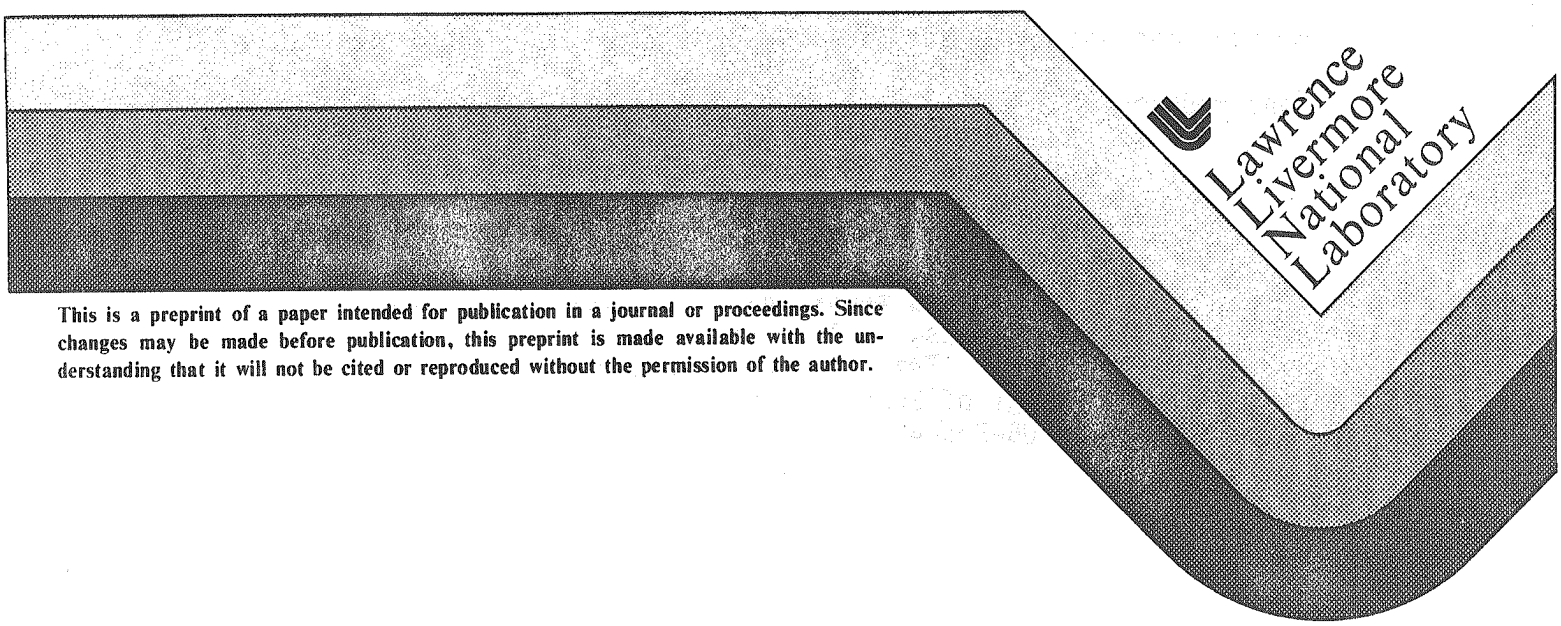
UCRL-87912
PREPRINT

Analysis of a Multigrid Method as an Iterative Technique for Solving Linear Systems

Anne Greenbaum

This paper was prepared for presentation at
1982 DOE NASIG Meeting
June 23-25, 1982
Los Alamos, N. M.

June 18, 1982



Analysis of a Multigrid Method as an Iterative Technique for Solving Linear Systems*

ABSTRACT

A general class of iterative methods is introduced for solving symmetric, positive definite linear systems. These methods use two different approximations to the inverse of the matrix of the problem, one of which involves the inverse of a smaller matrix. It is shown that the methods of this class reduce the error by a constant factor at each step and that under "ideal" circumstances this constant is equal to $\frac{\kappa'-1}{\kappa'+1}$, where κ' is the ratio of the largest eigenvalue to the $(J+1)^{st}$ eigenvalue of the matrix, J being the dimension of the smaller matrix involved. A multigrid method is presented as an example of a method of this class, and it is shown that while the multigrid method does not quite achieve this optimal rate of convergence, it does reduce the error at each step by a constant factor independent of the mesh spacing h.

*The author would like to thank the U. S. Dept. of Energy, Office of Basic Energy Sciences, Applied Mathematical Sciences Division for support of this work. This work was performed under the auspices of the U. S. Department of Energy by Lawrence Livermore Laboratory under Contract No. W-7405-Eng-48.

