AN ITERATIVE METHOD FOR LARGE SPARSE LINEAR SYSTEMS ON A VECTOR COMPUTER

V. RUGGIERO and E. GALLIGANI
Department of Mathematics, University of Ferrara, Ferrara, Italy

(Received 16 June 1989)

Abstract—In this paper we consider the arithmetic mean method for solving large sparse systems of linear equations. This iterative method converges for systems with coefficient matrices that are symmetric positive definite or positive real or irreducible L-matrices with a strong diagonal dominance. The method is very suitable for parallel implementation on a multiprocessor system, such as the CRAY X-MP. Some numerical experiments on systems resulting from the discretization, by means of the usual 5-point difference formulae, of an elliptic partial differential equation are presented.

1. THE ARITHMETIC MEAN METHOD

In many physical applications, one must solve an $n \times n$ system of linear algebraic equations

$$Ax = b,$$

where $A$ arises from a finite difference approximation to an elliptic partial differential equation. For this reason, the matrix $A$ is extremely large and sparse. We can express the matrix $A = (a_{ij})$ as the matrix sum

$$A = L + D + U,$$

where $D = \text{diag}\{a_{11}, a_{22}, \ldots, a_{nn}\}$ and $L$ and $U$ are respectively, strictly lower and upper triangular matrices.

If we remember that there exists a formal correspondence between the parabolic difference equations and the iterative methods for solving elliptic difference equations, the results in Refs [1, 2] suggest the following iterative method, called the method of the arithmetic mean, for approximating the solution $x$ of system (1):

$$
\begin{align*}
(L + (D + \rho W))x^{(1)} &= (\rho W - U)x^{(k)} + b, \\
((D + \rho W) + U)x^{(2)} &= (\rho W - L)x^{(k)} + b,
\end{align*}
$$

$$x^{(k+1)} = \frac{1}{2}(x^{(1)} + x^{(2)}),$$

where $x^{(0)}$ is an initial vector approximation to $x$, $\rho$ is a positive parameter and $W$ is a positive diagonal matrix.

Method (3) is characterized by having within its overall mathematical structure certain well-defined substructures that can be executed simultaneously. This feature makes method (3) ideally suited for implementation on a multiprocessor system with two or more vector processors; the lower triangular system and the upper triangular system in method (3) can be solved simultaneously on two different processors.

One set of conditions which guarantees the convergence of the iterative method (3) is described in the following theorems [1, 2].

Theorem 1

Let $A = (a_{ij})$ be an $n \times n$ irreducibly diagonally dominant real matrix with $a_{ii} \leq 0$, for all $i \neq j$ and $a_{ii} > 0$, $i = 1, 2, \ldots, n$. Let $W$ be a positive diagonal matrix. Then, the iterative method (3) is convergent for all $\rho > 0$.  

25
Proof. By hypothesis, \( A \) is a non-singular \( M \)-matrix \([3, \text{p. } 110]\). Since the matrices \( H_1 = L + (D + \rho W) \) and \( H_2 = (D + \rho W) + U \) are strictly diagonally dominant matrices with positive entries on the diagonal for \( \rho > 0 \) and with non-positive off-diagonal elements, then \([4, \text{Theorems } 1.8, 3.4, 3.10]\) \( H_1 \) and \( H_2 \) are non-singular \( M \)-matrices. The matrices \( K_1 = \rho W - U \) and \( K_2 = \rho W - L \) are non-negative. Since \( H_1 - K_1 = H_2 - K_2 = A \), we can write \( Q = \frac{1}{2} H_1^{-1} K_1 + \frac{1}{2} H_2^{-1} K_2 = I - \left( \frac{1}{2} H_1^{-1} + \frac{1}{2} H_2^{-1} \right)A \) or \( \frac{1}{2} H_1^{-1} + \frac{1}{2} H_2^{-1} = (I - Q)A^{-1} \). Now, the proof of the theorem runs parallel to a standard proof given in Ref. \([3, \text{p. } 119]\).

