Convergence of Block Iterative Methods Applied to Sparse Least-Squares Problems

W. Niethammer
Institut für Praktische Mathematik
Universität Karlsruhe
D-7500 Karlsruhe, West Germany

J. de Pillis*
Institut für Praktische Mathematik
Universität Karlsruhe
D-7500 Karlsruhe, West Germany
and
Department of Mathematics
University of California
Riverside, California 92321

and

R. S. Varga[†]
Institut für Praktische Mathematik
Universität Karlsruhe
D-7500 Karlsruhe, West Germany
and
Institute for Computational Mathematics
Kent State University
Kent, Ohio 44242

ABSTRACT

Recently, special attention has been given, in the mathematical literature, to the problems of accurately computing the least-squares solutions of very large-scale overdetermined systems of linear equations, such as those arising in geodetical

network problems. In particular, it has been suggested that one solve such problems, iteratively by applying the block-SOR (successive overrelaxation) iterative method to a consistently ordered block-Jacobi matrix that is weakly cyclic of index 3. Here, we obtain new results (Theorem 1), giving the exact convergence and divergence domains for such iterative applications. It is then shown how these results extend, and correct, the literature on such applications. In addition, analogous results (Theorem 2) are given for the case when the eigenvalues of the associated block-Jacobi matrix are nonnegative.

1. INTRODUCTION

There has been much recent interest in accurately computing the least-squares solutions of very large sparse overdetermined linear systems of equations. In geodetical network problems, for example, such overdetermined systems have the form

$$A\mathbf{x} \approx \mathbf{b}. \tag{1.1}$$

Here, A (the observation matrix) is a given real $m \times n$ matrix (i.e., $A \in \mathbb{R}^{m,n}$) with $m \geqslant n$, where it is assumed that A has full column rank n, and $\mathbf{b}$ is a given real vector with m components (i.e., $\mathbf{b} \in \mathbb{R}^m$). The least-squares solution of (1.1) is the unique vector $\mathbf{x}$ in $\mathbb{R}^n$ for which

$$\|\mathbf{b} - A\mathbf{x}\|_2 = \min_{\mathbf{y} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{y}\|_2 \qquad \left(\text{where} \quad \|\mathbf{u}\|_2^2 := \mathbf{u}^* \cdot \mathbf{u}\right). \tag{1.2}$$

We recommend the recent papers of Golub and Plemmons [3] and Plemmons [5], where extended bibliographies for such geodetic problems are given.

An equivalent formulation of the above least-squares problem is the following: determine vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{r} \in \mathbb{R}^m$ such that

$$\mathbf{r} + A\mathbf{x} = \mathbf{b}, \qquad A^T \mathbf{r} = \mathbf{0}. \tag{1.3}$$

Since A has full column rank n, we may assume that the rows of A have been permuted so that A has the block-partitioned form

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \tag{1.4}$$

where A_1, in $\mathbb{R}^{n,n}$, is nonsingular. With the vectors r and b of (1.3) partitioned conformally with respect to the partitioning of A in (1.4), i.e.,

$$r = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}, \quad b = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}, \quad \text{where} \quad \mathbf{v}, \mathbf{b}_1 \in \mathbb{R}^n, \quad \mathbf{w}, \mathbf{b}_2 \in \mathbb{R}^{m-n}, \quad (1.5)$$

the system of equations (1.3) can be expressed as the following linear system of $m + n$ equations in $m + n$ unknowns:

$$Cz = d, \tag{1.6}$$

where

$$C := \begin{bmatrix} A_1 & 0 & I \\ A_2 & I & 0 \\ 0 & A_2^T & A_1^T \end{bmatrix}, \quad z := \begin{bmatrix} \mathbf{x} \\ \mathbf{w} \\ \mathbf{v} \end{bmatrix}, \quad d := \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ 0 \end{bmatrix}. \tag{1.7}$$

Because A_1 is nonsingular, it can be easily verified that the $(m+n) \times (m+n)$ matrix C of (1.7) is also nonsingular.

