Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 9, n. 1, 2022.

Trabalho apresentado no XLI CNMAC, Unicamp - Campinas - SP, 2022.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics
Preprint

Um Algoritmo Inercial Inexato para Funções DC em Variedades de Hadamard

João S. Andrade[1]
Departmento de Matemática, CCN/UFPI, Teresina, PI
CSHNB/UFPI, Picos, PI
Jurandir de O. Lopes[2] João Carlos de O. Souza[3]
Departmento de Matemática, CCN/UFPI, Teresina, PI

Resumo. Neste trabalho, propomos uma versão inexata do algoritmo de ponto proximal inercial para diferença de funções convexas em variedades de Hadamard. Em cada subproblema resolvemos a condição de otimalidade de primeira ordem de forma aproximada, porém controlada por um erro. Sob condições razoáveis provamos que todo ponto de acumulação da sequência é um ponto crítico da função objetivo.

Palavras-chave. Método do ponto proximal, versão inexata, funções DC, variedades de Hadamard

1 Introdução

Neste artigo, estamos interessados em encontrar soluções (pontos críticos) de problemas de otimização DC em variedades de Hadamard. Esse problema é denotado da seguinte forma:

$$\min_{x\in M} f(x), \quad \text{com} \quad f(x) = g(x) - h(x), \tag{1}$$

onde $g, h : M \to \mathbb{R}$ são funções convexas continuamente diferenciáveis e $\inf\limits_{x\in M} f(x) > -\infty$.

O interesse pela teoria das funções DC aumentou muito nos últimos anos. No cenário euclidiano, existem muitos trabalhos dedicados a teoria das funções DC em diferentes contextos, veja por exemplo [3, 4, 8, 13] e referências contidas neles. No contexto Riemanniano, recentemente, Almeida et al. [1], Andrade et al. [2], e Souza e Oliveira [12] propuseram algoritmos para resolver o problema de encontrar pontos críticos da função objetivo f. Alguns trabalhos propuseram melhorias computacionais para resolver (1), veja por exemplo [1–4, 12].

O objetivo deste artigo é apresentar uma versão inexata do método do ponto proximal inercial para funções DC(IDCPPA) em variedades de Hadamard para resolver (1), considerado em Andrade et al. [2]. De um modo geral as versões inexatas de algoritmos de ponto proximal tendem a ser mais eficazes computacionalmente como observado por Rockafellar em [9], onde é proposto um algoritmo de ponto proximal inexato para encontrar zeros de operadores monótonos em espaços de Hilbert. Assim, alguns pesquisadores concentraram sua atenção nos algoritmos de pontos proximais inexatos, como Rockafellar [9], Solodov e Svaiter [11] e referências contidas neles.

Considerar versões inexatas em variedades Riemannianas tem uma importância maior, pois algoritmos de pontos proximais em variedades Riemannianas são métodos implícitos em essência, o custo de resolver os subproblemas exatos é bastante caro a cada etapa da iteração. Por isso para

[1]joaosa.mat@ufpi.edu.br
[2]jurandir@ufpi.edu.br
[3]joaocos.mat@ufpi.edu.br

DOI: 10.5540/03.2022.009.01.0315

ter um melhor desempenho computacional é importante considerar versões inexatas de métodos de ponto proximal, como feito em [12, 14, 16].

Este artigo está organizado da seguinte forma. Na Seção 2, apresentamos algumas notações, conceitos básicos e resultados de otimização em variedades Riemannianas que serão usados ao longo do artigo. Na seção 3 apresentamos o algoritmo do ponto proximal inercial inexato para funções DC e analisamos suas propriedades de convergência. Por fim, na seção 4 apresentamos algumas considerações finais.

2 Preliminares

Ao longo deste artigo, M será uma variedade de Hadamard, isto é, M denota uma variedade Riemanniana $n-$dimensional completa, simplesmente conexa de curvatura seccional não positiva, com a função distância $d : M \times M \to \mathbb{R}$. As notações, conceitos e resultados preliminares de geometria Riemanniana usados ao longo deste artigo podem ser encontrados em Sakai [10], Souza e Oliveira [12] e Udriste [15].

