Pesquisa Operacional (2022) 42: e253422 p.1-16
doi: 10.1590/0101-7438.2022.042.00253422
© 2022 Brazilian Operations Research Society
Printed version ISSN 0101-7438 / Online version ISSN 1678-5142
www.scielo.br/pope
ARTICLES

FRACTIONAL ORDER LOG BARRIER INTERIOR POINT ALGORITHM FOR POLYNOMIAL REGRESSION IN THE ℓ_p-NORM

Eliana Contharteze Grigoletto

Received June 18, 2021 / Accepted June 23, 2022

ABSTRACT. Fractional calculus is the branch of mathematics that studies the several possibilities of generalizing the derivative and integral of a function to noninteger order. Recent studies found in literature have confirmed the importance of fractional calculus for minimization problems. However, the study of fractional calculus in interior point methods for solving optimization problems is still new. In this study, inspired in applications of fractional calculus in many fields, was developed the so-called fractional order log barrier interior point algorithm by replacing some integer derivatives for the corresponding fractional ones on the first order optimality conditions of Karush-Kuhn-Tucker to solve polynomial regression models in the ℓ_p-norm for $1 < p < 2$. Finally, numerical experiments are performed to illustrate the proposed algorithm.

Keywords: nonlinear programming, polynomial regression, ℓ_p-norm, interior point method, fractional derivative.

1 INTRODUCTION

It is fundamentally important to make predictions based upon scientific data. The problem of fitting curves to data points has many practical applications Bard (1974); Sevaux & Mineur (2007); Chatterjee & Hadi (2012).

Given a set of m data points in $\mathbb{R}^2$: $\{(a_i,b_i)\}_{i=1}^m$, where a_i is an argument value and b_i a corresponding dependent value, with $a_i \neq a_j$ for all $i \neq j$. The curve fitting procedure tries to build a linear or nonlinear function $y = f(x)$, defined for all possible choices of x, that approximately fits the data set. The fitted curves to the data by f are most often chosen to be polynomials Süli & Mayers (2003).

Let $y = f(x)$ be a polynomial function of degree $n-1$ of the form $f(x) = x_0 + x_1 x + \ldots + x_{n-1} x^{n-1}$, the candidate function to fit data. The procedure to fit the polynomial function to data, in the ℓ_p-norm, determines the vector $x = (x_0, x_1, x_2, \ldots, x_{n-1})^{\mathsf{T}} \in \mathbb{R}^n$, where the superscript T represents transpose, that minimizes the ℓ_p-norm of the residual error as follows

$$\min \quad \Phi(x) = \|b - Ax\|_p, \tag{1}$$

São Paulo State University (Unesp), Avenida Universitária, 3780, Botucatu, SP, Brazil – E-mail: eliana.contharteze@unesp.br – http://orcid.org/0000-0003-4336-5387

where $\|\cdot\|_p$ denotes ℓ_p-norm, $b = (b_1, b_2, \ldots, b_m)^{\mathsf{T}} \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$ is a Vandermonde matrix, with rank $(A) = m$, which can be written as

$$A = \begin{pmatrix} 1 & a_1 & {a_1}^2 & \cdots & {a_1}^{n-1} \\ 1 & a_2 & {a_2}^2 & \cdots & {a_2}^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_m & {a_m}^2 & \cdots & {a_m}^{n-1} \end{pmatrix}. \tag{2}$$

This is a nonlinear (or linear) regression model if $n > 2$ (or $n = 2$). If $m \leq n$, there is an $(n-1)$th degree polynomial satisfying $f(a_i) = b_i$ $(i = 1, 2, \ldots, m)$. If $m > n$, the problem cannot be exactly solved and one needs to find the vector x in (1).

In many industrial applications, missing data or anomalous values can arise as errors of measurement instruments during the data generation. In this cases, polynomial regression models in the ℓ_p-norm, with $p \neq 2$, are more robust Forsythe (1972). It is important to choose an appropriate value for p and several criteria for the choice of p have been studied Rice & White (1964).

The calculus of integrals and derivatives of arbitrary order, named as fractional calculus, was conceptualized in connection with the infinitesimal calculus since 1695 Oldham & Spanier (1974). Some of the application areas include: viscoelasticity, signal processing, probabi- lity, statistics, electrochemistry, diffusion in porous media, fluid flow, backpropagation training of neural networks, fuzzy control method, and so on Kilbas et al. (2006); Dalir & Bashour (2010); Mohammadzadeh & Kayacan (2020); Wang et al. (2017a); Chen et al. (2017a); Grigoletto & Oliveira (2020). Also, its importance in minimization problems has been confirmed by recent works found in literature. For example, Chen et al. Chen et al. (2017b) presented the fractional order gradient methods (FOGMs) by writing the Riemann-Liouville and Caputo fractional derivatives as Taylor series, and Wang et al. Wang et al. (2017b) proposed the fractional gradient descent method employing the Caputo derivative for the backpropagation training of neural networks. However, the study of fractional calculus in the field of interior point methods for solving optimization problems is still new.

The goal of this study is to investigate the fractional order log barrier interior point algorithm, involving the Caputo fractional derivative. It is based on replacing some integer derivatives for the corresponding fractional ones on the first order optimality conditions for solving polynomial regression models in the ℓ_p-norm for $1 < p < 2$. Functions of the form $g(x) = (x+\kappa)^p$, with $\kappa > 0$ and $1 < p < 2$, arise when solving problem 1), according to the approach which will be discussed throughout this study. The Caputo fractional derivative, in addition to generalize the integer order derivative, is a useful tool to obtain the derivative for functions as $g(x)$, as p is on the interval $1 < p < 2$, then it will be a non integer number.

This paper is organized as follows. Preliminary concepts of fractional calculus are presented in Section 2. In Section 3, the fractional order log barrier interior point algorithm for solving polynomial regression models in the ℓ_p-norm is discussed. Numerical experiments are performed to illustrate the proposed algorithm in Section 4, and the Section 5 contains the conclusions.

2 PRELIMINARY CONCEPTS OF FRACTIONAL CALCULUS

Some basic concepts and definitions involving special functions and the Caputo fractional derivatives will be presented in this section.

Definition 1. The gamma function $\Gamma(z)$ is originally defined by Erdélyi et al. (1981); Andrews et al. (1999):

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \mathrm{d}t \qquad (z > 0). \tag{3}$$

For gamma function, the reduction formula

$$\Gamma(z+1) = z\Gamma(z), \tag{4}$$

holds. In particular, when $z = n \in \mathbb{N}$, then $\Gamma(n+1) = n!$.

