Lanczos-like method for the time-ordered exponential

Pierre-Louis Giscard[1] Stefano Pozza[2]

[1] Université du Littoral Côte d'Opale, Calais

[2] Charles University, Prague

MAT TRIAD
September 9, 2019

Time-ordered exponential

Let $t' \geq t \in I \subseteq \mathbb{R}$, $\mathsf{A}(t')$ a time dependent matrix. The time-ordered exponential is the unique solution $\mathsf{U}(t', t)$ of the system of coupled linear differential equations

$$\mathsf{A}(t')\mathsf{U}(t', t) = \frac{d}{dt'}\mathsf{U}(t', t), \quad \mathsf{U}(t, t) = \mathsf{Id}.$$

If $\mathsf{A}(\tau_1)\mathsf{A}(\tau_2) = \mathsf{A}(\tau_2)\mathsf{A}(\tau_1)$ for all $\tau_1, \tau_2 \in I$, then

$$\mathsf{U}(t', t) = \exp\left(\int_t^{t'} \mathsf{A}(\tau)\, d\tau\right).$$

Otherwise, U has no known explicit form in terms of A. It can be denoted by the time-ordering operator $\mathcal{T}$ as

$$\mathsf{U}(t', t) = \mathcal{T}\exp\left(\int_t^{t'} \mathsf{A}(\tau)\, d\tau\right).$$

[Dyson, 1952]

Time-ordered exponential

The time-ordering expression is more a notation than an explicit form as the action of the time-ordering operator is very difficult to evaluate.

- Applications: System dynamics (quantum physics); e.g., [Blanes & al., 2009]. Differential Riccati matrix equations (control theory, filter design); e.g., [Abou-Kandil et al., 2003].
- Classical approaches Magnus series [Magnus, 1954]: infinite series of increasing complexity (small convergence domain). Using Floquet theory: For A periodic, the solution is given in terms of a infinite system. See, e.g., [Blanes & al., 2009].
- Path-sum approach: Based on the *path-sum* decomposition of the graph whose A is the adjacency matrix [Giscard & al., 2015]. The expression has a finite number of terms, but it may be too complicate.

Outline

1. Strategy: model reduction + path-sum approach.
2. $*$-Lanczos algorithm for time-ordered exponentials.
3. Examples and outlook.

Warninig 1:

$*$-Lanczos works on the space of complex generalized functions on $I \times I$. We will not deal with approximations and finite-precision arithmetic. Our main goal is to give an exact expression for U in a finite number of steps.

Warninig 2:

Rounding errors deeply effects (classical) Lanczos by loss of orthogonality. We expect a similar behavior in any numerical implementation of $*$-Lanczos. This must be investigated before confidently rely on the method in a computational setting.

Strategy

Let $t' \geq t$ and let $\mathsf{A}_1(t',t)$ and $\mathsf{A}_2(t',t)$ be (doubly) time-dependent matrices. We define the convolution-like product

$$(\mathsf{A}_2 * \mathsf{A}_1)(t',t) := \int_t^{t'} \mathsf{A}_2(t',\tau)\mathsf{A}_1(\tau,t)\,\mathrm{d}\tau.$$

Consider the $*$-resolvent

$$\mathsf{R}_*(\mathsf{A}) := (\mathsf{Id}1_* - \mathsf{A})^{*-1} = \mathsf{Id}1_* + \sum_{k\geq 1} \mathsf{A}^{*k}$$

(it exists if every entry of A is bounded for $t', t \in I$). Then

$$\mathsf{U}(t',t) = \int_t^{t'} \mathsf{R}_*(\mathsf{A})(\tau,t)\, d\tau,$$

with $1_* = \delta(t'-t)$ the Dirac delta function; [Giscard & al., 2015].

Path-sum formulation

If we look at A as the adjacency matrix of the graph G with time-dependent weights, then the (i, j) element of the $*$-resolvent, $(R_*)_{ij}$, can be rewritten in terms of the simple paths from i to j in G; [Giscard & al., 2015].

This approach provides exact solutions and it is convergent when R_* exists. However, it suffers from drawbacks:

- It requires to find all the simple cycles and simple paths of the graph G (computationally expensive);
- The expression for U may be too complicated due to the structure of the graph G.