Seja $p \in M$ um ponto arbitrário, o espaço tangente de M em p é denotado por T_pM e o fibrado tangente de M é dado por $TM = \cup_{p\in M} T_pM$, que é naturalmente uma variedade. Para qualquer $x \in M$ podemos definir a inversa da aplicação exponencial $\exp^{-1}_x : M \to T_xM$ que é C^∞. Sendo $d(x, x') = \| \exp^{-1}_{x'} x\|$, a aplicação $\rho_{x'} : M \to \mathbb{R}$ definida por $\rho_{x'}(x) = \frac{1}{2}d^2(x, x')$ é C^∞ e seu gradiente em x é grad $\rho_{x'}(x) = -\exp^{-1}_{x'} x$; veja Sakai [10, Proposição 4.8 pagina 108].

Lembramos que um triângulo geodésico $\triangle(p_1, p_2, p_3)$ de uma variedade Riemanniana é um conjunto que consiste em três pontos p_1, p_2 e p_3 chamados *vértices* e três geodésicas mínimas $\gamma_i : [0, l_i] \to M$ ligando esses três pontos. O ângulo $\alpha_i := \angle(\gamma_i'(0), -\gamma_{i-1}'(l_{i-1}))$ é chamado *ângulo* interno do vértice correspondente.

Teorema 2.1. *(Teorema da Comparação para Triângulos) Seja $\triangle(p_1, p_2, p_3)$ um triângulo geodésico. Denote, para cada $i = 1, 2, 3$, por $\gamma_i : [0, l_i] \to M$ a geodésica ligando p_i a p_{i+1}, $l_i = L(\gamma_i)$ e $\alpha_i := \angle(\gamma_i'(0), -\gamma_{i-1}'(l_{i-1}))$. Então, $\alpha_1 + \alpha_2 + \alpha_3 \leq \pi$ e*

$$l_i^2 + l_{i+1}^2 - 2l_il_{i+1} \cos\alpha_{i+1} \leq l_{i-1}^2. \tag{2}$$

Em termos de distância e a aplicação exponencial, a desigualdade (2) pode ser reescrita como

$$d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2\langle \exp^{-1}_{p_{i+1}} p_i, \exp^{-1}_{P_{i+1}} p_{i+2}\rangle \leq d^2(p_{i-1}, p_i),$$

sendo $\langle \exp^{-1}_{p+1} p_i, \exp^{-1}_{p+1} p_{i+2}\rangle = d(p_{i+1}, p_i)d(p_{i+1}, p_{i+2}) \cos\alpha_{i+1}$.

Demonstração. Veja [10, Proposição 4.5 página 223]. □

Usando as propriedades do transporte paralelo e da aplicação exponencial, obtemos a seguinte proposição que será usada com frequência na convergência do método.

Proposição 2.1. *Seja M uma variedade de Hadamard. Sejam $x, y \in M$ e $\{x^k\}, \{y^k\} \subset M$ tais que $x^k \to x$ e $y^k \to y$. Então, as seguintes condições são válidas:*

(i) Para todo $z \in M$, temos $\exp^{-1}_{x^k} z \to \exp^{-1}_x z \quad e \quad \exp^{-1}_z x^k \to \exp^{-1}_z x$.

(ii) Se $v^k \in T_{x^k}M$ e $v^k \to v$, então $v \in T_xM$.

(iii) Para todo $u \in T_xM$, a função $F : M \to TM$ definida por $F(z) = P_{z,x}u$ para cada $z \in M$ é contínua em M.

(iv) $\exp^{-1}_{y^k} x^k \to \exp^{-1}_y x$.

Demonstração. Veja [5, Lema 1.1]. □

Uma função $f : M \to \mathbb{R}$ é convexa se, para todo segmento de geodésica $\gamma : [a, b] \to M$ ligando p a q em M, a composição $f \circ \gamma : [a, b] \to \mathbb{R}$ é convexa, isto é,

$$f(\gamma(t)) = (f \circ \gamma)((1-t)s_1 + ts_2) \leq (1-t)(f \circ \gamma)(s_1) + (f \circ \gamma)(s_2),$$

para todo $t \in [0, 1]$ e quaisquer $s_1, s_2 \in [a, b]$, tal que $\gamma(a) = p$ e $\gamma(b) = q$. Se a desigualdade acima for estrita para $s_1 \neq s_2$, então f é estritamente convexa. A função $f : M \to \mathbb{R}$ é fortemente convexa se existir uma constante positiva m, para todo segmento de geodésica $\gamma : [a, b] \to M$ ligando p a q em M, a composição $f \circ \gamma : [a, b] \to \mathbb{R}$ é fortemente convexa, isto é,

$$f(\gamma(t)) = (f \circ \gamma)((1-t)s_1 + ts_2) \leq (1-t)(f \circ \gamma)(s_1) + (f \circ \gamma)(s_2) - \frac{m}{2}\|\gamma'(t)\|(1-t)t$$

para todo $t \in [0, 1]$ e quaisquer $s_1, s_2 \in [a, b]$, tal que $\gamma(a) = p$ e $\gamma(b) = q$.

