



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO EM MATEMÁTICA

**Problema de Viabilidade Convexa em Variedades de
Hadamard**

José Márcio Machado de Brito

Teresina - 2020

José Márcio Machado de Brito

Dissertação de Mestrado:

Problema de Viabilidade Convexa em Variedades de Hadamard

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:

Prof. Dr. João Xavier da Cruz Neto

Co-Orientador:

Prof. Dr. João Carlos de Oliveira Souza

Teresina - 2020



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Problema de Viabilidade Convexa em Variedades de Hadamard

José Márcio Machado de Brito

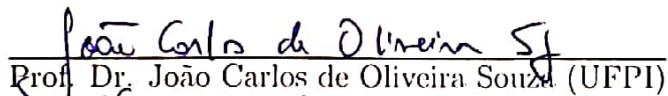
Esta Dissertação foi submetida como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**, outorgado pela Universidade Federal do Piauí.

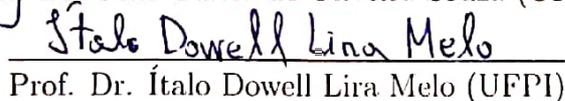
A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

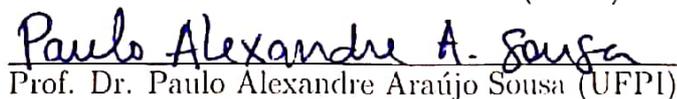
Dissertação aprovada em 20 de Fevereiro de 2020.

Banca Examinadora:


Prof. Dr. João Xavier da Cruz Neto - Orientador


Prof. Dr. João Carlos de Oliveira Souza (UFPI)


Prof. Dr. Ítalo Dowell Lira Melo (UFPI)


Prof. Dr. Paulo Alexandre Araújo Sousa (UFPI)

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

B862p Brito, José Márcio Machado de.

Problema de viabilidade convexa em variedades de Hadamard / José Márcio Machado de Brito. – Teresina, 2020.

57 f.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Matemática, 2020.

Orientador: Prof. Dr. João Xavier da Cruz Neto

Co-orientador: Prof. Dr. João Carlos de Oliveira Souza

1. Variedade Hadamard. 2. Curva Subgradiente. 3. Método do Gradiente. I. Título.

CDD 516.3

Bibliotecária : Caryne Maria da Silva Gomes / CRB3-1461

*Dedico à minha mãe Auzair, minha esposa Raynara
e aos meus irmãos Tiago, Ana e Marquim.*

Agradecimentos

A Deus, pela dádiva da vida e por colocar em meu caminho pessoas tão especiais.

A todos os meus familiares, em especial a minha mãe Maria Auzair, que em meio a tantas dificuldades sempre me apoiou e fez de tudo pela minha formação, à minha cunhada Raylla Gabrielly e aos meus irmãos Tiago Brito, João Marcos e Ana Márcia por todos os momentos de alegria que vocês me proporcionam.

A minha esposa Raynara Beatriz, por sua paciência, apoio, ajuda e compreensão em todos os momentos e por tornar meus dias mais felizes.

A todos os meus amigos pelo incentivo, força e companheirismo, em especial José Edilson, Erisvaldo Vieira, João Vinícius, Dário Severo, Renato Santos, Dieme Pereira, Jean Carlos, Thiago Mayson, Sandoel Vieira, Francimar Brito, Pedro Paulo, Pedro Rodrigues, Christopher, Igor Fontenele, Silas Augusto, Leonardo Silva, Antônio Wesley, Rafael Emanuel, Lucas Mendes, Francivaldo França, Lucas Cassiano, Fernandes Filho, Yldenilson, Severino Pereira, Ruan Diego, Kelvin Jhonsonm, Leandro Bittencourt, Luan Soares, Cícero Nadiel, Paulo Sérgio, João Santos e a todos os amigos que fiz na UESPI, UFPI e IMPA nos últimos anos. Suas amizades foram essências para o meu crescimento.

Meus sinceros agradecimentos ao meu orientador João Xavier que acompanha minha trajetória desde o ensino fundamental, quando eu participava das olimpíadas de matemática e me orientou durante toda a graduação e mestrado, e sempre que precisei estava disposto a conversar por horas, dando valiosos conselhos.

Aos professores e amigos Antônio Amaral (Amaral) e Raimundo Brito (Raimundim) por terem me apresentado as olimpíadas de matemática, por todos os anos de preparações olímpicas e por sempre estarem dispostos a ajudar.

Aos Professores do departamento de matemática da UFPI, em especial, aos professores Jurandir Oliveira, Vitaliano Amaral, João Carlos, Ítalo Dowell, Paulo Alexandre, Barnabé

Pessoa Lima, Marcos Travaglia e José Francisco pelo apoio, amizade e ensinamentos e por terem me acolhido nesta instituição. Também agradeço aos professores da UESPI, principalmente a José Arimatéa e Pedro Soares pela amizade e apoio durante toda a graduação.

Agradeço ao CNPq pelo apoio financeiro.

*“Lutar sempre, vencer talvez, desistir
jamais”.*

Maxwellton

Resumo

Neste trabalho mostramos a existência e unicidade de curvas subgradiente em variedades de Hadamard e apresentamos suas principais propriedades. A partir dessas propriedades, provamos a equivalência entre os conceitos de error bounds com comportamento moderado e a desigualdade de Kurdyka - Lojasiewicz. Como aplicação, usamos o método do gradiente para resolver um problema de viabilidade convexa em variedades de Hadamard.

Palavras-Chaves: Variedades de Hadamard, Curva Subgradiente, Desigualdade KL, Error Bounds, Método do Gradiente, Minimização Convexa.

Abstract

In this work, we show the existence and uniqueness of subgradient curves in Hadamard manifolds and present its main properties. Therefore, we prove the equivalence between the concepts of error bounds with moderate growth and Kurdyka-Lojasiewicz inequality. As application, we use the gradient method to solve a convex feasibility problem in Hadamard manifolds.

Keywords: Hadamard manifolds; subgradient curve; KL inequality; error bounds; gradient method; convex minimization.

Sumário

Resumo	v
Abstract	vi
1 Noções Preliminares	3
1.1 Variedades Riemannianas	3
1.2 Curvaturas	5
1.3 Variedades de Hadamard	7
1.4 Propriedades geométricas de uma Variedade de Hadamard	9
1.5 Função distância	12
2 Noções de Análise Convexa em Variedades de Hadamard	14
2.1 Conjuntos Convexos	14
2.2 Funções Convexas	16
2.3 Funções Coercivas	20
2.4 Resolvente de uma função convexa	22
2.5 Fluxo gradiente	23
3 Curva subgradiente e suas propriedades	27
3.1 Existência da curva subgradiente	27
3.2 Propriedades	28
4 Error bounds e desigualdade de Kurdyka - Łojasiewicz	36
4.1 A desigualdade de Kurdyka - Łojasiewicz	36
4.2 Error Bounds	38

Sumário	viii
<hr/>	
5 Problema de viabilidade convexa	43
5.1 Método do gradiente	43
5.2 Complexidade do método	46
Apêndice	47
A Análise de convergência do método do gradiente	48
A.1 Método do gradiente	48
A.2 Análise de convergência	51
Referências Bibliográficas	56

Introdução

Os resultados obtidos neste trabalho serão desenvolvidos em Variedades de Hadamard, ou seja, em uma variedade Riemanniana completa, simplesmente conexa e com curvatura seccional não positiva e será denotada por M .

Em muitas situações de otimização restrita, devemos minimizar uma função em um conjunto C , que é dado como uma interseção finita de conjuntos do tipo $C_i = \{x \in M; f_i(x) \leq c\}$, onde $f_i : M \rightarrow \mathbb{R}$ é uma função convexa, isto é, conjuntos convexos e fechados. Assim, antes de aplicar algum algoritmo de minimização, é necessário encontrar um ponto viável, isto é, um ponto de C . Essa tarefa é conhecida na literatura por *problema de viabilidade convexa*, mas precisamente, se C_1, C_2, \dots, C_m são conjuntos convexos e fechados tais que $C := \bigcap_{i=1}^m C_i \neq \emptyset$, estamos interessados em encontrar algum ponto $x \in C$.

Este problema foi bastante estudado nas últimas décadas e existem diversas aplicações em problemas da matemática e das ciências físicas, veja [3]. Encontra-se aplicações por exemplo em teoria de aproximação, que é frequentemente usada em estatística, equações diferenciais parciais e análise complexa. Também existem aplicações em reconstrução de imagens, utilizadas em tomografia computadorizada, processamento de sinais e microscopia eletrônica, que são essenciais nas áreas médicas.

Para resolver este problema em variedades de Hadamard, aplicaremos o método do gradiente na função $f : M \rightarrow \mathbb{R}$ dada por

$$f(x) = \frac{1}{2} \sum_{i=1}^m \alpha_i \text{dist}^2(x, C_i),$$

onde $\alpha_i > 0$ para todo $i = 1, 2, \dots, m$ e $\sum_{i=1}^m \alpha_i = 1$.

Para isso, precisamos que f satisfaça a desigualdade de Kurdyka-Lojasiewicz (ou abreviadamente, desigualdade KL). Em espaços de Hilbert, as noções de desigualdade KL e error bounds com comportamento moderado são equivalentes, como pode ser visto em [7].

A importância dessa equivalência vem do fato de que verificar a desigualdade KL é, em geral, mais difícil do que encontrar um error bounds, além disso, não costuma ser difícil obter resultados de complexidade de métodos de descida de primeira ordem a partir de um error bounds. Neste trabalho vamos generalizar a equivalência entre esses conceitos para Variedades de Hadamard.

A fim de que f tenha um error bounds local com comportamento moderado (e portanto satisfaça a desigualdade KL), vamos supor que os conjuntos C_1, C_2, \dots, C_m são linearmente limitadamente regular, isto é, que existe uma constante $\kappa > 0$ tal que

$$\text{dist}(x, C) \leq \kappa \max\{\text{dist}(x, C_i), i = 1, \dots, m\}, \forall x \in S,$$

para todo conjunto S limitado. Em espaços de Hilbert, esta é uma hipótese natural, pois se $\bigcap_{i=1}^m C_i$ tem interior não vazio, então os conjuntos C_1, C_2, \dots, C_m são linearmente regular, como mostrado em [7]. Este trabalho está organizado da seguinte forma:

No capítulo 1, apresentamos algumas definições, notações e resultados sobre geometria Riemanniana que darão suporte aos próximos capítulos.

No capítulo 2 introduzimos algumas noções de análise convexa em variedades de Hadamard, sobretudo aquelas necessárias para definir o resolvente, slope e fluxo gradiente de uma função convexa.

No capítulo 3, utilizamos o fluxo gradiente associado a uma função convexa $f : M \rightarrow \mathbb{R}$, para mostrar a existência de uma curva absolutamente contínua $x : [0, \infty) \rightarrow M$ satisfazendo $-x'(t) \in \partial f(x(t))$ para quase todo $t > 0$, isto é, que se move em uma direção de máxima descida de f , esta curva é chamada de *curva subgradiente*. Além disso, mostramos algumas propriedades interessantes sobre esta curva.

No capítulo 4, definimos a desigualdade de Kurdyka-Łojasiewicz, error bounds, comportamento moderado e usamos a existência da curva subgradiente e suas propriedades para mostrar a equivalência entre a desigualdade KL e error bounds com comportamento moderado.

No capítulo 5, usamos a relação entre a desigualdade KL e error bounds para resolver um problema de viabilidade convexa em variedades de Hadamard através do método do gradiente e obter a complexidade do método.

Capítulo 1

Noções Preliminares

Neste capítulo apresentaremos algumas definições, notações e resultados que serão utilizados nos capítulos seguintes. Não demonstraremos todos os resultados, mas deixaremos as devidas referências. Para a leitura do texto espera-se que o leitor tenha algumas noções de geometria, como a definição de Variedade Riemanniana, plano tangente, campos de vetores, geodésicas, aplicação exponencial e curvaturas. Todos esses conceitos podem ser encontrados em [14].