Idea

We want to get a matrix T so that

- T is tridiagonal, i.e, it corresponds to a graph which is a path (with self-loops);
- $R_*(\mathrm{T})_{11} = R_*(\mathrm{A})_{ij}$, i.e., they have the same $*$-moments

$$(T^{*k})_{11} = (A^{*k})_{ij}, \quad k = 0, 1, 2, \ldots .$$

In the simple case in which A is time-independent, a T with analogous properties is give as the output of the last iteration of Lanczos algorithm.

We will derive a Lanczos-like method based on an $*$-orthogonalization process.

Recall: Non-Hermitian Lanczos algorithm

Let A be a time-independent matrix and $\mathbf{v}, \mathbf{w}$ time-independent vectors. Our goal is to approximate

$$\mathbf{w}^H \exp(\mathrm{A})\, \mathbf{v}.$$

Consider the Krylov subspaces

$$\mathrm{span}\{\mathbf{v}, \mathrm{A}\,\mathbf{v}, \dots, \mathrm{A}^{n-1}\,\mathbf{v}\}, \quad \mathrm{span}\{\mathbf{w}, \mathrm{A}^H\mathbf{w}, \dots, (\mathrm{A}^H)^{n-1}\mathbf{w}\}.$$

Assuming not breakdowns, the non-Hermitian Lanczos computes the matrices

$$\mathrm{V}_n = [\mathbf{v}_0, \dots, \mathbf{v}_{n-1}], \quad \mathrm{W}_n = [\mathbf{w}_0, \dots, \mathbf{w}_{n-1}]$$

respectively basis of the Krylov subspaces, so that $\mathrm{W}_n^H \mathrm{V}_n = \mathrm{Id}$.

Non-Hermitian Lanczos algorithm

The tridiagonal matrix defined as

$$\mathrm{T}_n = \mathrm{W}_n^H \, \mathrm{A} \, \mathrm{V}_n,$$

satisfies the matching moment property

$$\mathbf{w}^H \mathrm{A}^k \mathbf{v} = \mathbf{e}_1^H \, \mathrm{T}_n^k \, \mathbf{e}_1, \quad k = 0, \dots, 2n-1.$$

Hence we get the approximation (model reduction)

$$\mathbf{w}^H \exp(\mathrm{A}) \, \mathbf{v} \approx \mathbf{e}_1^H \exp(\mathrm{T}_n) \, \mathbf{e}_1;$$

e.g., [Golub, Meurant, 2010].
In particular, for n equal to the size of A minus 1 (or for lucky or incurable breakdowns) we get exactness.

$*$-resolvent approximation

Now, $\mathrm{A}(t')$ is time-dependent and we want to approximate the i, j element of the time-ordered exponential of A, i.e.,

$$\mathrm{U}(t', t)_{i,j} = \mathbf{e}_i^H \mathcal{T} \exp\left(\int_t^{t'} \mathrm{A}(\tau)\, d\tau \right) \mathbf{e}_j = \int_t^{t'} \mathbf{e}_i^H \mathrm{R}_*(\mathrm{A})(\tau, t) \mathbf{e}_j \, d\tau.$$

Hence we want to approximate the value

$$\mathbf{w}^H \mathrm{R}_*(\mathrm{A})(t', t)\, \mathbf{v},$$

with $\mathbf{v}, \mathbf{w}$ time-independent vectors.

Time-dependent Krylov subspaces

Given the distributions $\alpha_0(t', t), \ldots, \alpha_k(t', t)$ (coefficients),

matrix $*$-polynomial: $$p(\mathsf{A})(t', t) := \sum_{j=0}^{k} \left(\mathsf{A}^{*j} * \alpha_j\right)(t', t),$$

dual $*$-polynomial: $$p^D(\mathsf{A})(t', t) := \sum_{j=0}^{k} \left(\overline{\alpha}_j * (\mathsf{A}^{*j})^H\right)(t', t),$$

and hence the time-dependent Krylov subspaces as

$$\mathcal{K}_n(\mathsf{A}, \mathbf{v})(t', t) := \{\, (\, p(\mathsf{A})\mathbf{v}) \mid p \text{ of degree } \leq n-1\},$$
$$\mathcal{K}_n^D(\mathsf{A}^H, \mathbf{w})(t', t) := \left\{\, \left(\mathbf{w}^H p^D(\mathsf{A}^H)\right) \mid p \text{ of degree } \leq n-1\right\}.$$

Note: To work with distribution we need to set
$\mathsf{A}(t', t) = \widetilde{\mathsf{A}}(t')\Theta(t' - t)$ and generalize the definition of $*$.