1. Introduction

Many iterative methods for solving a linear system, $Ax = b$, can be written in the general form

$$x^k = x^{k-1} + M^{-1}(b - Ax^{k-1}) , \quad k = 1,2,\ldots, \tag{1}$$

where x^0 is a given initial guess to the solution and each iterate x^k represents a hopefully improved approximation to the solution. The matrix M^{-1} is taken to be an approximation to A^{-1} such that the product of M^{-1} times an arbitrary vector is easy to compute. For example, taking M^{-1} to be $(\text{diag}(A))^{-1}$ gives Jacobi's method, and taking M^{-1} to be (lower triangle of $A)^{-1}$ gives the Gauss-Seidel method. The error $e^k \equiv A^{-1}b - x^k$ is then given by

$$e^k = (I-M^{-1}A)\, e^{k-1} , \tag{2}$$

and the norm of the error at the k^{th} step satisfies

$$||e^k|| \leq ||I-M^{-1}A|| \cdot ||e^{k-1}|| .$$

The error is reduced quickly if the norm of the matrix $I-M^{-1}A$ is much less than one.

Multigrid methods [cf. 1,2,3] can be written in the general form (1), where the iterates x^k represent quantities generated after a "coarse grid correction cycle" and a given number of "relaxation sweeps". That is, given an approximation x^{k-1} to the solution, the multigrid algorithm generates a new approximation $x^{k-1,0}$ via a formula of the form

$$x^{k-1,0} = x^{k-1} + B(b - Ax^{k-1}) \ , \tag{3}$$

where the matrix B represents a "coarse grid" approximation to A^{-1}. The method then generates a fixed number, say ℓ, of new approximations, $x^{k-1,j}$, $j=1,\ldots,\ell$, by performing "relaxation sweeps" of the form

$$x^{k-1,j} = x^{k-1,j-1} + C(b - Ax^{k-1,j-1}) \ , \quad j = 1, \ldots, \ell, \tag{4}$$

where the matrix C also represents an approximation to A^{-1}. If we denote by x^k the quantity $x^{k-1,\ell}$ of (4), then we find

$$x^k = x^{k-1} + [(I - (I-CA)^{\ell}(I-BA))A^{-1}] (b - Ax^{k-1}) \ , \tag{5}$$

and the error e^k satisfies

$$e^k = (I-CA)^{\ell}(I-BA)e^{k-1}. \tag{6}$$

Thus, for multigrid methods the iteration matrix, $I-M^{-1}A$ in (2), is of the special form $(I-CA)^{\ell}(I-BA)$, for certain matrices B and C. As with other iterative methods, the amount by which the error is reduced at a given step depends on the norm of this iteration matrix.

In the following section we analyze iterations of the form (3)-(4), where the matrix B is also of a special form, involving the inverse of a smaller matrix. In section 3 we examine a multigrid method (or, more precisely, a two-grid method) as an example of such an iteration and determine how close it comes to achieving a certain "optimal" convergence rate for methods of this class.

2. Analysis of a General Class of Iterative Methods

We will consider only problems in which the matrix A is symmetric and positive definite, and we will estimate the rate at which the A-norm of the error, $||e^k||_A \equiv \langle e^k, Ae^k \rangle^{1/2}$, is reduced. To avoid working directly with A-norms, we can define a modified error vector, $s^k \equiv A^{1/2}e^k$, and consider the rate at which the Euclidean norm of s^k (which is the A-norm of e^k) is reduced. From equation (6) we have

$$(7) \qquad s^k = (I-A^{1/2}CA^{1/2})^{\ell}(I-A^{1/2}BA^{1/2})\, s^{k-1} ,$$

and hence the Euclidean norm of s^k (denoted by $||\cdot||$) satisfies

$$(8) \qquad ||s^k|| \leq ||(I-A^{1/2}CA^{1/2})^{\ell}(I-A^{1/2}BA^{1/2})||\ ||s^{k-1}|| .$$

The quantity $||(I-A^{1/2}CA^{1/2})^{\ell}(I-A^{1/2}BA^{1/2})||$ in (8) is called the contraction number of the method and will be denoted by κ.

A simple bound for κ is given by

$$(9) \qquad \kappa \leq ||I-A^{1/2}CA^{1/2}||^{\ell} \cdot ||I-A^{1/2}BA^{1/2}|| .$$

Thus, if the matrices B and C are chosen so that $||I-A^{1/2}CA^{1/2}||$ and $||I-A^{1/2}BA^{1/2}||$ are both less than or equal to one, with one of these

norms strictly less than one, then the iteration (3)-(4) will converge to the solution and, moreover, the A-norm of the error will be reduced at each step. The class of iterative methods to be considered in this section employ such matrices B and C and hence are convergent whenever the matrix A is symmetric and positive definite.