1.1 Variedades Riemannianas

Denotaremos por M^n (Ou simplesmente M) uma variedade Riemanniana conexa de dimensão n e classe C^∞ , $\langle \cdot, \cdot \rangle$ sua métrica Riemanniana, com a correspondente norma denotada por $\| \cdot \|$, ∇ sua conexão Riemanniana, $T_p M$ o plano tangente a M no ponto $p \in M$ e $\mathfrak{X}(M)$ o conjunto dos campos de vetores C^∞ de M . A seguir, iremos relembrar algumas definições e proposições que serão usadas no transcorrer do texto.

Definição 1.1. *Seja $\alpha : [a, b] \rightarrow M$ uma curva diferenciável por partes. O comprimento da curva α , denotado por $L(\alpha)$, é definido por*

$$L(\alpha) = \int_a^b \|\alpha'(t)\| dt.$$

Definição 1.2. *Sejam M uma variedade e $f : M \rightarrow \mathbb{R}$ uma curva de classe C^1 . O gradiente de f , denotado por $\text{grad } f$, é o único campo $\text{grad } f \in \mathfrak{X}(M)$ definido por*

$$\langle \text{grad } f(p), v \rangle = f'(p) \cdot v.$$

Proposição 1.1. *Se $f, g : M \rightarrow \mathbb{R}$ são funções diferenciáveis, então*

(a) $\text{grad}(f + g) = \text{grad} f + \text{grad} g.$

(b) $\text{grad}(fg) = g \cdot \text{grad} f + f \cdot \text{grad} g.$

Demonstração. Veja [14]. □

Proposição 1.2. *Seja $f : M \rightarrow \mathbb{R}$ uma função diferenciável. Dados $p \in M$ e $v \in T_p M$, seja $\alpha : (-\varepsilon, \varepsilon) \rightarrow M$ uma curva suave tal que $\alpha(0) = p$ e $\alpha'(0) = v$. Então*

$$\langle \text{grad} f, v \rangle_p = \left. \frac{d}{dt} (f \circ \alpha)(t) \right|_{t=0}.$$

Proposição 1.3. *Se $f : M \rightarrow \mathbb{R}$ é uma função suave e $U \subset M$ é uma vizinhança coordenada, com campos coordenados $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$, então o gradiente de f é dado por*

$$\text{grad} f = \sum_{i,j=1}^n g^{ij} \frac{\partial f}{\partial x_j} \frac{\partial}{\partial x_i}.$$

onde as funções g_{ij} são os coeficiente da métrica e $[g^{ij}]$ é a matriz inversa de $[g_{ij}]$

Em particular, quando $M = \mathbb{R}^n$ tem-se que $g_{ij} = \delta_{ij}$ são os coeficientes da métrica euclidiana e

$$\text{grad} f = \sum_{i,j=1}^n \delta^{ij} \frac{\partial f}{\partial x_j} \frac{\partial}{\partial x_i} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial}{\partial x_i}.$$

Ou seja, no espaço euclidiano, coincide com a definição de gradiente no sentido usual.

Demonstração. Veja [14]. □

Definição 1.3. *Seja $f : M \rightarrow \mathbb{R}$ uma função diferenciável. A hessiana de f no ponto p é o operador linear $\text{Hess} f_p : T_p M \rightarrow T_p M$, definido por*

$$(\text{Hess} f)_p(v) = \nabla_v \text{grad} f.$$

Segue das propriedades da conexão Riemanniana que se $X \in \mathfrak{X}(M)$ é qualquer extensão de v em uma vizinhança de $p \in M$, então

$$(\text{Hess} f)_p(v) = (\nabla_X \nabla f)(p).$$

onde ∇ é a conexão Riemanniana de M .

Proposição 1.4. *Se $f : M \rightarrow \mathbb{R}$ é uma função suave e $p \in M$, então $(\text{Hess} f)_p : T_p M \rightarrow T_p M$ é um operador linear auto-adjunto.*

Demonstração. Veja [14]. □

1.2 Curvaturas

Vamos lembrar algumas definições a respeito de curvaturas que serão utilizadas ao longo do texto, por exemplo para definir variedades de Hadamard. As definições e resultados apresentados aqui podem ser encontrados em [14].

Definição 1.4 (Curvatura). *A curvatura R de uma variedade Riemanniana M é uma correspondência que associa a cada par $X, Y \in \mathfrak{X}(M)$ uma aplicação $R(X, Y) : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$ dada por*

$$R(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X, Y]} Z, \quad Z \in \mathfrak{X}(M),$$

onde ∇ é a conexão Riemanniana de M .

Observe que se $M = \mathbb{R}^n$, então $R(X, Y)Z = 0$ para todo $X, Y, Z \in \mathfrak{X}(\mathbb{R}^n)$.

Em termos de um sistema de coordenadas (U, x_1, \dots, x_n) em torno de $p \in M$. Seja $X_i = \frac{\partial}{\partial x_i}$. Como $[X_i, X_j] = 0$, temos

$$R(X_i, X_j)X_k = \nabla_{X_j} \nabla_{X_i} X_k - \nabla_{X_i} \nabla_{X_j} X_k.$$

E ainda, $\nabla_{X_i} X_k = \sum_l \Gamma_{ik}^l X_l$ e $\nabla_{X_j} X_k = \sum_l \Gamma_{jk}^l X_l$. Daí,

$$\begin{aligned} R(X_i, X_j)X_k &= \nabla_{X_j} \nabla_{X_i} X_k - \nabla_{X_i} \nabla_{X_j} X_k \\ &= \sum_l \nabla_{X_j} \Gamma_{ik}^l X_l - \sum_l \nabla_{X_i} \Gamma_{jk}^l X_l \\ &= \sum_l \left[X_j(\Gamma_{ik}^l) X_l + \Gamma_{ik}^l \nabla_{X_j} X_l \right] - \sum_l \left[X_i(\Gamma_{jk}^l) X_l + \Gamma_{jk}^l \nabla_{X_i} X_l \right] = \\ &= \sum_l \left[X_j(\Gamma_{ik}^l) X_l + \Gamma_{ik}^l \sum_s \Gamma_{jl}^s X_s - X_i(\Gamma_{jk}^l) X_l - \Gamma_{jk}^l \sum_s \Gamma_{il}^s X_s \right] = \\ &= \sum_s \left[\frac{\partial}{\partial x_j} (\Gamma_{ik}^s) - \frac{\partial}{\partial x_i} (\Gamma_{jk}^s) + \sum_l \left(\Gamma_{ik}^l \Gamma_{jl}^s - \Gamma_{jk}^l \Gamma_{il}^s \right) \right] X_s. \end{aligned}$$

Proposição 1.5. *A curvatura R de uma variedade Riemanniana satisfaz as seguintes propriedades:*

(i) R é $C^\infty(M)$ -bilinear em $\mathfrak{X}(M) \times \mathfrak{X}(M)$, isto é,

$$R(fX_1 + gX_2, Y_1) = fR(X_1, Y_1) + gR(X_2, Y_1),$$

$$R(X_1, fY_1 + gY_2) = fR(X_1, Y_1) + gR(X_1, Y_2),$$

$f, g \in C^\infty(M)$, $X_1, X_2, Y_1, Y_2 \in \mathfrak{X}(M)$.

(ii) Para todo par $X, Y \in \mathfrak{X}(M)$, o operador curvatura $R(X, Y) : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$ é $C^\infty(M)$ -linear, isto é,

$$R(X, Y)(Z + W) = R(X, Y)Z + R(X, Y)W,$$

$$R(X, Y)(fZ) = fR(X, Y)Z,$$

$$f \in C^\infty(M), Z, W \in \mathfrak{X}(M).$$

(iii) Vale a Primeira Identidade de Bianchi.

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0. \quad (1.1)$$

(iv) Para todo $X, Y, Z, T \in \mathfrak{X}(M)$ valem as seguintes propriedades de simetria:

$$(a) \langle R(X, Y)Z, T \rangle + \langle R(Y, Z)X, T \rangle + \langle R(Z, X)Y, T \rangle = 0$$

$$(b) \langle R(X, Y)Z, T \rangle = -\langle R(Y, X)Z, T \rangle$$

$$(c) \langle R(X, Y)Z, T \rangle = -\langle R(X, Y)T, Z \rangle$$

$$(d) \langle R(X, Y)Z, T \rangle = \langle R(Z, T)X, Y \rangle.$$

Demonstração. Veja [14]. □

A próxima definição é essencial para definir a curvatura seccional de uma variedade Riemanniana.

Proposição 1.6. *Seja $\sigma \subset T_p M$ um subespaço bi-dimensional do espaço tangente $T_p M$ e sejam $\mathbf{x}, \mathbf{y} \in \sigma$ dois vetores linearmente independentes. Então*

$$K_p(\mathbf{x}, \mathbf{y}) = \frac{\langle R(\mathbf{x}, \mathbf{y})\mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle^2},$$

não depende da escolha dos vetores $\mathbf{x}, \mathbf{y} \in \sigma$.

Demonstração. Veja proposição 3.1, página 104, [14]. □

Definição 1.5 (Curvatura Seccional). *Dado um ponto $\mathbf{p} \in M$ e um subespaço bi-dimensional $\sigma \subset T_p M$, o número real $K_p(\sigma) = K_p(\mathbf{x}, \mathbf{y})$, onde $\{\mathbf{x}, \mathbf{y}\}$ é uma base qualquer de σ é chamado curvatura seccional de σ em \mathbf{p} .*

1.3 Variedades de Hadamard

Definição 1.6. *Uma variedade Riemanniana M é (geodesicamente) completa se para todo $p \in M$, a aplicação exponencial, \exp_p , está definida para todo $v \in T_p M$, isto é, se as geodésicas $\gamma(t)$ que partem de p estão definidas para todos os valores do parâmetro $t \in \mathbb{R}$.*

Se M é uma variedade Riemanniana conexa e $p, q \in M$, então existe uma curva diferenciável por partes ligando p a q , isto é, existe uma curva $\alpha : [a, b] \rightarrow M$ com $\alpha(a) = p$ e $\alpha(b) = q$ e uma partição $a = t_0 < t_1 < \dots < t_n = b$ de $[a, b]$ tal que α é diferenciável em cada intervalo $[t_{i-1}, t_i]$. Usando este fato podemos definir uma distância em uma variedade Riemanniana, como segue:

Definição 1.7. *Seja M uma variedade Riemanniana conexa. Dados $p, q \in M$, definimos a distância $d(p, q)$ por*

$$d(p, q) = \inf\{L(\alpha); \alpha \text{ é uma curva diferenciável por partes ligando } p \text{ a } q\}$$

onde $L(\alpha)$ indica o comprimento da curva α .

Proposição 1.7. *Com a distância d , M é um espaço métrico, isto é:*

- (1) $d(p, q) \geq 0$ e $d(p, q) = 0 \Leftrightarrow p = q$.
- (2) $d(p, q) = d(q, p)$.
- (3) $d(p, r) \leq d(p, q) + d(q, r)$.

Demonstração. Veja Proposição 2.3, página 164, [14]. □

Observação 1.1. *A topologia induzida por d coincide com a topologia inicial de M .*

Definição 1.8. *Dizemos que uma função $f : M \rightarrow \mathbb{R}$ é lipschitziana quando existe um número real K tal que*

$$|f(x) - f(y)| \leq K \cdot d(x, y),$$

para todo $x, y \in M$. O número K é chamado de constante de Lipschitz.

Segue da definição que toda função lipschitziana é uniformemente contínua.

O teorema de Hopf e Rinow mostra algumas condições que são equivalentes à uma variedade Riemanniana M ser geodesicamente completa.

