Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 8, n. 1, 2021.

Trabalho apresentado no XL CNMAC, Evento Virtual - Co-organizado pela Universidade do Mato Grosso do Sul (UFMS).

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Um algoritmo para funções DC em variedades de Hadamard usando aproximação de Yosida

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

Resumo. Baseando-se na aproximação de Yosida, apresentamos um algoritmo para encontrar pontos críticos de funções DC em variedades de Hadamard. Mostramos que cada ponto de acumulação da sequência gerada por nosso algoritmo é ponto crítico da função objetivo.

Palavras-chave. Resolvente, funções DC, Variedades de Hadamard, Aproximação de Yosida.

1 Introdução

Nesse artigo, nosso interesse está focado no método do ponto proximal para encontrar uma solução (pontos críticos) de problemas de otimização envolvendo uma classe especial de funções não convexas que podem ser escritas como a diferença de duas funções convexas em variedades Hadamard. Denotamos este problema por

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

onde $g, h : M \longrightarrow \mathbb{R}$ são funções convexas continuamente diferenciáveis e f limitada inferiormente.

A função objetivo f é chamada de função DC, isto é, diferença de duas funções convexas. 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 exemplos [4, 6, 7, 9, 13], e recentemente Souza e Oliveira [12] e Almeida et al. [1] propuseram algoritmos no contexto Riemanniano.

O objetivo deste artigo é apresentar um algoritmo para encontrar pontos críticos de funções DC, o qual estende ao ambiente das variedades de Hadamard um algoritmo proposto por Moudafi [9] baseado no resolvente e na aproximação de Yosida dos campos gradientes das funções g e h.

2 Preliminares

Ao longo deste artigo, M denota uma variedade Riemanniana $n-$dimensional completa, simplesmente conexa de curvatura seccional não positiva (variedade de Hadamard), com a função

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

distância $d : M \times M \to \mathbb{R}$. Os resultados preliminares e conceitos de geometria Riemanniana usados ao longo deste artigo podem ser encontrados em [8, 11, 12].

Para qualquer $x \in M$ podemos definir a inversa da aplicação exponencial $\exp_x^{-1} : M \longrightarrow T_xM$ que é C^∞. Como $d(x, x') = \| \exp_{x'}^{-1} x\|$, a aplicação $\rho_{x'} : M \longrightarrow \mathbb{R}$ definida por $\rho_{x'}(x) = \frac{1}{2}d^2(x, x')$ é C^∞ e seu gradiente em x é $grad\ \rho_{x'}(x) = -\exp_{x'}^{-1} x$. Lembre-se de 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 e geodésicas mínimas $\gamma_i : [0, l_i] \longrightarrow M$ ligando esses três pontos.

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 $\gamma_i : [0, l_i] \longrightarrow M$ a geodésica ligando p_i a p_{i+1}, e $l_i = L(\gamma_i)$ 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

$$d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2\langle \exp_{p_{i+1}}^{-1} p_i, \exp_{P_{i+1}}^{-1} p_{i+2}\rangle \leq d^2(p_{i-1}, p_i) \tag{3}$$

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

Usando as propriedades do transporte paralelo e da aplicação exponencial, obtemos o seguinte resultado que será usado com frequência nas próximas seções.

Proposição 2.1. *Seja $x_0 \in M$ e $\{x_n\} \subset M$ tal que $x_n \to x_0$. Então,*
(i) Para todo $y \in M$, temos $\exp_{x_n}^{-1} y \to \exp_{x_0}^{-1} y$ e $\exp_y^{-1} x_n \to \exp_y^{-1} x_0$.
(ii) Se $\upsilon_n \in T_{x_n}M$ e $\upsilon_n \to \upsilon_0$, então $\upsilon_0 \in T_{x_0}M$.
(iii) Dados, $u_n, \upsilon_n \in T_{x_n}M$ e $u_0, \upsilon_0 \in T_{x_0}M$, se $u_n \to u_0$ e $\upsilon_n \to \upsilon_0$, então $\langle u_n, \upsilon_n\rangle \to \langle u_0, \upsilon_0\rangle$.
(iv) A função $F : M \longrightarrow TM$ definida por $F(x) = P_{x,x_0}u$ é contínua em M, $\forall x \in M, u \in T_{x_0}M$.

