Acceleration techniques for exponential integrators

Marina Popolizio
Dipartimento di Matematica, Università di Bari

popolizio@dm.uniba.it

Joint work with Valeria Simoncini, Università di Bologna

Problem

We want to approximate

$$y = \exp(A)v$$

with

- A large, symmetric, negative semidefinite
- $\sigma(A) \subset [\alpha, 0], \quad |\alpha| \gg 1$
- $v \in \mathbb{R}^n, \quad ||v|| = 1$

Problem

We want to approximate

$$y = \exp(A)v$$

with

- A large, symmetric, negative semidefinite
- $\sigma(A) \subset [\alpha, 0], \qquad |\alpha| \gg 1$
- $v \in \mathbb{R}^n, \quad ||v|| = 1$

Problem

We want to approximate

$$y = \exp(A)v$$

with

- A large, symmetric, negative semidefinite
- $\sigma(A) \subset [\alpha, 0], \qquad |\alpha| \gg 1$
- $v \in \mathbb{R}^n, \qquad ||v|| = 1$

Problem

We want to approximate

$$y = \exp(A)v$$

with

- A large, symmetric, negative semidefinite
- $\sigma(A) \subset [\alpha, 0], \qquad |\alpha| \gg 1$
- $v \in \mathbb{R}^n, \quad ||v|| = 1$

Applications

- exponential integrators for ODEs and PDEs
- dinamical systems
- control theory
- Markov chains

Applications

- exponential integrators for ODEs and PDEs
- dinamical systems
- control theory
- Markov chains

Applications

- exponential integrators for ODEs and PDEs
- dinamical systems
- control theory
- Markov chains

Applications

- exponential integrators for ODEs and PDEs
- dinamical systems
- control theory
- Markov chains

Example

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t)$$

$$u(x,0) = u_0, \; x \in \Omega$$

$$u(x,t) = \sigma(x), \; x \in \partial\Omega, \; t > 0$$

with $\mathcal{L}$ second order parabolic differential operator.

FD discretization in x

$$\frac{dw(t)}{dt} = Aw(t), \;\; w(0) = w_0$$

whose solution is

$$w(t) = \exp(tA)w_0$$

Example

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t)$$

$$u(x,0) = u_0, \; x \in \Omega$$

$$u(x,t) = \sigma(x), \; x \in \partial\Omega, \; t > 0$$

with $\mathcal{L}$ second order parabolic differential operator.

FD discretization in x

$$\frac{dw(t)}{dt} = Aw(t), \quad w(0) = w_0$$

whose solution is

$$w(t) = \exp(tA)w_0$$

Example

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t)$$

$$u(x,0) = u_0, \; x \in \Omega$$

$$u(x,t) = \sigma(x), \; x \in \partial\Omega, \; t > 0$$

with $\mathcal{L}$ second order parabolic differential operator.

FD discretization in x

$$\frac{dw(t)}{dt} = Aw(t), \;\; w(0) = w_0$$

whose solution is

$$w(t) = \exp(tA)w_0$$

Example

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t)$$

$$u(x,0) = u_0,\ \ x \in \Omega$$
$$u(x,t) = \sigma(x),\ \ x \in \partial\Omega,\ \ t > 0$$

with $\mathcal{L}$ second order parabolic differential operator.

FD discretization in x

$$\frac{dw(t)}{dt} = Aw(t),\ \ w(0) = w_0$$

whose solution is

$$w(t) = \exp(tA)w_0$$

Example

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t)$$

$$u(x,0) = u_0,\ x \in \Omega$$

$$u(x,t) = \sigma(x),\ x \in \partial\Omega,\ t > 0$$

with $\mathcal{L}$ second order parabolic differential operator.

FD discretization in x

$$\frac{dw(t)}{dt} = Aw(t),\ \ w(0) = w_0$$

whose solution is

$$w(t) = \exp(tA)w_0$$

Example II

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t) + N(u,t)$$

$$u(x,0) = u_0, \;\; x \in \Omega$$

$$u(x,t) = \sigma(x), \;\; x \in \partial\Omega, \;\; t > 0$$

with $\mathcal{L}$ differential operator.

$u(t_{n-1}) = u_{n-1}, \;\; t_{n-1} = (n-1)\Delta t$

Lawson-Euler scheme

$$u_n = \exp(\Delta t \; L) \; (u_{n-1} + N(u_{n-1}, t_{n-1}))$$

Example II

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t) + N(u,t)$$

$$u(x,0) = u_0, \; x \in \Omega$$
$$u(x,t) = \sigma(x), \; x \in \partial\Omega, \; t > 0$$

with $\mathcal{L}$ differential operator.

$u(t_{n-1}) = u_{n-1}, \; t_{n-1} = (n-1)\Delta t$

Lawson-Euler scheme

$$u_n = \exp(\Delta t \; L) \; (u_{n-1} + N(u_{n-1}, t_{n-1}))$$

Example II

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t) + N(u,t)$$

$$u(x,0) = u_0, \; x \in \Omega$$
$$u(x,t) = \sigma(x), \; x \in \partial\Omega, \; t > 0$$

with $\mathcal{L}$ differential operator.

$u(t_{n-1}) = u_{n-1}, \; t_{n-1} = (n-1)\Delta t$

Lawson-Euler scheme

$$u_n = \exp(\Delta t \; L) \left(u_{n-1} + N(u_{n-1}, t_{n-1})\right)$$

Example II

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t) + N(u,t)$$

$$u(x,0) = u_0,\ x \in \Omega$$
$$u(x,t) = \sigma(x),\ x \in \partial\Omega,\ t > 0$$

with $\mathcal{L}$ differential operator.

