The 3-Block KSOR Method for Full Rank Rectangular Systems

I. K. Youssef, Salwa M. Ali, M. A. Naser

Math Department, Faculty of Science, Ain Shams University, Cairo, Abbassia, Egypt

Email address:

Kaoud22@hotmail.com (I. K. Youssef), moathnaser87@hotmail.com (M. A. Naser)

Abstract: A new version of the KSOR method is considered to introduce the least squares solution of a full rank over determinant system of linear algebraic equations. The treatment depends on introducing an augmented non-singular square system through splitting the coefficient matrix A into two matrices A_1, A_2 with non-singular part, A_1. Accordingly, a new version of the 3-block SOR method is introduced, the 3-block KSOR method. Selection of the relaxation parameter which guarantees the convergence in the sense of reducing the spectral radius of the iteration matrix is considered. Application of the theoretical results to a numerical example has confirmed the expected behaviour of the 3-block KSOR method.

Keywords: SOR, KSOR, 3 Block SOR Method, Rectangular Systems

1. Introduction

Large linear systems of algebraic equations has appeared in many applications of science and engineering problems. The selection of the method of solution depends on the structure of the linear algebraic system under consideration. Iterative techniques especially the successive over-relaxation (SOR) method is highly recommended in the treatment of large sparse systems. The efficient use of the SOR methods depends on the selection of the relaxation parameters. The theory of SOR method for square nonsingular systems is well established, [1]. In 2012, Youssef introduced the KSOR method as a new version of the SOR method, [2] and in 2013 Youssef and Taha introduced the MKSOR versions of the MSOR method, [3]. We adapt the KSOR to treat full rank rectangular system.

Consider the linear system

$$\sum_{j=1}^{n} a_{ij}\ x_j = b_i, i = 1,2,\dots,n \qquad (1)$$

Or

$$Ax = b, \qquad (2)$$

where $A \in R^{n \times n}$, is nonsingular matrix and has no zero diagonal elements. The following splitting of the coefficient matrix A, is used in iterative treatments,

$$A = D - L - U \qquad (3)$$

where D, is the diagonal part, $-L$ is the strictly lower triangular part and $-U$ is the strictly upper triangular part of the matrix A.

The main idea in the KSOR is the assumption of using the current component in the evaluation process in addition to the use of most recent calculated components used in the SOR methods.

Let $x^{(s)}$ the s$^{\text{th}}$ approximation of solution of (2) then the KSOR, [2, 3] method is given by

$$x_i^{(s+1)} = \frac{1}{(1+\omega^*)}\left\{x_i^{(s)} + \frac{\omega^*}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1}a_{ij}x_j^{(s+1)} - \sum_{j=i+1}^{n}a_{ij}x_j^{(s)}\right)\right\}$$

$$i = 1,2,\dots,n, s = 0,1,2,\dots,\omega^* \in R-[-2,0]$$

Or

$$x_i^{(s+1)} = \frac{1}{(1+\omega^*)}x_i^{(s)} + \frac{\omega^*}{(1+\omega^*)}x_{GSi}^{(s+1)}, \qquad (4)$$

$$i = 1,2,\cdots,n, s = 0,1,2,\cdots,\omega^* \in R-[-2,0]$$

$x_{GSi}^{(s+1)}$ is the Gauss –Seidel solution.

The matrix formulation of KSOR method is

$$x^{(s+1)} = T_{KSOR}x^{(s)} + \left((1 + \omega^*)D - \omega^*L\right)^{-1}(\omega^*b), \quad (5)$$

where, T_{KSOR} is the iteration matrix

$$T_{KSOR} = \left((1 + \omega^*)D - \omega^*L\right)^{-1}(D + \omega^*U), \quad (6)$$

The main objective in this work is the treatment of the full rank rectangular system, [4, 5, 6].

$$Ax = b, \quad (7)$$

where $A \in R^{m \times n}, m > n, rank\ A = n, x \in R^n, b \in R^m$. such systems are known as over-determinant systems. In general, Over-determinant systems do not admit classical unique solutions, [7, 8]. The general treatment for such systems is to seek for solutions in a pre-described sense. For example one can write, [5, 7]

$$x = A^+b, \quad (8)$$

Where, A^+ is the well-known Moore Penrose generalized inverse.