Definition 2. The beta function $\mathrm{B}(z,w)$ is defined by Erdélyi et al. (1981); Andrews et al. (1999):

$$\mathrm{B}(z,w) = \int_0^\infty \xi^{z-1} (1+\xi)^{-z-w} \, \mathrm{d}\xi \qquad (z > 0; w > 0). \tag{5}$$

It is connected with the gamma function by the following relation

$$\mathrm{B}(z,w) = \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)} \qquad (z > 0; w > 0). \tag{6}$$

The fractional calculus basically defines several integral and derivative operators of arbitrary order. The Caputo fractional derivative is one of the fractional derivative operators Kilbas et al. (2006).

Definition 3. The Caputo left-sided fractional derivative with respect to x, of order $0 < \alpha < 1$, of functions f on a subset $\Omega = [a,b]$ of real axis $\mathbb{R} = (-\infty, \infty)$, where $f \in \mathrm{C}^1[a,b]$, denoted by $\left({}_x^{\mathrm{C}}\mathrm{D}_{a+}^{\alpha} f\right)(x)$, is given by

$$\left({}_x^{\mathrm{C}}\mathrm{D}_{a+}^{\alpha} f\right)(x) := \frac{1}{\Gamma(1-\alpha)} \int_a^x (x-\tau)^{-\alpha} f'(\tau) \, \mathrm{d}\tau \qquad (x \in \Omega). \tag{7}$$

Note that the Caputo fractional derivative is a nonlocal fractional derivative, because it depends on the choice of the order α, function $f(x)$, x, and also relies on the total effects on the interval $[a,x]$. Usually, this is called memory effect.

In particular, $\left({}_x^{\mathrm{C}}\mathrm{D}_{a+}^{1} f\right)(x) := f'(x)$.

Property 1. Let $g(x) = (x+\kappa)^p$ be defined for $-\kappa \leq x < \infty$, with $\kappa \geq 0$, and $1 < p < 2$, and let $0 < \alpha < 1$. Then

$$\left[{}_x^{\mathrm{C}}\mathrm{D}_{-\kappa+}^{\alpha} (x+\kappa)^p\right](x) = \frac{\Gamma(p+1)}{\Gamma(p-\alpha+1)} (x+\kappa)^{p-\alpha}. \tag{8}$$

Proof. Applying the Caputo left-sided fractional derivative (7) to the function $g(x)$ with $a = -\kappa$,

$$\left[{}_x^{\mathrm{C}}\mathrm{D}_{-\kappa+}^{\alpha} (x+\kappa)^p\right](x) = \frac{1}{\Gamma(1-\alpha)} \int_{-\kappa}^x (x-\tau)^{-\alpha} p(\tau+\kappa)^{p-1} \, \mathrm{d}\tau. \tag{9}$$

Using the change of variables $\xi = \dfrac{\tau+\kappa}{x-\tau}$, then:

$$\tau = \frac{\xi x - \kappa}{1+\xi}, \ \mathrm{d}\tau = \frac{(x+\kappa)}{(1+\xi)^2} \, \mathrm{d}\xi, \ \tau \to -\kappa \Rightarrow \xi \to 0 \quad \text{and} \quad \tau \to x \Rightarrow \xi \to \infty.$$

The integral (9) under this change of variables and by means of (4)–(6) takes the form

$$\begin{aligned}\left[{}_{x}^{\mathrm{C}}\mathrm{D}_{-\kappa+}^{\alpha}\,(x+\kappa)^{p}\right](x) &= \frac{p}{\Gamma(1-\alpha)}\int_{-\kappa}^{x}(x-\tau)^{-\alpha}(\tau+\kappa)^{p-1}\,\mathrm{d}\tau\\ &= \frac{p}{\Gamma(1-\alpha)}\int_{0}^{\infty}\left(\frac{x+\kappa}{1+\xi}\right)^{-\alpha}\left(\frac{\xi\,(x+\kappa)}{1+\xi}\right)^{p-1}\frac{(x+\kappa)}{(1+\xi)^{2}}\,\mathrm{d}\xi\\ &= \frac{p\,(x+\kappa)^{p-\alpha}}{\Gamma(1-\alpha)}\int_{0}^{\infty}\xi^{p-1}(1+\xi)^{\alpha-p-1}\,\mathrm{d}\xi\\ &= \frac{p\,(x+\kappa)^{p-\alpha}}{\Gamma(1-\alpha)}\,\mathrm{B}(p,1-\alpha)\\ &= \frac{p\,(x+\kappa)^{p-\alpha}}{\Gamma(1-\alpha)}\,\frac{\Gamma(p)\,\Gamma(1-\alpha)}{\Gamma(p-\alpha+1)}\\ &= \frac{\Gamma(p+1)}{\Gamma(p-\alpha+1)}\,(x+\kappa)^{p-\alpha}.\end{aligned}$$

□

3 FRACTIONAL ORDER LOG BARRIER INTERIOR POINT ALGORITHM

Raising $\Phi(x) = \|b - Ax\|_p$ to the power p, the problem (1) can be rewritten in the alternative form as follows

$$\min \quad \|b-Ax\|_p^p, \tag{10}$$

where $x \in \mathbb{R}^n$. The problem (10) is similar to the following nonlinear optimization problem

$$\begin{aligned}\min \quad & \|r\|_p^p\\ \text{s.t.} \quad & Ax + r = b,\\ & x \in \mathbb{R}^n,\end{aligned} \tag{11}$$

where $r = (r_1, r_2, \ldots, r_m)^{\mathsf{T}} \in \mathbb{R}^m$ is the residual vector of regression and ℓ_p-norm is defined of the form $\|r\|_p = \left(\sum_{i=1}^{m} |r_i|^p\right)^{\frac{1}{p}}$ $(1 \le p < \infty)$ and $\|r\|_\infty = \max |r_i|$.

The parameter values p most commonly used to the problem (11) are $p = 1$, $p = 2$ and $p \to \infty$. Linear programming procedures Charnes et al. (1955); Oliveira et al. (2000); Oliveira & Lyra (2004) can be used for $p = 1$ and $p \to \infty$ cases. For $p = 2$, the direct solution is given by $x = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b$. For other values of p, unconstrained minimization procedures can be used Dennis Jr. & Schnabel (1983); Li (1993); Cantante et al. (2012).

