Electronic Transactions on Numerical Analysis.
Volume 60, pp. 421–427, 2024.

ISSN 1068–9613.
DOI: 10.1553/etna_vol60s421

ETNA
Kent State University and
Johann Radon Institute (RICAM)

NEAR-OPTIMAL CONVERGENCE OF THE FULL ORTHOGONALIZATION METHOD*

TYLER CHEN† AND GÉRARD MEURANT‡

Abstract. We establish a near-optimality guarantee for the full orthogonalization method (FOM), showing that the *overall* convergence of FOM is nearly as good as GMRES. In particular, we prove that at every iteration k, there exists an iteration $j \leq k$ for which the FOM residual norm at iteration j is no more than $\sqrt{k+1}$ times larger than the GMRES residual norm at iteration k. This bound is sharp, and it has implications for algorithms for approximating the action of a matrix function on a vector.

Key words. Full Orthogonalization Method, GMRES

AMS subject classifications. 65F10

1. Introduction.

The *full orthogonalization method* (FOM) [20] and the *generalized minimal residual method* (GMRES) [22] are two Krylov subspace methods used for solving a real or complex non-symmetric[1] linear system of equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{1.1}$$

Assuming an initial guess $\mathbf{x}_0 = \mathbf{0}$, both FOM and GMRES produce iterates from the Krylov subspace

$$\mathcal{K}_k(\mathbf{A}, \mathbf{b}) := \operatorname{span}\{\mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{k-1}\mathbf{b}\}$$

but according to slightly different formulas. The FOM iterate is of particular interest because it is closely related to the Arnoldi method for matrix-function approximation for approximating $f(\mathbf{A})\mathbf{b}$, the action of a matrix function on a vector [9, 11, 15]. We expand on this connection in Section 3.3.

Denote by $\mathbf{Q}_k = [\mathbf{q}_1, \dots, \mathbf{q}_k]$ the orthonormal basis for the Krylov subspace $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ produced by the Arnoldi algorithm [2]. Define also the $(k+1) \times k$ upper-Hessenberg matrix $\mathbf{H}_{k+1,k}$ of coefficients produced by the Arnoldi algorithm, and recall the Arnoldi recurrence relation

$$\mathbf{A}\mathbf{Q}_k = \mathbf{Q}_{k+1}\mathbf{H}_{k+1,k}.$$

The FOM and GMRES iterates (with zero initial guesses $\mathbf{x}_0^{\mathrm{F}} = \mathbf{x}_0^{\mathrm{G}} = \mathbf{0}$) are respectively defined as

$$\mathbf{x}_k^{\mathrm{F}} := \|\mathbf{b}\|_2 \mathbf{Q}_k (\mathbf{H}_k)^{-1} \mathbf{e}_1, \qquad \mathbf{x}_k^{\mathrm{G}} := \|\mathbf{b}\|_2 \mathbf{Q}_k (\mathbf{H}_{k+1,k})^{\dagger} \mathbf{e}_1,$$

where $\mathbf{H}_k$ is $\mathbf{H}_{k+1,k}$ with the last row deleted, the symbol $\dagger$ indicates the pseudoinverse, and $\mathbf{e}_1 = [1, 0, \dots, 0]^{\mathsf{T}}$ is the first canonical basis vector [18].

Define the FOM and GMRES residual vectors

$$\mathbf{r}_k^{\mathrm{F}} := \mathbf{b} - \mathbf{A}\mathbf{x}_k^{\mathrm{F}}, \qquad \mathbf{r}_k^{\mathrm{G}} := \mathbf{b} - \mathbf{A}\mathbf{x}_k^{\mathrm{G}}.$$

The norms of these residuals can be used as a measure of how well the iterates $\mathbf{x}_k^{\mathrm{F}}$ and $\mathbf{x}_k^{\mathrm{G}}$ solve the linear system (1.1), and understanding their relationship is the aim of this paper.

*Received July 7, 2024. Accepted July 18, 2024. Published online on August 28, 2024. Recommended by Vladimir Druskin.

†New York University, New York, USA (tyler.chen@nyu.edu).

‡Paris, France (gerard.meurant@gmail.com).

