On the equivalence of SOR, SSOR and USSOR as applied to σ_1-ordered systems of linear equations

By M. S. Lynn*

The method of symmetric successive over-relaxation (SSOR) was proposed by Sheldon (1955). It has been analyzed by Habetler and Wachspress (1961) and extended to the method of unsymmetric successive over-relaxation (USSOR) by D'Sylva and Miles (1964). The latter showed that for suitable choice of relaxation parameters asymptotic rates of convergence precisely half of those of the familiar method of successive over-relaxation (SOR) may be obtained when the method is applied to σ_1-ordered systems of linear equations possessing Property A. The present paper shows that this factor is in fact spurious, and that, under the latter hypotheses, the methods of SOR, SSOR and USSOR are identical when applied to σ_1-ordered systems of equations. It is hence shown that Chebyshev accelerated SSR (SSOR with unity relaxation parameter) becomes, in this case, identical with the Chebyshev accelerated Gauss–Seidel method (Varga (1957)). The theoretical results of D'Sylva and Miles and of this paper and the vastly different behaviours of the σ_1- and σ_2-orderings are emphasized by means of numerical examples.

1. Suppose we wish to solve the system of equations

$$Ax = f, \tag{1.1}$$

where we assume that A is a symmetric, positive definite, $N \times N$ matrix and x, f are $N \times 1$ vectors. We shall assume that A has Property A and is σ_1-ordered (Young, 1954) so that it has the form

$$A = \begin{bmatrix} D_1 & -R \\ -R^T & D_2 \end{bmatrix}, \tag{1.2}$$

where R is an $m \times n$ sub-matrix and D_1, D_2 are $m \times m$ and $n \times n$ diagonal sub-matrices, respectively.

It should be remarked at this juncture, that if A has *block* Property A and is *block* σ_1-ordered (Arms, Gates, Zondek (1956)), then all our subsequent remarks as applied to point relaxation methods will be equally valid for the associated block relaxation methods (block successive over-relaxation, block symmetric successive over-relaxation, and so forth). For in the sequel, it can be simply assumed for this purpose that D_1 and D_2 are *block* diagonal matrices which will not alter the argument.

Now let D, L, U be the partitioned matrices defined by

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \; L = \begin{bmatrix} 0 & 0 \\ R^T & 0 \end{bmatrix}, \; U = \begin{bmatrix} 0 & R \\ 0 & 0 \end{bmatrix}. \tag{1.3}$$

Then the method of successive over-relaxation (SOR) is defined, as usual, by the sequence of vectors $\{y^{(k)}\}$ where

$$Dy^{(k+1)} = \omega(Ly^{(k+1)} + Uy^{(k)} + f) + (1-\omega)Dy^{(k)} \tag{1.4}$$
$$k = 0, 1, 2, \ldots,$$

where $y^{(0)}$ is arbitrary and ω $(0 < \omega < 2)$ is the relaxation parameter.

The method of unsymmetric successive over-relaxation (USSOR) proposed by D'Sylva and Miles (1963) is defined by the sequence of vectors $\{x^{(k)}\}$, where

$$Dx^{(2k+1)} = \omega_1(Lx^{(2k+1)} + Ux^{(2k)} + f) + (1-\omega_1)Dx^{(2k)}, \tag{1.5}$$

$$Dx^{(2k+2)} = \omega_2(Ux^{(2k+2)} + Lx^{(2k+1)} + f) + (1-\omega_2)Dx^{(2k+1)}, \tag{1.6}$$
$$k = 0, 1, 2, \ldots,$$

where $x^{(0)}$ is arbitrary and ω_1, ω_2 are (possibly) distinct relaxation parameters. When $\omega_1 = \omega_2$ the method becomes the symmetric successive over-relaxation (SSOR) method, without Chebyshev acceleration, proposed by Sheldon (1955). It has been shown (Habetler and Wachspress (1961)) that if Chebyshev acceleration is applied to SSOR the resultant method has a rate of convergence which, except under favourable circumstances, is usually far slower than the rate of convergence of optimized SOR with

$$\omega = \omega_b = 2/[1 + \sqrt{(1-\mu^2)}], \tag{1.7}$$

where μ denotes the spectral radius of $B = D^{-1}(L+U)$. It was also shown by D'Sylva and Miles (1963) that if ω_1, ω_2 are chosen such that

$$\omega_{12}\,(\text{say}) = \omega_1 + \omega_2 - \omega_1\omega_2 = \omega_b, \tag{1.8}$$

then USSOR has an optimized asymptotic convergence rate which is precisely half that of SOR, the factor $\frac{1}{2}$ appearing since one complete iteration of USSOR involves exactly twice as much computation as one iteration of SOR.

