MA3227 Numerical Analysis II

Lecture 7: Convergence of GMRES

Simon Etter

National University of Singapore

2019/2020

Convergence of GMRES

Recap GMRES

GMRES approximates $x = A^{-1}b$ by

$$x_k := p_{k-1}(A)\, b \qquad \text{where} \qquad p_{k-1} := \operatorname*{arg\,min}_{p_{k-1} \in \mathcal{P}_{k-1}} \big\| \big(A p_{k-1}(A) - I\big)\, b \big\|_2.$$

This x_k can be determined using

$$k \text{ matrix-vector products, } \mathcal{O}\big(Nk^2\big) \text{ other operations.}$$

Observations

$x_N = x$ because then we can choose $p_{N-1}(x)$ such that $p_{N-1}(\lambda) = \frac{1}{\lambda}$ for all eigenvalues λ of A and hence $p_{N-1}(A) = A^{-1}$.

Alternatively, for $k = N$ the Krylov subspace $K_N = \operatorname{span}\{b, Ab, \ldots, A^{N-1}b\}$ is N-dimensional and hence $K_N = \mathbb{R}^N$.

Unfortunately, for $k = N$ the runtime becomes $\mathcal{O}(N^3)$ which is even worse than sparse LU factorisation.

Convergence of GMRES

Discussion

Krylov methods are powerful if we can get $x_k \approx x$ already for $k \ll N$.

This lecture will discuss under what conditions on A (and to a lesser extent b) this is the case.

Error measure

In the following, we will provide bounds for $\|Ax_k - b\|_2$.

The main reason for doing so is that bounding $\|Ax_k - b\|_2$ is easier than bounding other error measures because GMRES explicitly minimises this quantity.

In applications, we are often interested in $\|x_k - x\|_2$.
We have the following bound:

$$\|x_k - x\|_2 \leq \|A^{-1}\|_2 \, \|Ax_k - b\|_2.$$

This bound may be quite loose for practical purposes, but it is the best we can do given the current setting.

Convergence of GMRES

Bounding the GMRES residual

Assume A has eigendecomposition $A = V\Lambda V^{-1}$. Then,

$$\begin{aligned}
\|Ax_k - b\|_2 &= \min_{p_{k-1}\in\mathcal{P}_{k-1}} \left\|\left(Ap_{k-1}(A) - I\right) b\right\|_2 \\
&\le \min_{p_{k-1}\in\mathcal{P}_{k-1}} \|V\|_2 \left\|\Lambda p_{k-1}(\Lambda) - I\right\|_2 \|V^{-1}\|_2 \|b\|_2 \\
&\le \kappa(V)\,\|b\|_2 \min_{p_{k-1}\in\mathcal{P}_{k-1}} \max_{\lambda_\ell} |\lambda_\ell\, p_{k-1}(\lambda_\ell) - 1|.
\end{aligned}$$

$\lambda_\ell = \Lambda[\ell,\ell]$ on last line are the eigenvalues of A.
$\kappa(V) = \|V\|_2\,\|V^{-1}\|_2$ is the condition number of V.
We have $\|\Lambda\|_2 = \max_\ell |\Lambda[\ell,\ell]|$ for any diagonal matrix Λ.

Conclusion

To get an asymptotic error estimate for $\|Ax_k - b\|_2$, we should study the behaviour of

$$\min_{p_{k-1}\in\mathcal{P}_{k-1}} \max_{\lambda_\ell} |\lambda_\ell\, p_{k-1}(\lambda_\ell) - 1|$$

as a function of k.
We next reformulate this problem to make it easier to argue about it.

Convergence of GMRES

GMRES minimisation problem, observation 1

$$p_{k-1} \in \mathcal{P}_{k-1} \quad \Longrightarrow \quad q_k(x) := x\, p_{k-1}(x) - 1 \in \mathcal{P}_k,\ q(0) = -1$$

$$q_k \in \mathcal{P}_k,\ q_k(0) = -1 \quad \Longrightarrow \quad q_k(x) := x\, p_{k-1}(x) - 1 \text{ for some } p_{k-1} \in \mathcal{P}_{k-1}.$$

Proof. First implication is obvious.
Second implication: $p_{k-1}(x) = \frac{q_k(x)+1}{x}$ is a polynomial since $q_k(0) = -1$.