Inequality (9), however, is too crude an estimate to provide much useful information about the rate of convergence. In fact, the methods to be considered are designed to use matrices B and C which complement each other in such a way that the norm of the matrix product in (8) is much less than the product of the norms in (9). Instead of inequality (9), we estimate ρ with the following bound:

$$(10) \qquad \rho \leq \max_{\substack{||y||=1 \\ y \in \text{range}(I-A^{1/2}BA^{1/2})}} ||(I-A^{1/2}CA^{1/2})^{\ell} y|| \cdot ||I-A^{1/2}BA^{1/2}|| .$$

If the range of $I-A^{1/2}BA^{1/2}$ is a restricted set of vectors on which the operator $(I-A^{1/2}CA^{1/2})^{\ell}$ is highly contractive, then the bound in (10) may be much smaller than that in (9).

We now define the form of the matrix B in iteration (3). Suppose A is an N by N matrix and J is some number less than N. Let P be some arbitrary N by J matrix and define a J by N matrix Q by

$$(11) \qquad Q = P^T .$$

Next define a J by J matrix A_J by

$$(12) \qquad A_J = Q\,A\,P ,$$

and take B to be the matrix

$$(13) \qquad B = P\,A_J^{-1}\,Q .$$

The hope is that the matrix P can be defined in such a way that the smaller matrix A_J is easier to invert than A and, in some sense, approximates A. We might then expect that the N by N matrix B would be a good approximation to A^{-1}. Without specifically defining the matrix P, we can use the relations (11)-(13) to determine the norm and range of the matrix $I-A^{1/2}BA^{1/2}$ in (10). (We are assuming, of course, that P is chosen in such a way that A_J is invertible.)

Theorem. If the matrix B is defined by (11)-(13), then we have

$$A^{1/2} \cdot \mathcal{R}(I-A^{1/2}BA^{1/2}) \subset n(Q) \ , \tag{14}$$

where $\mathcal{R}(\cdot)$ denotes the range and $n(\cdot)$ the null space of an operator. The norm of the matrix $I-A^{1/2}BA^{1/2}$ satisfies

$$||I-A^{1/2}BA^{1/2}|| = 1 \ . \tag{15}$$

Proof: Let y be a vector in the range of $I-A^{1/2}BA^{1/2}$. Then there is a vector z such that

$$y = (I-A^{1/2}BA^{1/2})\ z \ .$$

Multiplying by $QA^{1/2}$ and using the definition (13) of B gives

$$QA^{1/2}y = (Q - QAPA_J^{-1}Q)\ A^{1/2}z \ .$$

But from (12), QAP is just A_J, and so the right-hand side of this equation is zero, thus proving (14).

To establish (15), we note that $I-A^{1/2}BA^{1/2}$ is a symmetric matrix and hence its norm is the absolute value of its largest eigenvalue. If z is an eigenvector of $I-A^{1/2}BA^{1/2}$ with eigenvalue λ, then we can write

$$QA^{1/2}(I-A^{1/2}BA^{1/2})z = \lambda QA^{1/2}z = 0 \ .$$

It follows that either $\lambda = 0$ or $A^{1/2}z \in n(Q)$. In the latter case, $BA^{1/2}z$ is zero, and hence we have $\lambda = 1$. Thus, the eigenvalues of $I-A^{1/2}BA^{1/2}$ are 0 and 1, and (15) is proved.

Applying the theorem, inequality (10) becomes

$$(16) \qquad \rho \le \max_{\substack{\|y\|=1 \\ A^{1/2}y \in n(Q)}} \|(I-A^{1/2}CA^{1/2})^{\ell} y\| .$$

Thus, in designing a method of the form (3)-(4), (11)-(13), we would like to choose matrices Q and C for which the bound in (16) is small. For the purpose of analysis, we will assume that the matrix C is chosen to be <u>symmetric</u> (though some multigrid methods use nonsymmetric matrices C and hence will not be covered by our general theory without some modification). Let $d_1 \le \ldots \le d_N$ be the eigenvalues of $A^{1/2}CA^{1/2}$, with corresponding normalized eigenvectors $v_1, \ldots, v_N$. Let y be any vector with norm one. We can expand y and $(I-A^{1/2}CA^{1/2})^{\ell}y$ in terms of $v_1, \ldots, v_N$ as follows:

$$(17) \qquad y = \sum_{i=1}^{N} \langle y,v_i \rangle v_i , \qquad \sum_{i=1}^{N} \langle y,v_i \rangle^2 = 1$$

$$(18) \qquad (I-A^{1/2}CA^{1/2})^{\ell} y = \sum_{i=1}^{N} \langle y,v_i \rangle (1-d_i)^{\ell} v_i$$

$$(19) \qquad \|(I-A^{1/2}CA^{1/2})^{\ell} y\|^2 = \sum_{i=1}^{N} \langle y,v_i \rangle^2 (1-d_i)^{2\ell} .$$