[1] If $\mathbf{A}$ is symmetric, then GMRES is mathematically equivalent to MINRES [19], and if $\mathbf{A}$ is symmetric positive definite, then FOM is mathematically equivalent to the conjugate gradient method [14]. While this is relevant for an efficient implementation, it does not impact the exact arithmetic theory in this note.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

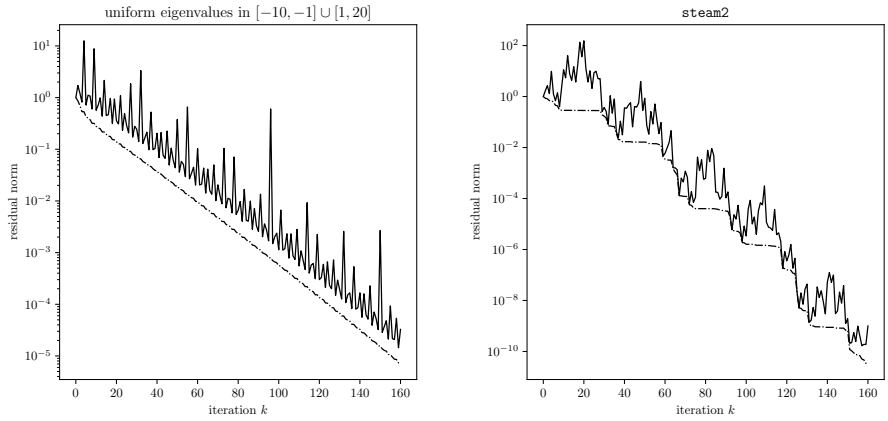

FIG. 1.1. *Residual norms for FOM* $\|\mathbf{r}_k^{\mathrm{F}}\|_2$ *(solid) and GMRES* $\|\mathbf{r}_k^{\mathrm{G}}\|_2$ *(dash-dot) for a symmetric matrix with eigenvalues in* $[-10,-1]\cup[1,20]$ *and the* `steam2` *matrix. All curves are normalized by* $\|\mathbf{b}\|_2$*. While the residual norms for FOM jump up and down, they exhibit a generally downward trend which mirrors the convergence of the GMRES residual norms.*

It is well known that the GMRES iterates satisfy a residual optimality guarantee:

$$\mathbf{x}_k^{\mathrm{G}} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{K}_k(\mathbf{A},\mathbf{b})} \|\mathbf{b}-\mathbf{A}\mathbf{x}\|_2. \tag{1.2}$$

Hence, the GMRES residual norms are non-increasing and are optimal among Krylov subspace methods. This optimality guarantee leads to well-known results on the rate of convergence of the residual norms in terms of quantities such as the condition number of $\mathbf{A}$ [21]. On the other hand, the FOM residual norms often appear oscillatory, with large jumps. In fact, it is easy to construct examples for which $\|\mathbf{r}_k^{\mathrm{F}}\|_2/\|\mathbf{r}_0^{\mathrm{F}}\|_2$ can be arbitrarily large [18]!

Typical examples of the convergence behavior of FOM and GMRES are illustrated in Figure 1.1. Here we consider a symmetric matrix with 500 eigenvalues equally spaced in $[-10,-1]\cup[1,20]$ and the `steam2` matrix from the Matrix Market [3]. In both cases we choose $\mathbf{b}$ as the all ones vector. We observe that while GMRES exhibits well-behaved non-increasing residual norms, the residual norms for FOM are highly oscillatory. In fact, at some iterations, the FOM residual norms are orders of magnitude larger than those of GMRES. Remarkably, however, FOM exhibits an "overall downward trend" mirroring the convergence of GMRES.

2. Convergence bounds. The purpose of this note is to establish the following bound for the FOM residual norms in terms of the optimal GMRES residual norms:

THEOREM 2.1. *For every* $k \geq 1$*,*

$$\min_{0\leq j\leq k} \|\mathbf{r}_j^{\mathrm{F}}\|_2 \leq \sqrt{k+1}\cdot\|\mathbf{r}_k^{\mathrm{G}}\|_2.$$

The theorem asserts that, while the FOM residual norm at any given iteration k can be arbitrarily large, the overall convergence of FOM is at most $\sqrt{k+1}$ times worse than that of GMRES. In light of the optimality of the GMRES iterates (1.2), this implies that the FOM iterates are, in a certain sense, near-optimal. While an immediate consequence of existing work, we have not been able to find this bound in the literature; see Section 3.1 for a discussion on past work. As we discuss in Section 3, Theorem 2.1 has implications for commonly used Krylov subspace methods for matrix functions.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