There is many interesting trials to reformulate the iterative techniques to be suitable for such systems, [9, 10].

The least squares technique is also one of the interesting approaches. The least squares solution of the system (7) is a vector $y \in R^n$ such that

$$\|b - Ay\|_2 = \min_x \|b - Ax\|_2, \quad (9)$$

It is well known that for full rank systems the least squares solution coincident with the generalized inverse solution, and the least squares solution is unique. An equivalent formulation of the least squares problem (9) can be written in the form, [11, 12, 13]:

Find $y \in R^n$ and $r \in R^m$ such that

$$r = b - Ay, A^T r = 0, \quad (10)$$

where $A \in R^{m \times n}, m > n, rank\ A = n, x \in R^n, b \in R^m$.

One can use the nonsingular part of the matrix A to write the system (10) in the following partitioned block form

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} y + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad (11)$$

$$[A_1^T \quad A_2^T]^T \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = 0, \quad (12)$$

where, $A_1 \in R^{n \times n}$, $rank\ A_1 = n$, $A_2 \in R^{(m-n) \times n}$ also b and r are partitioned conformably with A.

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad (13)$$

2. The Three Block Coefficient Matrix

The block formulation of the system (11-13) is

$$\begin{array}{l} A_1 y + r_1 = b_1 \\ A_2 y + r_2 = b_2 \\ A_1^T r_1 + A_2^T r = 0 \end{array}, \quad (14)$$

This formulation can be treated with different attitudes. We focus on the three block formulation. The system in (14) is known as the augmented system, it can be rearranged in the form

$$\hat{A}\,\hat{x} = \hat{b}, \quad (15)$$

Or, in full as

$$\begin{bmatrix} A_1 & 0 & I \\ A_2 & I & 0 \\ 0 & A_2^T & A_1^T \end{bmatrix} \begin{bmatrix} y \\ r_2 \\ r_1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}, \quad (16)$$

Where,

$$\hat{A} = \begin{bmatrix} A_1 & 0 & I \\ A_2 & I & 0 \\ 0 & A_2^T & A_1^T \end{bmatrix}, \hat{x} = \begin{bmatrix} y \\ r_2 \\ r_1 \end{bmatrix}, \text{and } \hat{b} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}, \quad (17)$$

The coefficient matrix $\hat{A}$ is square non-singular matrix. Accordingly the system (16) possesses a unique solution and suitable for the use of iterative techniques. Solving $(m + n) \times (m + n)$ system (16) gives the least squares solution y of the original system as understood from the following theorem

Theorem 1, [4, 6]: a vector $y \in R^n$ is a least squares solution for (7) if and only if the residual vector $r = b - Ax$ satisfies the orthogonality condition, $A^T r = 0$.

3. The Three Block KSOR Method

The three block SOR method was introduced by Chen in 1975, [5]. Plemmons [11], Niethammer, de Pillis and Varga [6], later developed this method and specify SOR convergence properties. The 3-block KSOR iterative method can be introduced in the same approach illustrated in (5) and (6) as follows:

Consider the three block matrix $\hat{A}$ given in (17) and consider the splitting

$$\hat{A} = D_3 - L_3 - U_3, \quad (18)$$

where

$$D_3 = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & A_1^T \end{bmatrix},$$

$$L_3 = \begin{bmatrix} 0 & 0 & 0 \\ -A_2 & 0 & 0 \\ 0 & -A_2^T & 0 \end{bmatrix}, U_3 = \begin{bmatrix} 0 & 0 & -I \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The three block KSOR method for system (7) can be defined as

$$z^{(s+1)} = T_3 z^{(s)} + C, \quad (19)$$

$$\begin{bmatrix} y^{(s+1)} \\ r_2^{(s+1)} \\ r_1^{(s+1)} \end{bmatrix} = T_3 \begin{bmatrix} y^{(s)} \\ r_2^{(s)} \\ r_1^{(s)} \end{bmatrix} + C, \quad (20)$$

