J Optim Theory Appl (2014) 163:884–899
DOI 10.1007/s10957-013-0511-5

Epsilon-Ritz Method for Solving a Class of Fractional Constrained Optimization Problems

Ali Lotfi · Sohrab Ali Yousefi

Received: 3 December 2013 / Accepted: 16 December 2013 / Published online: 29 January 2014
© Springer Science+Business Media New York 2014

Abstract In this paper, epsilon and Ritz methods are applied for solving a general class of fractional constrained optimization problems. The goal is to minimize a functional subject to a number of constraints. The functional and constraints can have multiple dependent variables, multiorder fractional derivatives, and a group of initial and boundary conditions. The fractional derivative in the problem is in the Caputo sense. The constrained optimization problems include isoperimetric fractional variational problems (IFVPs) and fractional optimal control problems (FOCPs). In the presented approach, first by implementing epsilon method, we transform the given constrained optimization problem into an unconstrained problem, then by applying Ritz method and polynomial basis functions, we reduce the optimization problem to the problem of optimizing a real value function. The choice of polynomial basis functions provides the method with such a flexibility that initial and boundary conditions can be easily imposed. The convergence of the method is analytically studied and some illustrative examples including IFVPs and FOCPs are presented to demonstrate validity and applicability of the new technique.

Keywords Epsilon method · Ritz method · Fractional derivative · Fractional variational problem · Fractional optimal control problem

1 Introduction

In this work, our focus is placed on developing an efficient approximate method for solving the fractional constrained optimization problems. A fractional constrained

A. Lotfi · S. A. Yousefi (✉)
Department of Mathematics, Shahid Beheshti University, G.C, Tehran, Iran
e-mail: s-yousefi@sbu.ac.ir

A. Lotfi
e-mail: ali292lotfi@gmail.com

optimization problem can be considered as an isoperimetric fractional variational problem (IFVP) or fractional optimal control problem (FOCP). Thus, it is worthwhile to find an efficient approximate method for solving such problems.

Fractional variational problems (FVPs), isoperimetric fractional variational problems, and FOCPs are three different types of fractional optimization problems. General optimality conditions have been developed for FVPs, IFVPs, and FOCPs. For instance in [1], the author has achieved the necessary optimality conditions for FVPs and IFVPs with Riemann-Liouville derivatives. Hamiltonian formulas for fractional optimal control problems with Riemann-Liouville fractional derivatives have been derived in [2,3]. In [4] the authors present necessary and sufficient optimality conditions for a class of FVPs with respect to Caputo fractional derivative. Agrawal [5] provides Hamiltonian formulas for FOCPs with Caputo fractional derivatives. Optimality conditions for fractional variational problems with functionals containing both fractional derivatives and integrals are presented in [6]. Such formulas are also developed for FVPs with other definitions of fractional derivatives in [7,8]. Agrawal [9] includes discussion about a General form of FVPs; the author claims that the derived equations are general form of Euler-Lagrange equations for problems with fractional Riemann-Liouville, Caputo, Riesz-Caputo, and Riesz-Riemann-Liouvile derivatives. Other generalizations of Euler-Lagrange equations for problems with free boundary conditions can be found in [10–13] as well. It is known that optimal solution of fractional variational and optimal control problems should satisfy Euler-Lagrange and Hamiltonian systems, respectively [1–3,5]. Hence, solving Euler-Lagrange equations and Hamiltonian systems leads to optimal solution of FVPs or FOCPs. Except for some special cases in FVPs [14], it is hard to find exact solution for Euler-Lagrange and Hamiltonian equations, specially when the problem has boundary conditions. Examples of numerical simulations for fractional optimal control problems with Riemann-Liouville fractional derivatives can be found in [2,3,15–17]. There also exist some numerical methods for solving fractional variational problems. For instance finite element method in [18,19] and fractional variational integrator in [20] are developed and applied for some classes of FVPs. In [21], the classical discrete direct method for solving variational problems is generalized for FVPs. Numerical simulations for FOCPs with Caputo fractional derivatives are developed in [5] and [22], where the author has solved the Hamiltonian equations approximately. A general class of FVPs in [23] and a class of FOCPs in [24] are solved directly without using Euler-Lagrange and hamiltonian formulas. Through the use of operational matrix of fractional integration and gauss quadrature formula, [25] presents approximate direct method for solving a class of FOCPs. An approximate method for solving FVPs with Lagrangian containing fractional integral operators is provided in [26]; the authors approximately transform fractional problem into regular problem by decomposing fractional integral operator with the finite series in terms of derivative operators. We refer readers interested in fractional calculus of variations to [27].

The epsilon method has been first introduced by Balakrishnan in [28]. Later, Frick [29,30] developed the method for solving optimal control problems. In this paper, we apply combination of Ritz and epsilon methods for solving fractional constrained optimization problems. These problems can also have a group of boundary conditions. Our development in Epsilon-Ritz method has the property that the approximate solutions satisfy all initial and boundary conditions of the problem. The implemented method reduces the given constrained optimization problem to the problem of finding optimal solution of a real value function. First, unknown functions are expanded with polynomial basis and unknown coefficients, then an algebraic function, which should be optimized with respect to its variables, in terms of unknown coefficients is achieved. We study the convergence of the approximate method and present numerical examples to illustrate the applicability of the new approach.

This paper is organized as follows. Section 2 presents problem formulation. In Sect. 3, epsilon method is applied to reduce constrained optimization problem of Sect. 2 to an unconstrained problem. In Sect. 4 we solve the unconstrained problem, constructed in Sect. 3, using Ritz method, to achieve an approximate solution of the main problem. Section 5 discusses on the convergence of the method presented in Sect. 4 and finally Sect. 6 reports numerical findings and demonstrates the accuracy of the numerical scheme by considering some test examples. Section 7 consists of a brief summary.

