Gauss-Seidel Iterative Methods for Rank Deficient Least Squares Problems

Davod Khojasteh Salkuyeh and Sayyed Hasan Azizi

Abstract—We study the semiconvergence of Gauss-Seidel iterative methods for the least squares solution of minimal norm of rank deficient linear systems of equations. Necessary and sufficient conditions for the semiconvergence of the Gauss-Seidel iterative method are given. We also show that if the linear system of equations is consistent, then the proposed methods with a zero vector as an initial guess converge in one iteration. Some numerical results are given to illustrate the theoretical results.

Keywords—rank deficient least squares problems, AOR iterative method, Gauss-Seidel iterative method, semiconvergence.

I. INTRODUCTION

Onsider the problem of computing the numerical solution to the system of linear equations

$$Ax = b, (1)$$

where $A \in \mathbb{C}^{m \times n}$ with $m \ge n$ and $\operatorname{rank}(A) = k < n$ and $b \in \mathbb{C}^m$. In [5], Miller and Neumann discussed successive overrelaxation (SOR) method to solve this problem. They partitioned the matrix A into four parts and then applied the SOR method to solve the new system. In their paper they consider the semiconvergence interval of the SOR methods and the optimal relaxation parameter which minimized the modulus of the controlling eigenvalue of the SOR iteration matrix. Also, they gave a method for transforming the solutions resulting from SOR iterative methods for the augmented systems to the solutions of the original problem. In [6], Tian extended Miller and Neumann's results to accelerated overrelaxation (AOR) iterative methods. Recently, Hung and Song [4] have considered semiconvergence of the AOR iterative methods for the least squares solution of minimal norm of rank deficient linear systems. They have given necessary and sufficient conditions for semiconvergence of the AOR and Jacobi overrelaxation (JOR) iterative methods. They also derived the optimum parameters and the associated convergence factor. In addition they proposed some AOR iterative methods induced by some different splittings of the augmented coefficient matrix \hat{A} of system (1).

In this paper, we study semiconvergence of the Gauss-Seidel iterative method for the least squares solution of minimal norm of rank deficient linear systems. We first give a necessary and sufficient conditions for the semiconvergence of the Gauss-Seidel iterative methods induced by splittings presented in [4]. Then, in the case that the system (1) is consistent, we show that the Gauss-Seidel iterative methods with zero vector as an initial guess converge in one iteration.

Throughout the paper, for $A \in \mathbb{C}^{m \times n}$, A^H , R(A), $\mathrm{rank}(A)$, $\sigma(A)$ and $\rho(A)$ denotes the conjugate transpose, the range space, the rank, the spectrum and the spectral radius of A, respectively. Moreover, $\mathbb{C}_k^{m \times n} = \{A \in \mathbb{C}^{m \times n} : \mathrm{rank}(A) = k\}$ and for $x = (x_1, \dots, x_n)^T \in \mathbb{C}^n$, $\|x\|_2$ denotes the 2-norm over \mathbb{C}^n , i.e., $\|x\|_2 = (\sum_{i=1}^{i=n} |x_i|^2)^{\frac{1}{2}}$.

This paper is organized as follows. In section 2, we review the block AOR iterative method for rank deficient least squares problem. Section 3 is devoted to the block Gauss-Seidel iterative method for rank deficient least squares problem. In section 4, we present some numerical results. Some concluding results are given in section 5.

