

Um Algoritmo Sem Derivadas para Problemas de Otimização de Menor Valor Ordenado

Anderson Ervino Schwertner¹

PMA/UEM, Maringá, PR

Francisco Nogueira Calmon Sobral²

DMA/UEM, Maringá, PR

Resumo. O problema de otimização de menor valor ordenado (LOVO) envolve minimizar o mínimo entre um número finito de valores de função sobre um conjunto viável. Possui diversas aplicações práticas, tais como alinhamento de proteínas, estimação robusta de parâmetros, otimização de portfólio, entre outros. Neste trabalho, estamos interessados no problema LOVO de otimização não linear restrito, sujeito a um conjunto convexo, fechado e não-vazio, onde cada função é do tipo *blackbox* e continuamente diferenciável. Neste sentido, apresentamos um algoritmo de região de confiança sem derivadas para problemas LOVO com restrições, o qual converge globalmente para pontos fracamente críticos sob condições adequadas.

Palavras-chave. Otimização de Menor Valor Ordenado, Otimização Sem Derivadas, Métodos de Região de Confiança, Convergência Global

1 Introdução

Considere o seguinte problema LOVO de otimização não linear restrito

$$\min_{x \in \Omega} f_{\min}(x) = \min_{x \in \Omega} \min\{f_1(x), \dots, f_r(x)\}, \quad (1)$$

onde $\Omega \subset \mathbb{R}^n$ é um conjunto convexo fechado e não-vazio, e $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, r$, são funções do tipo *blackbox* (caixa preta), para os quais não temos informações acerca de suas derivadas de primeira ou segunda ordem.

Problemas LOVO possuem diversas aplicações, dentre as quais podemos destacar alinhamento de proteínas [1, 2, 11, 13], estimação robusta de parâmetros [2, 4, 10, 16], otimização de portfólio [3, 11] e busca de padrões ocultos em conjuntos de dados [2]. Dadas as especificidades deste tipo de problema, diversos autores desenvolveram versões LOVO de algoritmos clássicos de otimização contínua, tais como os algoritmos de Levenberg-Marquardt [4, 10, 16], Lagrangiano Aumentado [2, 12], Região de Confiança [1] e Busca Linear [2, 13].

Este artigo é organizado como segue. Na Seção 2 introduzimos nosso algoritmo de região de confiança sem derivadas para problemas LOVO. Na Seção 3, discutimos aspectos de análise de convergência e estabelecemos resultados de convergência global. Por fim, realizamos algumas considerações finais na Seção 4.

Por conveniência de notação, $\|\cdot\|$ denota a norma euclidiana para vetores e matrizes, $\mathcal{I} = \{1, 2, \dots, r\}$ e $\bar{B}(x^*, \delta) = \{x \in \mathbb{R}^n \mid \|x - x^*\| \leq \delta\}$.

¹pg54133@uem.br

²fncsobral@uem.br

2 Algoritmo de Região de Confiança Sem Derivadas

No intuito de nos auxiliar na descrição do algoritmo e na exposição dos resultados teóricos, considere a seguinte definição.

Definição 2.1. [4, Definição 1] Dado um ponto viável $x \in \Omega$, definimos o conjunto de funções mínimo de f_{\min} em x como $\mathcal{I}_{\min}(x) := \{i \in \{1, \dots, r\} \mid f_i(x) = f_{\min}(x)\}$.

Consideramos as seguintes hipóteses acerca da função objetivo:

H1 Para cada índice $i \in \mathcal{I}$, a função $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ é continuamente diferenciável, e possui gradiente $\nabla f_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ Lipschitz contínuo, com constante de Lipschitz $L_i > 0$, em um domínio aberto e limitado $\mathcal{X} \subset \mathbb{R}^n$, suficientemente grande.

H2 A função f_i , para $i \in \mathcal{I}$, é limitada inferiormente em $\Omega \subset \mathbb{R}^n$, isto é, existe uma constante $M_i \in \mathbb{R}$ tal que $f_i(x) \geq M_i$, para todo $x \in \Omega$.

Observe que, se a Hipótese H2 é válida e $M := \min_{i \in \mathcal{I}} \{M_i\}$, então temos que $f_{\min}(x) \geq M$, para todo $x \in \Omega$.

