GMRES, pseudospectra, and Crouzeix's conjecture for shifted and scaled Ginibre matrices

Tyler Chen, Anne Greenbaum, Thomas Trogdon

Introduction

GMRES is a commonly used iterative algorithm for solving a linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$. At iteration k, GMRES outputs an approximation $p(\mathbf{A})\mathbf{b}$ to $\mathbf{A}^{-1}\mathbf{b}$ where p is a degree $k-1$ polynomial.

The GMRES algorithm produces approximations whose residual norms are minimal:

$$\|\mathbf{r}_k\| = \min_{\deg(p)<k} \|\mathbf{b} - \mathbf{A}p(\mathbf{A})\mathbf{b}\| = \min_{\substack{\deg(p)\leq k\\ p(0)=1}} \|p(\mathbf{A})\mathbf{b}\|. \tag{1}$$

GMRES on random systems. For $\sigma \in (0,1)$, what happens if we apply GMRES to the system

$$(\mathbf{I} + \sigma \mathbf{G}_N)\mathbf{x} = \mathbf{b},$$

where $\sqrt{N}\mathbf{G}_N$ is a $N \times N$ Ginibre matrix (i.e has iid complex Gaussian entries)?

In general, we can bound

$$\frac{\|\mathbf{r}_k\|}{\|\mathbf{b}\|} \leq \|p(\mathbf{A})\| \leq C(\Omega, \mathbf{A}) \min_{\substack{\deg(p)\leq k\\ p(0)=1}} \|p\|_\Omega. \tag{2}$$

where $\Omega \subset \mathbb{R}$ and

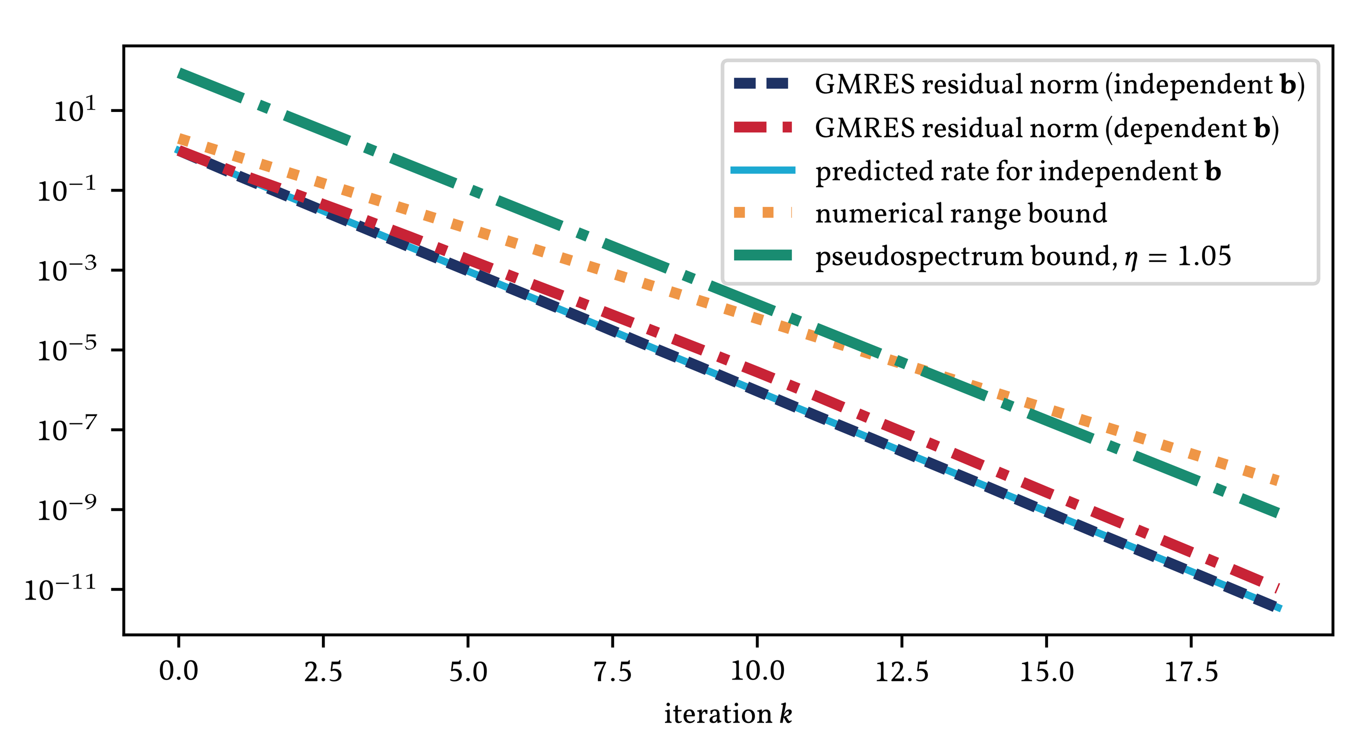
$$C(\Omega, \mathbf{A}) := \sup\{\|f\|_\Omega / \|f(\mathbf{A})\| : f \text{ analytic on } \Omega\}.$$

So it suffices to bound $C(\Omega, \mathbf{A})$ for any set Ω for which $\|p\|_\Omega$ can be bounded.

