On GMRES for singular EP and GP systems

Keiichi Morikuni[*]   Miroslav Rozložník[†]

Abstract

In this contribution, we study the numerical behavior of the Generalized Minimal Residual (GMRES) method for solving singular linear systems. It is known that GMRES determines a least squares solution without breakdown if the coefficient matrix is range-symmetric (EP), or if its range and nullspace are disjoint (GP) and the system is consistent. We show that the accuracy of GMRES iterates may deteriorate in practice due to three distinct factors: (i) the inconsistency of the linear system; (ii) the distance of the initial residual to the nullspace of the coefficient matrix; (iii) the extremal principal angles between the ranges of the coefficient matrix and its transpose. These factors lead to poor conditioning of the extended Hessenberg matrix in the Arnoldi decomposition and affect the accuracy of the computed least squares solution. We also compare GMRES with the range restricted GMRES (RR-GMRES) method. Numerical experiments show typical behaviors of GMRES for small problems with EP and GP matrices.

1 Introduction

Consider solving linear systems of equations

$$A\boldsymbol{x} = \boldsymbol{b}, \tag{1.1}$$

where $A \in \mathbb{R}^{n \times n}$ may be singular and $\boldsymbol{b} \in \mathbb{R}^n$ is not necessarily in $\mathcal{R}(A) = \{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{y} = A\boldsymbol{x}, \ \boldsymbol{x} \in \mathbb{R}^n\}$, the range of A. We say that $A\boldsymbol{x} = \boldsymbol{b}$ is consistent if $\boldsymbol{b} \in \mathcal{R}(A)$, and otherwise it is inconsistent. If (1.1) is inconsistent, instead of (1.1), it is natural to consider solving the least squares problem

$$\|\boldsymbol{b} - A\boldsymbol{x}\| = \min_{\boldsymbol{u} \in \mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{u}\|, \tag{1.2}$$

where $\|\cdot\|$ denotes the Euclidean norm. We call a minimizer $\boldsymbol{x} \in \operatorname{argmin}_{\boldsymbol{u} \in \mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{u}\|$ a least squares solution, which is not necessarily unique.

In order to analyze iterative methods for solving (1.1) in terms of the spaces associated with A, we give some required definitions and notations. Let $\mathcal{N}(A) = \{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \boldsymbol{0}\}$ be the nullspace of A. Then, we have $\mathcal{N}(A^{\mathsf{T}}) \oplus \mathcal{R}(A) = \mathcal{N}(A) \oplus \mathcal{R}(A^{\mathsf{T}}) = \mathbb{R}^n$, $\dim \mathcal{N}(A^{\mathsf{T}}) = \dim \mathcal{N}(A)$, and $\dim \mathcal{R}(A^{\mathsf{T}}) = \dim \mathcal{R}(A) = \operatorname{rank}(A)$, where $\oplus$ denotes the direct sum of subspaces. Let $r = \operatorname{rank}(A)$ and denote the singular value decomposition (SVD) of A by $U\Sigma V^{\mathsf{T}}$, where $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices $U^{\mathsf{T}}U = UU^{\mathsf{T}} = V^{\mathsf{T}}V = VV^{\mathsf{T}} = \mathrm{I}$, I is the identity matrix, $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r, 0, 0, \ldots, 0) \in \mathbb{R}^{n \times n}$, and σ_i is the ith largest nonzero singular value of A. Let $U = [U_1, U_2]$ and $V = [V_1, V_2]$, where the columns of $U_1 \in \mathbb{R}^{n \times r}$ and $U_2 \in \mathbb{R}^{n \times (n-r)}$ form orthonormal bases of $\mathcal{R}(A) = \mathcal{R}(U_1)$ and $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^{\mathsf{T}}) = \mathcal{R}(U_2)$, respectively, and the

columns of $V_1 \in \mathbb{R}^{n \times r}$ and $V_2 \in \mathbb{R}^{n \times (n-r)}$ form orthonormal bases of $\mathcal{N}(A)^\perp = \mathcal{R}(A^\mathsf{T}) = \mathcal{R}(V_1)$ and $\mathcal{N}(A) = \mathcal{R}(V_2)$, respectively.

We recall the definitions of generalized inverses. We call a matrix $X \in \mathbb{R}^{n \times n}$ the Moore-Penrose generalized inverse (pseudoinverse) of $A \in \mathbb{R}^{n \times n}$ if X satisfies the Penrose equations $AXA = A$, $XAX = X$, $(AX)^\mathsf{T} = AX$, and $(XA)^\mathsf{T} = XA$, denote it by $A^\dagger$, and have the identity $A^\dagger = V_1 \Sigma_r^{-1} U_1^\mathsf{T}$, where $\Sigma_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$. The condition number of A is denoted by $\kappa(A) = \|A\|\|A^\dagger\|$ [5, Definition 1.4.2]. The smallest nonnegative integer k such that $\mathrm{rank}(A^k) = \mathrm{rank}(A^{k+1})$ is called the index of A [10, Definition 7.2.1], and is denoted by $\mathrm{index}(A)$. In addition, $k \geq \mathrm{index}(A) \iff \mathcal{N}(A^k) \oplus \mathcal{R}(A^k) = \mathbb{R}^n$ [10, p. 121]. Let $\mathrm{index}(A) = 1$ and $X \in \mathbb{R}^{n \times n}$ be such that $AXA = A$, $XAX = X$, and $AX = XA$. Then, X is unique, and we call X the group inverse of A and denote it by $A^\#$. The group inverse of the matrix A is on $\mathcal{R}(A)$ equal to the inverse of the restriction of A to its range $\mathcal{R}(A)$ and admits $\mathcal{N}(A)$ as its nullspace [21, Theorem 2]. The group inverse can be characterized by the Jordan canonical form. Let S be a nonsingular matrix such that $J = S^{-1}AS$ is the Jordan canonical form of A. Then, $A^\# = SJ^\dagger S^{-1}$ holds. In particular, we have $\mathcal{R}(A^\#) = \mathcal{R}(A)$, $\mathcal{N}(A^\#) = \mathcal{N}(A)$, and $A^\#A = AA^\# = P_{\mathcal{R}(A), \mathcal{N}(A)}$ [4]. Here, $P_{\mathcal{R}(A), \mathcal{N}(A)}$ denotes the projection onto $\mathcal{R}(A)$ along $\mathcal{N}(A)$.

Now, we express solutions of eqs. (1.1) and (1.2). The vector $\boldsymbol{x}_* = A^\dagger \boldsymbol{b} = V_1 \Sigma_r^{-1} U_1^\mathsf{T} \boldsymbol{b}$ is called the minimum-norm least squares or pseudoinverse solution of (1.1) or (1.2), and it belongs to $\mathcal{R}(A^\mathsf{T})$. We next give the expressions of the residual of (1.2). Denote the orthogonal projector onto $\mathcal{R}(A)$ by $P_{\mathcal{R}(A)} = U_1 U_1^\mathsf{T}$ and that onto $\mathcal{N}(A^\mathsf{T})$ by $P_{\mathcal{N}(A^\mathsf{T})} = U_2 U_2^\mathsf{T}$. If $\boldsymbol{v}|_{\mathcal{S}} \in \mathbb{R}^n$ is the orthogonal projection of a vector $\boldsymbol{v} \in \mathbb{R}^n$ onto the subspace $\mathcal{S} \subseteq \mathbb{R}^n$, then for any $\boldsymbol{x}_0 \in \mathbb{R}^n$ and any $\boldsymbol{b} \in \mathbb{R}^n$, the corresponding residual of (1.2) is $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0 = \boldsymbol{b}|_{\mathcal{N}(A^\mathsf{T})} + \boldsymbol{b}|_{\mathcal{R}(A)} - A\boldsymbol{x}_0 = \boldsymbol{r}_* + \boldsymbol{r}_0|_{\mathcal{R}(A)}$, where $\boldsymbol{r}_* = \boldsymbol{b}|_{\mathcal{N}(A^\mathsf{T})} = P_{\mathcal{N}(A^\mathsf{T})}\boldsymbol{b} \in \mathcal{N}(A^\mathsf{T})$ is the least squares residual and $\boldsymbol{b}|_{\mathcal{R}(A)} = P_{\mathcal{R}(A)}\boldsymbol{b}$. In particular, we have $\boldsymbol{r}_* = \boldsymbol{0}$ for $\boldsymbol{b} \in \mathcal{R}(A)$. Assume $\mathrm{index}(A) = 1$. It holds that $\boldsymbol{x}_\# = A^\# \boldsymbol{b}$ is a solution of $A\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{b} \in \mathcal{R}(A)$ [10], and it is the unique solution of $A\boldsymbol{x} = \boldsymbol{b}$ in $\mathcal{R}(A)$. Furthermore, the minimum Euclidean norm solution of (1.2) satisfies $\boldsymbol{x}_* = P_{\mathcal{R}(A^\mathsf{T})}\boldsymbol{x}_\#$.

In this paper, we are interested in the numerical behavior of the Generalized Minimal Residual (GMRES) method [22] applied in particular to singular systems (1.1). In section 2, we give some well-known conditions under which GMRES determines a solution without breakdown when applied to certain classes of singular matrices. We discuss also a relation to the range-restricted GMRES (RR-GMRES) method proposed in [7]. In section 3, we examine the conditioning of the coefficient matrix A restricted to the Krylov subspaces that significantly influences the numerical behavior of GMRES. We consider first the case of EP (equal projection) or range-symmetric matrices and distinguish between the consistent and inconsistent cases showing that the consistent case is similar to the nonsingular case. Then we discuss the inconsistent EP case where GMRES suffers from an instability, since the convergence means ill-conditioned restriction of A to the Krylov subspaces. In section 4 we study the case of group (GP) matrices. We show that the numerical behavior of GMRES applied to such problems depends substantially on the principal angles between the subspaces $\mathcal{R}(A)$ and $\mathcal{R}(A^\mathsf{T})$. Surprisingly, difficulties can be expected for non-range-symmetric problems even for consistent systems. In section 5, we conclude the paper.

