Worst-case and ideal GMRES for a Jordan block [⋆]

Petr Tichý [1] and Jörg Liesen [1]

Abstract

We investigate the convergence of GMRES for an n by n Jordan block J. For each k that divides n we derive the exact form of the kth ideal GMRES polynomial and prove the equality

$$\max_{\|v\|=1} \min_{p \in \pi_k} \|p(J)v\| = \min_{p \in \pi_k} \max_{\|v\|=1} \|p(J)v\|,$$

where π_k denotes the set of polynomials of degree at most k and with value one at the origin, and $\|\cdot\|$ denotes the Euclidean norm. In other words, we show that for a Jordan block worst-case GMRES and ideal GMRES in these steps are the same. Moreover, we derive lower and upper bounds on the norm of the kth ideal GMRES polynomial in these steps. For the Jordan block with eigenvalue one, we present an explicit formula for its singular value decomposition and use it to improve the bound on the ideal GMRES residual norm in the considered steps k.

Key words: Krylov subspace methods, convergence behavior, ideal GMRES, worst-case GMRES, Jordan block, SVD.

1 Introduction

The GMRES method of Saad and Schultz [9] is a very popular iterative method for solving systems of linear algebraic equations $Ax = b$. Its convergence behavior has been investigated for many years, but a complete understanding

⋆ This work was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft. The work of the first author was also supported by GA AS CR under grant No. KJB1030306.

Email addresses: `tichy@math.tu-berlin.de` (Petr Tichý), `liesen@math.tu-berlin.de` (Jörg Liesen).

[1] Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany.

still remains elusive. A natural approach, that is common throughout numerical analysis, is to study the algorithm's *worst-case behavior*. Here for each iteration step k one needs to analyze the quantity

$$\max_{\|r_0\|=1} \min_{p \in \pi_k} \|p(A)r_0\| \,, \tag{1}$$

where π_k denotes the set of polynomials of degree at most k and with value one at the origin, $r_0 \equiv b - Ax_0$ is the initial residual, and $\|\cdot\|$ denotes the Euclidean norm. Changing the order of maximization and minimization in (1) leads to the *ideal GMRES* approximation problem [5],

$$\min_{p \in \pi_k} \|p(A)\| \,. \tag{2}$$

Because of the submultiplicativity property of the Euclidean norm, (2) is an upper bound for (1),

$$\max_{\|r_0\|=1} \min_{p \in \pi_k} \|p(A)r_0\| \le \min_{p \in \pi_k} \|p(A)\| \,. \tag{3}$$

To better understand the convergence behavior of GMRES, several researchers tried to identify cases in which (3) is an equality. The best known result of this type is that (3) is an equality whenever A is normal [4,7]. In addition, (3) is an equality for arbitrary A and $k = 1$ [4,7], for triangular Toeplitz matrices when the right hand side of (3) equals one [2], and also when the matrix $p_*(A)$ that solves the ideal GMRES approximation problem (2) has a simple maximal singular value [4, Lemma 2.4]. On the other hand, some examples of nonnormal matrices have been constructed, for which (3) is a sharp inequality [2,11]. Despite the existence of these counterexamples, it is still an open question whether (3) is an equality (or at least tight inequality) for larger classes of nonnormal matrices.

In this paper we investigate the ideal GMRES approximation problem and the inequality (3) for a general Jordan block, the prototype of a nonnormal matrix. Understanding of this case appears to be a prerequisite for the analysis of other classes of nonnormal matrices, particularly the general triangular Toeplitz matrices. We show that (3) is an equality whenever k divides n. For such k we also derive the exact form of the ideal GMRES polynomial, and bounds on the actual value of (2) (in this case coinciding with (1)). For the special case of a Jordan block with eigenvalue one we develop a very tight bound on (2) in the considered steps k.

Apart from providing general understanding of the GMRES convergence behavior, the investigation of the behavior for a (single) Jordan block can be interesting in practical applications. For example, discretized convection-diffusion operators can often be considered the direct sums of several nonsymmetric Toeplitz matrices, each of which close to a Jordan block. In such cases the initial phase of convergence of GMRES for the direct sum of Jordan blocks can be analyzed using knowledge about the GMRES convergence for a single Jordan block (see e.g. [8] for details).