The purpose of this paper is to show that, in the case where A is σ_1-ordered, this factor of $\frac{1}{2}$ is spurious, since the computation may be organized such that one

* *National Physical Laboratory, Teddington, Middlesex.*

Downloaded from https://academic.oup.com/comjnl/article/7/1/72/558762 by guest on 13 March 2024

complete sweep of USSOR involves precisely the same amount of arithmetic as one sweep of SOR. More surprising, however, is the fact that if one organizes the computation in this manner, USSOR and SOR become *precisely identical methods*, with $\omega = \omega_{12} = \omega_1 + \omega_2 - \omega_1\omega_2$ (hence providing a separate justification for (1.8)). As an immediate corollary, it follows that Chebyshev accelerated SSR ($\omega_1 = \omega_2 = \omega_{12} = 1$)* *becomes precisely identical with the Chebyshev accelerated Gauss–Seidel method* (Varga (1957), Sheldon (1959)) with a suitable choice of an initial vector.

Using Theorem 2 of Varga (1957), it follows that Chebyshev accelerated SSR, even with the economical computational scheme, can converge no faster than optimized SOR, when the σ_1-ordering is employed. It may seem as if this result is at variance with the results of Sheldon (1955) and of Habetler and Wachspress (1961) who indicated that rates of convergence faster than those obtained by optimized SOR may be achieved under certain conditions. It now appears that certainly one of these conditions is that an ordering other than the σ_1-ordering be employed; we shall consider this point further below.

2. Let, in (1.5) and (1.6),

$$\boldsymbol{x}^{(p)} = \begin{bmatrix} \boldsymbol{x}_1^{(p)} \\ \boldsymbol{x}_2^{(p)} \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{bmatrix} \quad p = 0, 1, 2, \ldots,$$

where $\boldsymbol{x}_1^{(p)}, \boldsymbol{f}_1$ are $m \times 1$ vectors, and $\boldsymbol{x}_2^{(p)}, \boldsymbol{f}_2$ are $n \times 1$ vectors. Then from (1.3), (1.5) and (1.6), we have for $k = 0, 1, 2, \ldots$,

$$\boldsymbol{D}_1\boldsymbol{x}_1^{(2k+1)} = \omega_1(\boldsymbol{R}\boldsymbol{x}_2^{(2k)} + \boldsymbol{f}_1) + (1 - \omega_1)\boldsymbol{D}_1\boldsymbol{x}_1^{(2k)}, \quad (2.1)$$

$$\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+1)} = \omega_1(\boldsymbol{R}^T\boldsymbol{x}_1^{(2k+1)} + \boldsymbol{f}_2) + (1 - \omega_1)\boldsymbol{D}_2\boldsymbol{x}_2^{(2k)}, \quad (2.2)$$

$$\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+2)} = \omega_2(\boldsymbol{R}^T\boldsymbol{x}_1^{(2k+1)} + \boldsymbol{f}_2) + (1 - \omega_2)\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+1)}, \quad (2.3)$$

$$\boldsymbol{D}_1\boldsymbol{x}_1^{(2k+2)} = \omega_2(\boldsymbol{R}\boldsymbol{x}_2^{(2k+2)} + \boldsymbol{f}_1) + (1 - \omega_2)\boldsymbol{D}_1\boldsymbol{x}_1^{(2k+1)}. \quad (2.4)$$

Let

$$\omega_{12} = \omega_1 + \omega_2 - \omega_1\omega_2; \; \omega_1' = \omega_{12}/\omega_1; \; \omega_2' = \omega_{12}/\omega_2. \quad (2.5)$$

