Computational methods for large-scale linear matrix equations and application to FDEs

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

Joint work with:

Tobias Breiten, University of Graz (A)

Martin Stoll, Max Planck Institute Magdeburg (G)

A motivational example with spatial fractional derivatives

Let $\Omega = (a_x, b_x) \times (a_y, b_y) \subset \mathbb{R}^2$. Consider

($u = u(x, y, t)$, $f = f(x, y, t)$, with $(x, y, t) \in \Omega \times (0, T]$)

$$\frac{du}{dt} - {}_xD_R^{\beta_1} u - {}_yD_R^{\beta_2} u = f,$$

$$u(x, y, t) = 0, \quad (x, y, t) \in \partial\Omega \times [0, T]$$

$$u(x, y, 0) = 0, \quad (x, y) \in \Omega$$

where $\beta_1, \beta_2 \in (1, 2)$ and

$${}_xD_R^{\beta_1} u = \frac{1}{\Gamma(-\beta_1)} \lim_{n_x \to \infty} \frac{1}{h_x^{\beta_1}} \sum_{k=0}^{n_x} \frac{\Gamma(k - \beta_1)}{\Gamma(k + 1)} u(x - (k-1)h_x, y, t)$$

$${}_yD_R^{\beta_2} u = \frac{1}{\Gamma(-\beta_2)} \lim_{n_y \to \infty} \frac{1}{h_y^{\beta_2}} \sum_{k=0}^{n_y} \frac{\Gamma(k - \beta_2)}{\Gamma(k + 1)} u(x, y - (k-1)h_y, t)$$

After finite-difference spatial discretization, implicit Euler yields

$$\mathbf{A}_{\beta_1,\tau} \mathbf{U}^{n+1} + \mathbf{U}^{n+1} \mathbf{B}_{\beta_2,\tau} = \mathbf{U}^n + \tau \mathbf{F}^{n+1}$$

to be solved for the **matrix** $\mathbf{U}^{n+1}$, at each time step t_{n+1}

Before we start. Kronecker products

- $(\mathbf{I} \otimes \mathbf{A})\mathbf{u} \leftrightarrow \mathbf{AU}$, where $\mathbf{u} = \mathrm{vec}(\mathbf{U})$
- $(\mathbf{B}^T \otimes \mathbf{I})\mathbf{u} \leftrightarrow \mathbf{UB}$
- $(\mathbf{I} \otimes \mathbf{A} + \mathbf{B}^T \otimes \mathbf{I})\mathbf{u} \leftrightarrow \mathbf{AU} + \mathbf{UB}$

Fractional calculus and Grünwald formulas

- Caputo fractional derivative (usually employed in time):

$$ {}_a^C D_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t \frac{f^{(n)}(s)ds}{(t-s)^{\alpha-n+1}}, \qquad n-1 < \alpha \le n $$

$\Gamma(x)$: Gamma function

- Left-sided Riemann-Liouville fractional derivative:

$$ {}_a^{RL} D_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_a^t \frac{f(s)ds}{(t-s)^{\alpha-n+1}}, \qquad a < t < b \qquad n-1 < \alpha \le n $$

- Right-sided Riemann-Liouville fractional derivative:

$$ {}_t^{RL} D_b^\alpha f(t) = \frac{(-1)^n}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_t^b \frac{f(s)ds}{(s-t)^{\alpha-n+1}}, \qquad a < t < b $$

- Symmetric Riesz derivative of order α:

$$ \frac{d^\alpha f(t)}{d|t|^\alpha} = {}_t D_R^\alpha f(t) = \frac{1}{2}\left({}_a^{RL} D_t^\alpha f(t) + {}_t^{RL} D_b^\alpha f(t) \right). $$

(Podlubny 1998, Podlubny etal 2009, Samko etal 1993)

Towards the numerical approximation of fractional derivatives

For the left-sided derivative it holds:

$$ {}_{a}^{RL}D_{x}^{\alpha} f(x,t) = \lim_{M\to\infty} \frac{1}{h^{\alpha}} \sum_{k=0}^{M} g_{\alpha,k} f(x-kh,t), \qquad h = \frac{x-a}{M} $$

with

$$ g_{\alpha,k} = \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} = (-1)^k \binom{\alpha}{k}. $$

Analogously for the right-sided derivative.