2 GMRES, worst-case GMRES and ideal GMRES

Let a linear system

$$Ax \;=\; b\,, \tag{4}$$

with a *nonsingular* matrix $A \in \mathbb{R}^{n\times n}$ and a right hand side vector $b \in \mathbb{R}^n$ be given. Suppose that we solve (4) with GMRES [9]. Starting from an initial guess x_0, this method computes the initial residual $r_0 \equiv b - Ax_0$ and a sequence of iterates $x_1, x_2, \ldots$, so that the kth residual $r_k \equiv b - Ax_k$ satisfies

$$\|r_k\| \;=\; \|p_k(A)\,r_0\| \;=\; \min_{p\in\pi_k} \|p(A)\,r_0\|\,. \tag{5}$$

The residual r_k is uniquely determined by the minimization condition (5) and satisfies the equivalent orthogonality condition

$$r_k \in r_0 + A\mathcal{K}_k(A, r_0)\,, \qquad r_k \perp A\mathcal{K}_k(A, r_0)\,. \tag{6}$$

Here $\mathcal{K}_k(A, r_0) = \mathrm{span}\{r_0, Ar_0, \ldots A^{k-1}r_0\}$ is the kth Krylov subspace generated by A and r_0, and $\perp$ means orthogonality with respect to the Euclidean inner product.

In this paper we are interested in the worst-case and the ideal GMRES behavior, cf. (1) and (2), respectively. The corresponding polynomials are formally defined in the following definition.

Definition 1 *Suppose that k and n are integers with $k < n$. Then the polynomial $\psi_k \in \pi_k$ is called the kth worst-case GMRES polynomial of $A \in \mathbb{R}^{n\times n}$, if there exists a vector $r_0^{(k)} \in \mathbb{R}^n$, $\|r_0^{(k)}\| = 1$, such that*

$$\|\psi_k(A) r_0^{(k)}\| = \min_{p\in\pi_k} \|p(A) r_0^{(k)}\| = \max_{\|v\|=1} \min_{p\in\pi_k} \|p(A)v\|.$$

Furthermore, the polynomial $\varphi_k \in \pi_k$ is called the kth ideal GMRES polyno-

mial of $A \in \mathbb{R}^{n\times n}$, if it satisfies

$$\|\varphi_k(A)\| = \min_{p\in\pi_k} \|p(A)\| = \min_{p\in\pi_k} \max_{\|v\|=1} \|p(A)v\| .$$

In case A is nonsingular, and $\|\varphi_k(A)\| \neq 0$, the ideal GMRES polynomial φ_k is uniquely determined [5, Theorem 2]. Furthermore, it is easy to see that the worst-case GMRES polynomial ψ_k also always exists. However, when $\|\psi_k(A)r_0^{(k)}\| < \|\varphi_k(A)\|$, it is unclear whether ψ_k is unique.

The following useful result follows from [4, Lemma 2.4].

Lemma 2 *If the ideal GMRES matrix $\varphi_k(A)$ of $A \in \mathbb{R}^{n\times n}$ has a simple maximal singular value, then $\varphi_k = \psi_k$, and $r_0^{(k)}$ is the right singular vector of $\varphi_k(A)$ corresponding to its maximal singular value.*

Furthermore, the following statements are equivalent:

1. *Equality holds in* (3).
2. *There exists a unit norm vector r_0 and a polynomial $\psi \in \pi_k$, such that ψ is the kth GMRES polynomial for A and r_0, and r_0 lies in the span of right singular vectors of $\psi(A)$ corresponding to its maximal singular value.*

In addition, if 2. holds, then $\psi = \psi_k = \varphi_k$.

The following result about ideal and worst-case GMRES for triangular Toeplitz matrices was derived in [2].

Lemma 3 *If $A \in \mathbb{R}^{n\times n}$ is an n by n triangular Toeplitz matrix, then*

$$\|\varphi_k(A)\| = 1 \quad \textit{if and only if} \quad \|\psi_k(A)r_0^{(k)}\| = 1. \tag{7}$$

In other words, ideal GMRES for A stagnates (up to step k) if and only if worst-case GMRES for A stagnates (up to step k), and in this case (3) is an equality.

While the ideal and worst-case GMRES residual norms for triangular Toeplitz matrices are the same in case of stagnation, it is in general unknown if they agree when ideal GMRES converges. In fact, this open question posed in [2, p. 722] to some extend motivated our work presented in this paper.

3 Considerations about J_λ and technical lemmas

We are concerned with the convergence behavior of ideal and worst-case GMRES for a real $n \times n$ Jordan block,

$$J_\lambda = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \equiv \lambda I_n + E_n \,. \tag{8}$$

If $\lambda = 0$, the polynomial numerical hull of J_λ of each degree contains the origin, which implies that ideal GMRES completely stagnates [1–3]. Hence of interest in our context is only the *nonsingular case*, i.e. $\lambda \neq 0$. Moreover, we may assume without loss of generality that $\lambda > 0$, since all results for this case can be easily transferred to the case $\lambda < 0$ using the transformation

$$J_\lambda = -I_n^\pm J_{-\lambda} I_n^\pm \,, \tag{9}$$

where

$$I_n^\pm \equiv \operatorname{diag}(1, -1, 1, -1, \ldots, (-1)^{n-1}) \;\in\; \mathbb{R}^{n\times n} \,. \tag{10}$$

Below we will make use of the *reverse identity* matrix defined by

$$I_n^R \equiv \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{n\times n} \,. \tag{11}$$

When it is clear from the context, we will skip the index n that denotes the matrix size.