Our interest in the reformulation (1.6) of (1.2) stems from the fact that the block-SOR (successive overrelaxation) iterative method can be conveniently applied to the solution of (1.6), an observation which was first made in Chen [2]. To define the iterative method, set $D := \text{diag}(C) = \text{diag}(A_1, I, A_1^T)$, so that D is a nonsingular block-diagonal matrix. The associated block-Jacobi matrix J for the matrix C of (1.6) is then given by

$$J := I - D^{-1}C = \begin{bmatrix} 0 & 0 & -A_1^{-1} \\ -A_2 & 0 & 0 \\ 0 & -A_1^{-T}A_2^T & 0 \end{bmatrix} =: \begin{bmatrix} 0 & 0 & B_1 \\ B_2 & 0 & 0 \\ 0 & B_3 & 0 \end{bmatrix}. \tag{1.8}$$

Next, on writing the block-Jacobi matrix J of (1.8) as the sum $J = L + U$ where

$$L := \begin{bmatrix} 0 & 0 & 0 \\ B_2 & 0 & 0 \\ 0 & B_3 & 0 \end{bmatrix}, \quad U := \begin{bmatrix} 0 & 0 & B_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{1.9}$$

the block-SOR iterative method, applied to (1.6), is, as usual, defined by

$$\mathbf{z}^{(m+1)} = \mathbf{z}^{(m)} + \omega\left\{L\mathbf{z}^{(m+1)} - \mathbf{z}^{(m)} + U\mathbf{z}^{(m)} + D^{-1}\mathbf{d}\right\} \qquad (m = 0, 1, \ldots),$$

(1.10)

where $\mathbf{z}^{(0)}$ is an arbitrary vector in $\mathbb{R}^{m+n}$, and where ω is the associated relaxation parameter. Equivalently, (1.10) can be expressed as

$$\mathbf{z}^{(m+1)} = \mathcal{L}_\omega \mathbf{z}^{(m)} + \omega(I - \omega L)^{-1}D^{-1}\mathbf{d} \qquad (m = 0, 1, \ldots), \quad (1.11)$$

where the block-SOR iterative matrix, $\mathcal{L}_\omega$, is defined as

$$\mathcal{L}_\omega := (I - \omega L)^{-1}\left\{(1 - \omega)I + \omega U\right\}. \tag{1.12}$$

For the convergence properties of the block-SOR iterative method (1.11), it is essentially to observe, as in Chen [2] and Plemmons [5], that the block-Jacobi matrix J of (1.8) is a *consistently ordered matrix, weakly cyclic of index* 3 (cf. Varga [6; 7, p. 101]). Moreover, from (1.8), it directly follows that

$$J^3 = \operatorname{diag}\left(-A_1^{-1}P^TPA_1; -PP^T; -P^TP\right) \qquad \left(\text{where} \quad P := A_2A_1^{-1}\right)$$

$$= \operatorname{diag}\left(B_1B_3B_2; B_2B_1B_3; B_3B_2B_1\right), \tag{1.13}$$

so that J^3 is similar to a real symmetric negative semidefinite matrix. Therefore the eigenvalues of J^3 lie in the real interval

$$I_- := \left[-\rho^3(J), 0\right]; \tag{1.14}$$

here, $\rho(J)$ denotes the spectral radius of J. Because J is a consistently ordered matrix that is weakly cyclic of index 3, the special case $p = 3$ of Varga ([6, Theorem 4] or [7, Theorem 4.3]) can be applied to deduce the following known relationship between the eigenvalues of $\mathcal{L}_\omega$ and those of J:

THEOREM A.   *If β is an eigenvalue of the block-Jacobi matrix J of (1.8), and if λ satisfies*

$$(\lambda + \omega - 1)^3 = \lambda^2\omega^3\beta^3, \tag{1.15}$$

then λ is an eigenvalue of the block-SOR iteration matrix $\mathscr{L}_\omega$ of (1.12). Conversely, if $\omega \neq 0$, if λ is a nonzero eigenvalue of $\mathscr{L}_\omega$, and if β satisfies (1.15), then β is an eigenvalue of J.

In the next section, our new results, concerning the *exact* convergence domain of the block-SOR iterative method, are stated (Theorem 1) for the block-SOR iterative method of (1.12), when the eigenvalues of J^3 are assumed to lie in the interval I_- of (1.14). As an important consequence of Theorem 1, applications of the block-SOR iterative method can be made even in cases when the associated block-Jacobi matrix is *divergent*, a case not treated heretofore in the literature. In analogy with Theorem 1, the exact convergence domain for the block-SOR iterative method (1.12) is stated (Theorem 2) for the case when the eigenvalues of J^3 are nonnegative, and connections with existing literature are made.

It should be noted that, in general, there will be *many* choices of A_1 (with A_1 nonsingular) possible in (1.4), and each choice clearly affects the spectral radius $\rho(J)$ of the associated block-Jacobi matrix of (1.8). Now, Theorem 1, which applies to *each* such choice of A_1, gives precise convergence and divergence regions (as a function of ω) for the associated block-SOR iterative method, these regions depending only on $\rho(J)$ [cf. (2.1)–(2.2)]. Thus, from a practical point of view, one is interested in techniques for selecting nonsingular matrices A_1 in (1.4) which minimize (or nearly minimize) the associated spectral radius $\rho(J)$ of the associated block-Jacobi matrix. The discussion of such practical techniques, which is beyond the scope of this paper, can be found in [2] and [5].

