Bull. Nov. Comp. Center, Num. Anal., 9 (2000), 59–67

An optimal multilevel method for computing the smallest eigenpair*

Maxim Larin

1. Introduction

In this paper, we are interested in computing the smallest eigenvalue and its corresponding eigenvector of a large symmetric positive definite matrix A, i.e., we find the first eigenpair $\{\lambda, \boldsymbol{u}\}$ such that

$$\begin{aligned} &A\boldsymbol{u} = \lambda \boldsymbol{u}, \quad A = A^T \in \boldsymbol{R}^{n\times n}, \quad \lambda > 0, \\ &\|\boldsymbol{u}\| = (\boldsymbol{u}, \boldsymbol{u})^{1/2} = 1, \end{aligned} \tag{1}$$

where $(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^T \boldsymbol{v}$. Moreover, without loss of generality we assume that λ is a simple and well separated eigenvalue since it may be derived by the special perturbation technique [1] which will not be analyzed in this paper. Typically, the matrix A is large and sparse. To be able to solve problem (1) with a reasonable computing time, one must use *an optimal* (or *nearly optimal*) technique, i.e., the methods for which the computational complexity grows *linearly* (or *almost linearly*) with increasing the problem size.

The usual methods for solving eigenvalue problems are often based on the effect of excitation of the smallest eigenvalues by repeated multiplication of the inverse matrix A^{-1} by a vector. This applies to such popular techniques as a subspace iteration, the Rayleigh quotient and the Lancoz method [2]. However, in the large-scale finite element problems, it is often desirable to avoid a costly inversion or, to be more precise, the exact factorization of the matrix A. A simple way is to use some iterative procedure instead of the direct method for solving the system whenever it is required in the algorithm. Note, also, that if we want to take an advantage of the Lancoz method, i.e., the exact three-diagonal form of the Ritz matrix, we have to take a special care in the solution method.

Today among a large body of literature published on this topic, one can distinguish two main directions. The first one is based on the implicit application of the multigrid methods to constructing the approximated inverse

*Supported by the NWO-NL-RF under Grant 047.008.007 and by the Russian Foundation for Basic Research under Grant 99-01-00579.

of A and their use as a preconditioner [3, 4], when the second one directly applies the main multigrid ideas of the fine grid relaxation and the coarse grid correction [5, 6]. The recent approach can also be split into two different parts similarly to the splitting of the basic multigrid methods into V- and F-cycles.

However, all the above approaches are treating this eigenvalue problem as purely algebraic, hence, we can lose some valuable information about the desired eigenpair. For example, the smallest eigenvector of the Laplace operator is very smooth, i.e., it can be well approximated on coarser grids, hence, one can use this information in the solution process. The full multigrid process, which takes the advantage of this smoothness, has been proposed by Brandt, McCormick, and Ruge [7, 8]. The idea of this method is in that we solve similar eigenvalue problems on a sequence of the finer grids using an interpolant of the solution on each level as the initial guess for the next one and improving it by the internal V-cycle multigrid method, i.e., suppressing the high-frequency oscillations arising as a result of the interpolation process.

In spite of the good numerical results and the fact that the method has been proposed in the early eighties, there had been no substantial further research and development in this direction for many years. It can be intuitively motivated by the following well-known fact that a standard (geometrical) multigrid process looses its optimal properties, when the coefficients become anisotropic or have large jumps. To overcome these difficulties, one can use special multigrid techniques, which successfully applied for solving a related linear system of equations, which arises as a result of the finite element approximations of similar elliptic boundary value problems. In particular, semi-coarsening and algebraic multigrid techniques allow us to construct optimal preconditioners for these boundary value problems.

The main advantages of the AMG method, which was developed by Brandt, McCormick, Ruge and Stuben [8–10], are its robustness and its applicability for complex geometric situations and, even, for problems with no geometrical or continuous background at all (as long as the given matrix satisfies certain conditions).

In the present study, we modify the above mentioned multigrid eigensolver using the algebraic multigrid technique, i.e., more exactly, we use the operator dependent interpolation and the Galerkin based coarse-grid correction process to construct a new matrix on the lower level, which has a structure similar to that of the original matrix. Moreover, its convergence analysis has not been given in [11]. Here we prove it under some natural assumptions on the restriction (or prolongation) matrices and the internal multigrid method.

