1992

A General Theory of Optimal p-Cyclic SOR

Apostolos Hadjidimos

Robert J. Plemmons

Report Number:
92-076
A GENERAL THEORY OF OPTIMAL p-CYCLIC SOR

Apostolos Hadjidimos
Robert L. Plemmons

CSD-TR-92-076
October 14, 1992
A General Theory of Optimal $p$-Cyclic SOR

Apostolos Hadjidimos* and Robert J. Plemmons†

October 4, 1992

Abstract

The convergence theory of the Successive Overrelaxation (SOR) iterative method for the solution of nonsingular linear systems $Ax = b$, when the matrix $A$ has a block $p \times p$ partitioned $p$-cyclic form, is well documented. However, when $A$ is singular the corresponding theory is far behind that for the nonsingular case. Our purpose in this paper is to extend the $p$-cyclic SOR theory to consistent singular systems and to apply the results to the solution of large scale systems arising, e.g., in queueing network problems in Markov analysis. Markov chains and queueing models lead to structured singular linear systems and are playing an increasing role in the understanding of complex phenomena arising in computer, communication and transportation systems.

For certain important classes of singular problems, we develop a convergence theory for $p$-cyclic SOR, and show how to repartition for optimal convergence. Results by Kontovasilis, Plemmons and Stewart on the new concept of convergence of SOR in an extended sense are rigorously analyzed and applied to the solution of periodic Markov chains with period $p = 2$. In addition, the use of $p$-cyclic SOR as a smoother for algebraic multigrid computations for queueing network problems is discussed.

*Department of Computer Science, Purdue University, West Lafayette, IN 47907. Research supported in part by the US Air Force under grant no. AFOSR-88-10243 and by NSF under grant no. CCR-86-19817.
†Department of Mathematics and Computer Science, Wake Forest University, P.O. Box 7388, Winston-Salem, NC 27109. Research supported by the US Air Force under grant no. AFOSR-91-0163 and by NSF under grant no. CCR-92-01105.
1 Introduction

Block iterative methods are suitable for the solution of large and sparse systems of linear equations having matrices that possess a special structure. Here we consider block p-cyclic SOR for arbitrary consistent systems of linear equations. Given

\[ Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad x, b \in \mathbb{R}^n \]  

and the usual block decomposition

\[ A = D - L - U \]  

where \( D, L \) and \( U \) are block diagonal, lower and upper triangular matrices respectively and \( D \) is nonsingular, the block SOR method for any \( \omega \neq 0 \) is defined as:

\[ D x^{(m)} = D x^{(m-1)} + \omega (L x^{(m)} - D x^{(m-1)} + U x^{(m-1)} + b), \quad m = 1, 2, \ldots \]  

The method can be equivalently described as

\[ x^{(m)} = \mathcal{L}_\omega x^{(m-1)} + c, \quad m = 1, 2, \ldots \]  

where

\[ \mathcal{L}_\omega = (D - \omega L)^{-1} [(1 - \omega)D + \omega U], \quad c = \omega (D - \omega L)^{-1} b. \]  

It is well known that, for nonsingular systems (1.1), SOR converges iff \( \rho(\mathcal{L}_\omega) < 1 \). The associated spectral convergence factor is then \( \rho(\mathcal{L}_\omega) \).
For arbitrary systems (1.1), very little is known about the optimal relaxation parameter \( \omega \) which minimizes \( \rho(\mathbf{L}_\omega) \) as a function of \( \omega \). However, considerable information is known for the situations where the matrix \( \mathbf{A} \) has a special block cyclic structure. For the important case of matrices with "Property A", Young [38] (see also [39]) discovered his famous result on the optimum \( \omega \). Here \( \mathbf{A} \) in (1.1) has a special two-by-two cyclic block form. Young's result was generalized by Varga [34] (see also [35]) to consistently ordered \( p \)-cyclic matrices. In this case it is assumed (without loss of generality) that \( \mathbf{A} \) has the partitioned block form

\[
\mathbf{A} = \begin{pmatrix}
\mathbf{A}_1 & 0 & 0 & \cdots & \mathbf{B}_1 \\
\mathbf{B}_2 & \mathbf{A}_2 & 0 & \cdots & 0 \\
0 & \mathbf{B}_3 & \mathbf{A}_3 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \mathbf{B}_p & \mathbf{A}_p
\end{pmatrix}
\]  

(1.6)

where each diagonal submatrix \( \mathbf{A}_i \) is square and nonsingular. With \( \mathbf{D} \) in (2) defined by \( \mathbf{D} \equiv \text{diag (A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p) \), the associated block Jacobi matrix \( \mathbf{J}_p \) defined by \( \mathbf{J}_p \equiv \mathbf{I} - \mathbf{D}^{-1} \mathbf{A} \), has the form

\[
\mathbf{J}_p = \begin{pmatrix}
0 & 0 & 0 & \cdots & \mathbf{C}_1 \\
\mathbf{C}_2 & 0 & 0 & \cdots & 0 \\
0 & \mathbf{C}_3 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \mathbf{C}_p & 0
\end{pmatrix}
\]  

(1.7)

where \( \mathbf{C}_i \equiv -\mathbf{A}_i^{-1} \mathbf{B}_i \), 1 \( \leq i \leq p \).

Matrices of the form (1.7) were defined by Varga [34] to be weakly cyclic of index \( p \), and in this case \( \mathbf{A} \) in (1.6) is termed \( p \)-cyclic and consistently ordered. For such matrices Varga proved the important relationship

\[
(\lambda + \omega - 1)^p = \lambda^{p-1} \omega^p \mu^p
\]  

(1.8)

between the eigenvalues \( \mu \) of \( \mathbf{J}_p \) and \( \lambda \) of \( \mathbf{L}_\omega \), generalizing in this way Young's relationship for \( p = 2 \). Assuming further that all eigenvalues of \( \mathbf{J}_p \) satisfy

\[
0 \leq \mu^p \leq \rho(\mathbf{J}_p^p) < 1,
\]

he showed that the optimum \( \omega \) value \( \omega_p \) is the unique positive solution of the equation
\[ \rho(J_p^p) = p^p(p - 1)^{1-p}(\omega - 1) \]  
\[ (1.9) \]

in the interval \((1, \frac{p}{p-1})\). This \(\omega_p\) yields a convergence factor equal to

\[ \rho(L_{\omega_p}) = (p - 1)(\omega_p - 1). \]  
\[ (1.10) \]

Similar results have been obtained (see [21, 26, 27, 36, 14]) for the case where the eigenvalues of \(J_p^p\) are nonpositive, that is,

\[-\left(-\frac{p}{p-2}\right)^p < -\rho(J_p^p) \leq \mu^p \leq 0.\]

Few applications of \(p\)-cyclic SOR have been found for \(p > 2\) (see, e.g., Berman and Plemmons [2]), but it turns out that, for example, least squares computations lead naturally to a 3-cyclic SOR iterative scheme. Markham, Neumann and Plemmons [22] have described a formulation of the problem leading to a 2-cyclic SOR method, obtained by repartitioning the coefficient matrix into a 2-cyclic form. They showed that this method always converges for sufficiently small values of the SOR parameter \(\omega\), in contrast to the 3-cyclic formulation, and that the 2-cyclic approach is asymptotically faster. Their result was extended by Galanis, Hadjidimos and Noutsos [14] to cover nonnegative and nonpositive spectra \(\sigma(J_p^p)\) for any \(p\) for certain classes of cyclic repartitionings. Recently, Pierce, Hadjidimos and Plemmons [29] have generalized and extended the technology of \(p\)-cyclic SOR by showing that if the spectrum of the \(p\)-th power, \(J_p^p\), of the block Jacobi matrix given in (1.7) is either nonpositive or nonnegative, then repartitioning a block \(p\)-cyclic matrix into a block \(q\)-cyclic form, \(q < p\), results in asymptotically faster SOR convergence for the same amount of work per iteration. As a consequence, 2-cyclic SOR is asymptotically optimal for SOR under these conditions. In particular, it follows that 2-cyclic is optimal for SOR applied to linear equality constrained least squares in the Kuhn–Tucker formulation since here the spectrum of \(J_3^3\) is nonpositive.

In general, the requirement that the spectrum of \(J_p^p\) be either nonpositive or nonnegative is critical. Eiermann, Niethammer and Ruttan [11] have shown by considering experimental and theoretical counterexamples that, without this requirement, 2-cyclic SOR is not always superior to \(p\)-cyclic SOR, \(p > 2\). Galanis and Hadjidimos [13] have now generalized all of this work for nonsingular systems by showing how to repartition a block \(p\)-cyclic
consistently ordered matrix for optimal SOR convergence for the general real case of the eigenvalues of $J_p$. One must keep in mind that all these convergence results hold only in an asymptotic sense. Golub and de Pillis [15] have pointed out that short-term convergence, i.e., error reduction in the early iterations, may only be controlled by reducing the spectral norm of the iteration matrix, while long-term or asymptotic convergence is generally improved by minimizing the spectral radius, as described above.