Proposição 1.8 (Hopf-Rinow). *Sejam M uma variedade Riemanniana conexa e $p \in M$. As afirmações abaixo são equivalentes:*

- (a) \exp_p está definida em todo o T_pM , isto é, M é geodesicamente completa.
- (b) Os conjuntos limitados e fechados em M são compactos.
- (c) M é completa como espaço métrico.

Além disso, cada uma das afirmações acima implicam que

- (d) Para todo $q \in M$ existe uma geodésica γ ligando p a q com comprimento de γ igual à $d(p, q)$.

Finalmente podemos definir uma Variedade de Hadamard:

Definição 1.9 (Variedade de Hadamard). *Dizemos que uma variedade Riemanniana é uma Variedade de Hadamard quando M é completa, simplesmente conexa e com curvatura seccional $K_p \leq 0$ para todo $p \in M$.*

Em geral, a aplicação exponencial é uma função de classe C^∞ definida apenas em uma vizinhança Ω da origem em T_pM e um difeomorfismo sobre sua imagem. No caso em que M é uma variedade de Hadamard, temos o seguinte resultado, mostrado por Élie Cartan e Jacques Hadamard.

Proposição 1.9 (Teorema de Cartan-Hadamard). *Se M uma variedade de Hadamard, então M é difeomorfa a \mathbb{R}^n , $n = \dim M$; mais precisamente, $\exp_p : T_pM \rightarrow M$ é um difeomorfismo.*

Demonstração. Veja teorema 3.1, página 165, [14]. □

Seja M uma variedade de Hadamard e $q \in M$. Pelo Teorema de Cartan-Hadamard podemos definir a inversa da aplicação exponencial $\exp_q^{-1} : M \rightarrow T_qM$ e obtém-se a seguinte relação entre distância Riemanniana e aplicação exponencial

$$d(p, q) = \|\exp_q^{-1} p\|. \tag{1.2}$$

De fato, dados $p, q \in M$, note que $\alpha : [0, 1] \rightarrow M$ dada por

$$\alpha(t) = \exp_p(t \cdot \exp_p^{-1} q) = \gamma(1, p, t \cdot \exp_p^{-1} q) = \gamma(t, p, \exp_p^{-1} q)$$

é uma geodésica que liga \mathbf{p} a \mathbf{q} com $\alpha'(0) = \exp_{\mathbf{p}}^{-1}\mathbf{q}$. Assim $\|\alpha'(t)\| = \|\alpha'(0)\| = \|\exp_{\mathbf{p}}^{-1}\mathbf{q}\|$ para todo $t \in [0, 1]$. Logo

$$d(\mathbf{p}, \mathbf{q}) = \int_0^1 \|\alpha'(t)\| dt = \int_0^1 \|\exp_{\mathbf{p}}^{-1}\mathbf{q}\| dt = \|\exp_{\mathbf{p}}^{-1}\mathbf{q}\|. \quad (1.3)$$

Exemplo 1.1. Denote por \mathbb{S}_{++}^n o cone das matrizes simétricas positivas definidas $n \times n$. É possível mostrar que o par $\mathbf{M} = (\mathbb{S}_{++}^n, \langle \cdot, \cdot \rangle)$, com a métrica induzida pela Hessiana Euclidiana de $\varphi(\mathbf{X}) = -\ln \det \mathbf{X}$, isto é,

$$\langle \mathbf{U}, \mathbf{V} \rangle_{\mathbf{X}} = \text{tr}(\mathbf{V}\varphi''(\mathbf{X})\mathbf{U}) = \text{tr}(\mathbf{V}\mathbf{X}^{-1}\mathbf{U}\mathbf{X}^{-1}), \quad \mathbf{X} \in \mathbf{M}, \quad \mathbf{U}, \mathbf{V} \in \mathbf{T}_{\mathbf{X}}\mathbf{M},$$

é uma variedade de Hadamard cuja a única geodésica ligando dois pontos $\mathbf{X}, \mathbf{Y} \in \mathbf{M}$ é dada por

$$\gamma(t) = \mathbf{X}^{\frac{1}{2}}(\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}})^t\mathbf{X}^{\frac{1}{2}},$$

veja [16, Teorema 1.2]. Além disso, as expressões da aplicação exponencial, sua inversa e do quadrado da função distância são dadas por

$$\exp_{\mathbf{X}} \mathbf{V} = \mathbf{X}^{\frac{1}{2}}e^{(\mathbf{X}^{-\frac{1}{2}}\mathbf{V}\mathbf{X}^{-\frac{1}{2}})}\mathbf{X}^{\frac{1}{2}}, \quad \exp_{\mathbf{X}}^{-1} \mathbf{Y} = \mathbf{X}^{\frac{1}{2}} \ln(\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}})\mathbf{X}^{\frac{1}{2}}$$

e

$$d^2(\mathbf{X}, \mathbf{Y}) = \text{tr}(\ln^2(\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}})) = \sum_{i=1}^n \ln^2 \lambda_i[\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}}],$$

onde $\lambda_i[\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}}]$ denota o i -ésimo autovalor da matriz $\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}}$.

1.4 Propriedades geométricas de uma Variedade de Hadamard

Equivalentemente, podemos definir variedades de Hadarmad comparando triângulos geodésicos em \mathbf{M} com triângulos em \mathbb{R}^2 , como pode ser visto em Bacak [2] em um contexto mais geral: Espaços de Hadamard.

Definição 1.10. Um triângulo geodésico $\Delta(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ em uma variedade Riemanniana \mathbf{M} é um conjunto formado por três segmentos de geodésicas minimizantes normalizadas

$$\gamma_1 : [0, l_1] \rightarrow \mathbf{M}, \quad \gamma_2 : [0, l_2] \rightarrow \mathbf{M}, \quad \gamma_3 : [0, l_3] \rightarrow \mathbf{M},$$

de modo que $\gamma_1(l_1) = \mathbf{p}_2 = \gamma_2(0)$, $\gamma_2(l_2) = \mathbf{p}_3 = \gamma_3(0)$ e $\gamma_3(l_3) = \mathbf{p}_1 = \gamma_1(0)$.

Os segmentos geodésicos são chamados de lados do triângulo geodésico e são denotados por $[p_1, p_2]$, $[p_2, p_3]$ e $[p_3, p_1]$.

Dizemos que $z \in \Delta(p_1, p_2, p_3)$ quando x pertence à algum dos segmentos geodésicos, digamos $x \in [p_1, p_2]$. O ponto de comparação de x no triângulo $\Delta(\bar{p}_1, \bar{p}_2, \bar{p}_3)$ é $\bar{x} = (1 - t)\bar{p}_1 + t\bar{p}_2$, onde $t = \frac{d(p_1, x)}{d(p_1, p_2)}$.

Os ângulos $\theta_1 = \angle(-\gamma'_3(l_3), \gamma'_1(0))$, $\theta_2 = \angle(-\gamma_1(l_1), \gamma'_2(0))$ e $\theta_3 = \angle(-\gamma'_2(l_2), \gamma'_3(0))$ são chamados de ângulos internos do triângulo $\Delta(p_1, p_2, p_3)$.

Dado um triângulo geodésico $\Delta(p, q, r)$ em M , existe um triângulo de comparação em \mathbb{R}^2 , isto é, três segmentos de reta $[\bar{p}, \bar{q}]$, $[\bar{q}, \bar{r}]$ e $[\bar{r}, \bar{p}]$ em \mathbb{R}^2 , tais que

$$d(p, q) = \|\bar{p} - \bar{q}\|, \quad d(q, r) = \|\bar{q} - \bar{r}\| \quad \text{e} \quad d(r, p) = \|\bar{r} - \bar{p}\|$$

como mostra a figura 1.1. O triângulo de comparação é único a menos de isometrias em \mathbb{R}^2 .



Figura 1.1: Triângulo geodésico e triângulo de comparação

Proposição 1.10. *Seja M uma variedade de Hadamard. Se $\Delta(p, q, r)$ é um triângulo geodésico e $x \in [p, r]$, $y \in [p, q]$, então*

$$d(x, y) \leq \|\bar{x} - \bar{y}\| \quad (\text{veja figura 1.2}).$$

onde \bar{x} e \bar{y} são os pontos de comparação correspondente no triângulo de comparação $\Delta(\bar{p}, \bar{q}, \bar{r})$.

Demonstração. Veja a definição 1.2.1 e proposição 1.3.3 em Bacak [2] e o corolário 3.1 em Cruz Neto et al.[11]. □



Figura 1.2: Triângulo geodésico e triângulo de comparação

O resultado a seguir, mostra algumas propriedades geométricas de uma variedade de Hadamard.

Proposição 1.11. *Sejam M uma variedade de Hadamard e a, b, c três pontos de M . Tais pontos determinam um único triângulo geodésico T de M com vértices a, b e c . Sejam α, β e γ os ângulos dos vértices a, b e c , respectivamente, e sejam A, B, C os comprimentos dos lados opostos aos vértices a, b, c , respectivamente. Então*

$$(i) \quad A^2 + B^2 - 2AB \cos \gamma \leq C^2 \quad (< C^2, \text{ se } K < 0)$$

$$(ii) \quad \alpha + \beta + \gamma \leq \pi \quad (< \pi, \text{ se } K < 0).$$

Demonstração. Veja Lema 3.1, página 286, [14]. □

Observe que no espaço euclidiano, ocorre a igualdade em (i) e (ii). Além disso, (i) é conhecido como lei dos cossenos.

Como consequência da lei dos cossenos, temos o seguinte resultado:

Proposição 1.12 ([13], proposição 1). *Sejam M uma variedade de Hadamard e $p \in M$. Então a aplicação $\exp_p^{-1} : M \rightarrow T_p M$ é lipschitziana com constante de lipschitz igual a 1.*

Demonstração. Dados $x, y \in M$, aplicamos a proposição 1.11(i) no triângulo geodésico $\Delta(x, p, y)$, obtendo

$$\begin{aligned} \|\exp_p^{-1}(x) - \exp_p^{-1}(y)\|^2 &= \|\exp_p^{-1}(x)\|^2 + \|\exp_p^{-1}(y)\|^2 - 2\langle \exp_p^{-1}(x), \exp_p^{-1}(y) \rangle \\ &= d(x, p)^2 + d(p, y)^2 - 2\langle \exp_p^{-1}(x), \exp_p^{-1}(y) \rangle \\ &\leq d(x, y)^2. \end{aligned}$$

Portanto, $\|\exp_p^{-1}(x) - \exp_p^{-1}(y)\| \leq d(x, y)$. □

1.5 Função distância

Como \exp_p^{-1} é uma aplicação C^∞ e $d(p, q) = \|\exp_p^{-1}q\|$, segue que a função

$$q \mapsto \frac{d^2(p, q)}{2}$$

também é C^∞ .

Neste seção, estamos interessados em encontrar o gradiente dessa função, mas antes disso, faremos uma breve revisão sobre a fórmula da primeira variação de comprimento.

Definição 1.11. *Seja $\alpha : [a, b] \rightarrow M$ uma curva diferenciável em uma variedade M . Uma variação de α é uma aplicação diferenciável $f : (-\epsilon, \epsilon) \times [a, b] \rightarrow M$ tal que $f(0, t) = \alpha(t)$ para todo $t \in [a, b]$.*

Para cada $s \in (-\epsilon, \epsilon)$, a curva $f_s : [a, b] \rightarrow M$ dada por $f_s(t) = f(s, t)$ é chamada de curva da variação.