Seja $f : M \to \mathbb{R}$ uma função convexa. O subdiferencial de f em $x \in M$ é definido por

$$\partial f(x) = \{u \in T_x M; \langle u, \exp_x^{-1} y\rangle \leq f(y) - f(x), \forall y \in M\}.$$

O vetor $u \in \partial f(x)$ é chamado o subgradiente de f em x. Em particular, se f é diferenciável, então $\partial f(x) = \{\text{grad } f(x)\}$ para todo $x \in M$ e $f : M \to \mathbb{R}$ é convexa se, e somente se, $f(\exp_p v) \geq f(p) + \langle \text{grad } f(p), v\rangle$, $\forall p \in M$, $\forall v \in T_p M$; ver por exemplo Andrade et al. [2].

Proposição 2.2. *Seja $\{x^k\}$ uma sequência limitada em M. Se a sequência $\{v^k\}$ é tal que $v^k \in \partial f(x^k)$ para cada $k \in \mathbb{N}$, então $\{v^k\}$ é também limitada.*

Demonstração. Veja [6, Proposição 3.2]. □

3 Algoritmo ponto proximal inercial inexato

Nesta seção, apresentaremos uma versão inexata do algoritmo de ponto proximal inercial para funções DC em variedades de Hadamard proposto por Andrade et al. [2].

Sem perda de generalidade podemos supor a seguinte hipótese sobre a componente DC h em (1):

(H): A função h é fortemente convexa em M com parâmetro $\rho > 0$, ou seja, para cada $v \in \partial h(x)$, temos que

$$h(y) \geq h(x) + \langle v, \exp_x^{-1} y\rangle + \frac{\rho}{2} d^2(x, y),\ \forall x, y \in M.$$

Com efeito, se não é válida (**H**) para uma certa função DC $f = \varphi - \phi$ podemos obter outra decomposição DC de f satisfazendo a condição **(H)** adicionando uma função fortemente convexa arbitrária $\psi : M \to \mathbb{R}$ às funções componentes, ou seja, $f = g - h$ com $g = \varphi + \psi$ e $h = \phi + \psi$. Podendo sempre tomar $\psi(\cdot) = \frac{1}{2} d^2(\cdot, y)$, que é uma função fortemente convexa pelo Corolário 3.1 em Cruz Neto [7].

A seguir, definiremos uma versão inexata do algoritmo IDCPPA proposto por Andrade et al. [2].

Algoritmo 1: IDCPPA Inexato

Passo 1: Dados um ponto inicial $x^0 \in M$, $\gamma_k \in [0, \frac{\rho}{2})$ e uma sequência limitada de números positivos $\{\mu_k\}$ tal que $0 < a \leq \mu_k \leq b$. Defina $x^{-1} = x^0$.

Passo 2: Dado $x^k \in M$, defina $d^k = \gamma_k \exp^{-1}_{x^k} x^{k-1}$. Defina

$$y^k = \exp_{x^k} \mu_k(w^k + d^k), \text{ com } w^k = \text{grad } h(x^k). \tag{3}$$

Passo 3: Encontre $x^{k+1} \in M$ e $e^{k+1} \in T_{x^{k+1}}M$ tais que

$$e^{k+1} = \text{grad } g(x^{k+1}) - \frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k, \tag{4}$$

onde $\|e^{k+1}\| \leq \eta d(x^{k+1}, x^k)$ e $\eta \in [0, \frac{1}{b})$.

Passo 4: Se $x^{k+1} = x^k$ e $d^k = 0$, pare. Caso contrário, faça $k := k+1$ e retorne ao Passo 2.

Observe que se $\eta = 0$, obtemos $e^{k+1} = 0$. Assim, (4) se torna $0 = \text{grad } g(x^{k+1}) - \frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k$ equivalente a

$$x^{k+1} = \arg\min_{x \in M} \{g(x) + \frac{1}{2\mu_k} d^2(x, y^k)\}.$$

Com isso, o Algoritmo 1 coincide com o Algoritmo IDCPPA em Andrade et al. [2].

A partir de agora, consideramos $\{x^k\}$ a sequência gerada pelo Algoritmo 1.