2.1 Estrutura de Região de Confiança

O algoritmo descrito nesta seção é baseado na estrutura de região de confiança apresentado em [7, 18]. A cada iteração $k \in \mathbb{N}$, consideramos o iterando corrente $x_k \in \Omega$ e escolhemos um índice $i_k \in \mathcal{I}_{\min}(x_k)$ conforme indicado pelo Algoritmo 2, presente na Seção 2.2. Na sequência, construímos um modelo quadrático para f_{i_k} , o qual assume a forma

$$m_k(x) = m_k(x_k + s) = b_k + \mathbf{g}_k^T s + \frac{1}{2} s^T \mathbf{H}_k s, \quad (2)$$

onde $s = x - x_k \in \mathbb{R}^n$, $b_k \in \mathbb{R}$, $\mathbf{g}_k \in \mathbb{R}^n$ e $\mathbf{H}_k \in \mathbb{R}^{n \times n}$ é uma matriz simétrica. Em [1], uma abordagem semelhante é realizada tomando $\mathbf{g}_k = \nabla f_{i_k}(x_k)$ e $\mathbf{H}_k = \nabla^2 f_{i_k}(x_k)$. Note que para cada iteração k está associado um único índice $i_k \in \mathcal{I}_{\min}(x_k)$, desse modo, ocultaremos a indicação de tal índice na definição do modelo e de outras expressões que dependem da escolha do índice.

Seguimos as ideias desenvolvidas em [14, 18] de empregar dois raios distintos: $\delta_k > 0$, relacionado à qualidade do modelo, e $\Delta_k > 0$, associado à região de confiança. Neste sentido, iremos assumir que o modelo $m_k(x)$ é construído satisfazendo a seguinte hipótese:

H3 Existe uma constante $\kappa_g > 0$ tal que, para todo $k \in \mathbb{N}$, $\|\nabla f_{i_k}(x) - \nabla m_k(x)\| \leq \kappa_g \delta_k$, para todo $x \in \mathcal{X} \cap \overline{B}(x_k, \delta_k)$, onde $i_k \in \mathcal{I}_{\min}(x_k)$ é o índice associado a m_k .

Considere a medida de estacionariedade em x_k para o problema de minimizar o modelo m_k sobre o conjunto Ω por $\pi_k = \|P_\Omega(x_k - \mathbf{g}_k) - x_k\|$, onde P_Ω denota a projeção ortogonal sobre Ω , a qual existe uma vez que o conjunto é convexo e fechado [7, 15].

A cada iteração $k \in \mathbb{N}$, dado um raio $\Delta_k > 0$, aplicamos o modelo (2) para encontrar uma solução aproximada $d_k \in \mathbb{R}^n$ para o subproblema de região de confiança

$$\begin{aligned} &\text{minimizar} && m_k(x_k + d) \\ &\text{sujeito a} && x_k + d \in \Omega \\ &&& \|d\| \leq \Delta_k, \end{aligned} \quad (3)$$

de modo a determinar um novo ponto $x_k + d_k \in \Omega$, onde esperamos que haja um decréscimo no valor de função. O ponto é aceito como novo iterando quando a razão entre as reduções real e prevista

$$\rho_k = \frac{f_{\min}(x_k) - f_{\min}(x_k + d_k)}{m_k(x_k) - m_k(x_k + d_k)} \quad (4)$$

é maior ou igual a uma constante fixada $\eta > 0$. Neste caso, o novo iterando é dado por $x_{k+1} = x_k + d_k$, o modelo é atualizado e o raio da região de confiança Δ_k possivelmente é aumentado. Por outro lado, se $\rho < \eta$, rejeitamos o passo e Δ_k é diminuído.

Para que possamos garantir a convergência global do algoritmo, assumiremos que a solução do subproblema de região de confiança (3) satisfaz a seguinte hipótese:

H4 A direção $d_k \in \mathbb{R}^n$, solução do subproblema de região de confiança (3), satisfaz a condição de decréscimo suficiente

$$m_k(x_k) - m_k(x_k + d_k) \geq \theta \pi_k \min \left\{ \frac{\pi_k}{1 + \|\mathbf{H}_k\|}, \Delta_k, 1 \right\},$$

onde $\theta > 0$ é uma constante que independe da iteração k e da escolha do índice $i_k \in \mathcal{I}_{\min}(x_k)$ associado a m_k .

2.2 Algoritmo

Nesta seção, apresentamos um algoritmo de região de confiança sem derivadas genérico para problemas LOVO, sem especificar as fases de atualização do modelo m_k ou a solução do subproblema (3). O algoritmo é baseado nas ideias discutidas em [1] e nos algoritmos propostos em [5, 18].

