Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 215 (2008) 14–27
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Symmetric block-SOR methods for rank-deficient least squares problems

M.T. Darvishi[a,*], F. Khani[b], S. Hamedi-Nezhad[a], B. Zheng[c]

[a]Department of Mathematics, Razi University, Kermanshah 67149, Iran
[b]Department of Mathematics, Ilam University, P.O. Box 69315516, Ilam, Iran
[c]School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, PR China

Received 25 September 2006; received in revised form 3 March 2007

Abstract

In this article, we develop symmetric block successive overrelaxation (S-block-SOR) methods for finding the solution of the rank-deficient least squares problems. We propose an S2-block-SOR and an S3-block-SOR method for solving such problems and the convergence of these two methods is studied. The comparisons between the S2-block and the S3-block methods are presented with some numerical examples.

MSC: 65F10

Keywords: Rank-deficient; Least squares problems; Symmetric block-SOR methods; *p*-Cyclic; Convergence

1. Introduction

A rank-deficient least squares problem is given by

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|_2, \tag{1}$$

where A is an $m \times n$ matrix with $m \geqslant n$ and $\operatorname{rank}(A) < n$. We regard this problem when we solve the actual problems in economics, statistics, differential equations, image and signal processing and genetics. Least squares problems have gotten more attention in application areas and also in applied mathematics. There are many works that discussed preconditioned iterative methods to solve full rank least squares problems.

The essential problem for iterative methods to solve the rank-deficient problem is the determination of rank A. The problem is simple in theory, but not in application. Björck and Yuan [2] proposed three algorithms to find linearly independent rows of the matrix A by LU factorization, Luo et al. [6] used the basic solution method (Benzi and Meyer called it the direct-projection method [1]) to find rank of A, and Silva and Yuan [9] applied the QR decomposition

* Corresponding author. Tel./fax: +98 831 4274569.
E-mail address: darvishi@razi.ac.ir (M.T. Darvishi).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2007.03.023

column-wise to find the set of linearly independent rows of A. Those methods motivate us to consider iterative methods for solving the rank-deficient problems.

Since there are many solutions for the rank-deficient problems, we are generally interested in just the minimum 2-norm solution. Miller and Neumann [7] proposed the 4-block-SOR method to solve the rank-deficient problems. They partitioned matrix A into four parts and then applied the SOR method to solve a new 4×4 block linear system. Santos et al. [8] studied the 3-block-SOR and 2-block-SOR methods to solve a new 3×3 block linear equation by preconditioning technique. Their block-SOR methods are different from Miller and Neumann's SOR method. Recently, Darvishi and Khosro-Aghdam [3] and so Zheng and Wang [14] proposed the symmetric SOR methods to find the least squares solution of minimal norm of system (1) based on the 4×4 block augmented system in [7].

In this paper, a new symmetric 2-block-SOR and a symmetric 3-block-SOR method to find the solution of (1) are presented. The convergence of these two methods is discussed and some comparisons between our method and the results obtained in Refs. [3,11,15] are given.

Throughout the paper, we always assume that matrix A with $\mathrm{rank}(A) = k < n$ has the partition

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \tag{2}$$

where $A_1 \in \mathbb{R}^{k \times n}$ is full row rank, and $A_2 \in \mathbb{R}^{(m-k) \times n}$.

2. Preparatory knowledge

We first briefly sketch the SSOR method for the consistent linear system

$$Ax = b, \tag{3}$$

where $n \times n$ matrix A is nonsingular and has nonzero diagonal elements. We consider the following splitting of A:

$$A = D - L - U, \tag{4}$$

where D is a diagonal matrix and L and U are strictly lower and upper triangular matrices, respectively.

Let $x^{(k)}$ be kth approximation of solution (3) by SOR method using splitting (4). In symmetric SOR we obtain $x^{(k+1/2)}$ as follows [5]:

$$x^{(k+1/2)} = (D - \omega L)^{-1}((1-\omega)D + \omega U)x^{(k)} + \omega(D - \omega L)^{-1} b$$

or

$$x^{(k+1/2)} = L_\omega x^{(k)} + C, \tag{5}$$

where

$$\begin{aligned} L_\omega &= (D - \omega L)^{-1}((1-\omega)D + \omega U) \\ &= I - \omega(D - \omega L)^{-1} A \end{aligned}$$

and

$$C = \omega(D - \omega L)^{-1} b.$$

And so by backward SOR we compute $x^{(k+1)}$ as follows [5]:

$$x^{(k+1)} = U_\omega x^{(k+1/2)} + \omega(D - \omega U)^{-1} b, \tag{6}$$

where

$$\begin{aligned} U_\omega &= (D - \omega U)^{-1}((1-\omega)D + \omega L) \\ &= I - \omega(D - \omega U)^{-1} A. \end{aligned}$$

We delete $x^{(k+1/2)}$ from (5) and (6) and then obtain the symmetric SOR as follows:

$$x^{(k+1)} = \mathscr{S}_\omega x^{(k)} + C, \tag{7}$$

where

$$\begin{aligned}\mathscr{S}_\omega &= U_\omega L_\omega \\ &= (D-\omega U)^{-1}((1-\omega)D+\omega L)(D-\omega L)^{-1}((1-\omega)D+\omega U), \\ &= (I-\omega(D-\omega U)^{-1}A)(I-\omega(D-\omega L)^{-1}A)\end{aligned}$$

and

$$\begin{aligned}C &= \omega(D-\omega U)^{-1}((1-\omega)D+\omega L)(D-\omega L)^{-1}b+\omega(D-\omega U)^{-1}b \\ &= \omega(2-\omega)(D-\omega U)^{-1}D(D-\omega L)^{-1}b.\end{aligned}$$

When A is partitioned into some block form with square and nonsingular block diagonal submatrices, we can present the block SSOR method in the similar manner.

