10.18605/2175-7275/cereus.v13n4p235-247

<< Recebido em: 16/11/2021 Aceito em: 30/11/2021. >>

ARTIGO ORIGINAL

REVISTA CEREUS
ISSN: 2175-7275

Optimization of Mathematical Functions Using Fractional Steepest Descent Method with Self-Adaptive Order

Otimização de Funções Matemáticas Usando o Método da Máxima Descida Fracionário com Ordem Auto Adaptativa

Fran Sérgio Lobato[1], Gustavo Barbosa Libotte[2], Gustavo Mendes Platt[3]

[1] Universidade Federal de Uberlândia - Uberlândia, MG, Brasil

E-mail: fslobato@ufu.br

[2] Laboratório Nacional de Computação Científica - Petrópolis, RJ, Brasil

E-mail: glibotte@lncc.br

[3] Universidade Federal do Rio Grande - Santo Antônio da Patrulha, RS, Brasil

E-mail: gmplatt@furg.br

ABSTRACT

Fractional calculus represents a mathematical framework that has numerous applications. In the context of optimization, it can be used to increase the performance of gradient-based methods. However, when the direction of the first order integer gradient is generalized using a fractional order, this approach may converge to a different solution from the one obtained by the classical Steepest Descent Method, making the application of this type of methodology complex. To solve this convergence issue, this contribution aims to propose an adaptive approach in which the fractional order is updated along the iterations so that at the end of the optimization process the fractional order is equal to one. For this purpose, the adaptive fractional order is defined from a new parameter, namely, the reduction rate. The results obtained with the optimization of two mathematical functions demonstrate the potential of the proposed methodology.

Keywords: Steepest Descent Method, Fractional Calculus, Fractional Order Update, Optimization.

RESUMO

O cálculo fracionário configura um conjunto de ferramentas matemáticas que apresentam inúmeras aplicações. No contexto da otimização, este pode ser empregado para aumentar o desempenho de métodos baseados em gradiente. Todavia, quando a direção do gradiente de primeira ordem inteira é generalizada por uma com ordem fracionária, essa abordagem pode convergir para uma solução diferente daquela obtida pelo tradicional Método da Máxima Descida, dificultando a aplicação deste tipo de metodologia. Para resolver esse problema de convergência, a presente contribuição tem por objetivo propor uma abordagem adaptativa em que a ordem fracionária é atualizada ao longo das iterações de forma que no final do processo de otimização a ordem fracionária seja igual a unidade. Para essa finalidade, a ordem fracionária adaptativa é definida a partir de um novo parâmetro, a saber, a taxa de redução. Os resultados obtidos com a otimização de duas funções matemáticas demonstram o potencial da metodologia proposta.

Palavras-chave: Método da Máxima Descida, Cálculo Fracionário, Atualização da Ordem Fracionária, Otimização.

1. INTRODUCTION

The classical optimization methods are based on the use of information about the search direction associated with the determination of step size to update the candidate vector. In this context, there are two possibilities to increase the performance of these methods in terms of convergence and computational cost. The first is to increase the information about the search direction by using higher-order derivatives, or more information about earlier steps. The second consists in the development of new strategies to dynamically update the step size in each iteration (VANDERPLAATS, 1984; EDGAR et al., 2001).

In terms of search direction, the main methods are based on the use of gradient vector, and the Hessian matrix can be computed. In this scenario, the Steepest Descent Method (SDM), which makes use of information about the gradient, is known as one of most used approaches for solving optimization problems. This is due to its intuitive principle, simple structure, and ease of implementation (VANDERPLAATS, 1984; EDGAR et al., 2001). These characteristics allow applications in different fields of science, among which we can mention: image noise elimination (PU et al., 2010), machine learning (LECUN et al., 2015), control (REN et al., 2019), and system identification (GE et al., 2019).