Seja $\mathfrak{X}(M)$ o conjunto de todos os campos vetoriais $A : M \longrightarrow 2^{TM}$ de valores múltiplos, de modo que $A(x) \subset T_xM$ para cada $x \in M$ e o domínio $D(A)$ de A é $D(A) = \{x \in M;\ A(x) \neq \emptyset\}$.

Definição 2.1. *Dizemos que $A \in \mathfrak{X}(M)$ é monótono se a seguinte condição é válida*

$$\langle u, \exp_x^{-1} y\rangle \leq \langle \upsilon, -\exp_y^{-1} x\rangle,\ \forall u \in A(x) \text{ e } \forall \upsilon \in A(y).$$

A função $f : M \longrightarrow \mathbb{R}$ é convexa se a composição $f \circ \gamma : [a, b] \longrightarrow \mathbb{R}$ é convexa, isto é, $(f\circ\gamma)((1-t)a+tb) \leq (1-t)(f\circ\gamma)(a)+t(f\circ\gamma)(b)$, para todo segmento de geodésica $\gamma : [a, b] \longrightarrow M$ $\forall a, b \in \mathbb{R}$ e $0 \leq t \leq 1$. O subdiferencial de uma função convexa $f : M \longrightarrow \mathbb{R}$ em x é definida por

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

Além disso, o subdiferencial de f é um campo monótono. Em particular, se f é diferenciável, então $\partial f(x) = \{grad\ f(x)\}$ para todo $x \in M$ e $f : M \longrightarrow \mathbb{R}$ é convexa se, e somente se, $f(\exp_p v) \geq f(p) + \langle grad\ f(p), v\rangle$, $p \in M$ e $v \in T_pM$ (ver por exemplo [2]).

Definição 2.2. *Seja $C \subset M$ um conjunto convexo e fechado. Uma aplicação $T : C \longrightarrow M$*
(i) é não expansiva se $d(Tx, Ty) \leq d(x, y)$, $\forall x, y \in C$.
(ii) é firmemente não expansiva se a aplicação $\phi : [0, 1] \longrightarrow [0, \infty)$ definida por

$$\phi(t) := d(\exp_x t \exp_x^{-1} Tx, \exp_y t \exp_y^{-1} Ty),\ \forall t \in [0, 1]$$

é não crescente para quaisquer $x, y \in C$. Em particular T é não expansiva.

Agora vamos definir o resolvente e aproximação de Yosida de um campo vetorial $A \in \mathfrak{X}(M)$.

Definição 2.3. *Dados $\lambda > 0$ e $A \in \mathfrak{X}(M)$, o resolvente e a aproximação de Yosida de A de ordem λ são as aplicações de múltiplos valores $J_\lambda^A : M \longrightarrow 2^M$ e $A_\lambda : M \longrightarrow 2^{TM}$ definidas respectivamente por $J_\lambda^A(x) := \{z \in M; x \in \exp_z \lambda A(z)\}$ e $A_\lambda(x) = -\frac{1}{\lambda} \exp_x^{-1} J_\lambda^A(x),\ \forall x \in M$.*

O seguinte resultado relaciona a não expansividade de J_λ^A e a monoticidade de $A \in \mathfrak{X}(M)$ (ver por exemplo [8]).

Proposição 2.2. *Seja $A \in \mathfrak{X}(M)$. Então, para todo $\lambda > 0$, o campo vetorial A é monótono se, e somente se, J_λ^A é de valor único e firmemente não expansivo.*

A seguinte proposição descreve o comportamento do resolvente J_λ^A quando $\lambda \to 0$ (ver por exemplo [8]).

Proposição 2.3. *Seja $A \in \mathfrak{X}(M)$ monótono. Então, para todo $\lambda > 0$,*
(i) para todo $x \in D(A_\lambda)$ $A_\lambda(x) \in P_{x,J_\lambda^A(x)} A(J_\lambda^A(x))$;
(ii) se $x \in D(J_\lambda^A) \cap D(A)$, então $\lim\limits_{\lambda \to 0} J_\lambda^A(x) = x$.

Consideremos os seguintes lemas técnicos, os quais podem ser encontrados, respectivamente em [2, 3] e [10].