For the rank-deficient least squares problem (1), we are naturally interested in getting the minimum 2-norm solution by the aforementioned block SSOR method. For this purpose, we need the following lemma which was given by Santos et al. [8].

Lemma 2.1 (*Santos et al. [8]*). *Assume that matrix A has structure* (2). *Then*

$$\mathbf{N}(A) = \mathbf{N}(A_1) \quad \text{and} \quad \mathbf{R}(A^{\mathrm{T}}) = \mathbf{R}(A_1^{\mathrm{T}}), \tag{8}$$

where $\mathbf{R}(A)$ *and* $\mathbf{N}(A)$ *are range and null of* A, *respectively*.

Since the minimum 2-norm solution x of the rank-deficient least squares problem (1) is in $\mathbf{R}(A^{\mathrm{T}})$, that is, in $\mathbf{R}(A_1^{\mathrm{T}})$ by Lemma 2.1, we can consider the transformation

$$x = A_1^{\mathrm{T}} y, \tag{9}$$

where $y \in \mathbb{R}^{k\times 1}$, to obtain the minimum 2-norm solution of problem (1). By substituting (9) into (1), we obtain the new system of the normal equations off problem (1) for rank-deficient case as follows:

$$A_1 A^{\mathrm{T}} A A_1^{\mathrm{T}} y = A_1 A^{\mathrm{T}} b. \tag{10}$$

By the structure of A in (2), we can rewrite system (1) as the following augmented system:

$$\begin{pmatrix} AA_1^{\mathrm{T}} & I \\ 0 & A_1A^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} y \\ r \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} \tag{11}$$

or

$$\begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & I \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & A_1A_2^{\mathrm{T}} & A_1A_1^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} y \\ r_2 \\ r_1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}, \tag{12}$$

where

$$r = b - Ax = b - AA_1^{\mathrm{T}} y = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

It is worth mentioning that the block coefficient matrix of Eq. (12) can be, respectively, viewed as block 3-cyclic and 2-cyclic matrices based on two kinds of splittings, see details in Sections 3 and 4. That is to say that the Jacobi matrices

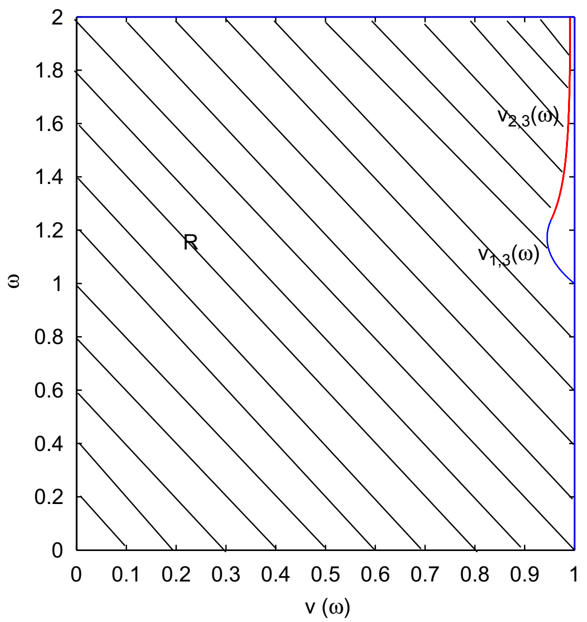

Fig. 1. Convergence region of the S3-block-SOR method .

J_3 and J_2 corresponding to these two splittings are, respectively, weakly 3-cyclic and 2-cyclic. With the condition that the block Jacobi matrix is weakly cyclic of index p, Varga et al. [12] presented the eigenvalue functional equation between the Jacobi matrix and the SSOR iteration matrix. Hence in the sequel we will discuss block SSOR method to solve system (1) mainly depending on the following theorem.

Theorem 2.2 (*Varga et al. [12]*). *Assume that the Jacobi iteration matrix B is weakly p-cyclic and so $0 < \omega < 2$. If λ is an eigenvalue of $\mathscr{S}_\omega$ for which $\lambda \neq (1-\omega)^2$, and if μ satisfies*

$$[\lambda - (1-\omega)^2]^p = \lambda[\lambda + 1 - \omega]^{p-2}(2-\omega)^2\omega^p\mu^p, \tag{13}$$

then μ is an eigenvalue of the Jacobi matrix B. Conversely, if μ is an eigenvalue of B and if λ satisfies (13) *with $\lambda \neq (1-\omega)^2$, then λ is an eigenvalue of $\mathscr{S}_\omega$.*

3. The symmetric 3-block-SOR method

In this part, we obtain the symmetric 3-block (S3-block)-SOR method.

3.1. S3-block-SOR algorithm

We set the coefficient matrix in system (12) as $\tilde{A}$: that is,

$$\tilde{A} = \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & I \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & A_1A_2^{\mathrm{T}} & A_1A_1^{\mathrm{T}} \end{pmatrix}. \tag{14}$$

Consider the following splitting of A:

$$\tilde{A} = D - L - U, \tag{15}$$

where $D=\mathrm{diag}(A_1A_1^{\mathrm{T}}, I, A_1A_1^{\mathrm{T}})$ and

$$L=\begin{pmatrix} 0 & 0 & 0\\ -A_2A_1^{\mathrm{T}} & 0 & 0\\ 0 & -A_1A_2^{\mathrm{T}} & 0\end{pmatrix},\quad U=\begin{pmatrix} 0 & 0 & -I\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

In this case, it is easy to know the Jacobi matrix $D^{-1}(L+U)$ is a weakly 3-cyclic matrix relative to the above splitting of $\tilde{A}$.