Proposição 3.1. *Suponha que o Algoritmo 1 termine na iteração k, então x^k é um ponto crítico de f.*

Demonstração. De (3) e (4), temos que $\frac{1}{\mu_k} \exp^{-1}_{x^k} y^k - d^k = \text{grad } h(x^k)$ e $e^{k+1} + \frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k = \text{grad } g(x^{k+1})$. Assim, se $x^{k+1} = x^k$ e $d^k = 0$, segue que $e^{k+1} = 0$. Então,

$$\frac{1}{\mu_k} \exp^{-1}_{x^k} y^k = \text{grad } h(x^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp^{-1}_{x^k} y^k = \text{grad } g(x^k).$$

Portanto, $\text{grad } f(x^k) = \text{grad } g(x^k) - \text{grad } h(x^k) = 0$, isto é, x^k é um ponto crítico de f. □

3.1 Análise de Convergência

Tendo em vista a Proposição 3.1, assumimos que $x^{k+1} \neq x^k$ para todo $k \in \mathbb{N}$, caso contrário, o algoritmo retorna um ponto crítico da função objetivo f. Além disso, iremos assumir que (1) tem solução, isto é, $S = \{x \in M; \text{grad } f(x) = 0\}$ o conjunto dos pontos críticos de f é não vazio.

A seguinte proposição desempenha um papel importante na análise de convergência do Algoritmo 1.

Proposição 3.2. *Suponha que a hipótese **(H)** é satisfeita. Então a seguinte desigualdade é válida:*

$$f(x^{k+1}) + \left(\frac{1-\eta b}{b} + \frac{\rho}{4}\right) d^2(x^{k+1}, x^k) \leq f(x^k) + \frac{\rho}{4} d^2(x^k, x^{k-1}), \quad \forall k \in \mathbb{N}.$$

Demonstração. De (3), temos que $\frac{1}{\mu_k} \exp^{-1}_{x^k} y^k - d^k = w^k = \text{grad } h(x^k)$. Por **(H)**, segue que

$$h(x^{k+1}) \geq h(x^k) + \langle \frac{1}{\mu_k} \exp^{-1}_{x^k} y^k - d^k, \exp^{-1}_{x^k} x^{k+1} \rangle + \frac{\rho}{2} d^2(x^k, x^{k+1}). \tag{5}$$

Usando agora (4), temos $e^{k+1} + \frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k = \text{grad } g(x^{k+1})$. Então, pela convexidade de g, temos

$$g(x^k) \geq g(x^{k+1}) + \langle e^{k+1} + \frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k, \exp^{-1}_{x^{k+1}} x^k \rangle. \tag{6}$$

DOI: 10.5540/03.2022.009.01.0315

Somando (5) e (6), obtemos

$$f(x^k) - f(x^{k+1}) + \langle d^k, \exp^{-1}_{x^k} x^{k+1}\rangle \geq$$

$$\geq \tfrac{1}{\mu_k}[\langle \exp^{-1}_{x^k} y^k, \exp^{-1}_{x^k} x^{k+1}\rangle + \langle \exp^{-1}_{x^{k+1}} y^k, \exp^{-1}_{x^{k+1}} x^k\rangle] + \langle e^{k+1}, \exp^{-1}_{x^{k+1}} x^k\rangle + \tfrac{\rho}{2} d^2(x^{k+1}, x^k). \quad (7)$$

Como provado em [2, Lema 3.1] usando o Teorema de Comparação de Triângulos, temos que

$$d^2(x^{k+1}, x^k) \leq \langle \exp^{-1}_{x^k} y^k, \exp^{-1}_{x^k} x^{k+1}\rangle + \langle \exp^{-1}_{x^{k+1}} y^k, \exp^{-1}_{x^{k+1}} x^k\rangle. \quad (8)$$

Usando (8) em (7), segue que

$$\langle e^{k+1}, \exp^{-1}_{x^{k+1}} x^k\rangle + (\frac{1}{\mu_k} + \frac{\rho}{2}) d^2(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}) + \langle d^k, \exp^{-1}_{x^k} x^{k+1}\rangle.$$

Sendo $\gamma_k \in [0, \frac{\rho}{2})$ e $\langle d^k, \exp^{-1}_{x^k} x^{k+1}\rangle \leq \frac{\gamma_k}{2} d^2(x^k, x^{k-1}) + \frac{\gamma_k}{2} d^2(x^{k+1}, x^k)$(veja [2, Lema 3.1]), obtemos

$$\langle e^{k+1}, \exp^{-1}_{x^{k+1}} x^k\rangle + \left(\frac{1}{\mu_k} + \frac{\rho}{2}\right) d^2(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}) + \frac{\rho}{4} d^2(x^k, x^{k-1}) + \frac{\rho}{4} d^2(x^{k+1}, x^k).$$