Definição 1.12. *Dado uma curva $\alpha : [a, b] \rightarrow M$ e uma variação $f : (-\epsilon, \epsilon) \times [a, b] \rightarrow M$, definimos o campo variacional de f por*

$$V(t) = \frac{\partial f}{\partial s}(0, t) \in T_{\alpha(t)}M.$$

Se $\alpha : [a, b] \rightarrow M$ é uma geodésica, f é uma variação de α e V é o campo variacional de f , então a fórmula de primeira variação de comprimento afirma que:

$$L'(0) := \left. \frac{\partial}{\partial s} L(f_s) \right|_{s=0} = \left\langle \frac{\alpha'(b)}{\|\alpha'(b)\|}, V(b) \right\rangle - \left\langle \frac{\alpha'(a)}{\|\alpha'(a)\|}, V(a) \right\rangle.$$

Dado $p \in M$, definimos a função $\rho_p : M \rightarrow \mathbb{R}$ por

$$\rho_p(q) = \frac{1}{2}d(p, q)^2.$$

Proposição 1.13. *O gradiente da função ρ_p no ponto q é dado por:*

$$\text{grad } \rho_p(q) = -\exp_p^{-1}q.$$

Demonstração. Dados $q, u \in M$, considere a curva $\alpha : [0, 1] \rightarrow M$ dada por

$$\alpha(t) = \exp_p(t \cdot \exp_p^{-1}q) = \gamma(1, p, t \exp_p^{-1}q) = \gamma(t, p, \exp_p^{-1}q).$$

Observe que α é uma geodésica ligando p a q , com $\alpha'(0) = \exp_p^{-1}q$ e $\alpha'(1) = -\exp_q^{-1}p$.

Considere a aplicação $f : (-\epsilon, \epsilon) \times [0, 1] \rightarrow M$ dada por

$$f(s, t) = f_s(t) = \exp_p(tv(s)) = \gamma(t, p, v(s)),$$

onde $\mathbf{v}(s) = \mathbf{v} + s\mathbf{w}$, $\mathbf{v} = \alpha'(0) = \exp_p^{-1}\mathbf{q}$ e \mathbf{w} satisfaz $d(\exp_p)_\mathbf{v}\mathbf{w} = \mathbf{u}$. Note que f é uma variação de α , pois f é diferenciável e $f(0, t) = \exp_p(t\mathbf{v}) = \exp_p(t \cdot \exp_p^{-1}\mathbf{q}) = \alpha(t)$. Agora considere a curva $\beta : (-\epsilon, \epsilon) \rightarrow M$ dada por $\beta(s) = f(s, 1) = \exp_p\mathbf{v}(s)$, daí

$$\beta(0) = f(0, 1) = \alpha(1) = \mathbf{q} \text{ e } \beta'(0) = \left. \frac{\partial}{\partial s} f(s, 1) \right|_{s=0} = d(\exp_p)_\mathbf{v}\mathbf{w} = \mathbf{u}.$$

Pela regra da cadeia,

$$d(\rho_p)_\mathbf{q}\mathbf{u} = \left. \frac{d}{ds} \rho_p(\beta(s)) \right|_{s=0} = \left. \frac{d}{ds} \left(\frac{1}{2} d(\mathbf{p}, \beta(s))^2 \right) \right|_{s=0} = d(\mathbf{p}, \mathbf{q}) \cdot \left. \frac{d}{ds} d(\mathbf{p}, \beta(s)) \right|_{s=0}.$$

Note que

$$d(\mathbf{p}, \beta(s)) = \|\exp_p^{-1}\beta(s)\| = \|\exp_p^{-1}(\exp_p\mathbf{v}(s))\| = \|\mathbf{v}(s)\| = \|f'_s(0)\| = L(f_s).$$

Assim,

$$\left. \frac{d}{ds} d(\mathbf{p}, \beta(s)) \right|_{s=0} = \left. \frac{d}{ds} L(f_s) \right|_{s=0} = \left\langle \frac{\alpha'(1)}{\|\alpha'(1)\|}, \mathbf{V}(1) \right\rangle - \left\langle \frac{\alpha'(0)}{\|\alpha'(0)\|}, \mathbf{V}(0) \right\rangle,$$

onde $\mathbf{V}(1) = \left. \frac{\partial}{\partial s} f(s, 1) \right|_{s=0} = \beta'(0) = \mathbf{u}$ e $\mathbf{V}(0) = \left. \frac{\partial}{\partial s} f(s, 0) \right|_{s=0} = 0$. Logo

$$d(\rho_p)_\mathbf{q}\mathbf{u} = d(\mathbf{p}, \mathbf{q}) \cdot \left. \frac{d}{ds} d(\mathbf{p}, \beta(s)) \right|_{s=0} = d(\mathbf{p}, \mathbf{q}) \left\langle \frac{-\exp_q^{-1}\mathbf{p}}{d(\mathbf{p}, \mathbf{q})}, \mathbf{u} \right\rangle = \langle -\exp_q^{-1}\mathbf{p}, \mathbf{u} \rangle.$$

□

Capítulo 2

Noções de Análise Convexa em Variedades de Hadamard

Neste capítulo, apresentaremos as noções básicas de análise convexa em variedades de Hadamard. Ao longo deste capítulo denotaremos uma variedade de Hadamard por M .

2.1 Conjuntos Convexos

Definição 2.1. *Seja M uma variedade de Hadamard. Um conjunto $C \subset M$ é convexo, se para quaisquer pontos $p, q \in C$ a única geodésica ligando p a q em M está contida em C .*

Se $M = \mathbb{R}^n$, a geodésica que liga os pontos p e q é o segmento de reta $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ dado por $\gamma(t) = (1 - t)p + tq$, assim um conjunto $C \subset \mathbb{R}^n$ é convexo quando para quaisquer $p, q \in C$ e $t \in [0, 1]$, tem-se

$$(1 - t)p + tq \in C.$$

Assim, a definição 2.1 é uma generalização natural da definição clássica de conjuntos convexos em \mathbb{R}^n .

Afim de definir o operador projeção, vamos provar o seguinte resultado:

Proposição 2.1. *Sejam C um conjunto convexo fechado de M e $q \in M$. Se $p_q \in C$ é tal que $d(q, p_q) \leq d(q, p')$ para todo $p' \in C$, então*

$$\langle \exp_{p_q}^{-1}q, \exp_{p_q}^{-1}p' \rangle \leq 0, \tag{2.1}$$

para todo $p' \in C$.

Demonstração. Se $q \in C$, tem-se $p_q = q$ e o resultado segue de imediato. Se $q \notin C$, considere a geodésica $\gamma : [0, 1] \rightarrow M$ dada por

$$\gamma(t) = \exp_q(t \cdot \exp_q^{-1} p_q).$$

Note que $\gamma(0) = q, \gamma(1) = p_q$ e $\gamma'(1) = -\exp_{p_q}^{-1} q$. Suponha por absurdo que exista $p_0 \in C$ tal que $\langle \exp_{p_q}^{-1} q, \exp_{p_q}^{-1} p_0 \rangle > 0$. Seja $\beta : (-\epsilon, 1 + \epsilon) \rightarrow M$ uma geodésica tal que $\beta(0) = p_q$ e $\beta(1) = p_0$, assim $\beta'(0) = \exp_{p_q}^{-1} p_0$. Note que a aplicação $\alpha : (-\epsilon, 1 + \epsilon) \times [0, 1] \rightarrow M$ dada por

$$\alpha(s, t) = \exp_q(t \cdot \exp_q^{-1} \beta(s)),$$

é uma variação de γ e seu campo variacional $V(t) = \frac{\partial \alpha}{\partial s}(0, t)$ satisfaz $V(0) = 0$ e $V(1) = \beta'(0) = \exp_{p_q}^{-1}(p_0)$. Daí

$$\begin{aligned} L'(0) &= \left. \frac{d}{ds} L(\alpha_s) \right|_{s=0} = \left\langle V(1), \frac{\gamma'(1)}{\|\gamma'(1)\|} \right\rangle - \left\langle V(0), \frac{\gamma'(0)}{\|\gamma'(0)\|} \right\rangle \\ &= \left\langle \exp_{p_q}^{-1} p_0, \frac{-\exp_{p_q}^{-1} q}{\|\exp_{p_q}^{-1} q\|} \right\rangle \\ &= -\frac{1}{\|\exp_{p_q}^{-1} q\|} \cdot \langle \exp_{p_q}^{-1} q, \exp_{p_q}^{-1} p_0 \rangle < 0. \end{aligned}$$

Logo existe $\delta > 0$ tal que $0 < s < \delta \Rightarrow L(\gamma) > L(\alpha_s)$, isto é, $d(q, p_q) > d(q, \beta(s))$. Além disso, como C é convexo, tem-se $\beta(s) \in C$ para todo $s \in [0, 1]$, que é um absurdo, pois por hipótese, tem-se $d(q, p_q) \leq d(q, p')$ para todo $p' \in C$. \square

Proposição 2.2. *Sejam M uma variedade de Hadamard e $C \subset M$ um conjunto convexo e fechado. Dado $q \in M$, existe um único ponto $p_q \in C$ tal que*

$$d(p_q, q) = \inf\{d(p, q); p \in C\}.$$

Demonstração. Dado $p_0 \in M$ o conjunto $A_{p_0} = \{p \in M; d(p, q) \leq d(p_0, q)\}$ é compacto, pois a aplicação $p \mapsto d(p, q)$ é contínua. Note que

$$\inf\{d(p, q); p \in C\} = \inf\{d(p, q); p \in C \cap A_{p_0}\}.$$

Como $C \cap A_{p_0}$ é compacto, existe $p_q \in C \cap A_{p_0}$ tal que $d(p_q, q) = \inf\{d(p, q); p \in C\}$.

Suponha que existam pontos $p_q, p'_q \in C$ satisfazendo

$$d(p_q, q) = d(p'_q, q) = \inf\{d(p, q); p \in C\}.$$

Considere o triângulo geodésico $\Delta(p_q, p'_q, q)$, pela proposição 2.1, segue que $\theta = \angle(\exp_{p_q}^{-1}q, \exp_{p_q}^{-1}p'_q) \geq \frac{\pi}{2}$ e $\theta' = \angle(\exp_{p'_q}^{-1}q, \exp_{p'_q}^{-1}p_q) \geq \frac{\pi}{2}$, como a soma dos ângulos internos de um triângulo geodésico é menor ou igual a π (ver proposição 1.11), tem-se $\theta = \theta' = \frac{\pi}{2}$, assim $\langle \exp_{p_q}^{-1}q, \exp_{p_q}^{-1}p'_q \rangle = 0$. Pela proposição 1.11, temos

$$d(p_q, q)^2 + d(p_q, p'_q)^2 - 2\langle \exp_{p_q}^{-1}q, \exp_{p_q}^{-1}p'_q \rangle \leq d(p'_q, q)^2.$$

Como $d(p_q, q) = d(p'_q, q)$, segue que $d(p_q, p'_q)^2 \leq 0$, logo $p_q = p'_q$. \square

O ponto p_q da proposição anterior é conhecido como a projeção de q sobre o conjunto C e é denotado por $P_C(p)$. Assim, dado um conjunto C convexo e fechado em M , fica bem definida a aplicação projeção $P_C : M \rightarrow M$, onde $P_C(x)$ é a projeção de x sobre C .

2.2 Funções Convexas

Definição 2.2. *Seja M uma variedade Hadamard. Dizemos que uma função $f : M \rightarrow \mathbb{R}$ é convexa quando para toda geodésica $\gamma : \mathbb{R} \rightarrow M$ a função $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$ for convexa, isto é, quando*

$$f(\gamma((1-t)a + tb)) \leq (1-t)f(\gamma(a)) + tf(\gamma(b)),$$

para todo $t \in [0, 1]$ e $a, b \in \mathbb{R}$. Se a desigualdade for estrita para todo $t \in (0, 1)$, dizemos que f é estritamente convexa.