We can obtain the symmetric 3-block-SOR method as follows:

$$\begin{pmatrix} y^{(k+1)}\\ r_2^{(k+1)}\\ r_1^{(k+1)}\end{pmatrix}=\mathscr{S}_\omega^{(3)}\begin{pmatrix} y^{(k)}\\ r_2^{(k)}\\ r_1^{(k)}\end{pmatrix}+\mathscr{C}, \tag{16}$$

where

$$\mathscr{S}_\omega^{(3)}=\begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & \omega I\\ 0 & I & 0\\ 0 & 0 & A_1A_1^{\mathrm{T}}\end{pmatrix}^{-1}\begin{pmatrix} (1-\omega)A_1A_1^{\mathrm{T}} & 0 & 0\\ -\omega A_2A_1^{\mathrm{T}} & (1-\omega)I & 0\\ 0 & -\omega A_1A_2^{\mathrm{T}} & (1-\omega)A_1A_1^{\mathrm{T}}\end{pmatrix}$$
$$\times\begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0\\ \omega A_2A_1^{\mathrm{T}} & I & 0\\ 0 & \omega A_1A_2^{\mathrm{T}} & A_1A_1^{\mathrm{T}}\end{pmatrix}^{-1}\begin{pmatrix} (1-\omega)A_1A_1^{\mathrm{T}} & 0 & -\omega I\\ 0 & (1-\omega)I & 0\\ 0 & 0 & (1-\omega)A_1A_1^{\mathrm{T}}\end{pmatrix}$$

and

$$\mathscr{C}=\omega(2-\omega)\begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & \omega I\\ 0 & I & 0\\ 0 & 0 & A_1A_1^{\mathrm{T}}\end{pmatrix}^{-1}\begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0\\ 0 & I & 0\\ 0 & 0 & A_1A_1^{\mathrm{T}}\end{pmatrix}$$
$$\times\begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0\\ \omega A_2A_1^{\mathrm{T}} & I & 0\\ 0 & \omega A_1A_2^{\mathrm{T}} & A_1A_1^{\mathrm{T}}\end{pmatrix}^{-1}\begin{pmatrix} b_1\\ b_2\\ 0\end{pmatrix}.$$

S3-block-SOR algorithm.

1. Given initial vector $y^{(0)}\in\mathbb{R}^{k\times 1}$,
2. Compute $r_1^{(0)}$ and $r_2^{(0)}$,
3. Iterate for $k=1,2,3,\ldots$ until "Convergence",

$$\begin{aligned} r_2^{(k+1)}&=\omega(\omega-1)(2-\omega)Py^{(k)}+\omega^2(2-\omega)PQ^{-1}(r_1^{(k)}-b_1)+(\omega-1)^2r_2^{(k)}\\ &\quad+\omega(2-\omega)b_2,\end{aligned}$$

$$r_1^{(k+1)}=\omega Q^{-1}T((\omega-1)r_2^{(k)}-r_2^{(k+1)})+(\omega-1)^2r_1^{(k)},$$

$$y^{(k+1)}=\omega Q^{-1}((\omega-1)r_1^{(k)}-r_1^{(k+1)})+\omega(2-\omega)Q^{-1}b_1+(\omega-1)^2y^{(k)},$$

where $Q=A_1A_1^{\mathrm{T}}\in\mathbb{R}^{k\times k}$ is a nonsingular matrix, $P=A_2A_1^{\mathrm{T}}$ and $T=A_1A_2^{\mathrm{T}}$.

3.2. Convergence region

Now we discuss the convergence of the S3-block-SOR method for rank deficient least squares problems. In this case, the Jacobi matrix J_3 is

$$J_3 = \begin{pmatrix} 0 & 0 & -(A_1A_1^{\mathrm{T}})^{-1} \\ -A_2A_1^{\mathrm{T}} & 0 & 0 \\ 0 & -(A_1A_1^{\mathrm{T}})^{-1}A_1A_2^{\mathrm{T}} & 0 \end{pmatrix}, \tag{17}$$

which is a weakly 3-cyclic matrix.

Lemma 3.2.1 (*Santos et al. [8]*). *The eigenvalues of J_3 in* (17) *lie in the real interval*

$$I_3 := [-\alpha^{2/3}, 0],$$

where $\alpha = \|A_2A_1^{\mathrm{T}}(A_1A_1^{\mathrm{T}})^{-1}\|_2$.

Set $p = 3$ and for a fixed $\omega \in (0, 2)$ in Eq. (13), let

$$f(\lambda) = g(\lambda) - \lambda[\lambda + 1 - \omega](2 - \omega)^2\omega^3\mu^3,$$

where

$$g(\lambda) = [\lambda - (1 - \omega)^2]^3.$$

Set $\Omega := \{\lambda \in C : |\lambda| \leqslant 1\}$, then $g(\lambda)$ has all its roots in the interior of Ω. As $\lambda + 1 - \omega \neq 0$, if

$$(2 - \omega)^2\omega^3\nu^3 < \min_{\lambda\in\partial\Omega} \frac{|\lambda - (\omega - 1)^2|^3}{|\lambda + 1 - \omega|}, \tag{18}$$

where $\nu = \rho(J_3) = \alpha^{2/3}$ and ∂ refers to the boundary of a set, then

