SCSC seminar 2.

Fruzsina Agocs : Convergence rate of GMRES with potential theory

① Introduction

- large-scale matrix problems $Ax = b$, $A \in \mathbb{C}^{N \times N}$
x unknown

solved by iterating in Krylov subspace:

$$K_n = \{ b, Ab, A^2b, \dots, A^{n-1}b \}$$

$\{\cdot\}$ is space spanned by $\cdot$

$n = 1, 2, 3, \dots$ iterations (subspace grows)

e.g. GMRES, conjugate gradients, QMR, ...

- generate estimates $x_0, x_1, \dots$ converge on $A^{-1}b$

with residuals $r_i = b - Ax_i$

could also consider $e_i = A^{-1}b - x_i$, $r_i = Ae_i$

but 1) r_i can be generated in an iteration, e_i can't

2) sometimes e_i large but solution as good as can be expected, better to consider r_i

- how quickly do $\frac{\|r_n\|}{\|r_0\|}$ go to 0?

2-norm

"initial" residual, $\approx b$(?)

- idea:

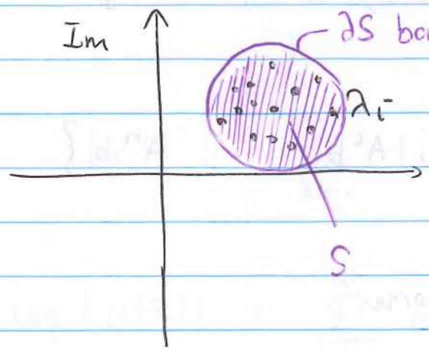
(*) $$\frac{\|r_n\|}{\|r_0\|} \approx \rho^n$$

geometric progression w/ rate ρ, which we get from potential theory spectrum, $\{\lambda_i\}$ set of eigenvalues of A in complex plane

② Link between ς and $\Lambda(A)$

assume we are given an S encircling $\Lambda(A)$; but picking S is part of the process!

- Let spectrum of A, $\Lambda(A)$, be approx. by a compact set $S \subset \mathbb{C}$, $\underline{0 \notin S}$!



- define polynomial $p(z) \in P_n$ on S

P_n: Set of poly. with degree at most n, norm. by $p(0) = 1$.

define

$$\|p\|_S = \max_{z \in S} |p(z)|$$

and

$$E_n = \min_{p \in P_n} \|p\|_S \qquad n = 1, 2, 3, \ldots$$

i.e. increasingly large degree polynomials, minimise their maxima on S

then

$\{E_n\}$ geometrically decrease with rate

$$\varsigma = \lim_{n \to \infty} \left(E_n(S)\right)^{1/n} \leq 1$$

"estimated asymptotic convergence rate"

same ς as ~~for~~ in $(*)$

strictly $<$ unless S completely surrounds the origin, separates it from point @ ∞

- how does this relate to GMRES?

Krylov subspace iterations deal with polynomials in A:

x_i live in $\{b, Ab, A^2b, \dots\}$

then

$r_n \in r_0 + \{Ab, A^2b, \dots, A^n b\}$

for $n=0$
to get correct norm:

equivalently,

$r_n = p(A) r_0$ with $p(0) = 1$,

so $p \in P_n$

GMRES tries to find $p(A)$ to minimise r_n, convergence will be quick if $p(A)$ is chosen to annihilate r_0 quickly.

intuitively, take A to be normal:

$A = UDU^*$, $D = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$
U unitary

then "smallness of $p(A)$":

$\|p(A)\| = \max_{z \in \Lambda(A)} |p(z)|$ $\Rightarrow \min_{p \in P_n} \|P(A)\| = \min_{p \in P_n} \left[\max_{z \in \Lambda(A)} \left(|p(z)| \right) \right] = \min_{p \in P_n} \|p\|_{\Lambda}$
matrix norm!

$\Lambda \approx S$

in general (for nonnormal A):

$\min_{p \in P_n} \|p(z)\|_{\Lambda(A)} \leq \min_{p \in P_n} \|p(A)\|$

therefore in general this would give us an upper bound on S, but more approximations to consider.

