ISSN 1346-5597

National Institute of Informatics

NII Technical Report

GMRES on singular systems revisited

Ken Hayami and Kota Sugihara

NII-2020-004E
Sept. 2020

URL: http://www.nii.ac.jp/ Tokyo 101-8430, JAPAN +81-3-4212-2000 FAX: +81-3-3556-1916

GMRES on singular systems revisited

Ken Hayami*and Kota Sugihara†

Abstract

In [Hayami K, Sugihara M. Numer Linear Algebra Appl. 2011; 18:449–469], the authors analyzed the convergence behaviour of the Generalized Minimal Residual (GMRES) method for the least squares problem $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2{}^2$, where $A \in \mathbf{R}^{n\times n}$ may be singular and $\boldsymbol{b} \in \mathbf{R}^n$, by decomposing the algorithm into the range $\mathcal{R}(A)$ and its orthogonal complement $\mathcal{R}(A)^\perp$ components. However, we found that the proof of the fact that GMRES gives a least squares solution if $\mathcal{R}(A) = \mathcal{R}(A^\mathrm{T})$ was not complete. In this paper, we will give a complete proof.

Keywords: Krylov subspace method, GMRES method, singular system, least squares problem.

1 Introduction

In Hayami, Sugihara[1], we showed in Theorem 2.6 that the Generalized Minimal Residual (GMRES) method of Saad, Schultz[2] gives a least squares solution to the least squares problem

$$\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2{}^2 \tag{1}$$

*National Institute of Informatics, and The Graduate University for Advanced Studies (SOKENDAI), 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan, e-mail: hayami@nii.ac.jp

†National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan, e-mail: sugihara@nii.ac.jp

where $A \in \mathbf{R}^{n\times n}$ may be singular, for all $\boldsymbol{b} \in \mathbf{R}^n$ and initial solution $\boldsymbol{x}_0 \in \mathbf{R}^n$ if and only if $\mathcal{R}(A) = \mathcal{R}(A^{\mathrm{T}})$, where $\mathcal{R}(A)$ is the range space of A. The theorem had been proved by Brown and Walker[3], but we gave an alternative proof by decomposing the algorithm into the $\mathcal{R}(A)$ component and $\mathcal{R}(A)^{\perp}$ component, thus giving a geometric interpretation to the range symmetry condition: $\mathcal{R}(A) = \mathcal{R}(A^{\mathrm{T}})$. However, we later realized that the proof is not so obvious as we stated. In this paper, we will give a complete proof.

We assume exact arithmetic, and the following notations will be used.

$V^{\perp}$: orthogonal complement of subspace V of $\mathbf{R}^n$.

For $X \in \mathbf{R}^{n\times n}$,

$\mathcal{R}(X)$: the range space of X, i.e., the subspace spanned by the column vectors of X,

$\mathcal{N}(X)$: the null space of X, i.e., the subspace of vectors $\boldsymbol{v} \in \mathbf{R}^n$ such that $X\boldsymbol{v} = \mathbf{0}$,

2 Convergence analysis of GMRES on singular systems

2.1 GMRES

The GMRES method of Saad, Schultz[2] applied to (1) is given as follows.

GMRES

Choose $\boldsymbol{x}_0$.
$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$
$\boldsymbol{v}_1 = \boldsymbol{r}_0/||\boldsymbol{r}_0||_2$
For $j = 1, 2, \cdots$ until satisfied do
    $h_{i,j} = (\boldsymbol{v}_i, A\boldsymbol{v}_j) \quad (i = 1, 2, \dots, j)$
    $\hat{\boldsymbol{v}}_{j+1} = A\boldsymbol{v}_j - \sum_{i=1}^{j} h_{i,j}\boldsymbol{v}_i$
    $h_{j+1,j} = ||\hat{\boldsymbol{v}}_{j+1}||_2$.   If $h_{j+1,j} = 0$, goto $*$.
    $\boldsymbol{v}_{j+1} = \hat{\boldsymbol{v}}_{j+1}/h_{j+1,j}$
End do
$*\, k := j$
Form the approximate solution

$\boldsymbol{x}_k = \boldsymbol{x}_0 + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k]\boldsymbol{y}_k$
where $\boldsymbol{y} = \boldsymbol{y}_k$ minimizes $||\boldsymbol{r}_k||_2 = ||\beta\boldsymbol{e}_1 - \overline{H}_k\boldsymbol{y}||_2$.

Here, $\overline{H}_k = [h_{i,j}] \in \mathbf{R}^{(k+1)\times k}$ is a Hessenberg matrix, i.e., $h_{i,j} = 0$ for $i > j+1$. $\beta = ||\boldsymbol{r}_0||_2$ and $\boldsymbol{e}_1 = [1, 0, \ldots, 0]^{\mathrm{T}} \in \mathbf{R}^{k+1}$. The method minimizes the residual norm $\|\boldsymbol{r}_k\|_2$, over the search space $\boldsymbol{x}_k = \boldsymbol{x}_0 + \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$, where $\mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} = \mathrm{span}\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0\}$, and $(\boldsymbol{v}_i, \boldsymbol{v}_j) = 0 \quad (i \neq j)$. Let $V_j = [v_1, \ldots, v_j]$. Then,

$$AV_j = V_{j+1}\overline{H}_j \tag{2}$$

holds.

The GMRES is said to break down when $h_{j+1,j} = 0$. Then,

$$AV_j = V_j H_j \tag{3}$$

holds, where $H_j \in \mathbf{R}^{j\times j}$ consists of the firet j rows of $\overline{H}_j$.