Where

$$T_3 = ((1 + \omega^*)D_3 - \omega^* L_3)^{-1}(D_3 + \omega^* U_3), \quad (21)$$

is the iteration matrix of three block the KSOR method.

$$T_3 = \begin{bmatrix} (1+\omega^*)A_1 & 0 & 0 \\ \omega^* A_2 & (1+\omega^*)I & 0 \\ 0 & \omega^* A_2^T & (1+\omega^*)A_1^T \end{bmatrix}^{-1} \begin{bmatrix} A_1 & 0 & -\omega^* I \\ 0 & I & 0 \\ 0 & 0 & A_1^T \end{bmatrix},$$

$$C = \omega^*((1+\omega^*)D_3 - \omega^* L_3)^{-1}\hat{b}, \quad (22)$$

The theoretical characterization of the KSOR methods is strongly dependent on the corresponding Jacobi method. The three block Jacobi iteration matrix is given by

$$J_3 = D_3^{-1}(L_3 + U_3), \quad (23)$$

$$J_3 = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & A_1^T \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 & -I \\ -A_2 & 0 & 0 \\ 0 & -A_2^T & 0 \end{bmatrix},$$

$$J_3 = \begin{bmatrix} 0 & 0 & -A_1^{-1} \\ -A_2 & 0 & 0 \\ 0 & -(A_2 A_1^{-1})^T & 0 \end{bmatrix}. \quad (24)$$

As in the three block SOR method the spectrum of J_3 is very important in the convergence of the 3-block KSOR method.

Theorem 2: The eigenvalues of J_3, lie in the real interval $I = [-\alpha^{2/3}, 0]$, where $\alpha = \|P\|_2 = \|A_2 A_1^{-1}\|_2$

Proof: suppose that μ is an eigenvalue of J_3 then μ^3 is an eigenvalue of $(J_3)^3$

From J_3 as defined in (24) one can easily see that

$$(J_3)^3 = \begin{bmatrix} -A_1^{-1}P^T P A_1 & 0 & 0 \\ 0 & -PP^T & 0 \\ 0 & 0 & -P^T P \end{bmatrix}, \quad (25)$$

Where $P = A_2 A_1^{-1}$

So that $(J_3)^3$ is similar to a block diagonal matrix with blocks $-PP^T$, which is real symmetric negative semidefinite matrix. Therefore

$$\rho(-PP^T) \le \mu^3 \le 0$$

$$-\|P\|_2^2 \le \mu^3 \le 0$$

$$-\alpha^2 \le \mu^3 \le 0$$

So the eigenvalues of J_3 lie in the real interval $I = [-\alpha^{2/3}, 0]$.

It is clear from this theorem that $\rho(J_3) < 1 \leftrightarrow \alpha < 1$.

Algorithm for three block KSOR method
- Set $y^{(0)}$.
- Compute $r_1^{(0)}$ and $r_2^{(0)}$.
- Compute the iterative parameter ω^*.
- Iterate for $s = 0, 1, 2, \cdots$. until convergence.

$$y^{(s+1)} = (1/(1+\omega^*))y^{(s)} + (\omega^*/(1+\omega^*))A_1^{-1}\left(b_1 - r_1^{(s)}\right),$$

$$r_2^{(s+1)} = (1/(1+\omega^*))r_2^{(s)} + (\omega^*/(1+\omega^*))(b_2 - A_2 y^{(s+1)}),$$

$$r_1^{(s+1)} = (1/(1+\omega^*))r_1^{(s)} - (\omega^*/(1+\omega^*))(A_1^T)^{-1}A_2^T r_2^{(s+1)}.$$

Theorem 3: consider the three block KSOR method of (19) for system of (7) and consider the parameter $\alpha = \|A_2 A_1^{-1}\|_2$. Then the three block KSOR method converges for parameter ω^* in some interval if and only if

$$\alpha < 3^{3/2} \approx 5.1961,$$

in particular, if

$$\alpha < 1,$$

then (19) converges for all ω^* in the interval

$$]-\infty, 2/(\alpha^{2/3} - 1)[\cup]0. \infty[, \quad (26)$$

and if $\alpha = 1$ Then (19) converges for all ω^* in the interval

$$]0, \infty[, \quad (27)$$

and if $1 < \alpha < 2^{3/2}$, then (19) converges for all ω^* in the interval

$$]0, 2/(\alpha^{2/3} - 1)[, \quad (28)$$

and if $2^{3/2} \le \alpha < 3^{3/2}$, then (19) converges for all ω^* in the interval

$$]\alpha^{2/3} - 2, 2/(\alpha^{2/3} - 1)[, \quad (29)$$

and it is diverging for other values of ω^*.