2. STATEMENT OF NEW RESULTS

With $\beta := \rho(J)$ denoting the spectral radius of the block-Jacobi matrix J of (1.8), our first result (whose proof is given in Section 3) is

THEOREM 1. *The block-SOR iterative method of (1.11), applied to the matrix equation (1.6), converges for*

$$0 < \omega < \omega_1(\beta) := \frac{2}{1+\beta} \qquad when \quad 0 \leqslant \beta \leqslant 2, \qquad (2.1)$$

converges for

$$\omega_2(\beta) := \frac{\beta-2}{\beta-1} < \omega < \omega_1(\beta) \qquad when \quad 2 \leqslant \beta < 3, \qquad (2.2)$$

and diverges for all other values of ω. The optimal relaxation factor $\omega_b = \omega_b(\beta)$ is the unique positive root of

$$4\beta^3\omega^3 + 27\omega - 27 = 0 \qquad (0 \leqslant \beta < 3), \tag{2.3}$$

and ω_b satisfies

$$\tfrac{1}{2} < \omega_b \leqslant 1 \qquad \text{for all} \quad 0 \leqslant \beta < 3. \tag{2.4}$$

Further, there holds

$$\rho(\mathscr{L}_{\omega_b}) = 2(1 - \omega_b) \qquad \text{for all} \quad 0 \leqslant \beta < 3. \tag{2.5}$$

It is clear from Theorem 1 that one can find values of ω for which the block-SOR iteration matrix $\mathscr{L}_\omega$ is *convergent*, even when the block-Jacobi matrix J is divergent, i.e., with $1 \leqslant \beta := \rho(J) < 3$, there are intervals in ω [cf. (2.1) and (2.2)] for which $\rho(\mathscr{L}_\omega) < 1$. In this respect, Theorem 1 extends what is known theoretically in the literature for such block-SOR applications. More important, however, is the fact that Theorem 1 greatly increases the *applicability* of this block-SOR iterative method to least-squares problems, such as those arising in geodetical network problems.

It is also important to note that Theorem 1 *corrects* results stated in the literature for such least-squares applications. Under the assumptions of Theorem 1, one finds in Plemmons [5, p. 166] the statement that for any $\rho(J) < 1$, the associated block-SOR iterative method converges for *all* ω satisfying

$$0 < \omega < \tfrac{3}{2} \qquad [0 \leqslant \rho(J) < 1], \tag{2.6}$$

whereas from (2.1) of Theorem 1, the correct statement is that the associated block-SOR iterative method converges in the subset of (2.6) consisting of all ω satisfying

$$0 < \omega < \frac{2}{1 + \rho(J)} \qquad [0 \leqslant \rho(J) < 1]. \tag{2.7}$$

The same error occurs in Berman and Plemmons [1, p. 179].

In contrast with the behavior of the familiar SOR iterative method which arises in applications to the numerical solution of positive definite matrix problems derived from elliptic boundary-value problems, we remark that it is now preferable to *underestimate*, rather than to overestimate, the optimum

relaxation factor $\omega_b(\beta)$, for the above applications to least-squares problems. This is particularly evident in Figure 2, where it is seen that even overestimating ω_b by a very small amount harms the associated rate of convergence of the block-SOR method more severely than does underestimating ω_b by the same amount. In this regard, we further remark that any overestimate of $\beta := \rho(J)$ yields, fortunately, an underestimate of $\omega_b(\beta)$, which can either be deduced from (2.3) of Theorem 1, or seen graphically in Figure 1.

It is interesting to note that Theorem 1 can also be derived from the results of Niethammer and Varga [4], where k-step iterative methods are studied from the point of view of summability theory. The idea there is to write the block-SOR iterative method as a three-step iterative method, in the same way as it was done in [4] for deriving two-step iterative methods for matrices that are weakly cyclic of index 2.

As a counterpart of Theorem 1, we now present results for the case where the eigenvalues of J^3 are nonnegative, i.e., the eigenvalues of J^3 lie in the interval

$$I_+ := \left[0, \rho^3(J)\right].\tag{2.8}$$

THEOREM 2. *Let a block-Jacobi matrix J be a consistently ordered matrix, weakly cyclic of index 3, such that the eigenvalues of J^3 are real and nonnegative. With $\beta := \rho(J)$, the associated block-SOR iterative method converges for*

$$0 < \omega < \omega_3(\beta) := \frac{\beta+2}{\beta+1}, \qquad when \quad 0 \leqslant \beta < 1,\tag{2.9}$$

and diverges for all other values of ω. The optimal relaxation factor $\omega_b = \omega_b(\beta)$ is the smallest positive root of

$$4\beta^3\omega^3 - 27\omega + 27 = 0 \qquad (0 \leqslant \beta < 1),\tag{2.10}$$

and ω_b satisfies

$$1 \leqslant \omega_b < \tfrac{3}{2} \qquad for\ all \quad 0 \leqslant \beta < 1.\tag{2.11}$$