Corollary
The GMRES minimisation problem may equivalently be formulated as

$$\min_{q_k \in \mathcal{P}_k} \max_{\lambda_\ell} \frac{|q_k(\lambda_k)|}{|q_k(0)|}.$$

Hence, we want $q_k \in \mathcal{P}_k$ such that $|q_k(\lambda_\ell)|$ is small relative to $|q_k(0)|$.

Convergence of GMRES

GMRES minimisation problem, observation 2

Eigenvalues λ_ℓ typically cluster in a set $\mathcal{E} \subset \mathbb{C}$.
The GMRES minimisation problem

$$\min_{q_k \in \mathcal{P}_k} \max_{\lambda_k} \frac{|q_k(\lambda_\ell)|}{|q(0)|}$$

may then be replaced by

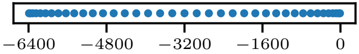
$$\min_{q_k \in \mathcal{P}_k} \max_{x \in \mathcal{E}} \frac{|q_k(x)|}{|q(0)|}$$

without losing much in sharpness.

Example

Recall eigenvalues $\lambda_\ell := (n+1)^2 \big(2\cos\big(\pi \frac{\ell}{n+1}\big) - 2\big)$ of discrete Laplacian.
These eigenvalues cluster in the interval $\mathcal{E} = [-4\,(n+1)^2, 0]$.

$$n = 39 \quad \longrightarrow \quad \lambda_\ell =$$



Convergence of GMRES

GMRES minimisation problem, conclusion

Summary of the above: there exists a $C \neq C(k)$ such that

$$\|Ax_k - b\|_2 \leq C \min_{q_k \in \mathcal{P}_k} \max_{x \in \mathcal{E}} \frac{|q_k(x)|}{|q_k(0)|}.$$

To get a good bound on $\|Ax_k - b\|_2$, we should hence look for a $q_k(x)$ which is as small as possible on $\mathcal{E}$ relative to $q_k(0)$.

It is hard to make rigorous statements about this problem for general $\mathcal{E}$. Instead, we will do the following:

- Develop some intuition for what properties of $\mathcal{E}$ make GMRES converge fast.
- Provide a rigorous estimate for the case $\mathcal{E} = [a, b]$.

Convergence of GMRES

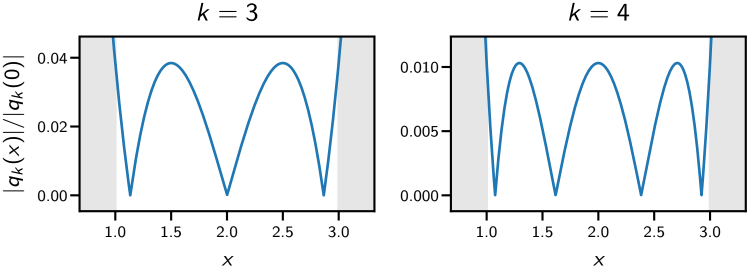
Qualitative convergence theory

Recipe for constructing $q_k(x)$ such that $\max_{x \in \mathcal{E}} \frac{|q_k(x)|}{|q_k(0)|}$ is small:

Choose $q_k(x) = \prod_{\ell=1}^{k} (x - x_\ell)$ with x_ℓ distributed over $\mathcal{E}$.

This choice ensures $q_k(x) \approx 0$ for $x \approx x_\ell$, i.e. each factor makes $q_k(x)$ in a small region around x_ℓ. If we add enough factors, we get that $q_k(x)$ is small throughout $\mathcal{E}$.

Example for $\mathcal{E} = [1, 3]$.



Convergence of GMRES

Qualitative convergence theory (continued)

Properties of $\mathcal{E}$ which make GMRES converge fast:

- $\mathcal{E}$ is small: few x_ℓ are enough to ensure that $\max_{x \in \mathcal{E}} |q_k(x)|$ is small.
- $\mathcal{E}$ is far away from 0: $|q_k(0)| \geq \big(\min_{x \in \mathcal{E}} |x|\big)^k$.

These two points are equivalent after scaling:

- $\mathcal{E} = [\frac{1}{n}, 1]$ is bounded ("small") but close to 0 for $n \to \infty$.
- $\mathcal{E} = [1, n]$ is unbounded ("large") but bounded away from 0 for $n \to \infty$.