Considering the unrestricted residual term r_i as the difference between two nonne- gative variables u_i and v_i: $r_i = u_i - v_i$, with u_i and v_i being defined by

$$u_i = \begin{cases} r_i, & \text{if } r_i \ge 0,\\ 0, & \text{if } r_i < 0,\end{cases} \qquad \text{and} \qquad v_i = \begin{cases} 0, & \text{if } r_i \ge 0,\\ -r_i, & \text{if } r_i < 0,\end{cases} \tag{12}$$

for all $i = 1, 2, \ldots, m$, then $|r_i| = u_i + v_i$ and the problem (11) can be converted into a convex programming problem Charnes et al. (1955); Cantante et al. (2012):

$$\begin{aligned}\min \quad & \sum_{i=1}^{m}(u_i + v_i)^p\\ \text{s.t.} \quad & Ax + u - v = b,\\ & x \in \mathbb{R}^n,\ u \in \mathbb{R}_+^m,\ v \in \mathbb{R}_+^m,\end{aligned} \tag{13}$$

where $\mathbb{R}_+^m$ denote the set of m-dimensional nonnegative vectors.

Interior point methods are widely used to solve convex optimization problems because of their good performance in practice Biegler (2010); Gondzio (2012); Lilian et al. (2016). By adding a logarithmic barrier function to the objective function in (13), the barrier problem is given by

$$\begin{array}{ll} \min & \sum_{i=1}^{m}\left[(u_i+v_i)^p-\mu\ln u_i-\mu\ln v_i\right] \\ \text{s.t.} & Ax+u-v=b, \\ & x\in\mathbb{R}^n, \end{array} \tag{14}$$

where $\mu>0$ is the barrier parameter. An optimal solution of (13) can be found by solving a series of barrier problems of the form (14) while μ is decreasing and going to zero. The Lagrangian function associated with the problem (14) is

$$\mathrm{L}\left(x,\lambda,u,v\right)=\sum_{i=1}^{m}\left[(u_i+v_i)^p-\mu\ln u_i-\mu\ln v_i\right]+\lambda^{\mathsf{T}}\left(Ax+u-v-b\right), \tag{15}$$

where $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_m)^{\mathsf{T}}\in\mathbb{R}^m$ is a Lagrange multiplier vector.

Let ϕ and φ be the functions given by

$$\phi\left(x,\lambda,u,v\right)=\sum_{i=1}^{m}\left[(u_i+v_i)^p\right], \tag{16}$$

and

$$\varphi\left(x,\lambda,u,v\right)=-\mu\sum_{i=1}^{m}\left(\ln u_i+\ln v_i\right)+\lambda^{\mathsf{T}}\left(Ax+u-v-b\right), \tag{17}$$

then the Lagrangian function can be written as

$$\mathrm{L}\left(x,\lambda,u,v\right)=\phi\left(x,\lambda,u,v\right)+\varphi\left(x,\lambda,u,v\right). \tag{18}$$

The fractional order log barrier interior point algorithm, involving the Caputo fractional derivative, is based on replacing some integer derivatives for the corresponding fractional ones on the first order optimality conditions ($\nabla\mathrm{L}=\nabla\phi+\nabla\varphi=0$), of the following form

$$\nabla^{\alpha}\phi+\nabla\varphi=\begin{bmatrix}\nabla_x\phi+\nabla_x\varphi\\ \nabla_\lambda\phi+\nabla_\lambda\varphi\\ \nabla_u^{\alpha}\phi+\nabla_u\varphi\\ \nabla_v^{\alpha}\phi+\nabla_v\varphi\end{bmatrix}=\begin{bmatrix}A^{\mathsf{T}}\lambda\\ Ax+u-v-b\\ g_\alpha-\mu U^{-1}1_m+\lambda\\ g_\alpha-\mu V^{-1}1_m-\lambda\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}, \tag{19}$$

where ∇_x, ∇_λ, ∇_u and ∇_v represent the gradient with respect to x, λ, u and v, respectively, ∇_u^{α} and ∇_v^{α} represent the fractional order gradient with respect to u and v, respectively, given by Caputo fractional derivative (7).

Furthermore, $1_m=(1,1,\ldots,1)^{\mathsf{T}}\in\mathbb{R}^m$, $U=\mathrm{diag}\left(u\right)$, $V=\mathrm{diag}\left(v\right)$, where $\mathrm{diag}\left(w\right)$ denote the diagonal matrix from a vector w, and $g_\alpha\in\mathbb{R}^m$. The ith component of the vector g_α is $\left(g_\alpha\right)_i=\left[{}_{u_i}^{\mathrm{C}}\mathrm{D}_{-v_i+}^{\alpha}\left(u_i+v_i\right)^p\right]\left(u_i\right)=\left[{}_{v_i}^{\mathrm{C}}\mathrm{D}_{-u_i+}^{\alpha}\left(u_i+v_i\right)^p\right]\left(v_i\right)$. By Property 1, for $i=1,2,\ldots,m$, $\left(g_\alpha\right)_i$ is given by

$$\left(g_\alpha\right)_i=\frac{\Gamma\left(p+1\right)}{\Gamma\left(p+1-\alpha\right)}\left(u_i+v_i\right)^{p-\alpha}\qquad\left(1<p<2;0<\alpha<1\right). \tag{20}$$

The nonlinear system of equations (19) can be rewritten in the alternative form as follows

$$\begin{bmatrix} A^{\mathsf{T}}\lambda \\ Ax+u-v-b \\ U\left(g_\alpha+\lambda\right) \\ V\left(g_\alpha-\lambda\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu_m \\ \mu_m \end{bmatrix}, \tag{21}$$

where $\mu_m = (\mu, \mu, \ldots, \mu)^{\mathsf{T}} \in \mathbb{R}^m$.