Building up the $*$-Lanczos process

We aim to build basis for $\mathcal{K}_n(\mathsf{A}, \mathbf{v})$ and $\mathcal{K}_n^D(\mathsf{A}^H, \mathbf{w})$

$$\mathsf{V}_n = [\mathbf{v}_0, \dots, \mathbf{v}_{n-1}], \quad \mathsf{W}_n^H = [\mathbf{w}_0, \dots, \mathbf{w}_{n-1}]^H$$

which are $*$-biorthonormal, i.e.,

$$\mathsf{W}_n^H * \mathsf{V}_n = \mathsf{I}1_*.$$

Assuming $\mathbf{w}^H\mathbf{v} = 1$, the first vectors are

$$\mathbf{v}_0 = \mathbf{v}\, 1_*, \quad \mathbf{w}_0^H = \mathbf{w}^H 1_*.$$

Building up the $*$-Lanczos process

Consider a vector $\widehat{\boldsymbol{v}}_1 \in \mathcal{K}_2(\mathsf{A}, \boldsymbol{v})$

$$\widehat{\boldsymbol{v}}_1 = \mathsf{A} * \boldsymbol{v}_0 - \boldsymbol{v}_0 * \alpha_0 = \mathsf{A}\boldsymbol{v} - \boldsymbol{v}\alpha_0.$$

Under the condition $\boldsymbol{w}_0^H * \widehat{\boldsymbol{v}}_1 = 0$, we get

$$\alpha_0 = \boldsymbol{w}_0^H * \mathsf{A} * \boldsymbol{v}_0 = \boldsymbol{w}^H \mathsf{A}\, \boldsymbol{v} \qquad (1)$$

Similarly,

$$\widehat{\boldsymbol{w}}_1^H = \boldsymbol{w}_0^H * \mathsf{A} - \alpha_0 * \boldsymbol{w}_0^H = \boldsymbol{w}^H \mathsf{A} - \alpha_0 \boldsymbol{w}^H,$$

with α_0 given by (1).

Building up the $*$-Lanczos process

The condition

$$\boldsymbol{w}_1^H * \boldsymbol{v}_1 = 1_*,$$

is satisfied definining

$$\boldsymbol{v}_1 = \widehat{\boldsymbol{v}}_1 * (\beta_1)^{*-1}, \quad \boldsymbol{w}_1 = \widehat{\boldsymbol{w}}_1,$$

with

$$\beta_1 = \widehat{\boldsymbol{w}}_1^H * \widehat{\boldsymbol{v}}_1 = \boldsymbol{w}^H \mathsf{A}^{*2} \boldsymbol{v} - \alpha_0^{*2}.$$

- We assume β_1 $*$-invertible;
- We proved that every non-identically null holonomic (D-finite) function is invertible.

Building up the $*$-Lanczos process

By induction, consider the vector

$$\widehat{\boldsymbol{v}}_n = \mathrm{A} * \boldsymbol{v}_{n-1} - \sum_{i=0}^{n-1} \boldsymbol{v}_i * \gamma_i.$$

The condition $\boldsymbol{w}_j^H * \widehat{\boldsymbol{v}}_n = 0$ for $j = 0, \dots, n-1$ gives

$$\gamma_j = \boldsymbol{w}_j^H * \mathrm{A} * \boldsymbol{v}_{n-1}, \quad j = 0, \dots, n-1.$$

