Some Convergence Results Related to the JOR Iterative Method for Symmetric, Positive-Definite Matrices

Firdaus E. Udwadia

Mechanical and Civil Engineering Departments
University of Southern California
Los Angeles, California 90089-1453

ABSTRACT

The Jacobi method for iteratively solving a set of linear algebraic equations is well known. However, it suffers from the drawback that it does not converge for all linear algebraic systems. In science and engineering, one often encounters algebraic systems of the form $Ax = b$ where the matrix A is symmetric and positive-definite. In this paper we show that a slight variant of the Jacobi method, namely the Jacobi overrelaxation (JOR) method, can be made to ensure convergence of the iterative scheme for such matrices. Thus we expand the applicability of JOR iterative methods to *all* symmetric, positive-definite matrices in a manner which is computationally convenient and simple. In particular, we show that this variant scheme can be explicitly determined without the need for any eigenvalue computations. We show that if eigenvalue computations are employed, one can obtain the maximum rate of convergence by a proper choice of parameters for the JOR scheme. This maximum rate of convergence is also explicitly obtained.

I. INTRODUCTION

There is an exhaustive literature on iterative methods for solving algebraic systems of linear equations (see, e.g., References [1, 2]). The solution of a linear system of equations express by the relation [1]

$$Ax = b \tag{1}$$

is provided, using the Jacobi decomposition of the matrix A, by the iterative scheme described by

$$Dx^{(n+1)} = b + Nx^{(n)}, \tag{2}$$

DOI:10.1016/0096-3003(92)90063-7
655 Avenue of the Americas, New York, NY 10010
0096-3003/92/$03.50

where the matrix A is decomposed as

$$A = D - N. \tag{3}$$

The matrix A is taken to be of dimension $n \times n$. The matrix D is a diagonal matrix whose diagonal elements are the same as those of A. It is well known that this scheme converges if and only if the spectral radius, $\rho(D^{-1}N)$, is less than unity (see, e.g., Reference [1]). This condition is not satisfied by all positive-definite, symmetric matrices, thus making the direct Jacobi method inapplicable for obtaining the solution to all such systems of equations. We show in this paper that by using a slight variant of the decomposition shown in Equation (3), namely by using the JOR scheme [2], we can guarantee convergence for *all* positive-definite, symmetric matrices A. In addition, we show that this variant is easy to implement computationally.

Linear systems of equations with positive-definite, symmetric matrices A are common in many areas of science and engineering. A significant area of application is the analysis of the vibratory behavior of structural and mechanical systems [3, 4]. While iterative schemes, like Gauss-Seidel and SOR, may be used in such problems, Jacobi-type methods hold out a tremendous advantage in that they lend themselves to improved physical interpretations of the dynamic vibratory behavior of such systems [4]. This is because Jacobi-type schemes can be interpreted in terms of the uncoupled modal responses of dynamic systems, and therefore they provide additional physical insights (which other schemes like SOR do not) into the dynamic response of structural and mechanical systems. It is for this reason that we concentrate in this paper on Jacobi-type iterative methods. Our aim here is to develop Jacobi-type iterative methods which are applicable to *all* positive-definite, symmetric matrices.

II. THE JACOBI OVERRELAXATION SCHEME

Consider an alternative decomposition of the matrix A as [2]

$$A = \alpha D - \{N - (1-\alpha)D\} =: D_1 - N_1, \tag{4}$$

where D is again a diagonal matrix whose diagonal entries are the same as those of A, and α is a parameter, as yet unspecified. Consider now the iterative scheme, JOR, given by

$$D_1 x^{(n+1)} = b + N_1 x^{(n)}. \tag{5}$$

Once again, convergence occurs if and only if the spectral radius, $\rho(D_1^{-1}N_1)$, is less than unity [1]. In this paper we show that we can simply and explicitly obtain values of the parameter α for which the JOR scheme represented by equations (4) and (5) always converges when A is symmetric and positive-definite.

The JOR scheme is described in Reference [2]. However, we provide here several new results which go beyond those in Reference [2]. In particular, we provide ways of explicitly determining the parameter α without the need for any eigenvalue computations, thus making the method more useful for large matrix computations. In addition, we obtain explicitly the value of α which provides the smallest value of the spectral radius, $\rho(D_1^{-1}N_1)$, along with the value of this radius.

III. CONVERGENCE RESULTS

We start with the following lemma.

LEMMA 1.

(a) $\rho(D_1^{-1}N_1)<1 \Leftrightarrow \rho(D_1^{-1}A)<2$;
(b) $\rho(D^{-1}N)<1 \Leftrightarrow \rho(D^{-1}A)<2$.