Lemma 4 *Suppose that T is any real Toeplitz matrix, and that $T = USV^T$ is its singular value decomposition. Then $V = I^R U \hat{I}^\pm$, where $\hat{I}^\pm$ denotes a diagonal matrix having entries 1 or -1 on its diagonal (not necessarily alternating).*

PROOF. Since each Toeplitz matrix T is symmetric with respect to its antidiagonal, it is easy to see that the matrix TI^R is symmetric. Hence there exists a unitary matrix U and a real diagonal matrix D, such that $TI^R = UDU^T$.

Now choose the matrix $\hat{I}^{\pm}$ such that the diagonal entries of $D\hat{I}^{\pm}$ are all positive. Since $(\hat{I}^{\pm})^2 = I$, we see that

$$T = U\,(D\hat{I}^{\pm})\,(\hat{I}^{\pm}U^T I^R) = U\,(D\hat{I}^{\pm})\,(I^R U\hat{I}^{\pm})^T\,,$$

which shows the assertion. □

Lemma 5 *Consider the real m by m Jordan block $H = \alpha I_m + \beta E_m$, and let*

$$H = USV^T \tag{12}$$

be its singular value decomposition. Then for all integers n and k with $m = n/k$, the singular value decomposition of the n by n matrix $B = \alpha I_n + \beta E_n^k$ is given by

$$B = (U \otimes I_k)(S \otimes I_k)(V \otimes I_k)^T\,. \tag{13}$$

PROOF. Define the n by n matrix P by

$$P \equiv [I_m \otimes e_1, \ldots, I_m \otimes e_k]\,.$$

This matrix is the symmetric permutation that maps $\mathrm{vec}(X)$ to $\mathrm{vec}(X^T)$ for each n by n matrix X (see [6, Definition 4.2.9] for the definition of the vec operator). Furthermore,

$$P^T BP = I_k \otimes H = I_k \otimes (USV^T) = (I_k \otimes U)(I_k \otimes S)(I_k \otimes V)^T\,,$$

and hence

$$\begin{aligned} B &= P(I_k \otimes U)(I_k \otimes S)(I_k \otimes V)^T P^T \\ &= [P(I_k \otimes U)P^T]\,[P(I_k \otimes S)P^T]\,[P(I_k \otimes V)P^T]^T \\ &= (U \otimes I_k)(S \otimes I_k)(V \otimes I_k)^T\,. \end{aligned}$$

In the last equation we have used [6, Corollary 4.3.10]. □

4 Main result

To formulate our main result, we need to define some quantities related to the step $k = 1$ of ideal GMRES for a Jordan block.

Definition 6 *If $J_\omega = \omega I_m + E_m$ is a real $m \times m$ Jordan block, we define*

$$\mu_\omega^{(m)} \equiv \arg\min_{\mu \in \mathbb{R}} \|I_m - \mu J_\omega\|\,, \qquad \varrho_\omega^{(m)} \equiv \|I_m - \mu_\omega^{(m)} J_\omega\|\,. \tag{14}$$

By definition, the polynomial $1 - \mu_\omega^{(m)} z$ is the first ideal GMRES polynomial of J_ω. Since for any matrix A and $k = 1$ equality holds in (3) (see [4,7]), $1 - \mu_\omega^{(m)} z$ is also the first worst-case GMRES polynomial of J_ω. Moreover, according to Lemma 2, the corresponding worst-case initial residual $r_0^{(1)}$ lies in the space of right singular vectors of the matrix $\varphi_1(J_\omega)$ corresponding to its maximal singular value. Since φ_1 is the GMRES polynomial for the matrix J_ω and the initial residual $r_0^{(1)}$, (6) implies that

$$\varphi_1(J_\omega) r_0^{(1)} \perp J_\omega r_0^{(1)}\,. \tag{15}$$

Next we state and prove the main theorem of this paper.