Lema 2.1. *Se $f : M \longrightarrow \mathbb{R}$ é uma função diferenciável tal que seu campo gradiente é Lipschitz contínuo com constante $L_1 \geq 0$, então $f(\exp_p v) \leq f(p) + \langle grad\ f(p), v\rangle + \frac{L_1}{2}\|v\|^2,\ \forall p \in M, v \in T_pM$. Além disso, se f for Lipschitz contínua com constante $L_2 \geq 0$, então $\|grad\ f(x)\| \leq L_2$ para todo $x \in M$.*

Lema 2.2. *Sejam $\{a_k\}$ e $\{\Gamma_k\}$ sequências de números reais não negativos tais que $a_{k+1} \leq a_k + \Gamma_k$. Se $\sum_{k=0}^{\infty} \Gamma_k < +\infty$, então $\{a_k\}$ é convergente.*

3 Algoritmo para funções DC usando aproximação de Yosida

Seja $f(x) = g(x) - h(x)$, onde g e h são funções convexas continuamente diferenciáveis e f limitada inferiormente. Apresentamos um algoritmo para encontrar pontos críticos de f, baseando-se na aproximação de Yosida do campo *grad h*. Para facilitar a notação vamos denotar o gradiente da função φ por $grad\ \varphi(x) := \varphi'(x)$. O algoritmo é dado pelo seguinte processo iterativo.

Algoritmo APPY:
Dado um ponto inicial $x^0 \in M$, e as sequências $\{\mu_k\}, \{\lambda_k\} \subset (0, +\infty)$. A sequência $\{x^k\}$ é dada da seguinte maneira: dado $x^k \in M$,

1. Calcular $y^k = \exp_{x^k}(\mu_k h'_{\lambda_k}(x^k))$, onde h'_{λ_k} é a aproximação de Yosida do campo h'.

2. Encontrar $x^{k+1} = \arg\min\limits_{x \in M}\{g(x) + \dfrac{1}{2\mu_k} d^2(x, y^k)\}$.

A boa definição das sequências $\{y^k\}$ e $\{x^k\}$ segue da Proposição 2.2 e do fato que a função $g(\cdot) + \frac{1}{2\mu_k} d^2(\cdot, y^k) : M \longrightarrow \mathbb{R}$ é estritamente convexa e 1-coerciva (ver Lema 4.1 em [5]).

Na seguinte proposição vamos assumir que a função h e seu gradiente são Lipschitz contínuos de mesma constante L, do contrário podemos tomar L o máximo das constantes de Lipschitz.

DOI: 10.5540/03.2021.008.01.0489

Proposição 3.1. *Seja* $\{x^k\}$ *a sequência gerada pelo algoritmo* ***APPY****. Se* h *e seu gradiente são Lipschitz contínuos de constante* L*, então*

$$f(x^{k+1}) \leq f(x^k) + \frac{L^3}{2}\lambda_k^2 - \frac{1}{\mu_k} d^2(x^{k+1}, x^k). \tag{4}$$

Prova: De fato, usando o Algoritmo **APPY**, temos que $\frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k = g'(x^{k+1})$. Então, pela convexidade de g, segue que

$$g(x^k) \geq g(x^{k+1}) + \left\langle \frac{1}{\mu_k} \exp^{-1}_{x^{k+1}} y^k, \exp^{-1}_{x^{k+1}} x^k \right\rangle. \tag{5}$$

Pela Proposição 2.3, temos que $\frac{1}{\lambda_k} \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^k = h'(J^{h'}_{\lambda_k}(x^k))$. Segue da convexidade de h

$$\begin{aligned} h(x^{k+1}) &\geq h(J^{h'}_{\lambda_k}(x^k)) + \left\langle \tfrac{1}{\lambda_k} \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^k, \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^{k+1} \right\rangle \\ &= h(x^k) + h(J^{h'}_{\lambda_k}(x^k)) - h(x^k) + \left\langle \tfrac{1}{\lambda_k} \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^k, \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^{k+1} \right\rangle. \end{aligned} \tag{6}$$