Bounding the quantity in (19), we have

$$(20) \qquad \max_{\|y\|=1} \|(I-A^{1/2}CA^{1/2})^{\ell} y\|^2 \le \max \{(1-d_1)^{2\ell}, (1-d_N)^{2\ell}\} .$$

Now suppose y satisfies the additional constraint $A^{1/2}y \in n(Q)$, or, equivalently, $A^{1/2}y \perp R(P)$. If P and C have been chosen in such a way that some of the eigenvectors of $A^{1/2}CA^{1/2}$, say, the first J eigenvectors, are of the form

$$(21) \qquad v_i = A^{1/2}Pz_i \ , \qquad i=1,\ldots,J \ ,$$

for some vectors z_i, then y has no components in the directions of these eigenvectors. We can say that these components were annihilated by the partial step (3) of the iteration. Expression (19) for $||(I-A^{1/2}CA^{1/2})^{\ell}y||^2$ then becomes

$$(22) \qquad ||(I-A^{1/2}CA^{1/2})^{\ell}\ y||^2 = \sum_{i=J+1}^{N} <y,v_i>^2\ (1-d_i)^{2\ell} \ ,$$

and the bound in (20) is replaced by

$$(23) \qquad \max_{\substack{||y||=1 \\ A^{1/2}y \in n(Q)}} ||(I-A^{1/2}CA^{1/2})^{\ell}\ y||^2 \leq \max\ \{(1-d_{J+1})^{2\ell},\ (1-d_N)^{2\ell}\} \ .$$

Thus, under the "ideal" conditions (21), the norm of the iteration matrix is effectively reduced from expression (20) -- which it would have been had we iterated using only steps of the form (4) -- to the possibly much smaller value in (23).

As an example, suppose C is taken to be of the form

$$C = \gamma I \ , \tag{24}$$

where the constant γ is chosen in some optimal or near optimal way. Then the eigenvectors of $A^{1/2}CA^{1/2}$ are just the eigenvectors of A. If $\lambda_1 \leq \ldots \leq \lambda_N$ are the eigenvalues of A, then the bound (20) becomes

$$\max_{||y||=1} ||(I-\gamma A)^{\ell} y||^2 \leq \max \{(1-\gamma\lambda_1)^{2\ell}, (1-\gamma\lambda_N)^{2\ell}\} \ . \tag{25}$$

To minimize this bound we take

$$\gamma = \frac{2}{\lambda_1+\lambda_N} \ ,$$

and (25) becomes

$$\max_{||y||=1} ||(I-\gamma A)^{\ell} y||^2 \leq (\frac{\lambda_N-\lambda_1}{\lambda_N+\lambda_1})^{2\ell} = (\frac{\kappa-1}{\kappa+1})^{2\ell} \ , \quad \kappa = \frac{\lambda_N}{\lambda_1} \ . \tag{26}$$

This is the usual bound for the method of steepest descent.

On the other hand, if the conditions (21) are satisfied, then the improved bound (23) holds, and for our example (23) becomes

$$(27)\qquad \max_{\substack{||y||=1 \\ A^{1/2}y\,\varepsilon\, n(Q)}} ||(I-\gamma A)^{\ell}\, y||^2 \le \max\{(1-\gamma\lambda_{J+1})^{2\ell},\ (1-\gamma\lambda_N)^{2\ell}\}\ .$$

To minimize this bound we take

$$\gamma = \frac{2}{\lambda_{J+1}+\lambda_N}\ ,$$

and (27) becomes

$$(28)\qquad \max_{\substack{||y||=1 \\ A^{1/2}y\,\varepsilon\, n(Q)}} ||(I-\gamma A)^{\ell}\, y||^2 \le \left(\frac{\lambda_N-\lambda_{J+1}}{\lambda_N+\lambda_{J+1}}\right)^{2\ell} = \left(\frac{\kappa'-1}{\kappa'+1}\right)^{2\ell}\ ,\qquad \kappa' = \frac{\lambda_N}{\lambda_{J+1}}\ .$$

Thus, for our example, the effective condition number of the iteration matrix has been reduced from $\kappa = \lambda_N/\lambda_1$ to $\kappa' = \lambda_N/\lambda_{J+1}$. If the latter ratio is much smaller, as is typically the case when the matrix A approximates a differential operator, then much faster convergence is achieved by using a partial step of the form (3) than by iterating only with steps of the form (4). A goal, then, in designing a method of the form described, with C of the form (24), is to choose a matrix P for which (21) is approximately satisfied. In the following section we define the matrix P according to a multigrid prescription and decide how nearly the ideal conditions (21) are approximated.