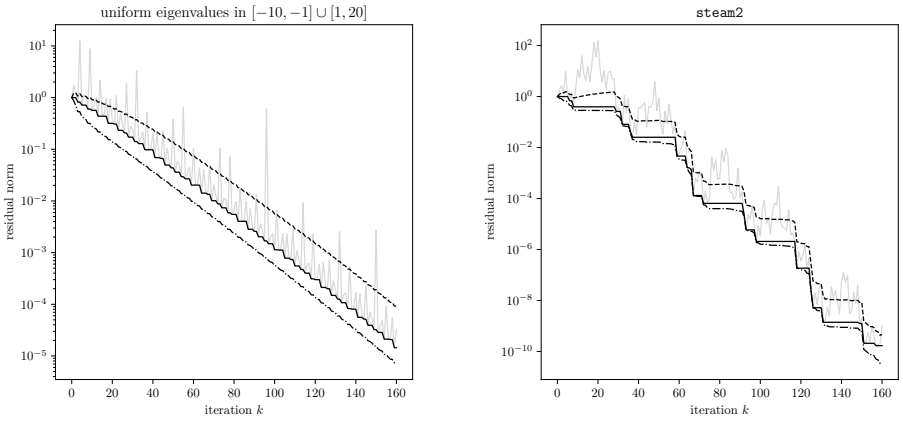

FIG. 2.1. *Best FOM residual so far* $\min_{j\le k}\|\mathbf{r}_j^{\mathrm{F}}\|_2$ *(solid), bound of Theorem 2.1 (dashed), and residual norms for FOM* $\|\mathbf{r}_k^{\mathrm{F}}\|_2$ *(solid grey) and GMRES* $\|\mathbf{r}_k^{\mathrm{G}}\|_2$ *(dash-dot). All curves are normalized by* $\|\mathbf{b}\|_2$*. While FOM exhibits oscillatory convergence, Theorem 2.1 ensures that the best FOM residual norm seen so far matches closely with the GMRES residual norm.*

In many practical situations, the GMRES residual norm is exponential in k, in which case the factor $\sqrt{k+1}$ is comparatively unimportant. In Figure 2.1, we show the examples from Figure 1.1 along with the bound from Theorem 2.1. To emphasize the "overall convergence" of FOM, we also display the smallest residual norm of FOM seen up to a given iteration k: $\min_{j\le k}\|\mathbf{r}_j^{\mathrm{F}}\|_2$. As expected, this quantity tracks very closely the convergence of GMRES.

Our proof makes use of a characterization of the GMRES residual norms in terms of the FOM residual norms. Define $\vartheta_1 = 1$, and for $k \ge 1$ define

$$\vartheta_{k+1} = -\frac{1}{h_{k+1,k}}\sum_{j=1}^{k}\vartheta_j h_{j,k},$$

where $h_{i,j}$ is the (i,j)-entry of $\mathbf{H}_{k+1,k}$. The FOM and GMRES residuals norms are related in the following sense:

THEOREM 2.2 (Theorem 3.12 in [18]). *For every* $k \ge 0$*,*

$$\frac{\|\mathbf{r}_k^{\mathrm{F}}\|_2}{\|\mathbf{r}_0^{\mathrm{F}}\|_2} = \frac{1}{|\vartheta_{k+1}|}, \qquad \frac{\|\mathbf{r}_k^{\mathrm{G}}\|_2}{\|\mathbf{r}_0^{\mathrm{G}}\|_2} = \left(\sum_{j=1}^{k+1}|\vartheta_j|^2\right)^{-1/2}.$$

The proof of Theorem 2.1 is an immediate consequence of Theorem 2.2.

Proof of Theorem 2.1. Note that $\mathbf{H}_k$ is singular if and only if $\vartheta_{k+1} = 0$. In this case, the FOM iterate is undefined, and we can define the FOM residual norm to be infinite. Therefore, defining $1/\infty = 0$, we have that

$$\frac{\|\mathbf{r}_k^{\mathrm{G}}\|_2}{\|\mathbf{r}_0^{\mathrm{G}}\|_2} = \left(\sum_{j=0}^{k}\frac{\|\mathbf{r}_0^{\mathrm{F}}\|_2^2}{\|\mathbf{r}_j^{\mathrm{F}}\|_2^2}\right)^{-1/2}.$$