When A is nonsingular, the iterates of GMRES converges to the solution for all $\boldsymbol{b}, \boldsymbol{x}_0 \in \mathbf{R}^n$ within at most n steps in exact arithmetic [2].

For the general case when A may be singular, we define the following.

2.2 A geometrical framework

In this section we will begin by giving geometric interpretations to the conditions $\mathcal{N}(A) = \mathcal{N}(A^{\mathrm{T}})$ and $\mathcal{R}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}$. This is done by decomposing the space $\mathbf{R}^n$ into $\mathcal{R}(A)$ and $\mathcal{R}(A)^\perp$.

Let $\mathrm{rank}A = \dim \mathcal{R}(A) = r > 0$, and

$$\boldsymbol{q}_1, \ldots, \boldsymbol{q}_r : \text{orthonormal basis of } \mathcal{R}(A), \tag{4}$$

$$\boldsymbol{q}_{r+1}, \ldots, \boldsymbol{q}_n : \text{orthonormal basis of } \mathcal{R}(A)^\perp, \tag{5}$$

$$Q_1 := [\,\boldsymbol{q}_1, \ldots, \boldsymbol{q}_r] \in \mathbf{R}^{n\times r}, \tag{6}$$

$$Q_2 := [\,\boldsymbol{q}_{r+1}, \ldots, \boldsymbol{q}_n] \in \mathbf{R}^{n\times(n-r)}, \tag{7}$$

so that,

$$Q := [\,Q_1, Q_2] \in \mathbf{R}^{n\times n} \tag{8}$$

is an orthogonal matrix satisfying

$$Q^{\mathrm{T}}Q = QQ^{\mathrm{T}} = \mathrm{I}_n, \tag{9}$$

where I_n is the identity matrix of order n.

Orthogonal transformation of the coefficient matrix A using Q gives

$$\tilde{A} := Q^{\mathrm{T}} A Q = \begin{bmatrix} {Q_1}^{\mathrm{T}} A Q_1 & {Q_1}^{\mathrm{T}} A Q_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix}, \tag{10}$$

since ${Q_2}^{\mathrm{T}} A Q = 0$. Here, $A_{11} := {Q_1}^{\mathrm{T}} A Q_1$ and $A_{12} := {Q_1}^{\mathrm{T}} A Q_2$.

In Hayami, Sugihara[1] we derived the following properties concerning the sub-matrices A_{11} and A_{12} in (10).

Theorem 1 A_{11} : *nonsingular* $\Longleftrightarrow \mathcal{R}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}$.

Lemma 1 $A_{12} = 0 \Longrightarrow A_{11}$: *nonsingular*

Theorem 2 $A_{12} = 0 \Longleftrightarrow \mathcal{R}(A) = \mathcal{R}(A^T) \Longleftrightarrow \mathcal{N}(A) = \mathcal{N}(A^T)$.

Now we will consider decomposing iterative algorithms into the $\mathcal{R}(A)$ and $\mathcal{R}(A)^{\perp}$ components. In order to do so, we will use the transformation

$$\tilde{\boldsymbol{v}} := Q^{\mathrm{T}} \boldsymbol{v} = [Q_1, Q_2]^{\mathrm{T}} \boldsymbol{v} = \begin{bmatrix} {Q_1}^{\mathrm{T}} \boldsymbol{v} \\ {Q_2}^{\mathrm{T}} \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}^1 \\ \boldsymbol{v}^2 \end{bmatrix},$$

$$\boldsymbol{v} = Q \tilde{\boldsymbol{v}} = [Q_1, Q_2] \begin{bmatrix} \boldsymbol{v}^1 \\ \boldsymbol{v}^2 \end{bmatrix} = Q_1 \boldsymbol{v}^1 + Q_2 \boldsymbol{v}^2,$$

cf. (4)-(9), to decompose a vector variable $\boldsymbol{v}$ in the algorithm. Here, $\boldsymbol{v}^1$ corresponds to the $\mathcal{R}(A)$ component $Q_1 \boldsymbol{v}^1$ of $\boldsymbol{v}$, and $\boldsymbol{v}^2$ corresponds to the $\mathcal{R}(A)^{\perp}$ component $Q_2 \boldsymbol{v}^2$ of $\boldsymbol{v}$.

For instance, the residual vector $\boldsymbol{r} := \boldsymbol{b} - A\boldsymbol{x}$ is transformed into

$$\tilde{\boldsymbol{r}} := Q^{\mathrm{T}} \boldsymbol{r} = Q^{\mathrm{T}} \boldsymbol{b} - Q^{\mathrm{T}} A Q (Q^{\mathrm{T}} \boldsymbol{x}),$$

or

$$\begin{bmatrix} \boldsymbol{r}^1 \\ \boldsymbol{r}^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}^1 \\ \boldsymbol{b}^2 \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^1 \\ \boldsymbol{x}^2 \end{bmatrix},$$

i.e.,

$$\begin{array}{rcl} \boldsymbol{r}^1 & = & \boldsymbol{b}^1 - A_{11} \boldsymbol{x}^1 - A_{12} \boldsymbol{x}^2 \\ \boldsymbol{r}^2 & = & \boldsymbol{b}^2. \end{array} \tag{11}$$

Hence, in the least squares problem (1), we have

$$\|\boldsymbol{b} - A\boldsymbol{x}\|_2{}^2 = \|\boldsymbol{r}\|_2{}^2 = ||\tilde{\boldsymbol{r}}||_2{}^2 = \|\boldsymbol{r}^1\|_2{}^2 + \|\boldsymbol{b}^2\|_2{}^2. \tag{12}$$

Note that it is not necessary to compute Q or to decompose the algorithm into the $\mathcal{R}(A)$ and $\mathcal{R}(A)^{\perp}$ components in practice. It is only for the theoretical analysis.