3. Analysis of a Specific Multigrid Method

Suppose the matrix A arises from a finite element approximation to a self-adjoint elliptic boundary value problem on a region Ω using a grid with N free nodes and maximum mesh spacing h. Assume that the differential operator L is of order 2m and that the finite element space S^h contains all admissible polynomials of degree r-1 or less, where $r \geq 2m$. For any N-vector v, let $v(x)$ denote the piecewise polynomial in S^h that is represented by the vector v. Assume that a coarser J-point grid (with elements of the same shape; e.g., triangles, rectangles, etc.) can be formed by deleting some nodes of the N-point grid and that the maximum mesh spacing on the J-point grid is ρh, $\rho \geq 1$. Again, for any J-vector w, let $w(x)$ denote the piecewise polynomial in the corresponding finite element space $S^{\rho h}$ for the coarse grid, that is represented by the vector w.

Define the coarse to fine interpolation matrix P of section 2, so that functions in $S^{\rho h}$ map into themselves in S^h; that is, so that the piecewise polynomial represented by any J-vector w is the same as that represented by the N-vector Pw:

$$w(x) = (Pw)(x) , \qquad \forall\, w \in R^J. \tag{29}$$

With these definitions, it is shown in [4] that the solution to a differential equation

$$(30) \qquad Lu(x) = f(x) , \qquad x \in \Omega$$

with the given boundary conditions, is related to the finite element approximation $z(x)$ in S^h by

$$(31) \qquad a(u(x)-z(x),u(x)-z(x))^{1/2} \leq c_1 \, h^m \, ||u(x)||_{2m} ,$$

assuming that $u(x)$ lies in $H^{2m}(\Omega)$, the $(2m)^{th}$ Sobolev space with norm $||\cdot||_{2m}$. Here c_1 is some constant independent of h and $a(\cdot,\cdot)$ is the energy inner product as defined in [4].

The differential equation (30) with the given boundary conditions is equivalent to the minimization problem: Find $u(x) \in H_E^m$ (the set of functions in the Sobolev space $H^m(\Omega)$ which satisfy the essential boundary conditions) to minimize

$$a(u(x),u(x)) - 2(u(x),\hat{f}(x)) .$$

The function $\hat{f}(x)$ is equal to $f(x)$ if the boundary conditions are homogeneous and is modified by an appropriate boundary term for nonhomogeneous boundary conditions. (See [4], sections 1.7-1.8.)

With appropriate smoothness conditions on the coefficients of the differential operator L, and assuming that $\hat{f}(x)$ lies in the Sobolev space $H^0(\Omega)$, it is further shown that

$$(32) \qquad ||u(x)||_{2m} \leq c_2 \, ||\hat{f}(x)||_0 \, , \qquad c_2 \text{ independent of } h \, .$$

We will make the additional assumption that if $\hat{f}(x)$ lies in the finite element space S^h, so that the finite element system for equation (30) is

$$(33) \qquad Az = \hat{f} \, ,$$

then the H^0-norm of the function $\hat{f}(x)$ is related to the Euclidean norm of the vector $\hat{f}$ by

$$(34) \qquad ||\hat{f}(x)||_0 \leq c_3 \, ||\hat{f}|| \, , \qquad c_3 \text{ independent of } h \, .$$

Combining inequalities (31)-(34) we have

$$(35) \qquad a(u(x)-z(x),u(x)-z(x))^{1/2} \leq c \, h^m \, ||Az|| \, ,$$

$$c = c_1 c_2 c_3 \text{ independent of } h.$$

Similarly, using these same assumptions, it can be shown that if $w(x)$ represents the finite element approximation in $S^{\rho h}$ for the differential equation (30), then $w(x)$ is related to $u(x)$ by

$$(36) \qquad a(u(x)-w(x),u(x)-w(x))^{1/2} \leq c\ \rho^m\ h^m\ ||Az|| \ .$$

We will use the relations (35) and (36) to determine how well the desired conditions (21) of section 2 are approximated, assuming that C is of the form (24). Again, letting $v_1,\ldots,v_N$ denote the eigenvectors of A, with $\lambda_1 \leq \ldots \leq \lambda_N$ the corresponding eigenvalues, we have the following theorem:

Theorem 3.1. If v is any vector in $\text{span}[v_1,\ldots,v_j]$, $j \leq N$, and if $||v|| \leq 1$, then v can be written in the form

$$(37) \qquad v = A^{1/2}Pw + \delta \ , \qquad \text{where} \quad ||\delta|| \leq c\ (1+\rho^m)\ h^m\ \lambda_j^{1/2} \ ,$$

for some J-vector w.