2 GMRES methods and its convergence for singular systems

GMRES for the linear system (1.1) with initial iterate $\boldsymbol{x}_0 \in \mathbb{R}^n$, independent of any particular implementation of the algorithm, determines the kth iterate $\boldsymbol{x}_k$ over $\boldsymbol{x}_0 + \mathcal{K}_k(A, \boldsymbol{r}_0)$ that minimizes $\|\boldsymbol{b} - A\boldsymbol{x}_k\|$, where $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ is the initial residual and $\mathcal{K}_k(A, \boldsymbol{r}_0) = \mathrm{span}\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0\}$ is the Krylov subspace of order k. Note that there exist $\boldsymbol{x}_k$ and hence $\boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k$ for all $k \geq 0$ but $\boldsymbol{x}_k$ may not be unique in the singular case. Denote $\mathcal{K}_k(A, \boldsymbol{r}_0)$ by $\mathcal{K}_k$ for simplicity. It is clear that $\mathcal{K}_k = \mathrm{span}\{\boldsymbol{r}_0\} + A\mathcal{K}_{k-1} \subset \mathrm{span}\{\boldsymbol{r}_*\} + \mathcal{R}(A)$

holds. If $\dim A\mathcal{K}_k = \dim \mathcal{K}_k$, then the problem

$$\|\boldsymbol{b} - A\boldsymbol{x}_k\| = \min_{\boldsymbol{z}\in\mathcal{K}_k} \|\boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{z})\| = \min_{\boldsymbol{z}\in\mathcal{K}_k} \|\boldsymbol{r}_0 - A\boldsymbol{z}\| = \|\boldsymbol{r}_0 - A\boldsymbol{z}_k\| \qquad (2.1)$$

has a unique solution $\boldsymbol{x}_k = \boldsymbol{x}_0 + \boldsymbol{z}_k$ and hence $\boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k \in \boldsymbol{r}_0 + A\mathcal{K}_k$ is uniquely determined.

General studies on Krylov subspace methods in the singular case were done in [16], [24], [31], [23]. Particular studies on GMRES-type methods in the singular case were done in [7], [25], [20], [26], [13]. See [8], [9], [14] for GMRES on ill-posed linear systems, and [32] for GMRES with preconditioning. See also [15] for GMRES and [1] for GMRES with preconditioning in Hilbert spaces.

The GMRES method for solving ill-posed problems was studied in [8], [9], [14]. The pre-conditioned GMRES method applied to singular systems was considered in [32] . See also [15] for GMRES and [1] for the preconditioned GMRES method applied in Hilbert spaces.

In the nonsingular case, GMRES determines the solution of $A\boldsymbol{x} = \boldsymbol{b}$ for all $\boldsymbol{b} \in \mathbb{R}^n$ and for all $\boldsymbol{x}_0 \in \mathbb{R}^n$ within n iterations. In the singular case, GMRES may fail to determine a solution of (1.1), and is said to break down at some step k if $\dim A\mathcal{K}_k < \dim \mathcal{K}_k$ or $\dim \mathcal{K}_k < k$ [6, p. 38]. Note that, in general, $\dim A\mathcal{K}_k \leq \dim \mathcal{K}_k \leq k$ holds for each k.

We give an explicit expression of the iterate $\boldsymbol{x}_k$ for GMRES using the Arnoldi decomposition $AQ_k = Q_{k+1}H_{k+1,k}$, $k = 1, 2, \ldots$, where the columns of $Q_k = [\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k]$ form an orthonormal basis of the Krylov subspace $\mathcal{K}_k$, and $H_{k+1,k} = (h_{i,j}) \in \mathbb{R}^{(k+1)\times k}$ is an extended Hessenberg matrix. Then the iterate is given by $\boldsymbol{x}_k = \boldsymbol{x}_0 + Q_k\boldsymbol{y}_k$ with $\boldsymbol{y}_k = \arg\min_{\boldsymbol{y}\in\mathbb{R}^k}\|\beta\boldsymbol{e}_1 - H_{k+1,k}\boldsymbol{y}\|$, where $\beta = \|\boldsymbol{r}_0\|$, $\boldsymbol{e}_1$ is the first column of the identity matrix and $\|\boldsymbol{b} - A\boldsymbol{x}_k\| = \|\boldsymbol{r}_0 - AQ_k\boldsymbol{y}_k\| = \|\beta\boldsymbol{e}_1 - H_{k+1,k}\boldsymbol{y}_k\|$.

It is clear that if $h_{i+1,i} \neq 0$ for $i = 1, 2, \ldots, k-1$, the breakdown does not occur until step $k-1$ of GMRES with $\dim A\mathcal{K}_i = i$, or $\text{rank}(H_{i+1,i}) = i$, $i = 1, 2, \ldots, k-1$. At breakdown of GMRES at step k with $h_{k+1,k} = 0$, one of the following cases holds [18, Appendix B] (cf. [6, Theorem 2.2]):

Case I. $\dim A\mathcal{K}_{k+1} = k < \dim \mathcal{K}_{k+1} = k + 1$, whereas $\text{rank}(H_{k,k}) = k - 1$.

Case II. $\dim A\mathcal{K}_k = k = \dim \mathcal{K}_{k+1} < k + 1$, whereas $\text{rank}(H_{k,k}) = k$ (GMRES determines a solution of $A\boldsymbol{x} = \boldsymbol{b}$ at step k).

Here, $H_{k,k} = (h_{i,j}) \in \mathbf{R}^{k\times k}$.

A variant of GMRES called the range restricted GMRES (RR-GMRES) method was proposed in [7]. RR-GMRES determines the kth iterate by minimizing the same objective function as GMRES over a different Krylov subspace

$$\|\boldsymbol{b} - A\boldsymbol{x}_k^{\mathrm{R}}\| = \min_{\boldsymbol{z}\in\mathcal{K}_k(A,A\boldsymbol{r}_0)} \|\boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{z})\| = \min_{\boldsymbol{z}\in\mathcal{K}_k(A,A\boldsymbol{r}_0)} \|\boldsymbol{r}_0 - A\boldsymbol{z}\| = \|\boldsymbol{r}_0 - A\boldsymbol{z}_k^{\mathrm{R}}\|.$$

It was shown in [11, Theorem A2] that if RR-GMRES applied to (1.2) breaks down at step m with $\text{rank}(A) = m - 1$ and $\dim A\mathcal{K}_m(A, A\boldsymbol{r}_0) = m - 1$, then it determines a solution of (1.2). Here, RR-GMRES is said to break down if $\dim A\mathcal{K}_k(A, A\boldsymbol{r}_0) < \dim \mathcal{K}_k(A, A\boldsymbol{r}_0)$ or $\dim \mathcal{K}_k(A, A\boldsymbol{r}_0) < k$.

We give an explicit expression of the RR-GMRES iterate $\boldsymbol{x}_k$ using the Arnoldi decomposition $AQ_k^{\mathrm{R}} = Q_{k+1}^{\mathrm{R}}H_{k+1,k}^{\mathrm{R}}$, $k = 1, 2, \ldots$, where the columns of $Q_k^{\mathrm{R}} = [\boldsymbol{q}_1^{\mathrm{R}}, \boldsymbol{q}_2^{\mathrm{R}}, \ldots, \boldsymbol{q}_k^{\mathrm{R}}]$ form an orthonormal basis of the Krylov subspace $\mathcal{K}_k(A, A\boldsymbol{r}_0)$ with the initial vector $\boldsymbol{q}_1^{\mathrm{R}} = A\boldsymbol{r}_0/\|A\boldsymbol{r}_0\|$, and $H_{k+1,k}^{\mathrm{R}} = (h_{i,j}^{\mathrm{R}}) \in \mathbb{R}^{(k+1)\times k}$ is an extended Hessenberg matrix. Then, the iterate is given by $\boldsymbol{x}_k^{\mathrm{R}} = \boldsymbol{x}_0 + Q_k^{\mathrm{R}}\boldsymbol{y}_k^{\mathrm{R}}$ with $\boldsymbol{y}_k^{\mathrm{R}} = \arg\min_{\boldsymbol{y}\in\mathbb{R}^k}\|(Q_{k+1}^{\mathrm{R}})^{\mathsf{T}}\boldsymbol{r}_0 - H_{k+1,k}^{\mathrm{R}}\boldsymbol{y}\|$, where

$$\begin{aligned}
\|\boldsymbol{b} - A\boldsymbol{x}_k^{\mathrm{R}}\|^2 &= \|\boldsymbol{r}_0 - AQ_k^{\mathrm{R}}\boldsymbol{y}_k^{\mathrm{R}}\|^2 \\
&= \|(Q_{k+1}^{\mathrm{R}})^{\mathsf{T}}\boldsymbol{r}_0 - H_{k+1,k}^{\mathrm{R}}\boldsymbol{y}_k^{\mathrm{R}}\|^2 + \|[\mathrm{I} - Q_{k+1}^{\mathrm{R}}(Q_{k+1}^{\mathrm{R}})^{\mathsf{T}}]\boldsymbol{r}_0\|^2 \\
&= \min_{\boldsymbol{y}\in\mathbb{R}^k} \|(Q_{k+1}^{\mathrm{R}})^{\mathsf{T}}\boldsymbol{r}_0 - H_{k+1,k}^{\mathrm{R}}\boldsymbol{y}\|^2 + \|[\mathrm{I} - Q_{k+1}^{\mathrm{R}}(Q_{k+1}^{\mathrm{R}})^{\mathsf{T}}]\boldsymbol{r}_0\|^2.
\end{aligned}$$

The last term is equal to the kth residual norm for the simpler GMRES method [29], which is not larger than the kth residual norm for RR-GMRES, i.e., $\|\boldsymbol{b} - A\boldsymbol{x}_k\| \leq \|\boldsymbol{b} - A\boldsymbol{x}_k^{\mathrm{R}}\|$. Note also that $\|H_{k,k-1}^{\mathrm{R}}\| = \|AQ_{k-1}^{\mathrm{R}}\| \leq \|AQ_k\| = \|H_{k+1,k}\|$ and

$$\sigma_k(H_{k+1,k}) \leq \min_{\boldsymbol{y} \in \mathbb{R}^{k-1} \setminus \{\mathbf{0}\}} \frac{\|AQ_{k-1}^{\mathrm{R}}\boldsymbol{y}\|}{\|Q_{k-1}^{\mathrm{R}}\boldsymbol{y}\|} = \sigma_{k-1}(H_{k,k-1}^{\mathrm{R}})$$

leading to an interesting bound $\kappa(H_{k,k-1}^{\mathrm{R}}) \leq \kappa(H_{k+1,k})$ for $k = 2, \ldots, n-1$, where $\sigma_k(\cdot)$ denotes the kth largest singular value of a matrix.