From (2.2) and (2.3)

$$\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+2)} = (\omega_2/\omega_1)\boldsymbol{D}_2(\boldsymbol{x}_2^{(2k+1)} - (1 - \omega_1)\boldsymbol{x}_2^{(2k)}) + (1 - \omega_2)\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+1)},$$

so that

$$\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+2)} = \omega_1'\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+1)} + (1 - \omega_1')\boldsymbol{D}_2\boldsymbol{x}_2^{(2k)} \quad (k = 0, 1, 2, \ldots). \quad (2.6)$$

Similarly, from (2.1) and (2.4) (with k replaced by $k - 1$ in the latter),

$$\boldsymbol{D}_1\boldsymbol{x}_1^{(2k+1)} = \omega_2'\boldsymbol{D}_1\boldsymbol{x}_1^{(2k)} + (1 - \omega_2')\boldsymbol{D}_1\boldsymbol{x}_1^{(2k-1)} \quad (k = 1, 2, \ldots). \quad (2.7)$$

* Without Chebyshev acceleration, SSR is the same as the "double-sweep" method originally proposed by Aitken (1950).

Hence from (2.2) and (2.6)

$$\boldsymbol{D}_2\boldsymbol{x}_2^{(2k+2)} = \omega_{12}(\boldsymbol{R}^T\boldsymbol{x}_1^{(2k+1)} + \boldsymbol{f}_2) + (1 - \omega_{12})\boldsymbol{D}_2\boldsymbol{x}_2^{(2k)}, \quad (k = 0, 1, 2, \ldots), \quad (2.8)$$

whilst from (2.4) and (2.7) (with $k - 1$ replacing k in the former),

$$\boldsymbol{D}_1\boldsymbol{x}_1^{(2k+1)} = \omega_{12}(\boldsymbol{R}\boldsymbol{x}_2^{(2k)} + \boldsymbol{f}_1) + (1 - \omega_{12})\boldsymbol{D}_1\boldsymbol{x}_1^{(2k-1)}, \quad (k = 1, 2, \ldots). \quad (2.9)$$

Now let

$$\boldsymbol{y}^{(k)} = \begin{bmatrix} \boldsymbol{x}_1^{(2k-1)} \\ \boldsymbol{x}_2^{(2k)} \end{bmatrix} \; (k = 0, 1, 2, \ldots). \quad (2.10)$$

Then from (1.3), (2.8) and (2.9),

$$\boldsymbol{D}\boldsymbol{y}^{(k+1)} = \omega_{12}(\boldsymbol{L}\boldsymbol{y}^{(k+1)} + \boldsymbol{U}\boldsymbol{y}^{(k)} + \boldsymbol{f}) + (1 - \omega_{12})\boldsymbol{D}\boldsymbol{y}^{(k)} \quad (k = 0, 1, 2, \ldots). \quad (2.11)$$

which is *precisely* the sequence of vectors we should obtain if we were to apply the method of SOR with starting vector

$$\boldsymbol{y}^{(0)} = \begin{bmatrix} \boldsymbol{x}_1^{(1)} \\ \boldsymbol{x}_2^{(2)} \end{bmatrix} \quad (2.12)$$

and relaxation factor $\omega = \omega_{12}$.

It is thus apparent that, in the method of USSOR and hence SSOR, it is only necessary to calculate the sequences of vectors $\{\boldsymbol{x}_1^{(2k+1)}\}$ and $\{\boldsymbol{x}_2^{(2k)}\}$ from (2.8) and (2.9) which converge to $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively, where

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix},$$

and that, in so economizing the computation, the method is *identical* with SOR.

It may readily be shown that, if $\omega_{12} \neq 1$, $\boldsymbol{x}_1^{(1)}$ and $\boldsymbol{x}_2^{(2)}$ may be chosen arbitrarily and independently by appropriately choosing the arbitrary starting vectors $\boldsymbol{x}_1^{(0)}$ and $\boldsymbol{x}_2^{(0)}$. On the other hand, if $\omega_{12} = 1$, so that $\omega_1 = \omega_2 = 1$, then $\boldsymbol{x}_1^{(1)}$ may be chosen arbitrarily, whilst

$$\boldsymbol{x}_2^{(2)} = \boldsymbol{D}_2^{-1}[\boldsymbol{R}^T\boldsymbol{x}_1^{(1)} + \boldsymbol{f}_2].$$

This is exactly the situation with the Gauss–Seidel iteration scheme (SOR with $\omega = 1$), and thus we see that Chebyshev accelerated SSR is identical with Chebyshev accelerated Gauss–Seidel.