Proof: Let $z = A^{-1/2}v$, and consider the minimization problem: Find u(x) in H_E^m to minimize

$$a(u(x),u(x)) - 2(u(x),(A^{1/2}v)(x)) \ .$$

The finite element system for this problem is

$$Az = A^{1/2}v \ ,$$

and so from (35), z(x) is related to u(x) by

$$(38) \qquad a(u(x)-z(x),u(x)-z(x))^{1/2} \leq c\ h^m\ ||A^{1/2}v|| \leq c\ h^m\ \lambda_j^{1/2} \ .$$

Similarly, if w(x) is the finite element approximation in $S^{\rho h}$ for this same minimization problem, then from (36) w(x) is related to u(x) by

$$(39) \qquad a(u(x)-w(x),u(x)-w(x))^{1/2} \leq c\ \rho^m\ h^m\ \lambda_j^{1/2}\ .$$

Combining (38) and (39) using the triangle inequality, we find that $z(x)$ is related to $w(x)$ by

$$(40) \qquad a(z(x)-w(x),z(x)-w(x))^{1/2} \leq c\ (1+\rho^m)\ h^m\ \lambda_j^{1/2}\ .$$

Now $w(x)$ is the same as $Pw(x)$, and from the definition of the finite element matrix A it can be shown that for any function $z(x)$ in S^h, $a(z(x),z(x))^{1/2} = ||z||_A$. Thus (40) becomes

$$||z - Pw||_A = ||v - A^{1/2}Pw|| \leq c\ (1+\rho^m)\ h^m\ \lambda_j^{1/2}\ ,$$

which is the desired result. QED.

We now use theorem 3.1 to bound the quantity on the right-hand side of (16), again assuming that the matrix C is of the form (24). We can write inequality (16) in the form

$$(41) \qquad \kappa^2 \leq \max_{\substack{||y||=1 \\ A^{1/2}y\ \varepsilon\ n(Q)}} \left(\sum_{i=1}^{j} <y,v_i>^2 (1-\gamma\lambda_i)^{2\ell} + \sum_{i=j+1}^{N} <y,v_i>^2 (1-\gamma\lambda_i)^{2\ell} \right) .$$

We will try to choose the number j so that the first sum in this expression is small, and we will then choose γ to approximately minimize the second sum. Taking γ to be

$$\gamma = \frac{2}{\lambda_N + \lambda_{j+1}} , \tag{42}$$

we can write

$$\kappa^2 \leq \max_{\substack{||y||=1 \\ A^{1/2}y \,\varepsilon\, n(Q)}} \left[\sum_{i=1}^{j} \langle y, v_i \rangle^2 + \left(\frac{\kappa'-1}{\kappa'+1}\right)^{2\ell} \left(1 - \sum_{i=1}^{j} \langle y, v_i \rangle^2\right) \right] , \tag{43}$$

$$\kappa' = \frac{\lambda_N}{\lambda_{j+1}} .$$

Theorem 3.2. If $A^{1/2}y \,\varepsilon\, n(Q)$ and $||y|| = 1$, then we have

$$\sum_{i=1}^{j} \langle y, v_i \rangle^2 \leq c\, (1+\rho^m)\, h^m\, \lambda_j^{1/2} . \tag{44}$$

Proof: The left-hand side of expression (44) is the square of the norm of the vector

$$v = \sum_{i=1}^{j} \langle y, v_i \rangle\, v_i ,$$

and according to theorem 3.1, v can be written in the form

$$v = A^{1/2} P w + \delta , \qquad ||\delta|| \leq c\, (1+\rho^m)\, h^m\, \lambda_j^{1/2} .$$

The condition $A^{1/2}y \,\varepsilon\, n(Q)$ implies that y is orthogonal to $A^{1/2}Pw$, and so we have

$$\langle y, v \rangle = \langle y, \delta \rangle .$$

Since $\langle y, v \rangle = \langle v, v \rangle$ we have

$$||v||^2 \leq ||\delta|| ,$$

from which the theorem follows. QED.