For given $\mu > 0$, if $\alpha = 1$ in the system (21), then it recovers the first order optimality conditions, since $\lim\limits_{\alpha\to1} (\nabla^\alpha \phi + \nabla\varphi) = \nabla\phi + \nabla\varphi = \nabla\mathrm{L}$, and has a unique solution $\left(x_\mu^*, \lambda_\mu^*, u_\mu^*, v_\mu^*\right)$. When the classical gradient is replaced by the fractional one ($0 < \alpha < 1$), then if the system (21) has a solution $\left(x_\mu^{\alpha*}, \lambda_\mu^{\alpha*}, u_\mu^{\alpha*}, v_\mu^{\alpha*}\right)$, it can be called fractional solution and $\lim\limits_{\alpha\to1} \left(x_\mu^{\alpha*}, \lambda_\mu^{\alpha*}, u_\mu^{\alpha*}, v_\mu^{\alpha*}\right) = \left(x_\mu^*, \lambda_\mu^*, u_\mu^*, v_\mu^*\right)$.

For given $\mu > 0$ and $0 < \alpha \leq 1$, suppose that the system (21) has a solution. Given an initial point $\left(x^0, \lambda^0, u^0, v^0\right)$ such that $u^0 \in \mathbb{R}^m_{++}$, $v^0 \in \mathbb{R}^m_{++}$, where $\mathbb{R}^m_{++}$ denote the set of m-dimensional positive vectors, and $Ax^0 + u^0 - v^0 = b$, by applying Newton method to the system (21), then the search direction $(\Delta x, \Delta\lambda, \Delta u, \Delta v)$ can be found by solving the following Newton system

$$\begin{bmatrix} 0 & A^{\mathsf{T}} & 0 & 0 \\ A & 0_m & I_m & -I_m \\ 0 & U & D_u & UH_\alpha \\ 0 & -V & VH_\alpha & D_v \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}, \tag{22}$$

where

$$\begin{aligned} r_1 &= -A^{\mathsf{T}}\lambda, \\ r_2 &= -Ax-u+v+b, \\ r_3 &= -U\left(g_\alpha+\lambda\right)+\mu_m, \\ r_4 &= -V\left(g_\alpha-\lambda\right)+\mu_m, \end{aligned} \tag{23}$$

I_m is the identity matrix of order m, $D_u = G_\alpha + \lambda I_m + UH_\alpha$, $D_v = G_\alpha - \lambda I_m + VH_\alpha$, where $G_\alpha = \operatorname{diag}(g_\alpha)$, $H_\alpha = \operatorname{diag}(h_\alpha)$, where $h_\alpha \in \mathbb{R}^m$ and for $i = 1, 2, \ldots, m$, the ith component of the vector h_α is given by

$$(h_\alpha)_i = \frac{\Gamma(p+1)}{\Gamma(p-\alpha)} (u_i+v_i)^{p-\alpha-1}. \tag{24}$$

The $(h_\alpha)_i$ component is obtained by evaluating the derivative

$$(h_\alpha)_i = \frac{\partial}{\partial u_i}\left[(g_\alpha)_i\right] = \frac{(p-\alpha)\Gamma(p+1)}{\Gamma(p+1-\alpha)} (u_i+v_i)^{p-\alpha-1} = \frac{\partial}{\partial v_i}\left[(g_\alpha)_i\right], \tag{25}$$

and taking (4) into account, $\Gamma(p+1-\alpha)$ can be rewritten in the form $(p-\alpha)\Gamma(p-\alpha)$. So, $(h_\alpha)_i$ in equation (24) can be obtained.

For given $\mu > 0$, the new iterate barrier parameter $\hat{\mu}$ is updated in the following form Wright (1996):

$$\hat{\mu} = \frac{\mu}{\beta}, \tag{26}$$

where $\beta > 1$ is added to control your decay and to improve the convergence process.

To keep off the next $\hat{u}$ and $\hat{v}$ in the border region is required to shorten the step size α_{uv} by introducing a parameter σ, with $0 < \sigma < 1$, which is often a value close to 1, as follows Biegler (2010); Vanderbei (2020):

$$\alpha_{uv} = \min\left\{\sigma\left(\min_{(\Delta u)_i<0} \frac{u_i}{|(\Delta u)_i|}\right), \sigma\left(\min_{(\Delta v)_i<0} \frac{v_i}{|(\Delta v)_i|}\right), \sigma\right\}. \tag{27}$$

Given a point (x,λ,u,v), the new iterate $\left(\hat{x},\hat{\lambda},\hat{u},\hat{v}\right)$ is given by

$$\hat{x} = x+\alpha_{uv}\Delta x, \qquad \hat{\lambda} = \lambda+\alpha_{uv}\Delta\lambda, \qquad \hat{u} = u+\alpha_{uv}\Delta u, \qquad \hat{v} = v+\alpha_{uv}\Delta v.$$

3.1 Search direction

The search direction (Δx, $\Delta\lambda$, Δu, Δv) is unique and can be obtained by solving the Newton system (22):

$$A^{\mathsf{T}}\Delta\lambda = r_1, \tag{28}$$

$$A\Delta x+\Delta u-\Delta v = r_2, \tag{29}$$

$$U\Delta\lambda+D_u\Delta u+UH_\alpha\Delta v = r_3, \tag{30}$$

$$-V\Delta\lambda+VH_\alpha\Delta u+D_v\Delta v = r_4. \tag{31}$$

Isolating Δu in (30),

$$\Delta u = D_u^{-1}\left(r_3-U\Delta\lambda-UH_\alpha\Delta v\right). \tag{32}$$

Substituting (32) into (31) and isolating Δv,

$$\Delta v = D_1{}^{-1}\left[r_4+V\Delta\lambda+VH_\alpha D_u^{-1}\left(U\Delta\lambda-r_3\right)\right], \tag{33}$$

where

$$D_1 = D_v-UV\left(H_\alpha\right)^2 D_u^{-1}. \tag{34}$$

Substituting (32) into (29),

$$A\Delta x+D_u^{-1}\left(r_3-U\Delta\lambda\right)-\left(UH_\alpha D_u^{-1}+I_m\right)\Delta v = r_2. \tag{35}$$

Now, substituting (33) into (35),

$$\Delta\lambda = D_2{}^{-1}\left[r_2-A\Delta x-D_u^{-1}r_3+D_1{}^{-1}\left(UH_\alpha D_u^{-1}+I_m\right)\left(r_4-VH_\alpha D_u^{-1}r_3\right)\right], \tag{36}$$

where

$$D_2 = -UD_u^{-1}-VD_1{}^{-1}\left(UH_\alpha D_u^{-1}+I_m\right)^2. \tag{37}$$

Then, Δx can be obtained by replacing $\Delta\lambda$ given by (36) into (28). In this case, Δx can be written as follows

$$\Delta x = \left(A^{\mathsf{T}}D_2{}^{-1}A\right)^{-1}\bar{r}, \tag{38}$$

where

$$\bar{r} = -r_1+A^{\mathsf{T}}D_2{}^{-1}\left[r_2-D_u^{-1}r_3+D_1{}^{-1}\left(UH_\alpha D_u^{-1}+I_m\right)\left(r_4-VH_\alpha D_u^{-1}r_3\right)\right]. \tag{39}$$