2 Statement of the Problem

Consider the following fractional constrained optimization problem:

$$\min \quad J[y_1, \dots, y_m] = \int_{t_0}^{t_1} F(t, y_1, \dots, y_m, \dots, {}_{t_0}^{C}D_t^{\alpha_r} y_r, \dots)\mathrm{d}t,$$

subject to

$$G_l(t, y_1, \dots, y_m, \dots, {}_{t_0}^{C}D_t^{\alpha_r} y_r, \dots) = 0, \quad l = 1, \dots, L,$$

where $n - 1 < \alpha_r \leq n$, $n \in \mathbb{N}$, L is the number of constraints, functions F and G are continuously differentiable with respect to all their arguments, and functional J is bounded from below, i.e., there exist $\lambda \in \mathbb{R}$ such that $J[y_1, \dots, y_m] \geq \lambda$. The fractional derivatives are defined in the Caputo sense

$${}_{t_0}^{C}D_t^{\alpha} y(t) = \frac{1}{\Gamma(n-\alpha)} \int_{t_0}^{t} (t-\tau)^{n-\alpha-1} y^{(n)}(\tau)\mathrm{d}\tau, \quad n - 1 < \alpha < n,$$

In cases when $\alpha = n$, the Caputo derivative is defined ${}_{t_0}^{C}D_t^{\alpha} y(t) = y^{(n)}(t)$. In the above problem functionals F and G can contain fractional derivatives for some of the variables y_j, $1 \leq j \leq m$, (not necessarily for all y_js) and each variable can have initial or boundary conditions.

3 Epsilon Method

Without loss of generality, we let $t_0 = 0$, $t_1 = 1$, and $t \in [0, 1]$ in the problem of Sect. 2.

$$\min \quad J[y_1, \ldots, y_m] = \int_0^1 F(t, y_1, \ldots, y_m, \ldots, {}_0^C D_t^{\alpha_r} y_r, \ldots) \mathrm{d}t \tag{1}$$

$$G_l(t, y_1, \ldots, y_m, \ldots, {}_0^C D_t^{\alpha_r} y_r, \ldots) = 0, \quad l = 1, \ldots, L, \tag{2}$$

where $n-1 < \alpha_r \leq n$, $1 \leq r \leq m$ and L is the number of constraints. In the problem (1) – (2) each variable y_j, $1 \leq j \leq m$, can be considered in the following three cases:

(i) y_j has no derivative, neither initial nor boundary conditions.

(ii) y_j has fractional derivative of order at most α_j and initial conditions

$$y_j^{(i)}(0) = y_{j0}^i, \quad 0 \leq i \leq \lceil \alpha_j \rceil - 1.$$

(iii) y_j has fractional derivative of order at most α_j and initial and boundary conditions

$$y_j^{(i)}(0) = y_{j0}^i, \quad y_j^{(i)}(1) = y_{j1}^i, \quad 0 \leq i \leq \lceil \alpha_j \rceil - 1.$$

In this paper, it is considered that the constrained problem (1) – (2) has minimum $\mu = J[y_\mu^1, \ldots, y_\mu^m]$ on

$$X = \{(y_1, \ldots, y_m) \in \prod_{j=1}^m E_j[0,1] : G_l(t, y_1, \ldots, y_m, \ldots, {}_0^C D_t^{\alpha_r} y_r, \ldots) = 0,\ 1 \leq l \leq L\},$$

where

$$E_j[0,1] = C[0,1],$$

when y_j belongs to the case (i),

$$E_j[0,1] = \{y(t) \in C^{\lceil \alpha_j \rceil}[0,1] : y_j^{(i)}(0) = y_{j0}^i,\ 0 \leq i \leq \lceil \alpha_j \rceil - 1\},$$

when y_j belongs to the case (ii), and

$$E_j[0,1] = \{y(t) \in C^{\lceil \alpha_j \rceil}[0,1] : y_j^{(i)}(0) = y_{j0}^i,\ y_j^{(i)}(1) = y_{j1}^i,\ 0 \leq i \leq \lceil \alpha_j \rceil - 1\},$$

when y_j belongs to the case (iii).

Note that here $(C^n[0,1], \| \cdot \|_n)$ is the Banach space

$$C^n[0,1] = \{f(t) : f^{(n)}(t) \in C[0,1]\},$$

where

$$\| f \|_n = \| f \|_\infty + \| f' \|_\infty + \cdots + \| f^{(n)} \|_\infty .$$

Consider the following optimization problem:

$$\min \quad J_\epsilon[y_1,\ldots,y_m] = \int_0^1 F(t, y_1,\ldots,y_m,\ldots,{}_0^C D_t^{\alpha_r} y_r,\ldots)\mathrm{d}t + \frac{1}{\epsilon}\sum_{l=1}^{L}\int_0^1 G_l^2(t, y_1,\ldots,y_m,\ldots,{}_0^C D_t^{\alpha_r} y_r,\ldots)\mathrm{d}t, \tag{3}$$

where $\epsilon > 0$ is given.

We solve the unconstrained problem (3) instead of the constrained optimization problem (1)–(2) for sufficiently small value of ϵ. Theorem 5.3 ensures that solving problem (3) with Ritz method leads to an approximate solution for the problem (1)–(2).

4 Ritz Approximation Method

Since Legendre Polynomials have been applied to approximate functions in the subsequent development, we state some basic properties of these polynomials. Of course it is possible to use other types of polynomials, such as Taylor, Bernstein, etc, for approximations.

4.1 Legendre Polynomials

The Legendre polynomials are orthogonal polynomials on the interval $[-1, 1]$ and can be determined with the following recurrence formula:

$$L_{i+1}(y) = \frac{2i+1}{i+1} y L_i(y) - \frac{i}{i+1} L_{i-1}(y), \quad i = 1, 2, 3, \ldots$$

where $L_0(y) = 1$ and $L_1(y) = y$. By the change of variable $y = 2t - 1$ we will have the well-known shifted Legendre polynomials. Let $p_m(t)$ be the shifted Legendre polynomials of order m which are defined on the interval $[0, 1]$ and can be determined with the following recurrence formula

$$p_{m+1}(t) = \frac{2m+1}{m+1}(2t-1)p_m(t) - \frac{m}{m+1}p_{m-1}(t), \quad m = 1, 2, 3, \ldots$$
$$p_0(t) = 1, \quad p_1(t) = 2t - 1.$$

We also have analytical form of the shifted Legendre polynomial of degree i, $p_i(t)$ as follows

$$p_i(t) = \sum_{k=0}^{i} (-1)^{i+k} \frac{(i+k)!t^k}{(i-k)!(k!)^2}, \quad i = 0, 1, 2, \ldots \quad .$$

In Sect. 4.2, we minimize functional (3) on the set of all polynomials that satisfy initial and boundary conditions of the problem (1) – (2). Lemma 4.1 shows that all such polynomials have the same form.