Convergence is the same in both cases (see next slide)

Convergence of GMRES

Thm: Scale-invariance of GMRES

Let $\mathcal{E} \subset \mathbb{C}$ and define $\theta\mathcal{E} = \{\theta x \mid x \in \mathcal{E}\}$ for $\theta \in \mathbb{R} \setminus \{0\}$. Then,

$$\min_{q_k \in \mathcal{P}_k} \max_{x \in \mathcal{E}} \frac{|q_k(x)|}{|q_k(0)|} = \min_{q_k \in \mathcal{P}_n} \max_{x \in \theta\mathcal{E}} \frac{|q_k(x)|}{|q_k(0)|}.$$

Proof.

Assume there exists $\theta \in \mathbb{R} \setminus \{0\}$ such that

$$\min_{q_k^{(1)} \in \mathcal{P}_k} \max_{x \in \mathcal{E}} \frac{|q_k^{(1)}(x)|}{|q_k^{(1)}(0)|} < \min_{q_k^{(\theta)} \in \mathcal{P}_n} \max_{x \in \theta\mathcal{E}} \frac{|q_k^{(\theta)}(x)|}{|q_k^{(\theta)}(0)|} = \min_{q_k^{(\theta)} \in \mathcal{P}_n} \max_{x \in \mathcal{E}} \frac{|q_k^{(\theta)}(\theta x)|}{|q_k^{(\theta)}(0)|}, \tag{1}$$

Let $q_k^{(1)}(x) = \prod_{\ell=1}^{k}(x - x_\ell)$ be the minimiser from the left-hand side.

Note that any $q_k \in \mathcal{P}_k$ is of the form $q_k(x) = C \prod_{\ell=1}^{k}(x - x_\ell)$ by the fundamental theorem of algebra, and we can assume $C = 1$ in our case because we are only interested in the ratio $q_k(x)/q_k(0)$.

Convergence of GMRES

Proof (continued).

Then, we obtain for $\tilde{q}_k^{(\theta)}(\theta x) = \prod_{\ell=1}^{k}(\theta x - \theta\, x_\ell) \in \mathcal{P}_k$ that

$$\begin{aligned}\max_{x\in\mathcal{E}} \frac{|\tilde{q}_k^{(\theta)}(\theta x)|}{|\tilde{q}_k^{(\theta)}(0)|} &= \max_{x\in\mathcal{E}} \frac{\prod_{\ell=1}^{k}|\theta x - \theta\, x_\ell|}{\prod_{\ell=1}^{k}|\theta x_\ell|} \\ &= \max_{x\in\mathcal{E}} \frac{\prod_{\ell=1}^{k}|x - x_\ell|}{\prod_{\ell=1}^{k}|x_\ell|} \quad = \max_{x\in\mathcal{E}} \frac{|q_k^{(1)}(x)|}{|q_k^{(1)}(0)|}.\end{aligned} \tag{2}$$

Equation (2) contradicts the assumption (1); hence we have

$$\min_{q_k\in\mathcal{P}_k} \max_{x\in\mathcal{E}} \frac{|q_k(x)|}{|q_k(0)|} \geq \min_{q_k\in\mathcal{P}_n} \max_{x\in\theta\mathcal{E}} \frac{|q_k(x)|}{|q_k(0)|}.$$

The bound in the other direction follows by applying the above bound to $\mathcal{E}' = \theta\mathcal{E}$ and $\theta' = \theta^{-1}$.

Convergence of GMRES

Qualitative convergence theory (continued)

Previous slides showed that GMRES converges fast if $\mathcal{E}$ is small relative to its distance to 0. Now we somewhat relax this condition.

Thm: Invariance of GMRES with respect to outliers

Assume $\mathcal{E} = \bar{\mathcal{E}} \cup \{\lambda^\star\}$, i.e. the eigenvalues of A are all contained in some set $\bar{\mathcal{E}}$ except for a single outlier eigenvalue $\lambda^\star \in \mathbb{C}$.

Then,

$$\min_{q_{k+1} \in \mathcal{P}_{k+1}} \max_{x \in \mathcal{E}} \frac{|q_{k+1}(x)|}{|q_{k+1}(0)|} \leq \left(\max_{x \in \bar{\mathcal{E}}} \frac{|x - \lambda^\star|}{|\lambda^\star|} \right) \left(\min_{q_k \in \mathcal{P}_k} \max_{x \in \bar{\mathcal{E}}} \frac{|\bar{q}_k(x)|}{|\bar{q}_k(0)|} \right).$$

In words:

- The convergence of GMRES on $\mathcal{E}$ is at most a constant factor worse than the convergence of GMRES on $\bar{\mathcal{E}}$.
- This constant factor $\max_{x \in \bar{\mathcal{E}}} \frac{|x-\lambda^\star|}{|\lambda^\star|}$ may be large if $\lambda^\star$ is close to 0.