All results mentioned thus far consider only nonsingular systems of equations of the form (1.1). For many applications, for example to Markov chains, the coefficient matrix $A$ will be singular. Hadjidimos [17] examined the singular case ($\det(A) = 0$ and $b \in \mathcal{R}(A)$). Under the assumptions that: (1) the Jacobi matrix $J = J_p$ is weakly cyclic of index $p$, (2) the eigenvalues of $J_p$ are nonnegative with $\rho(J) = 1$, and (3), that $J$ has either a simple unit eigenvalue or a multiple one associated with $1 \times 1$ Jordan blocks. Hadjidimos proved, among other results, that $\omega_p$ is the unique root of (1.9) (in the same interval as in the nonsingular case), where $\rho(J_p)$ has to be replaced by $\gamma(J_p)$, the maximum of the moduli of the eigenvalues of $J_p$, excluding those that have modulus 1, viz.

$$\gamma(J_p) \equiv \max \{|\lambda|, \lambda \in \sigma(J), |\lambda| \neq 1\}.$$ 

Recall now that, if $A$ is a singular irreducible $M$-matrix then $1 \in \sigma(L_\omega)$ for all $\omega$ and, the conditions for semiconvergence (see, e.g., [2]) become:

- $\rho(L_\omega) = 1$.
- Elementary divisors associated with 1 are linear, i.e., $\operatorname{rank}(I - L_\omega)^2 = \operatorname{rank}(I - L_\omega)$ or, equivalently, $\operatorname{index}(I - L_\omega) = 1$.
- If $\lambda \in \sigma(L_\omega)$ with $|\lambda| = 1$, then $\lambda = 1$, i.e., $\gamma(L_\omega) < 1$.

For consistency, we will use the term convergence to mean semiconvergence in the singular case.

The results we will obtain here on optimal $p$-cyclic SOR for consistent linear systems $Ax = b$ have applications to discrete ergodic Markov Chain problems with a transition probability matrix $P$ being cyclic with period $p$, as discussed in [20, 33]. In particular, Markov chains sometimes possess the property that the minimum number of transitions that must be made on leaving any state to return to that state, is a multiple of some integer
$p > 1$. These models are said to be periodic of period $p$, or $p$-cyclic of index $p$. Bonhoure, Dallery and Stewart [3] have shown that Markov chains that arise from queueing network models frequently possess this property.

Indeed, in the discrete case, the problem to be solved is

$$\pi^T P = \pi^T, \quad \|\pi\|_1 = 1$$  \hspace{1cm} (1.11)

or, equivalently,

$$(I - P^T)\pi^T = 0, \quad \|\pi\|_1 = 1,$$

where the element $\pi_i$ is the probability of being in state $i$ when the system reaches statistical equilibrium. It is immediate that, setting $A \equiv I - P^T$, and noting that if $P$ is a cyclic stochastic matrix with transpose of the form (1.7), the corresponding homogeneous stochastic problem has a matrix that is of the form (1.6) and the associated Jacobi matrix is $J_p = P^T$. Therefore, all the results of this paper carry over to $p$-cyclic Markov Chains, simply by replacing $J_p$ with $P^T$. In particular, the matrix $A$ is a singular $M$-matrix and is irreducible when the chain is ergodic. Thus the conditions for semiconvergence described earlier apply to this Markov chain application.

For homogeneous continuous time $p$-cyclic Markov chains with infinitesimal generator $Q$, considered in [3, 20, 33, 4, 18], we are interested in solving

$$\pi^T Q = 0, \quad \|\pi\|_1 = 1.$$  \hspace{1cm} (1.12)

Equation (1.12) may also be written in the form (1.11), where

$$P = Q\Delta t + I,$$  \hspace{1cm} (1.13)

if $\Delta t$ is sufficiently small. In the $p$-cyclic case, the infinitesimal generator matrix $Q$ in (1.12) is such that $Q^T$ has the block form (1.6).

Markov chains and queueing models thus lead to structured singular, irreducible linear systems of the type considered in this paper. Queueing models are playing an increasing role in the understanding of complex phenomena arising in computer, communication and transportation systems.

Our purpose in this paper is to extend the $p$-cyclic SOR theory to the singular case and to apply the results to the solution of large scale singular systems arising, e.g., in queueing network problems in Markov analysis. For certain important classes of singular problems, we provide a convergence
theory for $p$-cyclic SOR in §2, and show in §3 how to repartition for optimal convergence. Recent results by Kontovasilis, Plemmons and Stewart [20] on the new concept of convergence of SOR in an extended sense are rigorously analyzed in §4, for the important case where $J^2_2$ has all real eigenvalues with the same sign. In addition, the possible use of $p$-cyclic SOR as a smoother for algebraic multigrid computations for queueing network problems in Markov analysis is discussed along with other topics in §5.

2 The General $p$-Cyclic Case

For the determination of $\omega_p$ in the general $p$-cyclic singular case we begin our analysis by giving the corresponding result for the nonsingular case. This is stated in [11] or, in a more compact form, in Theorem 2.2 of [13]. More specifically:

**Lemma 2.1:** Suppose we are given a nonsingular system of the form (1.1), where $A$ is of the form (1.6) and the associated Jacobi iteration matrix $J_p$ is of the form (1.7). Let $\omega_p$ and $\rho_p$ denote the relaxation factor and the convergence factor, respectively, of the optimal $p$-cyclic SOR for which

$$\sigma(J^p_p) \subset [-\alpha^p, \beta^p], \quad -\alpha^p, \beta^p \in \sigma(J^p_p),$$

$$0 \leq \alpha < \frac{p}{p-2}, \quad 0 \leq \beta < 1.$$ 

(2.1)

Then $\omega_p$ and $\rho_p$ are determined from the equations

$$\left(\frac{\alpha_p + \beta_p}{2}\right) \omega^p - \frac{(\alpha_p + \beta_p)}{(\beta_p - \alpha_p)} (\omega - 1) = 0$$

(2.2)

and

$$\rho_p = \left(\frac{\alpha_p + \beta_p}{2}\right) (\omega_p - 1) = \left(\frac{(\alpha_p + \beta_p)}{2}\right)^p \omega_p$$

(2.3)

where $\omega_p$ is the unique positive root of (2.2) in

$$\left(\min\{1, 1 + \frac{\beta_p - \alpha_p}{\alpha_p + \beta_p}\}, \max\{1, 1 + \frac{\beta_p - \alpha_p}{\alpha_p + \beta_p}\}\right)$$

(2.4)

and where
Note: The limiting cases $\alpha = \beta (= 0$ or $\neq 0)$ lead to $\omega_p = 1$ and $\rho_p = \alpha^p = \beta^p$; while for $\alpha \neq 0$, $\beta = 0$ it is assumed that $\frac{\alpha}{\beta} = \infty$ and also for $p = 2$, $\frac{\rho}{\rho_{p-2}} = \infty$.

For the singular case we are studying here we recall a result from Theorem 3.1 of [17] which will be used in the sequel.

**Lemma 2.2:** If the block Jacobi matrix $J_p$ in (1.7) satisfies the assumption index $(I - J_p) = 1$, then for all

$$\omega \in (0, 2) \setminus \{p/(p - 1)\}$$

it follows that

$$\text{index}(I - L_\omega) = \text{index}(I - J_p) = 1.$$ 

**Note:** Lemma 2.2 can be extended to cover all $\omega \in (-\infty, \infty) \setminus \{0, p/(p - 1)\}$. For the general singular case of interest in this paper we assume that

$$\sigma(J_p^\omega) \subset [-\alpha^p, \beta^p] \cup \{1\},$$

(2.6)

with $\alpha$ and $\beta$ being defined as in (2.1). Thus, under the assumption that $\text{index}(I - J_p) = 1$ and with (2.6) replacing the first part of (2.1), the main result of this section is identically the same as that of Lemma 2.1. Evidently, in (2.3) the optimal semiconvergence factor $\gamma_p = \gamma(L_{\omega_p})$ must replace the optimal spectral radius $\rho_p = \rho(L_{\omega_p})$. This result extends that obtained in Theorem 3.3 of [17], where the nonnegative case ($\alpha = 0 \leq \beta < 1$) was treated, to that of the general real case of $\sigma(J_p^\omega)$. The proof of the resulting theorem stated next is given in the Appendix.

**Theorem 2.1:** Suppose we are given a possibly singular system of the form (1.1), where $A$ is of the form (1.6) and the associated Jacobi iterative matrix $J_p$ is of the form (1.7). Let $\omega_p$ and $\gamma_p$ denote the relaxation factor and the convergence factor, respectively, of the optimal $p$-cyclic SOR for which

$$\sigma(J_p^\omega) \subset [-\alpha^p, \beta^p] \cup \{1\}, \quad -\alpha^p, \beta^p \in \sigma(J_p^\omega),$$

$$0 \leq \alpha < \frac{p}{p-2}, \quad 0 \leq \beta < 1,$$

(2.7)

and, moreover, assume that $\text{index}(I - J_p) = 1$. Then $\omega_p$ and $\gamma_p$ are determined by equations (2.1) through (2.5), with $\gamma_p$ replacing $\rho_p$. 

8
3 Best Cyclic Repartitioning

As was mentioned in the Introduction, Markhan, Neumann and Plemmons [22] were the first who considered the problem of repartitioning a block 3-cyclic consistently ordered matrix into a 2-cyclic form for optimal SOR convergence. The most recent result on the general problem of the best cyclic repartitioning seems to be that obtained by Galanis and Hadjidimos [13]. It covers the case where the spectrum $\sigma(J_p)$ is real, under the assumptions (2.1), and where the conditions on $\alpha$ are relaxed to $0 \leq \alpha < \infty$. The result is given in Theorem 2.1 of [13]. In the following lemma we give the main part of Theorem 2.1 of [13] and provide its accompanying Table 1.