As mentioned by Wei et al. (2020), SDM can present a *zigzag* behavior, as well as a slow convergence in the neighborhoods of the optimal solution, which can affect its performance, especially in more complex case studies. To overcome these disadvantages, the development of approaches that associate fractional calculus to SDM have been increasingly explored. The interest in this type of strategy is basically due to the change in the dynamics of fractional differential models with the variation of the order inherent to these models (LIMA et al., 2021). So, as the dynamics of a model is influenced by its fractional order, why not use this principle to increase the convergence speed during an optimization method?. Based on this concept, algorithms based on SDM and with fractional order can be proposed (WEI et al., 2015; PU et al., 2015; LIU et al., 2020). However, Wei et al. (2015) highlight that first-order optimization methods based on the fractional context may not converge to the optimal solution, that is, varying the initial condition and the fractional order, the problem may not converge to the best solution reported in the literature. According to Wei et al. (2020) this is due to non-locality observed in fractional differential systems. Consequently, in the optimization context, these approaches depend on design variable vector, the initial condition and, obviously, the fractional order. Thus, it is not easy to guarantee that the fractional derivative in relation to the design variable vector is equal to zero when compared with the original problem with integer order (WEI et al., 2015).

In this contribution, the aim is to increase the chance of convergence during the application of Fractional Steepest Descent Method (FSDM), regardless of the fractional order considered. In this case, the Fractional Steepest Descent Method with Self-Adaptive Order (FSDM-θ) is proposed. In this new strategy, the fractional order in FSDM is updated during the iterative process by defining a parameter called reduction rate. This work is structured as follows. A brief description of classical SDM (with integer order) is presented in Section 2. The proposed methodology for updating the fractional order is described in Section 3. The results obtained with the application of the proposed methodology in two purely mathematical case studies are presented and discussed in Section 4. Finally, the last section presents the conclusion and the proposal for future work.

2. STEEPEST DESCENT METHOD WITH INTEGER ORDER

To minimize an unconstrained multidimensional function f, basically two main lines of research, namely, the classical (deterministic) and the heuristic (non-deterministic) can be found. The first is based on Variational Calculus to update a candidate for solving the optimization problem. The second is, generally, based on a population of candidates that is updated based on information about the interaction between individuals that are part of that population (EDGAR et al., 2001).

In general, methods based on information of derivatives consists of an iterative search characterized by the determination of a search direction through the execution of a one-dimensional search, and, consequently, obtaining the step size in this direction (VANDERPLAATS, 1984). In the case of the SDM with integer order, the recurrence relationship that characterizes this methodology is based on the following equation:

$$x^q = x^{q-1} - \eta\nabla\left(f(x^{q-1})\right) \quad (1)$$

where x^q and x^{q-1} represent the design variables vector (or decision) at iterations q and $q-1$, respectively, and η is a scalar that gives the step size along the search direction defined by the gradient of the function f. In summary, the search towards the optimal value is composed of two steps. The first consists in expressing the design variables vector from the definition of the search direction, that is, using information about the gradient vector $\nabla f(x)$ at the point corresponding to $q-1$. In the second step, by replacing the information contained in Eq. (1) in f, the search by the optimal solution becomes a one-dimensional search problem, where in each iteration an optimal scalar η (η*) is determined. This process continues until a certain stopping criterion is satisfied.

3. METHODOLOGY

As mentioned earlier, the aim of this contribution is to propose an optimization strategy in fractional context, where the fractional order is dynamically updated along the iterative process. Next, the Fractional Steepest Descent Method (FSDM) and the Fractional Steepest Descent Method with Self-Adaptive Order (FSDM-θ) are proposed. To represent the fractional contribution, the Caputo derivative $({}_C\mathrm{d}_x^{\alpha}F(x))$ of a generic function $F(x)$ with relation to the independent variable x, is defined as (CAPUTO, 1999):

$$ {}_C\mathrm{d}_x^{\alpha}F(x) = \frac{1}{\Gamma(y-\alpha)}\int_0^x (x-t)^{y-\alpha-1}\frac{\mathrm{d}^y F(t)}{\mathrm{d}t^y}\mathrm{d}t \tag{2} $$

where Γ is the *Gamma* function and$y$ is an integer, defined as $y = [\alpha] + 1$, where $[\alpha]$ is the integer part of the fractional order α.