If $\mathbf{G}_N$ is independent of $\mathbf{b}$, then we might be able to say more.

A simple experiment

We set $N = 1500$ and $\sigma = 1/4$ and consider the convergence of GMRES when the right hand side is random as well as when it is chosen to try to get the slowest possible convergence.



In this paper we show:

- If $\mathbf{b}$ is independent of $\mathbf{G}_N$, then we can exactly characterize the convergence (in the large N limit).
- If $\mathbf{b}$ is dependent on $\mathbf{G}_N$, then we can bound the convergence.

Independent right hand side

Theorem 1. *Let $\mathbf{G}_N$ be an $N \times N$ complex Ginibre matrix. Then, for $\sigma \in (0,1)$ and $\mathbf{b}$ independent of $\mathbf{G}_N$, the step k GMRES residual norm $\|\mathbf{r}_k^N\|$ for the linear system $(\mathbf{I} + \sigma\mathbf{G}_N)\mathbf{x} = \mathbf{b}$ satisfies*

$$\frac{\|\mathbf{r}_k^N\|}{\|\mathbf{b}\|} \xrightarrow[N\to\infty]{\text{prob.}} \left(\frac{1-\sigma^2}{1-\sigma^{2+2k}}\right)^{1/2} \sigma^k.$$

Argument sketch. Ginibre matrices are each unitarily invariant, so WLOG $\mathbf{b} = [1, 0, \ldots, 0]^\mathsf{T}$. The GMRES residual is also unitarily invariant, so we will find a unitary matrix which makes the problem easier to analyze.

Partition $\mathbf{G}_N$ as

$$\mathbf{G}_N = \frac{1}{\sqrt{2N}} \begin{bmatrix} g & \mathbf{y}^\mathsf{T} \\ \mathbf{x} & \mathbf{G} \end{bmatrix}.$$

Now, conditioning on the probability one event that $\|\mathbf{v}\| \neq 0$, define $\mathbf{U}$ as the Householder reflector

$$\mathbf{U} = \begin{bmatrix} 1 & 0^\mathsf{T} \\ 0 & \mathbf{F} \end{bmatrix}, \qquad \mathbf{F} = \mathbf{I} - \frac{\mathbf{v}\mathbf{v}^\mathsf{H}}{\mathbf{v}^\mathsf{H}\mathbf{v}}, \qquad \mathbf{v} = \|\mathbf{x}\|\mathbf{e}_1 - \mathbf{x}.$$

Then $\mathbf{U}\mathbf{e}_1 = \mathbf{e}_1$ and

$$\mathbf{U}\mathbf{G}_N\mathbf{U}^\mathsf{H} = \frac{1}{\sqrt{2N}} \begin{bmatrix} g & \mathbf{y}^\mathsf{T}\mathbf{F}^\mathsf{H} \\ \mathbf{F}\mathbf{x} & \mathbf{F}\mathbf{G}\mathbf{F}^\mathsf{H} \end{bmatrix}.$$

By construction, $\mathbf{F}\mathbf{x} = [\|\mathbf{x}\|, 0, \ldots, 0]^\mathsf{T}$. Moreover, since the real and imaginary parts of the entires of $\mathbf{x}$ are iid standard normal random variables,

$$\|\mathbf{x}\| \stackrel{\text{dist.}}{=} \chi(2(N-1)).$$

Inductively:

$$\mathbf{V}\mathbf{G}_N\mathbf{V}^\mathsf{H} \stackrel{\text{dist.}}{=} \frac{1}{\sqrt{2N}} \begin{bmatrix} \mathcal{N}_\mathbb{C}(0,2) & \cdots & \cdots & \cdots & \mathcal{N}_\mathbb{C}(0,2) \\ \chi(2(N-1)) & \ddots & & & \vdots \\ & \chi(2(N-2)) & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \chi(2(1)) & \mathcal{N}_\mathbb{C}(0,2) \end{bmatrix}.$$

In the large N limit,

$$\mathbf{A}_N = \mathbf{V}(\mathbf{I} + \sigma\mathbf{G}_N)\mathbf{V}^\mathsf{H} \xrightarrow[N\to\infty]{\text{prob.}} \begin{bmatrix} 1 & & & \\ \sigma & \ddots & & \\ & \ddots & \ddots & \\ & & \sigma & 1 \end{bmatrix} =: \mathbf{A}_\infty.$$

To complete the analysis, we analyze the residual norms for GMRES applied to the deterministic system $\mathbf{A}_\infty\mathbf{x} = [1, 0, \ldots, 0]^\mathsf{T}$. Note that Arnoldi applied to an upper-Hessenberg matrix and the first unit vector produces back that exact same upper-Hessenberg matrix. Using this, and classical formulas for the GMRES residual in terms of the upper-Hessenberg matrix produced by Arnoldi, we get the result.