Lemma 3.1: Let $J_p$ be the block Jacobi matrix (1.7) associated with the linear system (1.1), where $A$ has the $p$-cyclic consistently ordered form (1.6), $p \geq 3$, and let $\sigma(J_p)$ satisfy (2.1), where the bound $-\frac{p}{p-2}$ on $\alpha$ is replaced by $\infty$. Assume that $A$ is repartitioned into a block $q$-cyclic consistently ordered form ($2 \leq q < p$) and denote by $\omega_p$ and $\rho_p$ the relaxation factor and the spectral radius of the optimal $q$-cyclic SOR. Let $r$ be the value of $q$ that gives the best cyclic repartitioning; i.e., the smallest optimal spectral radius $\rho_q$. Then the value of $r$ is given in Table 1, where the quantities $\alpha_{\ell,\ell+1}$ and $\beta_{\ell,\ell+1}$ in the table are found from the expressions

$$\alpha_{\ell,\ell+1} = \left( \frac{2^\ell \ell^{1/2} (1+\rho) \beta^{\ell+1}}{1-\rho} \right)^{\ell/p},$$

$$\beta_{\ell,\ell+1} = \left( \frac{2^\ell \ell^{1/2} (1-\rho) \alpha^{\ell+1}}{1+\rho} \right)^{\ell/p}. \tag{3.1}$$

In (3.1), $\rho$ is the unique root, in $(0,1)$, of the equation

$$\beta^p(\ell + \rho)^{\ell+1} - (\ell + 1)^{\ell+1} \rho = 0, \tag{3.2}$$

for $\alpha_{\ell,\ell+1}$, and of the equation

$$\alpha^p(\ell - \rho)^{\ell+1} - (\ell + 1)^{\ell+1} \rho = 0, \tag{3.3}$$

for $\beta_{\ell,\ell+1}$. The values of $\omega_r$ and $\rho_r$ are determined via (2.2)–(2.5), where in all these formulas, $r$ replaces $p$ and then $\alpha^{p/r}$ and $\beta^{p/r}$ replace $\alpha$ and $\beta$, respectively.

Note: As was pointed out in Kontovasilis, Plemmons and Stewart [20], Lemma 3.1 is of a more general value since it also covers the case of complex eigenvalues $\mu \in \sigma(J_p)$, with $\mu \in \mathbb{C}$, provided $|\mu| \leq \min\{\alpha, \beta\}$.
Suppose now that in the singular case of the $p$–cyclic consistently ordered matrix $A$ \( \text{index}(I - J_p) = 1 \). Suppose also that $\sigma(J_p^p)$ is given by (2.7), in which $0 \leq \alpha < \frac{p}{p-2}$ has been replaced by $0 \leq \alpha < \infty$. Suppose also that $A$ is repartitioned into a block $q$–cyclic consistently ordered form. From the analysis in §2 it is obvious that Lemma 2.2 and therefore Theorem 2.1 apply to the singular case for $q = p$ provided $\alpha < \frac{p}{p-2}$. For Lemma 2.2 to apply for any $2 \leq q \leq p - 1$ one must have

$$\omega \in (0,2) \setminus \left\{ \frac{q}{q-1} \right\}.$$  

This, however, assumes that the relationship

$$\text{index}(I - J_q) = 1, \quad q = 2, \ldots, p - 1,$$  

is valid. There are certainly cases of vital practical importance where the implication (3.4) is a straightforward consequence of some further property of the matrix $A$. For example, (3.4) follows directly when $A$ is a singular irreducible $M$-matrix (see, e.g, [31]), as in the case of the matrix coefficient in the Markov Chain problem with $A = I - PT$, where $P$ is the transition probability matrix and the chain is ergodic. However, (3.4) is always true under the assumption index$(I - J_p) = 1$. This is stated in the following theorem whose proof is given in the Appendix.

**Theorem 3.1:** Let $J_p$ be the block Jacobi matrix (1.7) associated with the linear system (1.1), where $A$ has the $p$–cyclic consistently ordered form (1.6), $p \geq 3$, and let index$(I - J_p) = 1$. Assume that $A$ is repartitioned into a block $q$–cyclic consistently ordered form (2 $\leq q < p$) and denote by $J_q$ the block Jacobi matrix corresponding to the new repartitioning. Then (3.4) holds.

So, under no further assumption, Theorem 2.1 holds for any $q$, when $\alpha^{p/q} < \frac{q}{q-2}$. Hence, a statement, let us call it Theorem 3.2 (which will not be formally stated here), completely analogous to Lemma 3.1 holds true. We note that since Theorem 3.2 refers to the singular case instead of the optimal spectral radius $\rho_q$, one must use the optimal semiconvergence factor $\gamma_q = \gamma(L_{w_q})$ in its place. Thus from Theorem 3.2 we obtain formulas that tell us how to repartition $A$ for optimal SOR convergence in order to compute the stationary distribution vectors of Markov chains (see Table 1).
4 Optimal Extended SOR

In this section we provide a convergence analysis for extended SOR convergence for the important 2-cyclic consistently ordered case. Here we assume that $J_2^2$ has all real eigenvalues with the same sign. It is shown that small perturbations around the optimal $\omega$ in the extended SOR method affect the convergence factor much less than for the usual SOR method. This formally confirms the validity of observations about numerical tests showing this phenomenon reported in [20].

We first introduce the following notation from [20]:

\begin{align}
\alpha(\omega) &:= \max\{|\lambda| := f_1(\lambda, \omega) = 0\} \geq 1 \\
\vartheta(\omega) &:= \max\{|\lambda| := f_\mu(\lambda, \omega) = 0, \mu \in \sigma(J_p), |\mu| < 1\} \\
r(\omega) &:= \vartheta(\omega)/\alpha(\omega),
\end{align}

where

\begin{equation}
f_\mu(\lambda, \omega) := \lambda^p - \omega \mu \lambda^{p-1} + \omega - 1. \tag{4.2}
\end{equation}

In [20] a detailed analysis led to the determination of the optimal parameter(s) $\omega_p$ and therefore of the optimal convergence factor $r_p(\omega_p)$ of the extended SOR in the case of nonnegative spectra $\sigma(J_p)$. In this section we study in more detail the behavior of the asymptotic convergence factor of the extended SOR in both cases of the nonnegative ($\alpha = 0$) and the nonpositive ($\beta = 0$) spectra $\sigma(J_2^2)$ (see (2.7)). Also, the detailed study of the behavior of $r_2(\omega)$ around $\omega_p$ will subsequently allow us to explain the phenomenon observed in [20]; namely, that small perturbations around $\omega_p$ affect the convergence factor in the extended SOR much less than small perturbations around the corresponding $\omega_p$ in the usual SOR. In both cases to be studied the interval for $\omega$ will be considered as being $(-\infty, \infty) \setminus (0, 2)$, i.e.,

\begin{equation}
\omega \in (-\infty, 0] \cup [2, \infty) \tag{4.3}
\end{equation}

while from (4.1)

\begin{equation}
f_1(\lambda, \omega) = \lambda^2 - \omega \lambda + \omega - 1 = 0 \tag{4.4}
\end{equation}

and

\begin{equation}
\ldots
\end{equation}
\[ f_\mu(\lambda, \omega) = \lambda^2 - \omega \mu \lambda + \omega - 1 = 0, \]  
with \( \mu \in \sigma(J_2) \) and \( |\mu| \neq 1 \). From (4.1), (4.3) and (4.4) we immediately obtain that

\[ \alpha(\omega) = |\omega - 1| \geq 1 \]  
while from (4.1) and (4.5)

\[ \vartheta(\omega) := \frac{1}{2} \max |\omega \mu \pm (\mu^2 \omega^2 - 4\omega + 4)^{1/2}|. \]  

**4.1 The Nonnegative Case**

In this case it is assumed that \( \sigma(J_2) \) satisfies the following conditions

\[ \sigma(J_2^2) \subset [0, \beta^2] \cup \{1\}, \quad 0 \leq \beta < 1. \]  

Let then

\[ 0 \leq \mu^2 \leq \beta^2, \quad \mu \in \sigma(J_2) \]

, where without loss of generality, we may consider \( \mu \) such that \( 0 \leq \mu \leq \beta \). Denoting the discriminant in (4.7) by

\[ D(\mu) := \mu^2 \omega^2 - 4\omega + 4 \]  
we readily have

\[ D(\beta) \geq D(\mu) \geq D(0). \]

So, we distinguish the three subcases a) \( D(\beta) \leq 0 \), b) \( D(0) \geq 0 \), and c) \( D(\beta) \geq 0 \geq D(0) \) which are studied separately.

In subcase (a), (4.5) has two complex conjugate roots of modulus \( |\omega - 1|^{1/2} \) each and this is the case when

\[ \frac{2}{1 + (1 - \beta^2)^{1/2}} \leq \omega \leq \frac{2}{1 - (1 - \beta^2)^{1/2}}, \]  
whence

\[ \vartheta(\omega) = |\omega - 1|^{1/2}. \]

Therefore, setting
Table 2:

<table>
<thead>
<tr>
<th>( \omega )</th>
<th>-( \infty )</th>
<th>0</th>
<th>- 2</th>
<th>( \omega )</th>
<th>( \infty )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r^2(\omega) )</td>
<td>( \beta^2 )</td>
<td>( \rightarrow )</td>
<td>1</td>
<td>( - 1 )</td>
<td>( r^2(\omega) )</td>
</tr>
</tbody>
</table>

\[
 r^2(\omega) = \frac{\beta^2(\omega)}{\alpha^2(\omega)} = \frac{1}{|\omega - 1|} < 1, \quad (4.11)
\]

we conclude that the extended SOR converges for all

\[
 \omega \in (2, \frac{2}{1 - (1 - \beta^2)^{1/2}}].
\]

From (4.11) it is also readily seen that \( r^2(\omega) \) is a strictly decreasing function of \( \omega \). Specifically as \( \omega \) increases from 2 to \( 2/(1 - (1 - \beta^2)^{1/2}) \), \( r^2(\omega) \) decreases from 1 to \((1 - (1 - \beta^2)^{1/2})/(1 + (1 - \beta^2)^{1/2}) \). After the study of subcases (b) and (c) takes place (see Appendix), the main result for the non-negative case, which is stated below, confirms the corresponding one obtained in [20].