It is important to mention that the choice of this type of fractional derivative is due to the memory effect by means of a convolution between the integer order derivative and a power of time. In this case, the initial conditions for the fractional differential equations can be handled by using an analogy with the classical case (ordinary derivative) (GÓMEZ-AGUILAR et al., 2012).

3.1 Fractional Steepest Descent Method (FSDM)

In general, the FSDM consist of replace the integer order in gradient used in the SDM by an equivalent with fractional order (α), i.e.:

$$ x^q = x^{q-1} - \eta\nabla^{\alpha} f(x^{q-1}) \tag{3} $$

In this case, as described for the integer order SDM, if the vector x^q is replaced in the original objective function, it becomes dependent on α, x^{q-1}, and η. Thus, the optimal value for the step size (η*) can be computed from the simple application of the optimality condition or from one of various techniques for the one-dimensional search (VANDERPLAATS, 1984; EDGAR et al., 2001). With this optimal value, the vector x^q can be updated according to Eq. (3) until a certain stopping criterion is not met. The iterative procedure regarding the FSDM is described in Algorithm 1.

Algorithm 1: Fractional Steepest Descent Method

Input: Information about the problem (number of design variables, objective function) and method parameters (fraction order, initial estimative, strategy to update the size step, stopping criterion, tolerance)
1 . Start of the optimization process:
2. Initialize the counter (*q*)
3. **While** the stopping criterion is not satisfied do
4. Find the value of the (η^{*})
5. Update the value of the design variables by using Eq. (3)
6. *q* = *q* + *1*
7. **End While**
8. End
Output: Optimal solution and iterations number

3.2 Fractional Steepest Descent Method with Self-Adaptive Order (FSDM-θ)

As mentioned earlier, employing a fractional derivative in the optimization context implies the possibility of convergence to a solution that is not optimal, due to the non-locality of fractional models (WEI et al., 2015, 2020). To increase the chance of the FSDM to converge for an optimal solution, the FSDM-θ is proposed. In this new optimization strategy, the fractional order (α) value is dynamically updated throughout during the optimization process. In this case, at the end of optimization process, α tends to one, i.e., the fractional derivative becomes one with integer order. It is important to highlight that this is imposed to prevent the process from halting with an order other than one, which can make the process converge to a point that is not the optimal solution, as mentioned by WEI et al. (2015, 2020). For this purpose, the following criterion is defined:

$$\alpha^{q} = \begin{cases} 1 + |\theta^{q}|(\alpha^{q-1} - 1), & \text{if } \alpha > 1 \\ 1 + |\theta^{q}|(1 - \alpha^{q-1}), & \text{otherwise} \end{cases} \tag{4}$$

where α^{q} is the updated fractional order in function of the parameter θ^{q}, defined as being a reduction rate in relation to the objective function. Mathematically, this can be represented as:

$$\theta^{q} = \frac{f^{q}}{f_{\text{worst}}} \tag{5}$$

where f^{q} and f_{worst} represent the values of the objective function at the q-th iteration and considering an initial estimative previously defined, respectively.

In general, as the natural tendency for the value of the objective function is to be reduced along the iterations (for a minimization problem), the value of the reduction rate is also reduced, taking the value of α to converge to one, i.e., at the end of optimization process the value of the reduction rate tend to zero and the fractional derivative tend to one, increasing the chance of an optimal solution to be obtained. The major advantage of this approach is that it allows the reduction of the number of iterations while the optimal solution can be found from an arbitrary value for α. The iterative procedure regarding the FSDM-θ is presented in Algorithm 2.

Algorithm 2: Fractional Steepest Descent Method with Self-Adaptive Order

Input: Information about the problem (number of design variables, objective function) and method parameters (fraction order, initial estimative, strategy to update the size step, stopping criterion, tolerance)
1 . Start of the optimization process:
2. Initialize the counter (*q*)
3. **While** the stopping criterion is not satisfied do
    4. Find the value of the (η^*)
    5. Update the value of the design variables by using Eq. (3)
    6. Evaluate the value of the reduction rate by using Eq. (5)
    7. Update the value of the fractional order by using Eq. (4)
    8. *q = q + 1*
9. **End While**
10. End
Output: Optimal solution, iterations number, convergence rate, fractional order along the evolutionary process

4. RESULTS AND DISCUSSION

To assess the proposed methodology (FSDM-θ), two mathematical test cases are considered. The obtained results are compared with those obtained by using the SDM with integer order, the FSDM (with constant fractional order), and the Newton Method (NM). For this purpose, the stopping criterion considered is the sum of the absolute error (in terms of the design variables) less than a tolerance (10^{-6}), as well as the use of different initial estimative to evaluate the number of iterations (n_{iter}) required for each approach.