In the following, we present conditions under which GMRES determines a solution of (1.1). We start with the observation that in the case of $\mathcal{N}(A) \cap \mathcal{R}(A) \neq \{\mathbf{0}\}$, GMRES breaks down and fails to determine a solution.

Proposition 2.1. *If $\boldsymbol{b} \in \mathcal{R}(A)$ and $\mathbf{0} \neq \boldsymbol{r}_0 \in \mathcal{N}(A) \cap \mathcal{R}(A)$, then GMRES breaks down at step 1 without determining a solution of $A\boldsymbol{x} = \boldsymbol{b}$.*

Proof. Since $\boldsymbol{r}_0 \neq \mathbf{0}$, we have $\dim \mathcal{K}_1 = \dim \mathrm{span}\{\boldsymbol{r}_0\} = 1$. Since $\boldsymbol{r}_0 \in \mathcal{N}(A)$ gives $A\boldsymbol{r}_0 = \mathbf{0}$, we have $\dim A\mathcal{K}_1 = \dim \mathrm{span}\{A\boldsymbol{r}_0\} = \dim \mathrm{span}\{\mathbf{0}\} = 0$. Hence, $\dim A\mathcal{K}_1 < \dim \mathcal{K}_1$ holds. Therefore, GMRES breaks down at step 1 without determining a solution of $A\boldsymbol{x} = \boldsymbol{b}$. $\qquad\square$

Similarly to GMRES, RR-GMRES also breaks down at step 1 without determining a solution of $A\boldsymbol{x} = \boldsymbol{b}$ if $\boldsymbol{b} \in \mathcal{R}(A)$ and $\mathbf{0} \neq \boldsymbol{r}_0 \in \mathcal{N}(A) \cap \mathcal{R}(A)$. Therefore, we will restrict our attention to the cases of $\mathcal{N}(A) \cap \mathcal{R}(A) = \{\mathbf{0}\}$. The following statement holds.

Theorem 2.2 ([6, Theorem 2.6], [31, Theorem 3.2]). *If $\mathcal{N}(A) \cap \mathcal{R}(A) = \{\mathbf{0}\}$, then GMRES determines a solution of $A\boldsymbol{x} = \boldsymbol{b}$ without breakdown for all $\boldsymbol{b} \in \mathcal{R}(A)$ and for all $\boldsymbol{x}_0 \in \mathbb{R}^n$. The solution is $\boldsymbol{x}_\# + (\mathrm{I} - A^\# A)\boldsymbol{x}_0$.*

The condition $\mathcal{N}(A) \cap \mathcal{R}(A) = \{\mathbf{0}\}$ is equivalent to $\mathrm{rank}([U_1, V_2]) = n$, or $\mathcal{R}(U_1) \cap \mathcal{R}(V_2) = \{\mathbf{0}\}$ [10, Lemma 7.2.1]. Then, we have $P_{\mathcal{R}(A), \mathcal{N}(A)} = A^\# A = ff[U_1, \mathrm{O}][U_1, V_2]^{-1} = U_1(V_1^{\mathsf{T}} U_1)^{-1} V_1^{\mathsf{T}}$, because of [4, Exercise 30, p. 167] and

$$[U_1, V_2]^{-1} = \begin{bmatrix} (V_1^{\mathsf{T}} U_1)^{-1} & \mathrm{O} \\ -V_2 U_1 (V_1^{\mathsf{T}} U_1)^{-1} & \mathrm{I} \end{bmatrix} [V_1, V_2]^{\mathsf{T}}.$$

Thus, $\kappa(A^\# A) = \kappa(V_1^{\mathsf{T}} U_1)$ holds.

For a special class of singular matrices, GMRES determines a least squares solution.

Theorem 2.3 ([6, Theorem 2.4]). *If $\mathcal{R}(A) = \mathcal{R}(A^{\mathsf{T}})$, then GMRES determines a solution of $\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|$ without breakdown for all $\boldsymbol{b} \in \mathbb{R}^n$ and for all $\boldsymbol{x}_0 \in \mathbb{R}^n$.*

A matrix $A \in \mathbb{R}^{n \times n}$ satisfying $\mathcal{N}(A) \cap \mathcal{R}(A) = \{\mathbf{0}\}$, is called a GP (group) matrix. A GP matrix satisfying in addition $\mathcal{R}(A^{\mathsf{T}}) = \mathcal{R}(A)$, or equivalently $\mathcal{R}(U_1) = \mathcal{R}(V_1)$, is called an EP (equal projection) or range-symmetric matrix. Now, we characterize the GP and EP matrices in terms of their singular value decompositions. The matrix A can be decomposed into

$$A = U \begin{bmatrix} \Sigma_r K & \Sigma_r L \\ \mathrm{O} & \mathrm{O} \end{bmatrix} U^{\mathsf{T}} = V \begin{bmatrix} K \Sigma_r & \mathrm{O} \\ M \Sigma_r & \mathrm{O} \end{bmatrix} V^{\mathsf{T}}$$

with the identity $KK^{\mathsf{T}} + LL^{\mathsf{T}} = \mathrm{I}$, where $K = V_1^{\mathsf{T}} U_2$, $L = V_1^{\mathsf{T}} U_1$, and $M = V_2^{\mathsf{T}} U_1$, and the pseudoinverse matrix of A can be decomposed into

$$A^\dagger = U \begin{bmatrix} K^{\mathsf{T}} \Sigma_r^{-1} & \mathrm{O} \\ L^{\mathsf{T}} \Sigma_r^{-1} & \mathrm{O} \end{bmatrix} U^{\mathsf{T}} = V \begin{bmatrix} \Sigma_r^{-1} K^{\mathsf{T}} & \Sigma_r^{-1} M^{\mathsf{T}} \\ \mathrm{O} & \mathrm{O} \end{bmatrix} V^{\mathsf{T}}.$$

The equivalences for GP matrices

$$\mathcal{R}(U_1) \cap \mathcal{R}(V_2) = \{\mathbf{0}\} \iff \operatorname{rank}([U_1, V_2]) = n \iff V_1^\mathsf{T} U_1 (= K) \text{ is nonsingular}$$

follow from the equation

$$[V_1, V_2]^\mathsf{T} [U_1, V_2] = \begin{bmatrix} V_1^\mathsf{T} U_1 & \mathrm{O} \\ V_2^\mathsf{T} U_1 & \mathrm{I} \end{bmatrix}.$$

The group inverse of a GP matrix A can be decomposed into

$$A^\# = U \begin{bmatrix} K^{-1}\Sigma_r^{-1} & K^{-1}\Sigma_r^{-1}K^{-1}L \\ \mathrm{O} & \mathrm{O} \end{bmatrix} U^\mathsf{T} = V \begin{bmatrix} \Sigma_r^{-1}K^{-1} & \mathrm{O} \\ MK^{-1}\Sigma_r^{-1}K^{-1} & \mathrm{O} \end{bmatrix} V^\mathsf{T}.$$

See [2, Section 1]. The conditioning of $K = V_1^\mathsf{T} U_1$ is independent of the conditioning of A but it gives a difficulty in solving singular linear systems with GMRES. The EP case $\mathcal{R}(A^\mathsf{T}) = \mathcal{R}(A)$ is equivalent to that $K = V_1^\mathsf{T} U_1$ is orthogonal, since $L = V_1^\mathsf{T} U_2 = \mathrm{O}$.

Next, we characterize GP and EP matrices in terms of the principal angles. In the EP case $\mathcal{R}(A^\mathsf{T}) = \mathcal{R}(A)$, the matrix $V_1^\mathsf{T} U_1$ is orthogonal and the cosines of the principal angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^\mathsf{T})$ are all zero. In the GP case, since the columns of U_1 and V_1 form bases of $\mathcal{R}(A)$ and $\mathcal{R}(A^\mathsf{T})$, respectively, the cosines of the canonical angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^\mathsf{T})$ are the singular values of $V_1^\mathsf{T} U_1$ [12, section 1.2]. Hence, the condition number of $V_1^\mathsf{T} U_1$ is related to the extremal principal angles.

Note that due to $\|V_1^\mathsf{T} U_1\| \le 1$ all singular values of $V_1^\mathsf{T} U_1$ are less than or equal to 1 and the number of those equal exactly to 1 gives the dimension of $\mathcal{R}(U_1) \cap \mathcal{R}(V_1)$. So, if $\|V_1^\mathsf{T} U_1\| < 1$, then $\mathcal{R}(U_1) \cap \mathcal{R}(V_1) = \mathcal{R}(A) \cap \mathcal{R}(A^\mathsf{T}) = \{\mathbf{0}\}$. If a matrix $A \in \mathbb{R}^{n \times n}$ satisfies $\mathcal{R}(A) \cap \mathcal{R}(A^\mathsf{T}) = \{\mathbf{0}\}$, it is called a disjoint range (DR) matrix [3].