**Theorem 4.1:** Let \( \sigma(J_2) \) satisfy (4.8) and let index\((I - J_2) = 1 \). Then, the extended SOR converges for all

\[
 \omega \in (-\infty, 0) \cup (2, \infty).
\]

The optimal relaxation factor is given by

\[
 \omega_2 = \frac{2}{1 - (1 - \beta^2)^{1/2}} \quad (4.12)
\]

while for the (optimal) convergence factor there holds

\[
 1 > r^2(\omega) > r^2(\omega_2) = \frac{1 - (1 - \beta^2)^{1/2}}{1 + (1 - \beta^2)^{1/2}}, \quad \omega_2 \neq \omega \in (-\infty, 0) \cup (2, \infty). \quad (4.13)
\]

Furthermore, the behavior of \( r^2(\omega) \) is illustrated in Table 2.

**Note:** As is seen from Table 2 when \( \omega \to \pm \infty \), \( r^2(\omega) \to \beta^2 \); in other words, the extended SOR converges (in the limiting cases) as fast as the usual Gauss-Seidel method.
4.2 The Nonpositive Case

This time it is assumed that \( \sigma(J_2) \) satisfies

\[
\sigma(J_2^2) \subset [-\alpha^2, 0] \cup \{1\}, \quad 0 \leq \alpha,
\]

so \( \tilde{\sigma}(J_2) := \sigma(J_2) \setminus \{-1, 1\} \) is purely imaginary. Let then \( i\mu \in \tilde{\sigma}(J_2) \) be any eigenvalue of \( J_2 \), where without loss of generality we may assume that \( 0 \leq \mu \leq \alpha \) whence \(-\alpha^2 \leq -\mu^2 \leq 0\). Obviously (4.4) will remain the same, leading again to the expression (4.6) for \( \alpha(\omega) \), while (4.5) will become

\[
f_{i\mu} = \lambda^2 - i\omega \mu \lambda + \omega - 1 = 0
\]

and hence

\[
\psi(\omega) = \frac{1}{2} \max |i\omega \mu \pm (-\mu^2 \omega^2 - 4\omega + 4)^{1/2}|.
\]

The discriminant \( D(\mu) \) in (4.16) is now

\[
D(\mu) := -\mu^2 \omega^2 - 4\omega + 4
\]

implying that

\[
D(0) \geq D(\mu) \geq D(\alpha).
\]

Three subcases are considered again. Specifically, a) \( D(\alpha) \geq 0 \), b) \( D(0) \leq 0 \), and c) \( D(0) \geq 0 \geq D(\alpha) \). As in the nonnegative case the simplest of the three subcases, that is (a), will be examined in the sequel. For this we have \( D(\mu) \geq 0 \) for all \( \mu \in [0, \alpha] \), so \( (D(\mu))^{1/2} \) is real. In view of the purely imaginary nature of \( i\omega \mu \), (4.15) has two complex roots having the same imaginary parts and opposite in sign real parts. Hence, the two roots have equal moduli, consequently

\[
\psi(\omega) = |\omega - 1|^{1/2}.
\]

However, from \( D(\alpha^2) \geq 0 \) it follows that

\[
\omega \in \left[ \frac{2}{1 - (1 + \alpha^2)^{1/2}}, \frac{2}{1 + (1 + \alpha^2)^{1/2}} \right]
\]

which together with
\[ r^2(\omega) = \varphi^2(\omega) = \frac{1}{\alpha^2(\omega) \left| \omega - 1 \right|} < 1, \quad (4.18) \]

which implies that \( \omega \in (-\infty, 0) \cup (2, \infty) \), gives

\[ \omega \in \left[ \frac{2}{1 - (1 + \alpha^2)^{1/2}}, 0 \right). \]

It can be readily found out that in the previous interval \( r^2(\omega) \) strictly increases from the value \( \frac{(1 + \alpha^2)^{1/2} - 1}{(1 + \alpha^2)^{1/2} + 1} \) to the value 1.

After the examination of the subcases (b) and (c) (see Appendix) the main result of the present section can be stated as follows:

**Theorem 4.2:** Let \( \sigma(J_2) \) satisfy (4.14) and let \( \text{index}(I - J_2) = 1 \). Then, the extended SOR converges for all

1. \( \omega \in (-\infty, 0) \cup (2, \infty) \) iff \( \alpha < 1 \)
2. \( \omega \in (-\infty, 0) \) iff \( \alpha = 1 \)
3. \( \omega \in (2, \infty) \) iff \( \alpha > 1 \).

The optimal relaxation factor, in all three cases, is given by

\[ \omega_2 = \frac{2}{1 - (1 + \alpha^2)^{1/2}}, \quad (4.19) \]

while for the (optimal) convergence factor there holds

\[ r^2(\omega) > r^2(\omega_2) = \frac{(1 + \alpha^2)^{1/2} - 1}{(1 + \alpha^2)^{1/2} + 1}, \quad \omega_2 \neq \omega \in (-\infty, 0) \cup (2, \infty). \quad (4.20) \]

Moreover the behavior of \( r^2(\omega) \) in each one of the three cases is illustrated in Tables 3i, 3ii and 3iii, respectively.

**Note:** As in the nonnegative case when \( \omega \to \pm\infty \), \( r^2(\omega) \to \alpha^2 \) and the extended (SOR) converges (or diverges) as fast as the usual Gauss-Seidel method.
Table 3: (3i) ($\alpha < 1$); (3ii) ($\alpha = 1$); (3iii) ($\alpha > 1$);

### 3i

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$-\infty$</th>
<th>$\omega_2$</th>
<th>$0$</th>
<th>$-2$</th>
<th>$\frac{2}{1-\alpha}$</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2(\omega)$</td>
<td>$\alpha^2$</td>
<td>$\downarrow$</td>
<td>$r^2(\omega_2)$</td>
<td>$\nearrow$</td>
<td>$1$</td>
<td>$\downarrow$</td>
</tr>
</tbody>
</table>

### 3ii

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$-\infty$</th>
<th>$\omega_2 = -2(\sqrt{2} + 1)$</th>
<th>$0$</th>
<th>$-2$</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2(\omega)$</td>
<td>$1$</td>
<td>$\downarrow$</td>
<td>$r^2(\omega_2) = 3 - 2\sqrt{2}$</td>
<td>$\nearrow$</td>
<td>$1$</td>
</tr>
</tbody>
</table>

### 3iii

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$-\infty$</th>
<th>$\frac{2}{1-\alpha}$</th>
<th>$\omega_2$</th>
<th>$0$</th>
<th>$-2$</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2(\omega)$</td>
<td>$\alpha^2$</td>
<td>$\downarrow$</td>
<td>$1$</td>
<td>$\downarrow$</td>
<td>$r^2(\omega_2)$</td>
<td>$\nearrow$</td>
</tr>
</tbody>
</table>
4.3 The convergence factor around \( \omega_p \)

From the analysis in the proof of Theorem 4.1 regarding the nonnegative case, we have that

\[ r^2(\omega) = \frac{1}{\omega - 1}, \quad 2 < \omega \leq \omega_2 \]  

(4.21)

while

\[ r^2(\omega) = \frac{(\omega \beta + (D(\beta))^{1/2})^2}{4(\omega - 1)^2}, \quad \omega_2 \leq \omega < \infty. \]  

(4.22)

On differentiation, we obtain from (4.21) and (4.22) after some manipulation that

\[ \frac{\partial r^2(\omega)}{\partial \omega} = -\frac{1}{(\omega - 1)^2}, \quad 2 < \omega \leq \omega_2 \]  

(4.23)

and

\[ \frac{\partial r^2(\omega)}{\partial \omega} = \frac{(\omega \beta + (D(\beta))^{1/2})^2}{2(\omega - 1)^3} \left(-\beta + \frac{(2 - \beta^2)\omega - 2}{(D(\beta))^{1/2}}\right), \quad \omega_2 \leq \omega < \infty, \]  

(4.24)

respectively. Now, from (4.23) and (4.24) it can be readily obtained that

\[ \lim_{\omega \to \omega_2^-} \frac{\partial r^2(\omega)}{\partial \omega} = -\left(\frac{1 - (1 - \beta^2)^{1/2}}{1 + (1 - \beta^2)^{1/2}}\right)^2 \]  

(4.25)

and

\[ \lim_{\omega \to \omega_2^+} = +\infty, \]  

(4.26)

respectively. By straightforward calculations it can be found out that

\[ r^2(\omega_2 + \epsilon) > r^2(\omega_2 - \epsilon), \quad \epsilon \to 0^+. \]

In other words the tangent to the left branch of the graph of the curve \( r^2(\omega) \) at the point with abscissa \( \omega = \omega_2 \) intersects the \( \omega^- \) axis with an angle \( \theta \) such that

\[ \tan \theta = -\left(\frac{1 - (1 - \beta^2)^{1/2}}{1 + (1 - \beta^2)^{1/2}}\right)^2 \quad \text{and} \quad 145^0 < \theta < 180^0, \]
with
\[ \lim_{\beta \to 0^+} = 180^\circ \quad \text{and} \quad \lim_{\beta \to 1^-} = 145^\circ, \]
while the tangent to the right branch of the same graph at the same point is perpendicular to the \( w \)-axis. This situation is similar to the corresponding one in the classical SOR case when we consider small perturbations around the corresponding optimal \( \omega \) (see e.g., [35], [39]). So, in our present case of the extended SOR an underestimation rather than an overestimation of \( \omega_2 \) should be preferred in practice since it gives a smaller value for \( r^2(\omega) \). Moreover, one should always bear in mind that the values of \( r^2(\omega) \) at \( \omega_2 \) and as \( \omega \to \infty \) are \( \frac{1-(1-\beta^2)^{1/2}}{1+(1-\beta^2)^{1/2}} \) and \( \beta^2 \), respectively, meaning that even if we overestimate \( \omega_2 \) by a large amount the value of \( r^2(\omega) \) is not going to be affected very much. This fact makes the graph of \( r^2(\omega) \) look like a flat one around \( \omega_2 \) despite the fact that the two branches of the curve, at that very point, meet at an angle \( \phi \in (45^\circ, 90^\circ) \). The previous analysis explains in a satisfactory way the phenomenon observed by Kontovasilis, Plemmons and Stewart [20] for \( p \geq 3 \).