Theorem 2

Let \( A \) be a real \( n \times n \) matrix expressed in form (2). Assume that the symmetric matrix \( M = A + A^T \) is positive definite. If

\[
\rho_* = \max \left\{ \frac{|\mu_{\text{min}}|}{\lambda_{\text{min}}}, \frac{|\nu_{\text{min}}|}{\lambda_{\text{min}}} \right\}
\]

where \( \lambda_{\text{min}} \) is the smallest eigenvalue of \( M \), \( \mu_{\text{min}} \) is the algebraically smallest eigenvalue of \( M_1 = (D + U)(D + U)^T - LL^T \) and \( \nu_{\text{min}} \) is the algebraically smallest eigenvalue of \( M_2 = (L + D)(L + D)^T - UU^T \), then the iterative method (3) is convergent for all \( \rho > \rho_* \) and \( W = I \).

Proof. Since \( 0 < \lambda_{\text{min}} \leq z^T M z / z^T z \) and \( \mu_{\text{min}} \leq z^T M_1 z / z^T z \), \( \nu_{\text{min}} \leq z^T M_2 z / z^T z \), for any \( z \neq 0 \), we have for \( \rho > 0 \)

\[
z^T (\rho M + M_1) z \geq (\rho \lambda_{\text{min}} + \mu_{\text{min}}) z^T z,
\]

\[
z^T (\rho M + M_2) z \geq (\rho \lambda_{\text{min}} + \nu_{\text{min}}) z^T z.
\]

Thus, it is a simple matter to verify that for \( \rho > \rho_* \) the symmetric matrices \( \rho M + M_1 \) and \( \rho M + M_2 \) are positive definite. The matrix \( \rho M + M_1 \) may be written in the form

\[
\rho M + M_1 = (D + U + \rho I)Z_1(D + U + \rho I)^T,
\]

where

\[
Z_1 = \left( (D + U + \rho I)^{-1} A \right) + \left( (D + U + \rho I)^{-1} A \right)^T - \left( (D + U + \rho I)^{-1} A \right) \left( (D + U + \rho I)^{-1} A \right)^T.
\]

In the same way we have

\[
\rho M + M_2 = (D + L + \rho I)Z_2(D + L + \rho I)^T,
\]

where

\[
Z_2 = \left( (D + L + \rho I)^{-1} A \right) + \left( (D + L + \rho I)^{-1} A \right)^T - \left( (D + L + \rho I)^{-1} A \right) \left( (D + L + \rho I)^{-1} A \right)^T.
\]

Now, if \( Q_1 = (U + D + \rho I)^{-1}(\rho I - L) \) and \( Q_2 = (L + D + \rho I)^{-1}(\rho I - U) \), we have \( Q_1 Q_1^T = I - Z_1 \) and \( Q_2 Q_2^T = I - Z_2 \).

Since the matrices \( \rho M + M_1 \) and \( \rho M + M_2 \) are positive definite the spectral norm of \( Q_1, Q_2 \) is less than unity; thus, the iterative method (3) is convergent.

Theorem 3

Let \( A \) be a real \( n \times n \) symmetric positive definite matrix. Let \( A \) be expressed in form (2), \( A = L + D + L^T \), where \( D \) is a positive diagonal matrix and \( L \) is a lower triangular matrix. Let \( W \) be a positive diagonal matrix. Then, the iterative method (3) is convergent for all \( \rho > 0 \).

