Exponential Integrators using Matrix Functions: Krylov Subspace Methods and Chebyshev Expansion approximations

The HPC Approach

México, February 2018

Marlon Brenes
brenesnm@tcd.ie
https://www.tcd.ie/Physics/research/groups/quSys/people/Navarro/
Outline

• Exponential Integrators

• Brief Introduction: Chebyshev expansion for matrix functions

• Brief Introduction: Krylov subspace techniques

• HPC Approach:
  • Relevance and importance of an HPC approach
  • Parallelisation strategy

• Outlook
Exponential Integrators

Matrix functions
The problem

• Consider a problem of the type

\[ \frac{dw(t)}{dt} = Aw(t), \ t \in [0, T] \]

\[ w(0) = v, \ initial\ condition \]

• Its analytic solution is \( w(t) = e^{tA}v \)

• Things to consider:

  • \( A \) is a matrix

  • The exponential of the matrix is not really required, but merely it’s action on the vector \( v \)
Exponential Integrators

- Mathematical models of many physical, biological and economic processes systems of linear, constant-coefficient ordinary differential equations

- Growth of microorganisms, population, decay of radiation, control engineering, signal processing...

- More advanced: MHD (magnetohydrodynamics), quantum many-body problems, reaction-advection-diffusion equations...

Numerical approach

• The idea to use exponential functions for the matrix is not new
  • Was considered impractical…

• The development of Krylov subspace techniques to the action of the matrix exponential substantially changed the landscape

• Different types of solution evaluation for matrix exponentials:
  • ODE methods: numerical integration
  • Polynomial methods
  • Matrix decomposition methods

Numerical approach

• We’re interested in the case where $A$ is large and sparse

• A sole implementation may not be reliable for all types of problems

• Chebyshev expansion approach

• The technique of Krylov subspaces has been proven to be very efficient for many classes of problems

• Convergence is faster than applying the solution to linear systems in both techniques

Take-home message #1: There’s a big number of problems in science and engineering that can be tackled using exponential integrators and matrix functions.
Polynomial approximation: Chebyshev expansion
A brief introduction
Definition

• We intend to employ a polynomial expansion as an approximation

• Let us start with a definition of the matrix exponential by convergent power series:

\[ e^{tA} = I + tA + \frac{t^2 A^2}{2!} + \ldots \]

• An effective computation of the action of this operator on a vector is the main topic of this talk

Chebyshev expansion

• We intend to employ a **good** converging polynomial expansion

• Explicit computation of $e^{At}$ **has** to be avoided

• **Key component:** Efficient matrix-vector product operations

  • **Upside:** Efficient parallelisation and vectorisation, extremely simple approach

  • **Downside:** Requires computation of two eigenvalues, not so versatile
Chebyshev polynomials

- The polynomials are defined by a three-term recursion relationship in the interval $x \in [-1, 1]$: 
  
  $T_0(x) = 1$
  
  $T_1(x) = x$
  
  $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$

- $T_n(x)$ constitutes an orthogonal basis, therefore one can write:
  
  $f(x) = \sum_{n=0}^{\infty} b_n T_n(x) \approx \sum_{n=0}^{N} b_n T_n(x)$

- with $b_n = \frac{2 - \delta_n}{\pi} \int_{-1}^{1} \frac{f(y)T_n(y)}{\sqrt{1 - y^2}} dy$

Bessel coefficients for the particular case of the exponential function


Chebyshev polynomials

• We’re interested in applying the approximation to the action of the operator on a vector
• First step: Find the extremal eigenvalues of $\mathbf{A}$, more on this later

$$\lambda_{\text{min}} \quad \lambda_{\text{max}}$$

• Second step: Rescale operator such that it’s spectrum is bounded by $[-1,1]$

$$\mathbf{A}' = 2 \frac{\mathbf{A} - \lambda_{\text{min}} \mathbf{I}}{\lambda_{\text{max}} - \lambda_{\text{min}}} - \mathbf{I}$$

• Third step: Use the Chebyshev recursion relation

$$f(t\mathbf{A}')\mathbf{v} = e^{t\mathbf{A}'}\mathbf{v} \approx \sum_{n=0}^{N} b_n T_n(t\mathbf{A}')\mathbf{v}$$

• The recursion can be truncated up to a desired tolerance

Chebyshev recursion relation

• The recursion relation

\[ f(tA')v = e^{tA'}v \approx \sum_{n=0}^{N} b_n T_n(tA')v \]

• Goes as follows:

\[ \phi_0 = v \]
\[ \phi_1 = tA'v \]
\[ \phi_{n+1} = 2tA'\phi_n - \phi_{n-1} \]

• Then:

\[ f(tA')v \approx \sum_{n=0}^{N} b_n \phi_n \]

• Until desired tolerance

Chebyshev polynomials

- Why do we choose the Chebyshev polynomials as basis set? Why not another polynomial set?