Lemma 4.1 *Let $p(t)$ be a polynomial that satisfies the following conditions*

$$p^{(l)}(0) = y_0^l, \quad p^{(l)}(1) = y_1^l, \quad 0 \le l \le n,$$

then $p(t)$ has the following form

$$p(t) = \sum_{j=0}^{k} c_j t^{n+1}(t-1)^{n+1} p_j(t) + w(t),$$

where, $k \in Z^+$, $c_j \in \mathbb{R}$, and $w(t)$ is the Hermit interpolating polynomial of degree at most $2n+1$ that satisfies above conditions.

Proof Obviously we have

$$p(t) = p(t) - w(t) + w(t),$$

where

$$p^{(l)}(0) - w^{(l)}(0) = 0, \quad p^{(l)}(1) - w^{(l)}(1) = 0, \quad 0 \le l \le n.$$

So we have

$$p(t) - w(t) = t^{n+1}(t-1)^{n+1} s(t),$$

where $s(t) = \sum_{j=0}^{k} c_j p_j(t)$. □

Remark 4.1 Considering above lemma, it is easy to see that polynomial $p(t)$, that satisfies conditions

$$p^{(l)}(0) = y_0^l, \quad 0 \le l \le n,$$

has the form

$$p(t) = \sum_{j=0}^{k} c_j t^{n+1} p_j(t) + w(t),$$

where, $k \in Z^+$, $c_j \in \mathbb{R}$, and $w(t)$ is the Hermit interpolating polynomial that satisfies given conditions.

4.2 Approximation

Consider expansions $y_{j,\epsilon}^k(t)$, $1 \le j \le m$, in the following forms:

$$y_{j,\epsilon}^k(t) = {C_{j,\epsilon}^k}^T . \Psi_k(t), \quad \Psi_k(t) = \begin{pmatrix} p_0(t) \\ p_1(t) \\ \vdots \\ p_k(t) \end{pmatrix}, \quad C_{j,\epsilon}^k = \begin{pmatrix} c_{j,\epsilon}^0 \\ c_{j,\epsilon}^1 \\ \vdots \\ c_{j,\epsilon}^k \end{pmatrix}, \tag{4}$$

for when y_j belongs to the case (i).

$$y_{j,\epsilon}^k(t) = {C_{j,\epsilon}^k}^T.\Psi_k(t) + w_j(t), \quad \Psi_k(t) = \begin{pmatrix} p_0(t)t^{\lceil\alpha_j\rceil} \\ p_1(t)t^{\lceil\alpha_j\rceil} \\ \vdots \\ p_k(t)t^{\lceil\alpha_j\rceil} \end{pmatrix}, \quad C_{j,\epsilon}^k = \begin{pmatrix} c_{j,\epsilon}^0 \\ c_{j,\epsilon}^1 \\ \vdots \\ c_{j,\epsilon}^k \end{pmatrix}, \tag{5}$$

for when y_j belongs to the case (ii). Here, w_j is the Hermit interpolating polynomial that satisfies all initial conditions of y_j.

$$\begin{aligned} &y_{j,\epsilon}^k(t) = {C_{j,\epsilon}^k}^T.\Psi_k(t) + w_j(t), \\ &\Psi_k(t) = \begin{pmatrix} p_0(t)t^{\lceil\alpha_j\rceil}(t-1)^{\lceil\alpha_j\rceil} \\ p_1(t)t^{\lceil\alpha_j\rceil}(t-1)^{\lceil\alpha_j\rceil} \\ \vdots \\ p_k(t)t^{\lceil\alpha_j\rceil}(t-1)^{\lceil\alpha_j\rceil} \end{pmatrix}, \quad C_{j,\epsilon}^k = \begin{pmatrix} c_{j,\epsilon}^0 \\ c_{j,\epsilon}^1 \\ \vdots \\ c_{j,\epsilon}^k \end{pmatrix}, \end{aligned} \tag{6}$$

for when y_j belongs to the case (iii). In this case, w_j is the Hermit interpolating polynomial that satisfies all initial and boundary conditions of y_j.

Substituting $y_{j,\epsilon}^k$, $1 \le j \le m$, in (3), we achieve

$$\begin{aligned} J_\epsilon[C_{1,\epsilon}^k, \dots, C_{m,\epsilon}^k] = \int_0^1 &[F(t, y_{1,\epsilon}^k, \dots, y_{m,\epsilon}^k, \dots, {}_0^C D_t^{\alpha_r} y_{r,\epsilon}^k, \dots) \\ &+ \frac{1}{\epsilon}\sum_{l=1}^L G_l^2(t, y_{1,\epsilon}^k, \dots, y_{m,\epsilon}^k, \dots, {}_0^C D_t^{\alpha_r} y_{r,\epsilon}^k, \dots)]\mathrm{d}t, \end{aligned} \tag{7}$$

which is an algebraic function of unknowns $c_{j,\epsilon}^i$, $i = 0, 1, \dots, k$, $j = 1, \dots, m$. If $c_{j,\epsilon}^i$s be determined by minimizing function J_ϵ, then by (4) – (6) we achieve functions that approximate minimum value of J_ϵ in (7) and also satisfy all initial and boundary conditions of the problem. According to differential calculus, the following system of equations is the necessary condition of optimization for the function (7):

$$\frac{\partial J_\epsilon}{\partial c_{j,\epsilon}^i} = 0, \quad 1 \le j \le m, \quad 0 \le i \le k. \tag{8}$$

By solving the system (8), we can determine the minimizing values of $c_{j,\epsilon}^i$s, $i = 0, 1, \dots, k$, $j = 1, \dots, m$ for function (7). Hence, we achieve functions $y_{j,\epsilon}^k$, $1 \le j \le m$, by (4) – (6), which approximate minimum value of J by

$$J[C_{1,\epsilon}^k, \dots, C_{m,\epsilon}^k] = \int_0^1 F(t, y_{1,\epsilon}^k, \dots, y_{m,\epsilon}^k, \dots, {}_0^C D_t^{\alpha_r} y_{r,\epsilon}^k, \dots)\mathrm{d}t, \tag{9}$$

while

$$\| G_l(t, y^k_{1,\epsilon}, \ldots, y^k_{m,\epsilon}, \ldots, {}^C_0D^{\alpha_r}_t y^k_{r,\epsilon}, \ldots) \|_{L^2[0,1]} \simeq 0, \quad l = 1, \ldots, L,$$

and also satisfy all initial and boundary conditions of the problem.