Proposição 2.3. *Se $f_i : M \rightarrow \mathbb{R}$, $i = 1, 2, \dots, n$ são funções convexas e $c > 0$, então a função $f = \sum_{i=1}^n c_i f_i$ dada por $f(x) = \sum_{i=1}^n c_i f_i(x)$ também é convexa.*

Demonstração. Sejam $\gamma : \mathbb{R} \rightarrow M$ uma geodésica, $a, b \in \mathbb{R}$ e $t \in [0, 1]$. Por hipótese, tem-se:

$$f_i(\gamma((1-t)a + tb)) \leq (1-t)f_i(\gamma(a)) + tf_i(\gamma(b)), \quad i = 1, 2, \dots, n.$$

Daí,

$$c_i f_i(\gamma((1-t)a + tb)) \leq (1-t)c_i f_i(\gamma(a)) + t c_i f_i(\gamma(b)), \quad i = 1, 2, \dots, n. \quad (2.2)$$

Somando as n desigualdades em 2.2, segue o resultado. \square

Proposição 2.4. *Se $f : M \rightarrow \mathbb{R}$ é convexa, então f é localmente lipschitziana, e portanto, contínua.*

Demonstração. Ver teorema 3.6 e corolário 3.10, [21]. □

Definição 2.3. *Seja $f : M \rightarrow \mathbb{R}$ uma função convexa. Dado $p_0 \in M$, dizemos que $s \in T_{p_0}M$ é um subgradiente de f em p_0 se*

$$f(p) \geq f(p_0) + \langle s, \exp_{p_0}^{-1}(p) \rangle,$$

para todo $p \in M$. O conjunto dos subgradientes de f em p_0 é chamado de subdiferencial de f em p_0 e é denotado por $\partial f(p_0)$.

Observação 2.1. *Note que $0 \in \partial f(p)$ se, e somente se, p é ponto de mínimo de f .*

Proposição 2.5. *Se $f : M \rightarrow \mathbb{R}$ é uma função convexa e $p_0 \in M$, então $\partial f(p_0)$ é convexo, compacto e não vazio.*

Demonstração. Ver Teorema 4.5 e 4.6, [21] □

Para demonstrar a próxima proposição, precisaremos do seguinte lema:

Lema 2.1. *Sejam $x_0, y_0 \in M$ e sequências $\{x_n\}, \{y_n\}$ em M tais que $x_n \rightarrow x_0$ e $y_n \rightarrow y_0$. Então*

(i) $\exp_{x_n}^{-1}(y_n) \rightarrow \exp_{x_0}^{-1}(y_0)$.

(ii) *Dados $u_0, v_0 \in T_{x_0}M$ e $\{u_n\}, \{v_n\} \subset T_{x_n}M$, se $u_n \rightarrow u_0$, se $u_n \rightarrow u_0$ e $v_n \rightarrow v_0$, então*

$$\langle u_n, v_n \rangle \rightarrow \langle u_0, v_0 \rangle.$$

Demonstração. (i) Considere a sequência $\omega_n = \exp_{x_n}^{-1}(y_n)$, daí $\exp_{x_n}(\omega_n) = y_n$ para todo $n \in \mathbb{N}$, ou seja,

$$\gamma(1, x_n, \omega_n) = y_n \quad \forall n \in \mathbb{N}, \tag{2.3}$$

onde $\gamma : \mathbb{R} \times TM \rightarrow M$ é o fluxo geodésico de M . Usando a proposição 1.12, tem-se

$$\|\omega_n\| = \|\exp_{x_n}^{-1}(y_n)\| = \|\exp_{x_n}^{-1}(y_n) - \exp_{x_n}^{-1}(x_n)\| \leq d(x_n, y_n),$$

logo a sequência $\{\omega_n\}$ é limitada, pois $\{x_n\}$ e $\{y_n\}$ são convergentes. Se ω' é um ponto de acumulação de $\{\omega_n\}$, segue de (2.3) e da continuidade do fluxo geodésico que

$$\gamma(1, x_0, \omega') = y_0.$$

Ou seja, $\exp_{x_0}(\omega') = y_0 \Rightarrow \omega' = \exp_{x_0}^{-1}(y_0)$, portanto $\exp_{x_n}^{-1}(y_n) \rightarrow \exp_{x_0}^{-1}(y_0)$.

(ii) Dados $x, y \in M$, denotaremos por $P_{x,y}$ o transporte paralelo ao longo da única geodésica que liga x a y . Aplicando a desigualdade triangular e usando que o transporte paralelo é uma isometria, obtemos:

$$\begin{aligned} \|\langle u_n, v_n \rangle - \langle u_0, v_0 \rangle\| &\leq \|\langle u_n, v_n \rangle - \langle P_{x_0, x_n}(u_0), v_n \rangle\| \\ &\quad + \|\langle P_{x_0, x_n}(u_0), v_n \rangle - \langle P_{x_0, x_n}(u_0), P_{x_0, x_n}(v_0) \rangle\| \\ &= \|\langle u_n - P_{x_0, x_n}(u_0), v_n \rangle\| + \|\langle P_{x_0, x_n}(u_0), v_n - P_{x_0, x_n}(v_0) \rangle\| \\ &\leq \|u_n - P_{x_0, x_n}(u_0)\| \cdot \|v_n\| + \|P_{x_0, x_n}(u_0)\| \cdot \|v_n - P_{x_0, x_n}(v_0)\|. \end{aligned}$$

Como a função $x \mapsto P_{x_0, x}(u_0)$ é contínua (veja lema 2.4, [17]), e as sequências $\{v_n\}$ e $\{P_{x_0, x_n}(u_0)\}$ são limitadas, tem-se

$$\langle u_n, v_n \rangle \rightarrow \langle u_0, v_0 \rangle.$$

□

Proposição 2.6. *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Então*

(a) *Se $\{x_k\}$ uma sequência limitada em M e $\{v_k\}$ é tal que $v_k \in \partial f(x_k)$ para todo $k \in \mathbb{N}$, então $\{v_k\}$ também é limitada.*

(b) *Se a sequência $\{x_k, v_k\} \subset M$ é tal que $(x_k, v_k) \rightarrow (x, v) \in TM$ e $v_k \in \partial f(x_k)$ para todo $k \in \mathbb{N}$, então $v \in \partial f(x)$.*

Demonstração. (a) Suponha que $\{v_k\}$ não seja limitada, daí podemos supor sem perda de generalidade que $\|v_k\| \geq k$ para todo $k \in \mathbb{N}$. Como $v_k \in \partial f(x_k)$, tem-se

$$f(y) \geq f(x_k) + \langle v_k, \exp_{x_k}^{-1}y \rangle, \quad \forall y \in M. \quad (2.4)$$

Pela proposição 1.9, tem-se que $\exp_{x_k}^{-1} : M \rightarrow T_{x_k}M$ é um difeomorfismo, logo existe $y_k \in M$ tal que $\exp_{x_k}^{-1}y_k = \frac{v_k}{\|v_k\|}$. Note que

$$\exp_{x_k}^{-1}y_k = \frac{v_k}{\|v_k\|} \Rightarrow d(y_k, x_k) = 1, \quad \forall k \in \mathbb{N}.$$

Como $\{x_k\}$ é limitada, segue que y_k também é limitada. Logo existem subsequências convergentes $\{x_{k_j}\}$ e $\{y_{k_j}\}$ de $\{x_k\}$ e $\{y_k\}$ respectivamente.

Aplicando desigualdade 2.4 em $y = y_{k_j}$, obtemos

$$f(y_{k_j}) \geq f(x_{k_j}) + \left\langle v_{k_j}, \frac{v_{k_j}}{\|v_{k_j}\|} \right\rangle \Rightarrow k_j \leq \|v_{k_j}\| \leq f(y_{k_j}) - f(x_{k_j}) \quad \forall j \in \mathbb{N},$$

que é um absurdo, pois f é convexa e portanto contínua.

(b) Como $v_n \in \partial f(x_n)$ para todo $n \in \mathbb{N}$, tem-se

$$f(y) \geq f(x_n) + \langle v_n, \exp_{x_n}^{-1}(y) \rangle \quad \forall y \in M.$$

Fazendo $n \rightarrow +\infty$, segue do lema 2.1 que

$$f(y) \geq f(x_0) + \langle v, \exp_{x_0}^{-1}(y) \rangle \quad \forall y \in M.$$

Portanto $v \in \partial f(x)$. □

Para a definição a seguir, dados $q, r \in M$, considere $\gamma : [0, 1] \rightarrow M$ a única geodésica ligando q a r .

Definição 2.4. *Uma função $f : M \rightarrow \mathbb{R}$ é k -fortemente convexa quando existe uma constante $k > 0$ tal que*

$$f(\gamma(t)) \leq (1-t)f(q) + tf(r) - kt(1-t)d(q, r)^2,$$

para todo $q, r \in M$ e $t \in [0, 1]$.

Equivalentemente, uma função $f : M \rightarrow \mathbb{R}$ é k -fortemente convexa quando existe uma constante $k > 0$ tal que para toda geodésica $\gamma : \mathbb{R} \rightarrow M$ tem-se

$$f(\gamma((1-t)a + tb)) \leq (1-t)f(\gamma(a)) + tf(\gamma(b)) - kt(1-t)d(\gamma(a), \gamma(b))^2$$

para todo $t \in [0, 1]$ e $a, b \in \mathbb{R}$.

Segue da definição que toda função $f : M \rightarrow \mathbb{R}$ fortemente convexa é estritamente convexa.

Proposição 2.7. *Se $f : M \rightarrow \mathbb{R}$ é uma função k -fortemente convexa e $g : M \rightarrow \mathbb{R}$ é convexa, então a soma $f + g : M \rightarrow \mathbb{R}$ é k -fortemente convexa.*

Demonstração. Seja $\gamma : \mathbb{R} \rightarrow M$ uma geodésica. Dados $a, b \in \mathbb{R}$ e $t \in [0, 1]$, tem-se

$$\begin{aligned} (f + g)(\gamma((1-t)a + tb)) &= f(\gamma((1-t)a + tb)) + g(\gamma((1-t)a + tb)) \\ &\leq (1-t)f(\gamma(a)) + tf(\gamma(b)) - kt(1-t)d^2(\gamma(a), \gamma(b)) \\ &\quad + (1-t)g(\gamma(a)) + tg(\gamma(b)) \\ &\leq (1-t)(f + g)(\gamma(a)) + t(f + g)(\gamma(b)) \\ &\quad - kt(1-t)d^2(\gamma(a), \gamma(b)). \end{aligned}$$

Logo, $f + g$ é k -fortemente convexa. □

A próxima proposição foi demonstrada por Cruz Neto et al.[11] através do estudo de campos monótonos, onde provaram que a função $\rho_p(x) = \frac{d(p, x)^2}{2}$ é fortemente convexa, mas faremos uma demonstração alternativa.

Proposição 2.8. *Seja M uma variedade de Hadamard. Fixado $p \in M$, a função $x \mapsto d^2(p, x)$ é fortemente convexa, isto é, dados $q, r \in M$ e uma geodésica $\gamma : [0, 1] \rightarrow M$ ligando q a r , tem-se*

$$d^2(p, \gamma(t)) \leq (1-t)d^2(p, q) + td^2(p, r) - t(1-t)d^2(q, r),$$

para todo $t \in [0, 1]$.