3.2 Initialization

The initial point $\left(x^0, \lambda^0, u^0, v^0\right)$ is chosen so that $u^0 \in \mathbb{R}^m_{++}$, $v^0 \in \mathbb{R}^m_{++}$ and $Ax^0 + u^0 - v^0 = b$, and is given by Coleman & Li (1992); Oliveira & Cantante (2004):

$$
\begin{array}{rcll}
x^0 & = & (A^\intercal A)^{-1} A^\intercal b, & \\
r^0 & = & b - Ax^0, & \\
\lambda^0 & = & \dfrac{\delta r^0}{\left\| r^0 \right\|_\infty}, \text{ with } \delta = 0.975, & \\
u_i^0 & = & \begin{cases} r_i^0 + \varepsilon_1, & \text{if } r_i^0 \geq 0, \\ \varepsilon_1, & \text{if } r_i^0 < 0, \end{cases} & (\text{for } i = 1, 2, \ldots, m), \\
v_i^0 & = & \begin{cases} \varepsilon_1, & \text{if } r_i^0 \geq 0, \\ \varepsilon_1 - r_i^0, & \text{if } r_i^0 < 0, \end{cases} & (\text{for } i = 1, 2, \ldots, m),
\end{array}
\tag{40}
$$

where ε_1 is a positive value close to zero.

The x^0 choice is the direct solution to the problem (11) for $p = 2$, and has proved to be a good choice in numerical experiments performed for polynomial regression models in the ℓ_p-norm for values of p reasonably closer to 2 Cantante et al. (2012); Oliveira & Cantante (2004); Cantane (2004); Grigoletto (2011).

3.3 Termination criteria

The fractional order log barrier interior point algorithm terminates if at least one of the two following convergence criteria is satisfied:

$$
N = \frac{\|(\nabla^\alpha \phi + \nabla \varphi)\|}{(1 + \|x\| + \|u\| + \|v\| + \|\lambda\|)(2m)} \leq \varepsilon_2, \tag{41}
$$

where m is the number of rows of the matrix A, and

$$
\left|\hat{N} - N\right| \leq \varepsilon_3, \tag{42}
$$

where ε_2 and ε_3 are positive values close to zero, and $\hat{N}$ is the new iterate, given by

$$
\hat{N} = \frac{\left\|\left(\nabla^\alpha \hat{\phi} + \nabla \hat{\varphi}\right)\right\|}{\left(1 + \|\hat{x}\| + \|\hat{u}\| + \|\hat{v}\| + \left\|\hat{\lambda}\right\|\right)(2m)},
$$

where $\nabla^\alpha \hat{\phi} + \nabla \hat{\varphi}$ is obtained by considering the new iterates $x := \hat{x}$, $\lambda := \hat{\lambda}$, $u := \hat{u}$, $v := \hat{v}$ and $\mu := \hat{\mu}$ in equation (19).

3.4 Algorithm

The fractional order log barrier interior point algorithm works as follows. For given p, α, μ, β, σ, ε_1, ε_2, ε_3, δ, and initial point (x, λ, u, v) from (40), the Newton system (22) is solved, and its solution $(\Delta x, \Delta\lambda, \Delta u, \Delta v)$ is obtained. The step size α_{uv} is determined by (27) and the new iterate $\left(\hat{x}, \hat{\lambda}, \hat{u}, \hat{v}\right)$ is obtained. The new barrier parameter $\hat{\mu}$ is updated by (26). This procedure repeat until the termination criterion is satisfied.

In particular, when $\alpha = 1$, the fractional order log barrier interior point algorithm recover the classical log barrier interior point algorithm.

Algorithm 1 Fractional order log barrier interior point algorithm for polynomial regression models in the ℓ_p-norm.

```
Input: Parameters: $\sigma = 0.9995$, $\varepsilon_1 = 10^{-10}$, $\varepsilon_2 = 10^{-10}$, $\varepsilon_3 = 10^{-8}$, $\delta = 0.975$, and $\beta = 10$.
Choose the parameters: $\alpha$, $p$ and $n$;
Data: A and b;
Initialization: $\mu := \mu^0 = 0.001$. From (40) determine $x := x^0$, $\lambda := \lambda^0$, $u := u^0$ and $v := v^0$, and from (41) determine $N$.
while conditions (41) and (42) are both false do
    Solve the system (22) for $\Delta x$, $\Delta\lambda$, $\Delta u$, $\Delta v$ by means of equations described in Subsection 3.1;
    Determine the step size $\alpha_{uv}$ from (27);
    $x := x + \alpha_{uv}\Delta x$;
    $\lambda := \lambda + \alpha_{uv}\Delta\lambda$;
    $u := u + \alpha_{uv}\Delta u$;
    $v := v + \alpha_{uv}\Delta v$;
    $\mu := \frac{\mu}{\beta}$.
end while
Output: $x$.
```

4 NUMERICAL EXPERIMENTS

In order to illustrate the proposed fractional order log barrier interior point algorithm to solve polynomial regression models in the ℓ_p-norm, numerical experiments were performed to compare it with the classical log barrier interior point algorithm. The implementation of the fractional order log barrier interior point algorithm was performed under Windows 10 and Matlab (R2016a) running on a desktop with 2.20 GHz Intel Core i5-5200 central processing unit (CPU) and 4G random-access memory (RAM).

A data set containing daily interest rates observed over 40 years was used for the analysis of polynomial regressions. The data set contains 10958 observations: $\{(a_i, b_i)\}_{i=1}^{10958}$, where a_i represents the day of the week for a given date and b_i the interest rate (in percentage) for the specific day a_i. The a_i values were normalized to the $[0,1]$ interval of the real line $\mathbb{R}$ to avoid numerical stability problems of the algorithm Oliveira & Cantante (2004); Cantane (2004). Figure 1 shows the data set.