Further, there holds

$$\rho\left(\mathscr{L}_{\omega_b}\right) = 2(\omega_b - 1).\tag{2.12}$$

Interestingly enough, while the proof of (2.10)–(2.12) is given in Varga [6], the precise upper bound $\omega_3(\beta)$ for convergence in (2.9) is new. Previously, it had been shown in [6], under the assumptions of Theorem 2, that convergence of the associated block-SOR method holds in the *subset* of (2.9) defined by

$$0 < \omega < \tfrac{3}{2} \qquad [0 \leqslant \beta := \rho(J) < 1]. \qquad (2.13)$$

We now describe the results of both Theorems 1 and 2 in Figure 1. For convenience, we introduce the variable $\tilde{\beta}$, where $\tilde{\beta} := \rho(J)$ if the eigenvalues of J^3 are nonnegative (cf. Theorem 2), and where $\tilde{\beta} := -\rho(J)$ if the eigenvalues of J^3 are nonpositive (cf. Theorem 1). Thus, we obtain an open bounded region Ω in the $\tilde{\beta}$-ω plane, such that for each point in Ω (shown as the shaded region), the associated block-SOR iterative method, with relaxation factor ω [when applied to a system for which $\rho(J) = \tilde{\beta}$ or $\rho(J) = -\tilde{\beta}$] is convergent, and is divergent for all points in the complement of Ω. Also included in Figure 1 is the set of all optimum relaxation factors $\omega_b(\tilde{\beta})$, as a function of $\tilde{\beta}$, and the difference D of the sets of (2.6) and (2.7), as well as the difference E of the sets (2.9) and (2.13).

In Figure 2 is a plot of $\rho(\mathscr{L}_\omega(\tilde{\beta}))$ as a function of ω, for $\tilde{\beta} = -2.0$ and for $\tilde{\beta} = -2.5$. Here, one sees graphically that underestimating $\omega_b(\tilde{\beta})$ (for $\tilde{\beta} < 0$) is in general much superior to overestimating $\omega_b(\tilde{\beta})$.
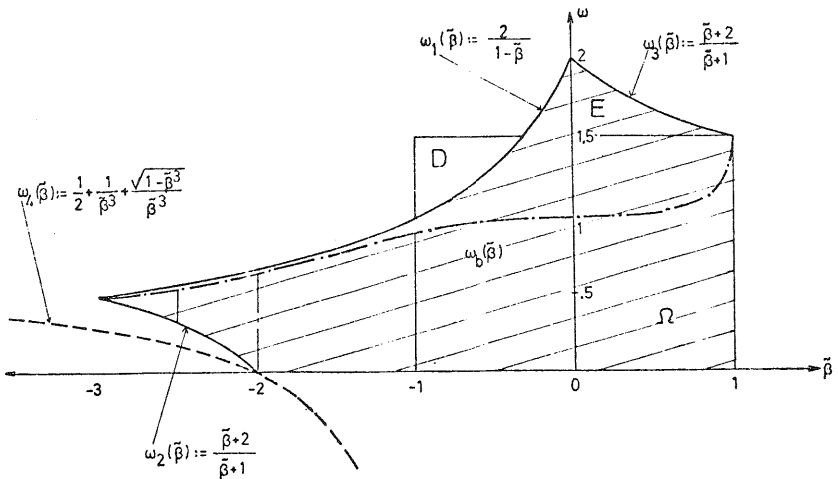

FIG. 1.

$$\rho(\mathscr{L}_\omega(\tilde\beta))$$

$\tilde\beta = -2.5$

$(.333,1.000)$ $(.571,1.000)$

$\leftarrow(.667,1.000)$

1.0

$.9$

$.8$ $(.570,.859)$

$.7$ $\tilde\beta = -2$

$.6$ $(.660,.681)$

$.5$

$.4$

$.3$

$.2$

$.1$

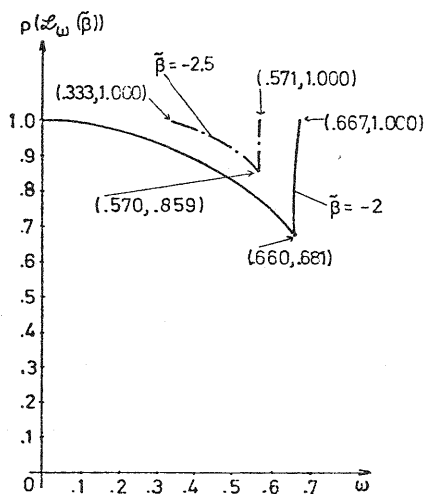0 $.1$ $.2$ $.3$ $.4$ $.5$ $.6$ $.7$ ω

FIG. 2.