Demonstração. Considere o triângulo geodésico $\Delta(p, q, r)$ e o triângulo de comparação $\Delta(\bar{p}, \bar{q}, \bar{r})$, de acordo com a proposição 1.10, tem-se

$$\begin{aligned} d(p, \gamma(t))^2 &\leq \|\bar{p} - \gamma(\bar{t})\|^2 \\ &= \|\bar{p} - ((1-t)\bar{q} + t\bar{r})\|^2 \\ &= \|\bar{p} - \bar{q}\|^2 - t\langle \bar{p} - \bar{q}, \bar{r} - \bar{q} \rangle - t\langle \bar{q} - \bar{p}, \bar{q} - \bar{r} \rangle + t^2\|\bar{q} - \bar{r}\|^2 \\ &= \|\bar{p} - \bar{q}\|^2 - t\langle \bar{p} - \bar{q}, \bar{r} - \bar{p} + \bar{p} - \bar{q} \rangle - t\langle \bar{q} - \bar{r} + \bar{r} - \bar{p}, \bar{q} - \bar{r} \rangle + t^2\|\bar{q} - \bar{r}\|^2 \\ &= (1-t)\|\bar{p} - \bar{q}\|^2 + t(-\langle \bar{p} - \bar{q}, \bar{r} - \bar{p} \rangle - \langle \bar{r} - \bar{p}, \bar{q} - \bar{r} \rangle) - (t-t^2)\|\bar{q} - \bar{r}\|^2 \\ &= (1-t)\|\bar{p} - \bar{q}\|^2 + t\|\bar{r} - \bar{p}\|^2 - t(1-t)\|\bar{q} - \bar{r}\|^2 \\ &= (1-t)d^2(p, q) + td^2(p, r) - t(1-t)d^2(q, r). \end{aligned}$$

□

Corolário 2.1. *A função $\rho_p : M \rightarrow \mathbb{R}$ dada por $\rho_p(q) = \frac{d(p, q)^2}{2}$ é fortemente convexa.*

Em Cruz Neto et al.[11], também foi provado que o campo gradiente $\text{grad } \rho_q$ é fortemente monótono com constante igual a 1 e

$$\langle \text{Hess}\rho_p(q) \cdot v, v \rangle \geq \|v\|^2, \quad \forall q \in M \quad \text{e} \quad \forall v \in T_q M. \quad (2.5)$$

2.3 Funções Coercivas

Definição 2.5. *Dizemos que uma função $f : M \rightarrow \mathbb{R}$ é coerciva quando $f(x) \rightarrow \infty$ sempre que $d(x, x_0) \rightarrow \infty$ para algum $x_0 \in X$.*

Proposição 2.9. *Se uma função $f : M \rightarrow \mathbb{R}$ é contínua e coerciva, então f admite pelo menos um ponto de mínimo.*

Demonstração. Sejam $x_0 \in M$ e $c = f(x_0)$, segue da continuidade de f que o conjunto $A = \{x \in X; f(x) \leq c\}$ é fechado. Além disso, A é limitado, caso contrário existiria uma sequência (x_n) em A tal que $d(x_n, x_0) \rightarrow \infty$, contrariando a coercividade de f . Assim, A é limitado e fechado. Pelo teorema de Hopf e Rinow (proposição 1.8), A é compacto. Como f é contínua e A é compacto, existe $x^* \in A$ tal que

$$f(x^*) \leq f(x), \quad \forall x \in A.$$

Se $x \notin A$, também tem-se $f(x) \geq c = f(x_0) \geq f(x^*)$. Portanto x^* é ponto de mínimo da função f . □

Corolário 2.2. *Se $f : M \rightarrow \mathbb{R}$ é estritamente convexa (logo contínua) e coerciva, então f admite exatamente um ponto de mínimo.*

Demonstração. Pela proposição 2.9 existe algum ponto de mínimo. Se existisse dois pontos de mínimo $x_1, x_2 \in M$, teríamos

$$f(\gamma(t)) < (1-t)f(x_1) + tf(x_2)$$

onde $\gamma : [0, 1] \rightarrow M$ é a geodésica ligando x_1 a x_2 e $t \in (0, 1)$. Como $f(x_1) = f(x_2) = \min\{f(x); x \in M\}$, tem-se

$$f(\gamma(t)) < (1-t)f(x_1) + tf(x_2) = \min\{f(x); x \in M\},$$

que é uma contradição. □

Proposição 2.10. *Se $f : M \rightarrow \mathbb{R}$ é uma função fortemente convexa, então f é coerciva.*

Demonstração. Fixado $p_0 \in M$, pela proposição 2.5, existe $s_0 \in \partial f(p_0)$. Dado $p \in M$, considere a geodésica $\gamma : [0, 1] \rightarrow M$ ligando p_0 a p . Assim,

$$\begin{aligned} f\left(\gamma\left(\frac{1}{2}\right)\right) &= f\left(\gamma\left(\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1\right)\right) \\ &\leq \frac{1}{2}f(\gamma(0)) + \frac{1}{2}f(\gamma(1)) - \frac{k}{4}d^2(\gamma(0), \gamma(1)) \\ &= \frac{1}{2}f(p_0) + \frac{1}{2}f(p) - \frac{k}{4}d^2(p_0, p). \end{aligned}$$

Por outro lado,

$$\begin{aligned} f(\gamma(1/2)) &\geq f(p_0) + \langle s_0, \exp_{p_0}^{-1}(\gamma(1/2)) \rangle \\ &\geq f(p_0) - \|s_0\| \cdot \|\exp_{p_0}^{-1}(\gamma(1/2))\| \\ &= f(p_0) - \|s_0\| \cdot \frac{d(p_0, p)}{2}. \end{aligned}$$

Logo

$$\frac{1}{2}f(p_0) + \frac{1}{2}f(p) - \frac{k}{4}d^2(p_0, p) \geq f(\gamma(1/2)) \geq f(p_0) - \|s_0\| \frac{d(p_0, p)}{2}.$$

Isto é,

$$f(p) \geq f(p_0) + d(p_0, p) \left(\frac{k}{2}d(p_0, p) - \|s_0\| \right). \quad (2.6)$$

Se $d(p_0, p) \rightarrow \infty$, por (2.6), segue que $f(p) \rightarrow \infty$. Portanto f é coerciva. \square

Proposição 2.11. *Se $f : M \rightarrow \mathbb{R}$ é k -fortemente convexa, então f admite exatamente um ponto de mínimo. Além disso, se x^* é o ponto de mínimo de f , tem-se*

$$f(x^*) + k \cdot d^2(x^*, y) \leq f(y) \quad (2.7)$$

para todo $y \in M$.

Demonstração. Pela proposição 2.10 e corolário 2.2, segue que f admite exatamente um ponto de mínimo. Para demonstrar (2.7), para cada $y \in M$, considere a geodésica $\gamma : [0, 1] \rightarrow M$ ligando x^* a y , como f é k -fortemente convexa, temos

$$f(x^*) < f(\gamma(t)) \leq (1-t)f(x^*) + tf(y) - kt(1-t)d^2(x^*, y)$$

para todo $t \in (0, 1)$. Daí

$$\begin{aligned} f(x^*) < (1-t)f(x^*) + tf(y) - kt(1-t)d^2(x^*, y) &\Rightarrow tf(x^*) + kt(1-t)d^2(x^*, y) < tf(y) \\ &\Rightarrow f(x^*) + k(1-t)d^2(x^*, y) \leq f(y). \end{aligned}$$

Fazendo $t \rightarrow 0^+$ na última desigualdade, obtemos (2.7). \square

2.4 Resolvente de uma função convexa

Sejam $f : M \rightarrow \mathbb{R}$ uma função convexa e λ uma número real positivo. Fixado $x \in M$, o corolário 2.1 e a proposição 2.7 garante que a função $y \mapsto f(y) + \frac{1}{2\lambda}d^2(y, x)$ é fortemente convexa e pela proposição 2.11 esta função admite exatamente um ponto de mínimo. Finalmente, podemos definir o resolvente de uma função convexa:

Definição 2.6. *Seja $f : M \rightarrow \mathbb{R}$ uma função convexa. Dado $\lambda > 0$, a aplicação $J_\lambda : M \rightarrow M$ dada por*

$$J_\lambda x := \arg \min_{y \in M} \left[f(y) + \frac{1}{2\lambda} d^2(x, y) \right] \text{ e } J_0 x := x, \quad (2.8)$$

é chamado de resolvente de f .

A título de curiosidade, enunciaremos a seguir dois importantes resultados envolvendo o resolvente de uma função convexa, para consultar a demonstração desses fatos, veja [2].

Proposição 2.12. *Sejam M uma variedade de Hadamard e f uma função convexa. Dado $\lambda > 0$, o resolvente J_λ de f é não expansivo, isto é,*

$$d(J_\lambda x_0, J_\lambda x_1) \leq d(x_0, x_1).$$

Demonstração. Veja Teorema 2.2.22, [2]. □

Proposição 2.13. *Sejam M uma variedade de Hadamard, $f : M \rightarrow \mathbb{R}$ uma função convexa e $x \in M$. Se $\{J_{\lambda_n} x\}_{n \in \mathbb{N}}$ é limitada para alguma sequência $\lambda_n \rightarrow \infty$, então $\arg \min f \neq \emptyset$ e*

$$\lim_{\lambda \rightarrow \infty} J_\lambda(x) = P_S(x),$$

onde $S = \arg \min f$ e P_S é a projeção sobre S .

Demonstração. Veja Teorema 2.2.25, [2]. □

2.5 Fluxo gradiente

Para mais detalhes sobre esta seção, consulte [2].

Dado uma aplicação $F : M \rightarrow M$, denote

$$F^n := \underbrace{F \circ \dots \circ F}_{n \text{ vezes}}.$$

Definição 2.7 (Slope de uma função). *Sejam $f : M \rightarrow \mathbb{R}$ uma função convexa e $x \in M$. Definimos o slope de f em x por*

$$|\partial f|(x) := \limsup_{y \rightarrow x} \frac{\max\{f(x) - f(y), 0\}}{d(x, y)}.$$

A definição de slope de uma função tem caráter local, porém a proposição a seguir mostra que ela pode ser equivalentemente obtida por um supremo global.

Proposição 2.14. *Seja $f : M \rightarrow \mathbb{R}$ uma função convexa. Então*

$$|\partial f|(x) = \sup_{y \in M \setminus \{x\}} \frac{\max\{f(x) - f(y), 0\}}{d(x, y)}.$$

Demonstração. Veja Lema 5.1.2, [2]. □

Observação 2.2. *Pela proposição 2.14, segue que x é um ponto de mínimo de uma função convexa $f : M \rightarrow \mathbb{R}$ se, e somente se, $|\partial f|(x) = 0$.*

Se $f : M \rightarrow \mathbb{R}$ é convexa, então o conjunto $\{\|g\|; g \in \partial f(x)\}$ admite mínimo, pois $\partial f(x)$ é compacto e a norma $\|\cdot\|$ é contínua. Para cada $x \in M$, denotamos por $\partial^0 f(x)$ o elemento de $\partial f(x)$ que possui menor norma.

Proposição 2.15 ([13], Corolário 1). *Seja $f : M \rightarrow \mathbb{R}$ uma função convexa. Então*

$$|\partial f|(x) \leq \|\partial^0 f(x)\|,$$

para todo $x \in M$.

Demonstração. Dados $x, y \in M$ e $v \in \partial f(x)$ com $f(x) - f(y) > 0$, segue da definição de subgradiente que $f(y) - f(x) \geq \langle v, \exp_x^{-1}y \rangle$, daí

$$f(x) - f(y) \leq -\langle v, \exp_x^{-1}y \rangle \leq \|v\| \cdot \|\exp_x^{-1}y\| = \|v\| \cdot d(x, y).$$

Pela proposição 2.14, segue que $|\partial f|(x) \leq \|v\|$, portanto $\|\partial f\|(x) \leq \|\partial^0 f(x)\|$. □

Em particular, a proposição 2.15 garante que $|\partial f|(x) < \infty$ para todo $x \in M$. Assim, fica bem definida a função $|\partial f| : M \rightarrow \mathbb{R}$.

A próxima proposição também é válida em espaços de Hadamard, como mostrado em [2], mas continuaremos nos restringindo a Variedades de Hadamard.