When $\alpha < 3^{3/2}$, the optimum relaxation parameter is given by

$$\omega_b^* = \frac{3\left[\left(1 + \sqrt{\frac{1+\alpha^2}{\alpha^2}}\right)^{1/3} + \left(1 - \sqrt{\frac{1+\alpha^2}{\alpha^2}}\right)^{1/3}\right]}{2\alpha^{2/3} - 3\left[\left(1 + \sqrt{\frac{1+\alpha^2}{\alpha^2}}\right)^{1/3} + \left(1 - \sqrt{\frac{1+\alpha^2}{\alpha^2}}\right)^{1/3}\right]},$$

the proof is similar to the proof which is introduced by Niethammer, de Pills and Varga [6].

4. Numerical Calculation

In order to illustrate the theoretical discussions, we introduce the following full rank over-determinant system

$$\begin{bmatrix} 1 & -2 & 3 & 1 \\ 2 & 1 & 0 & -3 \\ 5 & 1 & -2 & 4 \\ 3 & 3 & 2 & 1 \\ 3 & 1 & 3 & -2 \\ 7 & 2 & -2 & 1 \\ 4 & 1 & 5 & 2 \\ 8 & 4 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 8 \\ 9 \\ 5 \\ 8 \\ 12 \\ 17 \end{bmatrix},$$

with $rank A = 4 = n$. The unique least squares solution of this system is the vector

$$y = [1\ 1\ 1\ 1]^T,$$

Define the splitting of A as

$$A_1 = \begin{bmatrix} 1 & -2 & 3 & 1 \\ 2 & 1 & 0 & -3 \\ 5 & 1 & -2 & 4 \\ 3 & 3 & 2 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 3 & 1 & 3 & -2 \\ 7 & 2 & -2 & 1 \\ 4 & 1 & 5 & 2 \\ 8 & 4 & 0 & 5 \end{bmatrix},$$

and the splitting of b as

$$b_1 = \begin{bmatrix} 3 \\ 0 \\ 8 \\ 9 \end{bmatrix}, b_2 = \begin{bmatrix} 5 \\ 8 \\ 12 \\ 17 \end{bmatrix}.$$

The augmented linear system of order (12×12) is

$$\begin{bmatrix} 1 & -2 & 3 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 5 & 1 & -2 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 3 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 3 & 1 & 3 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 2 & -2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 1 & 5 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 8 & 4 & 0 & 5 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 7 & 4 & 8 & 1 & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 & 4 & -2 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 3 & -2 & 5 & 0 & 3 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 & -2 & 1 & 2 & 5 & 1 & -3 & 4 & 1 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \\ y \\ \\ \end{bmatrix} \\ \begin{bmatrix} \\ r_2 \\ \end{bmatrix} \\ \begin{bmatrix} \\ r_1 \\ \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 8 \\ 9 \\ 5 \\ 8 \\ 12 \\ 17 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

It is clear that the coefficient matrix is nonsingular, and the augmented system has a unique solution and $\alpha = \|A_2 A_1^{-1}\|_2 = 1.9810$.

The 3-block SOR method converges for all ω in the interval $\omega \in]0, 2/(1 + \alpha^{2/3})[=]0, 0.7760[$, and ω_b can be calculated to be, $\omega_b = 0.752, 0$.

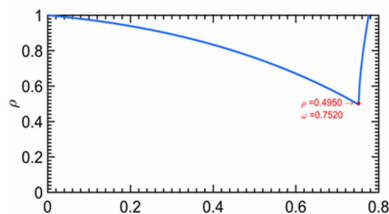Figure 1 illustrates the behavior of the spectral radius of the iteration matrix of the SOR method.