3. PROOFS

For the proof of Theorem 1, we begin with Equation (1.15) of Theorem A, in the form

$$(\lambda + \omega - 1)^3 = -\lambda^2\omega^3\beta^3, \tag{3.1}$$

where, from (1.14), we assume that $0 < \beta \leqslant \rho(J)$. Consider next the polynomial in ω

$$G(\omega) = G(\omega;\beta) := \beta^3\omega^3 + \tfrac{27}{4}\omega - \tfrac{27}{4}. \tag{3.2}$$

Clearly, $G(\omega)$ has, by Descartes's rule of signs, a unique positive zero $\hat\omega_b = \hat\omega_b(\beta)$, and as $G(0) = -\tfrac{27}{4}$ and $G(1) = \beta^3 \geqslant 0$, then evidently $\hat\omega_b(\beta)$ satisfies

$$0 < \hat\omega_b(\beta) \leqslant 1 \qquad [0 \leqslant \beta \leqslant \rho(J)] \tag{3.3}$$

as well as

$$\beta^3\hat\omega_b^3(\beta) = \tfrac{27}{4}[1 - \hat\omega_b(\beta)]. \tag{3.4}$$

Using (3.1), with $\omega = \hat{\omega}_b$, and (3.4), it follows that

$$\hat{\lambda}(\beta) := 2[\hat{\omega}_b(\beta) - 1] \tag{3.5}$$

is a zero of multiplicity two of the polynomial (3.1) in λ. Moreover, since the product of the zeros of (3.1) is, in this case, $-(\hat{\omega}_b - 1)^3$, we see that the third associated zero of (3.1) is necessarily smaller in modulus than $|\hat{\lambda}(\beta)|$.

Returning to (3.2), there is an $\hat{\omega}_b$ for each $\beta = \beta_j$, where β_j is an eigenvalue of the block-Jacobi matrix J. On differentiating (3.4) with respect to β, it follows that $\hat{\omega}_b(\beta)$ is a strictly decreasing function of $\beta > 0$, whence $|\hat{\lambda}(\beta)|$ is a strictly increasing function of $\beta > 0$. Since we are interested in $\rho(\mathscr{L}_\omega)$, we are then justified in considering (3.1) and (3.2) with $\beta = \rho(J)$. Consequently, with $\hat{\omega}_b := \hat{\omega}_b(\rho(J))$, it follows from (3.5) and (3.4) that

$$\rho(\mathscr{L}_{\hat{\omega}_b}) = 2(1 - \hat{\omega}_b). \tag{3.6}$$

Note from (3.4) that $\hat{\omega}_b(3) = \frac{1}{2}$, so that $\rho(\mathscr{L}_{\hat{\omega}_b(3)}) = 1$ from (3.6). Thus, since $\hat{\omega}_b(\beta)$ is a strictly decreasing function and $|\hat{\lambda}(\beta)|$ is a strictly increasing function of $\beta > 0$, it is evident that *only* in the interval $0 \leqslant \beta = \rho(J) < 3$ do we have $\rho(\mathscr{L}_{\hat{\omega}_b}) < 1$. Moreover, for each $\rho(J)$ with $0 \leqslant \rho(J) < 3$, there is an open interval $\Omega(\rho(J))$ in ω such that $\rho(\mathscr{L}_\omega) < 1$ for each $\omega \in \Omega(\rho(J))$.

We now set $\lambda := -z^3$ in (3.1), so that on taking cube roots in (3.1), we obtain the associated polynomial

$$g_3(z; \omega) := z^3 - \omega\beta z^2 + (1 - \omega), \qquad \text{where} \quad \beta := \rho(J) > 0. \tag{3.7}$$

[Note that if $\tilde{z}$ is any other solution of the cube root of $(-z^3 + \omega - 1)^3 = -z^6\omega^3\beta^3$, then $\tilde{z} = z\exp(\frac{2}{3}\pi ik)$, $k = 0, 1, 2$, so that $|\tilde{z}| = |z|$.] This brings us to

LEMMA 1.   *With $\hat{\omega}_b$ defined in (3.4), then*

(i) *for $0 < \omega < \hat{\omega}_b$, $g_3(z; \omega)$ has no positive real zeros;*
(ii) *for $\omega = \hat{\omega}_b$, $g_3(z; \omega)$ has a unique real positive zero of multiplicity two;*
(iii) *for $\hat{\omega}_b < \omega < 1$, $g_3(z; \omega)$ has precisely two positive real zeros;*
(iv) *for $\omega \geqslant 1$, $g_3(z; \omega)$ has precisely one positive real zero.*

Proof. (ii): For $\omega = \hat{\omega}_b$, the root $\hat{\lambda} < 0$, of multiplicity two, of equation (3.1) given by (3.5) yields, via $\hat{\lambda} = -\hat{z}^3$, a positive solution $\hat{z}$ of multiplicity two of $g_3(z; \omega_b)$.