II. A BRIEF REVIEW OF THE BLOCK AOR ITERATIVE METHOD FOR RANK DEFICIENT LEAST SQUARES PROBLEM

In this section we review the block AOR iterative method for rank deficient least squares problem proposed in [6] and more investigated in [4]. Let $A \in \mathbb{C}_k^{m \times n}$. It is well known that $y \in \mathbb{C}^n$ is the least squares solution to Eq. (1), that is,

$$||b - Ay||_2 = \min_{x \in \mathbb{C}^n} ||b - Ax||_2,$$

if and only if $A^H r = 0$ where r = b - Ay. Without loss of generality suppose that A has the form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where $A_{11} \in \mathbb{C}_k^{k \times k}$ and the remaining blocks of A are appropriate linear combinations of A_{11} with $A_{22} = A_{21}A_{11}^{-1}A_{12}$ [5]. We partition the vector y and r into

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
 and $\begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$

respectively in conformity with the partitioning of A and A^H . In this case, the augmented system can be written as

$$\tilde{A}z = \tilde{b},\tag{2}$$

D. K. Salkuyeh is with the Department of Mathematics, University of Mohaghegh Ardabil, P. O. Box. 179, Ardabil, Iran, e-mail: khojaste@uma.ac.ir, salkuyeh@yahoo.com.

S. H. Azizi is with the Department of Mathematics, University of Mohaghegh Ardabili, P. O. Box. 179, Ardabil, Iran, e-mail: m.azizi17@gmail.com.

where

$$\tilde{A} = \begin{pmatrix} A_{11} & 0 & I_k & A_{12} \\ A_{21} & I_{m-k} & 0 & A_{22} \\ 0 & A_{21}^H & A_{11}^H & 0 \\ 0 & A_{22}^H & A_{12}^H & 0 \end{pmatrix},$$

$$z = \begin{pmatrix} y_1 \\ r_2 \\ r_1 \\ y_2 \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \\ 0 \end{pmatrix}.$$

In [6], Tian split the matrix \tilde{A} into

$$\tilde{A} = D - L - U,\tag{3}$$

where

$$D = \begin{pmatrix} A_{11} & 0 & 0 & 0 \\ A_{21} & I_{m-k} & 0 & 0 \\ 0 & 0 & A_{11}^H & 0 \\ 0 & 0 & 0 & I_{n-k} \end{pmatrix},$$

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -A_{21}^H & 0 & 0 \\ 0 & -A_{22}^H & -A_{12}^H & 0 \end{pmatrix},$$

$$U = \begin{pmatrix} 0 & 0 & -I_k & -A_{12} \\ 0 & 0 & 0 & -A_{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{n-k} \end{pmatrix}.$$

Then, for $\omega \neq 0$, the AOR iterative method [3] is defined by

$$z^{(i+1)} = \mathcal{L}_{\gamma,\omega} z^{(i)} + c, \tag{4}$$

where

$$\mathcal{L}_{\gamma,\omega} = (D - \gamma L)^{-1} [(1 - \omega)D + (\omega - \gamma)L + \omega U]$$

$$= \begin{pmatrix} (1 - \omega)I & 0 \\ 0 & (1 - \omega)I \\ 0 & \omega(\gamma - 1)B^{H} \\ 0 & -\omega(\gamma - 1)^{2}A_{22}^{H} \end{pmatrix}$$

$$-\omega A_{11}^{-1} & -\omega C$$

$$\omega B & 0$$

$$(1 - \omega)I - \omega \gamma B^{H}B & 0$$

$$\omega(\gamma - 1)A_{12}^{H}(\gamma B^{H}B + I) & I$$

in which $B=A_{21}A_{11}^{-1}$, $C=A_{11}^{-1}A_{12}$ and $c=\omega(D-\gamma L)^{-1}\tilde{b}$. It is well known that the AOR iterative method (4) is semiconvergent from any initial vector $z^{(0)}$ if and only if the following three conditions hold: (see for example [1]) (a) $\rho(\mathcal{L}_{\gamma,\omega})=1$;

(b) If $\lambda \in \sigma(\mathcal{L}_{\gamma,\omega})$ with $|\lambda| = 1$, then $\lambda = 1$, i.e.,

$$\vartheta(\mathcal{L}_{\gamma,\omega}) = \max\{|\lambda|, \lambda \in \sigma(\mathcal{L}_{\gamma,\omega}), \lambda \neq 1\} < 1.$$

(c) Elementary divisors associated with 1 are linear, i.e.,

$$rank(I - \mathcal{L}_{\gamma,\omega}) = rank(I - \mathcal{L}_{\gamma,\omega})^2$$

or equivalently $I - \mathcal{L}_{\gamma,\omega}) = 1$.