Convergence of GMRES

Proof. Let

$$\bar{q}_k(x) = \operatorname*{arg\,min}_{\bar{q}_k \in \mathcal{P}_k} \max_{x \in \bar{\mathcal{E}}} \frac{|\bar{q}_k(x)|}{|\bar{q}_k(0)|} \qquad \text{and} \qquad q_{k+1}(x) = (x - \lambda^\star)\, \bar{q}_k(x).$$

Then $q_{k+1}(\lambda^\star) = 0$ and hence

$$\begin{aligned} \max_{x \in \mathcal{E}} \frac{|q_{k+1}(x)|}{|q_{k+1}(0)|} &= \max_{x \in \bar{\mathcal{E}}} \frac{|q_{k+1}(x)|}{|q_{k+1}(0)|} = \max_{x \in \bar{\mathcal{E}}} \frac{|x - \lambda^\star|}{|\lambda^\star|} \, \frac{|\bar{q}_k(x)|}{|\bar{q}_k(0)|} \\ &\le \left(\max_{x \in \bar{\mathcal{E}}} \frac{|x - \lambda^\star|}{|\lambda^\star|} \right) \left(\max_{x \in \bar{\mathcal{E}}} \frac{|\bar{q}_k(x)|}{|\bar{q}_k(0)|} \right). \end{aligned}$$

Convergence of GMRES

Quantitative convergence estimate

$$\min_{q_k \in \mathcal{P}_k} \max_{x \in [1,\kappa]} \frac{|q_k(x)|}{|q_k(0)|} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k .$$

Remarks

- By the scale-invariance of GMRES, this result applies to any interval $[a, b] \not\ni 0$ with $\kappa = \frac{b}{a}$.
- This result can be shown by choosing $q_k(x)$ as a shifted and scaled Chebyshev polynomial. See **[TB97]** for details.

Numerical illustration

See `gmres_convergence()`.

[TB97] L. N. Trefethen and D. Bau. *Numerical Linear Algebra.* Society for Industrial and Applied Mathematics (1997),

Convergence of GMRES

Recap: bound on GMRES residual

$$\|Ax_k - b\|_2 \leq \kappa(V)\, \|b\|_2 \min_{q_k \in \mathcal{P}_k} \max_{\lambda_\ell} \frac{|q_k(\lambda_\ell)|}{|q(0)|}$$

Previous slides discussed how to choose q_k such that the right hand side becomes as small as possible.

Surprisingly, it turns out that we can also play around with b and λ_ℓ to make the right-hand side even smaller.

The following slides will explain further.

Convergence of GMRES

GMRES with initial guess (modifying b)

Assume we have an initial guess $x_0 \approx x$ for the solution to $Ax = b$.

We can exploit the knowledge contained in x_0 by making the ansatz

$$A\,(x_0 + \Delta x) = b$$

and solving for Δx:

$$A\,\Delta x = b - Ax_0.$$

GMRES applied to the last equation will produce a sequence of estimates Δx_k which satisfy the bound

$$\begin{aligned} \|A\,(x_0 + \Delta x_k) - b\|_2 &= \|A\,\Delta x_k - (b - Ax_0)\|_2 \\ &\le \kappa(V)\,\|b - Ax_0\|_2 \min_{q_k \in \mathcal{P}_k} \max_{\lambda_\ell} \frac{|q_k(\lambda_\ell)|}{|q(0)|}. \end{aligned}$$

If $x_0 \approx x$, then $\|b - Ax_0\|_2$ is small and hence $\|A\,(x_0 + \Delta x_k) - b\|_2$ is small as well.

Convergence of GMRES

Restarted GMRES

One way to obtain an initial guess x_0 is to run GMRES with a small, a-priori determined number of steps k_{inner}. (A common choice is $k_{\text{inner}} = 20$.)

Iterating this idea yields the following algorithm.