$u(t_{n-1}) = u_{n-1},\ t_{n-1} = (n-1)\Delta t$

Lawson-Euler scheme

$$u_n = \exp(\Delta t\ L)\ (u_{n-1} + N(u_{n-1}, t_{n-1}))$$

Example II

Consider the problem

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t) + N(u,t)$$

$$u(x,0) = u_0,\ x \in \Omega$$
$$u(x,t) = \sigma(x),\ x \in \partial\Omega,\ t > 0$$

with $\mathcal{L}$ differential operator.

$u(t_{n-1}) = u_{n-1},\ t_{n-1} = (n-1)\Delta t$

Lawson-Euler scheme

$$u_n = \exp(\Delta t\ L)\ (u_{n-1} + N(u_{n-1}, t_{n-1}))$$

State of the art

- $\exp(A)v \approx g(A)v$

  with g approximation to exp

- $\exp(A)v \approx \exp(M)v$

State of the art

- $\exp(A)v \approx g(A)v$

  with g approximation to exp

- $\exp(A)v \approx \exp(M)v$

State of the art

- $\exp(A)v \approx g(A)v$

  with g approximation to exp

- $\exp(A)v \approx \exp(M)v$

  with M s.t.

  $M \approx A$

  $\exp(M)v$ easier to compute than $\exp(A)v$

State of the art

- $\exp(A)v \approx g(A)v$

  with g approximation to exp

- $\exp(A)v \approx \exp(M)v$

  with M s.t.

  $M \approx A$

  $\exp(M)v$ easier to compute than $\exp(A)v$

State of the art

- $\exp(A)v \approx g(A)v$

  with g approximation to exp

- $\exp(A)v \approx \exp(M)v$

  with M s.t.

  $M \approx A$

  $\exp(M)v$ easier to compute than $\exp(A)v$

Approximations to the exponential function

- Polynomial functions
- Rational functions

Approximations to the exponential function

- Polynomial functions
- Rational functions

Approximations to the exponential function

- Polynomial functions
- Rational functions

Partial fraction expansion

Let R_ν be a rational approximation to exp with ν distinct poles

$$y \approx R_\nu(A)v = \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v$$

$$(A - \xi_j I)x_j = v \qquad j = 1, \ldots, \nu$$

Partial fraction expansion

Let R_ν be a rational approximation to exp with ν distinct poles

$$y \;\approx\; R_\nu(A)v \;=\; \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v$$

$$(A - \xi_j I)x_j = v \qquad j = 1, \ldots, \nu$$

Partial fraction expansion

Let R_ν be a rational approximation to exp with ν distinct poles

$$y \;\approx\; R_\nu(A)v \;=\; \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v$$

$$(A - \xi_j I)x_j = v \qquad j = 1, \ldots, \nu$$

Approximations to A

- decompositions
- projection methods

Approximations to A

- decompositions

$$A = Q\Lambda Q^T \text{ then } \exp(A)v = Q\exp(\Lambda)Q^T v$$

- projection methods

Approximations to A

- decompositions

$$A = Q\Lambda Q^T \text{ then } \exp(A)v = Q\exp(\Lambda)Q^T v$$

- projection methods

Approximations to A

- decompositions

$$A = Q\Lambda Q^T \text{ then } \exp(A)v = Q\exp(\Lambda)Q^T v$$

- projection methods

Krylov subspace methods

Approximations to A

- decompositions

$$A = Q\Lambda Q^T \text{ then } \exp(A)v = Q\exp(\Lambda)Q^T v$$

- projection methods

Krylov subspace methods

Krylov subspace methods

$$\mathcal{K}_m(A, v) = \mathrm{span}\{v, Av, \ldots, A^{m-1}v\}$$

range(V_m) = $\mathcal{K}_m(A, v)$ $\qquad$ $V_m^* V_m = I$

T_m tridiagonal matrix

$$AV_m = V_m T_m + t_{m+1,m} v_{m+1} e_m^*$$

$$y = \exp(A)v \quad \approx \quad V_m \exp(H_m) e_1 = y_m$$

Krylov subspace methods

$$\mathcal{K}_m(A, v) = \mathrm{span}\{v, Av, \ldots, A^{m-1}v\}$$

range(V_m) = $\mathcal{K}_m(A, v)$ $\qquad V_m^* V_m = I$

T_m tridiagonal matrix

$$AV_m = V_m T_m + t_{m+1,m} v_{m+1} e_m^*$$

$$y = \exp(A)v \quad \approx \quad V_m \exp(H_m) e_1 = y_m$$

Krylov subspace methods

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

range(V_m) = $\mathcal{K}_m(A, v)$ $\qquad V_m^* V_m = I$

T_m tridiagonal matrix

$$AV_m = V_m T_m + t_{m+1,m} v_{m+1} e_m^*$$

$$y = \exp(A)v \quad \approx \quad V_m \exp(H_m) e_1 \equiv y_m$$

Problems with Krylov subspace methods:

The convergence depends on $\sigma(A)$ and can be slow

[Hochbruck-Lubich '97]

- problems for the memory requirements
- high computational costs

Problems with Krylov subspace methods:

The convergence depends on $\sigma(A)$ and can be slow

[Hochbruck-Lubich '97]

- problems for the memory requirements
- high computational costs

Problems with Krylov subspace methods:

The convergence depends on $\sigma(A)$ and can be slow

[Hochbruck-Lubich '97]

- problems for the memory requirements
- high computational costs

Our goal: *Acceleration techniques*

- *Shift-and-Invert*

  [Moret-Novati '04, van den Eshof-Hochbruck '06]

- *Partial Fraction Expansion*

- *Krylov plus Inverted Krylov*

  [Druskin-Knizhnerman '98, Simoncini '07]

- *Real valued* method for the linear systems of the partial fraction expansion
  [Axelsson-Kucherov '00]
  High computational costs

Our goal: *Acceleration techniques*

- *Shift-and-Invert*

  [Moret-Novati ’04, van den Eshof-Hochbruck ’06]

- *Partial Fraction Expansion*

- *Krylov plus Inverted Krylov*

  [Druskin-Knizhnerman ’98, Simoncini ’07]

- *Real valued* method for the linear systems of the partial fraction expansion
  [Axelsson-Kucherov ’00]
  High computational costs

Our goal: *Acceleration techniques*

- *Shift-and-Invert*

  [Moret-Novati ’04, van den Eshof-Hochbruck ’06]

- *Partial Fraction Expansion*

- *Krylov plus Inverted Krylov*

  [Druskin-Knizhnerman ’98, Simoncini ’07]

- *Real valued* method for the linear systems of the partial fraction expansion
  [Axelsson-Kucherov ’00]
  High computational costs

Our goal: *Acceleration techniques*

- *Shift-and-Invert*

  [Moret-Novati ’04, van den Eshof-Hochbruck ’06]

- *Partial Fraction Expansion*

- *Krylov plus Inverted Krylov*

  [Druskin-Knizhnerman ’98, Simoncini ’07]

- *Real valued* method for the linear systems of the partial fraction expansion
  [Axelsson-Kucherov ’00]
  High computational costs

Our goal: *Acceleration techniques*

- *Shift-and-Invert*

  [Moret-Novati '04, van den Eshof-Hochbruck '06]

- *Partial Fraction Expansion*

- *Krylov plus Inverted Krylov*

  [Druskin-Knizhnerman '98, Simoncini '07]

- *Real valued* method for the linear systems of the partial fraction expansion
  [Axelsson-Kucherov '00]
  High computational costs

Shift-and-Invert method

Fix $s > 0$ and apply the Lanczos method to $(I - sA)^{-1}$

$$(I - sA)^{-1} V_m = V_m T_m + \beta_m v_{m+1} e_m^*$$

T_m tridiagonal $\qquad V_m^* V_m = I$

$$y = \exp(A)v \quad \approx \quad y_{SI} \equiv V_m \exp\left(\frac{1}{s}(I - T_m^{-1})\right) e_1$$

Shift-and-Invert method

Fix $s > 0$ and apply the Lanczos method to $(I - sA)^{-1}$

$$(I - sA)^{-1} V_m = V_m T_m + \beta_m v_{m+1} e_m^*$$

T_m tridiagonal $\qquad V_m^* V_m = I$

$$y = \exp(A)v \quad \approx \quad y_{SI} \equiv V_m \exp\left(\frac{1}{s}(I - T_m^{-1})\right) e_1$$

Shift-and-Invert method

Fix $s > 0$ and apply the Lanczos method to $(I - sA)^{-1}$

$$(I - sA)^{-1} V_m = V_m T_m + \beta_m v_{m+1} e_m^*$$

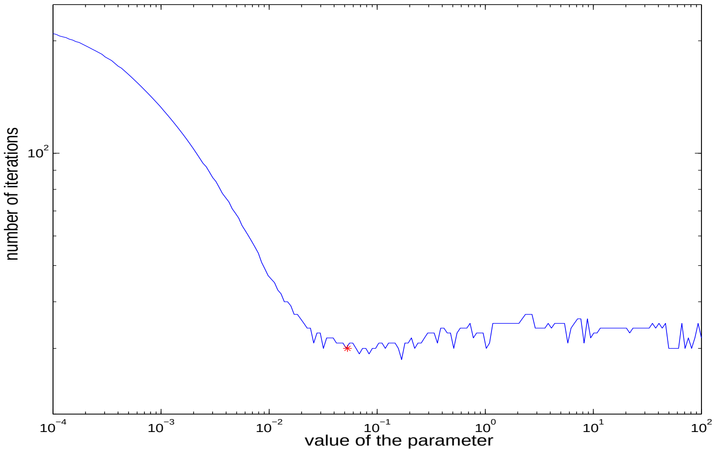
T_m tridiagonal $\qquad V_m^* V_m = I$

$$y = \exp(A)v \quad \approx \quad y_{SI} \equiv V_m \exp\left(\frac{1}{s}(I - T_m^{-1})\right) e_1$$

Numerical tests

D diagonal matrix 3375×3375, $\sigma(D) \subset [-2329.4, -22.597]$;

$b \in \mathbb{R}^{3375}, b = (1, ..., 1), \|b\| = 1$



Choice of the shift parameter

We combine

- Krylov subspace methods
- rational approximations to exp
- preconditioning

Choice of the shift parameter

We combine

- Krylov subspace methods
- rational approximations to exp
- preconditioning

Choice of the shift parameter

We combine

- Krylov subspace methods
- rational approximations to exp
- preconditioning

Choice of the shift parameter

We combine

- Krylov subspace methods
- rational approximations to exp
- preconditioning

Choice of the shift parameter

We combine

- Krylov subspace methods
- rational approximations to exp
- preconditioning

Equivalence of approaches:

[P. - Simoncini SIMAX'08]

- R_ν rational approximation to exp with ν distinct poles
- $y_{prec} = \tau_0 v + \sum_{j=1}^{\nu} \tau_j x_m^{(j)}$

  with $x_m^{(j)}$ solution of $(A - \frac{1}{s}I)^{-1}(A - \xi_j I)x = (A - \frac{1}{s}I)^{-1}v$

  in $\mathcal{K}_m((A - \frac{1}{s}I)^{-1}, v)$
- $y_{SI} = V_m R_\nu \left(\frac{1}{s}(I - T_m^{-1})\right) e_1$ [Moret-Novati, van den Eshof-Hochbruck]

then

$$y_{SI} = y_{prec}$$

Equivalence of approaches:

[P. - Simoncini SIMAX'08]

- R_ν rational approximation to exp with ν distinct poles

- $y_{prec} = \tau_0 v + \sum_{j=1}^{\nu} \tau_j x_m^{(j)}$

  with $x_m^{(j)}$ solution of $(A - \frac{1}{s}I)^{-1}(A - \xi_j I)x = (A - \frac{1}{s}I)^{-1}v$

  in $\mathcal{K}_m((A - \frac{1}{s}I)^{-1}, v)$

- $y_{SI} = V_m R_\nu\left(\frac{1}{s}(I - T_m^{-1})\right)e_1$ [Moret-Novati, van den Eshof-Hochbruck]

then

$$y_{SI} = y_{prec}$$

Equivalence of approaches:

[P. - Simoncini SIMAX'08]

- R_ν rational approximation to exp with ν distinct poles
- $y_{prec} = \tau_0 v + \sum_{j=1}^{\nu} \tau_j x_m^{(j)}$

  with $x_m^{(j)}$ solution of $(A - \frac{1}{s}I)^{-1}(A - \xi_j I)x = (A - \frac{1}{s}I)^{-1}v$

  in $\mathcal{K}_m\left((A - \frac{1}{s}I)^{-1}, v\right)$

- $y_{SI} = V_m R_\nu\left(\frac{1}{s}(I - T_m^{-1})\right)e_1$ [Moret-Novati, van den Eshof-Hochbruck]

then

$$y_{SI} = y_{prec}$$

Equivalence of approaches:

[P. - Simoncini SIMAX'08]

- R_ν rational approximation to exp with ν distinct poles

- $y_{prec} = \tau_0 v + \sum_{j=1}^{\nu} \tau_j x_m^{(j)}$

  with $x_m^{(j)}$ solution of $(A - \frac{1}{s}I)^{-1}(A - \xi_j I)x = (A - \frac{1}{s}I)^{-1}v$

  in $\mathcal{K}_m((A - \frac{1}{s}I)^{-1}, v)$

- $y_{SI} = V_m R_\nu\left(\frac{1}{s}(I - T_m^{-1})\right)e_1$ [Moret-Novati, van den Eshof-Hochbruck]

then

$$y_{SI} = y_{prec}$$

Equivalence of approaches:

[P. - Simoncini SIMAX'08]

- R_ν rational approximation to exp with ν distinct poles
- $y_{prec} = \tau_0 v + \sum_{j=1}^{\nu} \tau_j x_m^{(j)}$

  with $x_m^{(j)}$ solution of $(A - \frac{1}{s}I)^{-1}(A - \xi_j I)x = (A - \frac{1}{s}I)^{-1}v$

  in $\mathcal{K}_m\big((A - \frac{1}{s}I)^{-1}, v\big)$
- $y_{SI} = V_m R_\nu\left(\frac{1}{s}(I - T_m^{-1})\right)e_1$ [Moret-Novati, van den Eshof-Hochbruck]

then

$$y_{SI} = y_{prec}$$

Choice of the shift parameter

We have detected $s_{opt} = \frac{1}{p_{opt}}$ with p_{opt} s.t.

$$(A - p_{opt} I)^{-1}(A - \xi_j I)x = (A - p_{opt} I)^{-1} v$$

converges faster than

$$(A - \xi_j I)x = v, \quad \forall j = 1, \ldots, \nu$$

$\therefore s_{opt} = \frac{1}{|\xi_1|}$, ξ_1 pole with the largest imaginary part

Optimal values of the shift parameter

ν	3	4	5	6	7	8
s_{opt}	0.4134	0.2720	0.1988	0.1551	0.1264	0.1062
ν	9	10	11	12	13	14
s_{opt}	0.0914	0.0801	0.0711	0.0639	0.0580	0.0530
ν	15	16	17	18	19	20
s_{opt}	0.0488	0.0452	0.0421	0.0394	0.0369	0.0348

Choice of the shift parameter

We have detected $\color{red}{s_{opt}} = \frac{1}{p_{opt}}$ with p_{opt} s.t.

$$(A - p_{opt}I)^{-1}(A - \xi_j I)x = (A - p_{opt}I)^{-1}v$$

converges faster than

$$(A - \xi_j I)x = v, \quad \forall j = 1, \ldots, \nu$$

$\therefore s_{opt} = \frac{1}{|\xi_1|}, \quad \xi_1$ pole with the largest imaginary part