For stability reasons. Shifted version:

$$ {}_{a}^{RL}D_{x}^{\alpha} f(x,t) = \lim_{M\to\infty} \frac{1}{h^{\alpha}} \sum_{k=0}^{M} g_{\alpha,k} f(x-(k-1)h,t) $$

(Meerschaert & Tadjeran 2004)

A simple 1D model problem

$$\frac{du(x,t)}{dt} - {}_xD_R^{\beta} u(x,t) = f(x,t), \quad (x,t) \in (0,1) \times (0,T], \quad \beta \in (1,2)$$
$$u(0,t) = u(1,t) = 0, \qquad t \in [0,T]$$
$$u(x,0) = 0, \qquad x \in [0,1]$$

⋆ Time discretization. implicit Euler scheme of step size τ:

$$\frac{u^{n+1} - u^n}{\tau} - {}_xD_R^{\beta} u^{n+1} = f^{n+1},$$

where $u^{n+1} := u(x, t_{n+1})$, $f^{n+1} := f(x, t_{n+1})$

at time $t_{n+1} = (n+1)\tau$

⋆ Space discretization. **First** consider one-sided operator:

$$ {}_{a}^{RL}D_x^{\beta} u_i^{n+1} \approx \frac{1}{h_x^{\beta}} \sum_{k=0}^{i+1} g_{\beta,k} u_{i-k+1}^{n+1} $$

(n_x number of interior points in space), giving for all i,

$$ \mathbf{T}_{\beta}^{n_x} \mathbf{u}^{n+1} := \frac{1}{h_x^{\beta}} \begin{bmatrix} g_{\beta,1} & g_{\beta,0} & 0 & \cdots & 0 \\ g_{\beta,2} & g_{\beta,1} & g_{\beta,0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & g_{\beta,0} & 0 \\ g_{\beta,n_x-1} & \ddots & \ddots & g_{\beta,1} & g_{\beta,0} \\ g_{\beta,n_x} & g_{\beta,n_x-1} & \cdots & g_{\beta,2} & g_{\beta,1} \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ \vdots \\ \vdots \\ u_{n_x}^{n+1} \end{bmatrix} $$

A simple 1D model problem. Cont'd

For the Riesz derivative (left-right average) ${}_xD_R^{\beta}u(x,t)$:

$$ {}_xD_R^{\beta}u(x,t) \approx \mathbf{L}_{\beta}^{n_x}\mathbf{u}^{n+1} := \frac{1}{2}\left(\mathbf{T}_{\beta}^{n_x} + (\mathbf{T}_{\beta}^{n_x})^T\right)\mathbf{u}^{n+1} $$

(Meerschaert & Tadjeran 2006)

Therefore, the discretized version of the original equation reads:

$$ \left(\mathbf{I}^{n_x} - \tau\mathbf{L}_{\beta}^{n_x}\right)\mathbf{u}^{n+1} = \mathbf{u}^{n} + \tau\mathbf{f}^{n+1} $$

with $\left(\mathbf{I}^{n_x} - \tau\mathbf{L}_{\beta}^{n_x}\right)$ Toeplitz structure.

Numerical solution: direct FFT-based or preconditioned iterative solvers

The initial 2D (in space) problem

Let $\Omega = (a_x, b_x) \times (a_y, b_y) \subset \mathbb{R}^2$. Consider

$$\frac{du}{dt} - {}_xD_R^{\beta_1} u - {}_yD_R^{\beta_2} u = f,$$

$$u(x,y,t) = 0, \quad (x,y,t) \in \partial\Omega \times [0,T]$$

$$u(x,y,0) = 0, \quad (x,y) \in \Omega$$

where $\beta_1, \beta_2 \in (1,2)$

A similar discretization procedure yields:

$$\frac{1}{\tau}\left(\mathbf{u}^{n+1} - \mathbf{u}^n\right) = \underbrace{\left(\mathbf{I}^{n_y} \otimes \mathbf{L}_{\beta_1}^{n_x} + \mathbf{L}_{\beta_2}^{n_y} \otimes \mathbf{I}^{n_x}\right)}_{\mathbf{L}_{\beta_1,\beta_2}^{n_x n_y}} \mathbf{u}^{n+1} + \mathbf{f}^{n+1}.$$

where $h_x = \frac{b_x - a_x}{n_x + 1}$, $h_y = \frac{b_y - a_y}{n_y + 1}$ with $n_x + 2$ and $n_y + 2$ inner nodes.

At each time step, we need to solve

$$\left(\mathbf{I}^{n_x n_y} - \tau \mathbf{L}_{\beta_1,\beta_2}^{n_x n_y}\right) \mathbf{u}^{n+1} = \mathbf{u}^n + \tau \mathbf{f}^{n+1}.$$

The initial 2D (in space) problem. Cont'd

$$\left(\mathbf{I}^{n_x n_y} - \tau \mathbf{L}_{\beta_1,\beta_2}^{n_x n_y}\right) \mathbf{u}^{n+1} = \mathbf{u}^n + \tau \mathbf{f}^{n+1}$$

The coefficient matrix satisfies:

$$\mathbf{I}^{n_x n_y} - \tau \mathbf{L}_{\beta_1,\beta_2}^{n_x n_y} = \mathbf{I}^{n_y} \otimes \left(\frac{1}{2}\mathbf{I}^{n_x} - \tau \mathbf{L}_{\beta_1}^{n_x}\right) + \left(\frac{1}{2}\mathbf{I}^{n_y} - \tau \mathbf{L}_{\beta_2}^{n_y}\right) \otimes \mathbf{I}^{n_x}$$