- Because of the asymptotic property of the Bessel function!

  - When the order \( n \) of the polynomial becomes larger than the argument, the function decays \textbf{exponentially} fast

  - This means that in order to obtain a good approximation, an exponentially decreasing amount of terms are required in the expansion as a function of the argument (related to \( t, \lambda_{min} \) and \( \lambda_{max} \))

Take-home message #2:
The Chebyshev expansion approach provides a numerically stable and scalable approach at the cost of some restrictions of the problem.
Krylov subspace techniques to evaluate the solution

A brief introduction
Krylov subspace techniques

• We intend to employ a combination of a Krylov subspace technique and other known methods for matrix exponential

• Explicit computation of $e^{At}$ has to be avoided

• **Key component**: Efficient matrix-vector product operations

• **Upside**: Extremely versatile

• **Downside**: Storage of the subspace for large problems, “time scales”
Main idea

• Building a Krylov subspace of dimension $m$

$$\mathcal{K}_m = \text{span}\{v, Av, A^2v, \ldots, A^{m-1}v\}$$

• The idea is to approximate the solution to the problem by an element of $\mathcal{K}_m$

• In order to manipulate the subspace, it’s convenient to generate an orthonormal basis

$$V_m = [v_1, v_2, \ldots, v_m] \quad v_1 = v / \| v \|_2$$

• This can be achieved with the **Arnoldi algorithm**
Algorithm: Arnoldi

1. Initialize: Compute $v_1 = v/\|v\|_2$.
2. Iterate: Do $j = 1, 2, ..., m$
   (a) Compute $w := Av_j$
   (b) Do $i = 1, 2, \ldots, j$
      i. Compute $h_{i,j} := (w, v_i)$
      ii. Compute $w := w - h_{i,j}v_i$
   (c) Compute $h_{j+1,j} := \|w\|_2$ and $v_{j+1} := w/h_{j+1,j}$.

- Step 2-b is a modified Gram-Schmidt process.
- **Lanczos** can be applied for the case of symmetric matrices
Krylov subspace techniques

• The Arnoldi procedure produces a basis $V_m$ of $\mathcal{K}_m$ and an upper Hesseberg matrix $H_m$ of dimension $m \times m$ with coefficients $h_{ij}$

• We start by the relation given by

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_T^m$$

• Where $v_{m+1}$ satisfies $V_m^T v_{m+1} = 0$ and $e_m \in I_m$

• From which we obtain

$$H_m = V_m^T A V_m$$

• Therefore, $H_m$ is the projection of the linear transformation $A$ onto the Krylov subspace $\mathcal{K}_m$ with respect to the basis $V_m$

Gallopoulos, E. Saad, Y. ICS. 89′, 17—28 (1989).
Krylov subspace techniques

- Given that $H_m$ is a projection of the linear operator, an approximation can be made such that

$$e^{tA}v \approx \|v\|_2 V_m e^{tH_m} e_1$$

- The approximation is exact when the dimension of the Krylov subspace is equal to the dimension of the linear transformation

- Error

$$\|e^{tA}v - \|v\|_2 V_m e^{tH_m} e_1\| \leq 2\|v\|_2 \frac{(t\|A\|_2)^m e^{t\|A\|_2}}{m!}$$

Gallopoulos, E. Saad, Y. ICS. 89′, 17—28 (1989).
Krylov subspace techniques

• With this approach:
  • A **large sparse** matrix problem is approximated by a **small dense** matrix problem
  
• There are several methods to evaluate the small dense matrix exponential
  
  • Series methods, ODE methods, diagonalisation, matrix decomposition methods…
  
  • Padè methods

• This method can be used to compute $\lambda_{\text{min}}$ and $\lambda_{\text{max}}$ for the Chebyshev approach, very effective

Take-home message #3: Krylov subspace methods to evaluate the solution provides a more versatile and less restricted approach, at the expense of higher computational cost and memory consumption.
HPC Approach
Relevance and importance

• A platform for numerical calculations
  • Important for current research
• Undertake simulations efficiently
• Quantum physics, CFD, finite-element methods…
• Establish an instance where HPC approach has been used recently in scientific research
Parallelisation strategy

\[ y = Ax \]

- The linear operator is **sparse** (low density) and **huge** (big dimension, i.e., large amount of degrees of freedom)

- Approach:
  - Distribute the operator among processing elements
Take-home message #4:
Your numerical evaluation of the exponential integrator using matrix functions is only going to be as good as your implementation of the sparse matrix-vector product.
Outlook

- Two different approximation methods to target the solutions to matrix exponentials, which can be interpreted analytically as solutions to a particular set of differential equations

- We will discuss the practical approach using techniques, practices and libraries from the HPC perspective next session…
Thank you!