Applied Mathematics and Computation 314 (2017) 310–321

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Study on fractional order gradient methods

Yuquan Chen[a], Qing Gao[b], Yiheng Wei[a], Yong Wang[a,*]

[a] Department of Automation, University of Science and Technology of China, Hefei 230026, China
[b] School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2600, Australia

A R T I C L E I N F O

Keywords:
Fractional order calculus
Gradient methods
Iterative algorithms
Convergence capability

A B S T R A C T

In this paper, convergence capability of the conventional fractional order gradient methods (FOGMs) is analyzed and a new FOGM with guaranteed and faster convergence ability is proposed. In particular, we further consider the case that the derivative order varies between 1 and 2, and give some discussion on how to determine the derivative order in the algorithm settings. Simulation results from a number of numerical examples are provided to illustrate the approaches.

1. Introduction

Fractional order calculus is a natural generalization of classical integer order calculus and is particularly useful in studying engineering plants that behave anomalously such as diffusion process [1], electro-chemistry [2], viscoelasticity [3], and heat conduction [4]. It is also an important tool for improving the performance of many control schemes, such as sliding mode control [5–8] and adaptive control [9–11]. Basically, fractional order calculus reflects not only local characteristic at the evaluated time but also entire history of the function [12]. This brings new perspectives in developing new optimization algorithms which we mainly concern in this paper. An important application is the optimal control theory using fractional order calculus, see, e.g., [13,14]. The authors in [13] developed an extremum seeking control strategy with the aid of fractional order calculus and achieved a faster seeking speed compared with those using classical calculus. In [14], a fractional order minimum energy cognitive lighting control strategy was derived from the extremum seeking method in [13] and applied to a hybrid lighting system.

As a standard optimization algorithm, the gradient method (GM) has been widely used in many engineering applications like adaptive filtering [15,16], image processing [17–19], and system identification [20–22]. However, research of the fractional order gradient method (FOGM) is still in its infancy and deserves further investigation. In [23], the authors proposed a FOGM by using Caputo's fractional order derivative with a derivative order no more than 1 as the iterative direction, instead of an integer order derivative. It was found that a smaller weight noise can be achieved if a smaller derivative order is used, while the algorithm converges faster if a bigger derivative order is used. A similar idea can be found in [24] where a different Riemann–Liouville's fractional order derivative was used to develop a fractional steepest descent method. However, one cannot guarantee that the extreme point can be found using the method in [24] even if the algorithm there is indeed convergent. This shortcoming has been well overcome in [25]. Despite some minor errors with the calculation procedure, the method developed in [25] has been successfully applied in speech enhancement [26], noise suppression [27], and system identification [28,29] for many systems, such as CARMA systems [30], Hammerstein systems [31,32], and Box–Jenkins systems [33,34].

* Corresponding author.
E-mail address: yongwang@ustc.edu.cn (Y. Wang).

http://dx.doi.org/10.1016/j.amc.2017.07.023

It is worth pointing out that most existing work on FOGM in literature concerns the problem of quadratic function optimization only, due to the fact that fractional order derivative of a general function often has no explicit solutions. In addition, the convergence capability of the existing FOGMs depends on the initial integral point of fractional order derivative, and even the additional constant terms when using Riemann–Liouville's fraction order derivative, which is undesirable in practice. Motivated by these issues, in this paper we will propose an alternative FOGM that can be used in searching for the extreme point of a general convex function and the main contributions can be summarized as follows:

1) The reasons why some existing FOGMs cannot guarantee the convergence are carefully analyzed;
2) Two FOGMs with guaranteed convergence ability are proposed for Caputo's and Riemann–Liouville's derivative cases, respectively. The faster convergence ability of the proposed methods compared with the existing ones is also shown;
3) With the aid of infinite series representation, a simplified and convenient FOGM is developed;
4) The results in 1)–3) are extended to the case that the derivative order varies between 1 and 2; and
5) Some analyses on how the derivative order influences the performance of algorithms are provided.

The remainder of the article is organized as follows. Section 2 is devoted to math preparation and problem formulation. The key contributions are then presented in Sections 3 and 4, focusing on analyzing the existing FOGMs and developing new FOGMs using Caputo and Riemann–Liouville's definitions respectively. Some simulation examples are provided to demonstrate the effectiveness of the proposed methods in Section 5. The article is concluded in Section 6.

Note: In the following discussion, t^{α} with $t,\ \alpha \in \mathbb{R}$ always means the value of its real part.

2. Preliminaries and problem formulation

Definition 1. [35] For a scalar convex function $f(t)$ whose first order derivative is existing, if there exists a scalar $\lambda > 0$ such that

$$\left|f^{(1)}(x) - f^{(1)}(y)\right| \geq \lambda |x-y|, \tag{1}$$

for any x and y belong to the definition domain of $f(t)$, then $f(t)$ is said to be strong convex.

For any constant $n-1<\alpha<n,\ n\in\mathbb{N},$ the Caputo's derivative with order α for a smooth function $f(t)$ is given by

$${}_{t_0}^{\mathrm{C}}\mathscr{D}_t^{\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_{t_0}^{t} \frac{f^{(n)}(\tau)}{(t-\tau)^{\alpha-n+1}} \mathrm{d}\tau. \tag{2}$$

Alternatively, (2) can be rewritten in a form similar as the conventional Taylor series:

$${}_{t_0}^{\mathrm{C}}\mathscr{D}_t^{\alpha} f(t) = \sum_{i=n}^{\infty} \frac{f^{(i)}(t_0)}{\Gamma(i+1-\alpha)} (t-t_0)^{i-\alpha}. \tag{3}$$

The Riemann–Liouville's derivative with α for $f(t)$ is given by

$${}_{t_0}^{\mathrm{RL}}\mathscr{D}_t^{\alpha} f(t) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left[\frac{1}{\Gamma(n-\alpha)} \int_{t_0}^{t} \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} \mathrm{d}\tau\right], \tag{4}$$

or a series form as

$${}_{t_0}^{\mathrm{RL}}\mathscr{D}_t^{\alpha} f(t) = \sum_{i=0}^{\infty} \frac{f^{(i)}(t_0)}{\Gamma(i+1-\alpha)} (t-t_0)^{i-\alpha}. \tag{5}$$

Suppose $f(t)$ to be a smooth convex function with a unique extreme point t^*. It is well known that each iterative step of the conventional gradient method is formulated as

$$t_{k+1} = t_k - \rho f^{(1)}(t_k), \tag{6}$$

where $\rho > 0$ is the iteration step size.

The basic idea of FOGMs is then replacing the first order derivative in Eq. (6) by its fractional order counterpart, either using Caputo or Riemann–Liouville's definition. However, as we will show later that such a heuristic approach cannot guarantee the convergence capability of the algorithms. The key purposes of this paper will be explaining why divergence happens with the conventional FOGMs and proposing alternative FOGMs whose convergence can be guaranteed. Discussion will be given for the cases of using Caputo and Riemann–Liouville's definitions, respectively.

3. FOGM using Caputo's derivative

Let $f(t)$ be a smooth convex function with a unique extreme point t^*. In searching for the extreme point t^*, each iterative step of FOGM using Caputo's derivative definition is formulated as

$$t_{k+1} = t_k - \rho\, {}_{t_0}^{\mathrm{C}}\mathscr{D}_{t_k}^{\alpha} f(t), \tag{7}$$