Optimal values of the shift parameter

ν	3	4	5	6	7	8
s_{opt}	0.4134	0.2720	0.1988	0.1551	0.1264	0.1062
ν	9	10	11	12	13	14
s_{opt}	0.0914	0.0801	0.0711	0.0639	0.0580	0.0530
ν	15	16	17	18	19	20
s_{opt}	0.0488	0.0452	0.0421	0.0394	0.0369	0.0348

Choice of the shift parameter

We have detected $s_{opt} = \frac{1}{p_{opt}}$ with p_{opt} s.t.

$$(A - p_{opt} I)^{-1}(A - \xi_j I)x = (A - p_{opt} I)^{-1} v$$

converges faster than

$$(A - \xi_j I)x = v, \ \ \forall j = 1, \ldots, \nu$$

$\therefore s_{opt} = \frac{1}{|\xi_1|}$, ξ_1 pole with the largest imaginary part

Optimal values of the shift parameter

ν	3	4	5	6	7	8
s_{opt}	0.4134	0.2720	0.1988	0.1551	0.1264	0.1062
ν	9	10	11	12	13	14
s_{opt}	0.0914	0.0801	0.0711	0.0639	0.0580	0.0530
ν	15	16	17	18	19	20
s_{opt}	0.0488	0.0452	0.0421	0.0394	0.0369	0.0348

Choice of the shift parameter

We have detected $s_{opt} = \frac{1}{p_{opt}}$ with p_{opt} s.t.

$$(A - p_{opt} I)^{-1}(A - \xi_j I)x = (A - p_{opt} I)^{-1} v$$

converges faster than

$$(A - \xi_j I)x = v, \ \ \forall j = 1, \ldots, \nu$$

$\therefore s_{opt} = \frac{1}{|\xi_1|}$, ξ_1 pole with the largest imaginary part

Optimal values of the shift parameter

ν	3	4	5	6	7	8
s_{opt}	0.4134	0.2720	0.1988	0.1551	0.1264	0.1062
ν	9	10	11	12	13	14
s_{opt}	0.0914	0.0801	0.0711	0.0639	0.0580	0.0530
ν	15	16	17	18	19	20
s_{opt}	0.0488	0.0452	0.0421	0.0394	0.0369	0.0348

Choice of the shift parameter

We have detected $s_{opt} = \frac{1}{p_{opt}}$ with p_{opt} s.t.

$$(A - p_{opt} I)^{-1}(A - \xi_j I)x = (A - p_{opt} I)^{-1} v$$

converges faster than

$$(A - \xi_j I)x = v, \;\; \forall j = 1, \ldots, \nu$$

$\therefore s_{opt} = \frac{1}{|\xi_1|}$, ξ_1 pole with the largest imaginary part

Optimal values of the shift parameter

ν	3	4	5	6	7	8
s_{opt}	0.4134	0.2720	0.1988	0.1551	0.1264	0.1062
ν	9	10	11	12	13	14
s_{opt}	0.0914	0.0801	0.0711	0.0639	0.0580	0.0530
ν	15	16	17	18	19	20
s_{opt}	0.0488	0.0452	0.0421	0.0394	0.0369	0.0348

Krylov plus Inverted Krylov

[Druskin-Knizhnerman '98, Simoncini '07]

Fix k and m

$$\mathcal{K}_{k,m}(A, v) = \text{span}\{A^{-k+1}v, \ldots, A^{-1}v, v, Av, \ldots, A^{m-1}v\}$$

$$\mathcal{V}_m = [V_1, \ldots, V_m] \in \mathbb{R}^{n \times 2m}$$

with $V_i \in \mathbb{R}^{n \times 2}$

$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m \in \mathbb{R}^{2m \times 2m}$

$$\exp(A)v \approx \mathcal{V}_m \exp(\mathcal{T}_m) e_1$$

$$Aw_{i+1} = w_i, \quad i = 1, \ldots, k-1$$

Krylov plus Inverted Krylov

[Druskin-Knizhnerman ’98, Simoncini ’07]

Fix k and m

$$\mathcal{K}_{k,m}(A, v) = \mathrm{span}\{A^{-k+1}v, \ldots, A^{-1}v, v, Av, \ldots, A^{m-1}v\}$$

$$\mathcal{V}_m = [V_1, \ldots, V_m] \in \mathbb{R}^{n\times 2m}$$

with $V_i \in \mathbb{R}^{n\times 2}$

$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m \in \mathbb{R}^{2m\times 2m}$

$$\exp(A)v \approx \mathcal{V}_m \exp(\mathcal{T}_m)e_1$$

$$Aw_{i+1} = w_i, \quad i = 1, \ldots, k-1$$

Krylov plus Inverted Krylov

[Druskin-Knizhnerman '98, Simoncini '07]

Fix k and m

$$\mathcal{K}_{k,m}(A, v) = \mathrm{span}\{A^{-k+1}v, \ldots, A^{-1}v, v, Av, \ldots, A^{m-1}v\}$$

$$\mathcal{V}_m = [V_1, \ldots, V_m] \in \mathbb{R}^{n\times 2m}$$

with $V_i \in \mathbb{R}^{n\times 2}$

$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m \in \mathbb{R}^{2m\times 2m}$

$$\exp(A)v \approx \mathcal{V}_m \exp(\mathcal{T}_m)e_1$$

$$Aw_{i+1} = w_i, \quad i = 1, \ldots, k-1$$

Krylov plus Inverted Krylov

[Druskin-Knizhnerman '98, Simoncini '07]