$$\begin{aligned} |f(\lambda) - g(\lambda)| &= |\lambda + 1 - \omega|(2 - \omega)^2\omega^3\mu^3 \\ &\leqslant |\lambda + 1 - \omega|(2 - \omega)^2\omega^3\nu^3 \\ &< |\lambda - (1 - \omega)^2|^3 = |g(\lambda)| \end{aligned}$$

hold for any $\lambda \in \partial\Omega$. Therefore, it follows by Rouche's theorem (see for example [10]) that all roots of $f(\lambda)$ also locate in the interior of Ω, that is,

$$\rho(\mathscr{S}_\omega) < 1.$$

On substituting $\lambda = x + \mathrm{i}y$ into (18), where $x, y \in \mathbb{R}$ with $y^2 = 1 - x^2$, we can rewrite the inequality in (17) as follows:

$$(2 - \omega)^2\omega^3\nu^3 < \min_{x\in[-1,1]} h(x, \omega), \tag{19}$$

where

$$h(x, \omega) := \frac{[1 + (\omega - 1)^4 - 2(\omega - 1)^2x]^{3/2}}{[1 + (\omega - 1)^2 - 2(\omega - 1)x]^{1/2}}. \tag{20}$$

In the following theorem we obtain the convergence region of S3-block-SOR method (16).

Theorem 3.2.2. *Let $\tilde{A}$ be a* 3*-cyclic block matrix as in the form of* (14) *with Jacobi matrix* J_3. *Suppose that* $\rho(J_3) = v$. *Then* $\rho(\mathscr{S}_\omega) < 1$ *provided that* $(v, \omega) \in R$, *where* R *is the region in the* (v, ω)*-plane given by*

$$R = \begin{cases} 0 < \omega \leqslant 1, & 0 \leqslant v < 1, \\ 1 \leqslant \omega \leqslant \omega^*, & 0 \leqslant v < \dfrac{1 + (1-\omega)^2}{\sqrt[3]{(2-\omega)^2 w^4}} := v_{1,3}(\omega), \\ \omega^* \leqslant \omega < 2, & 0 \leqslant v < \dfrac{\sqrt{3(\omega - 1)}\,\sqrt[3]{\varphi(\omega) + 1}}{\sqrt[3]{2}\omega} := v_{2,3}(\omega), \end{cases} \tag{21}$$

where

$$\omega^* = \frac{2\sqrt{\varphi^* + 2}}{\sqrt{\varphi^* + 2} + \sqrt{\varphi^* - 2}}, \quad \varphi^* = \frac{3 + \sqrt{33}}{2}, \tag{22}$$

and

$$\varphi(\omega) = \omega - 1 + \frac{1}{\omega - 1}, \quad \omega \neq 1. \tag{23}$$

Proof. We only present a sketch of proof. The complete details of the analysis can be found in Hadjidimos and Neumann [4]. We begin by noting that for $\omega \neq 1$ the extremal points of the function $h(x, \omega)$ in (20) can be determined by observing that

$$\operatorname{sign}\left\{\frac{\partial h(x, \omega)}{\partial x}\right\} = \operatorname{sign}\{(\omega - 1)(x - \psi(\omega))\}, \tag{24}$$

where

$$\psi = \tfrac{1}{4}[-\varphi^2(\omega) + 3\varphi(\omega) + 2]. \tag{25}$$

Suppose now that $\omega \in (1, 2)$. Then $\psi(\omega) < 1$. If ω further satisfies $\omega^* \leqslant \omega < 2$, where ω^* is given in (22), then by (23), $2 < \varphi(\omega) \leqslant \varphi^*$. On regarding now $\psi(\omega)$ as a quadratic in $\varphi(\omega)$, it follows from (25) that for $\omega^* \leqslant \omega < 2$, $\psi(\omega) \geqslant -1$. For such ω we see from (24) that $h(x, \omega)$ as a function of x is strictly decreasing in the interval $[-1, \psi(\omega))$ and it is increasing in the interval $(\psi(\omega), 1)$. Thus

$$\min_{x \in [-1,1]} h(x, \omega) = h(\psi(\omega), \omega).$$

On substituting $x = \psi(\omega)$ in the expression for $h(x, \omega)$ given in (20) we obtain that the inequality in (19) holds for all nonnegative v satisfying $0 \leqslant v < v_{2,3}(\omega)$. We have thus verified that the constraint in (21) is valid. We can verify the remaining last constraints in (21) in a similar way. □

The shaded region in Fig. 1 provides a graphical illustration of the convergence region in the (v, ω)-plane which is specified by (21). Note that from region R in (21) and so from Fig. 1, we have $v < 1$.

4. The symmetric 2-block-SOR method

In this section we will investigate the S2-block-SOR method depending on another splitting of matrix $\tilde{A}$.

4.1. S2-block-SOR algorithm

Consider the following partitioning and splitting of matrix $\tilde{A}$ in (14):

$$\tilde{A} = \left(\begin{array}{cc|c} A_1 A_1^{\mathrm{T}} & 0 & I \\ A_2 A_1^{\mathrm{T}} & I & 0 \\ \hline 0 & A_1 A_2^{\mathrm{T}} & A_1 A_1^{\mathrm{T}} \end{array}\right) \tag{26}$$

and

$$\tilde{A} = D - L - U,$$

where

$$D_2 = \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0 \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & 0 & A_1A_1^{\mathrm{T}} \end{pmatrix}$$

and

$$L = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -A_1A_2^{\mathrm{T}} & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & 0 & -I \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then the corresponding Jacobi matrix J_2 is

$$J_2 = \begin{pmatrix} 0 & 0 & -(A_1A_1^{\mathrm{T}})^{-1} \\ 0 & 0 & P \\ 0 & -P^{\mathrm{T}} & 0 \end{pmatrix}, \tag{27}$$

where $P = A_2A_1^{\mathrm{T}}(A_1A_1^{\mathrm{T}})^{-1}$ and it is easy to check that J_2 is a weakly 2-cyclic matrix. So $\tilde{A}$ is a 2-cyclic matrix relative to the partitioning as shown in (26).