The initial 2D (in space) problem. Cont'd

$$\left(\mathbf{I}^{n_x n_y} - \tau \mathbf{L}^{n_x n_y}_{\beta_1,\beta_2}\right) \mathbf{u}^{n+1} = \mathbf{u}^n + \tau \mathbf{f}^{n+1}$$

The coefficient matrix satisfies:

$$\mathbf{I}^{n_x n_y} - \tau \mathbf{L}^{n_x n_y}_{\beta_1,\beta_2} = \mathbf{I}^{n_y} \otimes \left(\frac{1}{2}\mathbf{I}^{n_x} - \tau \mathbf{L}^{n_x}_{\beta_1}\right) + \left(\frac{1}{2}\mathbf{I}^{n_y} - \tau \mathbf{L}^{n_y}_{\beta_2}\right) \otimes \mathbf{I}^{n_x}$$

By "unfolding" the Kronecker products:

Sylvester matrix equation:

$$\left(\frac{1}{2}\mathbf{I}^{n_x} - \tau \mathbf{L}^{n_x}_{\beta_1}\right) \mathbf{U}^{n+1} + \mathbf{U}^{n+1} \left(\frac{1}{2}\mathbf{I}^{n_y} - \tau \mathbf{L}^{n_y}_{\beta_2}\right)^T = \mathbf{U}^n + \tau \mathbf{F}^{n+1}$$

Coefficient matrices are different and have Toeplitz structure

Time fractional derivative

For $\alpha \in (0,1)$ and $\beta \in (1,2)$ consider the problem

$$\begin{aligned} {}_0^C D_t^\alpha u(x,t) - {}_x D_R^\beta u(x,t) &= f(x,t), & (x,t) &\in (0,1)\times(0,T], \\ u(0,t) = u(1,t) &= 0, & t &\in [0,T], \\ u(x,0) &= 0, & x &\in [0,1], \end{aligned}$$

Discretizing the Caputo derivative in time, we obtain

$$\left((\mathbf{T}_\alpha^{n_t} \otimes \mathbf{I}^{n_x}) - (\mathbf{I}^{n_t} \otimes \mathbf{L}_\beta^{n_x}) \right) \mathbf{u} = \tilde{\mathbf{f}} \quad \Leftrightarrow \quad \mathbf{U}(\mathbf{T}_\alpha^{n_t})^T - \mathbf{L}_\beta^{n_x}\mathbf{U} = \mathbf{F}$$

$$\text{with} \qquad \mathbf{T}_\alpha^{n_t+1} := \tau^{-\alpha} \begin{bmatrix} g_{\alpha,0} & 0 & \cdots & \cdots & 0 \\ g_{\alpha,1} & g_{\alpha,0} & \ddots & & \vdots \\ \ddots & \ddots & \ddots & \ddots & \vdots \\ \ddots & \ddots & \ddots & g_{\alpha,0} & 0 \\ g_{\alpha,n_t} & \cdots & \cdots & g_{\alpha,1} & g_{\alpha,0} \end{bmatrix}$$

Note: dim($\mathbf{T}_\alpha^{n_t}$) may be much smaller than dim($\mathbf{L}_\beta^{n_x}$)

Numerical solution of the Sylvester equation

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}$$

Various settings:

- Small $\mathbf{A}$ and small $\mathbf{B}$: Bartels-Stewart algorithm
  1. Compute the Schur forms:
     $\mathbf{A}^* = URU^*$, $\mathbf{B} = VSV^*$ with R, S upper triangular;
  2. Solve $R^*\mathbf{Y} + \mathbf{Y}S = U^*\mathbf{G}V$ for $\mathbf{Y}$;
  3. Compute $\mathbf{U} = U\mathbf{Y}V^*$.

Numerical solution of the Sylvester equation

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}$$

Various settings:

- Small $\mathbf{A}$ and small $\mathbf{B}$: Bartels-Stewart algorithm
  1. Compute the Schur forms:
     $\mathbf{A}^* = URU^*$, $\mathbf{B} = VSV^*$ with R, S upper triangular;
  2. Solve $R^*\mathbf{Y} + \mathbf{Y}S = U^*\mathbf{G}V$ for $\mathbf{Y}$;
  3. Compute $\mathbf{U} = U\mathbf{Y}V^*$.