2.3 Decomposition of GMRES

Based on the above geometric framework, we will analyze GMRES for the case when A is singular, by decomposing it into the $\mathcal{R}(A)$ component and the $\mathcal{R}(A)^{\perp}$ component as follows.

Decomposed GMRES (general case)

$\mathcal{R}(A)$ component	$\mathcal{R}(A)^{\perp}$ component
$\boldsymbol{b}^1 = {Q_1}^{\mathrm{T}}\boldsymbol{b}$	$\boldsymbol{b}^2 = {Q_2}^{\mathrm{T}}\boldsymbol{b}$
Choose $\boldsymbol{x}_0$	
$\boldsymbol{x}_0^1 = {Q_1}^{\mathrm{T}}\boldsymbol{x}_0$	$\boldsymbol{x}_0^2 = {Q_2}^{\mathrm{T}}\boldsymbol{x}_0$
$\boldsymbol{r}_0^1 = \boldsymbol{b}^1 - A_{11}\boldsymbol{x}_0^1 - A_{12}\boldsymbol{x}_0^2$	$\boldsymbol{r}_0^2 = \boldsymbol{b}^2$
$\|\|\boldsymbol{r}_0\|\|_2 = \sqrt{\|\|\boldsymbol{r}_0^1\|\|_2{}^2 + \|\|\boldsymbol{b}^2\|\|_2{}^2}$	
$\boldsymbol{v}_1^1 = \boldsymbol{r}_0^1/\|\|\boldsymbol{r}_0\|\|_2$	$\boldsymbol{v}_1^2 = \boldsymbol{b}^2/\|\|\boldsymbol{r}_0\|\|_2$

For $j = 1, 2, \ldots$ until satisfied do

$$h_{i,j} = (\boldsymbol{v}_i^1, A_{11}\boldsymbol{v}_j^1 + A_{12}\boldsymbol{v}_j^2) \quad (i = 1, 2, \ldots, j)$$

$$\hat{\boldsymbol{v}}_{j+1}^1 = A_{11}\boldsymbol{v}_j^1 + A_{12}\boldsymbol{v}_j^2 - \sum_{i=1}^{j} h_{i,j}\boldsymbol{v}_i^1 \qquad \hat{\boldsymbol{v}}_{j+1}^2 = -\sum_{i=1}^{j} h_{i,j}\boldsymbol{v}_i^2$$

$$h_{j+1,j} = \sqrt{||\hat{\boldsymbol{v}}_{j+1}^1||_2{}^2 + ||\hat{\boldsymbol{v}}_{j+1}^2||_2{}^2}. \quad \text{If } h_{j+1,j} = 0, \text{ goto } * .$$

$$\boldsymbol{v}_{j+1}^1 = \hat{\boldsymbol{v}}_{j+1}^1/h_{j+1,j} \qquad \boldsymbol{v}_{j+1}^2 = \hat{\boldsymbol{v}}_{j+1}^2/h_{j+1,j}$$

End do

$* \, k := j$

Form the approximate solution

$$\boldsymbol{x}_k^1 = \boldsymbol{x}_0^1 + [\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_k^1]\,\boldsymbol{y}_k \qquad \boldsymbol{x}_k^2 = \boldsymbol{x}_0^2 + [\boldsymbol{v}_1^2, \ldots, \boldsymbol{v}_k^2]\,\boldsymbol{y}_k \tag{13}$$

where $\boldsymbol{y} = \boldsymbol{y}_k$ minimizes $||\boldsymbol{r}_k||_2 = ||\beta \boldsymbol{e}_1 - \overline{H}_k \boldsymbol{y}||_2$.

From the above decomposed form of GMRES, we obtain

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_j^1 \\ V_j^2 \end{bmatrix} = \begin{bmatrix} V_{j+1}^1 \\ V_{j+1}^2 \end{bmatrix} \overline{H}_j, \tag{14}$$

which is equivalent to (2), where $\left[V_j^l\right] = \left[v_1^l, \ldots, v_j^l\right]$ $(l = 1, 2)$.

When $h_{j+1,j} = 0$, (14) becomes

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_j^1 \\ V_j^2 \end{bmatrix} = \begin{bmatrix} V_j^1 \\ V_j^2 \end{bmatrix} H_j,$$

which is equivalent to (3).

In passing, when the system is consistent, i.e. $\boldsymbol{b} \in \mathcal{R}(A)$, then $\boldsymbol{b}^2 = Q_2^{\mathrm{T}} \boldsymbol{b} = \mathbf{0}$. Hence, in the $\mathcal{R}(A)^{\perp}$ component of the above decomposed algorithm, $\boldsymbol{r}_0^2 = \boldsymbol{b}^2 = \mathbf{0}$, $\boldsymbol{v}_1^2 = \mathbf{0}$. Thus, $\hat{\boldsymbol{v}}_l^2 = \mathbf{0}$ and $\boldsymbol{v}_l^2 = \mathbf{0}$ for $l = 1, \ldots, j+1$. Hence, $V_j^2 = 0$, $V_{j+1}^2 = 0$. Thus, (14) reduces to

$$A_{11} V_j^1 = V_{j+1}^1 \overline{H}_j.$$

(See section 2.5 of Hayami, Sugihara[1].)