5 Convergence

Let

$$E[0, 1] = \{f(t) \in C^n[0, 1] : f^{(j)}(0) = f^j_0, f^{(j)}(1) = f^j_1, j = 0, 1, \ldots, n-1\},$$

where f^j_0 and f^j_1 are given constant values. The following lemma plays an important role in our discussion. The lemma shows that polynomial functions of the metric space $E[0, 1]$ are dens.

Lemma 5.1 *Let $f(t) \in E[0, 1]$. There exist a sequence of polynomial functions $\{s_k(t)\}_{k\in N} \subset E[0, 1]$ such that $s_k \to f$ with respect to $\| . \|_n$.*

Proof [23]. □

Consider the normed space $(F_m[0, 1], \| . \|)$ as follows

$$F_m[0, 1] = \prod_{j=1}^{m} E_j[0, 1], \quad \| (y_1, \ldots, y_m) \| = \sum_{j=1}^{m} \| y_j \|_{E_j},$$

where $\| y_j \|_{E_j} = \| y_j \|_\infty$ when y_j belongs to the case (i), $\| y_j \|_{E_j} = \| y_j \|_{\lceil \alpha_j \rceil}$ when y_j belongs to the case (ii) or (iii).

Consider $H^k_m[0, 1]$ as follows:

$$H^k_m[0, 1] = \prod_{j=1}^{m} (E_j[0, 1] \bigcap < \{p_i\}^k_{i=0} >),$$

where $< \{p_i\}^k_{i=0} >$ is the Banach space generated by the Legendre polynomials of degree at most k. Of course $H^k_m[0, 1]$ is a subspace of $F_m[0, 1]$.

Let $y \in C^n[0, 1]$. For Caputo fractional derivative of order α, $n - 1 < \alpha \le n$, we have ${}^C_0D^\alpha_t y(t) \in C[0, 1]$ [31–33]. We also have

$$ {}^C_0D^\alpha_t y(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-s)^{n-\alpha-1} y^{(n)}(s)\mathrm{d}s,$$

$$| {}^C_0D^\alpha_t y(t) | \le \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-s)^{n-\alpha-1} | y^{(n)}(s) | \mathrm{d}s$$

$$\le \frac{\| y^{(n)} \|_\infty}{\Gamma(n-\alpha)} \int_0^t (t-s)^{n-\alpha-1}\mathrm{d}s = \frac{\| y^{(n)} \|_\infty t^{n-\alpha}}{\Gamma(n-\alpha)(n-\alpha)} \le \frac{\| y^{(n)} \|_\infty}{\Gamma(n-\alpha+1)}.$$

So

$$\| {}_0^C D_t^{\alpha} y \|_{\infty} \leq \frac{\| y^{(n)} \|_{\infty}}{\Gamma(n-\alpha+1)}, \quad n-1<\alpha \leq n. \tag{10}$$

Now consider (3) as functional $J_\epsilon : F_m[0,1] \to \mathbb{R}$. Lemma 5.2 shows that J_ϵ is continuous on it's domain. We use this important property later in Theorem 5.2. The following theorem from real analysis plays key role in the proof of Lemma 5.2.

Theorem 5.1 *Let f be continuous mapping of a compact metric space X into a metric space Y, then f is uniformly continuous.*

Proof [34]. □

Lemma 5.2 *The functional J_ϵ is continuous on $(F_m[0,1], \| . \|)$.*

Proof Let $(y_1^*, \dots, y_m^*) \in F_m[0,1]$. $\eta > 0$ is given. Consider $d > 0$ and

$$I = [0,1] \times [-L-d, L+d] \times \cdots \times [-L-d, L+d],$$

where $L = max\{\| y_1^* \|_\infty, \dots, \| y_m^* \|_\infty, \dots, \| {}_0^C D_t^{\alpha_r} y_r^* \|_\infty, \dots\}$.

Obviously $Y^*(t) = (t, y_1^*(t), \dots, y_m^*(t), \dots, {}_0^C D_t^{\alpha_r} y_r^*(t), \dots) \in I$, $t \in [0,1]$. $\gamma > 0$ is given. Let $\delta > 0$ and $\| (y_1, \dots, y_m) - (y_1^*, \dots, y_m^*) \| < \delta$, hence we have $\| y_j - y_j^* \|_{E_j} < \delta$, $1 \leq j \leq m$, and according to (10) it is easy to see that for small enough value of δ we have

$$Y(t) = (t, y_1(t), \dots, y_m(t), \dots, {}_0^C D_t^{\alpha_r} y_r(t), \dots) \in I,$$
$$| Y(t) - Y^*(t) | < \gamma, \quad t \in [0,1].$$

Since functions F and G_l, $l = 1, \dots, L$, are continuous on I and I is a compact set, according to Theorem 5.1, $R = F + \frac{1}{\epsilon} \sum_{l=1}^{L} G_l^2$ is uniformly continuous on I. So if $\gamma > 0$ be sufficiently small, then $| Y(t) - Y^*(t) | < \gamma$ implies that $| R(Y(t)) - R(Y^*(t)) | < \eta$, $t \in [0,1]$, and $| J_\epsilon[y_1, \dots, y_m] - J_\epsilon[y_1^*, \dots, y_m^*] | < \eta$. □

Theorem 5.2 *Let μ_ϵ be the minimum of the functional J_ϵ on $F_m[0,1]$ and also let $\hat{\mu}_{\epsilon,k}$ be the minimum of the functional J_ϵ on $H_m^k[0,1]$, then $\lim_{k\to\infty} \hat{\mu}_{\epsilon,k} = \mu_\epsilon$.*