3 GMRES and EP matrices

As was already noted, the GMRES iterate $\boldsymbol{x}_k = \boldsymbol{x}_0 + \boldsymbol{z}_k$ solves the least squares problem (2.1). Therefore, the restriction of A to the Krylov subspace $\mathcal{K}_k \subseteq \mathbb{R}^n$ denoted by $A|_{\mathcal{K}_k}$ plays an important role in the numerical behavior of GMRES. Indeed, the ill-conditioning of $A|_{\mathcal{K}_k}$ was studied and its condition number

$$\kappa(A|_{\mathcal{K}_k}) = \frac{\max_{\boldsymbol{z} \in \mathcal{K}_k \backslash \{\mathbf{0}\}} \|A\boldsymbol{z}\| \big/ \|\boldsymbol{z}\|}{\min_{\boldsymbol{z} \in \mathcal{K}_k \backslash \{\mathbf{0}\}} \|A\boldsymbol{z}\| \big/ \|\boldsymbol{z}\|}$$

was introduced by Brown and Walker in [6]. In practical computations, the iterate $\boldsymbol{x}_k$ is computed as $\boldsymbol{x}_k = \boldsymbol{x}_0 + Q_k \boldsymbol{y}_k$, where the columns of Q_k form an orthonormal basis of the Krylov subspace $\mathcal{K}_k$ and the vector $\boldsymbol{y}_k$ is a solution of the extended Hessenberg least squares problem $\min_{\boldsymbol{y} \in \mathbb{R}^k} \|\beta \boldsymbol{e}_1 - H_{k+1,k} \boldsymbol{y}\|$ (see section 2). The accuracy of $\boldsymbol{x}_k$ is thus affected directly by the conditioning of the matrix $H_{k+1,k}$, whereas the identity $\kappa(H_{k+1,k}) = \kappa(A|_{\mathcal{K}_k})$ follows from the identities

$$\{\max, \min\}_{\boldsymbol{z} \in \mathcal{K}_k \backslash \{\mathbf{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \{\max, \min\}_{\boldsymbol{w} \in \mathbb{R}^k \backslash \{\mathbf{0}\}} \frac{\|AQ_k\boldsymbol{w}\|}{\|Q_k\boldsymbol{w}\|} = \{\max, \min\}_{\boldsymbol{w} \in \mathbb{R}^k \backslash \{\mathbf{0}\}} \frac{\|H_{k+1,k}\boldsymbol{w}\|}{\|\boldsymbol{w}\|}.$$

Next, we give bounds on the extremal singular values of $H_{k+1,k}$. The norm of the matrix $H_{k+1,k}$ can be always bounded above by that of A

$$\|H_{k+1,k}\| = \max_{\boldsymbol{z} \in \mathcal{K}_k \backslash \{\mathbf{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} \le \max_{\boldsymbol{z} \in \operatorname{span}\{\boldsymbol{r}_*\} \oplus \mathcal{R}(A) \backslash \{\mathbf{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} \le \max_{\boldsymbol{z} \in \mathbb{R}^n \backslash \{\mathbf{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \|A\|.$$

This approach cannot be used to bound the kth (or smallest) singular value of $H_{k+1,k}$ due to

$$\sigma_k(H_{k+1,k}) = \min_{z \in \mathcal{K}_k \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} \geq \min_{z \in \mathrm{span}\{r_*\} \oplus \mathcal{R}(A) \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} \geq \min_{z \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} = 0 \qquad (3.1)$$

as the last equality holds for A singular.

In the consistent case, the condition number is bounded by $\kappa(H_{k+1,k}) \leq \kappa(A|_{\mathcal{R}(A)})$ from $\mathcal{K}_k \subseteq \mathcal{R}(A)$ and

$$\sigma_k(H_{k+1,k}) \geq \min_{z \in \mathcal{R}(A) \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|},$$

where $A|_{\mathcal{R}(A)}$ denotes the restriction of A to the range $\mathcal{R}(A)$. If A is an EP matrix $\mathcal{R}(A) = \mathcal{R}(A^\mathsf{T}) = \mathcal{N}(A)^\perp$, then

$$\min_{z \in \mathcal{R}(A) \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} = \min_{z \in \mathcal{R}(A^\mathsf{T}) \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} = \sigma_r(A) > 0$$

and

$$\kappa(A|_{\mathcal{R}(A)}) = \frac{\|A\|}{\min_{z \in \mathcal{R}(A) \setminus \{\mathbf{0}\}} \|Az\|/\|z\|} = \kappa(A).$$

Indeed, the consistent EP case is similar to the nonsingular case, and the condition number of the extended Hessenberg matrix $H_{k+1,k}$ is bounded by $\kappa(H_{k+1,k}) \leq \kappa(A)$ (cf. [33, Remark 3.2, Theorem 3.6]). Consequently, the rank deficiency of the least squares problem (2.1) cannot occur and GMRES will terminate if a solution is reached at some step with a degeneracy of the Krylov subspace at the next step.

In the inconsistent EP case, the equivalence $\mathcal{R}(A^\mathsf{T}) = \mathcal{R}(A) \iff \mathcal{N}(A^\mathsf{T}) = \mathcal{N}(A)$ shows that the nonzero least squares residual $r_* \in \mathcal{N}(A^\mathsf{T})$ belongs also to $\mathcal{N}(A)$ and

$$\sigma_k(H_{k+1,k}) \geq \min_{z \in \mathrm{span}\{r_*\} \oplus \mathcal{R}(A) \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} = 0.$$

It follows from (2.1) that the residual r_{k-1} at step $k-1$ belongs to the Krylov subspace $\mathcal{K}_k$ and satisfies $r_{k-1} - r_* \in \mathcal{R}(A)$. In addition, due to $Ar_* = \mathbf{0}$ we have

$$\begin{aligned}
\sigma_k(H_{k+1,k}) &= \min_{z \in \mathcal{K}_k \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} \leq \frac{\|Ar_{k-1}\|}{\|r_{k-1}\|} \\
&= \frac{\|A(r_{k-1} - r_*)\|}{\|r_{k-1}\|} \leq \|A\| \frac{\|r_{k-1} - r_*\|}{\|r_{k-1}\|}.
\end{aligned} \qquad (3.2)$$

This result was derived in a somewhat different form in [6, Theorem 2.5]. It is clear that in the inconsistent case, the least squares problem (2.1) becomes ill-conditioned as the GMRES iterate converges to a least squares solution. This situation is illustrated in Figure 3.1. Note that we also have

$$\sigma_k(H_{k+1,k}) = \min_{z \in \mathcal{K}_k \setminus \{\mathbf{0}\}} \frac{\|Az\|}{\|z\|} \leq \frac{\|Ar_0\|}{\|r_0\|} \leq \frac{\|A\| \|r_0|_{\mathcal{R}(A^\mathsf{T})}\|}{\|r_0\|} = \frac{\|A\| \|r_0|_{\mathcal{R}(A)}\|}{\|r_0\|}.$$

This bound indicates that if the norm of Ar_0 is too small, then $H_{k+1,k}$ becomes ill-conditioned and the inaccuracy can be expected at all subsequent steps of GMRES. Finally, since a symmetric matrix is an EP matrix, the above discussion also covers the MINRES method [19] applied to symmetric singular systems.
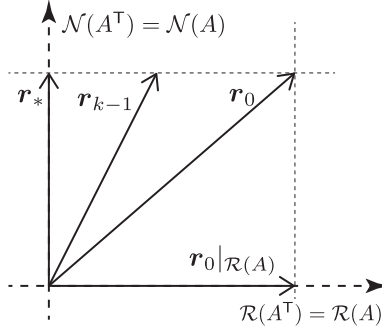
Figure 3.1: Geometric illustration of residual vectors in the EP case.

The conditioning of the extended Hessenberg matrix $H_{k+1,k}$ for GMRES and its relation to the conditioning of A are illustrated on small examples. First for simplicity, consider applying GMRES with $\boldsymbol{x}_0 = \boldsymbol{0}$ to $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}, \quad \varepsilon > 0. \tag{3.3}$$

The matrix A is EP, it has the range $\mathcal{R}(A) = \mathcal{R}(A^{\mathsf{T}}) = \mathrm{span}\{[1,0]^{\mathsf{T}}\}$ and the nullspace $\mathcal{N}(A) = \mathcal{N}(A^{\mathsf{T}}) = \mathrm{span}\{[0,1]^{\mathsf{T}}\}$, and its minimum nonzero singular value is $\sigma_1(A) = \min_{\boldsymbol{z} \in \mathcal{R}(A) \backslash \{\boldsymbol{0}\}} \|A\boldsymbol{z}\|/\|\boldsymbol{z}\| = 1$. The first two steps of the Arnoldi process for A and the initial vector $\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$ give the decomposition $AQ_2 = Q_2 H_{2,2}$, where

$$Q_2 = [\boldsymbol{q}_1, \boldsymbol{q}_2] = \frac{1}{\sqrt{1+\varepsilon^2}} \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & -1 \end{bmatrix}, \quad H_{2,2} = \frac{1}{1+\varepsilon^2} \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & \varepsilon^2 \end{bmatrix}.$$

Hence, $H_{2,2}$ is singular and we have $\sigma_1(A) = \sigma_1(H_{2,2}) = 1$. Solving $\min_{\boldsymbol{y} \in \mathbb{R}^2} \|\beta \boldsymbol{e}_1 - H_{2,2}\boldsymbol{y}\|$, where $\beta = \sqrt{1+\varepsilon^2}$, we have the minimum-norm solution $\boldsymbol{y}_2 = 1/\sqrt{1+\varepsilon^2}[1,\varepsilon]^{\mathsf{T}}$ and $\|\boldsymbol{y}_2\| = 1$. Therefore, $\|\boldsymbol{x}_2\| = \|\boldsymbol{y}_2\|$ and thus the norm of the iterate is not large, even if ε is very small. It is also clear that for $\varepsilon = 0$ the system (3.3) becomes consistent and then GMRES will deliver the minimum norm solution $\boldsymbol{x}_* = [1,0]^{\mathsf{T}}$ in one iteration.

For comparison, we also consider the 2×2 nonsingular ill-posed linear system $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \delta \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}, \quad \delta, \varepsilon > 0 \tag{3.4}$$

with the condition number $\kappa(A) = 1/\delta$, where δ is a small scalar and the right-hand side $\boldsymbol{b}$ is contaminated by the error $[0,\varepsilon]^{\mathsf{T}}$. The exact solution is $[1,0]^{\mathsf{T}}$ and the error-contaminated solution is $[1, \varepsilon/\delta]^{\mathsf{T}}$. GMRES applied to $A\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{x}_0 = \boldsymbol{0}$ gives the same Q_2 as the one for (3.3) but the different Hessenberg matrix

$$H_{2,2} = \frac{1}{1+\varepsilon^2} \begin{bmatrix} 1+\delta\varepsilon^2 & (1-\delta)\varepsilon \\ (1-\delta)\varepsilon & \delta+\varepsilon^2 \end{bmatrix},$$

whose condition number is the same as that of A. Thus, we have the upper bound $\|A^{-1}\boldsymbol{b}\| = \|\boldsymbol{x}_2\| = \|\boldsymbol{y}_2\| \le \|A^{-1}\|\|\boldsymbol{b}\| = \|H_{2,2}^{-1}\|\|\boldsymbol{r}_0\| = \sqrt{1+\varepsilon^2}/\delta$. This means that if $\varepsilon \ne 0$ and δ is small, then the iterates computed by GMRES can be inaccurate. In the exactly singular case with $\delta = 0$, this is not the case.