In Figure 1 the situation that was analyzed above is illustrated for \( \beta = 0.8 \). For comparison reasons both the graph of the spectral radius of the SOR iteration matrix \( \rho(L_\omega) \) in the interval \([1, 2]\) and the graph of \( r^2(\omega) \) in the interval \([2, 10]\) are juxtaposed. If \( \omega_2^c \) and \( \omega_2^e \) denote the optimal values of \( \omega \) in the classical and in the extended SOR case, respectively, then in the case of Figure 1 it is

\[ \omega_2^c = 1.25 \quad \text{and} \quad \omega_2^e = 5 \]

while
\[ \rho(L_{\omega_2^e}) = r^2(\omega_2^e) = 0.25. \]

A similar analysis to the the previous one based on the proof of Theorem 4.2 for the nonpositive case leads to the following results:
\[ \lim_{\omega \to \omega_2^c} \frac{\partial r^2(\omega)}{\partial \omega} = \left( \frac{(1 + \alpha^2)^{1/2} - 1}{(1 + \alpha^2)^{1/2} + 1} \right)^2 \quad (4.27) \]
and
Figure 1: Asymptotic convergence factor of the classical and extended SOR for the nonnegative case ($\beta = 0.8$).

\[
\lim_{\omega \rightarrow \omega_2^-} \frac{\vartheta r^2(\omega)}{\vartheta \omega} = -\infty, \quad (4.28)
\]

while

\[ r^2(\omega_2 - \epsilon) > r^2(\omega_2 + \epsilon), \quad \epsilon \rightarrow 0^+. \]

This time the tangent to the right branch of the graph of the curve $r^2(\omega)$ at the point $\omega = \omega_2$ and the $\omega-$ axis form an angle $\theta$ between $0^0$ and $45^0$ while the tangent to the left branch of the same graph at the same point is perpendicular to the $\omega-$ axis. Again we have a similar situation to that in the classical SOR case for small perturbations around the corresponding optimal $\omega$. In our present case of the extended SOR an overestimation rather than an underestimation of $\omega_2$ should be preferred in practice. Again the phenomenon of the flatness of the graph of the curve $r^2(\omega)$ around $\omega_2$ can be observed.

In Figure 2 the situation in the present case is illustrated with $\alpha = 0.8$. This time it is

\[ \omega_2^c \approx -7.127 \text{ and } \omega_2^c \approx 0.8770 \]

while

\[ r^2(\omega_2^c) = \rho(C_{\omega_2^c}) \approx 0.12305. \]

5 Applications to Markov Chains

As indicated in Section 1, applications of the results in this paper for the case of singular coefficient matrices $A$ include the solution to Markov Chain
problems. The problem to be solved is that of computing the probability distribution row vector $\pi$, where $\pi P = \pi$, $\|\pi\|_1 = 1$ or, equivalently, for $A = I - P^T$

$$A\pi^T = (I - P^T)\pi^T = 0, \|\pi\|_1 = 1.$$ 

Here $P$ is the stochastic transition probability matrix associated with the chain. The coefficient matrix $A$ is then a singular, irreducible (for the ergodic case) M-matrix. For large-scale Markov chains the use of iterative methods to compute $\pi$ is of prime importance and has been studied extensively, e.g., see the surveys in Berman and Plemmons [2], Courtois and Semal [10] and O’Leary [28].

A Nearly Completely Decomposable (NCD), or nearly uncoupled, stochastic matrix $P$ is one that can be assembled into the block form $P = (P_{ij})$ in which the diagonal blocks $P_{ii}, i = 1,\ldots,q$, are square and have components that are large compared with those of off-diagonal blocks. These matrices arise in problems whose components can be grouped into aggregates that are loosely connected to one another, e.g., Courtois [9] and Schweitzer [32]. Aggregation/disaggregation methods are an important class of algorithms which are used to compute the stationary probability vector $\pi$ of large scale Markov chains. For those chains that are NCD, iterative aggregation/disaggregation techniques can sometimes result in sequences which converge at surprisingly rapid rates.

It is clear that aggregation/disaggregation processes bear a close resemblance to multigrid or multilevel techniques. In a sense, aggregation corresponds to moving to a coarse level, while disaggregation corresponds to moving to a fine level. However, this relationship has evidently not been investigated in detail. Multigrid methods were first developed as fast solvers
for elliptic PDE. Surveys of these methods can be found in Briggs [5] and McCormick [23, 24]. In the last few years efforts have been made to extend the multilevel philosophy to problems without the geometrical background provided by PDEs. Purely algebraic algorithms have been developed that do not make use of geometrical neighborhood relations between gridpoints and unknowns. These algorithms are called Algebraic Multigrid (AMG) methods.

Multilevel families of methods arise from analyzing problems at various levels of granularity. For example, a course level approximation may be able to provide an approximate information for a more detailed model. A multilevel approach to the solution of a problem takes advantage of the convergence of discrete approximations as the refinement converges to infinity. The basic inner iteration used for AMG processes is typically some classical scheme—often point or block Gauss-Seidel [5, 23]. These iterations are called smoothers for the AMG process. It is possible to take advantage of the periodicity property of Markov chains when using aggregation/disaggregation schemes [3]. Thus p-cyclic iterations might very well be considered as a smoother for AMG methods applied to the solution of periodic Markov chains, but more work needs to be done on this topic.

Another possible application of p-cyclic iterations for Markov chains might be in conjunction with preconditioners for conjugate gradient methods e.g., [30]. Chan [8] has considered chains with overflow capacity and non-rectangular state spaces. He solves these problems by preconditioned conjugate gradient methods, making use of results from domain decomposition for elliptic partial differential equations. Separable preconditioners are considered. One might also consider the use of p-cyclic iterations for NCD Markov chains as a possible block preconditioning scheme. Again, more work on this topic needs to be done.

References


6 APPENDIX

Proof of Theorem 2.1:
Before we go on with the proof we state and prove the following lemma.
Lemma: For the function
\[
g(x) := \frac{(x + y)^p}{x}, \quad x \in (0, 1],
\]
where \(y \geq p - 1\), independent of \(x\), and \(p \geq 2\) there holds
\[
\min_{x \in (0, 1]} g(x) = (1 + y)^p.
\]

Proof: Differentiate (6.1) to obtain
\[
g'(x) = \frac{(x + y)^{p-1}}{x^2} ((p - 1)x - y) \leq 0,
\]
with equality holding iff \(x = 1\) and \(y = p - 1\). So, \(g(x)\) is strictly decreasing in \((0, 1]\) which implies the validity of (6.2).\(\Box\)

To simplify the notation we drop the index \(p\) from \(\alpha_p\) and \(\beta_p\) and distinguish two cases. The basic case \(\alpha \neq \beta\) and the trivial one \(\alpha = \beta \in (0, 1)\).
When \(\alpha \neq \beta\) it is either \(\beta > \alpha\) or \(\beta < \alpha\). In the former case we have from (2.4) \(\omega_p \in (1, 1 + \frac{\beta - \alpha}{\beta + \alpha})\), implying from (2.5) that
\[
0 < \omega_p - 1 < \frac{\beta - \alpha}{\beta + \alpha} \leq \frac{1}{p - 1}
\]
while in the latter case \(\omega_p \in (1 + \frac{\beta - \alpha}{\beta + \alpha}, 1)\) and therefore
\[
0 > \omega_p - 1 > \frac{\beta - \alpha}{\beta + \alpha} \geq -\frac{1}{p - 1}.
\]
From (6.3) and (6.4) we have \(\omega_p \in (\frac{p - 2}{p - 1}, \frac{p}{p - 1}) \setminus \{1\}\). So from Lemma 2.2, \(\text{index}(I - \mathcal{L}_{\omega_p}) = 1\). Also in both cases, in view of (6.3) and (6.4), there hold
\[
0 < |\omega_p - 1| < \frac{|\beta - \alpha|}{|\beta + \alpha|} \leq \frac{1}{p - 1}.
\]

Consider now the polynomial
\[ f(\lambda, \omega) := (\lambda + \omega - 1)^p - \omega^p \lambda^{p-1} \] (6.6)

which, if it is set equal to zero, gives as roots \( \lambda \neq 1 \) the images of the eigenvalues \( \mu = \exp(i2\pi q/p) \), \( q = 1(1)p - 1 \), via (1.8). Let us put

\[ y = \omega - 1, \quad y_p = \omega_p - 1, \quad z = \frac{\beta + \alpha}{\beta - \alpha} \] (6.7)

and divide out \( f(\lambda, \omega) \) by \( \lambda - 1 \) to obtain

\[ g(\lambda, y) := \frac{f(\lambda + y)}{\lambda - 1} = \lambda^{p-1} - \left[ (\frac{p}{2})y^2 + \ldots + (\frac{p}{p})y^p \right] \lambda^{p-2} - \ldots - \left[ (\frac{p}{p-1})y^{p-1} + (\frac{p}{p})y^p \right] \lambda - (\frac{p}{p})y^p. \] (6.8)