Fix k and m

$$\mathcal{K}_{k,m}(A,v) = \operatorname{span}\{A^{-k+1}v, \dots, A^{-1}v, v, Av, \dots, A^{m-1}v\}$$

$$\mathcal{V}_m = [V_1, \dots, V_m] \in \mathbb{R}^{n\times 2m}$$

with $V_i \in \mathbb{R}^{n\times 2}$

$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m \in \mathbb{R}^{2m\times 2m}$

$$\exp(A)v \approx \mathcal{V}_m \exp(\mathcal{T}_m) e_1$$

$$Aw_{i+1} = w_i, \quad i = 1, \dots, k-1$$

Krylov plus Inverted Krylov

[Druskin-Knizhnerman '98, Simoncini '07]

Fix k and m

$$\mathcal{K}_{k,m}(A, v) = \mathrm{span}\{A^{-k+1}v, \ldots, A^{-1}v, v, Av, \ldots, A^{m-1}v\}$$

$$\mathcal{V}_m = [V_1, \ldots, V_m] \in \mathbb{R}^{n\times 2m}$$

with $V_i \in \mathbb{R}^{n\times 2}$

$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m \in \mathbb{R}^{2m\times 2m}$

$$\exp(A)v \approx \mathcal{V}_m \exp(\mathcal{T}_m)e_1$$

$$Aw_{i+1} = w_i, \quad i = 1, \ldots, k-1$$

Krylov plus Inverted Krylov

[Druskin-Knizhnerman '98, Simoncini '07]

Fix k and m

$$\mathcal{K}_{k,m}(A, v) = \text{span}\{A^{-k+1}v, \ldots, A^{-1}v, v, Av, \ldots, A^{m-1}v\}$$

$$\mathcal{V}_m = [V_1, \ldots, V_m] \in \mathbb{R}^{n\times 2m}$$

with $V_i \in \mathbb{R}^{n\times 2}$

$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m \in \mathbb{R}^{2m\times 2m}$

$$\exp(A)v \approx \mathcal{V}_m \exp(\mathcal{T}_m)e_1$$

$$Aw_{i+1} = w_i, \quad i = 1, \ldots, k-1$$

Numerical Experiments

Consider $A \in \mathbb{R}^{10000 \times 10000}$ FD discretization of

$$\mathcal{L}(u) = (a(x,y)u_x)_x + (b(x,y)u_y)_y,$$

$$a(x,y) = 1 + y - x, \quad b(x,y) = 1 + x + x^2$$

$\sigma(A) \subset [-35424, -25.256]$

$y = \exp(\tau A)v, \ \tau = 0.1,$

v normally distributed random vector, $||v|| = 1$

Final accuracy $10^{-\nu}$

Numerical Experiments

Consider $A \in \mathbb{R}^{10000\times 10000}$ FD discretization of

$$\mathcal{L}(u) = (a(x,y)u_x)_x + (b(x,y)u_y)_y,$$

$$a(x,y) = 1 + y - x, \quad b(x,y) = 1 + x + x^2$$

$\sigma(A) \subset [-35424, -25.256]$

$y = \exp(\tau A)v, \ \tau = 0.1,$

v normally distributed random vector, $||v|| = 1$

Final accuracy $10^{-\nu}$

Numerical Experiments

Consider $A \in \mathbb{R}^{10000 \times 10000}$ FD discretization of

$$\mathcal{L}(u) = (a(x,y)u_x)_x + (b(x,y)u_y)_y,$$

$$a(x,y) = 1 + y - x, \quad b(x,y) = 1 + x + x^2$$

$\sigma(A) \subset [-35424, -25.256]$

$y = \exp(\tau A)v, \ \tau = 0.1,$

v normally distributed random vector, $||v|| = 1$

Final accuracy $10^{-\nu}$

Numerical Experiments

Consider $A \in \mathbb{R}^{10000 \times 10000}$ FD discretization of

$$\mathcal{L}(u) = (a(x,y)u_x)_x + (b(x,y)u_y)_y,$$

$$a(x,y) = 1 + y - x, \quad b(x,y) = 1 + x + x^2$$

$\sigma(A) \subset [-35424, -25.256]$

$y = \exp(\tau A)v, \; \tau = 0.1,$

v normally distributed random vector, $||v|| = 1$

Final accuracy $10^{-\nu}$

Numerical Experiments

Consider $A \in \mathbb{R}^{10000 \times 10000}$ FD discretization of

$$\mathcal{L}(u) = (a(x,y)u_x)_x + (b(x,y)u_y)_y,$$

$$a(x,y) = 1 + y - x, \quad b(x,y) = 1 + x + x^2$$

$\sigma(A) \subset [-35424, -25.256]$

$y = \exp(\tau A)v, \ \tau = 0.1,$

v normally distributed random vector, $||v|| = 1$

Final accuracy $10^{-\nu}$

We compare

- *PFE*: Computation of the Partial Fraction Expansion by explicitly solving each symmetric complex system
- *Standard Lanczos*: standard Lanczos approach
- *SI*: Shift-Invert Lanczos
- *KPIK*: Krylov Plus Inverted Krylov

We compare

- *PFE*: Computation of the Partial Fraction Expansion by explicitly solving each symmetric complex system