The paper is organized as follows. The description of the Full Algebraic MultiGrid method for EigenValue problems or, shortly, the FAMG-EV

method, and its components are presented in Section 2. In Section 3, the computational complexity and implementation details are discussed. The proof of the convergence behaviour is presented in Section 4. In the final section of the paper, the experimental results on standard test problems are presented.

2. The FAMG-EV method

The idea of the FAMG-EV method based on the idea of the well-known full multigrid approach, i.e., we find some approximate solution of a similar problem on the coarse level and using its interpolant as the initial guess for some internal iterative process on the next (finer) one.

2.1. Construction of $\{A^{(k)}\}$. Let $N_k \in \{1,\ldots,n_k\}$ be a set of indices of all unknowns (or nodes) on the level k, $k \geq 0$, and it is partitioned into two non-intersected subsets, e.g., $N_k^f \subset \{1,\ldots,n_k\}$ and $N_k^c \subset \{1,\ldots,n_k\}$ such that $N_k = N_k^c \cup N_k^f$ and $N_k^c \cap N_k^f = \emptyset$, using a heuristic algorithm, based on either the problem-dependent geometrical information, see [6, 11] and references therein, or the matrix-dependent numerical information [7, 9, 12, 13]. This partitioning yields 2-by-2 block matrix form of $A^{(k)}$, where the first group of unknowns corresponds to the indices in N_k^f and the second one forms a new set of indices on the next level, i.e., $N_{k+1} = N_k^c$. For the given partitioning, coefficients for the interpolation of the value $u_i^{(k)}$, $i \in N_k^f$, from $u^{(k+1)} = u_j^{(k)}$, $j \in N_k^c$, are computed using the matrix-dependent information. This defines a prolongation matrix $P_{k+1}^k \in R^{n_k \times n_{k+1}}$. The detailed choice of P_{k+1}^k used here is outline of the topic of this paper and is addressed to in [12–14]. Then the Galerkin based coarse-grid matrix is defined by

$$A^{(k+1)} = R_k^{k+1} A^{(k)} P_{k+1}^k, \qquad R_k^{k+1} = (P_{k+1}^k)^T, \tag{2}$$

where B^T is a transpose of B. As the corresponding matrix $A^{(k+1)}$ has usually the same properties as $A^{(k)}$, the coursing process can be repeated until a sufficiently coarse grid is attained, where the corresponding eigenvalue problem can be solved by a direct or an iterative method with sufficiently small computational cost in relation to n_0.

2.2. The basic method. Now we formulate the FAMG-EV method for solving the sequence of intermediate problems

$$A^{(k)} u^{(k)} = \lambda u^{(k)}, \qquad \|u^{(k)}\| = 1, \qquad k = L, L-1, \ldots, 1, 0, \tag{3}$$

beginning with the coarse level and including the finest level as the final stage.

In order to start the multilevel (recurrent) process, we have to compute the coarse level approximation $\{\lambda, v^{(L)}\}$ to the desired eigenpair $\{\lambda, u^{(L)}\}$ by solving the coarse level eigenvalue problem. Due to the fact that the order n_L of this system is sufficiently small, we can exactly compute it with a small number of arithmetical operations.

Now assume that we have an approximate solution $\{\lambda, v^{(k+1)}\}$ on the level $k+1$. The transition to the previous level k starts with the interpolation of the approximated eigenvector $v^{(k+1)}$ as follows

$$v^{(k+1)} = P_{k+1}^{k} v^{(k+1)}. \tag{4}$$

Next, in order to eliminate high-frequency oscillations, which arise as a result of the interpolation process (4), we apply μ times the internal algebraic multigrid method, which is based on the Full Approximation Scheme (FAS) approach (see, for example, [6]) and use the same sequence of matrices $\{A^{(k)}\}$. A specific nature of the internal solver will be discussed below. Usually, one or two internal multigrid sweeps is enough to suppress all (or almost all) undesired frequencies in the approximated solution. Finally, the iterate $v^{(k)}$ is normalized, and λ is updated by computing the corresponding Rayleigh quotient. This process is then repeated at the next finer level until the finest level ($k = 0$) is reached.