The author concurs with the results of D'Sylva and Miles (1963) who showed that σ_1-ordered SSOR is optimized when $\omega_1 = \omega_2 = 1$. For when $\omega_1 = \omega_2 = \bar{\omega}$ (say), then $\omega_{12} = \bar{\omega}(2 - \bar{\omega}) \leqslant 1$; thus from (2.11) SSOR is equivalent to SOR with $\omega = \bar{\omega}(2 - \bar{\omega}) \leqslant 1$, and is thus optimized when $\bar{\omega}(2 - \bar{\omega}) = 1$, and hence when $\omega_1 = \omega_2 = 1$ which proves our assertion.

This is perhaps surprising, for it is known that for the σ_2- or normal ordering (Young (1954)) it is possible under favourable circumstances for the optimal value of ω to be greater than unity (see Sheldon (1955), Habetler and Wachspress (1961) or the numerical experiments of Evans and Forrington (1963)) and that Chebyshev accelerated SSOR using this optimal value

Downloaded from https://academic.oup.com/comjnl/article/7/1/72/558762 by guest on 13 March 2024

of ω yields convergence rates superior to optimized SOR. Thus it appears that for SSOR, asymptotic rates of convergence may be greatly affected by different consistent orderings, which is unlike the situation for SOR.

To summarize the results thus far:

Theorem 1. *If A in* (1.1) *has* (*block*) *Property A and is* (*block*) σ_1*-ordered, then* (*block*) *USSOR and* (*block*) *SOR are identical with*

$$\omega = \omega_{12} = \omega_1 + \omega_2 - \omega_1\omega_2.$$

If $\omega_1 = \omega_2 = \bar{\omega}$ (*say*), *then* (*block*) *SSOR and* (*block*) *SOR are identical with*

$$\omega = \bar{\omega}(2 - \bar{\omega}) \leqslant 1.$$

If $\bar{\omega} = 1$, *so that* (*block*) *SSOR becomes* (*block*) *SSR, the latter is identical with* (*block*) *Gauss–Seidel. Hence Chebyshev accelerated SSR is identical with the Chebyshev accelerated Gauss–Seidel method.*

Now Varga ((1957), Theorem 2) proved that optimized SOR converges at least as fast as Chebyshev-accelerated Gauss–Seidel. Thus, using Theorem 1, we may state:

Theorem 2. *If A is symmetric positive-definite, has Property A and is* σ_1*- ordered, then optimized SOR converges at least as fast as Chebyshev-accelerated SSR and at the same rate as optimized USSOR when the economical scheme of* (2.7) *and* (2.8) *is employed. If the usual formulation of* (1.5) *and* (1.6) *is utilized, then optimized SOR converges at least twice as fast as the other two methods.*

It might be remarked that the requirement that A be symmetric, positive-definite may be replaced by the weaker requirement that it has the form

$$A = \begin{bmatrix} D_1 & K \\ M & D_2 \end{bmatrix}$$

where D_1, D_2 are (block) diagonal, and that the matrix

$$B = \begin{bmatrix} 0 & D_1^{-1}K \\ D_2^{-1}M & 0 \end{bmatrix}$$

has real eigenvalues and spectral radius less than unity.

3. Numerical example

We should thus expect that the behaviour of the SSOR method, with or without Chebyshev acceleration, is different according to whether it is applied to σ_1- or σ_2-ordered systems of equations (see Young (1954): for example, in solving the two-dimensional Dirichlet problem over the rectangle by finite-difference methods using a five-point star, the σ_2-ordering may be obtained (say) by sweeping through the mesh consecutively row by row, proceeding from left to right along each row. The σ_1-ordering, on the other hand, can be obtained by first processing the equations belonging to the (i, j)th-points such that $(i + j)$ is even, and then processing those equations belonging to points such that $(i + j)$ is odd.)