4.1 Test Case 1

Consider an unconstrained minimization problem given by Edgar et al. (2001):

$$f(x_1, x_2) = x_1^2 - 3x_1x_2 + 4x_2^2 + x_1 - x_2 \quad (6)$$

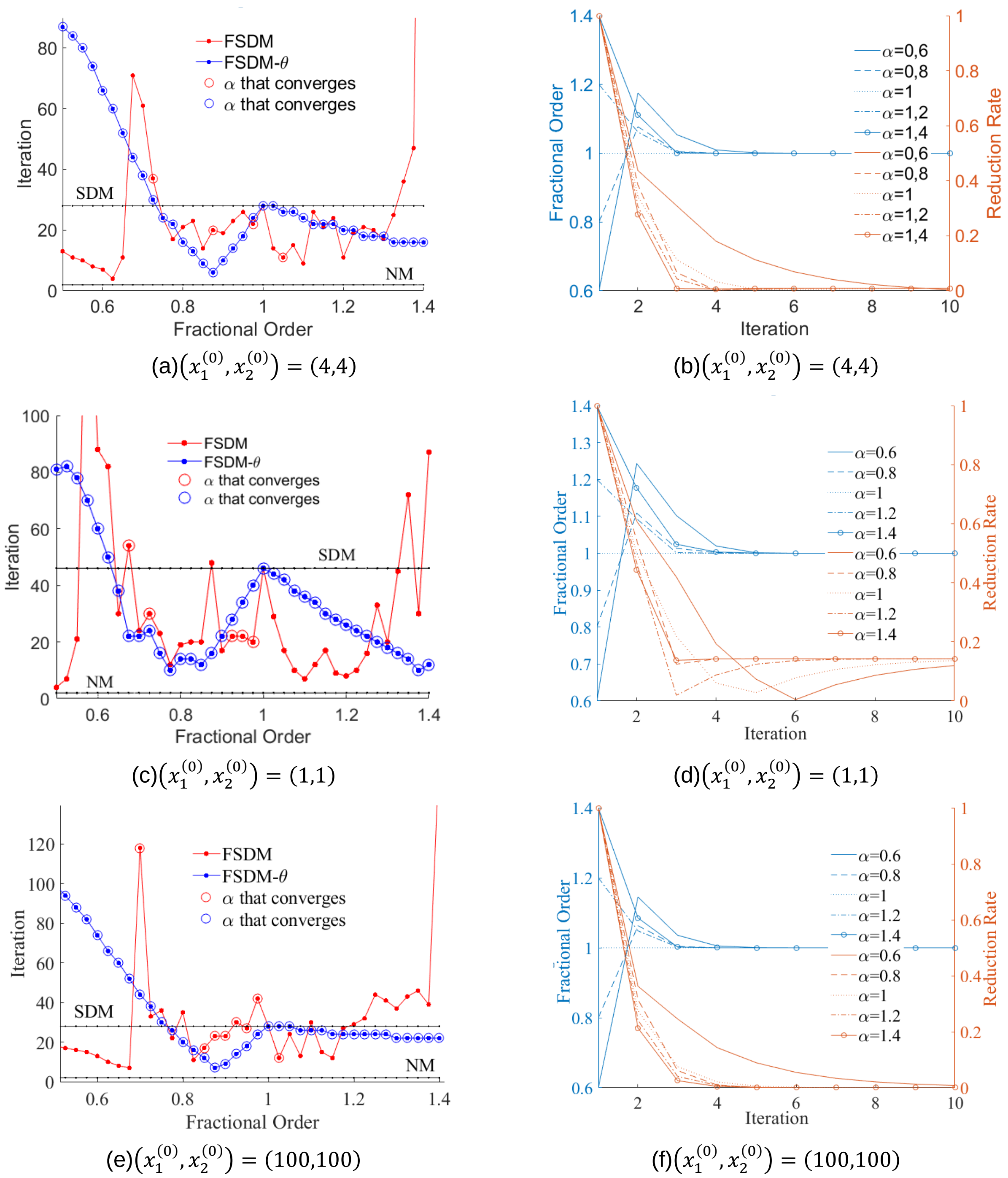
where x_1 and x_2 are the design variables and f is the objective function. This problem presents an optimal solution equal to $(x_1, x_2, f) = (-0{,}714283,\ -0{,}142856,\ -0{,}285714)$.