(iii): For $\hat{\omega}_b < \omega < 1$, $g_3(z; \omega)$ has at most two real positive zeros by Descartes's rule of signs. With $z =: \zeta(1 - \omega)^{1/3}$, then $g_3(z; \omega)$ can be written as

$$g_3(z; \omega) = (1 - \omega)\{\zeta^3 - \varepsilon(\omega)\zeta^2 + 1\} =: (1 - \omega)h_3(\zeta; \omega), \qquad (3.8)$$

where $\varepsilon(\omega)$ is defined by

$$\varepsilon(\omega) := \frac{\omega\beta}{(1 - \omega)^{1/3}}. \qquad (3.9)$$

Since

$$\frac{d\varepsilon(\omega)}{d\omega} = \frac{\beta(3 - 2\omega)}{3(1 - \omega)^{4/3}},$$

then $\varepsilon(\omega)$ is a *strictly increasing* function of ω in $0 \leqslant \omega < 1$. Since we have from (3.5) that $g_3(\hat{z}; \hat{\omega}_b) = 0$ where $\hat{z} := [2(1 - \hat{\omega}_b)]^{1/3}$, then $\zeta_0 = 2^{1/3}$ is a zero of multiplicity two of $h_3(2^{1/3}; \hat{\omega}_b)$, i.e.,

$$h_3(2^{1/3}; \hat{\omega}_b) = 0. \qquad (3.10)$$

From the monotonicity of $\varepsilon(\omega)$ in $\hat{\omega}_b < \omega < 1$, we conclude from (3.8) that $h_3(\zeta; \omega)$, and thus $g_3(z; \omega)$, have exactly two real positive zeros for $\hat{\omega}_b < \omega < 1$.

(iv): For $\omega \geqslant 1$, Descartes's rule of signs directly shows from (3.7) that $g_3(z; \omega)$ has exactly one positive real zero.

(i): Since from (3.10) we know that $\zeta_0 = 2^{1/3}$ is a zero of multiplicity two and that $h_3(0; \omega) = 1$ from (3.8), it follows from the monotonicity of $\varepsilon(\omega)$ that there are no positive real zeros for $0 < \omega < \hat{\omega}_b$. ∎

For fixed $\beta = \rho(J) > 0$, let us denote the three zeros $g_3(z; \omega)$ by z_1, z_2, and z_3, remembering that the z_i's are functions of ω and β. From $\prod_{i=1}^{3}(z - z_i) = z^3 - \omega\beta z^2 + (1 - \omega)$, we evidently have

$$z_1 + z_2 + z_3 = \omega\beta, \qquad (3.11)$$

$$z_1 z_2 + z_1 z_3 + z_2 z_3 = 0, \qquad (3.12)$$

$$z_1 z_2 z_3 = \omega - 1. \qquad (3.13)$$

Next, we examine the case for ω satisfying $0 < \omega < \hat{\omega}_b$ [where $\hat{\omega}_b$ is defined in (3.4)]. We know from Lemma 1(i) that there are no positive real zeros of $g_3(z; \omega)$. Replacing z by $-t$ in (3.7), we see that $g_3(z; \omega)$ has exactly one negative real zero, so that the two remaining zeros of $g_3(z; \omega)$ are necessarily complex. Thus, we can write, for $0 < \omega < \hat{\omega}_b$, that

$$z_1 < 0, \qquad z_2 := \eta e^{i\phi}, \ z_3 := \eta e^{-i\phi}, \quad \text{where } 0 < \phi < \pi. \qquad (3.14)$$

Using (3.11) and (3.12), we obtain

$$z_1 + 2\eta \cos \phi = \omega \beta, \qquad (3.15)$$

$$2z_1 \eta \cos \phi + \eta^2 = 0. \qquad (3.16)$$

On eliminating $2\eta \cos \phi$ from (3.15) and (3.16), then

$$\eta^2 = z_1^2 - \omega \beta z_1 > z_1^2, \qquad (3.17)$$

since $z_1 < 0$. In other words,

$$\eta = |z_2| = |z_3| > |z_1| \qquad \text{for } 0 < \omega < \omega_b. \qquad (3.18)$$