To prove the theorem it suffices to prove that \( g(\lambda, y_p) \) has zeros \( \lambda \) with moduli strictly less than \( \rho_p \), where

\[ \rho_p = \frac{(\beta + \alpha)}{(\beta - \alpha)}(\omega_p - 1) = zy_p = |z||y_p|. \] (6.9)

Obviously, \( \lambda = 0 \) is not a zero of (6.8) since in view of \( \alpha \neq \beta \) it is \( \omega_p \neq 1 \) and hence \( g(0, y_p) = -(\omega_p - 1)^p \neq 0 \). To show that \( |\lambda| < \rho_p \) we set

\[ \nu = \frac{\lambda}{zy_p}, \quad a = |z|, \quad b = |y_p| \] (6.10)

and will show that \( |\nu| < 1 \). For this we form the following polynomial in \( \nu \)

\[ h(\nu) := \frac{g(\lambda, y_p)}{(zy_p)^{p-1}} = \frac{g(z\nu y_p, y_p)}{(zy_p)^{p-1}} \]

\[ = \nu^{p-1} - \frac{1}{zy_p} \left[ (\frac{p}{2})y_p^2 + \ldots + (\frac{p}{p})y_p^p \right] \nu^{p-2} \]

\[ - \ldots - \frac{1}{(zy_p)^{p-1}}(\frac{p}{p})y_p^p. \] (6.11)

Since \( g(0, y_p) \neq 0 \), \( h(0) \neq 0 \), so we put \( h(\nu) = 0 \), solve for \( \nu^{p-1} \) in (6.11) and divide through by \( \nu^{p-2} \) to obtain
\[
\nu = \frac{1}{(zp)}[(\frac{p}{2})y_p^2 + \ldots + (\frac{p}{p})y_p^p] \\
+ \frac{1}{(zp)^2}[(\frac{p}{3})y_p^3 + \ldots + (\frac{p}{p})y_p^{p\frac{1}{2}}] \\
+ \ldots \\
+ \frac{1}{(zp)^{p-1}}(\frac{p}{p})y_p^{p\frac{1}{2}}.
\]

(6.12)

Suppose now that there exists a \( \nu \) satisfying (6.12) and such that \( |\nu| \geq 1 \). Then using (6.10) in (6.12) one obtains

\[
|\nu| \leq \frac{1}{ab}[(\frac{p}{2})b^2 + \ldots + (\frac{p}{p})b^p] \\
+ \frac{1}{a^2b}[(\frac{p}{3})b^3 + \ldots + (\frac{p}{p})b^p] \\
+ \ldots \\
+ \frac{1}{a^{p-1}b^{p-1}}(\frac{p}{p})b^p \\
= \frac{ab}{1-ab}[(\frac{1}{a^2} - \frac{b}{a})(\frac{p}{2}) + (\frac{1}{a^2} - \frac{b}{a})(\frac{p}{3}) \\
+ \ldots + (\frac{1}{a^p} - \frac{b^{p-1}}{a})(\frac{p}{p})] \\
= \frac{ab}{1-ab}[(1 + \frac{1}{a})^p - 1 - \frac{p}{a}] - \frac{1}{ab}[(1 + b)^p - 1 - pb]
\]

or

\[
|\nu| \leq 1 + \frac{b}{1 - ab}[(1 + a)^p - (1 + b)^p].
\]

(6.13)

Using (6.9) and (6.10), (6.13) becomes

\[
|\nu| \leq 1 + \frac{\rho_p}{(1 - \rho_p)a^p}[(1 + a)^p - (\rho_p + a)^p].
\]

(6.14)

Since \( \rho_p \in (0, 1) \) is a function of \( a = |z| \geq p - 1 \), then by virtue of the Lemma one has

\[
\frac{(\rho_p + a)^p}{\rho_p} \geq \min_{x \in (0,1]} \frac{(x + a)^p}{x} \geq \min_{x \in (0,1]} \frac{(x + a)^p}{x} = (1 + a)^p,
\]

28
where we note that the second inequality from the left is a strict one since the minimum on $(0,1]$ is attained at $x = 1$. Consequently, the difference in the brackets in (6.14) is strictly negative implying that $|\nu| < 1$. This contradicts our assumption that $|\nu| \geq 1$. Therefore no zero ($\lambda$) of $g(\lambda, y_p)$ of (6.8) can be in modulus greater than or equal to $\rho_p$ which effectively proves the theorem in the basic case $\alpha \neq \beta$. If, on the other hand, we have the trivial case $\alpha = \beta > 0$ then $\omega_p = 1$, implying that $g(\lambda, 1) = \frac{f(\lambda, 1)}{\lambda - 1} = \lambda^{p-1}$ whose (all) zeros are equal to 0 and are, therefore, strictly less than $\rho_p = \alpha^p = \beta^p \geq 0$, which concludes the proof of the theorem. □

**Proof of Theorem 3.1**

Let the block Jacobi matrix $J_p$ given in (1.7) satisfy the assumption $\text{index}(I - J_p) = 1$ and let $s$ be the multiplicity of the eigenvalue 1 of $J_p$. Each of the $s$ eigenvalues equal to 1 will be associated with $1 \times 1$ blocks in the Jordan canonical form of $J_p$. Let then

$$\Psi_{\beta_0}^{(\ell)} = [\psi_1^{(\ell)T}, \psi_2^{(\ell)T}, \ldots, \psi_p^{(\ell)T}]^T, \quad \ell = 1, 2, \ldots, s,$$  \hspace{1cm} (6.15)

be any one of the $s$ linearly independent eigenvectors of $J_p$ associated with the eigenvalue 1, where $\Psi_{\beta_0}^{(\ell)}$ has been partitioned in accordance with $J_p$. Then from $J_p \Psi_{\beta_0}^{(\ell)} = \Psi_{\beta_0}^{(\ell)}, \ell = 1, 2, \ldots, s$, it is readily obtained from (1.7) and (6.15) that

$$C_1 \psi_{1}^{(\ell)} = \psi_1^{(\ell)},$$

$$C_2 \psi_{2}^{(\ell)} = \psi_2^{(\ell)},$$

$$C_3 \psi_{3}^{(\ell)} = \psi_3^{(\ell)},$$

$$\vdots$$

$$C_{p-1} \psi_{p-2}^{(\ell)} = \psi_{p-1}^{(\ell)},$$

$$C_p \psi_{p}^{(\ell)} = \psi_p^{(\ell)},$$  \hspace{1cm} (6.16)

and from (6.16)
\[
\psi_1^{(\ell)} = C_1 C_p C_{p-1} \ldots C_2 \psi_1^{(\ell)}, \\
\psi_2^{(\ell)} = C_2 C_1 C_p \ldots C_3 \psi_2^{(\ell)}, \\
\psi_3^{(\ell)} = C_3 C_2 C_1 \ldots C_4 \psi_3^{(\ell)}, \\
\vdots \\
\psi_{p-1}^{(\ell)} = C_{p-1} C_{p-2} C_p \psi_{p-1}^{(\ell)}, \\
\psi_p^{(\ell)} = C_p C_{p-1} C_{p-2} \ldots C_1 \psi_p^{(\ell)}. 
\]

(6.17)

It is clear from (6.16) that all \(\psi_j^{(\ell)} \neq 0, j = 1, 2, \ldots, p\). For if \(\psi_j^{(\ell)} = 0\), for some \(j\), then from the \((j+1)^{st}\) equation in (6.16) it is obtained that \(\Psi_{j+1}^{(\ell)} = 0\), and from the remaining equations taken in a cyclic order, \(\psi_{j+2}, \ldots, \psi_{j+p-1} = 0\). (It is understood that if an index in either \(\psi_j^{(\ell)}\) or \(C_j\) exceeds \(p\) it will be considered as being modulo \(p\).) The result just obtained implies, in turn, that \(\Psi_{\beta_0}^{(\ell)} = 0\) which contradicts the fact that \(\Psi_{\beta_0}^{(\ell)}\) is an eigenvector of \(J_p\). Another implication of the previous result is that \(\psi_j^{(\ell)}\) is an eigenvector of the cyclic product \(C_{j} C_{j-1} \ldots C_{p} C_{1} \ldots C_{j-1}, j = 1, 2, \ldots, p\), with 1 as its corresponding eigenvalue. However, due to the cyclic nature of \(J_p\), besides 1, \(J_p\) will also have as eigenvalues of modulus 1 the numbers \(\beta_k = e^{i2k\pi/p}, k = 1, 2, \ldots, p-1\), with a multiplicity \(s\) each. This is an immediate consequence of Romanovski's Theorem (see Theorem 2.4 of [35]). Each of the aforementioned \(s \times (p-1)\) eigenvalues of modulus 1 will be associated with \(1 \times 1\) blocks in the Jordan canonical form of \(J_p\). This follows directly from Theorem 1 of Courtois and Semal [10] according to which to each \(\beta_k\), in view of (6.1), there corresponds an eigenvector of the form

\[
\psi_{\beta_k}^{(\ell)} = [\psi_1^{(\ell)} \beta_k^{-1} \psi_1^{(\ell)}, \ldots, \beta_k^{-(p-1)} \psi_1^{(\ell)}], 
\]

(6.18)

\[k = 1, 2, \ldots, p-1, \quad \ell = 1, 2, \ldots, s.\]

From the linear independence of \(\psi_{\beta_k}^{(\ell)}\)'s one concludes the linear independence of \(\psi_{\beta_k}^{(\ell)}\)'s for each \(k\) and also the linear independence of all \(s \times p\) eigenvectors which are associated with the \(s \times p\) eigenvalues of modulus 1. Consider now \(J_p\), given specifically by
\[
J_p^s = \begin{bmatrix}
C_1C_pC_{p-1}\ldots C_2 \\
C_2C_1C_p\ldots C_3 \\
C_3C_2C_1\ldots C_4 \\
\vdots \\
C_{p-1}C_pC_{p-2}\ldots C_3 \\
C_pC_{p-1}C_{p-2}\ldots C_1
\end{bmatrix}.
\]

\[ (6.19) \]