Table 1 present the obtained results by using SDM, NM, FSDM, and FSDM-θ strategies considering different values for α (for the last two algorithms) and initial condition $\left(x_1^{(0)}, x_2^{(0)}\right) = (4{,}4)$. In this table, it is possible to observe that, for different values of the fractional order, the FSDM converge to a non-optimal solution. This result is in accordance with Wei et al. (2015, 2020) and is due to the non-locality related to fractional models. On the other hand, FSDM-θ was always able to converge to the optimal solution, i.e., by employing a strategy for the dynamic update of the fractional order. Both SDM and NM algorithms always converged for the optimal solution without any difficulty. In terms of the number of iterations (n_{iter}), as expected, NM presents the best performance due to the use of information about the gradient vector and the Hessian matrix. However, when comparing the SDM, FSDM, and FSDM-θ algorithms, it is observed that the proposed methodology (FSDM-θ) can lead to a reduction in the number of iterations, in relation to SDM, unlike FSDM, in which not always the best tradeoff between convergence and number of iterations is obtained.

Table 1: Results obtained considering the SDM, NM, FSDM and FSDM-θ strategies by using different values for α in the first test case.

	α	x_1	x_2	f	n_{iter}
FSDM	0.6	4.925202	2.272114	13.988852	7
	0.8	−0.711257	−0.144257	−0.285684	21
	1.0	−0.714283	−0.142856	−0.285714	28
	1.2	−0.485301	−0.060624	−0.262721	11
	1.4	−0.023211	0.149174	−0.072447	299
FSDM-θ	0.6	−0.714283	−0.142855	−0.285714	66
	0.8	−0.714284	−0.142856	−0.285714	16
	1.0	−0.714283	−0.142856	−0.285714	28
	1.2	−0.714284	−0.142856	−0.285714	20
	1.4	−0.714285	−0.142857	−0.285714	16
SDM	–	−0.714284	−0.142856	−0.285714	28
NM	–	−0.714285	−0.142857	−0.285714	2

Figures 1(a,c,e) present the number of iterations required until convergence, in relation to the fractional order in FSDM and FSDM-θ, as well as its comparison with other required values by using SDM and NM considering different initial conditions: $\left(x_1^{(0)}, x_2^{(0)}\right) = (4,4)$, $\left(x_1^{(0)}, x_2^{(0)}\right) = (1,1)$,and $\left(x_1^{(0)}, x_2^{(0)}\right) = (100,100)$.



(a)$\left(x_1^{(0)}, x_2^{(0)}\right) = (4,4)$ (b)$\left(x_1^{(0)}, x_2^{(0)}\right) = (4,4)$

(c)$\left(x_1^{(0)}, x_2^{(0)}\right) = (1,1)$ (d)$\left(x_1^{(0)}, x_2^{(0)}\right) = (1,1)$

(e)$\left(x_1^{(0)}, x_2^{(0)}\right) = (100,100)$ (f)$\left(x_1^{(0)}, x_2^{(0)}\right) = (100,100)$

Figure 1: Convergence of the SDM, NM, FSDM, and FSDM-θ algorithms, fractional order, and reduction rate variations in the FSDM-θ for the first test case.