Theorem 7 *Let k, m, n be integers such that $m = n/k > 1$. Then*

$$\varphi(z) \equiv (1 - \mu_{\lambda^k}^{(m)} \lambda^k) + \mu_{\lambda^k}^{(m)} (\lambda - z)^k \tag{16}$$

is the kth ideal GMRES polynomial φ_k and also the kth worst-case GMRES polynomial ψ_k of the n by n Jordan block J_λ in (8). Moreover,

$$\varrho_{\lambda^k}^{(m)} = \|\varphi(J_\lambda)\| = \min_{p \in \pi_k} \max_{\|v\|=1} \|p(J_\lambda) v\| = \max_{\|v\|=1} \min_{p \in \pi_k} \|p(J_\lambda) v\|. \tag{17}$$

Furthermore, if $\lambda^k > \cos\left(\frac{\pi}{m+1}\right)$, then

$$\lambda^{-k} \cos\left(\tfrac{\pi}{m+1}\right) \;\le\; \|\varphi_k(J_\lambda)\| \;\le\; \lambda^{-k}, \tag{18}$$

and if $0 \le \lambda^k \le \cos\left(\frac{\pi}{m+1}\right)$, then $\|\varphi_k(J_\lambda)\| = 1$. Note that for $0 \le \lambda \le 1$ the upper bound in (18) *can be replaced by* 1.

PROOF. First note that by assuming $m > 1$ we exclude the trivial case of $n = k$ in which both (1) and (2) are equal to zero for any n by n matrix A.

We will prove (16)–(17) by constructing an explicit worst-case initial residual $r_0^{(k)}$ for the matrix J_λ, and by showing that φ is indeed the kth GMRES polynomial for J_λ and the constructed vector $r_0^{(k)}$.

Using the notation $\alpha \equiv 1 - \mu_{\lambda^k}^{(m)}\lambda^k$ and $\beta = (-1)^k \mu_{\lambda^k}^{(m)}$, we receive $\varphi(J_\lambda) = \alpha I_n + \beta E_n^k$. Furthermore, define the m by m Jordan block $H \equiv \alpha I_m + \beta E_m$. Then an easy computation shows that

$$H = \begin{cases} I_m - \mu_{\lambda^k}^{(m)} J_{\lambda^k}\,, & \text{if } k \text{ is odd,} \\ I_m^{\pm}\left(I_m - \mu_{\lambda^k}^{(m)} J_{\lambda^k}\right) I_m^{\pm}\,, & \text{if } k \text{ is even.} \end{cases} \tag{19}$$

By definition, $1 - \mu_{\lambda^k}^{(m)} z$ is the first ideal GMRES polynomial of the m by m Jordan block $J_{\lambda^k} = \lambda^k I_m + E_m$. As discussed after Definition 6, it is also the first worst-case GMRES polynomial of J_{λ^k}, i.e. there exists a right singular vector w of $I_m - \mu_{\lambda^k}^{(m)} J_{\lambda^k}$ corresponding to its maximal singular value such that

$$(I_m - \mu_{\lambda^k}^{(m)} J_{\lambda^k})w \perp J_{\lambda^k} w. \tag{20}$$

Define the vector v by $v \equiv w$ if k is odd and $v \equiv I_m^{\pm} w$ if k is even. Then it is easy to check that v is a right singular vector of H corresponding to its maximal singular value. Next, from (19) and (20) we see that

$$Hv \perp \begin{cases} J_{\lambda^k} v\,, & \text{if } k \text{ is odd,} \\ (I_m^{\pm} J_{\lambda^k} I_m^{\pm})\, v\,, & \text{if } k \text{ is even.} \end{cases} \tag{21}$$

Moreover, Lemma 5 and Lemma 4 imply that

$$\|H\| = \varrho_{\lambda^k}^{(m)} = \|\varphi(J_\lambda)\| \quad \text{and} \quad Hv = \pm \varrho_{\lambda^k}^{(m)} I_m^R v\,. \tag{22}$$

Lemma 5 shows that since v is a right singular vector of H corresponding to $\varrho_{\lambda^k}^{(m)}$, $v \otimes e_i$ for $i = 1, \ldots, k$ are right singular vectors of $\varphi(J_\lambda)$ corresponding to $\varrho_{\lambda^k}^{(m)}$. Define the vector

$$r_0 \equiv \gamma \sum_{i=1}^{k} (-\lambda)^{i-1}\, v \otimes e_i = \gamma\,(v \otimes e)\,, \tag{23}$$

where γ is chosen so that $\|r_0\| = 1$, and $e \equiv [1, -\lambda, \ldots, (-\lambda)^{k-1}]^T$. Clearly, r_0 is a right singular vector of the Toeplitz matrix $\varphi(J_\lambda)$ corresponding to $\varrho_{\lambda^k}^{(m)}$, so that Lemma 4 and (22) imply that

$$\begin{aligned} \varphi(J_\lambda) r_0 &= \gamma \varphi(J_\lambda)\,(v \otimes e) \\ &= \pm \gamma \varrho_{\lambda^k}^{(m)} I_n^R\,(v \otimes e) \\ &= \pm \gamma \varrho_{\lambda^k}^{(m)}\,((I_m^R v) \otimes (I_m^R e)) \\ &= \pm \gamma\,((Hv) \otimes (I_m^R e))\,. \end{aligned} \tag{24}$$

