ii} - f_i|^2$, subject to $L_{ii} \geq 1$, for some prescribed weights w_i . Since this is once again a series of local constrained minimization problems, the kinship of this approach to the general theme we are following seems evident. But will it work?

Clearly, we need to impose some restrictions on (4). In particular, W needs to be quite sparse so that L will remain sparse. Also, we need to rescale the equations and variables (or use appropriate weights) to make the coefficients corresponding to different equations (and different variables) comparable in size.

Discussion and preliminary results A central decision is how to choose the stencils associated with the rows of W , i.e., which equations should participate in the linear combination that makes up each equation in the transformed system, (4). Then, what criterion should be used to determine the stencil (i.e., nonzero elements) of L . One possibility is to use for row i the same stencil as the i th column of L . That is, to form row i of L , take a linear combination of all the equations which depend on variable i . With this choice, one possibility is simply to choose $L = W^T W$. Then, L is obviously a symmetric positive definite matrix, and there is a clear correspondence between equations and unknowns. But this is often far from optimal. (One exception is the Cauchy Reimann equations, where this is clearly an excellent choice.)

If we adopt the approach of solving local optimization problems to determine L , it makes sense not to include in the stencil associated with row i , all the equations where variable i appears, but rather only the equations where it is a dominant variable. This means that L_{ij} will be nonzero if L_{ii} is (nearly) the largest element (in absolute value) either in the j th row or in the i th column of L . The motivation is that if L_{ii} is relatively small then the inclusion of equation j in the linear combination is likely to cause more harm than good by increasing the number of relatively large off-diagonal terms. Also, this criterion ensures that every equation is included in at least one linear combination which is essential for making the transformed system equivalent to the original one. Note, however, that for L that dominates column i there might still exist some variable k for which $L_{kj} > L_{ii}$. In this case, we may also include in the stencil of W_i all equations in which variable k is dominant.

Thus far our experience has been that using this approach and applying AMG straightforwardly is unsatisfactory (except for simple problems such as the Cauchy-Reimann equations). For the Stokes system, for example, the coarsening process mixes up the different types of variables, and stencils are obtained which are not handled well by the standard AMG. We therefore examined a “grey-box” version, where we attach a “color” to each type of variable on the fine grid. Then, in the definition of the coarse grid, we apply the usual AMG coarsening process to each color separately (ignoring dependences on variables of different colors for this process.) This approach seems to hold some promise. Indeed, for the Stokes equations with staggered discretization and periodic boundary conditions, this approach can produce an exact decoupling of the equations (provided that the weights, w_i , also depend on the variable color, and they are chosen so as to promote a block-

triangular L). and the resulting system is solved essentially as fast as the Poisson problem. However, much work is yet needed before a reasonably general “grev-box”) solver for elliptic systems is achieved. We are currently testing and refining this approach for the Stokes system with boundaries and for linearized Navier-Stokes.

1.5 Concluding remarks

Some new algebraic coarsening methods have been studied, based on ideas related to the condition of fast convergence of compatible relaxation. A generalized (but semi-standard) AMG approach, where the prolongation operator is derived using compatible relaxation, has been formulated and tested with encouraging results. New methods of direct derivation (i.e., non-Galerkin) have also been introduced and preliminary tests have been made. Such methods may not be computationally competitive for relatively simple problems, but they can be made quite general. In particular they may be useful for creating re-usable coarse-grid equations, homogenization, etc.

An approach for applying AMG to systems of PDE has been introduced, whereby a pre-processing step is aimed at producing a system that can be solved by a fairly straightforward AMG solver. Some early tests show promise, but the details require significant further investigation and tests. Further details will be reported as this research continues to develop.

References

- [1] Achi Brandt, “General Highly Accurate Algebraic Coarsening”, *to appear in ETNA*.
- [2] Klaus Stueben, “Algebraic Multigrid (AMG): An Introduction with Applications”. GMD Report 53 (1999).