Algoritmo 1: ALGORITMO LOVO REGIÃO DE CONFIANÇA SEM DERIVADAS

Entrada: $x_0 \in \Omega$, $\beta > 0$, $0 < \delta_0 \leq \Delta_0$, $0 < \tau_1 < 1 \leq \tau_2 \leq \tau_3$, $\eta_1 \in (0, 1)$, $0 \leq \eta < \eta_1 \leq \eta_2$, $\Gamma = 0$, $1 \leq \Gamma_{\max} \in \mathbb{N}$.

```

1 para  $k = 0, 1, \dots$  faça
2     se  $k = 0$  então
3         Escolha  $i_0 \in \mathcal{I}_{\min}(x_0)$ .
4     senão
5         Atualize o índice  $i_k$  e os raios da região de amostra  $\delta_k$  e de confiança  $\Delta_k$  usando o Algoritmo 2.
6     fim
7     Construa o modelo  $m_k$ .
8     se  $\delta_k > \beta \pi_k$  então
9         Faça  $\delta_{k+1} = \tau_1 \delta_k$ ,  $d_k = 0$ ,  $\rho_k = 0$ ,  $x_{k+1} = x_k$ , e escolha  $\Delta_{k+1} \in [\delta_{k+1}, \Delta_k]$ .
10    senão
11        Encontre uma solução aproximada  $d_k$  para (4) satisfazendo a Hipótese H4.
12        Calcule  $\rho_k$  por (5).
13        se  $\rho_k \geq \eta$  então
14            Faça  $x_{k+1} = x_k + d_k$ .
15        senão
16            Faça  $x_{k+1} = x_k$ .
17        fim
18        Faça
19            
$$\delta_{k+1} = \begin{cases} \tau_1 \delta_k, & \text{se } \rho_k < \eta_1; \\ \tau_2 \delta_k, & \text{se } \rho_k > \eta_2 \text{ e } \|d_k\| = \Delta_k; \\ \delta_k, & \text{caso contrário;} \end{cases} \quad \text{e} \quad \Delta_{k+1} = \begin{cases} \tau_1 \Delta_k, & \text{se } \rho_k < \eta_1; \\ \tau_2 \Delta_k, & \text{se } \rho_k > \eta_2 \text{ e } \|d_k\| = \Delta_k; \\ \Delta_k, & \text{caso contrário.} \end{cases}$$

20        fim
21        Faça  $k = k + 1$ .
22 fim

```

Figura 1: Algoritmo 1 - LOVO Região de Confiança sem Derivadas.

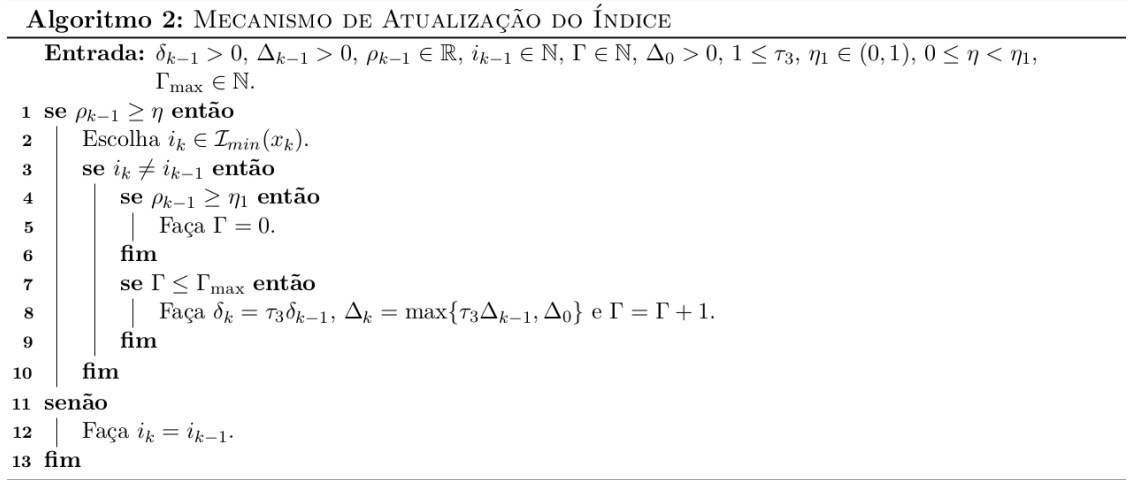


Figura 2: Algoritmo 2 - Mecanismo de Atualização do Índice.