We next show that

$$\varphi(J_\lambda)r_0 \;\perp\; J_\lambda^i r_0\,, \quad i=1,\ldots,k\,, \tag{25}$$

i.e. that φ is a GMRES polynomial for the matrix J_λ and the initial residual r_0. It is easy to see that (25) holds if and only if

$$\varphi(J_\lambda)r_0 \;\perp\; E_n^i J_\lambda r_0\,, \quad i=0,\ldots,k-1\,. \tag{26}$$

An elementary computation shows that, for $i=0,\ldots,k-1$,

$$E_n^i J_\lambda r_0 = \begin{cases} \gamma\left((J_{\lambda^k}v)\otimes e_{k-i}\right), & \text{if } k \text{ is odd,} \\ \gamma\left((J_{-\lambda^k}v)\otimes e_{k-i}\right), & \text{if } k \text{ is even.} \end{cases} \tag{27}$$

Using (24) and (27) we obtain

$$(\varphi(J_\lambda)r_0)^T E_n^i J_\lambda r_0 = \pm\,\gamma^2\,[(Hv)^T J_{\lambda^k} v]\,[e^T I_m^R e_{k-i}]$$

for k odd and

$$\begin{aligned}(\varphi(J_\lambda)r_0)^T E_n^i J_\lambda r_0 &= \pm\,\gamma^2\,[(Hv)^T J_{-\lambda^k} v]\,[e^T I_m^R e_{k-i}] \\ &= \mp\,\gamma^2\,[(Hv)^T (I_m^\pm J_{\lambda^k} I_m^\pm) v]\,[e^T I_m^R e_{k-i}]\end{aligned}$$

for k even. Finally, using (21) it is clear that (25) holds.

Summarizing, $\varphi(z)$ is the kth GMRES polynomial for the matrix J_λ and the initial residual r_0 that lies in the span of right singular vectors corresponding to the maximal singular value of $\varphi(J_\lambda)$. Then, according to Lemma 2, (17) holds.

The upper bound in (18) has been proven in [3]. The lower bound follows from results of [3] as well: Note that $\varrho_{\lambda^k}^{(m)}$ is the norm of the first ideal GMRES polynomial of the Jordan block $J_{\lambda^k} = \lambda^k I_m + E_m$. If λ^k is greater than $R_{1,m}$, the radius of the first degree polynomial numerical hull of J_{λ^k}, then

$$\min_{p\in\pi_1} \|p(J_{\lambda^k})\| \geq \frac{R_{1,m}}{\lambda^k}.$$

According to [3, Theorem 3.1], $R_{1,m} = \cos(\frac{\pi}{m+1})$. On the other hand, if $\lambda^k \leq R_{1,m}$, then the first degree polynomial numerical hull of J_{λ^k} contains zero and therefore $\varrho_{\lambda^k}^{(m)} = 1$, see e.g. [3] or [2, Theorem 2.8]. □

Remark 8 *Let n be even, $k = n/2$, and let $\lambda^k \geq R_{1,2} = \frac{1}{2}$. Based on Definition 6 and Theorem 7 it is possible to show that*

$$\mu_{\lambda^k}^{(2)} = \frac{1}{\lambda^k}\frac{4\lambda^{2k}-1}{4\lambda^{2k}+1}, \qquad \varrho_{\lambda^k}^{(2)} = \|\varphi_k(J_\lambda)\| = \frac{4\lambda^k}{4\lambda^{2k}+1}.$$

For $0 \leq \lambda^k \leq R_{1,2}$, it holds $\mu_{\lambda^k}^{(2)} = 0$ and $\varrho_{\lambda^k}^{(2)} = 1$, cf. [3].

Remark 9 *From (16) it is easy to see, that in case $\mu_{\lambda^k}^{(m)} \neq 0$, the k roots of φ_k are uniformly distributed on the circle around λ with radius $|(\mu_{\lambda^k}^{(m)})^{-1} - \lambda^k|^{1/k}$. If $\mu_{\lambda^k}^{(m)} = 0$, then $\varphi_k(z) = 1$.*

5 Observations for a general step k

In Section 4 we described the ideal GMRES behavior for a Jordan block J_λ and showed that equality holds in (3) when k divides n. We observed numerically that equality in fact holds for every k, but we were unable to prove such result.