In these figures, a similar behavior can be seen for the FSDM-θ, i.e., this approach was always able to find the best solution, regardless of the initial fractional order considered. Furthermore, for values of α greater than, approximately, 0.7, the proposed approach was always able to, at least, equal the performance of the SDM in terms of the n_{iter}. FSDM also always converged to a solution, regardless of the values of α considered. However, as seen in these figures, there were few times the FSDM converged to the optimal solution. In relation to NM, the FSDM-θ strategy, as expected, always had the worst performance in terms of the n_{iter}. Despite this, it is possible to observe in these figures that there is an optimal value for α that minimizes the number of iterations. Thus, even not using a second-order optimization method, it seems that the proposed methodology has good potential.

In Figures 1(b,d,f) the fractional order and the reduction rate variation in the FSDM-θ algorithm for different initial conditions are shown. In these figures, it is observed that as the optimization process progresses, the value of θ tends to zero, which makes the value of α tends to one, minimizing the chance of convergence to a value different from the optimal one. It is important to emphasize that, for the proposed equation (see Eq. (4)), the value of α, even starting from a value smaller than one, soon assumes a value greater than one, which is, throughout the optimization process, directed to the integer order.

4.2 Test Case 2

The second application also considers an unconstrained minimization problem given by Edgar et al.(2001):

$$f(x_1, x_2) = 5x_1^2 + x_2^2 + 2x_1x_2 - 12x_1 - 4x_2 + 8 \tag{7}$$

where x_1 and x_2 are the design variables and f is the objective function. This problem presents an optimal solution equal to $(x_1, x_2, f) = (1,1,0)$.

Table 2 presents the obtained results by using SDM, NM, FSDM, and FSDM-θ strategies considering different values for α (for the last two algorithms) and initial condition $\left(x_1^{(0)}, x_2^{(0)}\right) = (10,10)$. As observed in the previous test case, FSDM converges to a non-optimal solution, except for α equal to one. On the other hand, the remaining algorithms (FSDM-θ, SDM and NM) converged to the optimal solution. In terms of the number of iterations (n_{iter}), as expected, NM presents the best performance. However, when comparing FSDM and FSDM-θ, it is observed that, although the proposed methodology presents higher values for the number of iterations in comparison with FSDM, it always converged to the optimal solution.

Table 2: Results obtained considering the SDM, NM, FSDM, and FSDM-θ strategies by using different values for α in the second test case.

	α	x_1	x_2	f	n_{iter}
FSDM	0.6	0.899630	1.488201	0.190709	5
	0.8	0.930197	1.541759	0.242232	13
	1.0	0.999999	1.000000	0.000000	10
	1.2	1.024931	0.884773	0.010639	37
	1.4	0.814318	1.683908	0.386139	6
FSDM-θ	0.6	1.000000	0.999999	0.000000	8
	0.8	0.999999	0.999999	0.000000	8
	1.0	1.000000	1.000000	0.000000	10
	1.2	0.999999	1.000000	0.000000	6
	1.4	1.000000	1.000000	0.000000	12
SDM	–	0.999999	0.999999	0.000000	5
NM	–	0.999999	0.999999	0.000000	2

In Figures 2(a,c,e) the convergence analysis of the FSDM and FSDM-θ algorithms are presented, as well as their comparison with the values required by SDM and NM considering different initial conditions. In these figures, we can observe that the proposed methodology always converged to the optimal solution. Despite this, depending on the fractional order and the initial estimate, FSDM-θ can result in a higher number of iterations compared to SDM and FSDM, which presents difficulties in converging to the optimal solution. As expected, the NM always presented the best performance in terms of the number of iterations. However, it is possible to see in Figure 2(a,c,e) that there is an optimal value for α so that the number of iterations can be minimized, i.e., the computational cost can be reduced in comparison with NM. In terms of the reduction rate and the fractional order, as seen in Figures 2(b,d,f), there is a similar behavior in relation to the first case study, i.e., the value of θ tends to zero and, consequently, the value of α is close to one. As mentioned earlier, this characteristic allows that, at the end of the optimization process, the fractional derivative become an integer value, ensuring that the optimizer always converges to the optimal solution.