Returning to the general case when the system may be inconsistent, in Theorem 2 we gave a geometric interpretation: $A_{12} = 0$ to the condition: $\mathcal{N}(A) = \mathcal{N}(A^{\mathrm{T}})$. Now it is important to notice that if $A_{12} = 0$ holds, the decomposed GMRES further simplifies as follows.

Decomposed GMRES (Case $\mathcal{N}(A) = \mathcal{N}(A^{\mathbf{T}})$)

$\mathcal{R}(A)$ component | $\mathcal{R}(A)^{\perp}$ component

$\boldsymbol{b}^1 = {Q_1}^{\mathrm{T}}\boldsymbol{b}$ | $\boldsymbol{b}^2 = {Q_2}^{\mathrm{T}}\boldsymbol{b}$

Choose $\boldsymbol{x}_0$

$\boldsymbol{x}_0^1 = {Q_1}^{\mathrm{T}}\boldsymbol{x}_0$ | $\boldsymbol{x}_0^2 = {Q_2}^{\mathrm{T}}\boldsymbol{x}_0$

$\boldsymbol{r}_0^1 = \boldsymbol{b}^1 - A_{11}\boldsymbol{x}_0^1$ | $\boldsymbol{r}_0^2 = \boldsymbol{b}^2$

$||\boldsymbol{r}_0||_2 = \sqrt{||\boldsymbol{r}_0^1||_2{}^2 + ||\boldsymbol{b}^2||_2{}^2}$

$\boldsymbol{v}_1^1 = \boldsymbol{r}_0^1/||\boldsymbol{r}_0||_2$ | $\boldsymbol{v}_1^2 = \boldsymbol{b}^2/||\boldsymbol{r}_0||_2$

For $j = 1, 2, \ldots$ until satisfied do

$h_{i,j} = (\boldsymbol{v}_i^1, A_{11}\boldsymbol{v}_j^1) \quad (i = 1, 2, \ldots, j)$

$\hat{\boldsymbol{v}}_{j+1}^1 = A_{11}\boldsymbol{v}_j^1 - \sum_{i=1}^{j} h_{i,j}\boldsymbol{v}_i^1$ | $\hat{\boldsymbol{v}}_{j+1}^2 = -\sum_{i=1}^{j} h_{i,j}\boldsymbol{v}_i^2$

$h_{j+1,j} = \sqrt{||\hat{\boldsymbol{v}}_{j+1}^1||_2{}^2 + ||\hat{\boldsymbol{v}}_{j+1}^2||_2{}^2}$. If $h_{j+1,j} = 0$, goto $*$.

$\boldsymbol{v}_{j+1}^1 = \hat{\boldsymbol{v}}_{j+1}^1/h_{j+1,j}$ | $\boldsymbol{v}_{j+1}^2 = \hat{\boldsymbol{v}}_{j+1}^2/h_{j+1,j}$

End do

$* \, k := j$

Form the approximate solution

$\boldsymbol{x}_k^1 = \boldsymbol{x}_0^1 + [\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_k^1]\,\boldsymbol{y}_k$ | $\boldsymbol{x}_k^2 = \boldsymbol{x}_0^2 + [\boldsymbol{v}_1^2, \ldots, \boldsymbol{v}_k^2]\,\boldsymbol{y}_k$

where $\boldsymbol{y} = \boldsymbol{y}_k$ minimizes $||\boldsymbol{r}_k||_2 = ||\beta\boldsymbol{e}_1 - \overline{H}_k\boldsymbol{y}||_2$.

Then, (14) simplifies to

$$\begin{aligned} A_{11}V_j^1 &= V_{j+1}^1\overline{H}_j \\ 0 &= V_{j+1}^2\overline{H}_j. \end{aligned} \quad (15)$$

If further, $h_{j+1,j} = 0$, we have

$$\begin{aligned} A_{11}V_j^1 &= V_j^1 H_j \\ 0 &= V_j^2 H_j. \end{aligned} \quad (16)$$

Note here that the $\mathcal{R}(A)$ component of GMRES is "essentially equivalent" to GMRES applied to $A_{11}\boldsymbol{x}^1 = \boldsymbol{b}^1$, except for the scaling factors for $\boldsymbol{v}_j^1$. Note also that, from Lemma 1, $A_{12} = 0$ implies that A_{11} is nonsingular. From these observations, we concluded in Hayami, Sugihara[1] (Section 2.3, p. 454) that if $A_{12} = 0$, "arguments similar to Saad, Schultz[2] for GMRES on nonsingular systems imply that GMRES gives a least-squares solution for all $\boldsymbol{b}$ and $\boldsymbol{x}_0$".

However, we later found that the proof is not so obvious. The difficulty is that, although the Krylov basis $V_1 = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j]$ at step j of the GMRES is orthonormal, the corresponding $\mathcal{R}(A)$ component vecors $V_j^1 = \left[\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_j^1\right]$ are not necessarily orthogonal, and it is not even obvious that they are linearly independent. In the following, we give a complete proof of the statement. See also Sugihara, Hayami, Zheng[4], Theorem 1 for a related proof for the right-preconditioned MINRES method for symmetric singular systems.

First, we observe the following.

Lemma 2 *In the GMRES method, if* $\boldsymbol{r}_0 \neq \boldsymbol{0}$, $h_{i+1,i} \neq 0$ $(1 \leq i \leq j-1)$, *then* $\boldsymbol{v}_i^2 = c_i\boldsymbol{b}^2$ $(i = 1, \ldots, j)$, *i.e. all the* $\mathcal{R}(A)^\perp$ *components* $\boldsymbol{v}_i^2 (i = 1, \ldots, j)$ *are parallel to* $\boldsymbol{b}^2$.