Pela desigualdade de Cauchy-Schwarz, temos que

$$-\|e^{k+1}\| d(x^{k+1}, x^k) + \left(\frac{1}{\mu_k} + \frac{\rho}{4}\right) d^2(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}) + \frac{\rho}{4} d^2(x^k, x^{k-1}).$$

De $\|e^{k+1}\| \leq \eta d(x^{k+1}, x^k)$ e $0 < \mu_k \leq b$, chegamos a desigualdade procurada

$$\left(\frac{1-\eta b}{b} + \frac{\rho}{4}\right) d^2(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}) + \frac{\rho}{4} d^2(x^k, x^{k-1}).$$

□

Proposição 3.3. *Suponha que a hipótese **(H)** é satisfeita. Então* $\lim_{k\to\infty} d(x^{k+1}, x^k) = 0$.

Demonstração. Pelo Proposição 3.2, temos

$$\begin{aligned}\left(\frac{1-\eta b}{b}\right) \sum_{k=0}^{n} d^2(x^{k+1}, x^k) &\leq \sum_{k=0}^{n} \left[\left(f(x^k) + \frac{\rho}{4} d^2(x^k, x^{k-1})\right) - \left(f(x^{k+1}) + \frac{\rho}{4} d^2(x^{k+1}, x^k)\right)\right] \\ &= f(x^0) - \left(f(x^{n+1}) + \tfrac{\rho}{4} d^2(x^{n+1}, x^n)\right).\end{aligned}$$

Sendo $\frac{\rho}{4} d^2(x^{n+1}, x^n) \geq 0$, segue que

$$\left(\frac{1-\eta b}{b}\right) \sum_{k=0}^{n} d^2(x^{k+1}, x^k) \leq f(x^0) - f(x^{n+1}).$$

Sendo f limitada inferiormente, obtemos

$$\left(\frac{1-\eta b}{b}\right) \sum_{k=0}^{\infty} d^2(x^{k+1}, x^k) \leq f(x^0) - \lim_{n\to\infty} f(x^{n+1}) \leq f(x^0) - \inf_{x\in M} f(x) < +\infty.$$

Logo, $\lim_{k\to\infty} d(x^{k+1}, x^k) = 0$, pois $\dfrac{1-\eta b}{b} > 0$.

□

Observe que no Algoritmo 1, pela Proposição 2.2, se $\{x^k\}$ é limitada, então $\{w^k\}$ também é limitada. Como a aplicação exponencial é um difeomorfismo, também temos que $\{d^k\}$ e $\{y^k\}$ são limitadas.

Teorema 3.1. *Suponha que a hipótese **(H)** é satisfeita e $\{x^k\}$ é limitada. Então todo ponto de acumulação de $\{x^k\}$ é ponto crítico de f.*

Demonstração. Sejam x e y pontos de acumulação de $\{x^k\}$ e $\{y^k\}$, respectivamente. Então, considere duas subsequências $\{x^{k_j}\}$ e $\{y^{k_l}\}$ convergindo respectivamente para x e y, isto é, $x^{k_j} \to x$ e $y^{k_l} \to y$. Como $\{\mu_{k_j}\}$ é limitada, podemos assumir sem perda de generalidade que $\mu_{k_j} \to \mu$.

Segue da Proposição 3.3 que $x^{k_j+1} \to x$, $x^{k_j-1} \to x$ e $e^{k_j+1} \to 0$. Por (3), temos que

$$\frac{1}{\mu_{k_j}} \exp^{-1}_{x^{k_j}} y^{k_j} - d^{k_j} = \text{grad } h(x^{k_j}).$$

Pela Proposição 2.1 e sendo h continuamente diferenciável, obtemos

$$\frac{1}{\mu} \exp^{-1}_{x} y = \text{grad } h(x). \tag{9}$$

Analogamente, por (4), temos que

$$e^{k_j+1} + \frac{1}{\mu_{k_j}} \exp^{-1}_{x^{k_j+1}} y^{k_j} = \text{grad } g(x^{k_j+1}).$$

Usando a Proposição 2.1 e o fato que g é continuamente diferenciável, segue que

$$\frac{1}{\mu} \exp^{-1}_{x} y = \text{grad } g(x). \tag{10}$$

Sendo grad $f(x) =$ grad $g(x) -$ grad $h(x)$, segue de (9) e (10) que

$$\text{grad } f(x) = 0.$$

Portanto, x é ponto crítico de f. □

4 Considerações Finais

Neste artigo, consideramos um método de ponto proximal inercial inexato para calcular pontos críticos das funções DC (suaves) em variedades de Hadamard. Futuramente iremos analisar a eficácia computacional do método proposto com outros métodos. Além disso, iremos considerar uma versão não diferenciável do método proposto.