In [4], Huang and Song presented the following theorem concerning the semiconvergence of the iterative method (4).

Theorem 1. Let $\gamma \neq 0$. Then, the AOR iterative method (4) is semiconvergent if and only if the parameters γ and ω satisfy $\omega \in (0, \frac{2}{\sqrt{1+\bar{\mu}^2}})$ and $\gamma \in (\alpha(\bar{\mu}^2), \beta(\bar{\mu}^2))$, where $\bar{\mu} = \|B\|_2$ and

$$\alpha(z) = \frac{1}{z}(\omega - 2 + \omega z), \qquad \beta(z) = \frac{1}{\omega z}(2 - 2\omega + \frac{1}{2}\omega^2 + \frac{1}{2}\omega^2 z).$$

Moreover, if the AOR iterative method is semiconvergent, then the optimal pair of parameters γ_{opt} and ω_{opt} is determined by

$$\omega_{opt} = \gamma_{opt} = \frac{2}{1 + \sqrt{1 + \bar{\mu}^2}}.$$

It is well known that the AOR iterative method with $\omega=\gamma=1$ results in the Gauss-Seidel iterative method. In the next section we give more investigation of the Gauss-Seidel iterative method and give some results concerning its convergence.

III. BLOCK GAUSS-SEIDEL ITERATIVE METHOD FOR RANK DEFICIENT LEAST SQUARES PROBLEM

Let $\gamma=\omega=1$ and $\mathcal{G}=\mathcal{L}_{1,1}.$ Then, the Gauss-Seidel iterative method is defined by

$$z^{(i+1)} = \mathcal{G}z^{(i)} + f, (5)$$

where

$$\begin{split} \mathcal{G} &= (D-L)^{-1} U \\ &= \begin{pmatrix} 0 & 0 & -A_{11}^{-1} & -C \\ 0 & 0 & B & 0 \\ 0 & 0 & -B^H B & 0 \\ 0 & 0 & 0 & I \end{pmatrix}, \end{split}$$

is the Gauss-Seidel iteration matrix with $f = (D - L)^{-1}\tilde{b}$. Evidently, the spectrum of \mathcal{G} is

$$\sigma(\mathcal{G}) = \{0, 1\} \cup \sigma(-B^H B). \tag{6}$$

Now we state and prove the following theorem.

Theorem 2. The Gauss-Seidel iterative method is semiconvergence with any initial guess $z^{(0)}$ if and only if $||B||_2 < 1$.

Proof. Similar to Theorem 2.3 in [4], we have

$$\begin{split} I - \mathcal{G} &= \begin{pmatrix} I_k & 0 & 0 \\ 0 & I_{m-k} & 0 \\ 0 & 0 & I_k \\ 0 & 0 & 0 \end{pmatrix} \\ &\times \begin{pmatrix} I_k & 0 & A_{11}^{-1} & C \\ 0 & I_{m-k} & -B & 0 \\ 0 & 0 & I + B^H B & 0 \end{pmatrix} \\ &\equiv F.G, \end{split}$$

where

$$F = \begin{pmatrix} I_k & 0 & 0 \\ 0 & I_{m-k} & 0 \\ 0 & 0 & I_k \\ 0 & 0 & 0 \end{pmatrix},$$

$$G = \begin{pmatrix} I_k & 0 & A_{11}^{-1} & C \\ 0 & I_{m-k} & -B & 0 \\ 0 & 0 & I + B^H B & 0 \end{pmatrix}.$$