\( J_p^s \) has \( s \times p \) eigenvalues equal to 1 associated with \( s \times p \) linearly independent eigenvectors, namely the ones described in (6.15) and (6.18). This is deduced if one considers the Jordan canonical form for \( J_p \) and then that for \( J_p^s \). However, from (6.17) and (6.19) it is easily concluded that the \( s \) vectors \( \psi_1^{(\ell)} \) will give rise to \( s \) linearly independent eigenvectors of the form

\[ \psi_1^{(\ell)} = [\psi_1^{(\ell)T}, 0^T, \ldots, 0^T]^T \]

associated with eigenvalues equal to 1 of \( J_p^s \), similarly the \( s \) vectors \( \psi_2^{(\ell)} \) will give rise to another \( s \) linearly independent eigenvectors

\[ \psi_2^{(\ell)} = [0^T, \psi_1^{(\ell)T}, 0^T, \ldots, 0^T]^T \]

associated with eigenvalues equal to 1 of \( J_p^s \) and so on. Hence \( \psi_1^{(\ell)}, \psi_2^{(\ell)}, \ldots, \psi_p^{(\ell)} \), \( \ell = 1, 2, \ldots, s \), which are linearly independent eigenvectors, are associated with the \( s \times p \) eigenvalues of \( J_p^s \) equal to 1. Consequently all the eigenvalues of \( J_p^s \) equal to 1 are associated with \( 1 \times 1 \) blocks in the Jordan canonical form of \( J_p^s \) implying that

\[ \text{index}(I - J_p^s) = 1. \quad (6.20) \]

This also implies that the set of the \( s \) eigenvalues equal to 1 of each cyclic product in (6.17) that constitute the diagonal blocks of \( J_p^s \) in (6.19) are associated with \( 1 \times 1 \) blocks in the Jordan canonical form of the corresponding cyclic product \( C_jC_{j-1}\ldots C_1C_p\ldots C_{j+1}, j = 1, 2, \ldots, p. \) Therefore

\[ \text{index}(I - C_jC_{j-1}\ldots C_1C_p\ldots C_{j+1}) = 1, \quad j = 1, 2, \ldots, p. \quad (6.21) \]

Consider now the \( q \)-cyclic consistently ordered repartitioning of \( A \) in (1.6). To illustrate how our objective, namely (3.4), can be obtained we will give the
proof by means of a particular example and then the generalization follows easily. For this consider $p = 7$ in (1.6) and (1.7) and repartition $A$ in the way this matrix was repartitioned in [29] and also in [13]. Suppose then that $q = 3$ and that the particular repartitioning indicated below is considered

$$
A = \begin{bmatrix}
A_1 & & & & & & B_1 \\
B_2 & A_2 & & & & & \\
& B_3 & A_3 & & & & \\
& B_4 & A_4 & & & & \\
& & B_5 & A_5 & & & \\
& & & B_6 & A_6 & & \\
& & & & B_7 & A_7 & \\
\end{bmatrix}.
$$

(6.22)

It can be obtained that the block diagonal form of $J_3^3$ is as follows

$$
J_3^3 = \begin{bmatrix}
0 & C_1C_p \ldots C_3 & & & & & \\
& 0 & C_2C_1C_p \ldots C_3 & & & & \\
& & 0 & 0 & C_3C_2 \ldots C_6 & & \\
& & & 0 & 0 & C_4C_3C_2 \ldots C_6 & \\
& & & & 0 & 0 & C_5C_4C_3 \ldots C_6 \\
& & & & & 0 & C_6C_5 \ldots C_1 \\
& & & & & & 0 & C_7C_6C_5 \ldots C_1 \\
\end{bmatrix}.
$$

(6.23)

So, $J_3^3$ considered as a $7 \times 7$ block matrix is a block upper triangular matrix, with diagonal blocks $0, C_2C_1C_p \ldots C_3, 0, 0, C_5C_4C_3 \ldots C_6, 0, C_7C_6C_5 \ldots C_1$, respectively. From the results obtained on the cyclic products of $J_3^p$ ($p = 7$) in (6.19) and (6.21), it follows that $J_3^3$ has $s \times 3$ (instead of $s \times 7$) eigenvalues equal to 1 and these eigenvalues are associated with $1 \times 1$ blocks in the Jordan canonical form of $J_3$. Moreover, there are now $s \times 3$ (instead of $s \times 7$) linearly independent eigenvectors associated with these eigenvalues, specifically $\Psi^{(\ell)}_2, \Psi^{(\ell)}_5, \Psi^{(\ell)}_7, \ell = 1, 2, \ldots, s$. This implies that

$$\text{index}(I - J_3^3) = 1.
$$

(6.24)

If now $\text{index}(I - J_3) = r > 1$, then there would be at least one $r \times r$ block in the Jordan canonical form of $J_3$ associated with the eigenvalue 1 which, in turn, would imply that $\text{index}(I - J_3^3) = r > 1$, a contradiction to (6.24). Therefore $\text{index}(I - J_3) = 1$ and in general

32
index\( (I - J_q) = 1, \quad q = 2, 3, \ldots, p - 1, \) \hspace{1cm} \text{(6.25)}

which concludes the proof of the theorem. \( \square \)

**Proof of Theorem 4.1**

We are to examine subcases (b) and (c).

In subcase (b), \( D(0) \geq 0, \) implying \( \omega < 1, \) and therefore in view of (4.3), \( \omega \in (-\infty, 0). \) Hence, the largest of the two moduli in (4.7) is given by

\[ |\lambda| = \frac{1}{2}(-\omega \mu + (D(\mu))^{1/2}). \] \hspace{1cm} \text{(6.26)}

Introducing the symbol \( \sim \) to denote equality of signs of two expressions we readily obtain

\[ \partial|\lambda|/\partial \mu \sim -\omega(D(\mu))^{1/2} + \mu \omega^2 > 0. \]

This implies that

\[ \vartheta(\omega) = \max_\mu |\lambda| = \frac{1}{2}(-\omega \beta + (D(\beta))^{1/2}) \]

and since \( \alpha(\omega) = 1 - \omega \) it is obtained that

\[ r^2(\omega) = \frac{(-\omega \beta + (D(\beta))^{1/2})^2}{4(1 - \omega)^2}. \] \hspace{1cm} \text{(6.27)}

Since we are interested in finding out for which \( \omega \)'s, \( r^2(\omega) < 1, \) we have from (6.27) after some algebra takes place that \( r^2(\omega) < 1 \) is equivalent to

\[ (D(\beta))^{1/2} < 2(1 - \omega) + \omega \beta. \]

Since the right hand side of the inequality above is \( 2 - \omega (2 - \beta) > 0, \) we square both members to obtain equivalently that

\[ (1 - \omega)(1 - \beta^2) > 0 \]

which is always true. Hence the extended SOR converges for all \( \omega \in (-\infty, 0). \)

To study the behavior of \( r^2(\omega) \) we differentiate this function with respect to \( \omega \) to obtain, after some algebra,
\[ \frac{\partial r^2(\omega)}{\partial \omega} \sim \frac{\partial r(\omega)}{\partial \omega} \sim \]

\[ (1 - \omega)(-\beta + \frac{\beta^2 \omega - 2}{(D(\beta))^{1/2}}) + (-\omega \beta + (D(\beta))^{1/2}) \sim \]

\[ -\beta(D(\beta))^{1/2} + (1 - \omega)(\beta^2 \omega - 2) + D(\beta) \sim \]

\[ -\beta(D(\beta))^{1/2} + (2 - \omega(2 - \beta^2)) \sim \]

\[ (2(1 - \omega) + \beta^2 \omega)^2 - \beta^2 D(\beta) \sim (1 - \omega)(1 - \beta^2) > 0. \]

Therefore \( r^2(\omega) \) strictly increases in \((-\infty, 0)\) and especially \( \lim_{\omega \to -\infty} r^2(\omega) = \beta^2 \) while \( r^2(0) = 1 \). This concludes the study of subcase (b).

In subcase (c) we have \( D(\beta^2) > 0 \geq D(0) \) which together with (4.3) gives \( \omega \in \left[ \frac{2}{1 - (1 - \beta^2)^{1/2}}, \infty \right) \). It is obvious that in this subcase there always exists a unique value of \( \mu \in (0, \beta) \) denoted by \( \tilde{\mu} := 2(\omega - 1)^{1/2} \) such that \( D(\tilde{\mu}) = 0 \). If \( \mu \in [0, \tilde{\mu}] \) then for any fixed \( \omega \in \left[ \frac{2}{1 - (1 - \beta^2)^{1/2}}, \infty \right) \)

\[ \vartheta(\omega) = (\omega - 1)^{1/2} \]

while if \( \mu \in [\tilde{\mu}, \beta] \) the largest of the two moduli in (4.7) is

\[ |\lambda| = \frac{1}{2} (\omega \mu + (D(\mu))^{1/2}) \]

and therefore

\[ \partial |\lambda|/\partial \mu \sim \omega(D(\mu))^{1/2} + \mu \omega^2 > 0. \]

Consequently,

\[ \vartheta(\omega) = \max_\mu |\lambda| = \frac{1}{2} (\omega \beta + (D(\beta))^{1/2}). \quad (6.28) \]

Since for a fixed \( \omega \), \( |\lambda| \) as a function of \( \mu \) strictly increases in \([\tilde{\mu}, \beta]\), it is implied that out of the two expressions for \( \vartheta(\omega) \) found in this present subcase the expression in (A.28) is the one that gives it over all values \( \mu \in [0, \tilde{\mu}] \cup [\tilde{\mu}, \beta] = [0, \beta] \). Forming then \( r^2(\omega) \) and setting \( r^2(\omega) < 1 \), we have
After some algebra takes place we find out that the last inequality is equivalent to the following valid one \( 3 < 1 \). So, the extended SOR converges for all \( \omega \in \left[ \frac{2}{1-(1-\beta^2)^{1/2}}, \infty \right) \). The behavior of \( r^2(\omega) \) is determined by the sign of \( \partial r^2(\omega) / \partial \omega \). For this we have, after some algebra, that

\[
\partial r^2(\omega) / \partial \omega \sim (\omega - 1)(1 - \beta^2) > 0.
\]

Consequently, \( r^2(\omega) \) strictly increases in the interval of interest and there also holds that \( \lim_{\omega \to \infty} r^2(\omega) = \beta^2 \).

