GMRES with Preconditioning, Weighted norm and Deflation

Nicole Spillane, Daniel B. Szyld

Abstract

We consider the general problem of solving a linear system of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b};\ \mathbf{A} \in \mathbb{C}^{n\times n};\ \mathbf{b} \in \mathbb{C}^{n}.$$

The matrices $\mathbf{A}$ that we consider are non-singular, sparse and of high order n. For solving these matrices, GMRES [3, Chapter 6] is a natural choice. We address two fundamental and connected questions: How can the convergence of GMRES be predicted ? How can the convergence of GMRES be accelerated ? Our aim is to combine three ways of accelerating GMRES convergence:

- Weighting by a Hermitian positive definite (hpd) matrix $\mathbf{W}$: all inner products and norms in the GMRES algorithm are replaced by the ones induced by $\mathbf{W}$ (see [1]),
- Preconditioning by a non-singular matrix $\mathbf{H}$: GMRES is applied to the preconditioned problem $\mathbf{AHu} = \mathbf{b}$ with $\mathbf{x} = \mathbf{Hu}$ (see [3, Section 9.3]),
- Deflation by a projection operator $\mathbf{\Pi} := \mathbf{I} - \mathbf{AZ}(\mathbf{Y}^*\mathbf{AZ})^{-1}\mathbf{Y}^*$ (with $\mathbf{Y}, \mathbf{Z} \in \mathbb{C}^{n\times m}$): GMRES is applied to the projected problem $\mathbf{\Pi AHu} = \mathbf{\Pi b}$ (see [7, 4]). A suitable initialization is also performed that accounts for the part of the solution that has been projected away.

We refer to $\mathbf{W}$, $\mathbf{H}$ and $\mathbf{\Pi}$ as accelerators for GMRES. With words, the strategy is that the preconditioner $\mathbf{H}$ should be a *good* approximation of $\mathbf{A}^{-1}$, the deflation operator should handle the space where $\mathbf{H}$ does not *well* approximate $\mathbf{A}^{-1}$, and the weighted inner product should facilitate the analysis. In practice, identifying efficient accelerators requires a GMRES convergence bound where the influence of $\mathbf{H}$, $\mathbf{W}$ and $\mathbf{\Pi}$ is explicit. We prove in [6, Theorem 4.1] that the convergence rate is bounded by

$$\frac{\|\mathbf{r}_{i+1}\|^2_{\mathbf{W}}}{\|\mathbf{r}_i\|^2_{\mathbf{W}}} \leq 1 - \inf_{\mathbf{y}\in\mathrm{range}(\mathbf{\Pi})\setminus\{\mathbf{0}\}} \frac{|\langle \mathbf{\Pi AHy}, \mathbf{y}\rangle_{\mathbf{W}}|^2}{\|\mathbf{\Pi AHy}\|^2_{\mathbf{W}}\|\mathbf{y}\|^2_{\mathbf{W}}}.$$

Further Assumptions Major simplifications occur in the case where $\mathbf{A}$ is positive definite (*i.e.*, its Hermitian part is hpd), the preconditioner $\mathbf{H}$ is hpd, and the weight equals the preconditioner $\mathbf{W} = \mathbf{H}$. In this case (and with a technical assumption on the deflation operator), it holds that

$$\frac{\|\mathbf{r}_{i+1}\|^2_{\mathbf{H}}}{\|\mathbf{r}_i\|^2_{\mathbf{H}}} \leq 1 - \inf_{\mathbf{y}\in\mathrm{range}(\mathbf{AH\Pi})\setminus\{\mathbf{0}\}} \frac{|\langle \mathbf{A}^{-1}\mathbf{y}, \mathbf{y}\rangle|}{\langle \mathbf{y}, \mathbf{M}^{-1}\mathbf{y}\rangle} \times \frac{\lambda_{\min}(\mathbf{HM})}{\lambda_{\max}(\mathbf{HM})},$$

where $\mathbf{M} = 1/2(\mathbf{A} + \mathbf{A}^*)$ and $\mathbf{N} = 1/2(\mathbf{A} - \mathbf{A}^*)$ are the Hermitian and skew-Hermitian parts of $\mathbf{A}$, and the spectrum of $\mathbf{HM}$ is in the interval $[\lambda_{\min}(\mathbf{HM}), \lambda_{\max}(\mathbf{HM})]$.