Proof: From the above Decomposed GMRES(general case) (13), $\boldsymbol{v}_1^2 = \boldsymbol{b}^2/||\boldsymbol{r}_0||_2 = c_1\boldsymbol{b}^2$. Since $\hat{\boldsymbol{v}}_{j+1}^2 = -\sum_{i=1}^{j} h_{i,j}\boldsymbol{v}_i^2$ and $\boldsymbol{v}_{j+1}^2 = \hat{\boldsymbol{v}}_{j+1}^2/h_{j+1,j}$, by induction, we have $\boldsymbol{v}_i^2 = c_i\boldsymbol{b}^2$ $(i = 1, \ldots, j)$. □

Next, we prove the followi,ng.

Theorem 3 *In the GMRES method, assume* $\boldsymbol{r}_0 \neq \boldsymbol{0}$, $h_{i+1,i} \neq 0$ $(1 \leq i \leq j-1)$ *hold. If* $\boldsymbol{b} \in \mathcal{R}(A)$ $(\boldsymbol{b}^2 = 0)$, *then* rank$V_1^j = j$. *If* $\boldsymbol{b} \notin \mathcal{R}(A)$ $(\boldsymbol{b}^2 \neq \boldsymbol{0})$, *then* rank$V_j^1 = j - 1$ *or* j.

Proof: When $\boldsymbol{b} \in \mathcal{R}(A)$ ($\boldsymbol{b}^2 = 0$), from Lemma 2,

$$\tilde{V}_j = Q^T V_j = \begin{bmatrix} \boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_j^1 \\ \boldsymbol{0}, \ldots, \boldsymbol{0} \end{bmatrix}.$$

Hence, $\mathrm{rank} V_j^1 = \mathrm{rank} V_j = j$.

When $\boldsymbol{b} \notin \mathcal{R}(A)$ ($\boldsymbol{b}^2 \neq \boldsymbol{0}$), for $j = 1$, $\mathrm{rank} V_1^1 = \mathrm{rank}\,[\boldsymbol{v}_1^1] = 0$ or 1, depending on whether $\boldsymbol{v}_1^1 = \boldsymbol{0}$ or $\boldsymbol{v}_1^1 \neq \boldsymbol{0}$.

Let $j \geq 2$. From Lemma 2, and $c_1 = 1/||\boldsymbol{r}_0|| \neq 0$, we have

$$\tilde{V}_j = Q^T V_j = \begin{bmatrix} \boldsymbol{v}_1^1, & \ldots, & \boldsymbol{v}_j^1 \\ c_1 \boldsymbol{b}^2, & \ldots, & c_j \boldsymbol{b}^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_1^{1\prime}, & \boldsymbol{v}_2^{1\prime} & \ldots, & \boldsymbol{v}_j^{1\prime} \\ \boldsymbol{b}^2, & \boldsymbol{0}, & \ldots, & \boldsymbol{0} \end{bmatrix} S^{-1},$$

where

$$S = \begin{bmatrix} 1/c_1 & -c_2/c_1 & \cdots & -c_j/c_1 \\ & 1 & \cdots & 0 \\ & & \ddots & 0 \\ 0 & & & 1 \end{bmatrix} \in \mathbf{R}^{j \times j}$$

is nonsingular, and $\boldsymbol{v}_i^{1\prime} = \boldsymbol{v}_i^1 / c_1$ ($i = 1, \ldots, j$). Therefore,

$$\mathrm{rank} \begin{bmatrix} \boldsymbol{v}_1^{1\prime}, & \boldsymbol{v}_2^{1\prime} & \ldots, & \boldsymbol{v}_j^{1\prime} \\ \boldsymbol{b}_2, & \boldsymbol{0}, & \ldots, & \boldsymbol{0} \end{bmatrix} = \mathrm{rank} V_j = j.$$

Then, $\mathrm{rank}\,[\boldsymbol{v}_2{}', \ldots, \boldsymbol{v}_j{}'] = j - 1$, since if $\mathrm{rank}\,[\boldsymbol{v}_2{}', \ldots, \boldsymbol{v}_j{}'] < j - 1$, then

$$\mathrm{rank} \begin{bmatrix} \boldsymbol{v}_1^{1\prime}, & \boldsymbol{v}_2^{1\prime} & \ldots, & \boldsymbol{v}_j^{1\prime} \\ \boldsymbol{b}_2, & \boldsymbol{0}, & \ldots, & \boldsymbol{0} \end{bmatrix} < j.$$

Hence, $\mathrm{rank}\left[\boldsymbol{v}_1^1, \ldots \boldsymbol{v}_j^1\right] = \mathrm{rank}\left[\boldsymbol{v}_1^{1\prime}, \ldots, \boldsymbol{v}_j^{1\prime}\right] = j - 1$ or j. □

Note that Lemma 2 and Theorem 3 hold without assuming $A_{12} = 0$.

Next, we prove the following, which corresponds to the sufficiency of the condition in Theorem 2.6 of Hayami, Sugihara[1].

Theorem 4 *Assume $A_{12} = 0$. Then, GMRES determines a least squares solution of (1) for all $\boldsymbol{b}, \boldsymbol{x}_0 \in \mathbf{R}^n$.*

Proof: If $\boldsymbol{r}_0 = \boldsymbol{0}$, a (least squares) solution to (1) is obtained. Assume $\boldsymbol{r}_0 \neq \boldsymbol{0}$.