With theorem 3.2, then, assuming that the right-hand side of (44) is less than or equal to one, (43) becomes

$$(45) \qquad \rho^2 \le c\,(1+\rho^m)\,h^m\,\lambda_j^{1/2} + \left(\frac{\kappa'-1}{\kappa'+1}\right)^{2\ell}\left[1 - c\,(1+\rho^m)\,h^m\,\lambda_j^{1/2}\right], \qquad \kappa' = \frac{\lambda_N}{\lambda_{j+1}}.$$

Since the matrix A approximates a differential operator of order 2m, it can be shown that the largest eigenvalue of A is of order h^{-2m}, say,

$$\lambda_N = \alpha\, h^{-2m}, \qquad \alpha \text{ independent of } h.$$

Let $\beta > 0$ be some number less than α and independent of h, and choose j so that λ_j is the largest eigenvalue less than or equal to βh^{-2m}:

$$\lambda_j \le \beta\, h^{-2m}, \qquad \lambda_{j+1} > \beta\, h^{-2m}.$$

Then expression (45) becomes

$$(46) \qquad \rho^2 \le c\,(1+\rho^m)\,\beta^{1/2} + \left(\frac{\kappa'-1}{\kappa'+1}\right)^{2\ell}\left[1 - c\,(1+\rho^m)\,\beta^{1/2}\right], \quad \kappa' = \frac{\alpha}{\beta},$$

provided β is chosen so that the first term in (46) is less than one. We thus obtain a bound on ρ^2 that is strictly less than one and is independent of h. If β is chosen small enough, then the first term in (46) will be negligible compared to the second, and we will have established a bound on ρ^2 that is approximately equal to $\left(\frac{\kappa'-1}{\kappa'+1}\right)^{2\ell}$, where $\kappa' = \alpha/\beta$ is independent of h.

We have thus shown that by using a matrix C of the form (24), with γ chosen appropriately, we obtain a multigrid method (actually a two-grid method since we have assumed that the matrix A_J is inverted exactly) that reduces the A-norm of the error at each step by a certain constant factor independent of the mesh spacing h. The eigenvalues λ_N and λ_{j+1} in the expression for γ are usually not known, but if the constant γ is chosen in an optimal way at each step, as in the method of steepest descent, then the same result holds.

The analysis presented here is actually for a two-grid method, as we have assumed that the matrix A_J in (13) is inverted exactly. In a true multigrid method, the linear systems with coefficient matrix A_J would not be solved exactly, but instead would be solved to within some specified tolerance by using the multigrid method with still coarser grids. Since the linear systems would not be solved exactly, the error terms would have to enter into the analysis. For a precise analysis of true multigrid methods, we refer the reader to [3], where it is shown that if the equations on each of the coarser grids are solved only to within the truncation error of the finite element approximation for that grid, then the fine grid solution is obtained to within truncation error in O(N) operations. The results presented here are similar to those in [3] and also to those in [2], though we have presented them from a somewhat different point of view.

4. Numerical Examples and Different Forms for the Matrix C

Expression (46) provides the key to determining under what circumstances the multigrid method of section 3 will or will not converge quickly. If the constant c in (46) is large, then β will have to be chosen small in order to make the first term in expression (46) small. But then the ratio α/β will probably be large, resulting in a slow (but still independent of h) convergence rate for the multigrid method. The constant c is not usually known but can sometimes be estimated for model problems.

In Table 1 we show the results of using the algorithm of section 3 to solve the diffusion equation

$$-\nabla \cdot p(x,y)\nabla u = f \quad \text{on } (0,1)X(0,1)$$
$$u(x,0) = u(x,1) = u(0,y) = u(1,y) = 0$$

using piecewise bilinear finite elements ($r = 2$, $m = 1$) on a square 21 X 21 grid ($N = 21$, $h = 1/22$). An iteration consisted of one coarse grid correction cycle ($J = 10$, $\rho = 2$, $\rho h = 1/11$) followed by one ($\ell = 1$) steepest descent relaxation step. The coarse grid matrix A_J was inverted directly. The number of iterations required to reduce the Euclidean norm of the residual to the level 10^{-6} is given in Table 1 along with the number of iterations required by another popular algorithm -- the conjugate gradient method with diagonal scaling.

An estimate of the constant c for this problem involves the ratio $p_{max}^{1/2}/p_{min}$ [4, p. 47]. For c of moderate size, the multigrid method outperformed CG-DS in terms of number of iterations, for all problems tried ($p(x,y) = 1, 1+x+y, 1+x^2+y^2$). (Though we should point out that an "iteration" of a true multigrid method would involve solving to within truncation error an equation on the grid with spacing 2h, and hence would probably be considerably more expensive than a conjugate gradient iteration. On the other hand, the grid used in this experiment is rather coarse, and with smaller values of h the difference in number of iterations between the conjugate gradient method and the multigrid method is even greater in favor of the multigrid computation.) By taking $p(x,y)$ to be $.1+x^2+y^2$, however, we increased the ratio $p_{max}^{1/2}/p_{min}$ and greatly slowed the convergence rate of the multigrid method. An attempt to use the multigrid method with $p(x,y)$ equal to $.01+x^2+y^2$ failed to converge after 200 iterations. We also tried the multigrid method on a problem for which it really was not designed ($p(x,y) = 1$ if $x \leq .5$, 100 if $x > .5$) because the coefficients of the differential operator are not smooth and so the approximation theorems of section 3 do not hold. As might be predicted, the method performed very poorly and its convergence rate was not independent of h. Methods have been suggested for improving the performance of the multigrid method on such problems by using a different interpolation matrix, but we will not discuss these ideas here.