The matrix $I+B^HB$ is Hermitian positive definite. Therefore, $F\in\mathbb{C}_{m+r}^{(m+n)\times(m+k)}$ and $G\in\mathbb{C}_{m+r}^{(m+k)\times(m+n)}$. According to a theorem of Cline [2], $I-\mathcal{G})\leq 1$ if and only if $\det(GF)\neq 0$. Now, we have

$$\det(GF) = \det \begin{pmatrix} I & 0 & A_{11}^{-1} \\ 0 & I & -B \\ 0 & 0 & I + B^H B \end{pmatrix}$$
$$= \det(I + B^H B) > 0.$$

Obviously, the eigenvalues of $-B^HB$ are nonpositive. Therefore, 1 can not be an eigenvalue of $-B^HB$. Hence, the method is semiconvergent if and only if $\rho(-B^HB) < 1$. On the other hand it is easy to see that

$$\rho(-B^H B) = \|B\|_2^2.$$

This completes the proof. \Box

Theorem 3. Let $b \in R(A)$ and $z^{(0)} = 0$. Then, the vector $z^{(1)}$ computed by (5) provides the exact solution of Eq. (1).

Proof. If $z^{(0)} = 0$, then from (5) we have $z^{(1)} = f$. It is easy to see that

$$(D-L)^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 & 0 & 0 \\ -B & I & 0 & 0 \\ B^H B & -B^H & A_{11}^{-H} & 0 \\ 0 & 0 & -A_{12}^H A_{11}^{-H} & I \end{pmatrix}.$$

Therefore, we have

$$f = (D - L)^{-1}\tilde{b}$$

$$= \begin{pmatrix} A_{11}^{-1} & 0 & 0 & 0 \\ -B & I & 0 & 0 \\ B^{H}B & -B^{H} & A_{11}^{-H} & 0 \\ 0 & 0 & -A_{12}^{H}A_{11}^{-H} & I \end{pmatrix} \begin{pmatrix} b_{1} \\ b_{2} \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}^{-1}b_{1} \\ -Bb_{1} + b_{2} \\ B^{H}Bb_{1} - B^{H}b_{2} \\ 0 \end{pmatrix}.$$

We define

$$y = \left(\begin{array}{c} A_{11}^{-1}b_1 \\ 0 \end{array}\right).$$

Therefore, we obtain

$$r = b - Ay = \begin{pmatrix} 0 \\ b_2 - A_{21}A_{11}^{-1}b_1 \end{pmatrix} = \begin{pmatrix} 0 \\ b_2 - Bb_1 \end{pmatrix}.$$
 (7)

Obviously, system (1) is equivalent to

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= \begin{pmatrix} b_1 \\ b_2 - A_{21}A_{11}^{-1}b_1 \end{pmatrix}.$$

Hence, by the consistency of this system and $A_{22} = A_{21}A_{11}^{-1}A_{12}$, we have $b_2 = A_{21}A_{11}^{-1}b_1 = Bb_1$. Therefore, form (7) we have r = 0 and this completes the proof.

There is not any contradiction between theorems 1 and 3. Because, in Theorem 1 the minimization has been performed

with respect to ω and γ and the optimal parameters are independent of initial guess $z^{(0)}$.

Huang and Song in [4] proposed four other splittings of \tilde{A} as following:

$$\tilde{A} = \begin{pmatrix} A_{11} & 0 & 0 & 0 \\ 0 & I_{m-k} & 0 & 0 \\ 0 & A_{21}^H & A_{11}^H & 0 \\ 0 & 0 & 0 & I_{n-k} \end{pmatrix}$$

$$-\begin{pmatrix} 0 & 0 & 0 & 0 \\ -A_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -A_{22}^H & -A_{12}^H & 0 \end{pmatrix}$$

$$-\begin{pmatrix} 0 & 0 & -I_k & -A_{12} \\ 0 & 0 & 0 & -A_{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{n-k} \end{pmatrix}, (10)$$

It is easy to see that theorems 2 and 3 hold for all of these splittings.