Since $\boldsymbol{w}_j^H * \mathrm{A} \in \mathcal{K}_{j+1}^D(\mathrm{A}^H, \boldsymbol{w})$ we get

$$\gamma_j = 0, \quad j = 0, \dots, n-3.$$

Building up the $*$-Lanczos process

Assuming $\boldsymbol{v}_{-1} = \boldsymbol{w}_{-1} = 0$, we get the three-term recurrences

$$\boldsymbol{w}_n^H = \boldsymbol{w}_{n-1}^H * \mathsf{A} - \alpha_{n-1} * \boldsymbol{w}_{n-1}^H - \beta_{n-1} * \boldsymbol{w}_{n-2}^H,$$
$$\boldsymbol{v}_n * \beta_n = \mathsf{A} * \boldsymbol{v}_{n-1} - \boldsymbol{v}_{n-1} * \alpha_{n-1} - \boldsymbol{v}_{n-2},$$

with the coefficients given by

$$\alpha_{n-1} = \boldsymbol{w}_{n-1}^H * \mathsf{A} * \boldsymbol{v}_{n-1}, \quad \beta_n = \boldsymbol{w}_n^H * \mathsf{A} * \boldsymbol{v}_{n-1}.$$

- If β_n is not $*$-invertible we have a breakdown;
- We assume no breakdowns in this first study;
- The algorithm converges whenever $\mathbf{v}_n = 0$ or $\mathbf{w}_n = 0$ (lucky breakdown).

$*$-Lanczos Algorithm

Initialize: $\boldsymbol{v}_{-1} = \boldsymbol{w}_{-1} = 0$, $\boldsymbol{v}_0 = \boldsymbol{v}\, 1_*$, $\boldsymbol{w}_0^H = \boldsymbol{w}^H 1_*$.

$$\alpha_0 = \boldsymbol{w}^H \mathsf{A}\, \boldsymbol{v},$$

$$\boldsymbol{w}_1^H = \boldsymbol{w}^H \mathsf{A} - \alpha_0\, \boldsymbol{w}^H,$$

$$\widehat{\boldsymbol{v}}_1 = \mathsf{A}\, \boldsymbol{v} - \boldsymbol{v}\, \alpha_0,$$

$$\beta_1 = \boldsymbol{w}^H \mathsf{A}^{*2}\, \boldsymbol{v} - \alpha_0^{*2},$$

If β_1 is not $*$-invertible, then stop, otherwise,

$$\boldsymbol{v}_1 = \widehat{\boldsymbol{v}}_1 * \beta_1^{*-1},$$

For $n = 2, \dots$

$$\alpha_{n-1} = \boldsymbol{w}_{n-1}^H * \mathsf{A} * \boldsymbol{v}_{n-1},$$

$$\boldsymbol{w}_n^H = \boldsymbol{w}_{n-1}^H * \mathsf{A} - \alpha_{n-1} * \boldsymbol{w}_{n-1}^H - \beta_{n-1} * \boldsymbol{w}_{n-2}^H,$$

$$\widehat{\boldsymbol{v}}_n = \mathsf{A} * \boldsymbol{v}_{n-1} - \boldsymbol{v}_{n-1} * \alpha_{n-1} - \boldsymbol{v}_{n-2},$$

$$\beta_n = \boldsymbol{w}_n^H * \mathsf{A} * \boldsymbol{v}_{n-1},$$

If β_n is not $*$-invertible, then stop, otherwise,

$$\boldsymbol{v}_n = \widehat{\boldsymbol{v}}_n * \beta_n^{*-1},$$

end.

Tridiagonal matrix

$$\mathsf{T}_n := \begin{bmatrix} \alpha_0 & 1_* & & \\ \beta_1 & \alpha_1 & \ddots & \\ & \ddots & \ddots & 1_* \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix},$$

and $\mathsf{V}_n := [\boldsymbol{v}_0, \dots, \boldsymbol{v}_{n-1}]$ and $\mathsf{W}_n := [\boldsymbol{w}_0, \dots, \boldsymbol{w}_{n-1}]$. Then

$$\mathsf{A} * \mathsf{V}_n = \mathsf{V}_n * \mathsf{T}_n + (\boldsymbol{v}_n * \beta_n)\boldsymbol{e}_n^H$$
$$\mathsf{W}_n^H * \mathsf{A} = \mathsf{T}_n * \mathsf{W}_n^H + \boldsymbol{e}_n \boldsymbol{w}_n^H.$$

