JOURNAL OF COMPUTATIONAL PHYSICS 4, 419–423 (1969)

A Choice of Starting Vectors in Relaxation Methods

This note is concerned with the first step in the solution of difference equations by relaxation methods. A by-product of the relaxation parameter calculation is used to form an effective first guess with very little cost in additional computation.

I. Introduction

An accurate estimate of the spectral radius ρ, of the Jacobi iteration matrix is necessary to obtain the optimal asymptotic rate of convergence by any of the several variants of successive over-relaxation [1]. A survey of the literature reveals that two distinct points of view prevail on how this estimate should be made. One approach is to attack the eigenvalue problem directly and calculate the spectral radius as a preliminary to the main iteration [1, 2, 3]. In the second method, a value is assumed for ρ and the iteration is started. As the sequence of vectors proceeds toward the solution, a new estimate is obtained by comparing the theoretical and the observed rates of convergence [4, 5, 6].

Since all computation in the second method brings the vector closer to the desired solution, it seems to be the more attractive. However, in practice the direct approach often requires less computation overall because of the more accurate estimate of ρ. The relative merits of the two techniques apparently have not been settled to everyone's satisfaction. The purpose of this note is to demonstrate a combination of desirable features from both ideas by forming an initial guess using the eigenvector associated with ρ.

II. Computation of the Spectral Radius

To be specific, the following discussion is based on the assumption that a set of difference equations is to be solved by the Cyclic Chebyshev method.

That is, the equations are written in the form

$$\begin{bmatrix} D_1 & -B \\ -B^t & D_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} \tag{1}$$

where the coefficient matrix is symmetric and positive definite and its partitions D_1, D_2, and B are sparse band type matrices. The equations which describe the

iteration are taken from [1]. Beginning with an initial guess $\mathbf{x}_2^{(0)}$, we construct the sequence

$$\begin{aligned}
\mathbf{x}_1^{(2m+1)} &= \omega_{2m+1}\{D_1^{-1}B\mathbf{x}_2^{(2m)} + D_1^{-1}\mathbf{k}_1 - \mathbf{x}_1^{(2m-1)}\} + \mathbf{x}_1^{(2m-1)}, \qquad m \geqslant 0, \\
\mathbf{x}_2^{(2m+2)} &= \omega_{2m+2}\{D_2^{-1}B^t\mathbf{x}_1^{(2m+1)} + D_2^{-1}\mathbf{k}_2 - \mathbf{x}_2^{(2m)}\} + \mathbf{x}_2^{(2m)}, \qquad m \geqslant 0, \\
\omega_1 &= 1 \\
\omega_2 &= 1/(1 - \rho^2/2) \\
\omega_{n+1} &= 1/(1 - \rho^2\omega_n/4),
\end{aligned} \tag{2}$$

where ρ is a number in the interval $0 \leqslant \rho < 1$. For optimal asymptotic rate of convergence, ρ must be the spectral radius of

$$\begin{bmatrix} 0 & D_1^{-1}B \\ D_2^{-1}B^t & 0 \end{bmatrix}. \tag{3}$$

The two by two partitioning of this matrix allows us to carry out the calculation for ρ^2 using the smaller matrix $D_2^{-1}B^tD_1^{-1}B$. Several methods are available for computing the dominant eigenvalue λ of

$$\lambda D_2\mathbf{v}_2 = B^tD_1^{-1}B\mathbf{v}_2\,. \tag{4}$$

We choose the power method for two reasons: First, the computer program which performs the iteration (2) requires only minor modification to become the power method applied to (4). The algorithm takes advantage of the sparseness of the partitions and it is amenable to acceleration techniques [2], so the calculation is performed quite efficiently. The second reason for choosing the power method is that it yields an approximation to both the dominant eigenvalue and its associated eigenvector. The purpose to be served by this vector is set forth in the next section.

III. The Initial Guess

The iteration scheme (2) requires a first guess $\mathbf{x}_2^{(0)}$. A known solution of a similar problem or, for want of better information, the zero vector is often used for $\mathbf{x}_2^{(0)}$. In either case, we have the information at hand to modify $\mathbf{x}_2^{(0)}$, at trivial cost, so as to reduce the total number of iterations required to obtain a satisfactory solution of (1).

Equations (1) can be decoupled by a block elimination to give

$$(D_2 - B^t D_1^{-1} B)\mathbf{x}_2 = B^t D_1^{-1}\mathbf{k}_1 + \mathbf{k}_2 . \tag{5}$$

We denote the eigenvalues and corresponding eigenvectors of (4) by $\lambda_1 > \lambda_2 \geqslant \cdots \lambda_i > 0$ and $\{\mathbf{v}_{2i}\}$. The solution and the initial guess are expressed in terms of the D_2-orthogonal basis $\{\mathbf{v}_{2i}\}$.

$$\begin{aligned} \mathbf{x}_2 &= \sum_i \alpha_i \mathbf{v}_{2i} \\ \alpha_i &= \frac{\langle \mathbf{v}_{2i} , B^t D_1^{-1}\mathbf{k}_1 + \mathbf{k}_2 \rangle}{(1 - \lambda_i)\langle \mathbf{v}_{2i} , D_2 \mathbf{v}_{2i} \rangle} \end{aligned} \tag{6}$$

$$\begin{aligned} \mathbf{x}_2^{(0)} &= \sum_i \beta_i \mathbf{v}_{2i} \\ \beta_i &= \frac{\langle \mathbf{v}_{2i} , D_2 \mathbf{x}_2^{(0)} \rangle}{\langle \mathbf{v}_{2i} , D_2 \mathbf{v}_{2i} \rangle} . \end{aligned} \tag{7}$$