In the following numerical examples, we examine the accuracy of the GMRES iterate with respect to the degree of consistency of linear systems by using the test matrix and right hand

side vectors

$$A = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{128 \times 128}, \quad \boldsymbol{b} = \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\delta} \end{bmatrix}, \tag{3.5}$$

where $D = \operatorname{diag}(10^{\frac{0}{63}}, 10^{\frac{-4}{63}}, 10^{\frac{-8}{63}}, \ldots, 10^{-4}) \in \mathbb{R}^{64 \times 64}$, $\boldsymbol{\gamma} = [\gamma, \gamma, \ldots, \gamma]^{\mathsf{T}} \in \mathbb{R}^{64}$ and $\boldsymbol{\delta} = [\delta, \delta, \ldots, \delta]^{\mathsf{T}} \in \mathbb{R}^{64}$. Hence, A has the condition number 10^4, and $\boldsymbol{b} \notin \mathcal{R}(A) \iff \delta \neq 0$. The degree of inconsistency of the linear system $A\boldsymbol{x} = \boldsymbol{b}$ can be controlled by the ratio between γ and δ, as $\|\boldsymbol{b}|_{\mathcal{R}(A)}\| = 8\gamma$ and $\|\boldsymbol{b}|_{\mathcal{N}(A^{\mathsf{T}})}\| = 8\delta$ hold. Since $\mathcal{R}(A) = \mathcal{R}(A^{\mathsf{T}})$, GMRES should determine the least squares solution of $\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|$ for all $\boldsymbol{b} \in \mathbb{R}^{128}$ (Theorem 2.3). Throughout all our numerical experiments, we use GMRES and RR-GMRES with the Householder orthogonalization process [28] to ensure the best possible orthogonality among the Arnoldi basis vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k$ and we compute the kth residual $\boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k$ explicitly from $\boldsymbol{x}_k$ by solving the extended Hessenberg least squares problem $\min_{\boldsymbol{y} \in \mathbb{R}^k} \|\beta \boldsymbol{e}_1 - H_{k+1,k}\boldsymbol{y}\|$ with the Matlab backslash solver, which utilizes column pivoting. A mathematically equivalent solution of $\min_{\boldsymbol{y} \in \mathbb{R}^k} \|\beta \boldsymbol{e}_1 - H_{k+1,k}\boldsymbol{y}\|$ was presented in [17].

Figure 3.2 shows the relative residual norm $\|A^{\mathsf{T}}\boldsymbol{r}_k\|/\|A^{\mathsf{T}}\boldsymbol{b}\|$ versus the number of iterations of GMRES in the weakly inconsistent cases $(\gamma, \delta) = (1, 0)$, $(1, 10^{-12})$, $(1, 10^{-8})$, and $(1, 10^{-4})$ on the left, and in the strongly inconsistent cases $(\gamma, \delta) = (1, 1)$, $(10^{-4}, 1)$, $(10^{-8}, 1)$, and $(10^{-12}, 1)$ on the right. Note that the relative residual norm $\|A^{\mathsf{T}}\boldsymbol{r}_k\|/\|A^{\mathsf{T}}\boldsymbol{b}\|$ is associated with the normal equations $A^{\mathsf{T}}A\boldsymbol{x} = A^{\mathsf{T}}\boldsymbol{b}$, which are mathematically equivalent to the linear least squares problem $\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|$. Similarly, Figures 3.3 and 3.4 show the relative residual error norm $\|\boldsymbol{r}_k - \boldsymbol{r}_*\|/\|\boldsymbol{r}_k\|$ (cf. equation (3.2)) and the extremal singular values of A and $H_{k+1,k}$, respectively. If the inconsistency is small ($\delta \ll \gamma$), then GMRES is sufficiently accurate (Figure 3.2a); otherwise the relative residual norm $\|A^{\mathsf{T}}\boldsymbol{r}_k\|/\|A^{\mathsf{T}}\boldsymbol{b}\|$ stagnates before attaining the accuracy on the level of $u\kappa(A)$ (Figure 3.2b), where $u \simeq 1.1 \cdot 10^{-16}$ is the unit roundoff. In contrast to the nonsingular case, GMRES deteriorates not only due to the condition number of A but also due to the inconsistency measured here by $\delta > 0$. For strongly inconsistent systems with $\delta \gg \gamma$, i.e., for $\boldsymbol{r}_0$ close to $\mathcal{N}(A)$, even though $\|A^{\mathsf{T}}\boldsymbol{r}_k\|/\|A^{\mathsf{T}}\boldsymbol{b}\|$ is large and stagnates, and $\|\boldsymbol{r}_*\|$ and hence $\|\boldsymbol{r}_k\|$ are large, the residual $\boldsymbol{r}_k$ approaches $\boldsymbol{r}_*$. Figure 3.4b shows that for strongly inconsistent problems, $H_{k+1,k}$ has a condition number significantly larger than A, tends to become more ill-conditioned in the subsequent steps, and becomes numerically rank-deficient with $u\|H_{k+1,k}\|\|H_{k+1,k}^{\dagger}\| \geq 1$ as the iteration proceeds. In particular, for $\gamma = 0$ and $\delta = 1$, GMRES breaks down at step 1 but gives a least squares solution. Comparing Figures 3.3b and 3.4b, we see that in these cases the bound (3.2) gives a reasonably good upper estimate for the smallest singular value of $H_{k+1,k}$. The behavior described by this example is illustrated on a practical problem from a discretization of a partial differential equation with periodic boundary conditions in [6, Experiment 4.2].

A remedy for the ill-conditioning occurring in GMRES due to inconsistency is to form the Krylov subspace $\mathcal{K}_k(A, A\boldsymbol{r}_0)$ by starting with the initial vector $A\boldsymbol{r}_0$ in $\mathcal{R}(A)$ instead of $\boldsymbol{r}_0$ as is done in RR-GMRES. Note that, on the other hand, the RR-GMRES residual norm is always larger than or equal to the GMRES residual norm (see section 2). Similarly to the above, we show numerical results for RR-GMRES on the same inconsistent linear systems (3.5). Figures 3.5 to 3.7 show the same quantities as Figures 3.2 to 3.4 for RR-GMRES. For any inconsistency parameter $\delta > 0$, the condition number of $H_{k+1,k}^R$ is bounded above by the condition number of A and RR-GMRES is sufficiently accurate, as

$$\sigma_k(H_{k+1,k}^{\mathrm{R}}) = \min_{\boldsymbol{z} \in \mathcal{K}_k(A, A\boldsymbol{r}_0) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|}$$

$$\geq \min_{\boldsymbol{z} \in \mathcal{R}(A) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \min_{\boldsymbol{z} \in \mathcal{R}(A^{\mathsf{T}}) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \sigma_r(A)$$

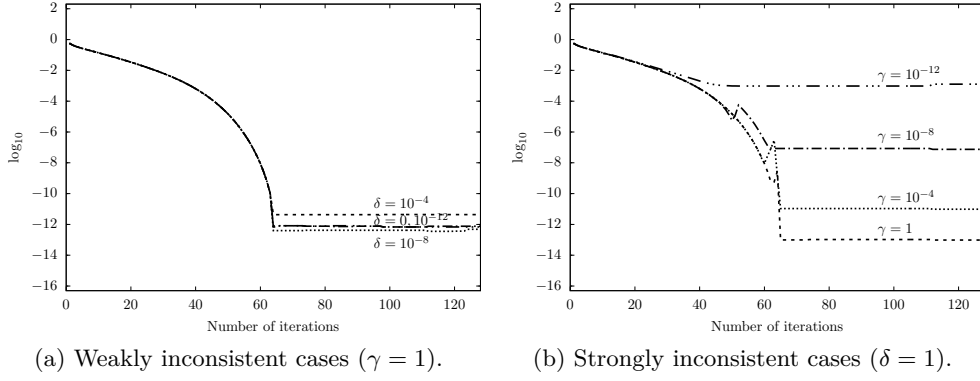for $\mathcal{R}(A) = \mathcal{R}(A^{\mathsf{T}})$. Thus, the accuracy of the RR-GMRES iterate is affected only by the

(a) Weakly inconsistent cases ($\gamma = 1$). (b) Strongly inconsistent cases ($\delta = 1$).

Figure 3.2: Relative residual norm $\|A^\mathsf{T} \boldsymbol{r}_k\| / \|A^\mathsf{T} \boldsymbol{b}\|$ for GMRES.

(a) Weakly inconsistent cases ($\gamma = 1$). (b) Strongly inconsistent cases ($\delta = 1$).
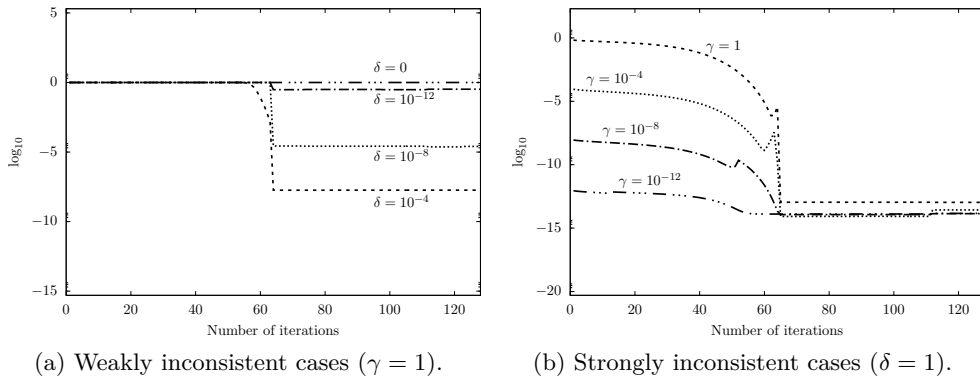
Figure 3.3: Relative residual error norm $\|\boldsymbol{r}_k - \boldsymbol{r}_*\| / \|\boldsymbol{r}_k\|$ for GMRES.