- Large $\mathbf{A}$ and small $\mathbf{B}$: Column decoupling
  1. Compute the decomposition $\mathbf{B} = WSW^{-1}$, $S = \mathrm{diag}(s_1, \ldots, s_m)$
  2. Set $\widehat{\mathbf{G}} = \mathbf{G}W$
  3. For $i = 1, \ldots, m$ solve $(\mathbf{A} + s_i I)(\widehat{\mathbf{U}})_i = (\widehat{\mathbf{G}})_i$
  4. Compute $\mathbf{U} = \widehat{\mathbf{U}}W^{-1}$

Numerical solution of the Sylvester equation

$$\mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{B} = \mathbf{G}$$

Various settings:

- Small $\mathbf{A}$ and small $\mathbf{B}$: Bartels-Stewart algorithm
  1. Compute the Schur forms:
     $\mathbf{A}^* = URU^*$, $\mathbf{B} = VSV^*$ with R, S upper triangular;
  2. Solve $R^*\mathbf{Y} + \mathbf{Y}S = U^*\mathbf{G}V$ for $\mathbf{Y}$;
  3. Compute $\mathbf{U} = U\mathbf{Y}V^*$.
- Large $\mathbf{A}$ and small $\mathbf{B}$: Column decoupling
  1. Compute the decomposition $\mathbf{B} = WSW^{-1}$, $S = \mathrm{diag}(s_1, \ldots, s_m)$
  2. Set $\widehat{\mathbf{G}} = \mathbf{G}W$
  3. For $i = 1, \ldots, m$ solve $(\mathbf{A} + s_i I)(\widehat{\mathbf{U}})_i = (\widehat{\mathbf{G}})_i$
  4. Compute $\mathbf{U} = \widehat{\mathbf{U}}W^{-1}$
- Large $\mathbf{A}$ and large $\mathbf{B}$: Iterative solution ($\mathbf{G}$ low rank)

Numerical solution of large scale Sylvester equations

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}$$

with $\mathbf{G}$ low rank

- Projection methods
- ADI (Alternating Direction Iteration)
- Data sparse approaches (structure-dependent)

Numerical solution of large scale Sylvester equations

$$\mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{B} = \mathbf{G}$$

with $\mathbf{G}$ low rank

- Projection methods
- ADI (Alternating Direction Iteration)
- Data sparse approaches (structure-dependent)

Projection methods

Seek $\mathbf{U}_k \approx \mathbf{U}$ of low rank:

$$\mathbf{U}_k = \begin{bmatrix} \\ \mathbf{U}_k^{(1)} \\ \\ \end{bmatrix} [\ (\mathbf{U}_k^{(2)})^*\ ]$$

with $\mathbf{U}_k^{(1)}, \mathbf{U}_k^{(2)}$ tall

Index k "related" to the approximation rank

Galerkin projection methods

$$\mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{B} = \mathbf{G}$$

Consider two (approximation) spaces Range(V) and Range(W) with V, W having orthonormal columns

$\star$ dim(Range(V))$\ll$ size$(\mathbf{A})$ and dim(Range(W))$\ll$ size$(\mathbf{B})$

- Write $\mathbf{U}_k = V\mathbf{Y}W^*$ with $\mathbf{Y}$ to be determined

Galerkin projection methods

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}$$

Consider two (approximation) spaces Range(V) and Range(W) with V, W having orthonormal columns

⋆ dim(Range(V))$\ll$ size$(\mathbf{A})$ and dim(Range(W))$\ll$ size$(\mathbf{B})$

- Write $\mathbf{U}_k = V\mathbf{Y}W^*$ with $\mathbf{Y}$ to be determined
- Impose Galerkin condition on the residual matrix $\mathbf{R}_k = \mathbf{AU}_k + \mathbf{U}_k\mathbf{B} - \mathbf{G}$:

$$V^*\mathbf{R}_kW = 0$$

condition corresponds to $(W \otimes V)^* \mathrm{vec}(\mathbf{R}_\mathrm{k}) = 0$

Galerkin projection methods

$$\mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{B} = \mathbf{G}$$

Consider two (approximation) spaces Range(V) and Range(W) with V, W having orthonormal columns

⋆ dim(Range(V))$\ll$ size$(\mathbf{A})$ and dim(Range(W))$\ll$ size$(\mathbf{B})$

- Write $\mathbf{U}_k = V\mathbf{Y}W^*$ with $\mathbf{Y}$ to be determined
- Impose Galerkin condition on the residual matrix $\mathbf{R}_k = \mathbf{A}\mathbf{U}_k + \mathbf{U}_k\mathbf{B} - \mathbf{G}$:

$$V^*\mathbf{R}_k W = 0$$

  condition corresponds to $(W \otimes V)^* \mathrm{vec}(\mathbf{R}_\mathrm{k}) = 0$