Somando as desigualdade (5) e (6), obtemos

$$\begin{aligned} f(x^{k+1}) &\leq f(x^k) + h(x^k) - h(J^{h'}_{\lambda_k}(x^k)) - \\ &- \tfrac{1}{\lambda_k} \left\langle \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^k, \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^{k+1} \right\rangle - \tfrac{1}{\mu_k} \left\langle \exp^{-1}_{x^{k+1}} y^k, \exp^{-1}_{x^{k+1}} x^k \right\rangle. \end{aligned}$$

Usando o Teorema de comparação para triângulos para $\triangle(y^k, x^k, x^{k+1})$ e $\triangle(y^k, x^{k+1}, x^k)$, onde $y^k = \exp_{x^k}(\mu_k h'_{\lambda_k}(x^k))$, obtemos

$$\begin{aligned} f(x^{k+1}) &\leq f(x^k) + h(x^k) - h(J^{h'}_{\lambda_k}(x^k)) - \tfrac{1}{\mu_k} d^2(x^{k+1}, x^k) \\ &- \tfrac{1}{\lambda_k} \left[\left\langle \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^k, \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^{k+1} \right\rangle + \left\langle \exp^{-1}_{x^k} J^{h'}_{\lambda_k}(x^k), \exp^{-1}_{x^k} x^{k+1} \right\rangle \right]. \end{aligned} \tag{7}$$

Considere agora os triângulos geodésicos $\triangle(x^k, J^{h'}_{\lambda_k}(x^k), x^{k+1})$ e $\triangle(J^{h'}_{\lambda_k}(x^k), x^{k+1}, x^k)$, então pelo Teorema de comparação para triângulos, temos de (7) que

$$f(x^{k+1}) \leq f(x^k) + h(x^k) - h(J^{h'}_{\lambda_k}(x^k)) - \tfrac{1}{\lambda_k} d^2(x^k, J^{h'}_{\lambda_k}(x^k)) - \tfrac{1}{\mu_k} d^2(x^{k+1}, x^k). \tag{8}$$

Usando o Lema 2.1 para $p = J^{h'}_{\lambda_k}(x^k)$, $v = \exp^{-1}_{J^{h'}_{\lambda_k}(x^k)} x^k$, temos que

$$h(x^k) - h(J^{h'}_{\lambda_k}(x^k)) - \frac{1}{\lambda_k} d^2(x^k, J^{h'}_{\lambda_k}(x^k)) \leq \frac{L^3}{2}\lambda_k^2. \tag{9}$$

Combinando (8) e (9), obtemos $f(x^{k+1}) \leq f(x^k) + \frac{L^3}{2}\lambda_k^2 - \frac{1}{\mu_k} d^2(x^{k+1}, x^k)$. □

Agora, vamos provar o resultado de convergência.

Teorema 3.1. *Seja $\{x^k\}$ a sequência gerada pelo Algoritmo **APPY**. Se h e seu gradiente são Lipschitz contínuos, $0 < \mu_1 \leq \mu_k \leq \mu_2$ e $\sum_{k=0}^{\infty} \lambda_k < \infty$, então*
(i) $\{f(x^k) - f^*\}$ é convergente, onde $f^* = \inf\{f(y); y \in M\}$.
(ii) $\lim_{k\to\infty} d(x^{k+1}, x^k) = 0$ e $\lim_{k\to\infty} \frac{1}{\mu_k} d(x^{k+1}, x^k) = 0$.
(iii) Todo ponto de acumulação de $\{x^k\}$ é ponto crítico de f.

Prova: i) Pela Proposição 3.1, temos que

$$f(x^{k+1}) - f^* \leq f(x^k) - f^* + \frac{L^3}{2}\lambda_k^2. \tag{10}$$

Definindo $a_k = f(x^k) - f^*$ e $\Gamma_k = \frac{L^3}{2}\lambda_k^2$, como $\sum_{k=0}^{\infty} \Gamma_k < \infty$ segue de (10) e do Lema 2.2 que $\{f(x^k) - f^*\}$ é convergente.

ii) Sendo $0 < \mu_1 \leq \mu_k \leq \mu_2$, segue pela Proposição 3.1 que

$$\frac{1}{\mu_2} d^2(x^{k+1}, x^k) \leq \frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq [(f(x^k) - f^*) - (f(x^{k+1}) - f^*)] + \frac{L^3}{2}\lambda_k^2.$$

Então, $\lim_{k\to\infty} d(x^{k+1}, x^k) = 0$ e $\lim_{k\to\infty} \frac{1}{\mu_k} d(x^{k+1}, x^k) = 0$.

iii) Seja $\bar{x}$ um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ subsequência tal que $x^{k_j} \to \bar{x}$. Segue de (ii) que $x^{k_j+1} \to \bar{x}$.