Hence T_n is the $*$-projection of A onto $\mathcal{K}_n(\mathsf{A}, \boldsymbol{v})$ along the direction of $\mathcal{K}_n^D(\mathsf{A}^H, \boldsymbol{w})$, i.e.,

$$\mathsf{T}_n = \mathsf{W}_n^H * \mathsf{A} * \mathsf{V}_n.$$

Matching moment

Matching moment property

Let $\mathsf{A}, \boldsymbol{w}, \boldsymbol{v}$ and T_n be as described above, then

$$\boldsymbol{w}^H(\mathsf{A}^{*k})\,\boldsymbol{v} = \boldsymbol{e}_1^H(\mathsf{T}_n^{*k})\,\boldsymbol{e}_1, \quad \text{for} \quad k = 0, \dots, 2n-1.$$

In particular, for n equal to the size of A minus 1 we get

$$\boldsymbol{w}^H(\mathsf{A}^{*k})\,\boldsymbol{v} = \boldsymbol{e}_1^H(\mathsf{T}_n^{*k})\,\boldsymbol{e}_1, \quad \text{for} \quad k = 0, 2, \dots,$$

and, hence,

$$\boldsymbol{w}^H R_*(\mathsf{A})\boldsymbol{v} = \boldsymbol{e}_1^H R_*(\mathsf{T}_n)\boldsymbol{e}_1.$$

$*$-Lanczos approximation

Consider the time-ordered exponential U_n given by

$$\mathsf{T}_\mathsf{n}(t',t)\mathsf{U}_\mathsf{n}(t',t) = \frac{d}{dt'}\mathsf{U}_n(t',t).$$

The matching moment property justify the use of the approximation

$$\boldsymbol{w}^H \mathsf{U}(t',t)\boldsymbol{v} \approx \boldsymbol{e}_1^H \mathsf{U}_n(t',t)\,\boldsymbol{e}_1 = \boldsymbol{e}_1^H \int_t^{t'} R_*(\mathsf{T}_n)(\tau,t)\,d\tau\,\boldsymbol{e}_1.$$

The path-sum method gives the finite $*$-continued fraction

$$R_*(\mathsf{T}_n)_{11} = \Big(1_* - \alpha_0 - \big(1_* - \alpha_1 - (\cdots * \beta_{n-1})^{*-1} \cdots * \beta_2\big)^{*-1} * \beta_1\Big)^{*-1},$$

see [Giscard & al. 2012, 2015].

Error bound?

Conjecture

If the entries of $\mathrm{A}(t')$ are smooth, then the $*$-Lanczos coefficients α_j and β_j are (usual) functions.

Proposition

Under the previous conjecture, the approximation error of the nth $*$-Lanczos can be bound by

$$\left| \mathbf{w}^H \mathrm{U}(t', t)\mathbf{v} - \int_t^{t'} \mathrm{R}_*(\mathrm{T}_n)_{1,1}(\tau, t)\, d\tau \right| \leq \left(C^{2n} + D_n^{2n}\right) \frac{(t'-t)^{2n}}{(2n)!},$$

with the finite coefficients

$$C := \sup_{t' \in I} \|\tilde{\mathrm{A}}(t')\|_\infty, \quad D_n := \sup_{t',t \in I^2} \max_{0 \leq j \leq n-1} \left\{ |\tilde{\alpha}_j(t', t)|, |\tilde{\beta}_j(t', t)| \right\}.$$

Example: time-dependent matrix

Consider the matrix

$$A = \begin{pmatrix} \cos(t') & 0 & 1 & 2 & 1 \\ 0 & \cos(t') - t' & 1 - 3t' & t' & 0 \\ 0 & t' & 2t' + \cos(t') & 0 & 0 \\ 0 & 1 & 2t' + 1 & t' + \cos(t') & t' \\ t' & -t' - 1 & -6t' - 1 & 1 - 2t' & \cos(t') - 2t' \end{pmatrix}.$$

The matrix does not commute with itself at different times and the corresponding differential system has no known analytical solution.