All the convergence and the optimal convergence results as well as the monotonicity behavior of the convergence factor \( r^2(\omega) \) in the present nonnegative case are summarized in formulas (4.12)-(4.13) and in Table 2. This concludes the proof of Theorem 4.1.\( \square \)

**Proof of Theorem 4.2**

Subcases (b) and (c) are the ones to be examined here. For subcase (b), \( D(0) \leq 0 \) and therefore \( \omega \in (2, \infty) \). From (4.15)-(4.17) both roots are purely imaginary and therefore

\[
|\lambda| = \frac{1}{2}(\omega \mu + (-D(\mu))^{1/2}).
\]

Since it is readily checked that

\[
\partial |\lambda| / \partial \mu \sim \omega \mu + (-D(\mu))^{1/2} > 0
\]

it is implied that

\[
\partial(\omega) = \frac{1}{2}(\omega \alpha + (-D(\alpha))^{1/2}).
\]

Consequently

\[
 r^2(\omega) = \frac{(\omega \alpha + (-D(\alpha))^{1/2})^2}{4(\omega - 1)^2}
\]

and requiring \( r^2(\omega) < 1 \) is equivalent to
\((-D(\alpha))^{1/2} < (2 - \alpha)\omega - 2.\) \hfill (6.31)

In view of \(\omega > 2\), necessary conditions for the latter inequality to hold are

\[ \alpha < 2 \quad \text{and} \quad \omega > \max\{2, 2/(2 - \alpha)\}. \] \hfill (6.32)

Under (6.32), (6.31) is equivalent to \((1 - \alpha)\omega - 2 > 0\) which holds true if and only if

\[ \alpha < 1 \quad \text{and} \quad \omega > \frac{2}{1 - \alpha} > 2 > \frac{2}{2 - \alpha}. \]

For the behavior of \(r^2(\omega)\) we have

\[ \frac{\partial r^2(\omega)}{\partial (\omega)} \sim -\alpha(-D(\alpha))^{1/2} - (\omega - 2) - \omega(1 + \alpha^2) < 0. \]

So \(r^2(\omega)\), as a function of \(\omega\), strictly decreases in the interval \((2, \infty)\) and especially \(\lim_{\omega \to \infty} r^2(\omega) = \alpha^2\). This limit implies that the extended SOR diverges on the whole interval if \(\alpha \geq 1\). However, if \(\alpha < 1\) the extended SOR will converge on \((\frac{2}{1 - \alpha}, \infty)\), where \(r^2(\frac{2}{1 - \alpha}) = 1\).

For sub case (c) we have \(D(0) \geq 0 \geq D(\alpha)\) from which together with (4.3) we obtain \(\omega \in (-\infty, \frac{2}{1-(1+\alpha^2)^{1/2}}]\). As in the corresponding subcase (c) of Theorem 4.1 we have that for \(\mu \in [0, \alpha]\) there exists a unique value \(\tilde{\mu}(:= -2(1-\omega)^{1/2})\) such that \(D(\tilde{\mu}) = 0\). If \(\mu \in [0, \tilde{\mu}]\) then for any fixed \(\omega\) in the previously found interval it will be \(D(\mu) \geq 0\). So (4.15) will have two complex zeros with equal moduli and therefore

\[ \partial(\omega) = (1 - \omega)^{1/2}. \]

If, on the other hand, \(\mu \in [\tilde{\mu}, \alpha]\) then \(D(\mu) \leq 0\) and the largest of the moduli of the two zeros of (4.15) will be

\[ |\lambda| = \frac{1}{2}(-\omega \mu + (-D(\mu))^{1/2}). \] \hfill (6.33)

Consequently

\[ \partial|\lambda|/\partial \mu \sim -\omega(-D(\mu)) + \omega^2 \mu > 0, \]
implying that \( \vartheta(\omega) \) in \([\bar{\mu}, \alpha]\) and, in view of the previous result for \( \mu \in [0, \bar{\mu}] \), in the whole interval \([0, \alpha]\), will be given by the expression in (6.33) with \( \mu = \alpha \). So

\[
\vartheta(\omega) = \frac{1}{2}(-\omega \mu + (-D(\alpha))^{1/2})
\]

and therefore

\[
r^2(\omega) = \frac{(-\omega \alpha + (-D(\alpha))^{1/2})^2}{4(1 - \omega)^2} < 1.
\]  

(6.34)

The inequality \( r^2(\omega) < 1 \) yields the equivalent one

\[
(-D(\alpha))^{1/2} < 2 - (2 - \alpha)\omega.
\]  

(6.35)

It is readily checked that this may hold for all \( \alpha \leq 2 \), since \( \omega < 0 \), or for all \( \alpha > 2 \) provided \( \omega \in \left(\frac{2}{2-\alpha}, \frac{2}{1-(1+\alpha^2)^{1/2}}\right) \). Under either of the last two conditions for (6.35) to hold we obtain by squaring \( 2 > (1 - \alpha)\omega \). So, finally, the inequality \( r^2(\omega) < 1 \) holds for

\[
1 < \alpha < \infty \quad \text{and} \quad \omega \in \left(\frac{2}{1-\alpha}, \frac{2}{1-(1+\alpha^2)^{1/2}}\right)
\]  

(6.36)

or

\[
0 < \alpha \leq 1 \quad \text{and} \quad \omega \in \left(-\infty, \frac{2}{1-(1+\alpha^2)^{1/2}}\right].
\]  

(6.37)

For the behavior of \( r^2(\omega) \) in the interval \((-\infty, \frac{2}{1-(1+\alpha^2)^{1/2}}]\) we differentiate the expression in (6.34) to obtain after some algebra

\[
\frac{\partial r^2(\omega)}{\partial \omega} = -\alpha(-D(\alpha))^{1/2} + \omega(2 + \alpha^2) - 2 < 0.
\]

This implies that \( r^2(\omega) \) is strictly decreasing in the whole interval. Moreover \( \lim_{\omega \to -\infty} r^2(\omega) = \alpha^2 \).

The convergence and the optimal convergence results as well as the monotonicity behavior of the convergence factor for the nonpositive case are presented in formulas (4.19)–(4.20) and Tables 3i–3iii. This concludes the proof of Theorem 4.2. \( \square \)
Table 1: Cyclicity \( r \) of the Best Repartitioning.

<table>
<thead>
<tr>
<th>Case</th>
<th>Value or domain of the ration ( \frac{\alpha}{\beta} )</th>
<th>Values of ( \ell ) and ( \alpha_{\ell,\ell+1} ) (or ( \beta_{\ell,\ell+1} )) if further subcases have to be considered</th>
<th>Further Subcases</th>
<th>Cyclicity ( r ) of the best repartitioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0</td>
<td>Determine the largest integer ( \ell \in {2, 3, \ldots, p-1} : (\frac{\ell-2}{\ell})^{\ell} \leq (\frac{3}{\alpha})^{p} ) and then ( \alpha_{\ell,\ell+1} ) from (3.1),(3.2)</td>
<td>A) ( \ell = k )</td>
<td>( \ell )</td>
</tr>
<tr>
<td></td>
<td>i) ( (0 = \alpha &lt; \beta &lt; 1) )</td>
<td></td>
<td></td>
<td>( \ell \leq 1 )</td>
</tr>
<tr>
<td></td>
<td>ii) ( (0, \frac{p-2}{p}) )</td>
<td>Determine the largest integers ( k \in {2, 3, \ldots, p} : (\frac{k-2}{k})^{k} \leq \frac{1}{\alpha^{p}} ), ( \ell \in {2, 3, \ldots, \min(p-1, k)} : (\frac{\ell-2}{\ell})^{\ell} \leq (\frac{3}{\alpha})^{p} ) and, if ( \ell &lt; k ), then ( \beta_{\ell,\ell+1} ) from (3.1),(3.3)</td>
<td>A) ( (\frac{\ell-2}{\ell})^{\ell/\beta} \leq \alpha &lt; \alpha_{\ell,\ell+1} )</td>
<td>( \ell )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B) ( \ell &lt; k )</td>
<td>( \ell \leq 1 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \ell \leq 1 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \ell \leq 1 )</td>
</tr>
<tr>
<td></td>
<td>III ( (0 \leq \alpha = \beta &lt; 1) )</td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
<tr>
<td></td>
<td>IV ( (1, \frac{p-2}{p-2}] )</td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
<tr>
<td></td>
<td>i) ( (\frac{p}{p-2}, \infty) )</td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
<tr>
<td></td>
<td>ii) ( (\alpha &gt; 0, \beta = 0) )</td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \ell, \ell+1^* )</td>
</tr>
</tbody>
</table>

* Either will do.
** Any will do. The optimal SOR is the Gauss-Seidel method.