- Substituting the residual $\mathbf{R}_k$ and $\mathbf{U}_k$:

$$V^*(\mathbf{A}V\mathbf{Y}W^* + V\mathbf{Y}W^*\mathbf{B} - \mathbf{G})W = 0$$

$$V^*\mathbf{A}V\mathbf{Y}W^*W + V^*V\mathbf{Y}W^*\mathbf{B}W - V^*\mathbf{G}W = 0$$

  that is

$$(V^*\mathbf{A}V)\mathbf{Y} + \mathbf{Y}(W^*\mathbf{B}W) - V^*\mathbf{G}W = 0$$

  Reduced (small) Sylvester equation

Galerkin projection methods. Cont'd

From large scale

$$\mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{B} = \mathbf{G}$$

to reduced scale

$$(V^*\mathbf{A}V)\mathbf{Y} + \mathbf{Y}(W^*\mathbf{B}W) = V^*\mathbf{G}W$$

so that $\mathbf{U}_k = V\mathbf{Y}W^* \approx \mathbf{U}$

Key issue: Choice of approximation spaces $\mathcal{V}$=Range(V), $\mathcal{W}$=Range(W)

Galerkin projection methods. Cont'd

From large scale

$$\mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{B} = \mathbf{G}$$

to reduced scale

$$(V^*\mathbf{A}V)\mathbf{Y} + \mathbf{Y}(W^*\mathbf{B}W) = V^*\mathbf{G}W$$

so that $\mathbf{U}_k = V\mathbf{Y}W^* \approx \mathbf{U}$

Key issue: Choice of approximation spaces $\mathcal{V}$=Range(V), $\mathcal{W}$=Range(W)

Desired features:

- "Rich" in problem information while of small size
- Nested, that is, $\mathcal{V}_k \subseteq \mathcal{V}_{k+1}$, $\mathcal{W}_k \subseteq \mathcal{W}_{k+1}$
- Memory/computational-cost effective

Krylov subspaces for Galerkin projection

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}, \qquad \mathbf{G} = \mathbf{G}_1\mathbf{G}_2^*$$

- Standard Krylov subspace

$$\mathcal{V}_k(\mathbf{A}, \mathbf{G}_1) = \mathrm{Range}([\mathbf{G}_1, \mathbf{AG}_1, \mathbf{A}^2\mathbf{G}_1, \ldots, \mathbf{A}^{k-1}\mathbf{G}_1])$$

And analogously,
$\mathcal{W}_k(\mathbf{B}^*, \mathbf{G}_2) = \mathrm{Range}([\mathbf{G}_2, \mathbf{B}^*\mathbf{G}_2, (\mathbf{B}^*)^2\mathbf{G}_2, \ldots, (\mathbf{B}^*)^{k-1}\mathbf{G}_2])$

Krylov subspaces for Galerkin projection

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}, \qquad \mathbf{G} = \mathbf{G}_1 \mathbf{G}_2^*$$

- Standard Krylov subspace

$$\mathcal{V}_k(\mathbf{A}, \mathbf{G}_1) = \mathrm{Range}([\mathbf{G}_1, \mathbf{A}\mathbf{G}_1, \mathbf{A}^2\mathbf{G}_1, \ldots, \mathbf{A}^{k-1}\mathbf{G}_1])$$

And analogously,
$\mathcal{W}_k(\mathbf{B}^*, \mathbf{G}_2) = \mathrm{Range}([\mathbf{G}_2, \mathbf{B}^*\mathbf{G}_2, (\mathbf{B}^*)^2\mathbf{G}_2, \ldots, (\mathbf{B}^*)^{k-1}\mathbf{G}_2])$

- Extended Krylov subspace

$$\begin{aligned} \mathcal{E}\mathcal{V}_k(\mathbf{A}, \mathbf{G}_1) &= \mathcal{V}_k(\mathbf{A}, \mathbf{G}_1) + \mathcal{V}_k(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{G}_1) \\ &= \mathrm{Range}([\mathbf{G}_1, \mathbf{A}^{-1}\mathbf{G}_1, \mathbf{A}\mathbf{G}_1, \mathbf{A}^{-2}\mathbf{G}_1, \mathbf{A}^2\mathbf{G}_1, \mathbf{A}^{-3}\mathbf{G}_1, \ldots,]) \end{aligned}$$

and analogously for $\mathcal{E}\mathcal{W}_k(\mathbf{B}^*, \mathbf{G}_2)$

Krylov subspaces for Galerkin projection

$$\mathbf{AU} + \mathbf{UB} = \mathbf{G}, \qquad \mathbf{G} = \mathbf{G}_1\mathbf{G}_2^*$$

- Standard Krylov subspace

$$\mathcal{V}_k(\mathbf{A}, \mathbf{G}_1) = \mathrm{Range}([\mathbf{G}_1, \mathbf{A}\mathbf{G}_1, \mathbf{A}^2\mathbf{G}_1, \ldots, \mathbf{A}^{k-1}\mathbf{G}_1])$$