Algorithm **FAMG-EV**:

Compute the coarse level approximation $\{\lambda, v^{(L)}\}$
for $k = L-1, \ldots, 1, 0$
  $v_0^{(k)} = P_{k+1}^{k} v^{(k+1)}$
  for $i = 1, \ldots, \mu$
    $v_i^{(k)} = \mathbf{VMG}(k, v_{i-1}^{(k)}, A^{(k)} - \lambda I_k, 0)$
  $v^{(k)} = v_\mu^{(k)} / \|v_\mu^{(k)}\|$
  $\lambda = (A^{(k)} v^{(k)}, v^{(k)}) / (v^{(k)}, v^{(k)})$

Here $\mathbf{VMG}(k, u, A, b)$ denotes the internal V-cycle multigrid method applied for solving the linear system $Ax = b$ at the level k using u as initial guess.

2.3. The internal multigrid solver. First of all, to explain the inner solver it is necessary to fix its goal. As stated before, once the vector $v^{(k)}$ has been interpolated by (4) at the level $k-1$, the vector $v^{(k-1)}$ is a *weak* approximation to $u^{(k-1)}$ to be sought for. Moreover, the eigenvalue problem is a *nonlinear* problem.

One possible way to solve all the above-mentioned problems is to use a nonlinear multigrid method as the full approximation scheme (FAS) (see,

for example, [6]), which responses quickly and qualitatively eliminates the high-frequency components of the corresponding error $u^{(k-1)} - v^{(k-1)}$.

Finally, note that the VMG method can be considered as eigenvalue solver itself.

Algorithm $\mathbf{VMG}(v^{(l)}, A^{(l)} - \lambda I_l, 0)$:

for $k = l, \ldots, L-1$

  for $i = 1, \ldots, \nu_1$ (*presmoothing*)

    $v_i^{(k)} = \mathbf{GSI}(A^{(k)} - \lambda I_k, v_{i-1}^{(k)}, \tau^{(k)})$

  $v^{(k)} = v_{\nu_1}^{(k)}$

  $\tau^{(k+1)} = R_k^{k+1}\tau^{(k)} + A^{(k+1)}R_k^{k+1}v^{(k)} - R_k^{k+1}A^{(k)}v^{(k)}$

  $v_0^{(k+1)} = R_k^{k+1}v^{(k)}$

Solve $A^{(L)}v^{(L)} = \lambda v^{(L)} + \tau^{(L)}$ (*coarse problem*)

for $k = L-1, \ldots, l$

  $v_0^{(k)} = v^{(k)} + P_{k+1}^k(v^{(k+1)} - R_k^{k+1}v^{(k)})$

  for $i = 1, \ldots, \nu_2$ (*postsmoothing*)

    $v_i^{(k)} = \mathbf{GSI}(A^{(k)} - \lambda I_k, v_{i-1}^{(k)}, \tau^{(k)})$

  $v^{(k)} = v_{\nu_2}^{(k)}$

Here $\mathbf{GSI}(A, y, b)$ denotes the Gauss–Seidel iteration step applied for solving the linear system $Ax = b$ using y as initial guess.

3. Computational complexity

In order to investigate the whole computational complexity of the FAMG-EV method we recall that the solution of the internal eigenvalue problem by the VMG method requires the highest computational costs. By its definition, such a solution breaks up into a set of problems with the matrices $A^{(k)}$ at all the levels.

Assume that the number of nodes decreases in a geometrical ratio with a factor ρ defined by

$$\frac{n_{k+1}}{n_k} = \rho_k \leq \rho < 1, \quad k = 0, 1, \ldots, L-1.$$

Now denoting by W the whole computational complexity of the VMG method and applying a standard recursive technique we obtain that

$$W = C(\mu, \nu_1, \nu_2, \rho) n_0 = O(n_0),$$

where $C(\mu, \nu_1, \nu_2, \rho)$ is a constant depending on the user-defined parameters only. Hence, the total computational costs of the FAMG-EV method are optimal.

4. Convergence analysis

The convergence analysis of the method is based on the natural partitioning of the process into the two main parts. At the first stage, we prove under some natural assumptions that the method is convergent and at the second stage we verify our assumptions.

Assume that

(A1) the internal multigrid method is convergent, i.e., $\|M^{(k)}\| \leq \eta < 1$, where $M^{(k)}$ is the matrix operator of the VMG method, whose explicit form will be discussed below;

(A2) $\|P_{k+1}^k\| = \|R_{k+1}^k\| \leq C$, i.e., the restriction (or prolongation) operator is bounded;

(A3) $\|\boldsymbol{u}^{(k+1)} - \boldsymbol{v}^{(k+1)}\|_{k+1} < \varepsilon_{k+1}$, $|\lambda - \lambda_{k+1}| < \delta_{k+1} \leq \lambda$, i.e., we obtain a *sufficiently* accurate solution at the previous level;

(A4) $\eta^\mu C^2 < 1$, i.e., the internal multigrid method is able to eliminate the high-frequency components of the error caused by the interpolation process.