Figure 1. *The behavior of the spectral radius of 3-block SOR.*

The 3-block KSOR method is convergent for all ω^* in the interval $\omega^* \in]0, 2/(\alpha^{2/3} - 1)[=]0, 3.4644[$,

by computing the ω_b^* by theorem 3, we get $\omega_b^* = 3.0350$,

Equivalently, it can be approximated from Figure 2, which shows the behavior of the spectral radius of the iteration matrix of the 3 block KSOR method.
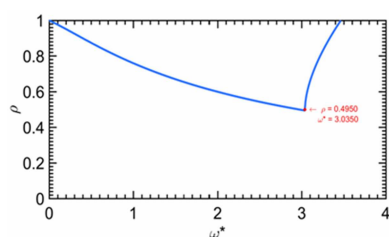

Figure 2. *The behavior of the spectral radius of 3-block KSOR.*

The table summarizes the results

Table 1. *Comparison between 3-blok SOR and KSOR.*

Method	ω_b	ρ	Number of iterations
3-block SOR	0.7520	0.4950	19
3-block KSOR	3.0350	0.4950	19

5. Conclusion

Square augmented systems obtained from full rank over determinant systems possesses unique solution, which can be obtained efficiently with iterative methods.

The structure of the three block system possesses a very nice property (large number of zeros) which illustrates the efficient use of iterative techniques. It is clear from the system (16) that for the first n equations each equation has at most (n+1) non-zero (at least (m-1) zeros), also this is the same for the next (m-n) equations and the last n equations each equation has at most m non-zeros.

The three block KSOR has the same interesting properties as the three block SOR.

The optimum values of the relaxation parameter can be calculated, and the calculated values in a good agreement with the spectral behavior of the iteration matrix

The calculated numerical results in the given example has confirmed the theoretical results.

From figures 1 and 2, the calculated optimum relaxation parameters is in a good agreement with the behavior of the spectral radius of the iteration matrices.

References

[1] D. M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, 1971.

[2] I. K. Youssef, On the Successive Overrelaxation Method, Journal of Mathematics and Statistics. 8 (2): 176-184, 2012.

[3] I. K. Youssef, A. A. Taha, On the Modified Successive Overrelaxation method, Applied Mathematics and Computation. 219, 4601-4613, 2013.

[4] T. L. Markham, M. Neumann, R. J. Plemmons, Convergence of a Direct-Iterative Method for Large-Scale Least-Squares Problems, Linear Algebra and Its Applications. 69: 155-167, 1985.

[5] V. A. Miller, M. Neumann, Successive Over Relaxation Methods for Solving the Rank Deficient Least Squares Problem, Linear Algebra Appl. 88-89: 533-557, 1987.

[6] W. Niethammer, J. de pillis, R. S. Varga, Convergence of Block Iterative Methods Applied to Sparse Least-Squares Problems, Linear Algebra and Its Applications. 58: 327-341, 1984.

[7] A. Bjorck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.

[8] A. Berman, R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.

[9]   C. H. Santos, B. P. B. Silva, J. Y. Yuan, Block SOR Methods for Rank-Deficient Least-Squares Problems, Journal of Computational and Applied Mathematics. 100: 1-9, 1998.

[10] M. T. Darvishi, F. Khani, S. Hamedi-Nezhad, B. Zheng, Symmetric Block-SOR Methods for Rank-Deficient Least Squares Problems, Journal of Computational and Applied Mathematics. 215: 14-27, 2007.

[11] R. J. Plemmons, Adjustment by Least Squares in Geodesy Using Block Iterative Methods for Sparse Matrices, in Proceeding of the U. S. Army Conference on Numerical Analysis and Computers, EL Paso, Texas, 1979.

[12] A. Berman, M. Neumann, Proper Splittings of Rectangular Matrices, SIAM J. Apll. math. 31: 307-312, 1976.

[13] C. L. Lawson, R. J. Hanson, Solving Least Squares Problems, SIAM, Philadelphia, 1995.