Proof For any given $\eta > 0$, let $(y_{1,\epsilon}^*, \dots, y_{m,\epsilon}^*) \in F_m[0,1]$ such that

$$J_\epsilon[y_{1,\epsilon}^*, \dots, y_{m,\epsilon}^*] < \mu_\epsilon + \eta.$$

Such $(y_{1,\epsilon}^*, \dots, y_{m,\epsilon}^*)$ exists by the properties of minimum. According to Lemma 5.2, J_ϵ is continuous on $(F_m[0,1], \| . \|)$ so we have

$$| J_\epsilon[y_1, \dots, y_m] - J_\epsilon[y_{1,\epsilon}^*, \dots, y_{m,\epsilon}^*] | < \eta, \tag{11}$$

provided that $\| (y_1, \dots, y_m) - (y_{1,\epsilon}^*, \dots, y_{m,\epsilon}^*) \| < \delta$. According to Weierstrass theorem [34] and Lemma 5.1, for large enough value of k there exist $(\gamma_1^k, \dots, \gamma_m^k) \in H_m^k[0,1]$ such that $\| (\gamma_1^k, \dots, \gamma_m^k) - (y_{1,\epsilon}^*, \dots, y_{m,\epsilon}^*) \| < \delta$. Moreover let $(y^k_{1,\epsilon},\ldots,y^k_{m,\epsilon})$ be the element of $H^k_m[0,1]$ such that $J_\epsilon[y^k_{1,\epsilon},\ldots,y^k_{m,\epsilon}]=\hat{\mu}_{\epsilon,k}$, then using (11) we have

$$\mu_\epsilon \le J_\epsilon[y^k_{1,\epsilon},\ldots,y^k_{m,\epsilon}] \le J_\epsilon[\gamma^k_1,\ldots,\gamma^k_m] < \mu_\epsilon + 2\eta.$$

Since $\eta>0$ is arbitrary, it follows that: $\lim_{k\to\infty}\hat{\mu}_{\epsilon,k}=\mu_\epsilon$. □

Theorem 5.3 *Let $\{\epsilon_j\}_{j\in N}\downarrow 0$ be a sequence of monotonically decreasing positive real numbers. Suppose $\hat{\mu}_{\epsilon_j,k}=J_{\epsilon_j}[y^k_{1,\epsilon_j},\ldots,y^k_{m,\epsilon_j}]$ be the minimum of the functional J_{ϵ_j} on $H^k_m[0,1]$, then for given $\eta>0$ there exist a sequence of natural numbers $\{k_j\}_{j\in N}$ such that*

$$\mu_{\epsilon_j,k_j} := J[y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j}] < \mu+\eta, \quad j\in N,$$

and

$$\lim_{j\to\infty} \| G_l(t, y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j},\ldots,{}^C_0D^{\alpha_r}_t y^{k_j}_{r,\epsilon_j},\ldots) \|_{L^2[0,1]}=0, \quad l=1,\ldots,L.$$

Proof $\eta>0$ is given. Suppose $\mu_{\epsilon_j}=J_{\epsilon_j}[y_{1,\epsilon_j},\ldots,y_{m,\epsilon_j}]$ be the minimum of the functional J_{ϵ_j} on $F_m[0,1]$. It is obvious that

$$\mu_{\epsilon_j} \le J_{\epsilon_j}[y^1_\mu,\ldots,y^m_\mu]=\mu, \quad j\in N. \tag{12}$$

On the other hand according to Theorem 5.2 we have

$$\lim_{k\to\infty}\hat{\mu}_{\epsilon_j,k}=\mu_{\epsilon_j}, \quad j\in N. \tag{13}$$

So considering (12) and (13), $\forall j\in N$ there exist $k_j\in N$ such that

$$J_{\epsilon_j}[y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j}] < \mu+\eta, \tag{14}$$

and we have

$$\mu_{\epsilon_j,k_j} = J[y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j}] \le J_{\epsilon_j}[y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j}] < \mu+\eta.$$

Now according to (14) and the assumption $J[y_1,\ldots,y_m]\ge\lambda$, it is easy to see that

$$0 \le \frac{1}{\epsilon_j}\sum_{l=1}^{L}\int_0^1 G_l^2(t, y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j},\ldots,{}^C_0D^{\alpha_r}_t y^{k_j}_{r,\epsilon_j},\ldots)\mathrm{d}t < \mu+\eta-\lambda, \quad j\in N. \tag{15}$$

Thus, (15) leads to: $\lim_{j\to\infty}\sum_{l=1}^{L}\int_0^1 G_l^2(t, y^{k_j}_{1,\epsilon_j},\ldots,y^{k_j}_{m,\epsilon_j},\ldots,{}^C_0D^{\alpha_r}_t y^{k_j}_{r,\epsilon_j},\ldots)\mathrm{d}t = 0$. □

6 Illustrative Test Problems

In this section we apply the method presented in Sect. 4.2 for solving the following test examples. The well-known symbolic software "Mathematica" has been employed for calculations and creating figures.

Example 6.1 Consider the one dimensional integer order FOCP

$$\min \quad J = \frac{1}{2}\int_0^1 [x^2(t) + u^2(t)]\mathrm{d}t,$$
$$\dot{x}(t) = -x(t) + u(t), \quad x(0) = 1.$$

For above problem there exist optimal solution

$$x(t) = \cosh(\sqrt{2}t) + \beta \sinh(\sqrt{2}t),$$
$$u(t) = (1 + \beta\sqrt{2})\cosh(\sqrt{2}t) + (\sqrt{2} + \beta)\sinh(\sqrt{2}t),$$
$$\beta = -\frac{\cosh(\sqrt{2}) + \sqrt{2}\sinh(\sqrt{2})}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})} \simeq -0.98,$$

and minimum value $J[x(t), u(t)] = 0.192909$ [35]. Let

$$G(t, x(t), u(t), \dot{x}(t)) = \dot{x}(t) + x(t) - u(t),$$

and $\epsilon = 0.00001$ in (3). Consider approximations (4) and (5), respectively, as $u_\epsilon^8(t) = \sum_{j=0}^8 u_\epsilon^j p_j(t)$ and $x_\epsilon^8(t) = \sum_{j=0}^8 x_\epsilon^j p_j(t)t + 1$. Substituting approximations $u_\epsilon^8(t)$ and $x_\epsilon^8(t)$ in functional $J_{0.00001}[x, u]$ to achieve function (7) and solving the system (8) we will have