This difference is indeed borne out by a simple numerical experiment, where the solution of the set of

Table 1. Comparison of spectral radii
(See text for explanation)

ω \ n	10			20		
	$\lambda_1(\omega)$	$\lambda_2(\omega)$	$\lambda(\omega)$	$\lambda_1(\omega)$	$\lambda_2(\omega)$	$\lambda(\omega)$
0·2	0·9824	0·9910	0·9910	0·9951	0·9974	0·9976
0·4	0·9624	0·9800	0·9799	0·9896	0·9943	0·9944
0·6	0·9424	0·9663	0·9679	0·9839	0·9899	0·9904
0·8	0·9267	0·9489	0·9469	0·9795	0·9853	0·9852
1·0	0·9206	0·9268	0·9206	0·9778	0·9780	0·9788
1·2	0·9267	0·8997	0·8803	0·9795	0·9685	0·9667
1·4	0·9424	0·8716	0·8060	0·9839	0·9545	0·9477
1·6	0·9624	0·8616	0·6000	0·9896	0·9357	0·9057
1·8	0·9824	0·9060	0·8000	0·9951	0·9293	0·8000

equations derived from a one-dimensional Dirichlet problem is obtained. In Table 1 we give the spectral radii of the iteration matrices of order 10×10 and 20×20 as follows:

(i) $\lambda_1(\omega)$ = the spectral radius of the iteration matrix of SSOR, σ_1-ordering.
(ii) $\lambda_2(\omega)$ = the square root (twice as much computation per iteration) of the spectral radius of the iteration matrix of SSOR, σ_2-ordering.
(iii) $\lambda(\omega)$ = the spectral radius of the SOR iteration matrix.

It will be seen that $\lambda_1(\omega)$ behaves according to the predicted theory with a minimum at $\omega = 1$, whereas the smallest value of $\lambda_2(\omega)$ is attained for a value of ω considerably greater than unity.

It is curiously interesting to calculate the asymptotic rates of convergence of Chebyshev accelerated SSOR. These are given in Table 2 together with the asymptotic rate of convergence for SOR. According to Theorem 2, the latter should be the same as the rate of convergence for Chebyshev accelerated SSOR, σ_1-ordering, with $\omega = 1$. The rate of convergence for Chebyshev accelerated SSOR, σ_2-ordering, using the optimal value,

Table 2. Optimized rates of convergence

METHOD	$n = 10$	$n = 20$
Chebyshev accelerated SSOR, σ_1-ordering ($\omega = 1$)	0·580	0·301
Chebyshev accelerated SSOR, σ_2-ordering ($\omega = \omega'$)	0·564 $\left\{\begin{matrix}\omega' = 1\cdot57 \\ \lambda_2(\omega') = 0\cdot8597\end{matrix}\right\}$	0·399 $\left\{\begin{matrix}\omega' = 1\cdot75 \\ \lambda_2(\omega') = 0\cdot9255\end{matrix}\right\}$
Optimized SOR ($\omega = \omega_b$)	0·580 ($\omega_b = 1\cdot5603$)	0·301 ($\omega_b = 1\cdot7406$)

Downloaded from https://academic.oup.com/comjnl/article/7/1/72/558762 by guest on 13 March 2024

Table 3. Ratios (θ) of asymptotic rates of convergence of Chebyshev accelerated SSOR (σ_2) to optimized SOR

n \ θ	EXPERIMENTAL	THEORETICAL (SHELDON)
10	0·97	0·94
20	1·33	1·29

ω', which minimizes $\lambda_2(\omega)$ is given by

$$R_2 = -\tfrac{1}{2}\log\left[\alpha + \sqrt{(\alpha^2 - 1)}\right], \qquad (3.1)$$

where

$$\alpha = 2/[\lambda_2(\omega')]^2 - 1 \qquad (3.2)$$

(see Sheldon (1955); Varga (1959)); we note that the spectral radius, $[\lambda_2(\omega')]^2$, appears in (3.2), and not its square root; the fact that twice as much computation is required per iteration is accounted for by the factor $\frac{1}{2}$ in (3.1).

We might remark that the ratio, θ, of the asymptotic rates of convergence of the Chebyshev accelerated SSOR (σ_2) method to optimized SOR is in keeping with the theorétical result of Sheldon ((1955), p. 109, equation (63)) which states that

$$\theta = \frac{1}{2}\sqrt{\frac{n}{\pi}}. \qquad (3.3)$$

This is illustrated in Table 3. θ is less than unity for $n = 10$ and greater than unity for $n = 20$. For this *particular* problem, $\theta \to \infty$ as $n \to \infty$, and accelerated SSOR is clearly superior.