- *Standard Lanczos*: standard Lanczos approach

- *SI*: Shift-Invert Lanczos

- *KPIK*: Krylov Plus Inverted Krylov

We compare

- *PFE*: Computation of the Partial Fraction Expansion by explicitly solving each symmetric complex system
- *Standard Lanczos*: standard Lanczos approach
- *SI*: Shift-Invert Lanczos
- *KPIK*: Krylov Plus Inverted Krylov

We compare

- *PFE*: Computation of the Partial Fraction Expansion by explicitly solving each symmetric complex system
- *Standard Lanczos*: standard Lanczos approach
- *SI*: Shift-Invert Lanczos
- *KPIK*: Krylov Plus Inverted Krylov

We compare

- *PFE*: Computation of the Partial Fraction Expansion by explicitly solving each symmetric complex system
- *Standard Lanczos*: standard Lanczos approach
- *SI*: Shift-Invert Lanczos
- *KPIK*: Krylov Plus Inverted Krylov

Direct methods				
ν	Standard Lanczos	KPIK	SI	PFE
	time (its.)	time (its.)	time (its.)	time
5	294.49 (406)	0.48 (6)	**0.47** (12)	0.74
8	293.68 (406)	0.72 (16)	**0.50** (13)	1.12
11	648.65 (485)	1.05 (26)	**0.62** (17)	1.59
14	- ($>$500)	1.41 (36)	**0.83** (23)	1.98

Iterative methods				
ν	Standard Lanczos	KPIK +PCG	SI +PCG	PFE +QMR
	time(its.)	time(out/avg in)	time(out/avg in)	time(avg its.)
5	294.49 (406)	**1.05** (6/21)	1.90 (9/21)	1.22 (22)
8	293.68 (406)	3.27 (16/35)	2.87 (11/27)	**2.08** (31)
11	648.65 (485)	6.26 (26/44)	5.09 (17/31)	**3.72** (39)
14	- ($>$500)	10.03 (36/53)	7.96 (23/37)	**4.85** (45)

Implementation enhancements

- *Relaxation strategy* for shift and invert

  [Simoncini-Szyld ’03, van den Eshof-Hochbruck ’06]

- *PFE+QMR+mono*

  Use one preconditioner for all systems

Implementation enhancements

- *Relaxation strategy* for shift and invert

  [Simoncini-Szyld ’03, van den Eshof-Hochbruck ’06]

  Fix the final tolerance $\epsilon > 0$

  for the j-th inner system use

  $$\eta_j = \frac{\epsilon}{\|\mathrm{err}_{j-1}\| + \epsilon},$$

  with err_{j-1} error for the $(j-1)$-th iteration.

- *PFE+QMR+mono*

  Use one preconditioner for all systems

Implementation enhancements

- *Relaxation strategy* for shift and invert

  [Simoncini-Szyld ’03, van den Eshof-Hochbruck ’06]

  Fix the final tolerance $\epsilon > 0$

  for the j-th inner system use

$$\eta_j = \frac{\epsilon}{\|\mathrm{err}_{j-1}\| + \epsilon},$$

  with err_{j-1} error for the $(j-1)$-th iteration.

- *PFE+QMR+mono*

  Use one preconditioner for all systems

Implementation enhancements

- *Relaxation strategy* for shift and invert

  [Simoncini-Szyld '03, van den Eshof-Hochbruck '06]

  Fix the final tolerance $\epsilon > 0$

  for the j-th inner system use

$$\eta_j = \frac{\epsilon}{\|\mathrm{err}_{j-1}\| + \epsilon},$$

  with err_{j-1} error for the $(j-1)$-th iteration.

- *PFE+QMR+mono*

  Use one preconditioner for all systems

Implementation enhancements

- *Relaxation strategy* for shift and invert

  [Simoncini-Szyld '03, van den Eshof-Hochbruck '06]

  Fix the final tolerance $\epsilon > 0$

  for the j-th inner system use

$$\eta_j = \frac{\epsilon}{\|\mathrm{err}_{j-1}\| + \epsilon},$$

  with err_{j-1} error for the $(j-1)$-th iteration.

- *PFE+QMR+mono*

  Use one preconditioner for all systems

Evolution problem

- Consider the problem

$$\begin{cases} \frac{\partial}{\partial t}u = \mathcal{L}(u), \quad (x,y) \in (0,1)^2, \quad 0 \leq t \leq \mathbf{T}, \\ u(0,x,y) = u_0 \end{cases}$$

  with $u_0 = 1$ and mixed b. c.
- FD discretization with respect to x and y
- The solution at $t = \mathbf{T}$ is

$$\exp(\mathbf{T}A)u_0 = \exp(\Delta t A) \cdots \exp(\Delta t A) u_0,$$

  with Δt time step.

$\rightarrow$ The performance depends on $||\Delta t A||$

Evolution problem

- Consider the problem

$$\begin{cases} \frac{\partial}{\partial t} u = \mathcal{L}(u), \quad (x,y) \in (0,1)^2, \quad 0 \le t \le \mathbf{T}, \\ u(0,x,y) = u_0 \end{cases}$$

with $u_0 = 1$ and mixed b. c.

- FD discretization with respect to x and y
- The solution at $t = \mathbf{T}$ is

$$\exp(\mathbf{T}A)u_0 = \exp(\Delta t A) \cdots \exp(\Delta t A) u_0,$$

with Δt time step.