We next describe observations from our numerical experiments (performed using the semidefinite programming package SDPT3 [10]) that might give some ideas of how to approach this open problem. Denote by d the greatest common divisor of n and k, $n_d = n/d$, $k_d = k/d$. Then the ideal GMRES matrix $\varphi_k(J_\lambda)$ has again a special structure determined by powers of the matrix E_n,

$$\varphi_k(J_\lambda) = c_0 I_n + \sum_{j=1}^{k_d} c_j (-E_n)^{j\,d}, \qquad \sum_{i=0}^{k_d} c_i \lambda^{i\,d} = 1,$$

and the polynomial φ_k has the form

$$\varphi_k(z) = c_0 + \sum_{j=1}^{k_d} c_j (\lambda - z)^{j\,d}.$$

In our experiments we always observed that $c_i > 0$, $i = 0, \ldots, k_d$, whenever $\|\varphi_k(J_\lambda)\| < 1$. Using the permutation matrix

$$P \equiv \left[I_{n_d} \otimes e_1, \ldots, I_{n_d} \otimes e_d\right],$$

it is possible to transform the matrix $\varphi_k(J_\lambda)$ to a block diagonal matrix with d identical diagonal blocks H of size n_d by n_d. The matrix H is a banded upper triangular Toeplitz matrix with c_0 on the diagonal and $(-1)^{i\,d} c_i$, $i = 1, \ldots, k_d$

on its superdiagonals. Similar as in (16), the coefficients $c_0, \dots, c_{k_d}$ minimize the norm of the k_dth ideal GMRES polynomial of the form

$$c_0 + \sum_{j=1}^{k_d} c_j (\lambda^d - z)^j, \qquad \sum_{i=0}^{k_d} c_i \lambda^{i\,d} = 1,$$

for the n_d by n_d Jordan block $J_{\lambda^d} = \lambda^d I_{n_d} + E_{n_d}$ (note that the greatest common divisor of n_d and k_d is 1). If one would be able to prove that indeed $c_i > 0$ for $i = 0, \dots, k_d$, then H would be known to be irreducible and thus would be known to have a simple maximal singular value. From this result the proof of Theorem 7 can be generalized to hold for each step k.

6 The special case of the eigenvalue one

For the Jordan block J_λ with the eigenvalue $\lambda = 1$, we are able to say more about the norm of the ideal GMRES polynomial in the considered steps $k = n/m$. Based on the singular value decomposition (SVD) of J_1 and on the eigenvalue interlacing property we can improve the bound (18). We first derive the SVD of J_1.

Theorem 10 *The SVD of J_1 is given by $J_1 = USV^T$, where*

$$V = \{v_{ij}\}_{i,j=1}^n, \qquad v_{ij} = \frac{2}{\sqrt{2n+1}} \sin\left(\frac{2i-1}{2n+1} j\pi\right), \tag{28}$$

$$U = \{u_{ij}\}_{i,j=1}^n, \qquad u_{ij} = \frac{2}{\sqrt{2n+1}} \sin\left(\frac{2i}{2n+1} j\pi\right), \tag{29}$$

and

$$S = \mathrm{diag}(\sigma_i), \qquad \sigma_i = 2\cos\left(\frac{i\pi}{2n+1}\right), \qquad i = 1, \dots, n. \tag{30}$$

Moreover, the matrices U and V satisfy

$$U = I_n^R V I_n^{\pm}. \tag{31}$$

PROOF. It is easy to check that the matrices U and V defined in (28) and (29) are orthogonal. The relation (31) follows from

$$(I_n^R U)_{ij} = \sin\left(\frac{2(n-i+1)}{2n+1} j\pi\right) = \sin\left(j\pi - \frac{2i-1}{2n+1} j\pi\right)$$

$$= (-1)^{j-1} \sin\left(\frac{2i-1}{2n+1} j\pi\right) = (VI_n^{\pm})_{ij}.$$

We will prove that $S = U^T J_1 V$ for U, S and V defined in (29), (30) and (28). Denote the elements of $J_1 V$ by f_{ij}, $i, j = 1, \ldots, n$. Then it holds for $i = 1, \ldots, n-1$ and $j = 1, \ldots, n$,

$$\begin{aligned} f_{ij} = v_{i,j} + v_{i+1,j} &= \frac{2}{\sqrt{2n+1}} \left[\sin\left(\frac{2i-1}{2n+1} j\pi\right) + \sin\left(\frac{2i+1}{2n+1} j\pi\right)\right] \\ &= \frac{2}{\sqrt{2n+1}} 2 \sin\left(\frac{2ij\pi}{2n+1}\right) \cos\left(\frac{j\pi}{2n+1}\right) \\ &= 2\cos\left(\frac{j\pi}{2n+1}\right) u_{ij}. \end{aligned} \tag{32}$$