Theorem. *Let the assumptions* (A1)–(A4) *be satisfied, then*

$$|\lambda_k - \lambda| < \varepsilon_{k+1}^2, \qquad \|\boldsymbol{u}^{(k)} - \boldsymbol{v}^{(k)}\|_k < \varepsilon_{k+1},$$

for all $k = L-1, \ldots, 1, 0$, *hence, the FAMG-EV method is monotonically convergent to solution* (1) *if a sufficiently accurate solution on the coarse level has been found.*

Proof. To write $M^{(k)}$, we first consider the two level cases. As usual, analysis of any multigrid like techniques is divided into the analysis of two independent parts: smoothing and the coarse-grid correction. In our case, we use a nonlinear Gauss–Seidel relaxation step as smoothing and the FAS technique for the coarse grid-correction which can be written as follows

$$M^{(k)} = S^{\nu_2} \left[I_k - P_{k+1}^k (A^{(k+1)} - \lambda I_{k+1})^{-1} \left(A^{(k+1)} R_k^{k+1} - R_k^{k+1} A^{(k)} \right) \right] S^{\nu_1}.$$

Now using the standard multigrid theory for the FAS method [6] we obtain

$$\|M^{(k)}\| \leq \eta < 1,$$

which proves the assumption (A1), and moreover, if

$$\mu > 2 \log_\eta C,$$

the assumption (A4) is fulfilled. Other assumptions are *natural* and are easy satisfied by the corresponding choice of prolongation matrices and the coarse-level solution method. □

5. Numerical results

To test the method, we first consider eigenvalue problem (1) for the small eigenpair $\{\lambda, v\}$ which corresponds to the piecewise-linear finite element discretization of the three-dimensional second order elliptic problems

$$-\Delta u = \lambda u \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \Gamma_D = \partial\Omega, \qquad \|u\| = 1,$$

in the cube domain $\Omega = [0,1]^3$ on a uniform Cartesian mesh $\mathcal{T}_h$ with the step size h.

For the present example, the AMG method yields a sequence of matrices similar to the one for the standard (geometrical) MG method. Moreover, we use the following generalized parameter ν defined by $\nu = \nu_1 = \nu_2$. The number of the coarse grid iterations ν_0 is equal to n_0.

The accuracy characteristics of the computed eigenpair $\{\lambda_h, v\}$ to the exact one $\{\lambda, u\}$ with respect to the parameters of the FMG method, $\lambda = 0.0072272627$, $N = 64$ are shown in the table below.

L	λ_h	$\lvert\lambda - \lambda_h\rvert$	$\|Av - \lambda_h v\|$
		$\mu = 1$	
5	$0.72376 \cdot 10^{-2}$	$0.10418 \cdot 10^{-4}$	$0.64687 \cdot 10^{-2}$
4	$0.72386 \cdot 10^{-2}$	$0.11352 \cdot 10^{-4}$	$0.69350 \cdot 10^{-2}$
3	$0.72489 \cdot 10^{-2}$	$0.21665 \cdot 10^{-4}$	$0.10321 \cdot 10^{-1}$
2	$0.79770 \cdot 10^{-2}$	$0.74975 \cdot 10^{-3}$	$0.14544 \cdot 10^{-1}$
		$\mu = 2$	
5	$0.72275 \cdot 10^{-2}$	$0.30742 \cdot 10^{-6}$	$0.14280 \cdot 10^{-2}$
4	$0.72273 \cdot 10^{-2}$	$0.72394 \cdot 10^{-7}$	$0.67310 \cdot 10^{-3}$
3	$0.72273 \cdot 10^{-2}$	$0.60540 \cdot 10^{-7}$	$0.60642 \cdot 10^{-3}$
2	$0.75394 \cdot 10^{-2}$	$0.31222 \cdot 10^{-3}$	$0.31084 \cdot 10^{-2}$

On the basis of the conducted experiments the following conclusions can be made:

1. Comparing the values of $|\lambda - \lambda_h|$ with the values of $\|A^{(k)} v^{(k)} - \lambda_h v^{(k)}\|$ for all Tables, one can see that the square root of the first one is proportional to the second, hence, the results of the theorem are numerically confirmed as well.