IV. NUMERICAL EXPERIMENTS

In this section, we give two examples to illustrate the theoretical results presented in the previous section.

Example 1. Let

$$A = \begin{pmatrix} -8 & 1 & 1 & -1\\ 1 & -8 & 1 & 1\\ 1 & -1 & 0 & \frac{2}{9}\\ 2 & 1 & -\frac{3}{7} & \frac{1}{9} \end{pmatrix},$$

and

$$b = \bar{b} = \begin{pmatrix} 7 \\ 5 \\ 3 \\ 4 \end{pmatrix}, \qquad b = \hat{b} = \begin{pmatrix} -12 \\ 33 \\ 5 \\ 2 \end{pmatrix}.$$

Then

$$B = \begin{pmatrix} -\frac{1}{9} & \frac{1}{9} \\ -\frac{17}{63} & -\frac{10}{63} \end{pmatrix}.$$

Here $\bar{b} \notin R(A)$, but we have $\|B\|_2 = 0.3163 < 1$. According to Theorem 2 the Gauss-Seidel iterative method to solve system (2) associated with \bar{b} is semiconvergent. On the other hand, we have $\hat{b} \in R(A)$, since we have $\hat{b} = A(1, -2, 7, 9)^T$. Therefore, according to Theorem 3 we expect that the Gauss-Seidel iterative with initial guess $z^{(0)} = 0$ converges in one iteration. In TABLE I, we compare the numerical results of the Gauss-Seidel method with that of AOR iterative method with optimal parameters. We use a zero vector as an initial guess and

$$E_k = \frac{\|A^H r^{(k)})\|_2}{\|A^H r^{(0)})\|_2} < 10^{-9},$$

as the stopping criterion, where $r^{(k)}=b-Ay^{(k)}$ in which $y^{(k)}$ is the approximation of the vector y computed at iteration k. We report the results for all of five splittings mentioned in this paper. In this table "GS" and "iters" stand for the Gauss-Seidel iterative method and the number of iterations for the convergence, respectively. As we observe, if $b \in R(A)$ then the Gauss-Seidel iterative method converges in 1 iteration (small reported errors are due to round-off error propagations). We also observe, in the case that $b \notin R(A)$ and both Gauss-Seidel and AOR iterative methods are convergent, the results of the AOR method with optimal parameters are better than that of the Gauss-Seidel iterative method.

Example 2. Let

$$A_{11} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

$$A_{12} = A_{21} = \frac{2}{25} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix},$$

$$A_{22} = \frac{8}{625} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix},$$

TABLE I Numerical results for Example 1.

| | Results for $b = \bar{b}$ | | | | | | |
|------------|---------------------------|------------|------------|-------------|-------------|--|--|
| | Split. (3) | Split. (8) | Split. (9) | Split. (10) | Split. (11) | | |
| AOR: iters | 7 | 7 | 7 | 7 | 7 | | |
| E_k | 2.40e-10 | 2.61e-10 | 2.28e-10 | 3.42e-10 | 3.42e-10 | | |
| GS: iters | 10 | 10 | 10 | 10 | 10 | | |
| E_k | 3.19e-10 | 3.19e-10 | 3.19e-10 | 3.19e-10 | 3.19e-10 | | |
| | Results for $b = \hat{b}$ | | | | | | |
| | Split. (3) | Split. (8) | Split. (9) | Split. (10) | Split. (11) | | |
| AOR: iters | 6 | 6 | 6 | 6 | 6 | | |
| E_k | 1.83e-10 | 1.83e-10 | 1.83e-10 | 2.55e-10 | 2.55e-10 | | |
| GS: iters | 1 | 1 | 1 | 1 | 1 | | |
| E_k | 1.90e-16 | 1.90e-16 | 1.90e-16 | 1.90e-16 | 1.90e-16 | | |