The S2-block-SOR method is defined as follows:

$$\begin{pmatrix} y^{(k+1)} \\ r_2^{(k+1)} \\ r_1^{(k+1)} \end{pmatrix} = \mathscr{S}_\omega^{(2)} \begin{pmatrix} y^{(k)} \\ r_2^{(k)} \\ r_1^{(k)} \end{pmatrix} + \mathscr{C}, \tag{28}$$

where

$$\begin{aligned} \mathscr{S}_\omega^{(2)} = {} & \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & \omega I \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & 0 & A_1A_1^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} (1-\omega)A_1A_1^{\mathrm{T}} & 0 & 0 \\ (1-\omega)A_2A_1^{\mathrm{T}} & (1-\omega)I & 0 \\ 0 & -\omega A_1A_2^{\mathrm{T}} & (1-\omega)A_1A_1^{\mathrm{T}} \end{pmatrix} \\ & \times \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0 \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & \omega A_1A_2^{\mathrm{T}} & A_1A_1^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} (1-\omega)A_1A_1^{\mathrm{T}} & 0 & -\omega I \\ (1-\omega)A_2A_1^{\mathrm{T}} & (1-\omega)I & 0 \\ 0 & 0 & (1-\omega)A_1A_1^{\mathrm{T}} \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} \mathscr{C} = {} & \omega(2-\omega) \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & \omega I \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & 0 & A_1A_1^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0 \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & 0 & A_1A_1^{\mathrm{T}} \end{pmatrix} \\ & \times \begin{pmatrix} A_1A_1^{\mathrm{T}} & 0 & 0 \\ A_2A_1^{\mathrm{T}} & I & 0 \\ 0 & \omega A_1A_2^{\mathrm{T}} & A_1A_1^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}. \end{aligned}$$

S2-block-SOR algorithm.

1. Given initial vector $y^{(0)} \in \mathbb{R}^{k\times 1}$,
2. Compute $r_1^{(0)}$ and $r_2^{(0)}$,

3. Iterate for $k = 1, 2, 3, \ldots$ until "Convergence",

$$\begin{aligned}
r_1^{(k+1)} &= \omega(2-\omega)P^{\mathrm{T}}((\omega-1)r_2^{(k)} - \omega b_2) + (\omega-1)^2 r_1^{(k)} \\
&\quad + \omega^2(2-\omega)P^{\mathrm{T}}P(b_1 - r_1^{(k)}), \\
r_2^{(k+1)} &= (\omega-1)^2 r_2^{(k)} + \omega(1-\omega)P r_1^{(k)} + \omega(2-\omega)b_2 \\
&\quad + \omega P(r_1^{(k+1)} + (\omega-2)b_1), \\
y^{(k+1)} &= (\omega-1)^2 y^{(k)} + \omega P((\omega-1)r_1^{(k)} + (2-\omega)b_1 - r_1^{(k+1)}),
\end{aligned}$$

where $P = (A_2 A_1^{\mathrm{T}})(A_1 A_1^{\mathrm{T}})^{-1}$ and $P \in \mathbb{R}^{(m-k)\times k}$.

4.2. Convergence region

Lemma 4.2.1. *Let μ be an eigenvalue of the Jacobi matrix J_2 in* (27). *Then the spectrum of J_2 is pure imaginary, that is*,

$$-\alpha^2 \leqslant \mu^2 \leqslant 0.$$

Proof. Since

$$J_2^2 = \begin{pmatrix} 0 & (A_1 A_1^{\mathrm{T}})^{-1}P & 0 \\ 0 & -PP^{\mathrm{T}} & 0 \\ 0 & 0 & -P^{\mathrm{T}}P \end{pmatrix} \tag{29}$$

as μ is an eigenvalue of J_2, then μ^2 is an eigenvalue of J_2^2; suppose that the corresponding eigenvector of μ^2 is $(x^{\mathrm{T}}, y^{\mathrm{T}}, z^{\mathrm{T}})^{\mathrm{T}}$. Then we have

$$J_2^2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mu^2 \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

hence, from (29) we have $-P^{\mathrm{T}}Pz = \mu^2 z$, then $-\rho(P^{\mathrm{T}}P) \leqslant \mu^2 \leqslant 0$ or $-\alpha^2 \leqslant \mu^2 \leqslant 0$. □

Lemma 4.2.2 (*Young [13]*). *Let x be any root of the real quadratic equation $x^2 - bx + c = 0$. Then $|x| < 1$ if and only if*

$$|c| < 1,$$

$$|b| < 1 + c,$$

where b and c are real.

We obtain the convergence region of the S2-block-SOR method for rank-deficient problem in the following theorem.

Theorem 4.2.3. *The S2-block-SOR method* (28) *for the rank-deficient least squares problem* (1) *converges if and only if*:

case 1: *when* $\alpha \geqslant 1$,

$$\omega \in \left(0, 1 - \sqrt{\frac{\alpha-1}{\alpha+1}}\right) \cup \left(1 + \sqrt{\frac{\alpha-1}{\alpha+1}}, 2\right);$$

case 2: *when* $\alpha < 1$, $0 < \omega < 2$, *where* $\alpha = \|A_2 A_1^{\mathrm{T}}(A_1 A_1^{\mathrm{T}})^{-1}\|_2$.