Usando a desigualdade triangular e a não expansividade de $J^{h'}_{\lambda_{k_j}}$, temos que

$$d(J^{h'}_{\lambda_{k_j}}(x^{k_j}), \bar{x}) \leq d(J^{h'}_{\lambda_{k_j}}(x^{k_j}), J^{h'}_{\lambda_{k_j}}(\bar{x})) + d(J^{h'}_{\lambda_{k_j}}(\bar{x}), \bar{x}) \leq d(x^{k_j}, \bar{x}) + d(J^{h'}_{\lambda_{k_j}}(\bar{x}), \bar{x}).$$

Como $\lim_{j\to\infty} J^{h'}_{\lambda_{k_j}}(\bar{x}) = \bar{x}$, pela Proposição 2.3, obtemos $\lim_{j\to\infty} J^{h'}_{\lambda_{k_j}}(x^{k_j}) = \bar{x}$. Assim, $\lim_{j\to\infty} h'(J^{h'}_{\lambda_{k_j}}(x^{k_j})) = h'(\bar{x}) = \bar{w}$.

Resta provar que $\lim_{j\to\infty} g'(x^{k_j+1}) = \bar{w}$, onde $g'(x^{k_j+1}) = \frac{1}{\mu_{k_j}} \exp^{-1}_{x^{k_j+1}} y^{k_j}$.

Afirmação: $\lim_{k\to\infty} \frac{1}{\mu_{k_j}} \exp^{-1}_{x^{k_j+1}} y^{k_j} = \bar{w}$. De fato, para $w^{k_j} = h'(x^{k_j})$

$$\|\frac{1}{\mu_{k_j}} \exp^{-1}_{x^{k_j+1}} y^{k_j} - P_{x^{k_j+1},\bar{x}}\bar{w}\| \leq \|\frac{1}{\mu_{k_j}} \exp^{-1}_{x^{k_j+1}} y^{k_j} - P_{x^{k_j+1},x^{k_j}} w^{k_j}\| + \|P_{x^{k_j+1},x^{k_j}} w^{k_j} - P_{x^{k_j+1},\bar{x}}\bar{w}\|.$$

1) Note que, $P_{x^{k_j+1},\bar{x}}\bar{w} = (P_{x^{k_j+1},x^{k_j}} \circ P_{x^{k_j},\bar{x}})\bar{w}$. Segue do fato que a aplicação transporte paralelo é uma isometria, que

$$\|P_{x^{k_j+1},x^{k_j}} w^{k_j} - P_{x^{k_j+1},\bar{x}}\bar{w}\| = \|P_{x^{k_j+1},x^{k_j}} w^{k_j} - P_{x^{k_j+1},x^{k_j}}(P_{x^{k_j},\bar{x}}\bar{w})\| = \|w^{k_j} - P_{x^{k_j},\bar{x}}\bar{w}\|.$$

Assim,

$$\lim_{j\to\infty} \|P_{x^{k_j+1},x^{k_j}} w^{k_j} - P_{x^{k_j+1},\bar{x}}\bar{w}\| = 0. \tag{11}$$

2) Veja agora que

$$\begin{aligned}
&\|\tfrac{1}{\mu_{k_j}}\exp^{-1}_{x^{k_j+1}} y^{k_j} - P_{x^{k_j+1},x^{k_j}} w^{k_j}\|^2 = \\
&= \tfrac{1}{\mu^2_{k_j}}\|\exp^{-1}_{x^{k_j+1}} y^{k_j}\|^2 + \tfrac{1}{\lambda^2_{k_j}}\|\exp^{-1}_{x^{k_j}} y^{k_j}\|^2 - 2\tfrac{1}{\mu_{k_j}}\left\langle \exp^{-1}_{x^{k_j+1}} y^{k_j}, P_{x^{k_j+1},x^{k_j}} w^{k_j}\right\rangle \qquad (12)\\
&= \tfrac{1}{\mu^2_{k_j}}\left(d^2(x^{k_j+1}, y^{k_j}) + d^2(y^{k_j}, x^{k_j})\right) - 2\tfrac{1}{\mu_{k_j}}\left\langle \exp^{-1}_{x^{k_j+1}} y^{k_j}, P_{x^{k_j+1},x^{k_j}} w^{k_j}\right\rangle.
\end{aligned}$$