3 Análise de Convergência

Nesta seção, apresentamos alguns resultados de convergência global sob as hipóteses estabelecidas anteriormente. Mais especificamente, mostramos que todo ponto de acumulação de uma sequência gerada pelo Algoritmo 1 é estacionário para algum índice $i \in \mathcal{I}$ e, conseqüentemente, é um ponto fracamente crítico para o problema (1).

Observe que estamos assumindo que o modelo m_k é reconstruído a cada iteração, assim como [18]. Em [17], discutimos como construir modelos satisfazendo as hipóteses adotadas neste trabalho, em especial, a Hipótese H3. Além disso, mostramos que modelos quadráticos determinados e subdeterminados construídos sobre conjuntos de amostra Λ -posicionados podem ser empregados. Do ponto de vista prático, a reconstrução dos modelos a cada iteração do algoritmo possui alto custo computacional, e pode ser substituída por estratégias mais eficientes, como os mecanismos para atualização de modelos apresentados em [8, 14].

O lema a seguir estabelece que a matriz hessiana do modelo m_k é limitada, sendo análogo a [18, Lema 1].

Lema 3.1. *Suponha que as Hipóteses H1 e H3 são válidas. Então, para toda iteração $k \in \mathbb{N}$, $\|\mathbf{H}_k\| \leq \kappa_H - 1$, onde $\kappa_H := 2\kappa_g + L + 1$ e $L := \max_{i \in \mathcal{I}}\{L_i\}$.*

De maneira análoga a [5, 18], considere os seguintes conjuntos de índices

$$\mathcal{S} = \{k \in \mathbb{N} \mid \rho_k \geq \eta\} \text{ e } \bar{\mathcal{S}} = \{k \in \mathbb{N} \mid \rho_k \geq \eta_1\}. \tag{5}$$

Note que $\bar{\mathcal{S}} \subset \mathcal{S}$. O próximo lema é análogo a [5, Lema 3.1] para problemas LOVO e afirma que se o raio da região de confiança é suficientemente pequeno, então a iteração será de sucesso.

Lema 3.2. *Suponha que as Hipóteses H1, H2 e H4 são válidas e considere o conjunto*

$$\mathcal{K} = \left\{ k \in \mathbb{N} : \Delta_k \leq \min \left\{ \frac{\pi_k}{\kappa_H}, \frac{(1 - \eta_1)\pi_k}{c}, \beta\pi_k, 1 \right\} \right\}, \tag{6}$$

onde $c := \frac{L + \kappa_g + \frac{\kappa_H}{2}}{\theta}$. Então, se $k \in \mathcal{K}$, temos que $k \in \bar{\mathcal{S}}$.

Como mencionado anteriormente, o raio da região de amostra δ_k controla a qualidade do modelo. Mais especificamente, a Hipótese H3 nos diz que quanto menor for o valor do raio de amostra, melhor o modelo representará a função f_{i_k} . Portanto, é razoável esperar que δ_k vá para zero. O lema a seguir é baseado em [18, Lema 3] e nos mostra que isso de fato ocorre.

Lema 3.3. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas. Então a sequência $\{\delta_k\}_{k \in \mathbb{N}}$ converge para zero.*

Agora, apresentamos um resultado de convergência fraca para o problema de minimizar o modelo m_k na região viável Ω .

Lema 3.4. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas. Então a sequência $\{\pi_k\}_{k \in \mathbb{N}}$ admite uma subsequência que converge para zero.*

O corolário a seguir é uma consequência direta do lema acima.

Corolário 3.1. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas. Então existe um índice $i \in \mathcal{I}$ que se repete para um número infinito de iterações tal que $\lim_{k \in \mathcal{J}} \pi_k = 0$, onde $\mathcal{J} = \{k \in \mathbb{N} \mid i_k = i\}$.*

Se assumirmos um decréscimo suficiente na função objetivo tomando $\eta > 0$ no Algoritmo 1, podemos provar que qualquer subsequência de $\{\pi_k\}_{k \in \mathbb{N}}$, na qual o mesmo índice $i \in \mathcal{I}$ é escolhido para um número infinito de iterações, converge para zero.

