Shift optimization for matrix exponential computing

Alexandr Katrutsa

Skolkovo Institute of Science and Technology

June 6, 2019

Problem statement

$$y(t) = \exp(-tA)v$$

- Integration of differential equations

$$\frac{dy}{dt} = -Ay, \quad y(0) = v$$

- Graph centrality measure
- Point cloud manifold descriptors

Shift-and-invert method

- Krylov methods are based on

$$\mathcal{K}_k(A, v) = \mathrm{span}(\, v, Av, A^2v, \ldots, A^{k-1}v \,)$$

- Approximate in the orthogonal Krylov subspace

$$\exp(-tA)v \approx V_m \exp(-tT_m)e_1$$

- Matrix exponential is mostly determined by the *smallest* eigenpairs
- Replace $\mathcal{K}_k(A, v)$ with $\mathcal{K}_k((I + \gamma A)^{-1}, v)$
- Shift parameter γ affects convergence

How to find optimal shift?

- Change parameter: $\delta = \gamma / t$
- Introduce loss function

$$\|r\|_\infty = \|Ay(t) + y'(t)\|_\infty$$

  and find δ that minimizes it
- Introduce method to update δ for online setting

Optimization methods

Main assumption: all vectors v come from the some fixed subset

- Optimize-and-run method

$$\delta^* = \arg\min_{\delta \in [0.01, 0.1]} \frac{1}{N} \sum_{i=1}^{N} \| r(\delta \mid A, v^{(i)}, K, \epsilon, t) \|_\infty,$$

- Incremental method: sequentially update δ for every next vector

Incremental method

- Compute not only residual, but also derivative estimate
- Update δ with this derivative estimate
- Take the next vector and use updated δ

Application to PDE integration

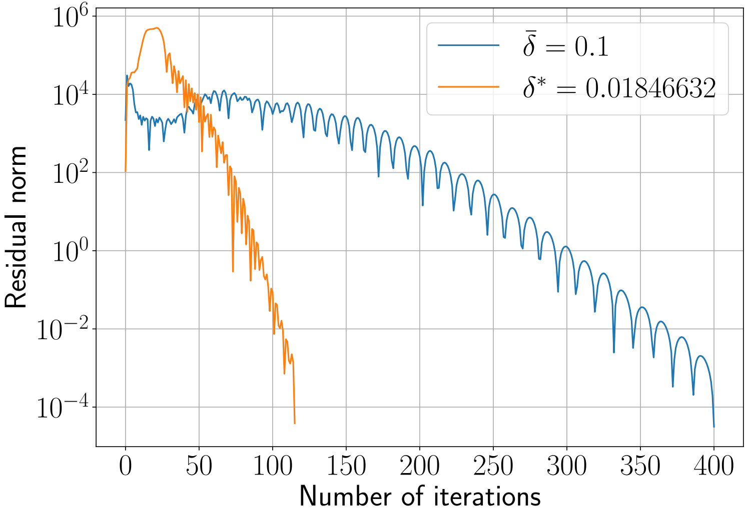
$$\frac{\partial u}{\partial t} = (D_1 u_x)_x + (D_2 u_y)_y, \quad (x, y) \in \Omega = [0, 1] \times [0, 1],$$
$$u(x, y, 0) = u_0,$$

$$D_1 = \begin{cases} 1000, & (x, y) \in \left[\frac{1}{4}, \frac{3}{4}\right] \times \left[\frac{1}{4}, \frac{3}{4}\right], \\ 0.1, & \text{otherwise}, \end{cases} \qquad D_2 = \frac{1}{2} D_1.$$

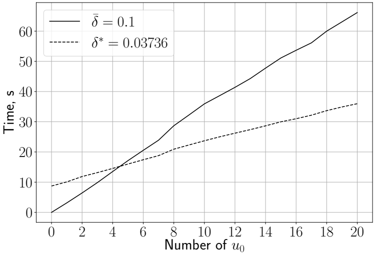
- $u_0 \sim \mathcal{N}(\mu, \Sigma)$
- μ – random vector from Ω
- $\Sigma = 0.05I$
- Second order finite difference discretization

Convergence speed up

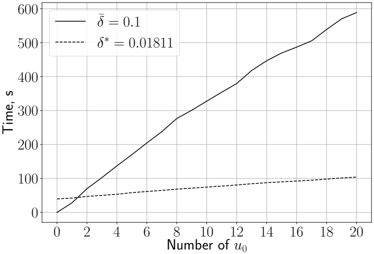
- $n = 300$, matrix size is 90000×90000
- $t = 2 \cdot 10^{-4}$



Total running time comparison



(a) $n = 200, t = 10^{-4}$

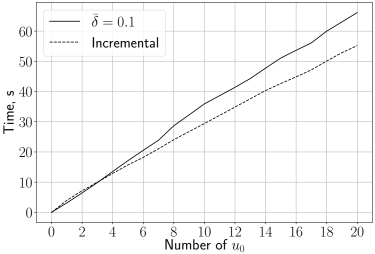


(b) $n = 200, t = 4 \cdot 10^{-4}$

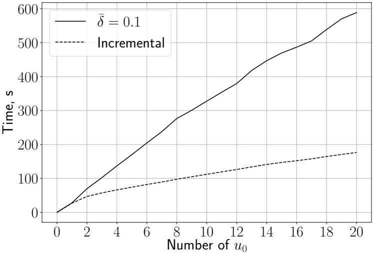
Comparison average number of iterations

	$\bar{\delta} = 0.1$	δ^*
$n = 200, t = 10^{-4}$	88.4	48.5
$n = 200, t = 4 \cdot 10^{-4}$	292.05	93.45
$n = 300, t = 10^{-4}$	237.75	100.6
$n = 300, t = 2 \cdot 10^{-4}$	517.7	150.75

Incremental method



(a) $n = 200, t = 10^{-4}$



(b) $n = 200, t = 4 \cdot 10^{-4}$

Summary

- Rational Krylov approximation of matrix exponential
- Residual norm as loss function to evaluate quality of shift
- Incremental method for online setting