From (3.13), we further deduce that

$$\left| z_2(\omega) \right|^2 = \eta^2 = \frac{1 - \omega}{-z_1(\omega)} \qquad \text{for } 0 < \omega < \omega_b. \qquad (3.19)$$

Clearly, as z_1 is a zero of $g_3(z; \omega)$, then

$$z_1^3 - \omega \beta z_1^2 + 1 - \omega = 0. \qquad (3.20)$$

Now, a straightforward calculation, based on (3.17) and (3.20), shows that $\eta = \eta(\omega)$ satisfies

$$\frac{d\eta}{d\omega} = \frac{2z_1 - \beta(2\omega - 1)}{2\eta(3z_1^2 - 2\omega \beta z_1)} \qquad \text{for } 0 < \omega < \hat{\omega}_b. \qquad (3.21)$$

Obviously, from (3.21) there follows

$$\frac{d\eta}{d\omega} < 0 \qquad \text{for} \quad \tfrac{1}{2} \leqslant \omega < \hat{\omega}_b. \tag{3.22}$$

Next, we see from (3.21) that $d\eta/d\omega = 0$ for

$$z_1 = \tilde{z}_1 := \frac{-\beta(1-2\omega)}{2}, \qquad \text{where} \quad 0 < \omega < \tfrac{1}{2}. \tag{3.23}$$

As this $\tilde{z}_1$ must satisfy (3.20) as well, a short calculation shows that the associated value of ω in (3.23) satisfies a quadratic equation whose only acceptable solution in the range $0 < \omega < \tfrac{1}{2}$ is

$$\omega_4(\beta) := \frac{1}{2} - \frac{1}{\beta^3} - \frac{\sqrt{1+\beta^3}}{\beta^3}. \tag{3.24}$$

Now, as $\omega_4(\beta) > 0$ only for $\beta > 2$, then $d\eta/d\omega = 0$ can occur in $0 < \omega < \tfrac{1}{2}$ only for $\beta > 2$. In other words, from (3.22), $d\eta/d\omega < 0$ for all $0 < \omega < \hat{\omega}_b$ and $0 < \beta \leqslant 2$, and $d\eta/d\omega < 0$ for all $\omega_4(\beta) < \omega < \hat{\omega}_b$ and $2 < \beta$.

Next, on assuming that $\eta = 1$, we deduce that $z_1 = \omega - 1$ from (3.19). On the other hand, as $z_1 = \omega - 1$ is by definition a zero of $g_3(z; \omega)$, we easily derive from (3.7) that $\eta = 1$ implies that

$$\omega = \omega_2(\beta) := \frac{\beta - 2}{\beta - 1} \qquad \text{for} \quad 0 < \omega < \omega_b. \tag{3.25}$$

As we have seen previously, $\hat{\omega}_b(3) = \tfrac{1}{2}$, and from (3.25), we similarly see that $\omega_2(3) = \tfrac{1}{2}$. Further, we see from (3.24) and (3.25) that $\omega_2(2) = \omega_4(2) = 0$; moreover, direct computation with (3.24) and (3.25) gives that

$$\omega_4(\beta) < \omega_2(\beta) < \tfrac{1}{2} \qquad \text{for all} \quad 2 < \beta < 3. \tag{3.26}$$

Thus, we have precisely one zero of $d\eta/d\omega$ in the interval $[0, \omega_2(\beta)]$ for $2 < \beta$, which occurs when $\omega = \omega_4(\beta)$. On the other hand, $\eta(0) = 1 = \eta(\omega_2(\beta))$, the latter holding for all $\beta > 2$, while $d\eta/d\omega < 0$ for $\omega = \omega_2(\beta)$, $\beta > 2$. Thus,

$$\eta(\omega) > 1 \qquad \text{for all} \quad 0 < \omega < \omega_2(\beta), \quad \beta > 2. \tag{3.27}$$

We have plotted the function ω_4 (with $\tilde{\beta}: = -\beta$) in Figure 1, in order to help interpret the above results graphically.

On putting together the above facts for the case $0 < \omega < \hat{\omega}_b$, and on relating these back to the variable λ by means of $\lambda: = -z^3$, we have thus established

LEMMA 2.   *The following are valid:*

$$\rho(\mathscr{L}_\omega) < 1 \quad \textit{for all } 0 < \omega < \hat{\omega}_b(\beta) \qquad \textit{when} \quad 0 \leqslant \beta \leqslant 2; \quad (3.28)$$

$$\rho(\mathscr{L}_\omega) < 1 \quad \textit{for all } \omega_2(\beta) < \omega < \hat{\omega}_b(\beta) \qquad \textit{when} \quad 2 \leqslant \beta < 3; \quad (3.29)$$

$$\rho(\mathscr{L}_\omega) > 1 \quad \textit{for all } 0 < \omega < \omega_2(\beta) \qquad \textit{when} \quad 2 < \beta. \quad (3.30)$$

Moreover, $\rho(\mathscr{L}_\omega)$ *is a decreasing function of* ω *when either* $0 < \omega < \hat{\omega}_b(\beta)$ *and* $\beta \leqslant 2$, *or* $\omega_4(\beta) < \omega < \hat{\omega}_b(\beta)$ *and* $\beta > 2$.