TABLE II
NUMERICAL RESULTS FOR EXAMPLE 2.

| | Results for $b = \bar{b}$ | | | | | | | |
|------------|---------------------------|------------|------------|-------------|-------------|--|--|--|
| | Split. (3) | Split. (8) | Split. (9) | Split. (10) | Split. (11) | | | |
| AOR: iters | 9 | 9 | 9 | 9 | 9 | | | |
| E_k | 1.73e-10 | 1.73e-10 | 1.73e-10 | 4.14e-11 | 4.14e-11 | | | |
| GS: iters | 13 | 13 | 13 | 13 | 13 | | | |
| E_k | 8.98e-10 | 8.98e-10 | 8.98e-10 | 8.98e-10 | 8.98e-10 | | | |
| | Results for $b = \hat{b}$ | | | | | | | |
| | Split. (3) | Split. (8) | Split. (9) | Split. (10) | Split. (11) | | | |
| AOR: iters | 6 | 6 | 6 | 6 | 6 | | | |
| E_k | 5.00e-010 | 5.00e-010 | 5.00e-010 | 5.01e-010 | 5.01e-010 | | | |
| GS: iters | 1 | 1 | 1 | 1 | 1 | | | |
| E_k | 6.06e-16 | 6.06e-16 | 6.06e-16 | 6.06e-16 | 6.06e-16 | | | |

and

$$b = \bar{b} = \begin{pmatrix} 1\\2\\1\\-1\\4\\2 \end{pmatrix}, \qquad b = \hat{b} = A \begin{pmatrix} 1\\1\\1\\1\\1\\1 \end{pmatrix} = \frac{1}{25} \begin{pmatrix} 27\\2\\\frac{58}{25}\\\frac{66}{25}\\\frac{82}{25}\\\frac{82}{25}\\6 \end{pmatrix}.$$

Then

$$B = \frac{2}{25} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix}$$

We have $\|B\|_2=0.4545<1$. Therefore, by Theorem 2 the Gauss-Seidel iterative method to solve system (2) associated with \bar{b} and \hat{b} is semiconvergent. Obviously, we have $\hat{b} \in R(A)$ and hence according to Theorem 3 we expect that the Gauss-Seidel iterative with initial guess $z^{(0)}=0$ converges in one iteration. Numerical results are given in TABLE II. Here, we mention that all of the assumptions are as Example 1. As we see, all of the observations of the Example 1 can be posed here.

V. CONCLUSION

We have considered the block Gauss-Seidel iterative method to solve rank deficient least squares problems. Two theorems concerning the semiconvergence of the method have been presented. Numerical results confirm the theoretical results. We have compared the numerical results of the block Gauss-Seidel iterative method with that of the AOR iterative method.

Numerical results show that in general the block Gauss-Seidel iterative method can not compete with the AOR iterative method with optimal parameters. But, in the case that $b \in R(A)$ the Gauss-Seidel iterative method converges only in one iteration, whereas this is not true for the AOR iterative methods. We also presented the numerical results for all of the splittings discussed in this paper. We have observed that for small problems there is no significant difference between these splittings.

REFERENCES

- A. Berman and R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, SIAM, Philadelphia 1994.
- [2] R. E. Cline, Inverses of rank invariant powers of a matrix, SIAM J. Numer. Anal. 5 (1968) 182-197.
- [3] A. Hadjidimos, Accelerated overrelaxation method, Math. Comput. 32 (1978) 149-157.
- [4] Y. Huang and Y. Song, AOR iterative methods for rank deficient least squares problems, J. Appl. Math. Comput. 26 (2008) 105-124.
- [5] V. A. Miller and M. Neumann, Successive overrelaxation methods for solving the rank deficient linear squares problem, Linear Algebra Appl 88/89 (1987) 533-557.
- [6] H. Tian, Accelerate overrelaxation methods for rank deficient linear systems, Appl. Math. Comput. 140 (2003) 485-499.