And analogously,
$\mathcal{W}_k(\mathbf{B}^*, \mathbf{G}_2) = \mathrm{Range}([\mathbf{G}_2, \mathbf{B}^*\mathbf{G}_2, (\mathbf{B}^*)^2\mathbf{G}_2, \ldots, (\mathbf{B}^*)^{k-1}\mathbf{G}_2])$

- Extended Krylov subspace

$$\begin{aligned} \mathcal{EV}_k(\mathbf{A}, \mathbf{G}_1) &= \mathcal{V}_k(\mathbf{A}, \mathbf{G}_1) + \mathcal{V}_k(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{G}_1) \\ &= \mathrm{Range}([\mathbf{G}_1, \mathbf{A}^{-1}\mathbf{G}_1, \mathbf{A}\mathbf{G}_1, \mathbf{A}^{-2}\mathbf{G}_1, \mathbf{A}^2\mathbf{G}_1, \mathbf{A}^{-3}\mathbf{G}_1, \ldots,]) \end{aligned}$$

and analogously for $\mathcal{EW}_k(\mathbf{B}^*, \mathbf{G}_2)$

- Rational Krylov subspace

$$\mathcal{RV}_k(\mathbf{A}, \mathbf{G}_1, \mathbf{s}) = \mathrm{Range}([\mathbf{G}_1, (\mathbf{A} - s_2\mathbf{I})^{-1}\mathbf{G}_1, \ldots, (\mathbf{A} - s_k\mathbf{I})^{-1}\mathbf{G}_1])$$

and analogously for $\mathcal{RW}_k(\mathbf{B}^*, \mathbf{G}_2, \mathbf{s})$

Krylov subspaces for Galerkin projection. Cont'd

In all these instances, spaces are nested

$$\mathcal{V}_k \subset \mathcal{V}_{k+1}, \qquad \mathcal{W}_k \subset \mathcal{W}_{k+1}$$

At step k:

$$(V_k^* \mathbf{A} V_k)\mathbf{Y} + \mathbf{Y}(W_k^* \mathbf{B} W_k) = V_k^* \mathbf{G} W_k$$

At step $k+1$:

$$(V_{k+1}^* \mathbf{A} V_{k+1})\mathbf{Y} + \mathbf{Y}(W_{k+1}^* \mathbf{B} W_{k+1}) = V_{k+1}^* \mathbf{G} W_{k+1}$$

$\Rightarrow$ Size of reduced problem increases

More general FDE settings. I

- Fractional derivatives both in time and 2D space:

$$ {}_0^C D_t^{\alpha} u - {}_x D_R^{\beta_1} u - {}_y D_R^{\beta_2} u = f, \quad (x, y, t) \in \Omega \times (0, T], $$

yields a linear system with a double tensor structure:

$$ \left(\mathbf{T}_{\alpha}^{n_t} \otimes \mathbf{I}^{n_x n_y} - \mathbf{I}^{n_t} \otimes \mathbf{L}_{\beta_1, \beta_2}^{n_x n_y} \right) \mathbf{u} = \tilde{\mathbf{f}} $$

with $\mathbf{L}_{\beta_1, \beta_2}^{n_x n_y} = \left(\mathbf{I}^{n_y} \otimes \mathbf{L}_{\beta_1}^{n_x} + \mathbf{L}_{\beta_2}^{n_y} \otimes \mathbf{I}^{n_x} \right)$

More general FDE settings. I

- Fractional derivatives both in time and 2D space:

$$ {}_0^C D_t^{\alpha} u - {}_x D_R^{\beta_1} u - {}_y D_R^{\beta_2} u = f, \quad (x,y,t) \in \Omega \times (0,T], $$

yields a linear system with a double tensor structure:

$$ \left(\mathbf{T}_{\alpha}^{n_t} \otimes \mathbf{I}^{n_x n_y} - \mathbf{I}^{n_t} \otimes \mathbf{L}_{\beta_1,\beta_2}^{n_x n_y} \right) \mathbf{u} = \tilde{\mathbf{f}} $$

with $\mathbf{L}_{\beta_1,\beta_2}^{n_x n_y} = \left(\mathbf{I}^{n_y} \otimes \mathbf{L}_{\beta_1}^{n_x} + \mathbf{L}_{\beta_2}^{n_y} \otimes \mathbf{I}^{n_x} \right)$

$\Rightarrow$ Numerical solution: tensor-train (TT) representation, and Alternating Minimal Energy Method (AMEN)

(Oseledets 2011, Tyrtyshnikov, Oseledets etal, Oseledets & Dolgov 2012)

More general FDE settings. II

- Variable coefficient case:

$$\frac{du}{dt} = p_+(x,y)\,{}^{\mathrm{RL}}_{\;0}D_x^{\beta_1}\,u + p_-(x,y)\,{}^{\mathrm{RL}}_{\;x}D_1^{\beta_1}\,u$$
$$+q_+(x,y)\,{}^{\mathrm{RL}}_{\;0}D_y^{\beta_2}\,u + q_-(x,y)\,{}^{\mathrm{RL}}_{\;y}D_1^{\beta_2}\,u + f, \qquad (x,y,t) \in \Omega \times (0,T]$$