(a) Weakly inconsistent cases ($\gamma = 1$). (b) Strongly inconsistent cases ($\delta = 1$).
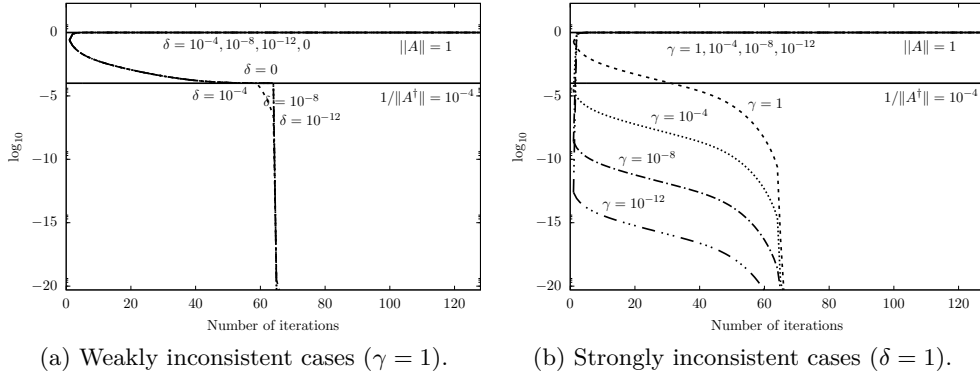
Figure 3.4: Extremal singular values of A and $H_{k+1,k}$ for GMRES.

condition number of A, even though the inconsistency increases or $\boldsymbol{r}_0$ approaches $\mathcal{N}(A)$. Hence, for inconsistent problems with EP matrices RR-GMRES is a successful alternative to GMRES.

4 GMRES and GP matrices

We have shown in section 3 that the condition number $\kappa(A|_{\mathcal{R}(A)})$ plays an important role in the behavior of GMRES and in the EP case we have $\kappa(A|_{\mathcal{R}(A)}) = \kappa(A)$. Thus for consistent problems with EP matrices, the condition number of A represents an upper bound for the condition number of $H_{k+1,k}$ due to $\kappa(H_{k+1,k}) \leq \kappa(A|_{\mathcal{R}(A)}) = \kappa(A)$ and the accuracy of the GMRES iterates is actually determined by the singular values of A. Consider now applying
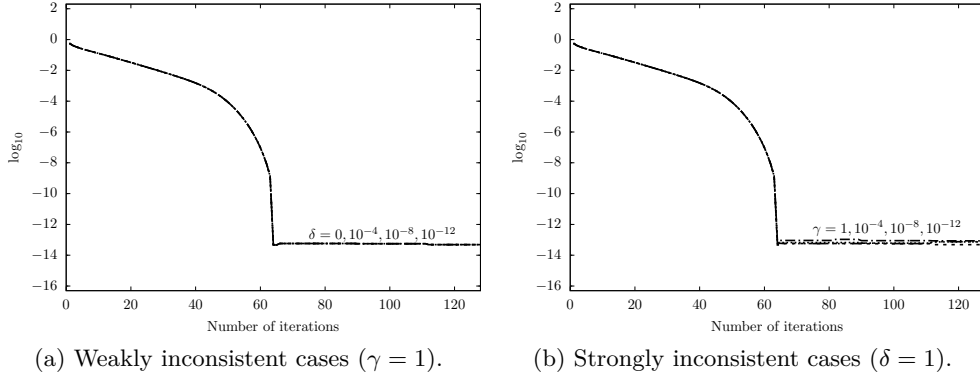
9

(a) Weakly inconsistent cases ($\gamma = 1$).   (b) Strongly inconsistent cases ($\delta = 1$).

Figure 3.5: Relative residual norm $\|A^\mathsf{T}\boldsymbol{r}_k^\mathrm{R}\|/\|A^\mathsf{T}\boldsymbol{b}\|$ for RR-GMRES.

(a) Weakly inconsistent cases ($\gamma = 1$).   (b) Strongly inconsistent cases ($\delta = 1$).
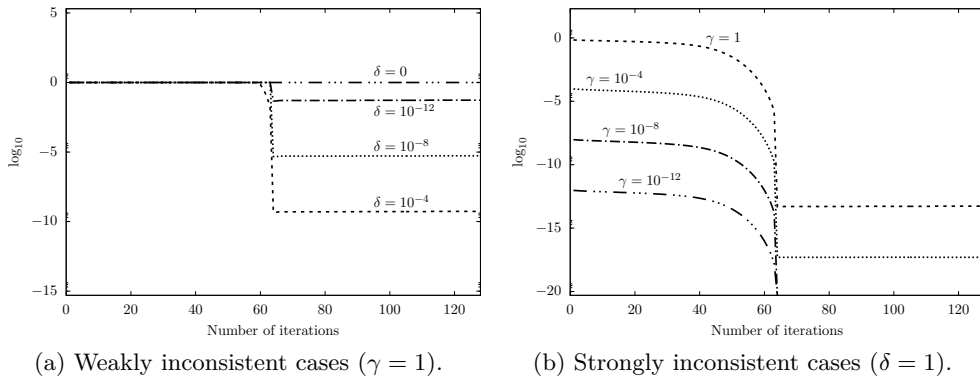
Figure 3.6: Relative residual error norm $\|\boldsymbol{r}_k^\mathrm{R} - \boldsymbol{r}_*\|/\|\boldsymbol{r}_k^\mathrm{R}\|$ for RR-GMRES.

(a) Weakly inconsistent cases ($\gamma = 1$).   (b) Strongly inconsistent cases ($\delta = 1$).
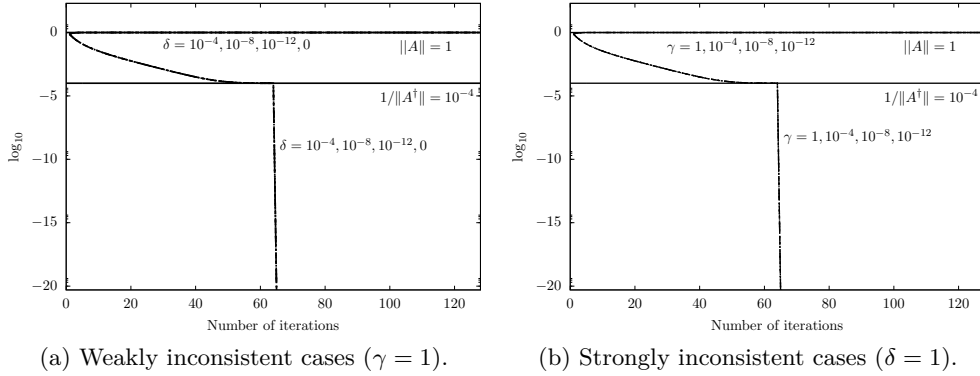
Figure 3.7: Extremal singular values of A and $H_{k+1,k}^\mathrm{R}$ for RR-GMRES.

GMRES to $A\boldsymbol{x} = \boldsymbol{b}$, where A is a GP matrix (Theorem 2.2). We will show that in the GP case $\kappa(A|_{\mathcal{R}(A)})$ can be significantly larger than $\kappa(A)$ and thus the condition number $\kappa(H_{k+1,k})$ can become larger than $\kappa(A)$ even in the consistent case. The accuracy of the GMRES iterates can be then affected by the inaccurate solution of the extended Hessenberg least squares problem that can be ill-conditioned even if A is well-conditioned.

According to Theorem 2.2, GMRES in the consistent GP case determines $\boldsymbol{x}_\# + (\mathrm{I} - A^\# A)\boldsymbol{x}_0$. The vector $\boldsymbol{x}_\#$ that belongs to $\mathcal{R}(A)$ can be related to the vector $\boldsymbol{x}_*$ that belongs to $\mathcal{R}(A^\mathsf{T})$ as follows:

$$\sigma_r(V_1^\mathsf{T} U_1)\|\boldsymbol{x}_\#\| \leq \|\boldsymbol{x}_*\| \leq \|\boldsymbol{x}_\#\|, \tag{4.1}$$

10

which follows from the identity $\boldsymbol{x}_* = P_{\mathcal{R}(A^{\mathsf{T}})}\boldsymbol{x}_{\#} = V_1 V_1^{\mathsf{T}} U_1 U_1^{\mathsf{T}} \boldsymbol{x}_{\#}$. Note that $\boldsymbol{x}_{\#}$ may have a large component in $\mathcal{N}(A)$, if the angle between $\mathcal{N}(A)$ and $\mathcal{R}(A)$ is small, which may affect the accuracy of GMRES iterates (see Figure 4.1), due to the ill-conditioning of the extended Hessenberg matrix $H_{k+1,k}$.

In the consistent case, the extremal singular values of $H_{k+1,k}$ can be bounded as

$$\sigma_1(H_{k+1,k}) \leq \max_{\boldsymbol{z} \in \mathcal{R}(A) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \max_{\boldsymbol{z} \in \mathbb{R}^r \setminus \{\boldsymbol{0}\}} \frac{\|U_1 \Sigma_r V_1^{\mathsf{T}} U_1 \boldsymbol{z}\|}{\|U_1 \boldsymbol{z}\|} \leq \|A\| \|V_1^{\mathsf{T}} U_1\|,$$

$$\sigma_k(H_{k+1,k}) \geq \min_{\boldsymbol{z} \in \mathcal{R}(A) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \min_{\boldsymbol{z} \in \mathbb{R}^r \setminus \{\boldsymbol{0}\}} \frac{\|U_1 \Sigma_r V_1^{\mathsf{T}} U_1 \boldsymbol{z}\|}{\|U_1 \boldsymbol{z}\|} \geq \sigma_r(A) \sigma_r(V_1^{\mathsf{T}} U_1). \tag{4.2}$$

Consequently, $\kappa(H_{k+1,k}) \leq \kappa(A)\kappa(V_1^{\mathsf{T}} U_1)$ is related to the extremal principal angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^{\mathsf{T}})$ (cf. [30, Theorem 2.1]). The lower bound (4.2) shows that, in the consistent case, the smallest singular value of $H_{k+1,k}$ can be smaller than the smallest nonzero singular value of A, depending on the smallest nonzero singular value of $V_1^{\mathsf{T}} U_1$. In addition, it is easy to see that $\sigma_k(H_{k+1,k})$ can be bounded by

$$\sigma_k(H_{k+1,k}) = \min_{\boldsymbol{z} \in \mathcal{K}_k \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} \leq \frac{\|A\boldsymbol{r}_0\|}{\|\boldsymbol{r}_0\|} \leq \frac{\|A\| \||\boldsymbol{r}_0|_{\mathcal{R}(A^{\mathsf{T}})}\|}{\|\boldsymbol{r}_0\|}. \tag{4.3}$$

Here, the last inequality is implied by the splitting $A\boldsymbol{r}_0 = A(\boldsymbol{r}_0|_{\mathcal{N}(A)} + \boldsymbol{r}_0|_{\mathcal{R}(A^{\mathsf{T}})}) = A\boldsymbol{r}_0|_{\mathcal{R}(A^{\mathsf{T}})}$. Although Brown and Walker mention in [6, p. 50] that the condition number of $A|_{\mathcal{K}_k}$ cannot become arbitrarily large through an unfortunate choice of $\boldsymbol{b}$ and $\boldsymbol{x}_0$, it is clear from (4.3) that if the residual $\boldsymbol{r}_0$ has a very small component in $\mathcal{R}(A^{\mathsf{T}})$ then the condition number of $H_{k+1,k}$ can be large for any singular matrix A.