PROOF. (a): We first note that the matrix $D_1^{-1}A$ has the same eigenvalues as $D_1^{-1/2}AD_1^{-1/2}$. Thus, the matrix D_1^{-1} is positive definite, because A is positive definite.

Now let λ be any eigenvalue of $D_1^{-1}N_1$, and y the corresponding eigenvector. Since $N_1 = D_1 - A$, we get

$$D_1^{-1}Ay = (1-\lambda)\,y. \tag{6}$$

Thus the eigenvalue of $D_1^{-1}A$, denoted μ, is obtained as

$$\mu = 1-\lambda. \tag{7}$$

If for every eigenvalue λ of $D_1^{-1}N_1$ we have $|\lambda|<1$, this implies [by Equation (7)] that for every eigenvalue μ of $D_1^{-1}A$, $\mu<2$. Also, $\mu>0$, since A is positive-definite.

Conversely, when $0<\mu<2$, Equation (7) requires that $|\lambda|<1$.

(b): The proof follows *mutatis mutandis* from part (a) by replacing D_1 by D and N_1 by N. ■

THEOREM 1. *The iterative JOR scheme converges for all values of* α, *such that*

$$\alpha > \left[\frac{\rho(D^{-1}A)}{2}\right] = \frac{\lambda_{\max}(D^{-1}A)}{2}. \tag{8}$$

PROOF. By Lemma 1 convergence is assured if $\rho(D_1^{-1}A) < 2$. But $\rho(D_1^{-1}A) = (1/\alpha)\rho(D^{-1}A)$. Hence whenever α satisfies the relation (8), convergence is assured. ■

We have thus shown that we can always find a positive number α, which is a function of the matrix A, such that the JOR scheme converges to the exact solution of Equation (1). However, this value of α (which also controls the rate of convergence) depends on the determination of the largest eigenvalue of the matrix $D^{-1}A$. We now provide estimates, which do not need any eigenvalue computation, of this maximal eigenvalue of $D^{-1}A$. We start by denoting the matrix $D^{-1/2}AD^{-1/2}$ by A'.

LEMMA 2. $\rho(D^{-1}A) < n$, *where n is the dimension of the square matrix* A.

PROOF. We note that the eigenvalues of $D^{-1}A$ are all real and positive and identical to those of A'. But the elements of the matrix A' are given by

$$a'_{ij} := [A']_{i,j} = \begin{cases} 1, & i = j, \\ \dfrac{a_{ij}}{\sqrt{a_{ii}a_{jj}}}, & i = j, \end{cases} \qquad \text{where} \quad [A]_{ij} = a_{ij}. \tag{9}$$

Since A is a positive definite matrix, the modulus of each of the off-diagonal terms [see Equation (9)] of A' is, by definition [5], less than unity. Using Gerschgorin's theorem [5], we then find that for any eigenvalue λ of $D^{-1}A$, $|\lambda - 1| < n - 1$. Hence the result. ■

THEOREM 2. *The JOR iterative scheme converges for all symmetric, positive-definite matrices A when $\alpha > n/2$, where n is the dimension of A.*

PROOF. Using Theorem 1 and Lemma 2, the result if obvious. ■

A tighter bound on the spectral radius of the matrix $D^{-1}A$ may be obtained by once again noting that the eigenvalues of $D^{-1}A$ and A' are the same. Denote by $|A'|$ the matrix obtained by taking the absolute value of each element of the matrix A'. We then have the following result.

LEMMA 3. $\rho(|A'|) \leqslant \max_i \sum_{j=1}^{n} |a'_{ij}| =: \gamma.$

PROOF. The result follows directly from Gerschgorin's theorem [5]. ■

LEMMA 4. $\rho(D^{-1}A) \leqslant \gamma < n$, *where γ is defined in Lemma* 3.

PROOF. $\rho(D^{-1}A) = \rho(A') \leqslant \rho(|A'|) \leqslant \gamma < n.$ ■

THEOREM 3. *The JOR iterative scheme always converges for $\alpha > \gamma/2$, where γ is defined in Lemma* 3.

PROOF. Using Lemma 4 and Theorem 1, the result if obvious. ■

The rate of convergence of the scheme proposed here depends on the spectral radius of the matrix $D_1^{-1}N_1 = I - D_1^{-1}A$. We now present a result related to the optimal value of the parameter α such that the spectral radius of this matrix is a minimum.