Lema 3.5. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas e $\eta > 0$. Seja $i \in \mathcal{I}$ um índice escolhido para um número infinito de iterações. Então, $\lim_{k \in \mathcal{J}} \pi_k = 0$, onde $\mathcal{J} = \{k \in \mathbb{N} \mid i_k = i\}$.*

O teorema a seguir é o resultado principal deste trabalho. Ele estabelece um resultado de convergência global fraca para pontos estacionários de primeira ordem. Adotamos a medida de criticalidade dada por $\|\mathcal{P}_\Omega(x - \nabla f(x)) - x\|$, proposto por [6] para problemas de otimização envolvendo funções contínuas f em um conjunto viável convexo Ω . Esta medida de criticalidade também é mencionada em [7, Seção 12.1.4], dada sua interpretação geométrica simples, e é empregada em [5, 9, 18] para estabelecer resultados de convergência global.

Teorema 3.1. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas. Seja $\{x_k\}_{k \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo 1. Então,*

1. Se $\eta = 0$, então

$$\liminf_{k \rightarrow \infty} \|\mathcal{P}_\Omega(x_k - \nabla f_{i_k}(x_k)) - x_k\| = 0.$$

Além disso, existe um índice $i \in \mathcal{I}$ escolhido para um número infinito de iterações tal que

$$\lim_{k \in \mathcal{J}} \|\mathcal{P}_\Omega(x_k - \nabla f_{i_k}(x_k)) - x_k\| = 0,$$

onde $\mathcal{J} = \{k \in \mathbb{N} \mid i_k = i\}$.

2. Se $\eta > 0$ e $i \in \mathcal{I}$ é um índice qualquer escolhido para um número infinito de iterações, então

$$\lim_{k \in \mathcal{J}} \|\mathcal{P}_\Omega(x_k - \nabla f_{i_k}(x_k)) - x_k\| = 0,$$

onde $\mathcal{J} = \{k \in \mathbb{N} \mid i_k = i\}$.

3. Se $i \in \mathcal{I}$ é um índice escolhido para um número infinito de iterações e $x_* \in \mathbb{R}^n$ é um ponto de acumulação para a subsequência $\{x_k\}_{k \in \mathcal{J}} \subseteq \{x_k\}_{k \in \mathbb{N}}$ gerada pelo Algoritmo 1, onde $\mathcal{J} = \{k \in \mathbb{N} \mid i_k = i\}$, então $i \in \mathcal{I}_{\min}(x_*)$.

O Teorema 3.1 estabelece que se $\eta > 0$ e $x_* \in \mathbb{R}^n$ é um ponto de acumulação de uma sequência $\{x_k\}_{k \in \mathbb{N}}$ gerada pelo Algoritmo 1, então é possível construir uma subsequência $\{x_k\}_{k \in \mathcal{J}}$ que converge para x_* , onde $\mathcal{J} = \{k \in \mathbb{N} \mid i_k = i\}$, e x_* satisfaz uma condição necessária de otimalidade do tipo gradiente projetado para o problema

$$\begin{aligned} & \text{minimizar} && f_i(x) \\ & \text{sujeito a} && x \in \Omega, \end{aligned} \tag{7}$$

isto é, x_* é um ponto estacionário de primeira ordem para o problema (7) (ver [6] e [7, p. 450]).

A definição e os corolários a seguir nos ajudam a compreender o Teorema 3.1 da perspectiva da teoria clássica dos problemas LOVO.

Definição 3.1. [2, p. 05] *Dado $x_* \in \Omega \subseteq \mathbb{R}^n$ e estabelecida uma condição necessária de otimalidade, a qual denotaremos por CNO, para o problema (7), dizemos que*

1. x_* é fortemente crítico se satisfaz CNO para (7), para todo $i \in I_{\min}(x_*)$.
2. x_* é fracamente crítico se existe $i \in I_{\min}(x_*)$ tal que x_* satisfaz CNO para (7).

Corolário 3.2. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas e $\eta > 0$. Se $x_* \in \mathbb{R}^n$ é um ponto de acumulação de uma sequência $\{x_k\}_{k \in \mathbb{N}}$ gerada pelo Algoritmo 1, então x_* é fracamente crítico.*

Corolário 3.3. *Suponha que as Hipóteses H1, H2, H3 e H4 são válidas. Se o Algoritmo 1 gera uma sequência $\{x_k\}_{k \in \mathbb{N}}$ que converge para $x_* \in \mathbb{R}^n$, então x_* é fracamente crítico.*

4 Considerações Finais

Neste artigo introduzimos uma nova classe de métodos para otimização de menor valor ordenado, ao considerarmos uma abordagem baseada em otimização sem derivadas. Apresentamos um algoritmo de região de confiança sem derivadas para problemas LOVO *blackbox* com restrições, o qual é baseado em [1, 5, 18].