In the following, we give illustrative examples that lead to ill-conditioned extended Hessenberg matrix $H_{k+1,k}$ in GMRES. First, we consider GMRES with $\boldsymbol{x}_0 = \boldsymbol{0}$ applied to $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} \varepsilon & 1 \\ 0 & 0 \end{bmatrix}, \quad A^{\#} = \begin{bmatrix} 1/\varepsilon & 1/\varepsilon^2 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{4.4}$$

The matrix A has the following ranges and nullspaces

$$\mathcal{R}(A) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \quad \mathcal{N}(A) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ -\varepsilon \end{bmatrix} \right\},$$

$$\mathcal{R}(A^{\mathsf{T}}) = \operatorname{span}\left\{ \begin{bmatrix} \varepsilon \\ 1 \end{bmatrix} \right\}, \quad \mathcal{N}(A^{\mathsf{T}}) = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$
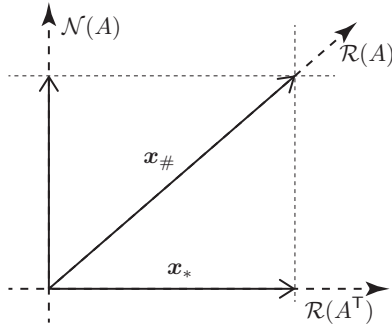

Figure 4.1: Geometric illustration of solution vectors in the GP case.

It is clear that for $\varepsilon = 0$ the matrix A is DR (see section 2). In addition, the initial vector $\boldsymbol{r}_0$ satisfies $\boldsymbol{r}_0 \in \mathcal{N}(A) \cap \mathcal{R}(A)$ and thus the GMRES method breaks down at step 1. Now suppose that $0 < \varepsilon \ll 1$. Then, the matrix A is GP but not EP, i.e., $\mathcal{R}(A^{\mathsf{T}}) \neq \mathcal{R}(A)$. Since $\sigma_1(V_1^{\mathsf{T}} U_1) = \varepsilon/\sqrt{1+\varepsilon^2}$, we have

$$\min_{\boldsymbol{z} \in \mathcal{R}(A) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \varepsilon, \qquad \min_{\boldsymbol{z} \in \mathcal{R}(A^{\mathsf{T}}) \setminus \{\boldsymbol{0}\}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|} = \sqrt{1+\varepsilon^2}.$$

The smallest singular value of $H_{2,1}$ is significantly smaller than the smallest nonzero singular value of A, $\sigma_1(H_{2,1}) = \varepsilon \ll \sqrt{1+\varepsilon^2} = \sigma_1(A)$. Indeed, the components of $H_{2,1}$ are $H_{2,1} = [\varepsilon, 0]^{\mathsf{T}}$. Furthermore, by solving $\min_{\boldsymbol{y} \in \mathbb{R}^1} \|\beta \boldsymbol{e}_1 - H_{2,1}\boldsymbol{y}\|$ with $\beta = 1$, we have $\boldsymbol{y}_1 = 1/\varepsilon$, i.e., $\boldsymbol{y}_1$ has a large component. We see that $\boldsymbol{x}_1 = Q_1 \boldsymbol{y}_1$ and $\boldsymbol{x}_1 = \boldsymbol{x}_\# = A^\# \boldsymbol{b} = [1/\varepsilon, 0]^{\mathsf{T}}$ for $\boldsymbol{b} = [1, 0]^{\mathsf{T}}$. Therefore, $\|\boldsymbol{x}_1\|$ becomes very large even if the condition number of A and the norm of the right-hand side are small. Thus, the vector $\boldsymbol{x}_\#$ contains a large component in $\mathcal{N}(A)$, whereas $\boldsymbol{x}_* = 1/(1+\varepsilon^2)[\varepsilon, 1]^{\mathsf{T}}$, and the inequalities (4.1) are satisfied.

In the following, we also consider the nonsingular ill-posed linear system $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} \varepsilon & 1 \\ 0 & \delta \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ -\varepsilon \end{bmatrix} \quad \delta, \varepsilon > 0, \tag{4.5}$$

where δ is a small parameter and the right-hand side $\boldsymbol{b}$ is contaminated by the error $[0, -\varepsilon]^{\mathsf{T}}$. The extremal singular values of A are $\sigma_1(A) \simeq 1$ and $\sigma_2(A) \simeq \delta\varepsilon$. Indeed, the condition number of A is bounded as

$$\kappa(A) = \left(\frac{1 + \delta^2 + \varepsilon^2 + \sqrt{(1+\delta^2+\varepsilon^2)^2 - 4\delta^2\varepsilon^2}}{1 + \delta^2 + \varepsilon^2 - \sqrt{(1+\delta^2+\varepsilon^2)^2 - 4\delta^2\varepsilon^2}} \right)^{1/2} \geq \frac{1}{\delta\varepsilon} \tag{4.6}$$

Then, GMRES applied to $A\boldsymbol{x} = \boldsymbol{b}$ with the error-free right-hand side $\boldsymbol{b} = [1, 0]^{\mathsf{T}}$ and $\boldsymbol{x}_0 = \boldsymbol{0}$ gives the same Arnoldi vector $\boldsymbol{q}_1$ and the same extended Hessenberg matrix $H_{2,1}$ as the ones for (4.4). GMRES applied to $A\boldsymbol{x} = \boldsymbol{b}$ with (4.5) give

$$Q_2 = \frac{1}{\sqrt{1+\varepsilon^2}} \begin{bmatrix} 1 & \varepsilon \\ -\varepsilon & 1 \end{bmatrix}, \quad H_{2,2} = \frac{1}{1+\varepsilon^2} \begin{bmatrix} \delta\varepsilon^2 & 1+\varepsilon^2 - \delta\varepsilon \\ \delta\varepsilon & \varepsilon(1+\varepsilon^2) + \delta \end{bmatrix}.$$

Hence, we have the upper bound $\|A^{-1}\boldsymbol{b}\| = \|\boldsymbol{x}_2\| = \|\boldsymbol{y}_2\| \leq \|A^{-1}\|\|\boldsymbol{b}\| = \|H_{2,2}^{-1}\|\|\boldsymbol{r}_0\| = \sqrt{1+\varepsilon^2}/(\delta\epsilon)$. This means that GMRES breaks down in the exact singular case, whereas, in the ill-posed case, GMRES does not break down but the iterates computed by GMRES can be inaccurate as δ, ϵ go to zero. In both exactly singular case and nonsingular ill-posed case, the iterates computed by GMRES will be inaccurate due to the ill-conditioning of $H_{2,2}$.

The above behavior of GMRES for singular linear systems with respect to the condition number of $V_1^{\mathsf{T}} U_1$ is illustrated on numerical examples $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} D & I \\ O & O \end{bmatrix} \in \mathbb{R}^{128 \times 128}, \quad \boldsymbol{b} = \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{0} \end{bmatrix}, \tag{4.7}$$

$D = \operatorname{diag}(d_1, d_2, \ldots, d_{64}) \in \mathbb{R}^{64 \times 64}$ is a diagonal matrix whose diagonal entries have the so-called Strakoš distribution [27]

$$d_1 = 1, \quad d_{64} = 10^{-\rho}, \quad d_i = d_{64} + \frac{64 - i}{63}(d_1 - d_{64}) \cdot 0.7^{i-1}, \quad i = 2, 3, \ldots, 63,$$

and $\boldsymbol{f} = (f_i) \in \mathbb{R}^{64}$ has the entries $f_i = 10^{-(64-i)\rho/63}$, $j = 1, 2, \ldots, 64$. Note that A is GP but not EP. This setting gives well-conditioned A with $\kappa(A) = \sqrt{2/(10^{-2\rho} + 1)} \simeq \sqrt{2}$ and ill-conditioned $V_1^{\mathsf{T}} U_1$ for $\kappa(V_1^{\mathsf{T}} U_1) = 10^\rho \sqrt{(10^{-2\rho} + 1)/2} \simeq 10^\rho/\sqrt{2}$ for $\rho \gg 1$. Furthermore,

the norms of vectors $\boldsymbol{f}$, $D\boldsymbol{f}$, $D^2\boldsymbol{f}$ decrease and this reduction is pronounced as the value of ρ increases.

Figures 4.2 and 4.3 show the relative residual norm and the smallest singular value of $H_{k+1,k}$ and $H^{\mathrm{R}}_{k+1,k}$ versus the number of iterations for GMRES and RR-GMRES, respectively, applied to the above linear systems with $\rho = 1$, 4, 8, and 12. As the value of ρ increases, the condition number of the extended Hessenberg matrix increases, and the accuracy of the relative residual for both GMRES and RR-GMRES is significantly lost. It is clear from our experiments that while RR-GMRES does help in the inconsistent EP case by starting with a vector in $\mathcal{R}(A)$ to construct the Krylov subspace, in the GP case both GMRES and RR-GMRES may not give accurate solutions, even in the consistent case, when the condition number of $V_1^{\mathsf{T}}U_1$ is large.