③ Deriving S via potential theory

- given set $S \subset \mathbb{C}$, $0 \notin S$,
- consider a polynomial

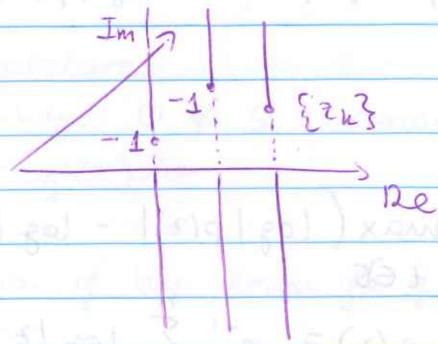
$$p(z) = \prod_{k=1}^{n} (z - z_k) \qquad \text{take abs}$$

$$|p(z)| = \prod_{k=1}^{n} |z - z_k| \qquad \text{take logs}$$

$$\log|p(z)| = \sum_{k=1}^{n} \log|z - z_k|$$

this is the potential associated with a collection of charges of strength -1 at $\{z_k\}$ in 2D (cylindrical symm)

prog:

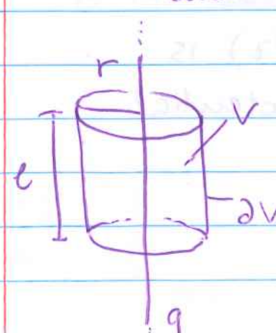


for a single, infinite line of charge, use gauss' theorem:

$$\int_{\partial V} \underline{E} \cdot \underline{\hat{n}}\, dS \overset{\text{Gauss}}{=} \int_V (\underline{\nabla} \cdot \underline{E})\, dV \overset{\text{use Maxwell I.}}{=} \int_V q\, dV$$

electric field strength; charge density; ignore units

also $\underline{E} = -\underline{\nabla}\varphi$ $\quad \varphi$ potential



$$E \cdot 2\pi r \cdot l = q\, l$$

q is charge per unit length

$$E \sim \frac{1}{r}$$

$$\varphi \sim \log r$$

r is radial distance from charge

(but could also just solve $\nabla^2 \varphi = 0$ in 2D)

- we wish to minimise

$\| \, |p(z)| / \underbrace{|p(0)|}_{} \, \|_S$ by ~~choosing~~ distributing the $\{z_k\}$ in $\mathbb{C}$

↳ ensures $p(0) = 1$.

this is equivalent to minimising $\| \log|p(z)| - \log|p(0)| \, \|_S$

which is (by maximum modulus principle) equivalent to minimising $\| \log|p(z)| - \log|p(0)| \, \|_{\partial S}$ ← boundary of S

i.e. we want

$$\min \Big[\max_{z \in \partial S} \big(\log|p(z)| - \log|p(0)| \big) \Big].$$

instead consider $n \to \infty$ and

$$g(z) = n^{-1} \sum_{k=1}^{n} \log|z - z_k| + C \qquad (**)$$

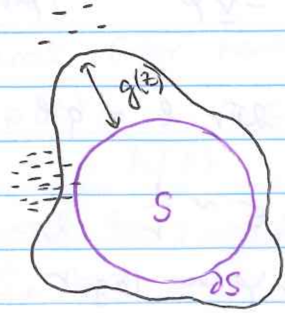
and seek $\min \Big[\max_{z \in S} g(z) - g(0) \Big]$

- so, physically: we want a ~~charge~~ distribution of charges $\{z_k\}$ that minimises their maximum potential on the boundary ∂S

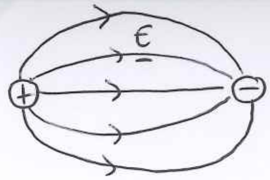
- but this happens if ∂S is the surface of a conductor (or many conductors connected electrically)

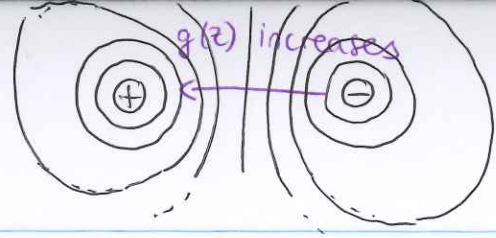
"prog":

radar plot g $g(z)$ on surface ∂S



$g(z)$ is potential





equipotential lines

−ve charge is $g(z)$ minimum

where $g(z)$ is small, lots of −ve charge
where $g(z)$ is large, not much −ve charge
therefore:

- we can always decrease the ~~minimal~~ max. potential on ∂S by drawing −ve charges away from low-$g(z)$ spots, therefore minimal max $g(z)$ on ∂S occurs if

$$g(z) = \tilde{C} \text{ on } \partial S$$

- minimal max $g(z)$ on ∂S (min $\tilde{C}$) occurs if <u>all</u> −ve charge is on ∂S.