In looking at Table 1, one is struck by the consistency with which the conjugate gradient method with diagonal scaling performed on all of these problems. Had the conjugate gradient method been used without diagonal scaling, it too would have exhibited the type of behavior displayed by the multigrid method -- slowing down on problems for which the diagonal elements of A varied greatly. Perhaps the multigrid method should be used with a relaxation step that also involves diagonal scaling, such as Jacobi's method or the steepest descent method with diagonal scaling or even conjugate gradients with diagonal scaling.

The theory that we have given, however, applies only when the matrix C is of the form a constant times the identity. The finite element theory tells us (theorem 3.1) that some of the eigenvectors of A can be well approximated by vectors of the form $A^{1/2}Pw$. But to establish fast convergence we need to know that some eigenvectors of $A^{1/2}CA^{1/2}$ are approximated by vectors of this form. Even for C of the simple form $C = (\text{diag}(A))^{-1}$, we have not yet established a general result relating the eigenvectors of $A^{1/2}CA^{1/2}$ to vectors of the desired form.

Despite the lack of theory, we tried combining the multigrid method with CG-DS relaxation steps, and the results are shown in Table 2. Here we combined one coarse grid correction cycle with

either one, two, or three conjugate gradient steps, and shown in the table are the number of fine grid steps required to reduce the residual to 10^{-6}. Similar results were obtained using the steepest descent method with diagonal scaling as the relaxation method. By using diagonal scaling, all of the problems except the last were made essentially as easy for the multigrid method as the Poisson equation ($p = 1$). Numerical experiments indicated a convergence rate independent of h, so it is conjectured that some eigenvectors of $A^{1/2}CA^{1/2}$ are approximated by vectors of the form $A^{1/2}Pw$, just as some eigenvectors of A are so approximated. For the last problem, in which the theory of section 3 does not apply, the extra coarse grid correction cycles still reduced slightly the number of iterations required by the conjugate gradient method to achieve convergence. This is probably because at least some eigenvectors of the iteration matrix are approximated by vectors of the desired form, but not closely enough to give a convergence rate independent of h.

Acknowledgment

The author would like to thank Steve McCormick for many helpful discussions and for first suggesting that the A-norm of the error be analyzed in the multigrid method.

p(x,y)	# Iterations	
	MG-SD	CG-DS
1	13	44
$1 + x + y$	28	44
$1 + x^2 + y^2$	30	44
$.1 + x^2 + y^2$	140	49
$.01 + x^2 + y^2$	> 200	49
$1, \quad x \leq .5$ $100, \quad x > .5$	> 200	46

Table 1. Multigrid with Steepest Descent (MG-SD) Vs. Conjugate Gradients with Diagonal Scaling (CG-DS) for the Equation $-\nabla \cdot p(x,y)\nabla u = f$.

p(x,y)	# Fine Grid Steps		
	MGCGDS(1)	MGCGDS(2)	MGCGDS(3)
1	13	11	18
$1 + x + y$	13	11	19
$1 + x^2 + y^2$	13	11	17
$.1 + x^2 + y^2$	13	11	18
$.01 + x^2 + y^2$	13	11	18
$1, \quad x \leq .5$ $100, \quad x > .5$	35	30	31

Table 2. Multigrid with One, Two, or Three Steps of Conjugate Gradients with Diagonal Scaling (MGCGDS(·)) for the Equation $-\nabla \cdot p(x,y)\nabla u = f$.

References:

[1] A. Brandt. Multilevel Adaptive Solutions to Boundary Value Problems. Math. of Comp. 31, 1977, pp. 333-390.

[2] S. McCormick and J. Ruge. Multigrid Methods for Variational Problems. to appear in SIAM J. Num. Anal.

[3] R. Nicolaides. On the ℓ^2 Convergence of an Algorithm for Solving Finite Element Equations. Math. of Comp. 31, 1977, pp. 892-906.

[4] G. Strang and G. Fix. An Analysis of the Finite Element Method. Prentice-Hall, Englewood Cliffs, N. J. 1973.

Technical Information Department · Lawrence Livermore Laboratory
University of California · Livermore, California 94550