Para o triângulo geodésico $\triangle(x^{k_j+1}y^{k_j}x^{k_j})$, temos

$$\frac{1}{\mu^2_{k_j}}[d^2(x^{k_j+1}, y^{k_j}) + d^2(y^{k_j}, x^{k_j})] - 2\frac{1}{\mu^2_{k_j}}\langle \exp^{-1}_{y^{k_j}} x^{k_j+1}, \exp^{-1}_{y^k_j} x^{k_j}\rangle \leq \frac{1}{\mu^2_{k_j}} d^2(x^{k_j+1}, x^{k_j}). \qquad (13)$$

Note que,
(I)

$$\begin{aligned}
\langle \exp^{-1}_{y^{k_j}} x^{k_j+1}, \exp^{-1}_{y^{k_j}} x^{k_j}\rangle &= \langle P_{x^{k_j+1},y^{k_j}} \exp^{-1}_{y^{k_j}} x^{k_j+1}, P_{x^{k_j+1},y^{k_j}} \exp^{-1}_{y^{k_j}x^{k_j}}\rangle \\
&= \langle -\exp^{-1}_{x^{k_j+1}} y^{k_j}, P_{x^{k_j+1}}, y^{k_j} \exp^{-1}_{y^{k_j}} x^{k_j}\rangle.
\end{aligned}$$

(II) Sendo $P_{x^{k_j+1},y^{k_j}} = P_{x^{k_j+1},x^{k_j}} \circ P_{x^{k_j},y^{k_j}}$, temos

$$\begin{aligned}
P_{x^{k_j+1},y^{k_j}}(\exp^{-1}_{y^{k_j}} x^{k_j}) &= (P_{x^{k_j+1},x^{k_j}} \circ P_{x^{k_j},y^{k_j}})(\exp^{-1}_{y^{k_j}} x^{k_j}) \\
&= P_{x^{k_j+1},x^{k_j}}(P_{x^{k_j},y^{k_j}}(\exp^{-1}_{y^{k_j}} x^{k_j})) \\
&= P_{x^{k_j+1},x^{k_J}}(-\exp^{-1}_{x^{k_J}} y^{k_j}) \\
&= -\mu_{k_j} P_{x^{k_j+1},x^{k_j}}(\tfrac{1}{\mu_{k_j}}\exp^{-1}_{x^{k_j}} y^{k_j}) = -\mu_{k_j} P_{x^{k_j+1},x^{k_j}} w^{k_j}.
\end{aligned}$$

Segue de (I) e (II) que $\langle \exp^{-1}_{y^{k_j}} x^{k_j+1}, \exp^{-1}_{y^{k_j}} x^{k_j}\rangle = \mu_{k_j}\langle \exp^{-1}_{x^{k_j+1}} y^{k_j}, P_{x^{k_j+1},x^{k_j}} w^{k_j}\rangle$. Usando essa igualdade em (13), obtemos

$$\frac{1}{\mu^2_{k_j}}[d^2(x^{k_j+1}, y^{k_j}) + d^2(y^{k_j}, x^{k_j})] - 2\frac{1}{\mu_{k_j}}\langle \exp^{-1}_{x^{k_j+1}} y^{k_j}, P_{x^{k_j+1},x^{k_j}} w^{k_j}\rangle \leq \frac{1}{\mu^2_{k_j}} d^2(x^{k_j+1}, x^{k_j}). \qquad (14)$$

Segue de (12) e (14), que $\|\frac{1}{\mu_{k_j}}\exp^{-1}_{x^{k_j+1}} y^{k_j} - P_{x^{k_j+1},x^{k_j}} w^{k_j}\|^2 \leq [\frac{1}{\mu_{k_j}} d(x^{k_j+1}, x^{k_j})]^2$. Sendo que $\lim_{j\to\infty} \dfrac{1}{\mu_{k_j}} d(x^{k_j+1}, x^{k_j}) = 0$, obtemos

$$\lim_{j\to\infty} \|\frac{1}{\mu_{k_j}}\exp^{-1}_{x^{k_j+1}} y^{k_j} - P_{x^{k_j+1},x^{k_j}} w^{k_j}\|^2 = 0. \qquad (15)$$