5 Conclusions

In this paper we have considered the behavior of the GMRES method for solving a linear system $A\boldsymbol{x} = \boldsymbol{b}$, where A is singular. We have discussed two classes of singular matrices (EP and GP) satisfying the conditions under which GMRES converges to a least squares solution and to the group inverse solution, respectively. We have distinguished between the consistent and inconsistent cases and showed that the conditioning of the extended Hessenberg least squares problem can significantly affect the accuracy of approximate solutions computed by GMRES in finite precision arithmetic.

It appears that the consistent EP case is similar to the nonsingular case. The rank deficiency of the extended Hessenberg least squares problem does not occur and GMRES converges to the accurate approximate solution and terminates with a degeneracy of the Krylov space in the
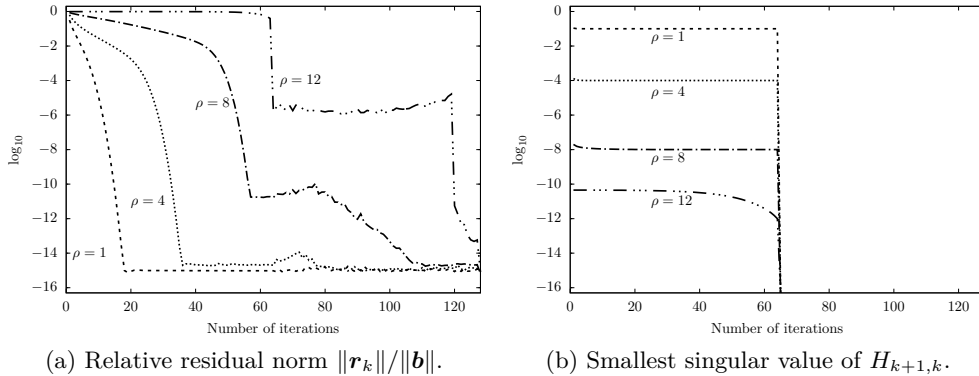

(a) Relative residual norm $\|\boldsymbol{r}_k\|/\|\boldsymbol{b}\|$.   (b) Smallest singular value of $H_{k+1,k}$.

Figure 4.2: GMRES on (4.7) with $\boldsymbol{b} = [\boldsymbol{f}^{\mathsf{T}}, \boldsymbol{0}^{\mathsf{T}}]^{\mathsf{T}}$ and different $\kappa(V_1^{\mathsf{T}}U_1) \simeq 10^\rho/\sqrt{2}$.

(a) Relative residual norm $\|\boldsymbol{r}_k^{\mathrm{R}}\|/\|\boldsymbol{b}\|$.   (b) Smallest singular value of $H^{\mathrm{R}}_{k+1,k}$.
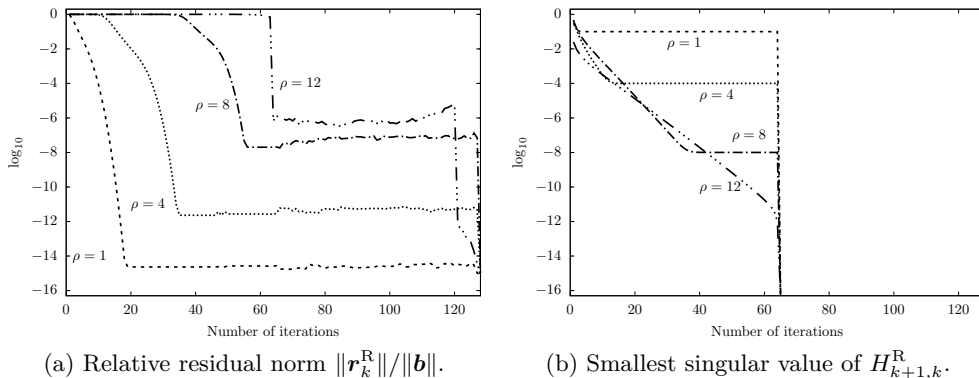
Figure 4.3: RR-GMRES on (4.7) with $\boldsymbol{b} = [\boldsymbol{f}^{\mathsf{T}}, \boldsymbol{0}^{\mathsf{T}}]^{\mathsf{T}}$ and different $\kappa(V_1^{\mathsf{T}}U_1) \simeq 10^\rho/\sqrt{2}$.

next step. If the coefficient matrix is EP, but system $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, then despite the theoretical guarantee for convergence to the least squares solution, the extended Hessenberg least squares problem becomes seriously ill-conditioned and this may lead to very inaccurate approximate solutions computed by GMRES. This happens when the distance of the initial residual to the nullspace is too small or when the residual vector converges gradually to the least squares residual. For such cases, a remedy is to use RR-GMRES.

It is known that if the coefficient matrix is GP and the system is consistent, then theoretically GMRES converges to the solution given by the group inverse of A. We have shown, however, that the extended Hessenberg least squares problem can be ill-conditioned even in the consistent case. Indeed, the conditioning of the extended Hessenberg matrix $H_{k+1,k}$ in GMRES depends not only on the conditioning of the coefficient matrix A (as it is in the case of consistent EP problems) but also on the smallest principal angle between the spaces $\mathcal{R}(A^\mathsf{T})$ and $\mathcal{R}(A)$ that can be quite large. In such cases, both GMRES and RR-GMRES may compute inaccurate approximate solutions.

We believe that under conditions guaranteeing the convergence of GMRES to the generalized least squares solution considered in [31], our results can be extended to singular systems with a general index(A). Note also that in this paper we assume only exact arithmetic and our considerations form a groundwork for future work on rounding error analysis.

We would like to point out that here we have considered the behavior of GMRES applied to exactly singular problems. The extension of our results to GMRES applied to almost singular (or numerically singular) linear systems is far from straightforward as it was also illustrated by our small examples.

Acknowledgments

References

[1] O. Awono and J. Tagoudjeu, *A preconditioned minimal residual solver for a class of linear operator equations*, Comput. Methods Appl. Math., 10 (2010), pp. 119–136.

[2] O. M. Baksalary and G. Trenkler, *Characterizations of EP, normal, and Hermitian matrices*, Linear and Multilinear Algebra, 56 (2008), pp. 299–304.

[3] ——, *On disjoint range matrices*, Linear Algebra Appl., 435 (2011), pp. 1222–1240.

[4] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Springer-Verlag, New York, 2nd ed., 2003.

[5] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[6] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.

[7] D. Calvetti, B. Lewis, and L. Reichel, *GMRES-type methods for inconsistent systems*, Linear Algebra Appl., 316 (2000), pp. 157–169.

[8] ——, *GMRES, L-curves, and discrete ill-posed problems*, BIT, 42 (2002), pp. 44–65.

[9] ——, *On the regularizing properties of the GMRES method*, Numer. Math., 91 (2002), pp. 605–625.

[10] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, SIAM, Philadelphia, 2009.

[11] Z.-H. Cao and M. Wang, *A note on Krylov subspace methods for singular systems*, Linear Algebra Appl., 350 (2002), pp. 285–288.

[12] F. Chatelin, *Eigenvalues of Matrices*, SIAM, Philadelphia, Revised ed., 2012.

[13] X. Du and D. B. Szyld, *Inexact GMRES for singular linear systems*, BIT, 48 (2008), pp. 511–531.

[14] L. Eldén and V. Simoncini, *Solving ill-posed linear systems with GMRES and a singular preconditioner*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1369–1394.

[15] M. G. Gasparo, A. Papini, and A. Pasquali, *Some properties of GMRES in Hilbert spaces*, Numer. Funct. Anal. Optim., 29 (2008), pp. 1276–1285.

[16] I. C. F. Ipsen and C. D. Meyer, *The idea behind Krylov methods*, Amer. Math. Monthly, 105 (1998), pp. 889–899.

[17] Z. Liao, K. Hayami, and J.-F. Yin, *A stabilized GMRES method for solving inconsistent underdetermined least squares problems*, in 22nd Meeting of the Japan SIAM Special Interest on Algorithm for Matrix/Eigenvalue Problems and their Applications, 2016.

[18] K. Morikuni and K. Hayami, *Convergence of inner-iteration GMRES methods for rank-deficient least squares problems*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 225–250.

[19] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[20] L. Reichel and Q. Ye, *Breakdown-free GMRES for singular systems*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 1001–1021.

[21] P. Robert, *On the group-inverse of a linear transformation*, J. Math. Anal. Appl., 22 (1968), pp. 658–669.

[22] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[23] O. Schneider, *Krylov Subspace Methods and their Generalizations for Solving Singular Linear Operator Equations with Applications to Continuous Time Markov Chains*, PhD thesis, Technischen Universität Bergakademie Freiberg, 2005.

[24] A. Sidi, *A unified approach to Krylov subspace methods for the Drazin-inverse solution of singular nonsymmetric linear systems*, Linear Algebra Appl., 298 (1999), pp. 99–113.

[25] ——, *DGMRES: A GMRES-type algorithm for Drazin-inverse solution of singular nonsymmetric linear systems*, Linear Algebra Appl., 335 (2001), pp. 189–204.

[26] L. Smoch, *Spectral behaviour of GMRES applied to singular systems*, Adv. Comput. Math., 27 (2007), pp. 151–166.

[27] Z. Strakoš, *On the real convergence rate of the conjugate gradient method*, Linear Algebra Appl., 154-156 (1991), pp. 535–549.

[28] H. F. Walker, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 152–163.

[29] H. F. WALKER AND L. ZHOU, *A simpler GMRES*, Numer. Linear Algebra Appl., 1 (1994), pp. 571–581.

[30] Y.-M. WEI, *Perturbation analysis of singular linear systems with index one*, Int. J. Comput. Math., 74 (2000), pp. 483–491.

[31] Y.-M. WEI AND H. WU, *Convergence properties of Krylov subspace methods for singular linear systems with arbitrary index*, J. Comput. Appl. Math., 114 (2000), pp. 305–318.

[32] N.-M. ZHANG, *A note on preconditioned GMRES for solving singular linear systems*, BIT, 50 (2010), pp. 207–220.

[33] N.-M. ZHANG AND Y.-M. WEI, *Solving EP singular linear systems*, Int. J. Comput. Math., 81 (2004), pp. 1395–1405.