Proof. It follows immediately from (13) for $p = 2$ that

$$[\lambda - (1-\omega)^2]^2 = -\lambda(2-\omega)^2\omega^2\tilde{\mu}^2, \tag{30}$$

for $\tilde{\mu}^2 = -\mu^2$ and $\tilde{\mu}^2 \in [0, \alpha^2]$. We now rewrite (30) as

$$\lambda^2 - \lambda[2(1-\omega)^2 - (2-\omega)^2\omega^2\tilde{\mu}^2] + (1-\omega)^4 = 0.$$

By Lemma 4.2.2, $|\lambda| < 1$ if and only if

$$(1-\omega)^4 < 1 \tag{31}$$

and

$$|2(1-\omega)^2 - (2-\omega)^2\omega^2\tilde{\mu}^2| < (1-\omega)^4 + 1. \tag{32}$$

From (31), it follows that $0 < \omega < 2$, and it follows from (32) that

$$\omega^2\tilde{\mu}^2 \leqslant \omega^2\alpha^2 < \left(\frac{1+(1-\omega)^2}{2-\omega}\right)^2,$$

that is,

$$\omega^2 - 2\omega + \frac{2}{1+\alpha^2} > 0. \tag{33}$$

From (33) and $0 < \omega < 2$, we easily know $|\lambda| < 1$, that is, $\rho(\mathscr{S}_\omega^{(2)}) < 1$ if and only if case 1 or case 2 holds. □

5. Numerical examples

In this section we present some numerical examples. All runs are performed in MATLAB on an Intel Celeron 600 (256M RAM) Windows 2000 system and terminated if

$$\|X_{k+1} - X_k\|_2 < 1E - 4.$$

To compare our results with other methods we select the test problems (Examples 1, 2) from other papers. Though the sizes of equation $Ax = b$ in Examples 1 and 2 are too small, they can show that our methods are efficient and better than the other known ones. The comparison between the S2-block-SOR and S3-block-SOR is also illustrated by these examples.

Example 1. Consider the following system:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Its least squares solution of minimal norm is $[0.5\ \ 0.5\ \ 0.5]^{\mathrm{T}}$, and the coefficient matrix is of rank 2.

Let $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ be the block form (2), where $A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. By taking the initial vector $y^{(0)} = [0\ \ 0]^{\mathrm{T}}$ and using the different values of ω in the convergence interval, we obtain the numerical results by the S2-block-SOR and S3-block-SOR methods shown in Tables 1 and 2, respectively. The convergence region

Table 1
S2-block-SOR

Relaxation parameter	Iterations	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
$\omega = \frac{1}{4}$	16	0.5001	0.5001	0.5001
	17	0.5000	0.5000	0.5000
	18	0.5000	0.5000	0.5000
$\omega = \frac{1}{2}$	6	0.5002	0.5002	0.5002
	7	0.5000	0.5000	0.5000
	8	0.5000	0.5000	0.5000
$\omega = 1$	Divergent	—	—	—
$\omega = \frac{3}{4}$	29	0.5001	0.5001	0.5001
	30	0.5000	0.5000	0.5000
	31	0.5000	0.5000	0.5000
$\omega = \frac{5}{4}$	33	0.5001	0.5001	0.5001
	34	0.5000	0.5000	0.5000
	35	0.5000	0.5000	0.5000
$\omega = \frac{3}{2}$	7	0.4999	0.4999	0.4999
	8	0.5000	0.5000	0.5000
	9	0.5000	0.5000	0.5000
$\omega = \frac{7}{4}$	15	0.5002	0.5002	0.5002
	16	0.5000	0.5000	0.5000
	17	0.5000	0.5000	0.5000

Table 2
S3-block-SOR

Relaxation parameter	Iterations	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
$\omega = \frac{1}{4}$	32	0.5001	0.5001	0.5001
	33	0.5000	0.5000	0.5000
	34	0.5000	0.5000	0.5000
$\omega = \frac{1}{2}$	16	0.4999	0.4999	0.4999
	17	0.5000	0.5000	0.5000
	18	0.5000	0.5000	0.5000
$\omega = \frac{3}{4}$	10	0.5001	0.5001	0.5001
	11	0.5000	0.5000	0.5000
	12	0.5000	0.5000	0.5000
$\omega = \frac{5}{4}$	Divergent	—	—	—
$\omega = \frac{3}{2}$	14	0.4999	0.4999	0.4999
	15	0.5000	0.5000	0.5000
	16	0.5000	0.5000	0.5000
$\omega = \frac{7}{4}$	20	0.5001	0.5001	0.5001
	21	0.5000	0.5000	0.5000
	22	0.5000	0.5000	0.5000

for S2-block-SOR is $(0, 1) \cup (1, 2)$ and for S3-block-SOR we show the convergence region in Fig. 1. In Tables 1 and 2 we show the iterations with a different parameter ω under the termination condition $\|X_{k+1} - X_k\|_2 < 1E - 4$. Obviously, S2-block-SOR has less iterations than S3-block-SOR method under the same termination condition.