~~physical picture:~~
- this is why $0 \notin S$: cannot shield origin

~~at $g(z)$ =~~

- explanation of the form of $g(z)$ in $(**)$:

  - n^{-1} at front means each charge has strength $-n^{-1}$, overall charge in the system is 1 (will see why later)

  - $+ C$ is to ensure $g(z)\big|_{\partial S} = \tilde{C} = 0$ by convention
    this is an arbitrary (gauge) choice

- if ∂S is surface of conductor(s), then mathematical task is to solve:

$$\nabla^2 g = 0 \quad \text{outside } S$$

boundary cond. $\left\{ \begin{array}{l} g\big|_{\partial S} = 0 \\ \text{as } |z| \to \infty, \; g \to \log|z| \quad \text{(due to unit charge in system)} \end{array} \right.$

- then, δ is given by:

$$\boxed{\delta = e^{-g(0)}}$$

proof:

from before, $\delta = \lim_{n \to \infty} (E_n(S))^{1/n}$

$$= \lim_{n\to\infty} \left\{ \min_{p \in P_n} \left[\max_{z \in S} \left(|p(z)| / |p(0)| \right) \right] \right\}^{1/n}$$

$$= \lim_{n\to\infty} \left\{ \min_{p} \left[\max_{z \in S} \left(e^{\sum_{k=1}^{n} \log|z - z_k| - \log|z_k|} \right) \right] \right\}^{1/n}$$

// use defn. of $g(z)$: $g(z) = n^{-1} \sum_{k=1}^{n} \log|z - z_k| + c$ //

$$= \lim_{n\to\infty} \left\{ \min_{p} \left[\max_{z \in S} \left(e^{n(g(z) - g(0))} \right) \right] \right\}^{1/n}$$

// but we saw that the minimal maximum of $g(z)$ occurs ~~~~ on ∂S, and ~~~~ when $g(z)$ on ∂S is 0 //

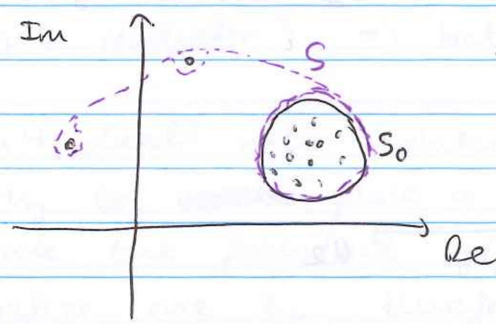
$$= \lim_{n\to\infty} \left\{ e^{-n g(0)} \right\}^{1/n} = e^{-g(0)}.$$

- caveats:

  - we took $n \to \infty$, but IRL, GMRES has a finite # of iterations $\Rightarrow$ finite # of charges. ρ estimated with a finite no. of charges cannot be worse than real ρ (because minimal max $g(z)$ on surface is not as small as it could be with $n \to \infty$, so ~~g(z) not as small as it could be~~ E_n are larger than they could be)

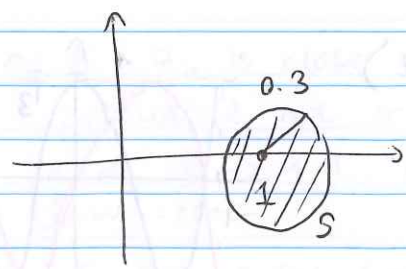
  - we took $S \approx \Lambda(A)$ instead of $\Lambda(A)$

  - charge distribution doesn't work for isolated charges, e.g.



so should treat these separately

- examples: 1) calculate ρ for disk S:



a) guess $p(z)$ with minimal max on S:

$$p(z) = |1-z|^n$$

$$p(0) = 1 \checkmark$$

$$\|p\|_S = 0.3^n$$

$$E_n \overset{?}{=} \|p\|_S = 0.3^n \Rightarrow \rho = 0.3$$

b) calculate from potential theory:

line of charge at $z = 1$ has potential

$$g(z) = \gamma_0 \log|1-z| + c$$

but we have $g\big|_{\partial S} = 0$ and $\lim\limits_{|z|\to\infty} g = \log|z|$

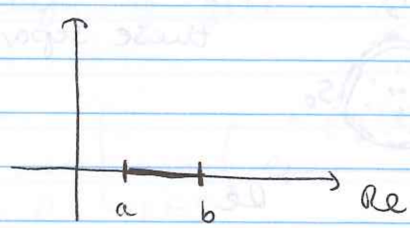
$$\Downarrow \qquad\qquad\qquad \Downarrow$$

$$c = -\log 0.3 \qquad\qquad \gamma_0 = 1$$

$$g(z) = \log|1-z| - \log 0.3$$

$$\Rightarrow \rho = e^{-g(0)} = 0.3 .$$

2) ~~calculate~~ Estimate ρ for S being a segment:

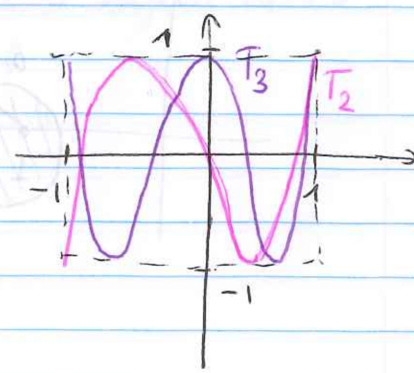


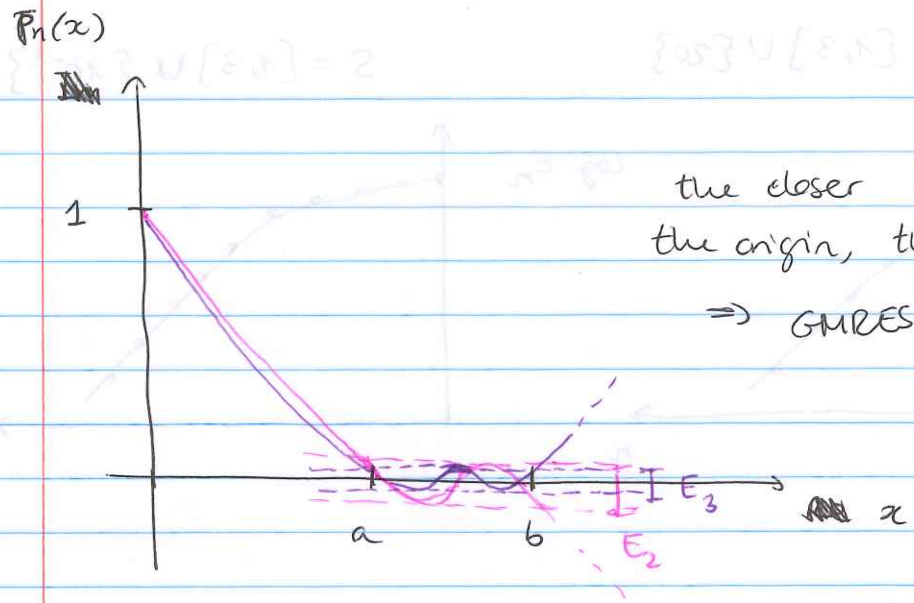
try Chebyshev polynomial T_n as $p(z)$

in $[-1, 1]$, $T_n(z) = \cos(n \arccos z)$

outside of $[-1, 1]$, $T_n(z) \sim \cosh$

$\Rightarrow$ blows up





the closer $[a,b]$ is to the origin, the larger the E_n, ρ

$\Rightarrow$ GMRES is slower

- further comments:

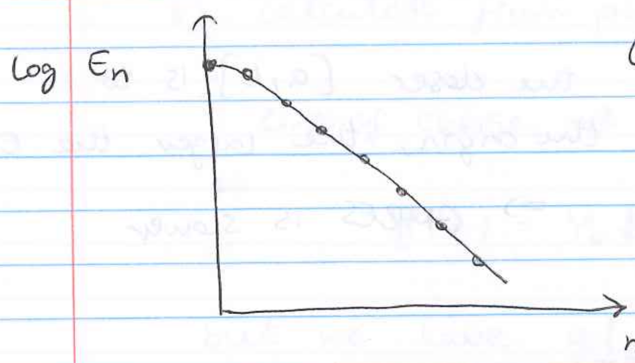
  - the smaller and further S is from the origin, the smaller ρ (think in terms of a larger $g(0)$, because we're further away from a -vely charged conductor) $\Rightarrow$ better convergence

  - haven't dealt with isolated eigenvalues: clearly to place a charge there to calculate the potential $g(z)$, we have to sacrifice one z_k, therefore at least one iteration. But not enough:

    ~ if z_k is close to σ, then it will decrease potential at o a lot, therefore $g(0) \to 0$ and $\rho \to 1$! Bad convergence for at least a few iters.

    ~ if z_k is far from σ, barely makes a difference to $g(0)$ and overall ρ:

$S = [1,3] \cup \{20\}$



$S = [1,3] \cup \{10^{-3}\}$