Applied Mathematics and Computation 273 (2016) 604–623

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Fast tensor product solvers for optimization problems with fractional differential equations as constraints

Sergey Dolgov [a,*], John W. Pearson [b], Dmitry V. Savostyanov [c], Martin Stoll [a]

[a] Numerical Linear Algebra for Dynamical Systems, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany

[b] School of Mathematics, Statistics and Actuarial Science, University of Kent, Cornwallis Building (East), Canterbury, Kent, CT2 7NZ, United Kingdom

[c] School of Computing, Engineering and Mathematics, University of Brighton, Moulsecoomb, Brighton, BN2 4GJ, United Kingdom

ARTICLE INFO

Keywords:
Fractional calculus
Iterative solvers
Sylvester equations
Preconditioning
Low-rank methods
Tensor equations
Schur complement

ABSTRACT

Fractional differential equations have recently received much attention within computational mathematics and applied science, and their numerical treatment is an important research area as such equations pose substantial challenges to existing algorithms. An optimization problem with constraints given by fractional differential equations is considered, which in its discretized form leads to a high-dimensional tensor equation. To reduce the computation time and storage, the solution is sought in the tensor-train format. We compare three types of solution strategies that employ sophisticated iterative techniques using either preconditioned Krylov solvers or tailored alternating schemes. The competitiveness of these approaches is presented using several examples with constant and variable coefficients.

1. Introduction

While the study of derivatives of arbitrary order is a long-standing subject area [31], its use in science and engineering has soared over recent years. Fractional calculus is often used due to the inadequateness of traditional schemes to describe certain phenomena, such as anomalous diffusion, anaelasticity [16] and viscoelasticity [45,88]. The applications include electrical circuits [40,73], electro-analytical chemistry [84], biomechanics [28], and image processing [91].

Over the last decade many researchers have worked on efficient numerical schemes for the discretization and solution of fractional differential equations (FDEs). Historically, the finite difference-based discretization techniques are arguably the most popular [51,52,71–73]. Adomian decomposition should be mentioned as a popular semi-analytical approach [1], although its use is limited. Recently, (discontinuous) finite element schemes for FDEs have also received considerable attention [17,59,90]. Since the fractional differential operators are, in fact, integral operators, the matrix of the corresponding linear system is usually dense, and the numerical complexity and storage grow rapidly with the grid size, especially for the problems posed in higher dimensions (e.g. three spatial plus a temporal dimension). To perform computations faster, we compress the solution in the low-rank format. This approach has already been applied to fractional calculus in [13,75].

Most commonly, the fractional calculus literature focuses on the solution of the equation itself, the so-called 'direct problem'. In this paper we consider the 'inverse problem', namely the computation of the forcing term (right-hand side of the FDE), that is

* Corresponding author. Tel.: +49 391 6110 450; fax: +49 391 6110 453.
E-mail addresses: sergey.v.dolgov@gmail.com, dolgov@mpi-magdeburg.mpg.de (S. Dolgov), J.W.Pearson@kent.ac.uk (J.W. Pearson), d.savostyanov@brighton.ac.uk (D.V. Savostyanov), stollm@mpi-magdeburg.mpg.de (M. Stoll).

http://dx.doi.org/10.1016/j.amc.2015.09.042
0096-3003/

best suited to describing a desired property or measured data. For this we study an optimization problem with constraints given by FDEs. In the context of partial differential equations (PDEs), problems of this type are often referred to as PDE-constrained optimization problems and have been studied extensively over the last decades (see [38,85] for introductions to the field). Optimal control problems for FDEs have previously been studied in literature such as [2,3,55,56,65,74]. However, they were mostly considering one-dimensional spatial domains, since the direct treatment of higher dimensions was too expensive. To overcome computational challenges, in this paper we rely on the recent advances in the development of numerical algorithms and solvers, particularly on data-sparse low-rank formats.

The goal of our paper is to present efficient numerical methods that allow the fast and accurate solution of the large optimization problem at hand. The paper is organized as follows. In Section 2.1 we recall some of the most important definitions needed for fractional derivatives. This is followed by Section 2.2 where we introduce the basic optimization problem subject to fractional differential equations posed in an increasing number of dimensions, along with the discretization of both the objective function and the differential equation. Section 3 presents three strategies that are well-suited to solving the discretized problem. This is followed by a discussion of numerical algorithms in Section 4 where we introduce the tensor-train format and several iterative solvers, either of Krylov subspace type or using an alternating framework. The effectiveness of our approach is shown in Section 5 where we compare our solvers using several numerical experiments. We also present results for the more challenging variable coefficient case and observe satisfying results.

Another approach that has recently been studied by Burrage et al. is to consider a matrix function approach to solve the discretized system (see [15] for details). Our work here is motivated by some recent results in [72] where the discretization via finite differences is considered in a purely algebraic framework.

2. Fractional calculus and Grünwald formulae

In this section we briefly recall the concept of fractional derivatives, and use this to state the matrix systems that result from discretizing the problems we consider using a finite difference method. The literature on fractional derivatives is vast and we refer to [18,31,35,54,69,70] for general introductions to this topic.

2.1. The fractional derivative

In fractional calculus there are several definitions of fractional derivatives. The Caputo and the Riemann–Liouville fractional derivatives [69] are among the most commonly used in applications and we use this section to briefly recall their definitions.

For a function $f(t)$ defined on an interval $[a, b]$, the Caputo derivative of real order α with $n-1 < \alpha < n$, $n \in \mathbb{N}$, is defined as the following integral

$$ {}_{a}^{C}D_t^{\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t \frac{d^n f(s)}{ds^n} \frac{ds}{(t-s)^{\alpha-n+1}}, $$

assuming that it is convergent (see [19,31,51,69,78] for more details). Based on the discussion in [72], the Caputo derivative is frequently used for the derivative with respect to time. The left-sided Riemann–Liouville derivative of real order α with $n-1 < \alpha < n$, is defined by

$$ {}_{a}^{RL}D_t^{\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_a^t \frac{f(s)ds}{(t-s)^{\alpha-n+1}}, $$

for $a < t < b$. The right-sided Riemann–Liouville fractional derivative is given by

$$ {}_{t}^{RL}D_b^{\alpha} f(t) = \frac{(-1)^n}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_t^b \frac{f(s)ds}{(s-t)^{\alpha-n+1}}, $$

for $a < t < b$. Finally, the symmetric Riesz derivative of order α is the half-sum of the left and right-side Riemann–Liouville derivative, i.e.,

$$ {}^{R}D_t^{\alpha} f(t) = \frac{1}{2}\left({}_{a}^{RL}D_t^{\alpha} f(t) + {}_{t}^{RL}D_b^{\alpha} f(t) \right). $$

In this work we do not advocate a particular method that is most suitable for the description of a natural phenomenon: we simply want to illustrate that the above formulations of FDEs, when coupled with certain types of discretization approaches, lead to similar structures on the discrete level. Our goal is to give guidelines and offer numerical schemes for the efficient and accurate solution of problems of various forms. For a discussion on the smoothness assumptions of the function $f(\cdot)$ we refer to [69], i.e., the Riemann–Liouville formulation requires a weaker differentiability assumption in contrast to the n-times differentiability required otherwise (cf. [69, Chapter 2.3]).

2.2. Model problems

In this section we introduce some FDE-constrained optimization problems. Consider the classical misfit problem, where we want to minimize the difference between the *state* y and the desired state (or *observation*) $\bar{y}$, with an additional regularization