Canceling $\|\mathbf{r}_0^{\mathbf{F}}\|_2 = \|\mathbf{r}_0^{\mathbf{G}}\|_2$, we therefore find

$$\|\mathbf{r}_k^{\mathrm{G}}\|_2 = \left(\sum_{j=0}^{k}\frac{1}{\|\mathbf{r}_j^{\mathrm{F}}\|_2^2}\right)^{-1/2}. \tag{2.1}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Thus, bounding each term in the sum by the maximum, it follows that

$$\|\mathbf{r}_k^{\mathrm{G}}\|_2 \geq \left((k+1) \max_{0\leq j\leq k} \frac{1}{\|\mathbf{r}_j^{\mathrm{F}}\|_2^2} \right)^{-1/2} = \frac{1}{\sqrt{k+1}} \cdot \min_{0\leq j\leq k} \|\mathbf{r}_j^{\mathrm{F}}\|_2,$$

which proves the result. ∎

One may wonder whether the pre-factor $\sqrt{k+1}$ in Theorem 2.1 is necessary. We show that no better value is possible:

THEOREM 2.3. *For every $k \geq 1$ there exists a matrix $\mathbf{A}$ and a vector $\mathbf{b}$ for which*

$$\min_{j\leq k} \|\mathbf{r}_j^{\mathrm{F}}\|_2 = \sqrt{k+1} \cdot \|\mathbf{r}_k^{\mathrm{G}}\|_2.$$

Proof. It is well-known that any sequence of residual norms is possible for FOM; that is, given any sequence of non-zero positive numbers $f_0, f_1, \ldots, f_k$, there exists a matrix $\mathbf{A}$ and vector $\mathbf{b}$ (of dimension $n = k+1$) for which $\|\mathbf{r}_j^{\mathrm{F}}\|_2 = f_j$ [18, Theorem 3.15]. In particular, there exists a matrix $\mathbf{A}$ and vector $\mathbf{b}$ for which $\|\mathbf{r}_j^{\mathrm{F}}\|_2^2 = 1$, for $j = 0, 1, \ldots, n$. In this case (2.1) implies

$$\|\mathbf{r}_k^{\mathrm{G}}\|_2^2 = \frac{1}{\sum_{j=0}^{k} 1/\|\mathbf{r}_j^{\mathrm{F}}\|_2^2} = \frac{1}{k+1},$$

which, since $\min_{0\leq j\leq k} \|\mathbf{r}_j^{\mathrm{F}}\|_2 = 1$, gives the result. ∎

3. Discussion.

3.1. Previous work. Many works have studied the relation between convergence of FOM and GMRES [4, 7, 12, 16, 17, 6, etc.], particularly with the goal of understanding how to efficiently estimate the residual/error norms in practice.

Perhaps the best-known [4, 7] relation between the FOM and GMRES residual norms is

$$\|\mathbf{r}_k^{\mathrm{F}}\|_2 = \frac{\|\mathbf{r}_k^{\mathrm{G}}\|_2}{\sqrt{1 - \left(\|\mathbf{r}_k^{\mathrm{G}}\|_2 / \|\mathbf{r}_{k-1}^{\mathrm{G}}\|_2\right)^2}}, \tag{3.1}$$

which can be obtained by rearranging (2.1). Informally, (3.1) says that at iterations where GMRES makes good progress (i.e., for which $\|\mathbf{r}_k^{\mathrm{G}}\|_2 \ll \|\mathbf{r}_{k-1}^{\mathrm{G}}\|_2$), the FOM residual norm is close to the GMRES residual norm, and at iterations where GMRES stagnates (i.e., for which $\|\mathbf{r}_k^{\mathrm{G}}\|_2 \approx \|\mathbf{r}_{k-1}^{\mathrm{G}}\|_2$), the FOM residual norm is very large. However, it does not provide any information about at which, if any, iterations GMRES makes good progress. As such, it is unclear from (3.1) that FOM's overall convergence must track that of GMRES.

There are at least two bounds reminiscent of Theorem 2.1 in that they aim to understand the "overall convergence" of FOM rather than the convergence at every iteration [13, 5]. While [13] uses entirely different techniques, [5] uses (3.1) to argue that if GMRES is making good progress, then there must be some iteration for which $\|\mathbf{r}_j^{\mathrm{G}}\|_2$ is substantially smaller than $\|\mathbf{r}_{j-1}^{\mathrm{G}}\|_2$ and hence for which $\|\mathbf{r}_k^{\mathrm{G}}\|_2 \approx \|\mathbf{r}_{k-1}^{\mathrm{G}}\|_2$. Both bounds are weaker than Theorem 2.1 in that they (i) hold only for symmetric matrices and (ii) are not in terms of the GMRES residual norm but rather the best approximation of zero on sets of the form $[a, b] \cup [c, b]$ by a polynomial taking the value one at the origin, where $b < 0 < c$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