For the numerical results below, a comparison of the fractional order log barrier interior point algorithm for different values of the order α considers: "It.", the number of iterations; "Res. Err.", the residual error, given by $\|r\|_p = \|b - Ax\|_p$; and "Time (s)", the CPU time in seconds.

The number of iterations and the value for the residual error for different values of α and p, obtained from the fractional order log barrier interior point algorithm for linear regression ($n = 2$) in the ℓ_p-norm, are shown in Tables 1–3. If the algorithm does not converge (divergence of Newton's method for solving the nonlinear system of equations (21): $\|b - Ax^{k+1}\|_p \geq \|b - Ax^k\|_p$), or if the algorithm fails (ill conditioned matrix), then the results will be shown in Tables 1–3 as "-" or "$*$", respectively. When $\alpha = 1$, the results of the classical log barrier interior point algorithm are recovered.

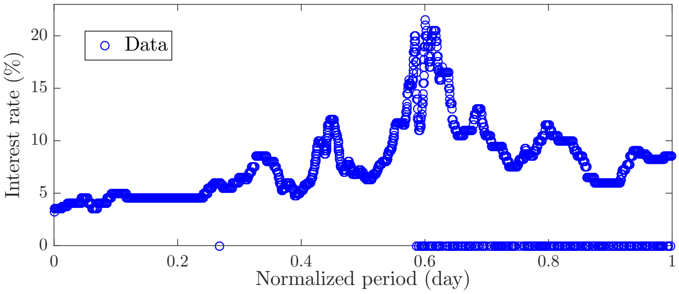

Figure 1 – Daily changes in the interest rate.

Table 1 – Numerical results for linear regression in the ℓ_p−norm ($p = 1.1, 1.2, 1.3$) for different values of α.

	$p = 1.1$		$p = 1.2$		$p = 1.3$	
α	It.	Res. Err.	It.	Res. Err.	It.	Res. Err.
0.1	–	–	–	–	–	–
0.2	5	10250	–	–	–	–
0.3	19	10118	5	5293	–	–
0.4	11	10297	20	5251	5	3044
0.5	$*$	$*$	14	5313	20	3024
0.6	3	10357	17	5324	11	3055
0.7	3	10358	3	5339	6	3064
0.8	5	10358	3	5339	3	3065
0.9	3	10359	5	5339	3	3065
1	2	10360	3	5340	5	3066

The algorithm failed for $\alpha = 0.5$ and $p = 1.1$, as can be seen in Table 1, but for example, for $\alpha = 0.55$ and $p = 1.1$, the algorithm converges.

According to the Tables 1–3, the fractional order log barrier interior point algorithm for linear regression models in the ℓ_p−norm ($p = 1.1, 1.2, \ldots, 1.7$) yield smaller residual error when $\alpha = p - 0.8$. It is a relationship between p and α, which can also be written as $p - \alpha = 0.8$. In this cases, the algorithm takes more iterations until convergence.

For linear regression in the $\ell_{1.2}$−norm, a smaller residual error, given by 5242, was obtained with 16 iterations when $\alpha = 0.39$ ($p - \alpha = 0.81$). For $\alpha = 0.36$, the residual error is 5256 and the number of iterations is 22. For $\alpha = 0.37$, the residual error is 5252 and the number of iterations is 28. For $\alpha = 0.38$, the residual error is 5247 and the number of iterations is 14.

For linear regression in the $\ell_{1.8}$−norm, a smaller residual error, given by 509.967, was obtained with 26 iterations when $\alpha = 0.96$ ($p - \alpha = 0.84$). For $\alpha = 0.94$, the residual error is 510.011 and the number of iterations is 28. For $\alpha = 0.95$, the residual error is 509.987 and the number of iterations is 16. For $\alpha = 0.97$, the residual error is 510.015 and the number of iterations is 17.

Table 2 – Numerical results for linear regression in the ℓ_p−norm ($p = 1.4$, 1.5, 1.6) for different values of α.

	$p = 1.4$		$p = 1.5$		$p = 1.6$	
α	It.	Res. Err.	It.	Res. Err.	It.	Res. Err.
$0.1, 0.2, 0.3, 0.4$	–	–	–	–	–	–
0.5	5	1905	–	–	–	–
0.6	18	1899	5	1274	–	–
0.7	11	1910	18	1272	5	900.8
0.8	123	1909	9	1278	19	900.3
0.9	3	1915	3	1280	11	902
1	3	1915	3	1280	3	903.5

Table 3 – Numerical results for linear regression in the ℓ_p−norm ($p = 1.7$, 1.8, 1.9) for different values of α.

	$p = 1.7$		$p = 1.8$		$p = 1.9$	
α	It.	Res. Err.	It.	Res. Err.	It.	Res. Err.
$0.1, 0.2, \ldots, 0.7$	–	–	–	–	–	–
0.8	5	665.5	–	–	–	–
0.9	18	665.3	6	510.2	–	–
1	11	666.2	19	510.1	26	403.4

For linear regression in the $\ell_{1.9}$−norm, the algorithm does not converge for $\alpha \leq 0.92$. For $\alpha = 0.93$, the residual error is 403.48 and the number of iterations is 4. For $\alpha = 0.95$, the residual error is 403.44 and the number of iterations is 5. For $\alpha = 0.97$, the residual error is 403.42 and the number of iterations is 6. For $\alpha = 0.99$, the residual error is 403.41 and the number of iterations is 6. When $\alpha = 0.99$, the algorithm takes fewer iterations than when $\alpha = 1$ (See Table 3), and the residual error is very close to the residual error obtained when $\alpha = 1$.

The number of iterations, the residual error and the CPU time for $(n-1)$th degree polynomial regressions ($n = 2$, 5, 10, 15), in the $\ell_{1.3}$−norm, for different values of α, obtained from the fractional order log barrier interior point algorithm, are given in the following tables.

The smallest residual errors for $n = 2$ and $n = 5$ were obtained with a greater number of iterations when $\alpha = 0.5$ (see Table 4), while for $n = 10$ and $n = 15$ (see Table 5), the residual errors were smallest for $\alpha = 0.4$.

The performance of the fractional order log barrier interior point algorithm appears to be computationally consistent. For example, the values of the residual error at each iteration k for linear regression in the $\ell_{1.3}$−norm are shown in Figure 2.

In the Figure 2, one can observe that the use of any one of the fractional order derivatives ($\alpha = 0,4, 0.5, \ldots, 0.9$) produces smaller residual errors than with the use of the integer order derivative ($\alpha = 1$). The use of the fractional derivatives does not interfere in the iterations running times, that compare similarly.

