Numerical methods for matrix functions

SF2524 - Matrix Computations for Large-scale Systems

Lecture 15: Krylov methods for matrix functions

Problem

In this lecture we wish to compute

$$f(A)b,$$

where $A \in \mathbb{R}^{n \times n}$ is a given large sparse matrix.

* Derive on Black board *

Cauchy integral definition leads to

$$f(A)b \quad = \quad \left(\frac{-1}{2i\pi} \oint_\Gamma f(z)(A - zI)^{-1}\, dz \right) b = \frac{-1}{2i\pi} \oint_\Gamma f(z)(A - zI)^{-1} b\, dz$$

How do we compute?

$$(A - zI)^{-1} b \qquad (\star)$$

Note: $(\star)$ is a shifted linear system of equations:

$$(A - zI)x = b.$$

We will solve the shifted linear system using an Arnoldi method.

The rest of this lecture

1. Arnoldi's method for shifted systems
2. GMRES-variant (FOM) for shifted systems
3. Use Cauchy definition $\Rightarrow$ Krylov method for matrix functions
4. Application to exponential integrators

Shift invariance of Krylov subspaces

$$\mathcal{K}_n(A, b) = \mathcal{K}_n(A - \sigma I, b)$$

Proof idea: Find a non-singular R such that
$[b, \ldots, A^{n-1}b]R = [b, (A - \sigma I)b, \ldots, (A - \sigma I)^n b]$

Recall: $W = VR$ and R non-singular and $w_1, \ldots, w_m$ linear independent
$\Rightarrow \text{span}(w_1, \ldots, w_m) = \text{span}(v_1, \ldots, v_m)$

What happens with the Arnoldi factorization? *On black board*

Arnoldi factorization for a shifted matrix

Suppose we have an Arnoldi factorization

$$AQ_m = Q_{m+1}\underline{H}_m \qquad (\star)$$

Lemma

Suppose $Q_m \in \mathbb{C}^{n\times m}$, $\underline{H}_m \in \mathbb{C}^{(m+1)\times m}$ is an Arnoldi factorization $(\star)$ associated with $\mathcal{K}_m(A, b)$. Then, for any $\sigma \in \mathbb{C}$, $Q_m \in \mathbb{C}^{n\times m}$ and $\underline{H}_m - \sigma I_{m+1,m}$ is an Arnoldi factorization associated with $\mathcal{K}_m(A - \sigma I, b)$,

$$(A - \sigma I)Q_m = Q_{m+1}(\underline{H}_m - \sigma I_{m+1,m}). \qquad (\star\star)$$

where

$$I_{m+1,m} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{(m+1)\times m}.$$

FOM - almost GMRES for linear system

We now wish to solve linear systems:

$$Cx = b$$

(where we later set $C = A - \sigma I$.)

Derive on blackboard

Full Orthogonalization Method (FOM)

- Compute an Arnoldi factorization $AQ_n = Q_{n+1}\underline{H}_n$
- Compute `z=H(1:n,1:n)\e1` $\Leftrightarrow z = H_n^{-1} e_1$
- Compute approximation $\tilde{x} = Q_n z \|b\|$

Only slight difference in GMRES z=H(1:n+1,1:n)\e1.
Convergence very similar to GMRES.

Relationship with GMRES

- GMRES corresponds to $(AQ_n)^T (A\tilde{x} - b) = 0$ (lecture 8)
- FOM corresponds to $Q_n^T (A\tilde{x} - b) = 0$