Since the entire spectral decomposition of $D_2 - B^t D_1 B$ is not available, the solution cannot be obtained through Eq. (6). However, the essential ingredients of its first term are at hand which we now use to adjust (7).

$$\begin{aligned} \mathbf{x}_2^{(0)} &= \mathbf{x}_2^{(0)} - \delta \mathbf{v}_{21} \\ \delta &= \beta_1 - \alpha_1 . \end{aligned} \tag{8}$$

It is unnecessary to compute $D_2\mathbf{v}_{21}$ which appears in the denominators of α_1 and β_1 because its equivalent from the right hand side of (4) will be saved from the eigenvalue calculation. Thus, we have

$$\delta = \frac{\lambda_1 \langle \mathbf{v}_{21} , D_2 \mathbf{x}_2^{(0)} \rangle}{\langle \mathbf{v}_{21} , B^t D_1^{-1} B \mathbf{v}_{21} \rangle} - \frac{\lambda_1}{1 - \lambda_1} \frac{\langle \mathbf{v}_{21} , B^t D_1^{-1}\mathbf{k}_1 + \mathbf{k}_2 \rangle}{\langle \mathbf{v}_{21} , B^t D_1^{-1} B \mathbf{v}_{21} \rangle}$$

which requires only the matrix vector multiplication $D_2\mathbf{x}_2^{(0)}$ and three inner products as significant extra expense.

IV. Numerical Experiments

This technique has a disadvantage common to many which are concerned with the early part of a sequence: It is difficult to make a quantitative prediction of its effectiveness which will hold in general. In fact, for the unlikely case where $\alpha_1 = \beta_1$ at the beginning, nothing at all is gained. A more pertinent theoretical objection to our strategy is that a completely accurate calculation of $\mathbf{v}_{21}$ and adjustment of $\mathbf{x}_2^{(0)}$ will cause the subsequent iteration (2) to take place in a subspace in which $\lambda_1 = \rho^2$ is no longer the optimal parameter.

In practice, our eigenvector is not computed with great accuracy. The use of Eq. (8) is recommended on the grounds that the initial error vector may be shortened, sometimes quite appreciably, in one trivial step and the discrimination in Eq. (2) against the $\mathbf{v}_{21}$ direction caused by underestimating ρ will have its effect reduced.

The calculation of an initial guess in the manner discussed above has been included in a computer program which solves the finite difference analog of Poisson's equation in two dimensions using the one-line Cyclic Chebyshev method. Comparisons were made between the zero vector and the result of (8) in terms of the number of iterations required to reduce the l_1 norm of the residual to a prescribed value. These experiments indicate that the calculated guess is indeed worthwhile. No comparisons were made with the indirect approach of [4, 5, 6].

Two comparisons were made with a model region consisting of a square with a 15×15 uniform mesh superimposed on it. Poisson's equation was solved with zero boundary conditions and a unit load at the center point. Convergence was reached after 30 iterations with a zero guess and after 21 iterations starting with the calculated guess. With a right hand side of the differential equation set to zero, boundary conditions of zero on three sides of the square, and the boundary condition on the fourth side taken to be unity, the number of iterations were 30 and 23 respectively. A third test was conducted on a more realistic problem; that of representing the field in a cylindrical image tube. The total number of discrete points was 616, on a mesh which was non-uniform in both directions. Laplace's equation was solved with mixed type boundary conditions. Convergence was reached after 63 and 58 iterations respectively for the two starting vectors. For all three problems, the stopping criterion was that the l_1 norm of the initial residual vector be reduced by a factor of 10^{-6}.

The more modest results of the third case may be more typical in practical problems. In any case, the fact that any such improvement in performance may be obtained so simply makes this technique worthwhile.

REFERENCES

1. R. S. VARGA, "Matrix Iterative Analysis." Prentice Hall, Inc. Englewood Cliffs, N. J., 1962.
2. L. A. HAGEMAN AND R. B. KELLOGG, Estimating optimum overrelaxation parameters. *Math. Comp.* **22**, 60–68 (1968).
3. A. K. RIGLER, Estimation of the successive overrelaxation factor. *Math. Comp.* **19**, 302–307 (1965).
4. G. E. FORSYTHE AND J. ORTEGA, Attempts to determine the optimum factor for successive overrelaxation. Proc. Int. Conf. Information Processing, UNESCO, 1960.
5. B. A. CARRÉ, The determination of the optimum accelerating factor for successive overrelaxation. *Computer J.* **4**, 73–78 (1961).
6. H. E. KULSRUD, A practical technique for the determination of the optimum relaxation factor of the successive overrelaxation method. *Comm. A.C.M.* **4**, 184–187 (1961).

RECEIVED: July 1, 1968

A. K. RIGLER

Quantum Electronics Research and Development,
Westinghouse Electric Corporation,
Pittsburgh, Pennsylvania 15235