Assume $\boldsymbol{b} \in \mathcal{R}(A)$. Then, from Theorem 3, $\mathrm{rank} V_1^j = j$. Since $\mathrm{rank} V_1^j \leq r = \mathrm{rank} A$, there exists a $j \leq r$, such that $h_{i+1,i} \neq 0$ ($1 \leq i \leq j-1$), $h_{j+1,j} =$

0. Then from (16), we have $A_{11}V_j^1 = V_j^1 H_j$. Since A_{11} is nonsingular, $\mathrm{rank} A_{11}V_1^j = j$. Then, $j = \mathrm{rank} V_j^1 H_j \le \min(j, \mathrm{rank} H_j)$, where $\mathrm{rank} H_j \le j$. Hence, $\mathrm{rank} H_j = j$, and H_j is nonsingular. Note that

$$\begin{aligned} \boldsymbol{r}_j^1 &= \boldsymbol{b}^1 - A_{11}\boldsymbol{x}_j^1 = \boldsymbol{b}^1 - A_{11}\left(\boldsymbol{x}_0^1 + V_j^1\boldsymbol{y}_j\right) = \boldsymbol{r}_0^1 - A_{11}V_j^1\boldsymbol{y}_j \\ &= \beta\boldsymbol{v}_1^1 - V_j^1 H_j\boldsymbol{y}_j = V_j^1\left(\beta\boldsymbol{e}_1 - H_j\boldsymbol{y}_j\right), \end{aligned} \tag{17}$$

where $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbf{R}^j$. Hence, a least squares solution is obtained at step j $(j \le r)$ for $\boldsymbol{y}_j = \beta {H_j}^{-1}\boldsymbol{e}_1$, for which $\boldsymbol{r}_j^1 = \boldsymbol{0}$.

Next, assume $\boldsymbol{b} \notin \mathcal{R}(A)$. Then, in the proof of Theorem 3, $\mathrm{rank} A = r \ge \mathrm{rank} V_j^1 = j$ or $j-1$, which implies that there exists $j \le r+1$ such that $h_{i+1,i} \ne 0$ $(1 \le i \le j-1)$, $h_{j+1,j} = 0$.

(As in Point a and b in the proof of Theorem 1 in Sugihara et al.[4]), since $V_j^2 H_j = 0$ from (16), if H_j is nonsingular, $V_j^2 = [\boldsymbol{v}_1^2, \ldots, \boldsymbol{v}_j^2] = 0$. However, since $\boldsymbol{b} \notin \mathcal{R}(A)$, $\boldsymbol{b}^2 \ne \boldsymbol{0}$, so that $\boldsymbol{v}_1^2 = \boldsymbol{b}^2/\|\boldsymbol{r}_0\|_2 \ne \boldsymbol{0}$. Hence, H_j is singular, and there exists $\boldsymbol{w} \ne \boldsymbol{0}$ such that $H_j\boldsymbol{w} = \boldsymbol{0}$. (In fact, $\mathrm{rank} H_j = j-1$, since $h_{i+1,i} \ne 0$ $(1 \le i \le j-1)$.) Then, from (16), $V_j^1 H_j \boldsymbol{w} = A_{11}V_j^1\boldsymbol{w} = \boldsymbol{0}$. Since A_{11} is nosingular, $V_j^1\boldsymbol{w} = \boldsymbol{0}$, $\boldsymbol{w} \ne \boldsymbol{0}$. Hence, $\mathrm{rank} V_j^1 = j-1$. Then, a least squares solution is obtained at step j if and only if $H_j\boldsymbol{y}_j - \beta\boldsymbol{e}_1 \in \mathcal{N}(V_j^1)$. Since $\mathrm{rank} V_j^1 + \dim\mathcal{N}(V_j^1) = j$, $\dim\mathcal{N}(V_j^1) = 1$. Let $\mathcal{N}(V_1^j) = \{c\boldsymbol{\nu}^j\}$, where $c \in \mathbf{R}$, $\boldsymbol{\nu} \ne \boldsymbol{0} \in \mathbf{R}^j$. Let

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \boldsymbol{\nu}_2 \end{bmatrix} \ne \boldsymbol{0} \in \mathbf{R}^j,\ \nu_1 \in \mathbf{R},\ \boldsymbol{\nu}_2 \in \mathbf{R}^{j-1},\ \text{and } H_j = \begin{bmatrix} {\boldsymbol{h}_{11}}^{\mathrm{T}} & h_{1j} \\ H_{21} & \boldsymbol{h}_{22} \end{bmatrix},$$

where ${\boldsymbol{h}_{11}}^{\mathrm{T}} = [h_{11}, \ldots, h_{1,j-1}]$,

$$H_{21} = \begin{bmatrix} h_{21} & \cdots & h_{2,j-1} \\ & \ddots & \vdots \\ 0 & & h_{j,j-1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{h}_{22} = \begin{bmatrix} h_{2j} \\ \vdots \\ h_{jj} \end{bmatrix}.$$

where H_{21} is nonsingular since

$$h_{i+1,i} \ne 0 \quad (1 \le i \le j-1). \tag{18}$$

Note the following:

$$\begin{aligned} & \text{A least squares solution is obtained at step } j \\ \Longleftrightarrow\ & \exists\boldsymbol{y} \text{ such that } H_j\boldsymbol{y} - \beta\boldsymbol{e}_1 = c\boldsymbol{\nu} \\ \Longleftrightarrow\ & \exists\boldsymbol{y}_1, y_j \text{ such that } \begin{cases} {\boldsymbol{h}_{11}}^{\mathrm{T}}\boldsymbol{y}_1 + h_{1j}y_j = \beta + c\nu_1 \\ H_{21}\boldsymbol{y}_1 + y_j\boldsymbol{h}_{22} = c\boldsymbol{\nu}_2 \end{cases} \\ \Longleftrightarrow\ & \left(h_{1j} - {\boldsymbol{h}_{11}}^{\mathrm{T}}{H_{21}}^{-1}\boldsymbol{h}_{22}\right)y_j = \beta + c\left(\nu_1 - {\boldsymbol{h}_{11}}^{\mathrm{T}}{H_{21}}^{-1}\boldsymbol{\nu}_2\right), \end{aligned}$$

where

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ y_j \end{bmatrix} \quad \text{and} \quad \boldsymbol{y}_1 = \begin{bmatrix} y_1 \\ \vdots \\ y_{j-1} \end{bmatrix} \in \mathbf{R}^{j-1}.$$