Now consider shifted system:

$$(A - \sigma I)x = b$$

FOM for shifted systems

Compute an Arnoldi factorization $AQ_n = Q_{n+1}\underline{H}_n$

Compute `z=(H(1:n,1:n)-`σI`)\e1` $\Leftrightarrow z = (H_n - zI)^{-1}e_1$

Compute approximation $\tilde{x} = Q_n z\|b\|$

Note: Step 1 is independent of σ and the Step 2-3 can be done for many σ without carrying out Arnoldi method:

$$x \approx \tilde{x} = Q_n(H_n - \sigma I)^{-1}e_1\|b\|.$$

On black board

Cauchy integral definition and use FOM-approximation:

$$\begin{aligned} f(A)b &= \frac{-1}{2i\pi} \oint_\Gamma f(z)(A - zI)^{-1} b \, dz \\ &\approx \frac{-1}{2i\pi} \oint_\Gamma f(z) Q_n (H_n - zI)^{-1} e_1 \|b\| \, dz \\ &= Q_n \frac{-1}{2i\pi} \oint_\Gamma f(z)(H_n - zI)^{-1} \, dz \, e_1 \|b\| \\ &= Q_n f(H_n) e_1 \|b\| \end{aligned}$$

Krylov approximation of matrix functions

$$f(A)b \approx f_n = Q_n f(H_n) e_1 \|b\|$$

Error analysis of Krylov approximation

On black board

Theorem

Suppose $A \in \mathbb{C}^{n\times n}$ is a normal matrix and suppose $\Omega \subset \mathbb{C}$ is a convex compact set such that $\lambda(A) \subset \Omega$. Let f_m be the Krylov approximation of $f(A)b$. Then,

$$\|f(A)b - f_m\| \leq 2\|b\| \min_{p\in P_{m-1}} \max_{z\in\Omega} |f(z) - p(z)|.$$

Favorable situations (fast convergence):

- $f(z)$ can be well approximated with low-order polynomials
- $\lambda(A)$ and $\lambda(H_k)$ are clustered such that Ω can be chosen small. (Note. Not relative clustering)

* Examples *

Application to exponential integrators
PDF lecture notes 4.4.3

We already know that the initial value problem

$$y'(t) = Ay(t), \;\; y(0) = y_0$$

has the solution

$$y(t) = \exp(tA)y_0.$$

What about more general ODEs?

Problem

We wish to numerically solve the initial value problem using matrix functions:

$$y'(t) = g(y(t)), \;\; y(0) = y_0.$$

Lemma (Explicit solution linear inhomogeneous ODE)

The linear inhomogeneous ODE with right-hand side $g(y) = g_1(y) := Ay + b$ *and*

$$y'(t) = Ay(t) + b = g_1(y(t)), \quad y(0) = y_0, \tag{1}$$

has a solution explicitly given by

$$y(t) = y_0 + t\varphi(tA)g_1(y_0). \tag{2}$$

The matrix function φ is called a φ-function

$$\phi(z) = \frac{e^z - 1}{z}$$

* plot phi-function + matlab demo *

Substitute in the nonlinear problem and repeat reasoning:

Definition (Forward Euler exponential integrator)

Let $0 = t_0 < t_1 < \cdots < t_N$. The *forward Euler exponential integrator* generate the approximations $y_k \approx y(t_k)$, $k =, \ldots, N$ defined as

$$y_{k+1} = y_k + h_k \varphi(h_k A_k) g(y_k) \tag{3}$$

where $h_k = t_{k+1} - t_k$ and $A_k := g'(y_k)$.

Properties:

- Exact for the linear inhomogeneous case (1), and one step can be proven to be second order in h in the general case.
- Requires the computation of $\varphi(hA)g(y_k)$ in every step. Suitable to be used with matrix functions.

Step-length trade-off

We want

Trade-off of time-step h

- small $h \Rightarrow$ small Krylov error;
- small $h \Rightarrow$ small time-stepping error; but
- large h, because to reach a specific time-point quicker.

In practice: Try to balance Krylov error and time-step error with error estimates and increase to specific tolerance.

More elaborate example in Lecture notes PDF.

It's been a pleasure to teach this course. Thanks!

Exam preparation information

- "Sometimes you need to work harder not smarter"
  $\Rightarrow$ Solve many problems as preparation:
  - old exams
  - selected wiki problems 2016
  - selected wiki problems 2017 (will come soon)
  - Lanczos quiz
- Read problem formulation carefully:
  e.g. "Show" means "prove" (not matlab code)
- Correction more strict than wiki correction
- No calculator, notes, phones, books, etc allowed

Good luck on the exam

Please fill out the course evaluation (later)