$$x_\epsilon^0 = -1.00579, x_\epsilon^1 = 0.328916, x_\epsilon^2 = -0.0457689, x_\epsilon^3 = 0.00495371,$$
$$x_\epsilon^4 = -0.000365183, x_\epsilon^5 = 0.0000250667, x_\epsilon^6 = -1.26624 \times 10^{-6},$$
$$x_\epsilon^7 = 5.96762 \times 10^{-8}, x_\epsilon^8 = -1.47141 \times 10^{-11},$$
$$u_\epsilon^0 = -0.166106, u_\epsilon^1 = 0.186771, u_\epsilon^2 = -0.0264288, u_\epsilon^3 = 0.00609059,$$
$$u_\epsilon^4 = -0.000372721, u_\epsilon^5 = 0.0000479288, u_\epsilon^6 = -1.87093 \times 10^{-6},$$
$$u_\epsilon^7 = 1.66111 \times 10^{-7}, u_\epsilon^8 = 1.57738 \times 10^{-8},$$
$$J[x_\epsilon^8, u_\epsilon^8] = 0.192909, \quad \| G(t, x_\epsilon^8, u_\epsilon^8, \dot{x}_\epsilon^8) \|_{L^2[0,1]} = 1.42204 \times 10^{-12}.$$

Figure 1 shows the approximate and exact solutions of the problem.

Example 6.2 Consider the following IFVP

$$\min \quad J[y_1, y_2] = \int_0^1 \left({}_0^C D_t^{\frac{1}{2}} y_1 + 2 {}_0^C D_t^{\frac{1}{2}} y_2 - \frac{8t^{\frac{3}{2}}}{3\sqrt{\pi}} - \frac{15\sqrt{\pi}t^2}{8} \right)^2 \mathrm{d}t,$$

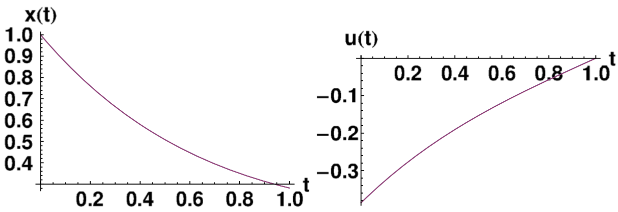

Fig. 1 Exact and approximate values of state and control variables for example 6.1

Table 1 Absolute errors in example 6.2

ϵ	$\mu_{\epsilon,4}$	$E_{\epsilon,4}$
0.1	4.23615×10^{-8}	5.32907×10^{-13}
0.01	4.23615×10^{-8}	5.68434×10^{-14}
0.001	4.23615×10^{-8}	4.9738×10^{-14}

subject to

$$\int_0^1 [5(y_1(t)-1)^2 + 6y_2^2(t)]\mathrm{d}t = 2,$$
$$y_1(0)=1, \quad y_1(1)=2, \quad y_2(0)=0, \quad y_2(1)=1.$$

For the given problem, we have minimizing functions $y_1(t)=t^2+1$ and $y_2(t)=t^{\frac{5}{2}}$ with $J[y_1,y_2]=0$. The problem is solved, substituting approximations (6) with $w_1(t)=1+t$, $w_2(t)=t$ and $k=4$ in (7) and solving the system (8). Table 1 shows $\mu_{\epsilon,k}$ and $E_{\epsilon,k}=\mid \int_0^1 [5(y_{1,\epsilon}^k(t)-1)^2+6{y_{2,\epsilon}^k}^2(t)]\mathrm{d}t - 2 \mid$, for $k=4$ and $\epsilon = 0.1, 0.01, 0.001$.

Example 6.3 Consider the IFVP

$$\min \quad J[y] = \int_0^1 \left({}_0^C D_t^{\frac{1}{2}} y(t) + {}_0^C D_t^{\frac{3}{2}} y(t) - \frac{15\sqrt{\pi} t}{8} - \frac{15\sqrt{\pi} t^2}{16} \right)^2 \mathrm{d}t,$$

subject to

$$\int_0^1 y^2(t)\mathrm{d}t = \frac{1}{6},$$
$$y(0)=0, \quad y'(0)=0, \quad y(1)=1, \quad y'(1)=\frac{5}{2}.$$

For the given problem we have $y(t)=t^{\frac{5}{2}}$ as minimizing function with $J[y]=0$. Applying approximation (6) with $w(t)=t^2+\frac{1}{2}t^2(t-1)$ and solving the system (8), approximate solution for the problem is achieved. Table 2 shows values of $\mu_{\epsilon,k}$ and $E_{\epsilon,k}=\mid \int_0^1 {y_\epsilon^k}^2(t)\mathrm{d}t - \frac{1}{6} \mid$ for different values of ϵ and basis functions k.

Table 2 Absolute errors in example 6.3

k	ϵ	$\mu_{\epsilon,k}$	$E_{\epsilon,k}$
3	0.01	7.35703×10^{-6}	0.0000186063
5	0.001	1.43755×10^{-6}	1.80355×10^{-6}
7	0.0001	3.9976×10^{-7}	9.10605×10^{-8}

Example 6.4 Consider the two dimensional integer order FOCP

$$\begin{aligned}\min \quad J &= \frac{1}{2}\int_0^1 [x_1^2(t) + x_2^2(t) + u^2(t)]\mathrm{d}t,\\ \dot{x_1}(t) &= -x_1(t) + x_2(t) + u(t),\\ \dot{x_2}(t) &= -2x_2(t),\\ x_1(0) &= 1, \quad x_2(0) = 1.\end{aligned}$$