Here note that

$$\begin{bmatrix} \mathrm{I} & \mathbf{0} \\ -\boldsymbol{h}_{11}{}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} H_{21}{}^{-1} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathrm{I} \\ 1 & \mathbf{0}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_{11}{}^{\mathrm{T}} & h_{1j} \\ H_{21} & \boldsymbol{h}_{22} \end{bmatrix} = \begin{bmatrix} \mathrm{I} & H_{21}{}^{-1}\boldsymbol{h}_{22} \\ \mathbf{0}^{\mathrm{T}} & h_{1j} - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{h}_{22} \end{bmatrix}.$$

Since $\det H_j = 0$, $\quad h_{1j} - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{h}_{22} = 0$. Thus,

$$\begin{aligned} & \text{A least squares solution is obtained at step } j \\ \Longleftrightarrow\ & \beta = c\left(\nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{\nu}_2\right) \\ \Longleftrightarrow\ & \nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{\nu}_2 \neq 0 \end{aligned}$$

since $\beta \neq 0$. Hence, if $\nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{\nu}_2 \neq 0$, a least squares solution is obtained at step j. If $\nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{\nu}_2 = 0$, a least squares solution is not obtained at step j. Note that

$$\begin{bmatrix} \mathrm{I} & \mathbf{0} \\ -\boldsymbol{h}_{11}{}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} H_{21}{}^{-1} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathrm{I} \\ 1 & \mathbf{0}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_{11}{}^{\mathrm{T}} & \nu_1 \\ H_{21} & \boldsymbol{\nu}_2 \end{bmatrix} = \begin{bmatrix} \mathrm{I} & H_{21}{}^{-1}\boldsymbol{\nu}_2 \\ \mathbf{0}^{\mathrm{T}} & \nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{\nu}_2 \end{bmatrix}.$$

Hence, if $\nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1}\boldsymbol{\nu}_2 = 0$,

$$\operatorname{rank} \begin{bmatrix} \boldsymbol{h}_{11}{}^{\mathrm{T}} & \nu_1 \\ H_{21} & \boldsymbol{\nu}_2 \end{bmatrix} = j - 1,$$

since $\operatorname{rank} H_{21} = j - 1$. Hence,

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \boldsymbol{\nu}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_{11}{}^{\mathrm{T}} \\ H_{21} \end{bmatrix} \boldsymbol{s},$$

where $\boldsymbol{s} \neq \mathbf{0}$. Then,

$$\mathbf{0} = V_j^1 \boldsymbol{\nu} = V_j^1 \begin{bmatrix} \boldsymbol{h}_{11}{}^{\mathrm{T}} \\ H_{21} \end{bmatrix} \boldsymbol{s} = V_j^1 H_j \begin{bmatrix} \mathrm{I}_{j-1} \\ \mathbf{0}^{\mathrm{T}} \end{bmatrix} \boldsymbol{s} = A_{11} V_j^1 \begin{bmatrix} \mathrm{I}_{j-1} \\ \mathbf{0}^{\mathrm{T}} \end{bmatrix} \boldsymbol{s}.$$

Since A_{11} is nonsingular,

$$V_j^1 \begin{bmatrix} \mathrm{I}_{j-1} \\ \mathbf{0}^{\mathrm{T}} \end{bmatrix} \boldsymbol{s} = \left[\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_{j-1}^1\right] \boldsymbol{s} = \mathbf{0},$$

where $\boldsymbol{s} \neq \mathbf{0}$. Hence, $\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_{j-1}^1$ are linearly dependent and $\mathrm{rank} V_{j-1}^1 = \mathrm{rank}\left[\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_{j-1}^1\right] \leq j-2$, but $\mathrm{rank} V_j^1 = \mathrm{rank}\left[\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_{j-1}^1, \boldsymbol{v}_j^1\right] = j-1$. Hence, we have $\mathrm{rank} V_{j-1}^1 = j-2$.

Next, we will use an induction argument on ℓ, where $1 \leq \ell \leq j-2$. Note

$$h_{i+1,i} \neq 0 \quad (1 \leq i \leq \ell). \tag{19}$$

Let $\mathrm{rank} V_{\ell+1}^1 = \ell$ where $V_{\ell+1}^1 \in \mathbf{R}^{r \times (\ell+1)}$. Since $\mathrm{rank} V_{\ell+1}^1 + \dim \mathcal{N}(V_{\ell+1}^1) = \ell + 1$, we have $\dim \mathcal{N}(V_{\ell+1}^1) = 1$. Hence, let $\mathcal{N}(V_{\ell+1}^1) = \{c\boldsymbol{\nu}\}$, where $c \in \mathbf{R}$, and

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \boldsymbol{\nu}_2 \end{bmatrix} \neq \mathbf{0} \in \mathbf{R}^{\ell+1},\ \nu_1 \in \mathbf{R},\ \boldsymbol{\nu}_2 \in \mathbf{R}^l.$$

Noting that, $A_{11} V_\ell^1 = V_{\ell+1}^1 \overline{H}_\ell$, similarly to (15), we have

$$\boldsymbol{r}_\ell^1 = \beta \boldsymbol{v}_1^1 - A_{11} V_\ell^1 \boldsymbol{y} = V_{\ell+1}^1 \left(\beta \boldsymbol{e}_1 - \overline{H}_\ell \boldsymbol{y}\right),$$

where $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbf{R}^{\ell+1}$.