Table 3
Spectral radius of four methods for different values of relaxation parameter

ω	S2-block-SOR	2-block-SOR	S3-block-SOR	3-block-SOR
0.45	0.3025	0.5500	0.7015	0.8100
0.5859	1.0005	0.4141	0.6518	0.6722
0.8	1.7623	0.8319	1.0759	0.3965
0.83	1.8287	1.0092	1.2666	0.7420
0.86	1.8830	1.1326	1.4320	0.9915
1.52	0.2704	5.6162	1.2807	6.6506
1.6	0.3600	6.2625	0.7023	6.7651

Table 4
Number of iterations for S2-block-SOR and symmetric SOR methods

ω	S2-block-SOR	Symmetric SOR
0.3	22	25
ω_{opt}	12	14
0.4	15	18
0.5	11	17

Example 2. Consider the following inconsistent system:

$$\begin{cases} 2x_2 + 3x_2 - 5x_3 = 0, \\ 4x_1 + 5x_2 + 3x_3 = 12, \\ 7x_1 + 6x_2 - 9x_3 = 4, \\ 6x_1 + 8x_2 - 2x_3 = 5. \end{cases}$$

The coefficient matrix A is of rank 3. The least squares solution of minimal norm is $(2.5400 \quad -0.7267 \quad 1.0467)^{\mathrm{T}}$. For this problem we have

$$\alpha = \|A_2 A_1^{\mathrm{T}} (A_1 A_1^{\mathrm{T}})^{-1}\|_2 = 1.4142,$$

hence the convergence region of S2-block-SOR method is

$$(0, 0.5858) \cup (1.4142, 2)$$

and the convergence region of 2-block- and 3-block-SOR methods, respectively, are $0 < \omega < 0.8284$ and $0 < \omega < 0.8850$. In Table 3 we obtain the special radius of four iterative methods, namely, S2-block-SOR, 2-block-SOR, S3-block-SOR and 3-block-SOR methods, for different values of relaxation parameter. As we can see from this table if ω lies in the convergence regions of all methods, the spectral radius of S2-block-SOR and S3-block-SOR methods is less than the spectral radius of other methods. And so the spectral radius of S2-block-SOR method is less than the spectral radius of S3-block-SOR method if ω lies in the convergence region of the methods.

Darvishi and Khosro-Aghdam [3] solved this problem by symmetric SOR method; they obtained the optimal value of relaxation parameter as

$$\omega_{\text{opt}} = 0.46898994354.$$

We solve this problem by symmetric SOR method and S2-block-SOR method. The number of iterations for different values of relaxation parameter are reported in Table 4. Table 4 shows the S2-block-SOR method is better than symmetric SOR method.

Example 3. We now consider a larger rank-deficient inconsistent linear system $Ax = b$, where A is a 20×12 random matrix as follows:

$$\begin{pmatrix}
0.9501 & 0.2311 & 0.6068 & 0.4860 & 0.8913 & 0.7621 & 0.4565 & 0.0185 & 0.8214 & 0.4447 & 2.3894 & 0.2311\\
0.6154 & 0.7919 & 0.9218 & 0.7382 & 0.1763 & 0.4057 & 0.9355 & 0.9169 & 0.4103 & 0.8936 & 1.2744 & 0.7919\\
0.0579 & 0.3529 & 0.8132 & 0.0099 & 0.1389 & 0.2028 & 0.1987 & 0.6038 & 0.2722 & 0.1988 & 1.0910 & 0.3529\\
0.0153 & 0.7468 & 0.4451 & 0.9318 & 0.4660 & 0.4186 & 0.8462 & 0.5252 & 0.2026 & 0.6721 & 1.3771 & 0.7468\\
0.8381 & 0.0196 & 0.6813 & 0.3795 & 0.8318 & 0.5028 & 0.7095 & 0.4289 & 0.3046 & 0.1897 & 2.3449 & 0.0196\\
0.1934 & 0.6822 & 0.3028 & 0.5417 & 0.1509 & 0.6979 & 0.3784 & 0.8600 & 0.8537 & 0.5936 & 0.6046 & 0.6822\\
0.4966 & 0.8998 & 0.8216 & 0.6449 & 0.8180 & 0.6602 & 0.3420 & 0.2897 & 0.3412 & 0.5341 & 2.4576 & 0.8998\\
0.7271 & 0.3093 & 0.8385 & 0.5681 & 0.3704 & 0.7027 & 0.5466 & 0.4449 & 0.6946 & 0.6213 & 1.5793 & 0.3093\\
0.7948 & 0.9568 & 0.5226 & 0.8801 & 0.1730 & 0.9797 & 0.2714 & 0.2523 & 0.8757 & 0.7373 & 0.8686 & 0.9568\\
0.1365 & 0.0118 & 0.8939 & 0.1991 & 0.2987 & 0.6614 & 0.2844 & 0.4692 & 0.0648 & 0.9883 & 1.4913 & 0.0118\\
0.5828 & 0.4235 & 0.5155 & 0.3340 & 0.4329 & 0.2259 & 0.5798 & 0.7604 & 0.5298 & 0.6405 & 1.3813 & 0.4235\\
0.2091 & 0.3798 & 0.7833 & 0.6808 & 0.4611 & 0.5678 & 0.7942 & 0.0592 & 0.6029 & 0.0503 & 1.7055 & 0.3798\\
0.4154 & 0.3050 & 0.8744 & 0.0150 & 0.7680 & 0.9708 & 0.9901 & 0.7889 & 0.4387 & 0.4983 & 2.4104 & 0.3050\\
0.2140 & 0.6435 & 0.3200 & 0.9601 & 0.7266 & 0.4120 & 0.7446 & 0.2679 & 0.4399 & 0.9334 & 1.7732 & 0.6435\\
0.6833 & 0.2126 & 0.8392 & 0.6288 & 0.1338 & 0.2071 & 0.6072 & 0.6299 & 0.3705 & 0.5751 & 1.1068 & 0.2126\\
0.4514 & 0.0439 & 0.0272 & 0.3127 & 0.0129 & 0.3840 & 0.6831 & 0.0928 & 0.0353 & 0.6124 & 0.0530 & 0.0439\\
0.6085 & 0.0158 & 0.0164 & 0.1901 & 0.5869 & 0.0576 & 0.3676 & 0.6315 & 0.7176 & 0.6927 & 1.1902 & 0.0158\\
0.0841 & 0.4544 & 0.4418 & 0.3533 & 0.1536 & 0.6756 & 0.6999 & 0.7275 & 0.4784 & 0.5548 & 0.7490 & 0.4511\\
0.1210 & 0.4508 & 0.7159 & 0.8928 & 0.2731 & 0.2548 & 0.8656 & 0.2324 & 0.8049 & 0.9084 & 1.2621 & 0.4508\\
0.2319 & 0.2393 & 0.0498 & 0.0784 & 0.6408 & 0.1909 & 0.8439 & 0.1739 & 0.1708 & 0.9943 & 1.3314 & 0.2393
\end{pmatrix},$$