$\rightarrow$ The performance depends on $||\Delta t A||$

Evolution problem

- Consider the problem

$$\begin{cases} \frac{\partial}{\partial t} u = \mathcal{L}(u), \quad (x, y) \in (0,1)^2, \quad 0 \leq t \leq \mathbf{T}, \\ u(0, x, y) = u_0 \end{cases}$$

  with $u_0 = 1$ and mixed b. c.

- FD discretization with respect to x and y
- The solution at $t = \mathbf{T}$ is

$$\exp(\mathbf{T}A)u_0 = \exp(\Delta t A) \cdots \exp(\Delta t A) u_0,$$

  with Δt time step.

$\rightarrow$ The performance depends on $||\Delta t A||$

Evolution problem

- Consider the problem

$$\begin{cases} \frac{\partial}{\partial t} u = \mathcal{L}(u), \quad (x, y) \in (0,1)^2, \quad 0 \le t \le \mathbf{T}, \\ u(0, x, y) = u_0 \end{cases}$$

  with $u_0 = 1$ and mixed b. c.

- FD discretization with respect to x and y
- The solution at $t = \mathbf{T}$ is

$$\exp(\mathbf{T}A)u_0 = \exp(\Delta t A) \cdots \exp(\Delta t A) u_0,$$

  with Δt time step.

$\rightarrow$ The performance depends on $||\Delta t A||$

Evolution problem

- Consider the problem

$$\begin{cases} \frac{\partial}{\partial t} u = \mathcal{L}(u), \quad (x, y) \in (0,1)^2, \quad 0 \leq t \leq \mathbf{T}, \\ u(0, x, y) = u_0 \end{cases}$$

  with $u_0 = 1$ and mixed b. c.

- FD discretization with respect to x and y
- The solution at $t = \mathbf{T}$ is

$$\exp(\mathbf{T}A)u_0 = \exp(\Delta t A) \cdots \exp(\Delta t A) u_0,$$

  with Δt time step.

$\rightarrow$ The performance depends on $||\Delta t A||$

Crank-Nicolson method

The solution to

$$\frac{du(t)}{dt} = Au(t), \quad u(0) = u_0$$

is

$$\left(I - \frac{\Delta t}{2} A\right) u_{n+1} = \left(I + \frac{\Delta t}{2} A\right) u_n$$

Crank-Nicolson method

The solution to

$$\frac{du(t)}{dt} = Au(t), \;\; u(0) = u_0$$

is

$$\left(I - \frac{\Delta t}{2}A\right)u_{n+1} = \left(I + \frac{\Delta t}{2}A\right)u_n$$

Evolution problem

Final accuracy: 10^{-6}; $T = 0.1$

grid (nx,ny)	Δt	Standard Lanczos	Crank-Nicolson +PCG	SI +PCG relax	PFE+QMR mono
(90,90)	1e-04	5.70	45.15	135.12	†
	5e-04	3.04	12.29	33.67	†
	1e-03	2.44	7.86	20.15	†
	5e-03	2.58	†	7.42	16.16
	1e-02	3.94	†	5.41	9.56
	5e-02	52.85	†	2.81	2.79
	1e-01	187.31	†	2.10	**1.62**
(120,120)	1e-04	14.16	78.63	339.99	†
	5e-04	7.03	25.92	80.39	†
	1e-03	6.10	18.90	46.49	†
	5e-03	7.53	†	17.80	38.71
	1e-02	12.96	†	13.06	23.35
	5e-02	266.67	†	6.69	7.12
	1e-01	902.46	†	4.80	**3.64**

Conclusions

- For rational functions SI is a preconditioning technique for the linear systems of the PFE
- *A-priori* selection of the shift parameter
- Numerical comparisons of several accelerating techniques
- For the 2D evolution problem SI+PCG and PFE+QMR are very effective, allowing large time steps

Conclusions

- For rational functions SI is a preconditioning technique for the linear systems of the PFE
- *A-priori* selection of the shift parameter
- Numerical comparisons of several accelerating techniques
- For the 2D evolution problem SI+PCG and PFE+QMR are very effective, allowing large time steps

Conclusions

- For rational functions SI is a preconditioning technique for the linear systems of the PFE
- *A-priori* selection of the shift parameter
- Numerical comparisons of several accelerating techniques
- For the 2D evolution problem SI+PCG and PFE+QMR are very effective, allowing large time steps

Conclusions

- For rational functions SI is a preconditioning technique for the linear systems of the PFE
- *A-priori* selection of the shift parameter
- Numerical comparisons of several accelerating techniques
- For the 2D evolution problem SI+PCG and PFE+QMR are very effective, allowing large time steps

Future work

- Non symmetric matrices
- Accelerating other functions

Future work

- Non symmetric matrices
- Accelerating other functions

Some references

- M. Hochbruck and C. Lubich, *On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal., 1997*
- L. Lopez and V. Simoncini, *Analysis of projection methods for rational function approximation to the matrix exponential , SIAM J. Numer. Anal., 2006*
- C. Moler and C. Van Loan, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, SIAM Review, 2003*
- I. Moret and P. Novati, *RD-Rational Approximations of the Matrix Exponential, BIT, Numerical Mathematics, 2004*
- M. P. and V. Simoncini *Acceleration Techniques for Approximating the Matrix Exponential Operator, SIAM J. Matrix Anal. Appl., to appear*
- J. van den Eshof and M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential , SIAM J. Sci. Comput., 2006*