Algorithm 1 Restarted GMRES

```
x_0 = 0
for k = 1, 2, ... do
    Approximately solve A Δx_k = b − A x_{k−1} using k_inner GMRES steps.
    Update x_k = x_{k−1} + Δx_k.
end for
```

The advantage of restarting is that it avoids the $\mathcal{O}(k^2)$ scaling of non-restarted GMRES.

The disadvantage is that it may converge (much) more slowly.

See `restarted_gmres_good()` and `restarted_gmres_bad()`.

Convergence of GMRES

Preconditioning (modifying λ_k)

Instead of $Ax = b$, we can solve

- $P^{-1}Ax = P^{-1}b$ (left preconditioning), or
- $(AP^{-1})(Px) = b$ and $x = P^{-1}(Px)$ (right preconditioning)

with P some invertible matrix (preconditioner).

GMRES applied to these modified systems converges in fewer iterations if the eigenvalues of $P^{-1}A$ or AP^{-1} are "nicer" than those of A.
In particular, GMRES converges in a single iteration if we choose $P = A$.

On the other hand, each iteration is more expensive because each Arnoldi iteration now requires us to evaluate two matrix-vector products $P^{-1}\,(Av)$ or $A\,(P^{-1}v)$ instead of just Av.
In particular, $P = A$ useless because if we could compute $P^{-1}v = A^{-1}v$ cheaply we would not be thinking about using GMRES in the first place.

We conclude from the above that a good preconditioner should balance

$$P^{-1} \approx A^{-1} \qquad \text{and} \qquad P^{-1}v \text{ is easy to evaluate.}$$

Convergence of GMRES

Remark

The discussion on the previous slide assumed that we evaluate $P^{-1}Av$ or $AP^{-1}v$ as two matrix-vector products $P^{-1}(Av)$ or $A(P^{-1}v)$.

It is also possible to evaluate $P^{-1}A$ or AP^{-1} once and then evaluate $(P^{-1}A)\,v$ or $(AP^{-1})\,v$ as a single matrix-vector product.

The second approach is rarely used in practice because we typically have a fast matrix-vector product for A and P^{-1} but not for AP^{-1} or $P^{-1}A$.

For this reason, the two-matrix-vector-products strategy is typically faster than the one-matrix-vector-product strategy.

Convergence of GMRES

Incomplete LU (ILU) preconditioning

Recall: the problem with LU factorisation is excessive fill-in.

Idea: only store a limited number of fill-in entries and discard the rest.

The resulting factors $\tilde{L}, \tilde{U}$ may be inaccurate ($\tilde{L}\tilde{U} - A$ may be large), but they may be accurate enough that $P = \tilde{L}\tilde{U}$ is a good preconditioner.

Two versions of ILU are in widespread use:

- ILU(c): only allow $(\tilde{L} + \tilde{U})[i, j] \neq 0$ if i, j are connected by a fill path of length $\leq c + 1$.
- ILU(τ): only allow $(\tilde{L} + \tilde{U})[i, j] \neq 0$ if $|(L + U)[i, j]| > \tau$.

I will demonstrate ILU preconditioning in Lecture 9.

Convergence of GMRES

Summary

- GMRES residual bound:

$$\|Ax_k - b\|_2 \leq C \min_{q_k \in \mathcal{P}_k} \max_{\lambda_\ell} \frac{|q_k(\lambda_\ell)|}{|q_k(0)|} \tag{3}$$

  where λ_ℓ are the eigenvalues of A and C is some constant which depends on A and b but not on k.
- The right-hand side of (3) is the same if we scale the eigenvalues, and it is asymptotically independent of "outlier eigenvalues".
- We have an explicit bound on the polynomial minimisation problem if the eigenvalues cluster in an interval $\mathcal{E} = [1, \kappa]$:

$$\min_{q_k \in \mathcal{P}_k} \max_{x \in [1,\kappa]} \frac{|q_k(x)|}{|q_k(0)|} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k .$$

  You do not have to remember the precise form of this bound, but you should remember that larger κ means slower convergence.

Convergence of GMRES

Summary (continued)

- GMRES with initial guess:

$$\Delta x \approx A^{-1}(b - Ax_0) \quad \Longleftrightarrow \quad x \approx x_0 + \Delta x.$$

- Preconditioning:

$$P^{-1}Ax = P^{-1}b \qquad \text{or} \qquad (AP^{-1})(Px) = b \text{ and } x = P^{-1}(Px).$$