(separable coefficients, e.g., $p_+(x,y) = p_{+,1}(x)p_{+,2}(y)$)

After the Grünwald-Letnikov discretization, we obtain $\mathbf{A}\mathbf{u} = \mathbf{f}$ with

$$\mathbf{A} = \mathbf{T}_\alpha^{n_t} \otimes \mathbf{I}^{n_2 n_1} - \mathbf{I}^{n_t} \otimes \mathbf{I}^{n_2} \otimes \left(\mathbf{P}_+\mathbf{T}_{\beta_1} + \mathbf{P}_-\mathbf{T}_{\beta_1}^\top\right) - \mathbf{I}^{n_t} \otimes \left(\mathbf{Q}_+\mathbf{T}_{\beta_2} + \mathbf{Q}_-\mathbf{T}_{\beta_2}^\top\right) \otimes \mathbf{I}^{n_1}$$

$\mathbf{T}_\beta$: one-sided derivative

$\mathbf{P}_+, \mathbf{P}_-, \mathbf{Q}_+, \mathbf{Q}_-$: diagonal matrices with the grid values of p_+, p_-, q_+, q_-

More general FDE settings. II

- Variable coefficient case:

$$\frac{du}{dt} = p_+(x,y)\, {}^{\mathrm{RL}}_{\ 0}D^{\beta_1}_{x}\, u + p_-(x,y)\, {}^{\mathrm{RL}}_{\ x}D^{\beta_1}_{1}\, u$$
$$+q_+(x,y)\, {}^{\mathrm{RL}}_{\ 0}D^{\beta_2}_{y}\, u + q_-(x,y)\, {}^{\mathrm{RL}}_{\ y}D^{\beta_2}_{1}\, u + f, \qquad (x,y,t) \in \Omega \times (0,T]$$

(separable coefficients, e.g., $p_+(x,y) = p_{+,1}(x)p_{+,2}(y)$)

After the Grünwald-Letnikov discretization, we obtain $\mathbf{A}\mathbf{u} = \mathbf{f}$ with

$$\mathbf{A} = \mathbf{T}^{n_t}_{\alpha} \otimes \mathbf{I}^{n_2 n_1} - \mathbf{I}^{n_t} \otimes \mathbf{I}^{n_2} \otimes \left(\mathbf{P}_+\mathbf{T}_{\beta_1} + \mathbf{P}_-\mathbf{T}^{\top}_{\beta_1}\right) - \mathbf{I}^{n_t} \otimes \left(\mathbf{Q}_+\mathbf{T}_{\beta_2} + \mathbf{Q}_-\mathbf{T}^{\top}_{\beta_2}\right) \otimes \mathbf{I}^{n_1}$$

$\mathbf{T}_{\beta}$: one-sided derivative

$\mathbf{P}_+, \mathbf{P}_-, \mathbf{Q}_+, \mathbf{Q}_-$: diagonal matrices with the grid values of p_+, p_-, q_+, q_-

$\Rightarrow$ Various numerical strategies:

- TT format

- GMRES solver with Sylvester-equation-based preconditioner

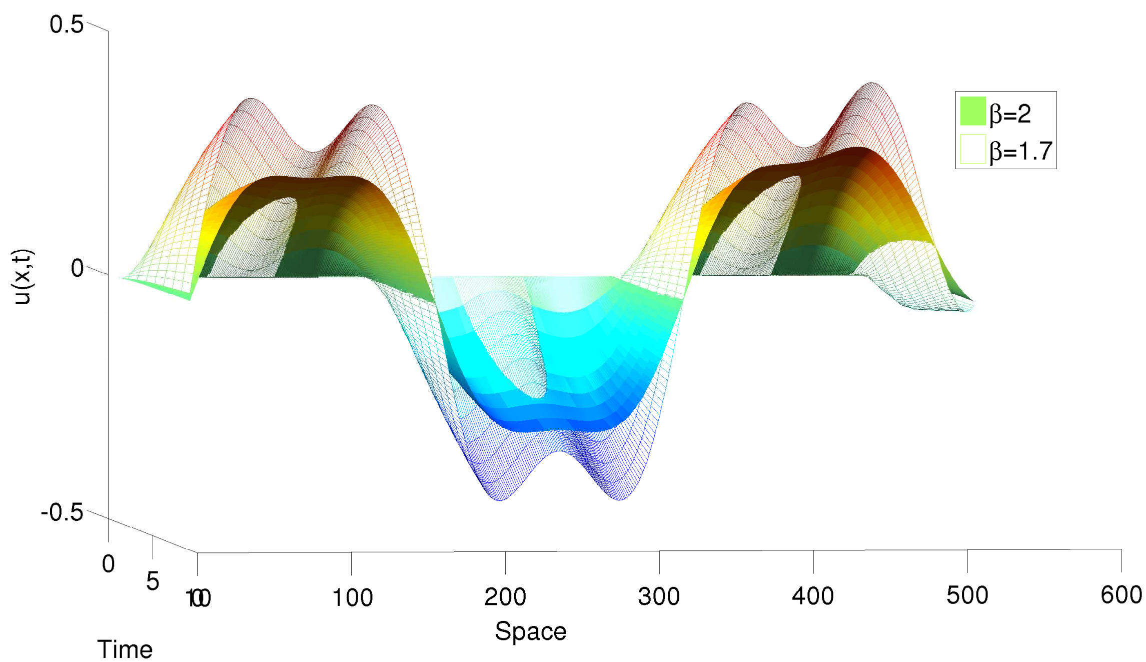
- Sylvester solver for special cases (our example)