Proof. We write \( H_1 = L + (D + \rho W) \), \( K_1 = \rho W - L^T \), \( H_2 = (D + \rho W) + L^T \) and \( K_2 = \rho W - L \); then, we have \( H_1 - K_1 = H_2 - K_2 = H_1^{-1} - K_1^{-1} = H_2^{-1} - K_2^{-1} = A \). Since the matrix \( (i = 1, 2) \)

\[
P = \frac{1}{2} [(H_i + K_i) + (H_i + K_i)^T] = D + 2 \rho W,
\]

is symmetric positive definite, the theorem follows as a consequence of a standard proof of convergence for \( P \) — regular splittings \([3, \text{p. } 123; 5]\).
When \( A \) is a symmetric positive definite matrix a special form of method (3) is
\[
(D + \omega L)x^{(1)} = ((1 - \omega)D - \omega L^T)x^{(k)} + \omega b, \\
(D + \omega L^T)x^{(2)} = ((1 - \omega)D - \omega L)x^{(k)} + \omega b,
\]
\[
x^{(k+1)} = \frac{1}{2}(x^{(1)} + x^{(2)}).
\]

A proof identical to that given above for Theorem 3 ensures that the iterative method (4) is convergent for \( 0 < \omega < 2 \). If, in addition, \( A \) is "2-cyclic", it is possible to derive an optimal value of \( \omega \) [6].

2. NUMERICAL EXAMPLES

In the following examples we apply method (3) to two algebraic systems which arise from an elliptic boundary value problem and from an eigenvalue problem, respectively.

**Example 1**

Consider the elliptic partial differential equation
\[
L[\phi] = -\phi_{xx} - \phi_{yy} + \alpha \phi_x + \beta \phi_y + \gamma \phi = f, \quad \text{in } \Omega,
\]
with boundary conditions \( \phi = g \) on \( \partial \Omega \). Here, \( \Omega = (0, 1) \times (0, 1) \), \( \alpha \) and \( \beta \) are given positive constants, \( \gamma = 1/x^2 + 1/\beta^2 \) and the functions \( f = f(xy) \) and \( g = g(xy) \) are chosen so that the exact solution is \( \phi(xy) = e^{-x^2} e^{-y^2} \). This problem is discretized by the second order accurate finite difference method with mesh spacings in both directions equal to \( h = 1/(m + 1) \), where \( m = 100 \). If \( h \alpha < 2 \) and \( h \beta < 2 \), the resulting coefficient matrix \( A \), of order \( n = m^2 \), satisfies the conditions of Theorems 1 and 2. Besides, \( A \) is an M-matrix such that each element in the upper triangular part is greater than or equal to the respective element in the lower triangular part; that is, the matrix \( A \) has property \( N \) [7].

This is an example which allows us to study the influence of asymmetry on the solution behaviour of the iterative methods.

We have solved on the multivector computer CRAY X-MP/48 the resulting linear system for different values of \( \alpha \) and \( \beta \), using the arithmetic mean method (3) (with \( W = I \) and \( \rho = 1 \)) and the biconjugate gradient method [8].

Results appear in Table 1; \( k^\ast \) is the number of iterations for an error \( \|x^{(k^\ast)} - x\|_\infty \) given in the last column of the table (\( x^{(0)} = 0 \)).

This example illustrates a situation for which the arithmetic mean method (3) produces a significantly better convergence behaviour than the biconjugate gradient method for strongly asymmetric problems; method (3) is less efficient than the biconjugate gradient method for nearly symmetric problems.

**Example 2**

In the study of "essentially positive" dynamic systems that arise in solving the initial-boundary value problem for the diffusion–convection equation by the method of lines using centred—in space—difference equations, one must determine the eigenvalue of maximal modulus of a positive \( n \times n \) matrix \( A^{-1} \), where \( A \) is large and sparse [9]. Namely, we wish to find the Perron root \( \mu_1 \) and