THEOREM 4. *The spectral radius of the matrix $D_1^{-1}N_1$ is a minimum when the parameter α is given by*

$$\alpha_{\text{opt}} = \frac{\lambda_{\max}(D^{-1}A) + \lambda_{\min}(D^{-1}A)}{2}. \tag{10}$$

The spectral radius for this value of α, *i.e.* $\alpha = \alpha_{\text{opt}}$, *is*

$$\rho_{\min}\left(D_{\mathrm{I}}^{-1}N_{\mathrm{I}}\right) = \frac{\lambda_{\max}\left(D^{-1}A\right) - \lambda_{\min}\left(D^{-1}A\right)}{\lambda_{\max}\left(D^{-1}A\right) + \lambda_{\min}\left(D^{-1}A\right)} \tag{11}$$

PROOF. We note that

$$\lambda_i\left(D_{\mathrm{I}}^{-1}N_{\mathrm{I}}\right) = \lambda_i\left(I - D_{\mathrm{I}}^{-1}A\right) = 1 - \lambda_i\left(D_{\mathrm{I}}^{-1}A\right) = 1 - \frac{\lambda_i\left(D^{-1}A\right)}{\alpha}, \tag{12}$$

where by $\lambda_i(*)$ we mean the ith eigenvalue of $*$. The value of α that causes $\rho(D_{\mathrm{I}}^{-1}N_{\mathrm{I}})$ to be a minimum therefore requires us to find α such that we minimize,

$$\rho_{\min}\left(D_{\mathrm{I}}^{-1}N_{\mathrm{I}}\right) := \min_{\alpha} \max_{i} \left| 1 - \frac{\lambda_i\left(D^{-1}A\right)}{\alpha} \right|. \tag{13}$$

This minimization can be geometrically thought of as a *rescaling* of the eigenvalues $\lambda_i(D^{-1}A)$ by a constant α such that the rescaled extreme outliers among them lie closest to unity. Noting that the eigenvalues of $D^{-1}A$ are always positive, this can be achieved uniquely by placing unity at the center of the rescaled interval which these eigenvalues span. This requires that we satisfy the following condition:

$$\frac{\lambda_{\max}\left(D^{-1}A\right)}{\alpha_{\text{opt}}} - 1 = 1 - \frac{\lambda_{\min}\left(D^{-1}A\right)}{\alpha_{\text{opt}}}. \tag{14}$$

From this relation, the result in Equation (10) follows. We note in passing that α_{opt} always satisfies (8).

Again, looking at it geometrically, for $\alpha = \alpha_{\text{opt}}$ the spectral radius is obtained as

$$\rho\left(D_{\mathrm{I}}^{-1}N_{\mathrm{I}}\right) = \max_{i} \left| 1 - \frac{\lambda_i\left(D^{-1}A\right)}{\alpha_{\text{opt}}} \right|. \tag{15}$$

The maximum of the right-hand side, because of our rescaling, will occur for both $\lambda_i(D^{-1}A) = \lambda_{\max}(D^{-1}A)$ and $\lambda_i(D^{-1}A) = \lambda_{\min}(D^{-1}A)$. Using these values in Equation (15), we obtain the result given in Equation (11). ■

Thus the fastest convergence occurs when the value of α is the arithmetic mean of the highest and lowest eigenvalues of $D^{-1}A$. The spectral radius then is always less than unity and is the ratio of the difference of the highest and lowest eigenvalues to their sum.

It should be pointed out that for large $(n \sim 1000)$ sparse matrices A, which often arise in engineering applications, good estimates of the highest and lowest eigenvalues of $D^{-1/2}AD^{-1/2}$ can be found very efficiently. For example, one can use the conjugate-gradient method to obtain a much lower-dimensional tridiagonal matrix, from which estimates of these eigenvalues can then be rapidly determined.

IV. A NUMERICAL EXAMPLE

Consider the positive-definite, symmetric matrix A given by

$$A = \begin{bmatrix} 0.3 & 0.2 & 0.3 & 0.4 & 0.2 \\ 0.2 & 1.0 & 1.0 & 0.2 & 0.3 \\ 0.3 & 1.0 & 1.5 & 0.4 & 0.2 \\ 0.4 & 0.2 & 0.4 & 1.0 & 0.1 \\ 0.2 & 0.3 & 0.2 & 0.1 & 0.5 \end{bmatrix}.$$

The spectral radius $\rho(D^{-1}A) = 2.713$; the smallest eigenvalue of $D^{-1}A$ is 0.1167. Thus by Lemma 1, the standard Jacobi iterative method would fail to converge. However, the JOR scheme will converge, according to Theorem 1, for all $\alpha > 1.357$. We note, however, that application of Theorem 1 requires knowledge of the largest eigenvalue of $D^{-1}A$. The estimates provided by Theorems 2 and 3, which avoid such eigenvalue computations, yield $\alpha > 2.5$ and $\alpha > 1.58$ respectively. The optimal value of α for which the rate of convergence is maximal is found, using Theorem 4, to be 1.415.