Dependent right hand side

Theorem 2. *Let $\mathbf{G}_N$ be an $N \times N$ complex Ginibre matrix. Then, for $\sigma \in (0,1)$ and $\mathbf{b}$ possibly depending on $\mathbf{G}_N$, the step k GMRES residual norm $\|\mathbf{r}_k^N\|$ for the linear system $(\mathbf{I} + \sigma\mathbf{G}_N)\mathbf{x} = \mathbf{b}$ satisfies, for all $\eta > 1$*

$$\frac{\|\mathbf{r}_k^N\|}{\|\mathbf{b}\|} \lesssim \eta\mathfrak{e}(\eta)^{-1/2}(\sigma\eta)^k \quad \text{and} \quad \frac{\|\mathbf{r}_k\|}{\|\mathbf{b}\|} \lesssim 2(\sqrt{2}\sigma)^k \quad \text{a.s. as } N \to \infty.$$

Pseudospectrum. We ca take Ω in (2) to be the ϵ-pseudospectrum

$$\Lambda_\epsilon(\mathbf{A}) := \{z \in \mathbb{C} : \|R(z, \mathbf{A})\| \geq \epsilon^{-1}\}$$

where $R(z, \mathbf{A}) := (z\mathbf{I} - \mathbf{A})^{-1}$ is the resolvent. For any $\epsilon > 0$, using the fact that $\|R(z, \mathbf{A})\| = \epsilon^{-1}$ for $z \in \partial\Lambda_\epsilon$,

$$\|f(\mathbf{A})\| = \left\|\frac{1}{2\pi i}\int_{\partial\Lambda_\epsilon} f(z)R(z, \mathbf{A})\mathrm{d}z\right\| \leq \frac{\operatorname{len}(\partial\Lambda_\epsilon(\mathbf{A}))}{2\pi\epsilon}\|f\|_{\Lambda_\epsilon}.$$

Using existing bounds [5, 1, 3], we can characterize $\Lambda_\epsilon(\mathbf{G}_N)$:

Theorem 3. *Let $\mathbf{G}_N$ be a complex Ginibre matrix of size $N \times N$. Then for any $r > 1$, with $\eta = \mathfrak{e}(\eta)^{-1/2}$,*

$$d_H(\Lambda_\eta(\mathbf{G}_N), \mathcal{D}(0, r)) \xrightarrow[N\to\infty]{\text{a.s.}} 0.$$

Here $\mathfrak{e} : [1, \infty) \to [0, \infty)$ is a deterministic increasing function with $\mathfrak{e}(1) = 0$ which is defined explicitly.

Numerical range. We can also take Ω as the numerical range

$$W(\mathbf{A}) := \{\mathbf{v}^\mathsf{H}\mathbf{A}\mathbf{v} : \|\mathbf{v}\| = 1\}.$$

It is known that $C(W(\mathbf{A}), \mathbf{A}) \leq 1 + \sqrt{2}$ for all $\mathbf{A}$ [4] and conjectured that $C(W(\mathbf{A}), \mathbf{A}) \leq 2$ for all $\mathbf{A}$. We use [2] to prove a slightly weaker version of the conjecture for Ginibre matrices:

Theorem 4. *Let $\mathbf{G}_N$ be a complex Ginibre matrix of size $N \times N$. Then,*

$$C(W(\mathbf{G}_N), \mathbf{G}_N) \lesssim 2 \quad \text{a.s. as } N \to \infty.$$

References

[1] Z. Bai and J. W. Silverstein. "No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Information-Plus-Noise Type Matrices". In: *Random Matrices: Theory and Applications* 01.01 (Jan. 2012), p. 1150004.

[2] T. Caldwell, A. Greenbaum, and K. Li. "Some Extensions of the Crouzeix–Palencia Result". In: *SIAM Journal on Matrix Analysis and Applications* 39.2 (Jan. 2018), pp. 769–780.

[3] G. Cipolloni, L. Erdős, and D. Schröder. "Optimal lower bound on the least singular value of the shifted Ginibre ensemble". In: *Probability and Mathematical Physics* 1.1 (Nov. 2020), pp. 101–146.

[4] M. Crouzeix and C. Palencia. "The Numerical Range is a $(1+\sqrt{2})$-Spectral Set". In: *SIAM Journal on Matrix Analysis and Applications* 38.2 (Jan. 2017), pp. 649–655.

[5] R. B. Dozier and J. W. Silverstein. "On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices". In: *Journal of Multivariate Analysis* 98.4 (Apr. 2007), pp. 678–694.