Referências

[1] Y. T. Almeida et al. "A modified proximal point method for DC functions on Hadamard manifolds". Em: **Computational Optimization and Applications** 76.3 (2020), pp. 649–673. DOI: 10.1007/s10589-020-00173-3..

[2] J. S. Andrade, J. O. Lopes e J. C. O. Souza. "Um algoritmo inercial para funções DC em variedades de Hadamard". Em: **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics** 8.1 (2021). DOI: 10.5540/03.2021.008.01.0488.

DOI: 10.5540/03.2022.009.01.0315

[3] F. A. J. Artacho, R. M. T. Fleming e P. T. Vuong. "Accelerating the DC algorithm for smooth functions". Em: **Mathematical Programming** 169.1 (2018), pp. 95–118. DOI: 10.1007/s10107-017-1180-1.

[4] F. A. J. Artacho e P. T. Vuong. "The boosted difference of convex functions algorithm for nonsmooth functions". Em: **SIAM Journal on Optimization** 30.1 (2020), pp. 980–1006. DOI: 10.1137/18M123339X.

[5] E. E. A. Batista, G. C. Bento e O. P. Ferreira. "An extragradient-type algorithm for variational inequality on Hadamard manifolds". Em: **ESAIM: Control, Optimisation and Calculus of Variations** 26 (2020), p. 63. DOI: 10.1051/cocv/2019040.

[6] G. C. Bento e J. G. Melo. "Subgradient method for convex feasibility on Riemannian manifolds". Em: **Journal of Optimization Theory and Applications** 152.3 (2012), pp. 773–785. DOI: doi.org/10.1007/s10957-011-9921-4.

[7] J. X. Cruz Neto, O. P. Ferreira e L. R. L. Pérez. "Contributions to the study of monotone vector fields". Em: **Acta Mathematica Hungarica** 94.4 (2002), pp. 307–320. DOI: 10.1023/A:1015643612729.

[8] W. Oliveira e M. P. Tcheou. "An inertial algorithm for DC programming". Em: **Set-Valued and Variational Analysis** 27.4 (2019), pp. 895–919. DOI: 10.1007/s11228-018-0497-0.

[9] R. T. Rockafellar. "Monotone operators and the proximal point algorithm". Em: **SIAM journal on control and optimization** 14.5 (1976), pp. 877–898. DOI: doi.org/10.1137/0314056.

[10] T. Sakai. **Riemannian geometry**. Vol. 149. American Mathematical Soc., 1996. ISBN: 0821802844.

[11] M. V. Solodov e B. F. Svaiter. "A hybrid approximate extragradient–proximal point algorithm using the enlargement of a maximal monotone operator". Em: **Set-Valued Analysis** 7.4 (1999), pp. 323–345. DOI: 10.1023/A:1008777829180.

[12] J. C. O. Souza e P.R. Oliveira. "A proximal point algorithm for DC fuctions on Hadamard manifolds". Em: **Journal of Global Optimization** 63.4 (2015), pp. 797–810. DOI: 10.1007/s10898-015-0282-7.

[13] W. Sun, R. J. B. Sampaio e M. A. B. Candido. "Proximal point algorithm for minimization of DC function". Em: **Journal of computational Mathematics** (2003), pp. 451–462.

[14] G. Tang e N. Huang. "An inexact proximal point algorithm for maximal monotone vector fields on Hadamard manifolds". Em: **Operations Research Letters** 41.6 (2013), pp. 586–591. DOI: doi.org/10.1016/j.orl.2013.08.003.

[15] C. Udriste. **Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and its Applications**. Vol. 297. Dordrecht, Netherlands: Kluwer Academic Publishers, 1994. ISBN: 9789048144402.

[16] J. Wang et al. "Convergence analysis of inexact proximal point algorithms on Hadamard manifolds". Em: **Journal of Global Optimization** 61.3 (2015), pp. 553–573. DOI: 10.1007/s10898-014-0182-2.