Let

$$\overline{H}_\ell = \begin{bmatrix} {\boldsymbol{h}_{11}}^{\mathrm{T}} \\ H_{21} \end{bmatrix},$$

where ${\boldsymbol{h}_{11}}^{\mathrm{T}} = [h_{11}, \ldots, h_{1\ell}]$, and

$$H_{21} = \begin{bmatrix} h_{21} & \cdots & h_{2\ell} \\ & \ddots & \vdots \\ 0 & & h_{\ell+1,\ell} \end{bmatrix},$$

where H_{21} is nonsingular due to (19).

Then, note the following:

$$\begin{aligned}
& \text{A least squares solution is obtained at step } \ell \\
\Longleftrightarrow\ & \exists \boldsymbol{y} \text{ such that } \boldsymbol{r}_\ell^1 = V_{\ell+1}^1 \left(\beta \boldsymbol{e}_1 - \overline{H}_\ell \boldsymbol{y}\right) = \mathbf{0} \\
\Longleftrightarrow\ & \exists \boldsymbol{y} \text{ such that } \beta \boldsymbol{e}_1 - \overline{H}_\ell \boldsymbol{y} \in \mathcal{N}(V_{\ell+1}^1) \\
\Longleftrightarrow\ & \exists \boldsymbol{y} \text{ such that } \left\{ \begin{array}{rcl} \beta - {\boldsymbol{h}_{11}}^{\mathrm{T}} \boldsymbol{y} & = & c\nu_1 \\ -H_{21} \boldsymbol{y} & = & c\boldsymbol{\nu}_2 \end{array} \right. \\
\Longleftrightarrow\ & \nu_1 - {\boldsymbol{h}_{11}}^{\mathrm{T}} {H_{21}}^{-1} \boldsymbol{\nu}_2 \neq 0
\end{aligned}$$

Hence, if $\nu_1 - {\boldsymbol{h}_{11}}^{\mathrm{T}} {H_{21}}^{-1} \boldsymbol{\nu}_2 \neq 0$, a least squares solution is obtained at step ℓ.

If $\nu_1 - \boldsymbol{h}_{11}{}^{\mathrm{T}} H_{21}{}^{-1} \boldsymbol{\nu}_2 = 0$, a least squares solution is not obtained at step ℓ, and

$$\begin{vmatrix} \nu_1 & \boldsymbol{h}_{11}{}^{\mathrm{T}} \\ \boldsymbol{\nu}_2 & H_{21} \end{vmatrix} = 0.$$

Since H_{21} is nonsingular and $\boldsymbol{\nu} \neq \boldsymbol{0}$,

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \boldsymbol{\nu}_2 \end{bmatrix} = \overline{H}_\ell \boldsymbol{s},$$

where $\boldsymbol{s} \neq \boldsymbol{0} \in \mathbf{R}^\ell$. Then,

$$A_{11} V_\ell^1 \boldsymbol{s} = V_{\ell+1}^1 \overline{H}_\ell \boldsymbol{s} = V_{\ell+1}^1 \boldsymbol{\nu} = \boldsymbol{0}.$$

Since A_{11} is nonsingular, $\mathrm{rank} V_\ell^1 \leq \ell - 1$. But since $\mathrm{rank} V_{\ell+1}^1 = \ell$, $\mathrm{rank} V_\ell^1 = \ell - 1$.

Thus, by induction on ℓ, a least squares solution is obtained at step ℓ $(2 \leq \ell \leq j)$, or $\mathrm{rank} V_1^1 = \mathrm{rank}\,[\boldsymbol{v}_1^1] = 0$, so that $\boldsymbol{v}_1^1 = \boldsymbol{0}$. Then, $\boldsymbol{r}_1^1 = \beta \boldsymbol{v}_1^1 - A_{11} \boldsymbol{v}_1^1 y = \boldsymbol{0}$, so a least squares solution is obtained at step 1.

Hence, if $h_{i+1,i} \neq 0$ $(1 \leq i \leq j-1)$, $h_{j+1,j} = 0$, a least squares solution is obtained by step j $(j \leq r+1)$. $\square$

The necessity of the condition $A_{12} = 0$ for GMRES to determine a least squares solution of (1) for all $\boldsymbol{b}, \boldsymbol{x}_0 \in \mathbf{R}^n$ was proved in Theorem 2.6 of Hayami and Sugihara[1].

References

[1] Hayami K, Sugihara M. A geometric view of Krylov subspace methods on singular systems. Numer Linear Algebra Appl. 2011; 18:449–469.

[2] Saad Y, Schultz MH. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J Sci Statist Comput. 1986; 7:856–869.

[3] Brown P, Walker HF. GMRES on (nearly) singular systems. SIAM J Matrix Anal Appl. 1997; 18:37–51.

[4] Sugihara K, Hayami K, Zheng, N. Right preconditioned MINRES for singular systems. Numer Linear Algebra Appl. 2020; 27:e2277. https://doi.org/10.1002/nla.2277