and b is the following vector:

$$b^{\mathrm{T}} = (5 \quad 9 \quad 8 \quad 7 \quad 3 \quad 1 \quad 8 \quad 8 \quad 4 \quad 8 \quad 7 \quad 5 \quad 3 \quad 5 \quad 8 \quad 1 \quad 3 \quad 9 \quad 4 \quad 7).$$

We have $\operatorname{rank}(A) = 10$. The least squares solution of minimal norm of this system is

$$x^{\mathrm{T}} = (0.5800, 1.6046, 4.9961, -1.6180, -1.8187, -2.8791, 0.2481, 0.5204,$$
$$-1.5181, 4.0515, 1.3587, 1.6046).$$

We solve this system using S2-block-SOR and 2-block-SOR methods and the following initial vector:

$$y^{(0)} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^{\mathrm{T}}.$$

The convergence region of S2-block-SOR method is

$$\omega \in (0, 0.0444) \cup (1.9556, 2)$$

and the convergence region of 2-block-SOR method is $\omega \in (0, 0.0868)$; the 3-block-SOR method diverges for all ω. Number of iterations (IT) and CPU time (in second) of S2-block- and 2-block-SOR methods are reported in Table 5 for some values of ω and it is easy to see that S2-block-SOR is better than 2-block-SOR method.

6. Conclusion

The S2-block-SOR and S3-block-SOR methods are simple and powerful techniques to find the least squares solution of minimal norm of rank-deficient linear systems. Its simplicity lies in the fact that one parameter is presented. Full exploitation of the presence of this parameter will provide us with methods which will converge faster than any other methods, e.g., SOR, SSOR and AOR. The determination of optimum value of the parameter needs further studies.

Table 5
Number of iterations and CPU time for S2-block- and 2-block-SOR methods, using initial vector $y^{(0)}$

	S2-block-SOR		2-block-SOR	
ω	IT	CPU	IT	CPU
0.0100	861	0.3910	1723	0.5310
0.0200	428	0.2040	857	0.2500
0.0300	284	0.1285	569	0.1880
0.0400	212	0.1100	424	0.1560
0.0443	191	0.0780	383	0.0945

References

[1] M. Benzi, C.D. Meyer, A direct projection method for sparse linear systems, SIAM J. Comput. 16 (1995) 1159–1176.
[2] A. Björck, J.Y. Yuan, Preconditioner for least squares problems by LU decomposition, Electron. Trans. Numer. Anal. 8 (1998) 26–35.
[3] M.T. Darvishi, R. Khosro-Aghdam, Symmetric successive overrelaxation methods for rank deficient linear systems, Appl. Math. Comput. 173 (2006) 404–420.
[4] A. Hadjidimos, M. Neumann, Precise domains of convergence for the block SSOR method associated with p-cyclic matrices, Preliminary Report, Department of Mathematics, University of Connecticut, 1986.
[5] A. Hadjidimos, A. Yeyios, The symmetric accelerated overrelaxation (SAOR) method, Math. Comput. Simulat. XXIV (1982) 72–76.
[6] Z. Luo, B.P.B. Silva, J.Y. Yuan, Direct projection methods, Internat. J. Comput. Math. 76 (1–2) (2001) 517–535.
[7] V.A. Miller, M. Neumann, Successive overrelaxation methods for solving the rank deficient least squares problem, Linear Algebra Appl. 88–89 (1987) 533–557.
[8] C.H. Santos, B.P.B. Silva, J.Y. Yuan, Block SOR methods for rank-deficient least-squares problems, J. Comput. Appl. Math. 100 (1998) 1–9.
[9] B.P.B. Silva, J.Y. Yuan, Preconditioner for least squares problems by QR decomposition, Technical Report, Department of Mathematics, UFPR, Brazil, 1997.
[10] D.O. Tall, Functions of a complex variable, Library of mathematics, Routledge and Kegan Paul, London, 1970.
[11] H. Tian, Accelerated overrelaxation methods for rank deficient linear systems, Appl. Math. Comput. 140 (2003) 485–499.
[12] R.S. Varga, W. Niethammer, D-Y. Cai, P-cyclic matrices and the symmetric successive overrelaxation method, Linear Algebra Appl. 58 (1984) 425–439.
[13] D.M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, 1971.
[14] B. Zheng, K. Wang, Symmetric successive overrelaxation method for solving the rank deficient linear least squares problem, Appl. Math. Comput. 169 (2005) 1305–1323.
[15] B. Zheng, K. Wang, On accelerate overrelaxation methods for rank deficient linear systems, Appl. Math. Comput. 173 (2) (2006) 951–959.