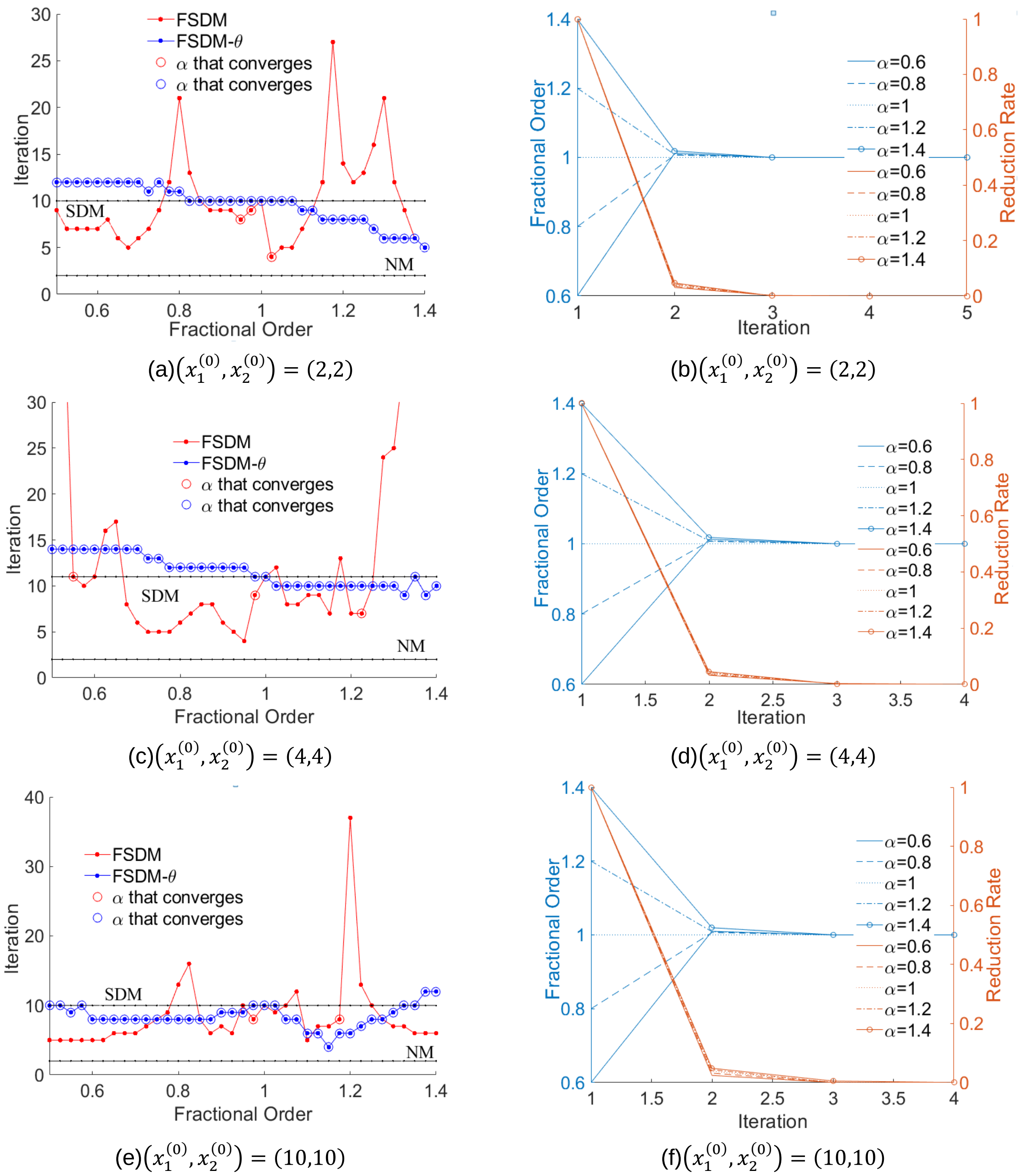


Figure 2: Convergence of the SDM, NM, FSDM, and FSDM-θ algorithms, fractional order, and reduction rate variations in the FSDM-θ for the second test case.