For $i = n$ and $j = 1, \ldots, n$ we obtain

$$f_{nj} = v_{nj} = \frac{2}{\sqrt{2n+1}} \sin\left(\frac{2n-1}{2n+1} j\pi\right). \tag{33}$$

Since

$$\begin{aligned} \sin\left(\frac{2n-1}{2n+1} j\pi\right) &= \sin\left(j\pi - \frac{2j\pi}{2n+1}\right) \\ &= (-1)^{j-1} \sin\left(\frac{2j\pi}{2n+1}\right) \\ &= (-1)^{j-1} 2 \sin\left(\frac{j\pi}{2n+1}\right) \cos\left(\frac{j\pi}{2n+1}\right) \\ &= 2\sin\left(j\pi - \frac{j\pi}{2n+1}\right) \cos\left(\frac{j\pi}{2n+1}\right) \\ &= 2\sin\left(\frac{2nj\pi}{2n+1}\right) \cos\left(\frac{j\pi}{2n+1}\right), \end{aligned}$$

(33) can be written in the form

$$f_{nj} = 2\cos\left(\frac{j\pi}{2n+1}\right) u_{nj}\,. \tag{34}$$

Summarizing, the formulas (32) and (34) imply that $J_1 V = US$, and hence $J_1 = USV^T$. Since S is a diagonal matrix with positive diagonal elements, $J_1 = USV^T$ is the SVD of J_1. □

From Theorem 7 it is clear, that the numbers $\mu_1^{(m)}$ and $\varrho_1^{(m)}$ play an important role in the kth ideal GMRES step if $m = n/k$. In the following theorem we bound the value $\varrho_1^{(m)}$.

Theorem 11 *Consider the m by m Jordan block J_1. Then*

$$\cos\left(\frac{\pi}{2m}\right) \leq \varrho_1^{(m)} \leq \cos\left(\frac{\pi}{2m+1}\right). \tag{35}$$

PROOF. The upper bound in (35) follows from

$$\begin{aligned}\|\varphi_1(J_1)\| \leq \|I_m - \frac{1}{2}J_1\| &= \frac{1}{2}\|I_m - E_m\| = \frac{1}{2}\|I_m + E_m\| \\ &= \frac{1}{2}\|J_1\| = \frac{1}{2}\sigma_1(J_1) = \cos\left(\frac{\pi}{2m+1}\right),\end{aligned}$$

where $\sigma_1(J_1)$ is known from Theorem 10. For $\mu \in \mathbb{R}$, define the polynomial

$$p_\mu(z) \equiv 1 - \mu z.$$

We will investigate the value of $\|p_\mu(J_1)\|$. The norm of the matrix $p_\mu(J_1)$ is the square root of the maximal eigenvalue of the matrix

$$p_\mu(J_1)^T p_\mu(J_1) = \begin{bmatrix} \gamma_\mu & -\beta_\mu & & \\ -\beta_\mu & \alpha_\mu & \ddots & \\ & \ddots & \ddots & -\beta_\mu \\ & & -\beta_\mu & \alpha_\mu \end{bmatrix},$$

where

$$\alpha_\mu \equiv \mu^2 + (1-\mu)^2, \quad \beta_\mu \equiv (1-\mu)\,\mu, \quad \gamma_\mu \equiv (1-\mu)^2.$$

Next, define the m by m matrix $T_{\mu,m}$,

$$T_{\mu,m} \equiv \begin{bmatrix} \alpha_\mu & -\beta_\mu & & \\ -\beta_\mu & \alpha_\mu & \ddots & \\ & \ddots & \ddots & -\beta_\mu \\ & & -\beta_\mu & \alpha_\mu \end{bmatrix}.$$

Denote by $\chi_{\mu,m}$ and $\tau_{\mu,m}$ the characteristic polynomials of $p_\mu(J_1)^T p_\mu(J_1)$ and $T_{\mu,m}$,

$$\chi_{\mu,m}(\xi) \equiv \det(\xi I_m - p_\mu(J_1)^T p_\mu(J_1)), \quad \tau_{\mu,m}(\xi) \equiv \det(\xi I_m - T_{\mu,n}).$$