Também discutimos resultados de convergência global para o algoritmo, adotando hipóteses comuns para métodos de região de confiança sem derivadas, assim como tais resultados podem ser interpretadas do ponto de vista da teoria clássica dos problemas LOVO.

Futuros trabalhos podem estudar aspectos de convergência local, e de convergência global para pontos fortemente críticos. Assim como estender a abordagem adotada neste artigo para problemas de Valor Ordenado (OVO), o estudo de problemas OVO com ruídos.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Referências

- [1] R. Andreani, J. M. Martínez e L. Martínez. “Trust-region superposition methods for protein alignment”. Em: **IMA journal of numerical analysis** 28.4 (2008), pp. 690–710. DOI: 10.1093/imanum/drn012.

- [2] R. Andreani et al. “Low order-value optimization and applications”. Em: **Journal of Global Optimization** 43.1 (2009), pp. 1–22. DOI: 10.1007/s10898-008-9280-3.
- [3] E. G. Birgin et al. “Low order-value approach for solving VaR-constrained optimization problems”. Em: **Journal of Global Optimization** 51.4 (2011), pp. 715–742. DOI: 10.1007/s10898-011-9656-7.
- [4] E. V. Castelani et al. “A robust method based on LOVO functions for solving least squares problems”. Em: **Journal of Global Optimization** 80.2 (2021), pp. 387–414. DOI: 10.1007/s10898-020-00970-4.
- [5] P. D. Conejo et al. “Global convergence of trust-region algorithms for convex constrained minimization without derivatives”. Em: **Applied Mathematics and Computation** 220 (2013), pp. 324–330. DOI: 10.1016/j.amc.2013.06.041.
- [6] A. R. Conn, N. I. M. Gould e P. L. Toint. “Global convergence of a class of trust region algorithms for optimization with simple bounds”. Em: **SIAM Journal on Numerical Analysis** 25.2 (1988), pp. 433–460.
- [7] A. R. Conn, N. I. M. Gould e P. L. Toint. **Trust-region methods**. SIAM, 2000. ISBN: 0-89871-460-5.
- [8] A. R. Conn, K. Scheinberg e L. N. Vicente. **Introduction to derivative-free optimization**. SIAM, 2009. ISBN: 978-0-898716-68-9.
- [9] A. R. Conn et al. “Convergence properties of minimization algorithms for convex constraints using a structured trust region”. Em: **SIAM Journal on Optimization** 6.4 (1996), pp. 1059–1086.
- [10] E. S. Duran. “Uma classe de métodos do tipo Levenberg-Marquardt com passos múltiplos para problemas de Otimização de Menor Valor Ordenado”. Dissertação de mestrado. UEM, 2020.
- [11] J. M. Martínez. “Generalized order-value optimization”. Em: **Top** 20.1 (2012), pp. 75–98. DOI: 10.1007/s11750-010-0169-1.
- [12] J. M. Martínez. “Order-value optimization and new applications”. Em: **ICIAM 07: 6th International Conference on Industrial and Applied Mathematics, Zürich, Switzerland, 16-20 July 2007: Invited Lectures**. European Mathematical Society. 2009, pp. 279–296. ISBN: 9783037190562.
- [13] L. Martínez, R. Andreani e J. M. Martínez. “Convergent algorithms for protein structural alignment”. Em: **BMC bioinformatics** 8.1 (2007), pp. 1–15. DOI: 10.1186/1471-2105-8-306.
- [14] M. J. D. Powell. “The BOBYQA algorithm for bound constrained optimization without derivatives”. Em: **Cambridge NA Report NA2009/06, University of Cambridge, Cambridge** (2009), pp. 26–46.
- [15] A. A. Ribeiro e E. W. Karas. **Otimização contínua: Aspectos Teóricos e Computacionais**. Cengage Learning, 2013. ISBN: 978-85-221-1501-3.
- [16] A. E. Schwertner. “O método de Levenberg-Marquardt para problemas de otimização de menor valor ordenado”. Dissertação de mestrado. UEM, 2019.
- [17] A. E. Schwertner e F. N. C. Sobral. **On complexity constants of linear and quadratic models for derivative-free trust-region algorithms**. arXiv:2205.11358. 2021.
- [18] A. Verdério et al. “On the construction of quadratic models for derivative-free trust-region algorithms”. Em: **EURO Journal on Computational Optimization** 5.4 (2017), pp. 501–527. DOI: 10.1007/s13675-017-0081-7.