For above problem we have optimal solution

$$\begin{aligned}x_1(t) &= -\frac{3}{2}\mathrm{e}^{-2t} + 2.48164\mathrm{e}^{-\sqrt{2}t} + 0.018352\mathrm{e}^{\sqrt{2}t},\\ x_2(t) &= \mathrm{e}^{-2t},\\ u(t) &= \frac{\mathrm{e}^{-2t}}{2} - 1.02793\mathrm{e}^{-\sqrt{2}t} + 0.0443056\mathrm{e}^{\sqrt{2}t},\end{aligned}$$

and minimum value $J[x_1, x_2, u] = 0.431984$ [17]. Let

$$\begin{aligned}G_1(t, x_1(t), x_2(t), u(t), \dot{x_1}(t), \dot{x_2}(t)) &= \dot{x_1}(t) + x_1(t) - x_2(t) - u(t),\\ G_2(t, x_1(t), x_2(t), u(t), \dot{x_1}(t), \dot{x_2}(t)) &= \dot{x_2}(t) + 2x_2(t),\end{aligned}$$

and $\epsilon = 0.00001$ in (3). Approximations (4) for $u(t)$ and (5) for $x_1(t)$ and $x_2(t)$ with $w_1(t) = w_2(t) = 1$ are considered. Substituting the approximations in functional $J_{0.00001}[x_1, x_2, u]$ as (7) and solving the system (8), we achieve the following approximate values for the problem:

$$\begin{aligned}J[x_{1,\epsilon}^5, x_{2,\epsilon}^5, u_\epsilon^5] &= 0.431987\\ \| G_1(t, x_{1,\epsilon}^5, x_{2,\epsilon}^5, u_\epsilon^5, \dot{x}_{1,\epsilon}^5, \dot{x}_{2,\epsilon}^5) \|_{L^2[0,1]} &= 2.5596 \times 10^{-12},\\ \| G_2(t, x_{1,\epsilon}^5, x_{2,\epsilon}^5, u_\epsilon^5, \dot{x}_{1,\epsilon}^5, \dot{x}_{2,\epsilon}^5) \|_{L^2[0,1]} &= 0.\end{aligned}$$

Figure 2 demonstrates exact and approximate solutions of the problem.

Example 6.5 Consider the two dimensional FOCP

$$\min \quad J = \int_0^1 \left[\left(x_1(t) - 1 - t^{\frac{3}{2}}\right)^2 + \left(x_2(t) - t^{\frac{5}{2}}\right)^2 + \left(u(t) - \frac{3\sqrt{\pi}}{4}t + t^{\frac{5}{2}}\right)^2\right]\mathrm{d}t,$$

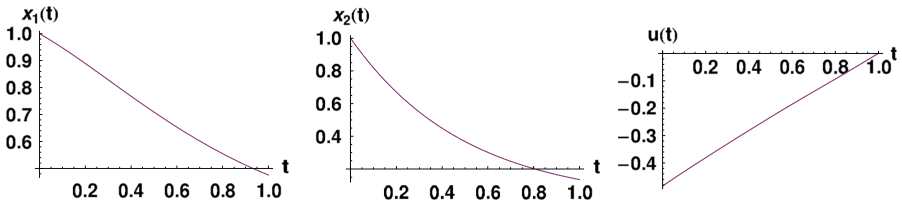

Fig. 2 Exact and approximate values of state and control variables in example 6.4

Table 3 Absolute errors in example 6.5

k	ϵ	$\mu_{\epsilon,k}$	$E^k_{1,\epsilon}$	$E^k_{2,\epsilon}$
3	0.01	0.000101378	0.00150624	0.000596015
5	0.001	0.000053424	0.000166997	0.0000687333
8	0.0001	8.0027×10^{-6}	0.0000423463	0.0000145219

subject to

$$
\begin{aligned}
{}^C_0D^{\frac{1}{2}}_t x_1(t) &= x_2(t) + u(t),\\
{}^C_0D^{\frac{1}{2}}_t x_2(t) &= x_1(t) + \frac{15\sqrt{\pi}}{16}t^2 - t^{\frac{3}{2}} - 1,\\
x_1(0) = 1,& \quad x_2(0) = 0.
\end{aligned}
$$

For above problem optimal solution $x_1(t) = 1 + t^{\frac{3}{2}}$, $x_2(t) = t^{\frac{5}{2}}$, $u(t) = \frac{3\sqrt{\pi}}{4}t - t^{\frac{5}{2}}$ and minimum value $J[x_1, x_2, u] = 0$ are available. Let

$$
\begin{aligned}
&G_1\big(t, x_1(t), x_2(t), u(t), {}^C_0D^{\frac{1}{2}}_t x_1(t), {}^C_0D^{\frac{1}{2}}_t x_2(t)\big) = {}^C_0D^{\frac{1}{2}}_t x_1(t) - x_2(t) - u(t),\\
&G_2\big(t, x_1(t), x_2(t), u(t), {}^C_0D^{\frac{1}{2}}_t x_1(t), {}^C_0D^{\frac{1}{2}}_t x_2(t)\big)\\
&\quad = {}^C_0D^{\frac{1}{2}}_t x_2(t) - x_1(t) - \frac{15\sqrt{\pi}}{16}t^2 + t^{\frac{3}{2}} + 1.
\end{aligned}
$$

We solve the problem with considering approximation (4) for $u(t)$ and (5) for $x_1(t)$ and $x_2(t)$ with $w_1(t) = 1$ and $w_2(t) = 0$, respectively. Table 3 shows values of approximate minimum $\mu_{\epsilon,k}$ and

$$
\begin{aligned}
E^k_{1,\epsilon} &= \| G_1\big(t, x^k_{1,\epsilon}, x^k_{2,\epsilon}, u^k_\epsilon, {}^C_0D^{\frac{1}{2}}_t x^k_{1,\epsilon}, {}^C_0D^{\frac{1}{2}}_t x^k_{2,\epsilon}\big) \|_{L^2[0,1]},\\
E^k_{2,\epsilon} &= \| G_2\big(t, x^k_{1,\epsilon}, x^k_{2,\epsilon}, u^k_\epsilon, {}^C_0D^{\frac{1}{2}}_t x^k_{1,\epsilon}, {}^C_0D^{\frac{1}{2}}_t x^k_{2,\epsilon}\big) \|_{L^2[0,1]},
\end{aligned}
$$

for different values of ϵ and k.

7 Conclusions

An approximate method based upon epsilon and Ritz methods is developed for solving a general class of fractional constrained optimization problems. First, by applying the epsilon method, the constrained optimization problem is reduced to an unconstrained problem, then using Ritz method with special type of polynomial basis functions, the unconstrained optimization problem is reduced to the problem of finding optimal solution of a real value function. The proposed polynomial basis functions have great flexibility in satisfying initial and boundary conditions. The convergence of the method is extensively discussed and illustrative test examples including IFVPs and FOCPs are presented to demonstrate efficiency of the new technique.