Example: time-dependent matrix

We get

$$T_5=\begin{pmatrix} \cos(t')\Theta & \delta & 0 & 0 & 0 \\ \frac{1}{2}(t'^2-t^2)\Theta & \cos(t)\Theta & \delta & 0 & 0 \\ 0 & t(t'-t)\Theta & \widetilde{\alpha}_2(t',t)\Theta & \delta & 0 \\ 0 & 0 & -\frac{1}{2}(3t^2-4tt'+t'^2)\Theta & \widetilde{\alpha}_3(t',t)\Theta & \delta \\ 0 & 0 & 0 & (-2t^2+3tt'-t'^2)\Theta & \widetilde{\alpha}_4(t',t)\Theta \end{pmatrix},$$

with

$$\widetilde{\alpha}_2(t',t) = (t'-t)\sin(t) + \cos(t),$$
$$\widetilde{\alpha}_3(t',t) = \frac{1}{2}\left(4(t'-t)\sin(t) - \left((t-t')^2-2\right)\right)\cos(t),$$
$$\widetilde{\alpha}_4(t',t) = \frac{1}{6}\left(\left((t-t')^2-18\right)(t-t')\sin(t) + \left(6-9(t-t')^2\right)\cos(t)\right),$$

Here $\Theta = \Theta(t'-t)$ (Heaviside function) and $\delta = \delta(t'-t)$ (Dirac delta function).

Example: time-dependent matrix

With the $*$-biorthonormal basis

$$V_5 = \begin{pmatrix} \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta^{(3)} & -2\delta^{(4)} \\ 0 & 0 & 0 & 0 & \delta^{(4)} \\ 0 & 0 & \delta'' & -\delta^{(3)} & \delta^{(4)} \\ 0 & \delta' & -2\delta'' & 2\delta^{(3)} & -3\delta^{(4)} \end{pmatrix},$$

and

$$W_5^H = \begin{pmatrix} \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & \Theta & 2\Theta & \Theta \\ 0 & (t'-t)\Theta & (t'-t)\Theta & (t'-t)\Theta & 0 \\ 0 & \frac{1}{2}(t-t')^2\Theta & (t-t')^2\Theta & 0 & 0 \\ 0 & 0 & -\frac{1}{6}(t-t')^3\Theta & 0 & 0 \end{pmatrix}.$$

The Dirac delta derivatives are coming from:

$$\beta_1^{*-1} = \frac{1}{t}\delta'(t'-t) * \delta'(t'-t), \qquad \beta_2^{*-1} = \frac{1}{t'}\delta'(t'-t) * \delta'(t'-t)$$

$$\beta_3^{*-1} = \frac{1}{t}\Theta(t'-t) * \delta^{(3)}(t'-t), \quad \beta_4^{*-1} = \frac{t'}{t^2}\Theta(t'-t) * \delta^{(3)}(t'-t).$$

Summarizing

$*$-Lanczos is able to express the solution of a coupled linear differential equations with non-constant coefficients.
It presents several difficulties ($*$-inverses, integral solutions, . . .).
Nevertheless,

- It is convergent in a finite number of steps;
- It express the solution with a finite formulation (as the path-sum method);
- It admits an expression in terms of a simple finite continued fraction.

Such a general approach has been so far out of reach.

Outlook: numerical implementation

There exists an isometry Φ between the algebra of generalized functions depending on two time variables and the algebra of "time continuous" operators (for which the time variables t' and t serve as line and row indices).

Once time is discretized, Φ sends the $*$-product to the ordinary matrix product. Most importantly, Φ sends usual functions of two times ($f(t',t) = \tilde{f}(t',t)\Theta(t',t)$) to *lower triangular matrices*, the matrix inverse of which is the image under Φ of the $*$-inverse of $f(t',t)$.

This is a natural time-discretization strategy that we will investigate in a future work.

Other open issues

- Convergence properties of the approximation (conjecture);
- Breakdown and near-breakdown issues;
- Computational approach: loss of orthogonality, reorthogonalization, stability, inexact approaches, better error estimates, . . . ;
- Numerical tests;
- Application of this $*$-strategy to other analogous problems.

More information:
P-L. G., S. P., *Lanczos-like method for the time-ordered exponential*, soon on arXiv!, http://karlin.mff.cuni.cz/∼pozza/

Other open issues

- Convergence properties of the approximation (conjecture);
- Breakdown and near-breakdown issues;
- Computational approach: loss of orthogonality, reorthogonalization, stability, inexact approaches, better error estimates, . . . ;
- Numerical tests;
- Application of this $*$-strategy to other analogous problems.

More information:
P-L. G., S. P., *Lanczos-like method for the time-ordered exponential*, soon on arXiv!, http://karlin.mff.cuni.cz/~pozza/

Thank you for your attention!