(see also talk by Mariarosa Mazza)

Numerical experiments

Goals:

- Test robustness wrt time and space discretization parameters
- Suitability for different orders of differentiation



1D (space) pb, constant coeff. Two different values for β with zero initial condition and zero Dirichlet boundary condition. (510 and 10 space and time points, resp.)

1D problem. Constant/variable coeffs.

$f = 80\sin(20x)\cos(10x)$

Variable coefficient case: $p_+ = \Gamma(1.2)x_1^{\beta_1}$ and $p_- = \Gamma(1.2)(2-x_1)^{\beta_1}$.

Solvers:

⋆ CG preconditioned by Strang circulant precond (constant coeff)

⋆ GMRES preconditioned by approx constant coeff oper. (variable coeff)

⋆ Convergence tolerance 10^{-6}, Time-step $\tau = h_x/2$.

	Constant Coeff.		Variable Coeff.	
n_x	$\beta = 1.3$	$\beta = 1.7$	$\beta = 1.3$	$\beta = 1.7$
32768	6.0 (*0.43*)	7.0 (*0.47*)	6.8 (*0.90*)	6.0 (*0.81*)
65536	6.0 (*0.96*)	7.0 (*0.97*)	6.4 (*1.93*)	6.0 (*1.75*)
131072	6.0 (*1.85*)	7.0 (*2.23*)	5.9 (*3.93*)	5.9 (*3.89*)
262144	6.0 (*7.10*)	7.1 (*8.04*)	5.4 (*12.78*)	6.0 (*13.52*)
524288	6.0 (*15.42*)	7.8 (*19.16*)	5.1 (*25.71*)	6.0 (*27.40*)
1048576	6.0 (*34.81*)	8.0 (*41.76*)	4.9 (*51.02*)	6.3 (*62.57*)

2D (space) problem. Constant/variable coeffs.

$F = 100\sin(10x)\cos(y) + \sin(10t)xy$

Variable coefficient case:
$p_+ = \Gamma(1.2)x^{\beta_1}, p_- = \Gamma(1.2)(2-x)^{\beta_1}, q_+ = \Gamma(1.2)y^{\beta_2}, q_- = \Gamma(1.2)(2-y)^{\beta_2}$

Solvers:

⋆ Extended Krylov projection method (constant coeff)

⋆ GMRES preconditioned by Sylvester solver (variable coeff)

⋆ Convergence tolerance 10^{-6} Time-step $\tau = h_x/2$. solution after 8 time steps

		Variable Coeff.		Constant Coeff.	
n_x	n_y	$\beta_1 = 1.3$ $\beta_2 = 1.7$ it(CPUtime)	$\beta_1 = 1.7$ $\beta_2 = 1.9$ it(CPUtime)	$\beta_1 = 1.3$ $\beta_2 = 1.7$ it(CPUtime)	$\beta_1 = 1.7$ $\beta_2 = 1.9$ it(CPUtime)
1024	1024	2.3 (*19.35*)	2.5 (*15.01*)	2.9 (*9.89*)	2.5 (*18.51*)
1024	2048	2.8 (*47.17*)	2.9 (*22.25*)	3.0 (*23.07*)	3.2 (*22.44*)
2048	2048	3.0 (*76.72*)	2.6 (*36.01*)	2.0 (*51.23*)	2.3 (*34.43*)
4096	4096	3.0 (*171.30*)	2.6 (*199.82*)	2.0 (*164.20*)	2.2 (*172.24*)

Conclusions

- FDEs provide challenging linear algebra settings
- The two fields are evolving in parallel
- Effective numerical solution for representative large problems

References

T. Breiten, V. Simoncini, M. Stoll, *Low-rank solvers for fractional differential equations*, ETNA (2016) v.45, pp.107-132.

V. Simoncini, *Computational methods for linear matrix equations*, SIAM Review, Sept. 2016.