Proposição 2.16 (Existência do Fluxo). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Se $x_0 \in M$ e $t \in [0, \infty)$, então existe o seguinte limite*

$$S_t(x_0) := \lim_{n \rightarrow \infty} J_{\frac{t}{n}}^{(n)}(x_0).$$

Além disso,

$$d(S_t(x_0), J_{\frac{t}{n}}^{(n)}x_0) \leq \frac{t}{\sqrt{2n}} |\partial f|(x_0).$$

Demonstração. Veja Teorema 5.1.6, página 99, [2]. □

Proposição 2.17. *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Então $(S_t)_t$ é um semigrupo fortemente contínuo de aplicações não expansivas, isto é,*

1. *A aplicação $t \mapsto S_t x_0$, onde $t \in [0, \infty)$, é contínua, em particular, $\lim_{t \rightarrow 0^+} S_t x = x$.*
2. *$S_t(S_s x) = S_{s+t} x$ para todo $s, t \in [0, \infty)$.*
3. *$d(S_t x, S_t y) \leq d(x, y)$, para todo $t \in [0, \infty)$ e $x, y \in M$.*

Demonstração. Veja proposição 5.1.8, [2]. □

Na verdade, a aplicação $t \mapsto S_t x_0$ é mais do que contínua, ela é localmente lipschitz em $(0, \infty)$ e lipschitz em $[t_0, \infty)$, onde t_0 é um número real positivo, como mostrado na proposição 5.1.10, [2].

Proposição 2.18. *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Dado $x_0 \in M$, seja $x(t) = S_t x_0$ para todo $t \in [0, \infty)$. Então a função $t \mapsto x(t)$ é absolutamente contínua em $(0, \infty)$ e satisfaz*

$$\frac{1}{2} \frac{d}{dt} [d^2(y, x(t))] + f(x(t)) \leq f(y) \quad (2.9)$$

para quase todo $t \in (0, \infty)$ e para todo $y \in M$. Reciprocamente, se uma curva absolutamente contínua $z : (0, \infty) \rightarrow M$ com $\lim_{t \rightarrow 0^+} z(t) = x_0$ satisfaz (2.9), então $z(t) = x(t)$ para todo $t \in (0, \infty)$.

Demonstração. Veja Teorema 5.1.11, [2] □

A proposição abaixo mostra uma importante propriedade das curvas absolutamente contínuas, para mais detalhes, consultar [8].

Proposição 2.19. *Seja M uma variedade Riemanniana conexa. Para qualquer curva absolutamente contínua $\gamma : [a, b] \rightarrow M$ a derivada γ' existe para quase todo $t \in [a, b]$ e $\|\gamma'(t)\|$ é integrável. Em particular,*

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Demonstração. Veja proposição 3.7, [8]. □

É importante lembrar também que uma função $F : [a, b] \rightarrow \mathbb{R}$ é absolutamente contínua se, e somente se. F' existe para quase todo ponto em (a, b) , F' é integrável e

$$F(x) - F(a) = \int_a^x F'(t) dt,$$

para todo $a \leq x \leq b$.

A proposição a seguir mostra uma importante propriedade do slope de uma função e será usada no próximo capítulo para mostrar que a curva $x(t)$ é diferenciável a direita para todo $t \geq 0$.

Proposição 2.20. *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Se $x_0 \in M$ e $x(t) = S_t x_0$, então para cada $t \in (0, \infty)$, tem-se*

1. $|\partial f|(x(t)) = \lim_{h \rightarrow 0^+} \frac{d(x(t+h), x(t))}{h},$
2. $|\partial f|(x(t)) = \lim_{h \rightarrow 0^+} \frac{f(x(t)) - f(x(t+h))}{d(x(t+h), x(t))},$
3. $|\partial f|^2(x(t)) = \lim_{h \rightarrow 0^+} \frac{f(x(t)) - f(x(t+h))}{h}.$

Demonstração. Veja Teorema 5.1.13, [2]. □

Proposição 2.21. *Sejam M uma variedade de Hadamard, $f : M \rightarrow \mathbb{R}$ uma função convexa e $x(t) = S_t x_0$. Então a função $t \mapsto \|\partial f\|(x(t))$ é não crescente e*

$$|\partial f|(x(t))^2 \leq |\partial f|(y)^2 + \frac{1}{t^2} d^2(y, x_0).$$

Demonstração. Veja teorema 5.1.14, [2]. □

Capítulo 3

Curva subgradiente e suas propriedades

3.1 Existência da curva subgradiente

Sejam H um espaço de Hilbert, $f : H \rightarrow \mathbb{R}$ uma função convexa diferenciável e $x_0 \in H$. Considere o seguinte problema

$$\begin{cases} -x'(t) = \text{grad } f(x(t)) \\ x(0) = x_0, \end{cases} \quad (3.1)$$

onde $x : (0, \infty) \rightarrow H$ é uma curva diferenciável. Em outras palavras, estamos procurando uma curva x que parte de x_0 e se move na direção de máxima descida da função f .

Neste capítulo, estamos interessados em estudar um problema mais geral do que esse, para posteriormente estudar a relação entre error bounds e a desigualdade de Kurdyka-Lojasiewicz. Ao invés de trabalhar com funções convexas diferenciáveis em espaços de Hilbert, trabalharemos com funções convexas (não necessariamente diferenciável) em Variedades de Hadamard, para isso usaremos o subdiferencial de f ao invés de $\text{grad } f$, mais precisamente, sejam M uma variedade de Hadamard, $f : M \rightarrow \mathbb{R}$ uma função convexa e $x_0 \in M$, considere o seguinte problema

$$\begin{cases} -x'(t) \in \partial f(x(t)) \\ x(0) = x_0, \end{cases} \quad (3.2)$$

onde $x : [0, \infty) \rightarrow M$ é uma curva em M .

Como a função f não é diferenciável, não é de se esperar que exista uma curva diferenciável x satisfazendo (3.2) para todo $t \in [0, \infty)$, mas os próximos resultados, que

podem ser encontrados em [13], mostra a existência de uma curva absolutamente contínua satisfazendo (3.2) para quase todo $t \in [0, \infty)$.

Note que (3.2) de fato é uma generalização de (3.1), pois todo espaço de Hilbert é uma variedade de Hadamard com curvatura seccional nula e se f é diferenciável, tem-se $\partial f(x) = \{\text{grad } f(x)\}$.

Proposição 3.1 ([14], Lema 1). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Para cada $x \in M$, existe uma única curva absolutamente contínua $\chi : [0, \infty) \rightarrow M$ tal que*

$$\begin{cases} -\chi'(t) \in \partial f(\chi(t)) \\ \chi(0) = x_0, \end{cases} \quad (3.3)$$

para quase todo $t > 0$.

Demonstração. Pela proposição 2.18, existe uma única curva absolutamente contínua $\chi : [0, \infty) \rightarrow M$ tal que $\chi(0) = x_0$ e

$$\frac{1}{2} \frac{d}{dt} [d^2(y, \chi(t))] + f(\chi(t)) \leq f(y),$$

para todo $y \in M$. Por outro lado, pela proposição 1.13, tem-se

$$\frac{1}{2} \frac{d}{dt} [d^2(y, \chi(t))] = \frac{d}{dt} [\rho_y(\chi(t))] = \langle \chi'(t), \text{grad } \rho_y(\chi(t)) \rangle = -\langle \chi'(t), \exp_{\chi(t)}^{-1}(y) \rangle.$$

Daí,

$$\langle -\chi'(t), \exp_{\chi(t)}^{-1}(y) \rangle + f(\chi(t)) \leq f(y),$$

para todo $y \in M$. Portanto $-\chi'(t) \in \partial f(\chi(t))$. □

3.2 Propriedades

No transcorrer do texto, denotaremos por $\chi(t)$ a solução dada pela proposição 3.1, que é chamada de curva subgradiente. Os próximos resultados mostram algumas propriedades de $\chi(t)$, e podem ser encontrados em Cruz Neto et al.[13].

Proposição 3.2 ([13], Lema 2). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Se $\chi : [0, \infty) \rightarrow M$ é a solução de (3.3), então*

$$(f \circ \chi)'(t) = -\|\chi'(t)\|^2,$$

para quase todo $t > 0$. Em particular a função $t \mapsto f(\chi(t))$ é decrescente.

Demonstração. Como $-x'(t) \in \partial f(x(t))$ para quase todo $t > 0$, temos:

$$f(x(t+h)) - f(x(t)) \geq \langle -x'(t), \exp_{x(t)}^{-1} x(t+h) \rangle. \quad (3.4)$$

Também tem-se $-x'(t+h) \in \partial f(x(t+h))$, assim

$$f(x(t+h)) - f(x(t)) \leq \langle x'(t+h), \exp_{x(t+h)}^{-1} x(t) \rangle. \quad (3.5)$$

Dividindo por $h > 0$ ambos os lados das desigualdades (3.4) e (3.5), obtemos

$$\frac{f(x(t+h)) - f(x(t))}{h} \geq \left\langle -x'(t), \frac{\exp_{x(t)}^{-1} x(t+h)}{h} \right\rangle, \quad (3.6)$$

$$\frac{f(x(t+h)) - f(x(t))}{h} \leq \left\langle x'(t+h), \frac{\exp_{x(t+h)}^{-1} x(t)}{h} \right\rangle. \quad (3.7)$$

Considere as curvas $\alpha : (-\epsilon, \epsilon) \rightarrow T_{x(t)}M$ e $\beta : (-\epsilon, \epsilon) \rightarrow M$ definidas por

$$\alpha(h) = \exp_{x(t)}^{-1}(x(t+h)) \text{ e } \beta(h) = \exp_{x(t)}(\alpha(h)) = x(t+h).$$

Note que $\alpha(0) = 0$ e $\beta'(0) = x'(t)$. Lembrando que $d(\exp_p)_0 = \text{Id}$, obtemos

$$\beta'(0) = d(\exp_{x(t)})_{\alpha(0)}(\alpha'(0)) = \alpha'(0) \Rightarrow \alpha'(0) = x'(t).$$

Podemos reescrever (3.6) da seguinte forma

$$\frac{f(x(t+h)) - f(x(t))}{h} \geq - \left\langle x'(t), \frac{\alpha(h) - \alpha(0)}{h} \right\rangle.$$

Da última desigualdade, segue que

$$\lim_{h \rightarrow 0^+} \frac{f(x(t+h)) - f(x(t))}{h} \geq - \langle x'(t), \alpha'(0) \rangle = -\|x'(t)\|^2.$$

Mostraremos agora a desigualdade oposta. Pela proposição 1.13, tem-se

$$\text{grad } \rho_{x(t)}(x(t+h)) = -\exp_{x(t+h)}^{-1}(x(t)).$$

Logo,

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{-\exp_{x(t+h)}^{-1}(x(t))}{h} &= \lim_{h \rightarrow 0^+} \frac{\text{grad } \rho_{x(t)}(x(t+h)) - \text{grad } \rho_{x(t)}(x(t))}{h} \\ &= (\rho_{x(t)} \circ x)''(t) \\ &= \text{Hess } \rho_{x(t)} \cdot x'(t). \end{aligned}$$

Fazendo $h \rightarrow 0^+$ na desigualdade 3.7 e usando (2.5), obtemos

$$\lim_{h \rightarrow 0^+} \frac{f(x(t+h)) - f(x(t))}{h} \leq -\langle x'(t), \text{Hess } \rho_{x(t)}(x(t)) \cdot x'(t) \rangle \leq -\|x'(t)\|^2.$$

Portanto,

$$\lim_{h \rightarrow 0^+} \frac{f(x(t+h)) - f(x(t))}{h} = -\|x'(t)\|^2.$$

Como f é convexa, segue da proposição 2.4 que f é localmente lipschitziana e portanto $f \circ x$ é localmente absolutamente contínua, isso conclui a demonstração. \square

Proposição 3.3 ([13], Lema 3). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Seja $x : [0, \infty) \rightarrow M$ uma solução de (3.3), então*

$$x'(t) = -\partial^0 f(x(t)), \quad (3.8)$$

para quase todo $t > 0$. Além disso, a função $x \mapsto \|\partial^0 f(x)\|$ é semicontínua inferiormente (lsc).