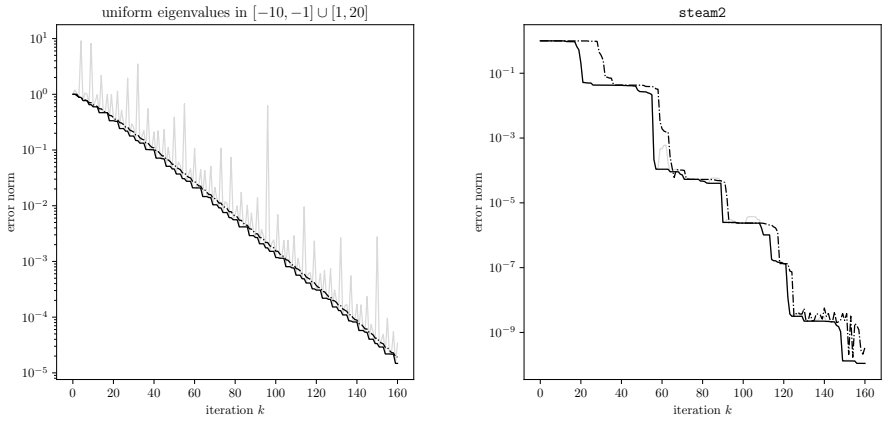

FIG. 3.1. *Best FOM error so far* $\min_{j\le k}\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}_k^{\mathrm{F}}\|_2$ *(solid), error for FOM* $\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}_k^{\mathrm{F}}\|_2$ *(solid grey) and GMRES* $\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}_k^{\mathrm{G}}\|_2$ *(dash-dot). All curves are normalized by* $\|\mathbf{A}^{-1}\mathbf{b}\|_2$*. Note that the FOM convergence is slightly better than GMRES, indicating FOM may be preferable to GMRES in some situations.*

3.2. Error norm. We may wish to obtain a solution with small *error* norm $\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}\|_2$. It is typically not clear whether GMRES or FOM is preferable, and in many cases the error norms of FOM are smaller than those of GMRES. In Figure 3.1 we have illustrated the error norms for the examples from Figure 1.1 and observe that the FOM errors are slightly better.

Using that $\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}=\mathbf{A}^{-1}(\mathbf{b}-\mathbf{A}\mathbf{x})$, Theorem 2.1 implies

$$\min_{j\le k}\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}_j^{\mathrm{F}}\|_2 \le \sqrt{k+1}\cdot\kappa(\mathbf{A})\cdot\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{x}_k^{\mathrm{G}}\|_2.$$

Therefore, so long as $\mathbf{A}$ is reasonably well-conditioned, the FOM error norms can, on the whole, never be significantly worse than those of GMRES. The error norms of FOM and GMRES can be estimated a posteriori to determine when to stop [16].

3.3. Connection to matrix functions. For convenience we will assume $\mathbf{A}$ is diagonalizable. FOM is closely related to the Arnoldi method for matrix function approximation (Arnoldi-FA) [9, 11, 15], which approximates $f(\mathbf{A})\mathbf{b}$ with the iterates $\mathbf{Q}_k f(\mathbf{H}_k)\mathbf{e}_1$. This is arguably the most widely used Krylov subspace method for approximating $f(\mathbf{A})\mathbf{b}$.

Commonly, $f(x)$ can be expressed in the form

$$f(x)=\int_\Gamma g(x)(x-z)^{-1}\mathrm{d}x, \tag{3.2}$$

for some weight function $g(x)$ and contour Γ. Common examples of functions which can be written in the form (3.2) include analytic functions (exponential, indicator function for a region in the complex plane) for which Γ is a contour in the complex plane, Stieltjes functions (inverse square root, logarithm) for which Γ is a subset of the real line, and rational functions with simple poles, which are a special case of Stieltjes functions; see for instance [10]. Arnoldi-FA is widely used to compute the action of the corresponding matrix functions on a vector in applications throughout the computational sciences, including quantum chromodynamics, differential equations, machine learning, etc. [10, 15].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

When $f(x)$ can be written as (3.2) at the eigenvalues of $\mathbf{A}$ and $\mathbf{H}_k$, the Arnoldi-FA error is expressed as

$$f(\mathbf{A})\mathbf{b} - \mathbf{Q}_k f(\mathbf{H}_k)\mathbf{e}_1 = \int_\Gamma \left[(\mathbf{A} - z\mathbf{I})^{-1}\mathbf{b} - \mathbf{Q}(\mathbf{H}_k - z\mathbf{I})^{-1}\mathbf{e}_1\right] \mathrm{d}\mu(z).$$

Note that $\mathbf{Q}(\mathbf{H}_k - z\mathbf{I})^{-1}\mathbf{e}_1$ is the FOM approximation to the linear system $(\mathbf{A} - z\mathbf{I})\mathbf{x} = \mathbf{b}$. Hence, understanding the behavior of FOM is important to understand the behavior of Arnoldi-FA.