<table>
<thead>
<tr>
<th>Method</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( k^\ast )</th>
<th>( |x^{(0)} - x|_\infty )</th>
</tr>
</thead>
<tbody>
<tr>
<td>arith. mean</td>
<td>1</td>
<td>1</td>
<td>8926</td>
<td>( 1 \times 10^{-3} )</td>
</tr>
<tr>
<td>bicon. gradient</td>
<td>1</td>
<td>1</td>
<td>349</td>
<td>( 2 \times 10^{-4} )</td>
</tr>
<tr>
<td>arith. mean</td>
<td>10</td>
<td>10</td>
<td>3525</td>
<td>( 1 \times 10^{-4} )</td>
</tr>
<tr>
<td>bicon. gradient</td>
<td>10</td>
<td>10</td>
<td>375</td>
<td>( 1 \times 10^{-4} )</td>
</tr>
<tr>
<td>arith. mean</td>
<td>20</td>
<td>50</td>
<td>750</td>
<td>( 5 \times 10^{-4} )</td>
</tr>
<tr>
<td>bicon. gradient</td>
<td>20</td>
<td>50</td>
<td>518</td>
<td>( 1 \times 10^{-4} )</td>
</tr>
<tr>
<td>arith. mean</td>
<td>50</td>
<td>50</td>
<td>371</td>
<td>( 2 \times 10^{-4} )</td>
</tr>
<tr>
<td>bicon. gradient</td>
<td>50</td>
<td>50</td>
<td>3071</td>
<td>( 3 \times 10^{-4} )</td>
</tr>
<tr>
<td>arith. mean</td>
<td>100</td>
<td>100</td>
<td>234</td>
<td>( 1 \times 10^{-7} )</td>
</tr>
</tbody>
</table>
its associated eigenvector $x_i$ of $A^{-1}$ for a given $A$: $A^{-1}x_i = \mu_i x_i$ with $x_i > 0$ and $\mu_i$ equal to the spectral radius of $A^{-1}$. We apply the power iteration method to find $\mu_1$ and $x_1$. Each step of this method involves the solution for an unknown vector $x$ of the matrix equation $Ax = b$. This linear system is solved by an iterative method. This is an example which allows us to study the effect of the “inner” iteration on convergence of the “outer” iteration to $x_1$; a significant reduction in the total effort can often be achieved by proper coordination of the “inner” and “outer” iterations.

Due to the inability of the current theory to predict the convergence rate achieved by the power method when the iterative method (3) is used for the “inner” iteration, some numerical studies are essential. An $n \times n$ test-matrix $A$ is generated by the discretization on $\Omega = (0, 1) \times (0, 1)$, using second order accurate finite difference equations, of the differential operator $L[\varphi]$ of formula (4). Mesh spacings in both directions are equal to $h = 1/(m + 1)$ where $m = 64$; $\alpha, \beta, \gamma$ in equation (4) are given non-negative constants with $h\alpha < 2$ and $h\beta < 2$. We have $n = m^2$. Results appear in Table 2. In this table $\tau$ indicates the number of “outer” iterations requested to make the max-norm of difference between successive “outer” iterations less than the convergence parameter $\epsilon$ and $itt$ indicates the maximum number of “inner” iterations, per “outer” iteration, requested to make the max-norm of difference between successive “inner” iterations less than a tolerance $\tau$. We have $\alpha = \beta = 10, \gamma = 401, \rho = 1, W = I(\mu_1 = 2.12439 \ldots 10^{-3})$. The initial vector $x(0)$, in the first “outer” iteration, is $(1 0 \ldots 0)^T$.

If we perform enough “inner” iterations per “outer” iteration, we obtain an accurate result. However, the comparison between the case $\tau = 10^{-9}$, $\epsilon = 10^{-5}$ and the case $\tau = 10^{-7}$, $\epsilon = 10^{-6}$ is very demonstrative. If we consider $\alpha = \beta = 0.1, \gamma = 1.04$ the matrix $A$ is nearly symmetric; thus, the number of “inner” iterations is high. For $\tau = 10^{-6}$ and $\epsilon = 10^{-5}$, we have $itt = 1447$ and $it = 9$. The computed eigenvalue is $4.8122 \times 10^{-2}$ ($\mu_1 = 4.81225 \ldots 10^{-2}$).

The computer-time on CRAY X-MP/48 for solving these problems utilizing two processors is $1.6 \times 10^{-3}$ s per each “inner” iteration.

REFERENCES