Por (11) e (15), pela continuidade da norma e da aplicação transporte paralelo $\lim_{j\to\infty} \dfrac{1}{\mu_{k_j}}\exp^{-1}_{x^{k_j+1}} y^{k_j} = \bar{w}$. Assim, $\lim_{j\to\infty} g'(x^{k_j+1}) = g'(\bar{x}) = \bar{w}$.

Portanto, $f'(\bar{x}) = g'(\bar{x}) - h'(\bar{x}) = 0$, ou seja, x é ponto crítico de f. $\square$

DOI: 10.5540/03.2021.008.01.0489

4 Conclusões

Neste artigo, consideramos um método para calcular pontos críticos das funções DC (suaves) em variedades de Hadamard, baseando-se na aproximação de Yosida do campo gradiente de função convexa. Tal algoritmo estende para o ambiente de variedades Riemannianas o algoritmo proposoto por Moudafi [9] em espaços de Hilbert. Foi provado que todo ponto de acumulação da sequência é ponto crítico da função DC f. Como um trabalho futuro, uma generalização do método para diferença de campos monótonos pode ser considerado.

Referências

[1] Almeida, Y.T., Cruz Neto, J.X., Oliveira,P.R. and Souza, J.C.O. A modified proximal point method for DC functions on Hadamard manifolds, *Computational Optimization and Applications*, 76:649–673, 2020. DOI: 10.1007/s10589-020-00173-3.

[2] Bento, G.C., Ferreira, O.P. and Melo. J.G. Iteration-Complexity of Gradient, Subgradient and Proximal Point Methods on Riemannian Manifolds, *Journal of Optimization Theory and Applications*, 173:548–562, 2017. DOI: 10.1007/s10957-017-1093-4.

[3] Boumal, N. *An introduction to optimization on smooth manifolds*, 2020. https://www.epfl.ch/schools/sb/education/sma/

[4] Elhilali Alaoui, A. Caractrisation des fonctions D.C. (Characterization of D. C. functions), *Ann. Sci. Math. Quebec*, 20: 1-13, 1996.

[5] Ferreira, O.P. and Oliveira, P.R. Proximal point algorithm on Riemannian manifolds, *Optimization*, 51:257-270, 2002. DOI:10.1080/02331930290019413.

[6] Hiriart-Urruty, J.B. Generalized differentiabity, duality and optimization for problems dealing with difference of convex functions, *Convexity and Duality in Optimization*, 37-70, 1985. DOI:10.1007/978-3-642-45610-7_3.

[7] Hiriart-Urruty, J.B. From convex optimization to nonconvex optimization. Necessary and sufficient conditions for global optimality, *Nonsmooth Optimization and Related Topics*, Springer US, chapter 13, pages 219-239, 1989.

[8] Li, C., López, G., Martín-Márquez, V. and Wang, J. H. Resolvents of Set Valued Monotone Vector Fields in Hadamard Manifolds, *Set-Valued and Variational Analysis*, 19:361–383, 2011. DOI 10.1007/s11228-010-0169-1.

[9] Moudafi, A. On critical points of the difference of two maximal monotone operators, *Afrika Matematika*, 26:457–463, 2015. DOI 10.1007/s13370-013-0218-7.

[10] Polyak, B.T. *Introduction to Optimization.* Optimization Software Inc., New York, 198.

[11] Sakai, T. *Riemannian Geometry. Translations of Mathematical Monographs, volume 149.* American Mathematical Society, Providence, 1996.

[12] Souza,J.C.O. and Oliveira,P.R. A proximal point method for DC functions on Hadamard manifolds, *Jornal of Global Optimization*, 63:797-810, 2015. DOI:10.1007/s10898-015-0282-7.

[13] Sun, W., Sampaio, R.J.B. and Candido, M.A.B. Proximal point algorithm for minimization of DC Functions, *Journal of Computational Mathematics*, 21:451-462, 2003.

DOI: 10.5540/03.2021.008.01.0489