The asymptotic rates of convergence are illustrated graphically in Figs. 1 and 2.

The author would like to acknowledge with gratitude Dr. D. W. Martin of the National Physical Laboratory for several stimulating conversations and helpful suggestions.

The above work was carried out as part of the research programme of the National Physical Laboratory, and is published by permission of the Director of the Laboratory.

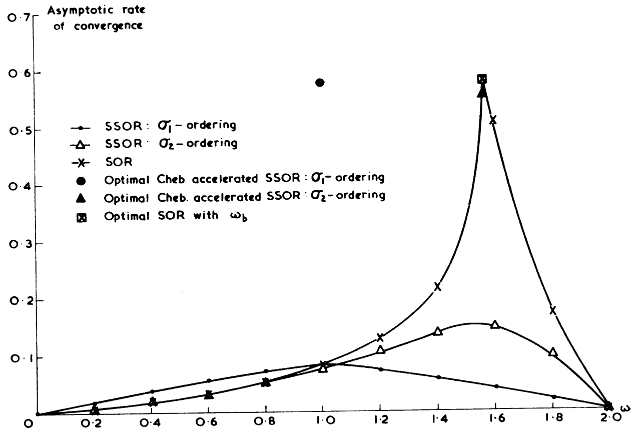

Fig. 1.—Asymptotic rates of convergence: n = 10

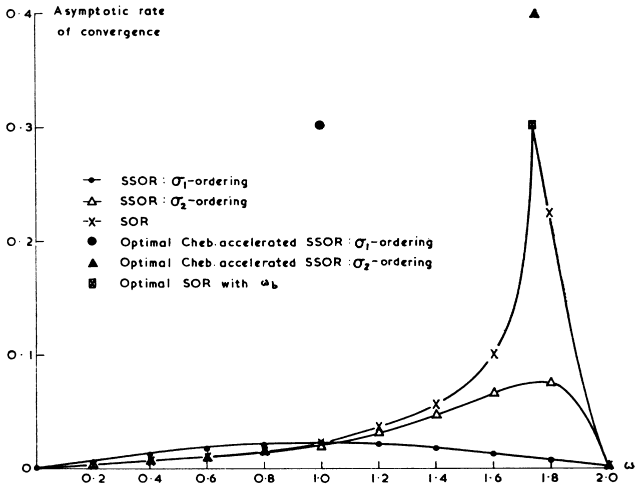

Fig. 2.—Asymptotic rates of convergence: n = 20

References

Aitken, A. C. (1950). "Studies in practical mathematics V. On the iterative solution of a system of linear equations," *Proc. Roy. Soc. Edinburgh* A, Vol. 63, p. 52.

Arms, R. J., Gates, L. D., and Zondek, B. (1956). "A method of block iteration," *J. Soc. Indust. Appl. Math.*, Vol. 4, p. 220.

D'Sylva, E., and Miles, G. A. (1964). "The S.S.O.R. iteration scheme for equations with σ_1-ordering," *The Computer Journal*, Vol. 6, p. 366.

Evans, D. J., and Forrington, C. V. D. (1963). "An iterative method for optimizing symmetric successive over-relaxation," *The Computer Journal*, Vol. 6, p. 271.

Habetler, G. J., and Wachspress, E. L. (1961). "Symmetric successive over-relaxation in solving diffusion difference equations," *Math. Comp.*, Vol. 15, p. 356.

Sheldon, J. W. (1955). "On the numerical solution of elliptic difference equations," *M.T.A.C.*, Vol. 9, p. 101.

Sheldon, J. W. (1959). "On the spectral norm of several iterative processes," *J. Assoc. Comp. Mach.*, Vol. 6, p. 494.

Varga, R. S. (1957). "A comparison of the successive over-relaxation method and semi-iterative methods using Chebyshev polynomials," *J. Soc. Indust. Appl. Math.*, Vol. 5, p. 39.

Varga, R. S. (1962). *Matrix Iterative Analysis*, New Jersey: Prentice-Hall.

Young, D. (1954). "Iterative methods for solving partial difference equations of elliptic type," *Trans. Amer. Math. Soc.*, Vol. 76, p. 92.

Downloaded from https://academic.oup.com/comjnl/article/7/1/72/558762 by guest on 13 March 2024