5. CONCLUSIONS

In this work, an approach for the dynamic update of the fractional order applied to the Steepest Descent Method was presented. This new approach, called the Fractional Steepest Descent Method with Self-Adaptive Order (FSDM-θ), is based on the definition of the reduction rate to update the fractional order. The results obtained with the application of

FSDM-θ in two mathematical functions demonstrate that the proposed methodology has the potential to improve the performance of FSDM, both in terms of convergence and number of iterations. In this case, these characteristics can be improved in relation to the classical Newton Method without, necessarily, computing the Hessian matrix.

As proposals for future works, we intend to test this methodology in problems with constraints and other characteristics, assess other fractional derivative types, as well as their numerical approximations, and propose other strategies for updating the fractional order.

REFERENCES

CAPUTO, M. Lectures on Seismology and Rheological Tectonics, Lecture Notes, Universit`a "La Sapienza", Dipartimento di Fisica, Roma, 1999.

EDGAR, T. F., HIMMELBLAU, D. M., & LASDON, L. S. **Optimization of Chemical Processes**. McGraw-Hill, New York, 2001.

YI-FEI, P., JI-LIU, Z., & XIAO, Y. Fractional Differential Mask: A Fractional Differential-Based Approach for Multiscale Texture Enhancement. IEEE **Transactions on Image Processing**, 19(2), 491–511, 2010. https://doi.org/10.1109/TIP.2009.2035980.

LECUN, Y., BENGIO, Y., & HINTON, G. **Deep Learning. Nature,** 521, 436–444, 2015. https://doi.org/10.1038/nature14539

GE, Z., DING, F., XU, L., ALSAEDI, A., & HAYAT, T. Gradient-based Iterative Identification Method for Multivariate Equation-Error Auto Regressive Moving Average Systems using the Decomposition Technique. **Journal of the Franklin Institute**, 356(3), 1658–1676, 2019. https://doi.org/10.1016/j.jfranklin.2018.12.002

GÓMEZ-AGUILAR, J. F., ROSALES-GARCÍA, J. J., BERNAL-ALVARADO, J. J., CÓRDOVA-FRAGA, T., & GUZMÁN-CABRERA, R., Fractional Mechanical Oscillators. **Revista Mexicana de Física,** 58. 348-352, 2012.

LIMA, J. V. C. F., LOBATO, F. S., & STEFFEN JR, V. Solution of Fractional Optimal Control Problems by using Orthogonal Collocation and Multi-objective Optimization Stochastic **Fractal Search. Advances in Computational Intelligence,** 1(4), 3, 2021.https://doi.org/10.1007/s43674-021-00003-x

LIU, J., ZHAI, R., LIU, Y., LI, W., WANG, B., & HUANG, L. A Quasi Fractional Order Gradient Descent Method with Adaptive Step Size and its Application in System Identification. **Applied Mathematics and Computation**, 393, 125797, 2021. https://doi.org/10.1016/j.amc.2020.125797

PU, Y. F., ZHOU, J. L., ZHANG, Y., ZHANG, N., HUANG, G., & SIARRY, P. Fractional Extreme Value Adaptive Training Method: Fractional Steepest Descent Approach. IEEE

Transactions on Neural Networks and Learning Systems, 26(4), 653–662, 2015. https://doi.org/10.1109/TNNLS.2013.2286175

REN, Z., XU, C., ZHOU, Z., WU, Z., & CHEN, T. Boundary Stabilization of a Class of Reaction–Advection–Diffusion Systems via a Gradient-Based Optimization Approach. **Journal of the Franklin Institute**, 356(1), 173–195, 2019. https://doi.org/10.1016/j.jfranklin.2018.10.013

VANDERPLAATS, G. N. **Numerical Optimization Techniques for Engineering Design: with Applications**. McGraw-Hill Book Company, New York. 1984.

WEI, Y., KANG, Y., YIN, W., & WANG, Y. Generalization of the Gradient Method with Fractional Order Gradient Direction. **Journal of the Franklin Institute**, 357(4), 2514–2532, 2012. https://doi.org/10.1016/j.jfranklin.2020.01.008

WEI, Y., SUN, Z., HU, Y., & WANG, Y. Online Parameter Estimation Based on Gradient Algorithm for Fractional Order Systems. **Journal of Control and Decision**, 2(4), 219–232, 2015. https://doi.org/10.1080/23307706.2015.1050466