Table 4 – Numerical results for polynomial regressions ($n = 2, 5$) in the $\ell_{1.3}-$norm.

	$n = 2$			$n = 5$		
α	It.	Time (s)	Res. Err.	It.	Time (s)	Res. Err.
0.1, 0.2, 0.3	–	–	–	–	–	–
0.4	5	0.0690	3044.3	16	0.1839	2433.6
0.5	20	0.1561	3024.8	20	0.2306	2423.5
0.6	11	0.1033	3055	17	0.2090	2431.8
0.7	6	0.0727	3064.1	5	0.0800	2441.4
0.8	3	0.0379	3065.7	5	0.0849	2443.3
0.9	3	0.0492	3065.8	3	0.0679	2443.7
1	5	0.0725	3065.9	3	0.0718	2443.8

Table 5 – Numerical results for polynomial regressions ($n = 10, 15$) in the $\ell_{1.3}-$norm.

	$n = 10$			$n = 15$		
α	It.	Time (s)	Res. Err.	It.	Time (s)	Res. Err.
0.1, 0.2, 0.3	–	–	–	–	–	–
0.4	2	0.0622	2213.4169	1	0.0555	1953.1475
0.5	1	0.0317	2220.5557	1	0.0583	1953.4556
0.6	1	0.0337	2221.8459	1	0.055	1953.9432
0.7	1	0.0320	2222.0088	1	0.0698	1954.1043
0.8	1	0.0312	2222.2097	1	0.0526	1953.8845
0.9	1	0.0318	2222.2880	1	0.0584	1953.8840
1	1	0.0314	2222.2941	1	0.0697	1953.8973

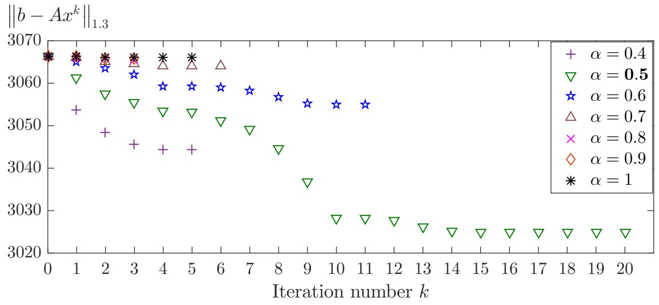

Figure 2 – Linear regression in the $\ell_{1.3}-$norm.

The approximate solution for $(n-1)$th degree polynomial regression in the ℓ_p-norm, obtained at the kth iteration of the fractional order log barrier interior point algorithm and given by $x^k = \left(x_0^k, x_1^k, x_2^k, \ldots, x_{n-1}^k\right)$, provides the coefficients of the $(n-1)$th degree polynomial that approximately fits the data set, given by $f(x) = x_0 + x_1 x + \ldots + x_{n-1} x^{n-1}$. In particular, the approximate solutions x^k, obtained at the kth iteration of the fractional order log barrier interior point algorithm, for $(n-1)$th degree polynomial regressions ($n = 2, 5, 10, 15$) in the $\ell_{1.3}-$norm, with smaller residual errors are:

$$x^{20} = \begin{bmatrix} x_0^{20} \\ x_1^{20} \end{bmatrix} = \begin{bmatrix} 4.5E+00 \\ 5.9E+00 \end{bmatrix} \quad (n=2; \alpha = 0.5),$$

i.e., at the 20th iteration of the algorithm with $\alpha = 0.5$, the linear regression ($n = 2$) provides the polynomial coefficients from the above vector x^{20}. So, the linear regression is given by $f(x) = 4.5 + 5.9x$.

$$x^{20} = \begin{bmatrix} x_0^{20} \\ x_1^{20} \\ x_2^{20} \\ x_3^{20} \\ x_4^{20} \end{bmatrix} = \begin{bmatrix} 5.3E+00 \\ -3.4E+01 \\ 1.9E+02 \\ -2.8E+02 \\ 1.2E+02 \end{bmatrix} \quad (n=5; \alpha = 0.5),$$

$$x^{2} = \begin{bmatrix} x_0^{2} \\ x_1^{2} \\ x_2^{2} \\ x_3^{2} \\ x_4^{2} \\ x_5^{2} \\ x_6^{2} \\ x_7^{2} \\ x_8^{2} \\ x_9^{2} \end{bmatrix} = \begin{bmatrix} 1.8E+00 \\ 1.7E+02 \\ -3.7E+03 \\ 3.5E+04 \\ -1.6E+05 \\ 4.6E+05 \\ -7.6E+05 \\ 7.2E+00 \\ -3.7E+05 \\ 7.9E+04 \end{bmatrix} \quad (n=10; \alpha = 0.4),$$

$$x^{1} = \begin{bmatrix} x_0^{1} \\ x_1^{1} \\ x_2^{1} \\ x_3^{1} \\ x_4^{1} \\ x_5^{1} \\ x_6^{1} \\ x_7^{1} \\ x_8^{1} \\ x_9^{1} \\ x_{10}^{1} \\ x_{11}^{1} \\ x_{12}^{1} \\ x_{13}^{1} \\ x_{14}^{1} \end{bmatrix} = \begin{bmatrix} 3.6E+00 \\ 2.6E+00 \\ -1.9E+02 \\ 1.3E+04 \\ -2.1E+05 \\ 1.4E+06 \\ -5.6E+06 \\ 1.1E+07 \\ -1.3E+07 \\ 5.9E+06 \\ -2.6E+06 \\ 1.2E+07 \\ -1.9E+07 \\ 1.2E+07 \\ -2.8E+06 \end{bmatrix} \quad (n=15; \alpha = 0.4).$$

The polynomials from the solutions described are shown in Figure 3.

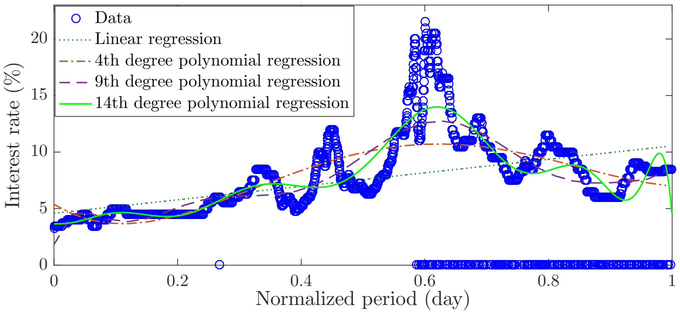

Figure 3 – Polynomial regressions in the $\ell_{1.3}-$norm.