Empirically, while Arnoldi-FA can sometimes have oscillatory convergence, it tends to follow a general downward trend approximately matching the best possible approximation from the Krylov subspace [1]. In fact, we are unaware of any examples for which this is not the case. Near-optimality guarantees have been proved in the symmetric case for the matrix exponential [8] and certain rational functions [1]. Theorem 2.1 is a near-optimality guarantee for Arnoldi-FA with $f(x) = 1/x$.

4. Conclusion. We have shown that the FOM residual norms are nearly as good as the GMRES residual norms in an overall sense. This provides theoretical justification for the use of FOM on linear systems of equations, as well as insight into the remarkable convergence of the Arnoldi method for matrix function approximation.

Acknowledgements. We thank the editor and anonymous referees for their helpful feedback which improved the paper.

REFERENCES

[1] N. AMSEL, T. CHEN, A. GREENBAUM, C. MUSCO, AND C. MUSCO, *Near-optimal approximation of matrix functions by the Lanczos method*, Preprint on arXiv, 2023. https://arxiv.org/abs/2303.03358
[2] W. E. ARNOLDI, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.
[3] R. F. BOISVERT, R. POZO, K. REMINGTON, R. F. BARRETT, AND J. J. DONGARRA, *Matrix market: a web resource for test matrix collections.*, in Quality of Numerical Software, R. F. Boisvert, ed., IFIP Advances in Information and Communication Technology 1997, Springer, New York, 1997, pp. 125–137.
[4] P. N. BROWN, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 58–78.
[5] T. CHEN, A. GREENBAUM, C. MUSCO, AND C. MUSCO, *Error bounds for Lanczos-based matrix function approximation*, SIAM J. Matrix Anal. Appl., 43 (2022), pp. 787–811.
[6] D. CHIN-LUNG FONG AND M. SAUNDERS, *CG versus MINRES: an empirical comparison*, Sultan Qaboos Univ. J. Sci. [SQUJS], 16 (2012), pp. 44–62.
[7] J. CULLUM AND A. GREENBAUM, *Relations between Galerkin and norm-minimizing iterative methods for solving linear systems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 223–247.
[8] V. DRUSKIN, A. GREENBAUM, AND L. KNIZHNERMAN, *Using nonorthogonal Lanczos vectors in the computation of matrix functions*, SIAM J. Sci. Comput., 19 (1998), pp. 38–54.
[9] V. DRUSKIN AND L. KNIZHNERMAN, *Two polynomial methods of calculating functions of symmetric matrices*, USSR Comput. Math. Math. Phys., 29 (1989), pp. 112–121.
[10] A. FROMMER AND V. SIMONCINI, *Matrix functions*, in Model Order Reduction: Theory, Research Aspects and Applications, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds., vol. 13 of Mathematics in Industry, Berlin, 2008, Springer, pp. 275–303.
[11] E. GALLOPOULOS AND Y. SAAD, *Efficient solution of parabolic equations by Krylov approximation methods*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1236–1264.
[12] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.
[13] A. GREENBAUM, V. DRUSKIN, AND L. KNIZHNERMAN, *On solving indefinite symmetric linear systems by means of the Lanczos method*, Zh. Vychisl. Mat. Mat. Fiz., 39 (1999), pp. 371–377.
[14] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436.
[15] N. J. HIGHAM, *Functions of Matrices*, SIAM, Philadelphia, 2008.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[16] G. MEURANT, *Estimates of the norm of the error in solving linear systems with FOM and GMRES*, SIAM J. Sci. Comput., 33 (2011), pp. 2686–2705.
[17] ———, *On the residual norm in FOM and GMRES*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 394–411.
[18] G. MEURANT AND J. DUINTJER TEBBENS, *Krylov Methods for Nonsymmetric Linear Systems–From Theory to Computations*, Springer, Cham, 2020.
[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] Y. SAAD, *Krylov subspace methods for solving large unsymmetric linear systems*, Math. Comp., 37 (1981), pp. 105–126.
[21] ———, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2003.
[22] 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.