2. From the first part of each Table one can see that if we use the maximal number of possible levels, then we lose the convergence to the solution, since the solution of the coarse eigenvalue problem does not represent a real approximation to the smallest eigenvalue. Indeed, in this case, on the coarse grid we have the problem for one (!) unknown only. Naturally, it is not sufficient to compute some valuable coarse-grid approximation to the smallest eigenvector, hence, the further work at intermediate levels does not considerably improve this approximation.

3. Decreasing the number of levels used here results in the necessity to increase ν, and hence, the computational costs are also increased. Thus, the number of levels has to be as maximal as possible. Analogously to the standard AMLI technique, it can be recommended to use $\frac{2}{3}L_{\max}$ levels, where $L_{\max}$ is the maximal number of possible levels.

4. Comparing the results for $\mu = 1$ with the corresponding results for $\mu = 2$, in all the tables, one can see that in the second case the accuracy of the approximated solution is better, i.e., the increasing number of the internal VMG steps results in improving the accuracy of the approximated eigenpair $\{\lambda_h, \boldsymbol{v}\}$.

As the final conclusion, we have found that the FAMG-EV method, applied for computing the smallest eigenvalue and its eigenvector, leads to an iterative eigensolver with an optimal order of the computational complexity and the monotone convergence behaviour. However, the method is applied well only for separated eigenpair $\{\lambda, \boldsymbol{u}\}$, i.e., the method does not work properly for the problems with multiplicity eigenvalues. It seems that it is not a strong restriction on the presented method, since the proposed technique with corresponding block type modifications can be adopted for such problems and, also, for the problem of computing the number of smallest eigenvalues and their eigenvectors. The future investigation will be developed in this way.

Acknowledgements. The author would like to thank Dr. K. Stuben and C.-A. Thole for the initiation of the present study and the fruitful discussion of numerical results.

References

[1] Dyakonov E., Knyazev A. On an iterative method for finding lower eigenvalues // Rus. J. Num. Anal. Math. Modelling. – 1992. – № 7. – P. 473–486.

[2] Parlett B. The Symmetric Eigenvalue Problem. – Prentice-Hall, Inc., 1980.

[3] Cai Z., Mandel J., McCormick S. Multigrid methods for nearly singular linear equations and eigenvalue problems // SIAM J. Numer. Anal. – 1997. – Vol. 34, № 1. – P. 178–200.

[4] Knyazev A.V. Preconditioned eigensolvers – an oxymoron? // Electronic Transaction on Numerical Analysis. – 1998. – Vol. 7. – P. 104–123.

[5] Brandt A., McCormick S., Ruge J. Multigrid methods for differential eigenproblems // SIAM J. Sci. Stat. Comput. – 1983. – Vol. 4, № 2. – P. 244–260.

[6] Hackbusch W. Multigrid Methods and Applications. – Springer-Verlag, 1985.

[7] Brandt A. Algebraic multigrid theory: The symmetric case // Appl. Math. Comp. – 1986. – Vol. 19. – P. 23–56.

[8] Brandt A., McCormick S., Ruge J. Algebraic multigrid (AMG) for sparse matrix equations // Sparsity and its Applications (D.J.Evans, ed.).– Cambridge: Cambridge Univ. Press, 1984.

[9] Ruge J., Stuben K. Efficient solution of finite diference and finite elemenet equations by algebraic multigrid (AMG) // Multigrid Methods for Integral and Differencial Equations (D. J. Paddon and H. Holstein, eds.) / The Institute of Mathematics and its Applications Conference Series. – Oxford: Claredon Press, 1985. – P. 169–212.

[10] Stuben K. Algebraic Multigrid (AMG): Experiences and comparisions // Appl. Math. Comp. – 1983. – Vol. 13. – P. 419–452.

[11] Shaidurov V.V. Multigrid Methods for Finite Elements. – Kluwer Academic Publ., 1995.

[12] Stuben K. Algebraic Multigrid (AMG): An Introduction with Applications / GMD Report 53. – Germany: Scholss Birlinghoven, Sankt Augustin, March 1999.

[13] Bank R., Wagner C. Multilevel ILU decomposition // Numerische Mathematik. – 1999. – Vol. 82. – P. 543–576.

[14] Ruge J., Stuben K. Algebraic Multigrid (AMG) // Multigrid Methods (S. McCormick, ed.) / SIAM. Fronties in Applied Mathematics. – Philadelphia, 1986. – Vol. 5.