We now consider the case when $\hat{\omega}_b < \omega < 1$. From Lemma 1(iii), we know that in this range of ω,

$$z_1(\omega) < 0 \quad \text{and} \quad 0 < z_2(\omega) < z_3(\omega). \quad (3.31)$$

Moreover, from $g_3(z; \omega)$ of (3.7), it follows that if $\hat{\omega}_b < \omega' < \omega < 1$, then $z_2(\omega) < z_2(\omega')$ and $z_3(\omega') < z_3(\omega)$. Also, since $g_3(z_1(\omega'); \omega) < 0$, then $z_1(\hat{\omega}_b) < z_1(\omega') < z_1(\omega) < 0$, so that $z_3(\omega)$ is the largest (in modulus) zero of $g_3(z; \omega)$ for $\hat{\omega}_b \leqslant \omega \leqslant 1$. Moreover,

$$z_3(\omega) \text{ is a strictly increasing function of } \omega \text{ in } \hat{\omega}_b < \omega. \quad (3.32)$$

We next consider the case when $\omega \geqslant 1$. From Lemma 1(iv), $g_3(z; \omega)$ has one positive zero, and on replacing z by $-t$, we see that $g_3(z; \omega)$ has no negative real zeros. Thus, the zeros of $g_3(z; \omega)$ in this case can be expressed as

$$z_1(\omega) = \eta e^{i\theta}, \qquad z_2(\omega) = \eta e^{-i\theta}, \qquad z_3(\omega) > 0 \quad \text{for } 1 \leqslant \omega, \quad (3.33)$$

where $0 < \theta < \pi$. Now, in the same way we derived (3.17), we get

$$\eta^2 = z_3^2 - \omega\beta z_3 < z_3^2 \qquad \text{for} \quad 1 \leqslant \omega. \quad (3.34)$$

Next, assume $z_3 = 1$. Then from (3.7), we obtain that ω must satisfy

$$\omega = \omega_1(\beta) := \frac{2}{1+\beta}. \tag{3.35}$$

But, as (3.32) holds for all $\hat{\omega}_b < \omega$, we have deduced

LEMMA 3.   *The following are valid:*

$$\rho(\mathscr{L}_\omega) < 1 \quad \textit{for all} \quad \hat{\omega}_b < \omega < \omega_1(\beta), \tag{3.36}$$

and

$$\rho(\mathscr{L}_\omega) \geqslant 1 \quad \textit{for all} \quad \omega_1(\beta) \leqslant \omega. \tag{3.37}$$

Moreover, $\rho(\mathscr{L}_\omega)$ is a strictly increasing function of ω for all $\hat{\omega}_b < \omega$.

To conclude the proof of Theorem 1, we remark that it is well known (cf. [7, Theorem 3.5]) that the block-SOR iterative method is necessarily divergent for $\omega \leqslant 0$ or for $\omega \geqslant 2$. With this and Lemmas 2 and 3, Theorem 1 is thus proved. ∎

The proof of Theorem 2 follows along the same lines as above, if Lemma 1 in [6] is used in place of Lemma 1 of this section.

REFERENCES

1   A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.
2   Y. T. Chen, Iterative methods for linear least-squares problems, Ph.D. Dissertation, Dept. of Computer Science, Univ. of Waterloo, Waterloo, Ontario, Canada, 1975.
3   G. H. Golub and R. J. Plemmons, Large-scale geodetic least-squares adjustment by dissection and orthogonal decomposition, *Linear Algebra Appl.* 34:3–27 (1980).
4   W. Niethammer and R. S. Varga, The analysis of k-step iterative methods for linear systems from summability theory, *Numer. Math.* 41:177–206 (1983).
5   R. J. Plemmons, Adjustments by least squares in geodesy using block iterative methods for sparse matrices, in *Proceedings of the Annual U.S. Army Conference on Numerical Analysis and Computers*, 1979, pp. 151–186.
6   R. S. Varga, p-cyclic matrices: A generalization of the Young-Frankel successive overrelaxation scheme, *Pacific J. Math.* 9:617–628 (1959).
7   R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.