In Figure 1, the parameter $\beta = \alpha / \rho(D^{-1}A)$ is plotted versus the spectral radii of $D_1^{-1}N_1$ (solid line) and $D_1^{-1}A$ (dashed line). We note that when β has a value of 0.5, the spectral radius of $D_1^{-1}N_1$ is unity. With increasing β, the spectral radius decreases and then increases, as expected, back to unity as $\beta \to \infty$. On the other hand, the spectral radius of $D_1^{-1}A$ (dashed line) is seen to decrease monotonically with increasing β, again as expected. The value of β when the spectral radius of $D_1^{-1}N_1$ is a minimum is obtained, using Theorem 4, as 0.5215; the spectral radius for this value of β is found, by Equation (11), to be 0.9174. The numerical computations performed show that this optimal value of β, as also the spectral

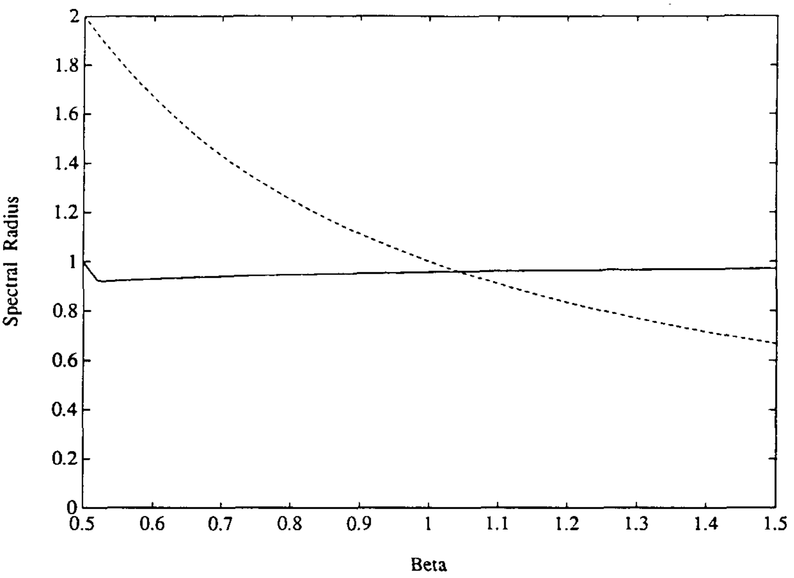

FIG. 1. The spectral radii $\rho(D_I^{-1}N_1)$ and $\rho(D_I^{-1}A)$ as functions of β. The radius $\rho(D_I^{-1}A)$, shown in the dashed line, monotonically decreases with increasing β, while $\rho(D_I^{-1}N_1)$, shown in the solid line, decreases first and then asymptotically approaches unity for large values of β. The optimal value of β for which the convergence is fastest is 0.5125.

radius corresponding to this value of β, are both correctly predicted by Theorem 4.

IV. CONCLUSIONS

In this paper we have investigated the JOR scheme for solving algebraic linear systems of equations with symmetric, positive-definite matrices. In essence, the scheme employs a decomposition of the matrix A as $A = \alpha D - N_1$. The scheme is shown to converge to the exact solution of Equation (1) for *all* positive-definite and symmetric matrices A as long as the parameter α is larger than half the largest eigenvalue of $D^{-1}A$. Next, we provide an estimate of this largest eigenvalue which avoids any eigenvalue computations. We thereby require α to be greater than $\gamma/2$, where the parameter γ is easy to compute directly from matrix A. Thus our results extend the applicability of JOR iterative methods to the class of all positive-definite, symmetric matrices by explicitly providing values of the parameter α, which are simple to obtain from the matrix A and which ensure convergence.

If eigenvalue computations are permitted, we can explicitly determine the optimal value of α which ensures the fastest convergence. This value of α turns out to be the arithmetic mean of the highest and the lowest eigenvalues of the matrix $D^{-1}A$. The rate of convergence for this value of α is also explicitly provided in this paper.

REFERENCES

1. R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, 1962.
2. D. M. Young, *Iterative Solution of Large Linear Systems*, Academic, 1971.
3. F. E. Udwadia, Some Analytical Results on the Iterative Approaches to the Determination of the Response of Nonclassically Damped Systems with Symmetric, Positive Definite Damping Matrices, Univ. of Southern California Report, Dept. of Mechanical Engineering, Dec. 1989.
4. F. E. Udwadia and R. Esfandiari, Nonclassically damped dynamic systems: An iterative approach, *J. Appl. Mech.* 57:423–433 (June 1990).
5. B. Noble, *Applied Linear Albegra*, Prentice-Hall, 1969.