Convergence without deflation Setting $\mathbf{\Pi} = \mathbf{I}$ (no deflation) and with an identity from [2] it is proved in [5, Theorem 4.3] that

$$\frac{\|\mathbf{r}_{i+1}\|^2_{\mathbf{H}}}{\|\mathbf{r}_i\|^2_{\mathbf{H}}} \leq 1 - \frac{1}{1 + \rho(\mathbf{M}^{-1}\mathbf{N})^2} \times \frac{\lambda_{\min}(\mathbf{HM})}{\lambda_{\max}(\mathbf{HM})}$$

where $\rho(\cdot)$ denotes the spectral radius of a matrix. The residuals are bounded with respect to two quantities. The first is the condition number of $\mathbf{HM}$, a measure of whether $\mathbf{H}$ is a good preconditioner for the hpd matrix $\mathbf{M}$. The second is the spectral radius of $\mathbf{M}^{-1}\mathbf{N}$, a measure of how non-Hermitian the problem is. The takeaway is that fast convergence is guaranteed if the problem is mildly non-Hermitian and $\mathbf{H}$ is a good preconditioner for $\mathbf{M}$. The bound also has important consequences for parallel computing and the analysis of domain decomposition methods.

A new deflation space [6, Theorem 6.3] When the problem is significantly non-Hermitian (in terms of $\rho(\mathbf{M}^{-1}\mathbf{N})$), we propose to combine Hermitian preconditioning with spectral deflation.

Under the same assumptions as above, we choose the matrices $\mathbf{Z}$ and $\mathbf{Y}$ in the characterization of the projection operator $\mathbf{\Pi}$ as follows. First, we denote by $(\lambda_j, \mathbf{z}^{(j)}) \in \mathrm{i}\mathbb{R} \times \mathbb{C}^n$ (for $j = 1, \ldots, n$) the eigenpairs of the generalized eigenvalue problem $\mathbf{N}\mathbf{z}^{(j)} = \lambda_j \, \mathbf{M}\mathbf{z}^{(j)}$. Then, with a chosen threshold $\tau > 0$ we select for the deflation operator, the highest frequency eigenvectors, by setting

$$\operatorname{span}(\mathbf{Z}) := \operatorname{span}\{\mathbf{z}^{(j)}; |\lambda_j| > \tau\} \text{ and } \mathbf{Y} = \mathbf{HAZ}.$$

Then the convergence of weighted, preconditioned and deflated GMRES is bounded by

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{H}}^2}{\|\mathbf{r}_i\|_{\mathbf{H}}^2} \leq 1 - \frac{1}{(1+\tau^2)} \times \frac{\lambda_{\min}(\mathbf{HM})}{\lambda_{\max}(\mathbf{HM})}.$$

Numerical illustrations show that preconditioning the Hermitian part in a way that is scalable leads to overall scalability and that spectral deflation accelerates convergence when the problems become more strongly non-Hermitian.

References

[1] A. Essai. Weighted FOM and GMRES for solving nonsymmetric linear systems. *Numer. Algorithms*, 18(3-4):277–292, 1998.

[2] C. R. Johnson. Inequalities for a complex matrix whose real part is positive definite. *Trans. Am. Math. Soc.*, 212:149–154, 1975.

[3] Y. Saad. *Iterative methods for sparse linear systems.* Philadelphia, PA: SIAM Society for Industrial and Applied Mathematics, 2nd ed. edition, 2003.

[4] K. M. Soodhalter, E. de Sturler, and M. E. Kilmer. A survey of subspace recycling iterative methods. *GAMM-Mitt.*, 43(4):29, 2020. Id/No e202000016.

[5] N. Spillane. Hermitian preconditioning for a class of non-Hermitian linear systems. *SIAM J. Sci. Comput.*, 46(3):a1903–a1922, 2024.

[6] N. Spillane and D. B. Szyld. New convergence analysis of GMRES with weighted norms, preconditioning, and deflation, leading to a new deflation space. *SIAM J. Matrix Anal. Appl.*, 45(4):1721–1745, 2024.

[7] J. M. Tang, R. Nabben, C. Vuik, and Y. A. Erlangga. Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods. *J. Sci. Comput.*, 39:340–370, 2009.