5 CONCLUSIONS

The fractional order log barrier interior point algorithm for polynomial regression models in the ℓ_p-norm $(1 < p < 2)$, obtained by replacing some integer derivatives for the corresponding fractional ones on the first order optimality conditions, was investigated in this paper. This algorithm appears to be computationally consistent but depending on the fractional order α, the algorithm does not converge or fails. The numerical experiments showed that the use of the fractional derivatives can be beneficial for solving optimization problems. In future, further studies on this subject could be undertaken.

Acknowledgments

The author thanks the reviewers for their numerous helpful suggestions.

References

ANDREWS GE, ASKEY R & ROY R. 1999. *Special Functions*. Cambridge: Cambridge University Press.

BARD Y. 1974. *Nonlinear Parameter Estimation*. New York: Academic Press.

BIEGLER LT. 2010. *Nonlinear Programming*. Philadelphia: SIAM.

CANTANE DR. 2004. *Métodos de Pontos Interiores Aplicados ao Problema de Regressão pela Norma Lp*. São Carlos: Dissertação de Mestrado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo.

CANTANTE DR, GRIGOLETTO EC & OLIVEIRA ARL. 2012. Método de pontos interiores barreira logarítmica preditor-corretor especializado para o problema de regressão pela norma Lp. *Tendências em Matemática Aplicada e Computacional – TEMA*, **13**: 219–231.

CHARNES A, COOPER WW & FERGUSON R. 1955. Optimal estimation of executive compensation by linear programming. *Management Sci.*, **1**: 138–151.

CHATTERJEE S & HADI AS. 2012. *Regression Analysis by Example*. Hoboken: Wiley.

CHEN Y, GAO Q, WEI Y & WANG Y. 2017a. Study on fractional order gradient methods. *Appl. Math. Comput.*, **314**: 310–321.

CHEN Y, GAO Q, WEI Y & WANG Y. 2017b. Study on fractional order gradient methods. *Appl. Math. Comput.*, **314**: 310–321.

COLEMAN TF & LI Y. 1992. A globally and quadratically convergent affine scaling method for linear ℓ_1 problems. *Math. Programming*, **56**: 189–222.

DALIR M & BASHOUR M. 2010. Applications of fractional calculus. *Appl. Math. Sci.*, **4**: 1021–1032.

DENNIS JR JE & SCHNABEL RB. 1983. *Numerical Methods for Unconstrained Optimization*. NJ: Prentice-Hall.

ERDÉLYI A, MAGNUS W, OBERHETTINGER F & TRICOMI F. 1981. *Higher Transcendental Functions*. vol. I-III. Melbourne, Florida: Krieger Pub.

FORSYTHE AB. 1972. Robust estimation of straight line regression coefficients by minimizing pth power deviations. *Technometrics*, **14**: 159–166.

GONDZIO J. 2012. Interior point methods 25 years later. *Eur. J. Oper. Res.*, **218**: 587–601.

GRIGOLETTO EC. 2011. *Implementação Eficiente dos Métodos de Pontos Interiores Especializados para o Problema de Regressão pela Norma Lp*. Campinas: Dissertação de Mestrado, Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas.

GRIGOLETTO EC & OLIVEIRA ARL. 2020. Fractional order gradient descent algorithm. In: *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics - XXXIX CNMAC*. p. 010387. São Carlos: SBMAC.

KILBAS AA, SRIVASTAVA HM & TRUJILLO JJ. 2006. *Theory and Applications of Fractional Differential Equations*. Amsterdam: Elsevier.

LI Y. 1993. A globally convergent method for ℓ_p problems. *SIAM J. Optimiz.*, **3**: 609–629.

LILIAN FB, OLIVEIRA ARL & GHIDINI CTLS. 2016. Use of continued iteration on the reduction of iterations of the interior point method. *Pesqui. Oper.*, **36**: 487–501.

MOHAMMADZADEH A & KAYACAN E. 2020. A novel fractional-order type-2 fuzzy control method for online frequency regulation in ac microgrid. *Eng. Appl. Artif. Intell.*, **90**: 103483.

OLDHAM KB & SPANIER J. 1974. *The Fractional Calculus*. New York: Academic Press.

OLIVEIRA ARL & CANTANTE DR. 2004. Método de pontos interiores aplicados ao problema de regressão pela norma Lp. *Tendências em Matemática Aplicada e Computacional – TEMA*, **5**: 281–291.

OLIVEIRA ARL & LYRA C. 2004. Interior point methods for the polynomial ℓ_∞ fitting problems. *Int. Trans. Oper. Res.*, **11**: 309322.

OLIVEIRA ARL, NASCIMENTO MA & LYRA C. 2000. Efficient implementation and benchmark of interior point methods for the polynomial ℓ_1 fitting problem. *Comput. Statist. Data Anal.*, **35**: 119–135.

RICE JR & WHITE JS. 1964. Norms for Smoothing and Estimation. *SIAM Review*, **6**: 243–256.

SEVAUX M & MINEUR Y. 2007. A curve-fitting genetic algorithm for a styling application. *Eur. J. Oper. Res.*, **179**: 895–905.

SÜLI E & MAYERS DF. 2003. *An Introduction to Numerical Analysis*. New York: Cambridge University Press.

VANDERBEI RJ. 2020. *Linear Programming: Foundations and Extensions*. Springer Nature.

WANG J, WEN Y, GOU Y, YE Z & CHEN H. 2017a. Fractional-order gradient descent learning of BP neural networks with Caputo derivative. *Neural Netw.*, **89**: 19–30.

WANG J, WEN Y, GOU Y, YE Z & CHEN H. 2017b. Fractional-order gradient descent learning of BP neural networks with Caputo derivative. *Neural Netw.*, **89**: 19–30.

WRIGHT SJ. 1996. *Primal–Dual Interior–Point Methods*. Philadelphia, USA: SIAM.

How to cite
GRIGOLETTO EC. 2022. Fractional order log barrier interior point algorithm for polynomial regression in the ℓ_p-norm. *Pesquisa Operacional*, **42**: e253422. doi: 10.1590/0101-7438.2022.042.00253422.