References

1. Agrawal, O.M.P.: Formulation of Euler-Lagrange equations for fractional variational problems. J. Math. Anal. Appl. **272**, 368–379 (2002)
2. Agrawal, O.M.P.: A general formulation and solution scheme for fractional optimal control problems. Nonlinear Dyn. **38**, 323–337 (2004)
3. Agrawal, O.M.P., Baleanu, D.: A Hamiltonian formulation and a direct numerical scheme for fractional optimal control problems. J. Vib. Cont. **13**, 1269–1281 (2007)
4. Almeida, R., Torres, D.F.M.: Necessary and sufficient conditions for the fractional calculus of variations with Caputo derivatives. Commun. Nonlinear Sci. Numer. Simul. **16**, 1490–1500 (2011)
5. Agrawal, O.M.P.: A formulation and numerical scheme for fractional optimal control problems. J. Vib. Cont. **14**, 1291–1299 (2008)
6. Almeida, R., Torres, D.F.M.: Calculus of variations with fractional derivatives and fractional integrals. Appl. Math. Lett. **22**, 1816–1820 (2009)
7. Agrawal, O.M.P.: Fractional variational calculus in terms of Riesz fractional derivatives. J. Phys. Math. Theo. **40**, 6287–6303 (2007)
8. Almeida, R., Torres, D.F.M.: Fractional variational calculus for nondifferentiable functions. Comput. Math. Appl. **61**, 3097–3104 (2011)
9. Agrawal, O.M.P.: Generalized variational problems and Euler-Lagrange equations. Comput. Math. Appl. **59**, 1852–1864 (2010)
10. Agrawal, O.M.P.: Fractional variational calculus and transversality conditions. J. Phys. Math. Gen. **39**, 10375–10384 (2006)
11. Agrawal, O.M.P.: Generalized Euler-Lagrange equations and transversality conditions for FVPs in terms of the Caputo Derivative. J. Vib. Cont. **13**, 1217–1237 (2007)
12. Malinowska, A.B., Torres, D.F.M.: Generalized natural boundary conditions for fractional variational problems in terms of the Caputo derivative. Comput. Math. Appl. **59**, 3110–3116 (2010)
13. Yousefi, S.A., Dehghan, M., Lotfi, A.: Generalized Euler-Lagrange equations for fractional variational problems with free boundary conditions. Comput. Math. Appl. **62**, 987–995 (2011)
14. Baleanu, D., Trujillo, J.J.: On exact solutions of class of fractional Euler-Lagrange equations. Nonlinear Dyn. **52**, 331–335 (2008)
15. Baleanu, D., Defterli, O., Agrawal, O.M.P.: A central difference numerical scheme for fractional optimal control problems. J. Vib. Cont. **15**, 583–597 (2009)
16. Tricaud, C., Chen, Y.Q.: An approximate method for numerically solving fractional order optimal control problems of general form. Comput. Math. Appl. **59**, 1644–1655 (2010)
17. Yousefi, S.A., Lotfi, A., Dehghan, M.: The use of a Legendre multiwavelet collocation method for solving the fractional optimal control problems. J. Vib. Cont. **17**, 1–7 (2011)
18. Agrawal, O.M.P.: A general finite element formulation for fractional variational problems. J. Math. Anal. Appl. **337**, 1–12 (2008)
19. Agrawal, O.M.P., Mehedi Hasan, M., Tangpong, X.W.: A numerical scheme for a class of parametric problem of fractional variational calculus. J. Comput. Nonlinear Dyn. **7**, 021005–1 (2012)
20. Wang, D., Xiao, A.: Fractional variational integrators for fractional variational problems. Commun. Nonlinear Sci. Numer. Simul. **17**, 602–610 (2012)
21. Pooseh, S., Almeida, R., Torres, D.F.M.: Discrete direct methods in the fractional calculus of variations. Comput. Math. Appl. **66**, 668–676 (2013)
22. Agrawal, O.M.P.: A quadratic numerical scheme for fractional optimal control problems. J. Dyn. Sys. Measur. Cont. **130**, 011010-1-011010-6 (2008)

23. Lotfi, A., Yousefi, S.A.: A numerical technique for solving a class of fractional variational problems. J. Comput. Appl. Math. **237**, 633–643 (2013)
24. Lotfi, A., Dehghan, M., Yousefi, S.A.: A numerical technique for solving fractional optimal control problems. Comput. Math. Appl. **62**, 1055–1067 (2011)
25. Lotfi, A., Yousefi, S.A., Dehghan, M.: Numerical solution of a class of fractional optimal control problems via the Legendre orthonormal basis combined with the operational matrix and the Gauss quadrature rule. J. Comput. Appl. Math. **250**, 143–160 (2013)
26. Pooseh, S., Almeida, R., Torres, D.F.M.: Approximation of fractional integrals by means of derivatives. Comput. Math. appl. **64**, 3090–3100 (2012)
27. Malinowska, A.B., Torres, D.F.M.: Introduction to the Fractional Calculus of Variations. Imp. Coll. Press, London (2012)
28. Balakrishnan, A.V.: On a new computing technique in optimal control. SIAM J. Cont. **6**, 149–173 (1968)
29. Frick, P.A.: An integral formulation of the ϵ-problem and a new computational approach to control function optimization. J. Optim. Theo. Appl. **13**, 553–581 (1974)
30. Frick, P.A., Stech, D.J.: Epsilon-Ritz method for solving optimal control problems: usefull parallel solution method. J. Optim. Theo. Appl. **79**, 31–58 (1993)
31. Samko, S.G., Kilbas, A.A., Marichev, O.I.: Fractional Integrals and Derivatives. Theory and Applications. Gordon and Breach Sci. Publ, Langhorne (1993)
32. Podlubny, I.: Fractional Differential Equations. Academic Press, San Diego (1999)
33. Kilbas, A.A., Srivastava, H.M., Trujillo, J.J.: Theory and Applications of Fractional Differential Equations. North-Holland Mathematics Studies, Amsterdam (2006)
34. Royden, H.L.: Real Analysis. Macmillan Publishing Company, U.S.A (1988)
35. Datta, K.B., Mohan, B.M.: Orthogonal Functions in Systems and Control. World Scientific, Singapore, New Jersey, London, Hong Kong (1995)