It is not hard to see that

$$\chi_{\mu,m}(\xi) = \tau_{\mu,m}(\xi) + \mu^2 \tau_{\mu,m-1}(\xi).$$

Using results of classical polynomial theory, the roots of polynomials $\tau_{\mu,m}$ and $\tau_{\mu,m-1}$ interlace. Therefore, the maximal root of $\chi_{\mu,m}$ must lay between maximal roots of $\tau_{\mu,m}$ and $\tau_{\mu,m-1}$ (between the maximal eigenvalues of $T_{\mu,m}$ and $T_{\mu,m-1}$). The eigenvalues $\lambda^{(j)}_{\mu,m}$, $j = 1, \ldots, m$, of the matrix $T_{\mu,m}$ are given by the formula

$$\begin{aligned}
\lambda^{(j)}_{\mu,m} &= \alpha_\mu - 2\beta_\mu \cos\left(\frac{j\pi}{m+1}\right) \\
&= \mu^2 + (1-\mu)^2 - 2(1-\mu)\mu\cos\left(\frac{j\pi}{m+1}\right) \\
&= 1 - 2(1-\mu)\mu\left[1 + \cos\left(\frac{j\pi}{m+1}\right)\right] \\
&= 1 - 4(1-\mu)\mu\sin^2\left(\frac{j\pi}{2(m+1)}\right) \\
&= 1 - 4\beta_\mu \sin^2\left(\frac{j\pi}{2(m+1)}\right)
\end{aligned}$$

and the maximal root of $\chi_{\mu,m}$ lies in the closed interval

$$\left[1 - 4\beta_\mu \sin^2\left(\frac{\pi}{2m}\right), 1 - 4\beta_\mu \sin^2\left(\frac{\pi}{2(m+1)}\right)\right]. \tag{36}$$

The lower bound (and also the upper bound) of the interval (36) is the smallest one for $\mu = 1/2$ (take derivatives with respect to μ to find the extrema). Since

$$1 - 4\beta_{\frac{1}{2}} \sin^2\left(\frac{\pi}{2m}\right) = 1 - \sin^2\left(\frac{\pi}{2m}\right) = \cos^2\left(\frac{\pi}{2m}\right),$$

it must hold

$$\|\varphi_1(J_1)\| \geq \cos\left(\frac{\pi}{2m}\right).$$

□

Let k, n, m be as in Theorem 7 and let $\lambda = 1$. Then Theorems 7 and 11 imply

$$\cos\left(\frac{\pi}{2m}\right) \leq \|\varphi_k(J_1)\| \leq \cos\left(\frac{\pi}{2m+1}\right). \tag{37}$$

Moreover, our numerical experiments predict that

$$\mu_1^{(m)} = \frac{m+1}{2m+1}. \tag{38}$$

Unfortunately, we were unable to prove (38) theoretically. It is not hard to determine the value $\varrho_1^{(m)}$ exactly some values of m, e.g.,

$$\varrho_1^{(2)} = \frac{4}{5}, \qquad \varrho_1^{(3)} = \frac{2\sqrt{7}+1}{7}.$$

However, for a general value of m, determining of $\varrho_1^{(m)}$ seems to be a nontrivial problem.

7 Conclusions

In this paper we investigate the ideal GMRES approximation problem and the inequality (3) for a general Jordan block. We show that (3) is an equality whenever k divides n. For such k we also derive the exact form of the ideal GMRES polynomial, and bounds on the actual value of (1). Moreover, these bounds are improved for the special case of a Jordan block with eigenvalue one. Our numerical experience indicates that (3) is indeed an equality for each k if A is a Jordan block, but a complete proof of such result remains the subject of further work.

Acknowledgment

The authors wish to thank J. D. Tebbens for his helpful comments and suggestions.

References

[1] V. Faber, A. Greenbaum, and D. E. Marshall, *The polynomial numerical hulls of Jordan blocks and related matrices*, Linear Algebra Appl., 374 (2003),

pp. 231–246.

[2] V. Faber, W. Joubert, E. Knill, and T. Manteuffel, *Minimal residual method stronger than polynomial preconditioning*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 707–729.

[3] A. Greenbaum, *Some theoretical results derived from polynomial numerical hulls of Jordan blocks*, ETNA, (to appear).

[4] A. Greenbaum and L. Gurvits, *Max-min properties of matrix factor norms*, SIAM J. Sci. Comput., 15 (1994), pp. 348–358.

[5] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.

[6] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[7] W. Joubert, *A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput., 15 (1994), pp. 427–439.

[8] J. Liesen and Z. Strakoš, *GMRES convergence analysis for a convection-diffusion model problem*, SIAM J. Sci. Comput., (to appear).

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

[10] K. Toh, M. Todd, and R. Tütüncü, *SDPT3 – a Matlab software package for semidefinite programming, version 2.3. Interior point methods.* The software is available on the site `http://www.math.nus.edu.sg/~mattohkc/sdpt3.html`, June 2001.

[11] K.-C. Toh, *GMRES vs. ideal GMRES*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.