Demonstração. Como $f \circ x$ é localmente absolutamente contínua, existe $\delta > 0$ tal que $f \circ x$ restrita a $[0, \delta]$ é absolutamente contínua, pela proposição 3.2, tem-se

$$f(x(t+\delta)) - f(x(t)) = \int_t^{t+\delta} (f \circ x)'(s) ds = - \int_t^{t+\delta} \|x'(s)\|^2 ds.$$

Dado $v \in \partial f(x(t))$, segue da definição de subgradiente que

$$\int_t^{t+\delta} \|x'(s)\|^2 ds = -(f(x(t+\delta)) - f(x(t))) \leq -\langle v, \exp_{x(t)}^{-1}(x(t+\delta)) \rangle \leq \|v\| \cdot d(x(t), x(t+\delta)).$$

Por outro lado,

$$d(x(t), x(t+\delta)) \leq \int_t^{t+\delta} \|x'(s)\| ds.$$

Logo,

$$\frac{\int_t^{t+\delta} \|x'(s)\|^2 ds}{\delta} \leq \|v\| \cdot \frac{\int_t^{t+\delta} \|x'(s)\| ds}{\delta}.$$

Fazendo $\delta \rightarrow 0^+$, segue do teorema de diferenciação de Lebesgue que

$$\|x'(t)\|^2 \leq \|v\| \cdot \|x'(t)\|.$$

Assim $\|x'(t)\| \leq \|v\|$ para todo $v \in \partial f(x(t))$, e portanto $\|x'(t)\| \leq \partial^0 f(x(t))$. Como $-x'(t) \in \partial f(x(t))$, também temos $\|x'(t)\| \geq \partial^0 f(x(t))$, logo $\|x'(t)\| = \partial^0 f(x(t))$.

Mostraremos agora que $\|\partial^0 f(x)\|$ é semicontínua inferiormente, isto é, se $x_n \rightarrow x$, então

$$\|\partial^0 f(x)\| \leq \liminf_{n \rightarrow \infty} \|\partial^0 f(x_n)\|.$$

Seja $\{x_{n_k}\}$ uma subsequência de $\{x_n\}$ tal que

$$\lim_{n_k \rightarrow \infty} \|\partial^0 f(x_{n_k})\| = \liminf_{n \rightarrow \infty} \|\partial^0 f(x_n)\|.$$

Pela proposição 2.6, podemos supor que $\partial^0 f(x_{n_k}) \rightarrow v \in \partial f(x)$, portanto

$$\|\partial^0 f(x)\| \leq \|v\| = \lim_{n_k \rightarrow \infty} \|\partial^0 f(x_{n_k})\| = \liminf_{n \rightarrow \infty} \|\partial^0 f(x_n)\|.$$

□

A proposição 3.3 não garante que a igualdade obtida em (3.8) é válida para todo $t \geq 0$, já que $x(t)$ não necessariamente é diferenciável. Porém, o próximo resultado mostra que $x(t)$ é diferenciável a direita para todo $t \geq 0$ e vale uma igualdade similar a (3.8).

Proposição 3.4 ([13], proposição 5). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Se $x(t)$ é solução de (3.3), então $x(t)$ é diferenciável a direita em todo $t \geq 0$, e*

$$x'(t^+) = -\partial^0 f(x(t))$$

para todo $t \geq 0$.

Demonstração. Considere a curva $\alpha : (-\epsilon, \epsilon) \rightarrow T_{x(t)}M$ dada por

$$\alpha(h) = \exp_{x(t)}^{-1}(x(t+h)).$$

Assim como na demonstração da proposição 3.2, se existir a derivada a direita de α em $t = 0$, então também existe $x'(0^+)$ e além disso $x'(0^+) = \alpha(0^+)$. Note que, para $h > 0$, tem-se

$$\frac{\alpha(h) - \alpha(0)}{h} = \frac{\exp_{x(t)}^{-1}(x(t+h)) - \exp_{x(t)}^{-1}(x(t))}{h} = \frac{\exp_{x(t)}^{-1}(x(t+h))}{h}.$$

Logo, basta mostrar que

$$\lim_{h \rightarrow 0^+} \left\| \frac{\exp_{x(t)}^{-1}(x(t+h))}{h} + \partial^0 f(x(t)) \right\| = 0.$$

Para isso, veja que

$$\begin{aligned} \left\| \frac{\exp_{x(t)}^{-1}(x(t+h))}{h} + \partial^0 f(x(t)) \right\|^2 &= \frac{1}{h^2} d^2(x(t), x(t+h)) + \|\partial^0 f(x(t))\|^2 \quad (3.9) \\ &+ \frac{2}{h} \langle \exp_{x(t)}^{-1}(x(t+h)), \partial^0 f(x(t)) \rangle. \end{aligned}$$

Pela definição de subgradiente e pela proposição 3.2, segue que

$$\begin{aligned}
 \langle \exp_{x(t)}^{-1}(x(t+h)), \partial^0 f(x(t)) \rangle &\leq f(x(t+h)) - f(x(t)) \\
 &= \int_t^{t+h} (f \circ x)'(s) \, ds \\
 &= - \int_t^{t+h} \|x'(s)\|^2 \, ds \\
 &\leq - \int_t^{t+h} \|\partial^0 f(x(s))\|^2 \, ds.
 \end{aligned}$$

Sendo a função $x \mapsto \|\partial^0 f(x)\|$ semicontínua inferiormente, para cada $\epsilon > 0$, existe $\delta > 0$ tal que

$$d(y, x) < \delta \Rightarrow \|\partial^0 f(y)\| > \|\partial^0 f(x)\| - \epsilon.$$

Daí, para h suficientemente pequeno, temos

$$- \int_t^{t+h} \|\partial^0 f(x(s))\|^2 \, ds \leq - \int_t^{t+h} (\|\partial^0 f(x)\| - \epsilon)^2 \, ds = -h(\|\partial^0 f(x)\| - \epsilon)^2.$$

Logo

$$\langle \exp_{x(t)}^{-1}(x(t+h)), \partial^0 f(x(t)) \rangle \leq -h \cdot \|\partial^0 f(x)\|^2. \quad (3.10)$$

Por (3.9) e (3.10), temos

$$\begin{aligned}
 \left\| \frac{\exp_{x(t)}^{-1}(x(t+h))}{h} + \partial^0 f(x(t)) \right\|^2 &\leq \frac{d^2(x(t), x(t+h))}{h^2} + \|\partial^0 f(x)\|^2 - \frac{2}{h} h(\|\partial^0 f(x)\| - \epsilon)^2 \\
 &= \frac{d^2(x(t), x(t+h))}{h^2} - (\|\partial^0 f(x)\| - \epsilon)^2.
 \end{aligned}$$

Lembre-se que na proposição 2.20 provamos que

$$|\partial f|(x(t)) = \lim_{h \rightarrow 0^+} \frac{d(x(t+h), x(t))}{h}$$

e na proposição 2.15 mostramos que $|\partial f|(x) \leq \|\partial^0 f(x)\|$, assim

$$\lim_{h \rightarrow 0^+} \left\| \frac{\exp_{x(t)}^{-1}(x(t+h))}{h} + \partial^0 f(x(t)) \right\|^2 \leq \|\partial^0 f(x)\|^2 - (\|\partial^0 f(x)\| - \epsilon)^2.$$

Fazendo $\epsilon \rightarrow 0^+$, segue o resultado. □

Na proposição 2.15 provamos que $|\partial f|(x) \leq \|\partial^0 f(x)\|$, mostraremos a seguir que na verdade ocorre a igualdade.

Corolário 3.1 ([13], corolário 2). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa, então*

$$|\partial f|(x) = \|\partial^0 f(x)\|, \quad \forall x \in M.$$

Além disso, $f \circ x$ é diferenciável a direita e $(f \circ x)'(t^+) = -\|x'(t^+)\|^2$.

Demonstração. Dado $x \in M$, seja $x(t)$ a solução de (3.3), com $x(0) = x$. Fixado $t \geq 0$, considere a curva $\alpha : [0, \epsilon] \rightarrow T_{x(t)}M$ dada por $\alpha(h) = \exp_{x(t)}^{-1}(x(t+h))$. Daí, $\alpha(0) = 0$ e $x(t+h) = \exp_{x(t)}(\alpha(h))$, como $\exp_{x(t)} : T_{x(t)}M \rightarrow M$ é um difeomorfismo, segue da proposição 3.4 que

$$-\partial^0 f(x(t)) = x'(t^+) = d(\exp_{x(t)})_{\alpha(0)}(\alpha'(0^+)) = \alpha'(0^+).$$

Por outro lado, usando a proposição 2.20, tem-se

$$\|\alpha'(0^+)\| = \lim_{h \rightarrow 0^+} \frac{\|\exp_{x(t)}^{-1}(x(t+h))\|}{h} = \lim_{h \rightarrow 0^+} \frac{d(x(t+h), x(t))}{h} = |\partial f|(x(t)).$$

Logo $|\partial f|(x(t)) = \|\partial^0 f(x(t))\|$, e portanto, $|\partial f|(x) = \|\partial^0 f(x)\|$.

Novamente pela proposição 2.20, podemos concluir que a função $(f \circ x)(t)$ é diferenciável a direita para todo $t \geq 0$ e

$$(f \circ x)'(t^+) = -|\partial f|^2(x(t)) = -\|\partial^0 f(x(t))\|^2 = -\|x'(t^+)\|^2.$$

□

A seguir, temos mais uma propriedade da curva subgradiente $x(t)$. Ao longo do texto usaremos a notação $S := \arg \min f = \{x \in M; f(x) \leq f(z) \quad \forall z \in M\}$.

Proposição 3.5 ([13], lema 5). *Sejam M uma variedade de Hadamard, $f : M \rightarrow \mathbb{R}$ uma função convexa. Se $p \in S$, então a função $t \mapsto d(p, x(t))$ é decrescente.*

Em particular existe

$$\lim_{t \rightarrow +\infty} d(x(t), p).$$

Demonstração. Usando a proposição 2.18 e que $p \in S$, temos

$$\frac{1}{2} \frac{d}{dt} [d^2(p, x(t))] + f(x(t)) \leq f(p) \leq f(x(t)) \Rightarrow \frac{d}{dt} [d^2(p, x(t))] \leq 0.$$

Assim, a função $t \mapsto d(p, x(t))$ é decrescente, em particular existe

$$\lim_{t \rightarrow +\infty} d(x(t), p).$$

□

Proposição 3.6. *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa limitada inferiormente. Se $\mathbf{x}(t)$ é solução de (3.3), então a função $t \mapsto \|\mathbf{x}'(t^+)\|$ é não crescente, e*

$$\lim_{t \rightarrow +\infty} \|\mathbf{x}'(t^+)\| = 0.$$

Demonstração. Suponha que

$$\inf\{\|\mathbf{x}'(t^+)\|; t \in [0, \infty)\} = c > 0.$$

Sejam $s, t \in [0, \infty)$, com $s < t$. Pelo corolário 3.1, tem-se

$$f(\mathbf{x}(t)) - f(\mathbf{x}(s)) = \int_s^t (f \circ \mathbf{x})'(\tau) d\tau = - \int_s^t \|\mathbf{x}'(\tau)\| d\tau \leq -c(t - s) < 0.$$

Daí, $(f \circ \mathbf{x})(t)$ é decrescente e $\lim_{t \rightarrow +\infty} (f \circ \mathbf{x})(t) = -\infty$, que é um absurdo, pois por hipótese f é limitada inferiormente. Isso mostra que

$$\inf\{\|\mathbf{x}'(t^+)\|; t \in [0, \infty)\} = 0. \quad (3.11)$$

Usando a proposição 3.4 e o corolário 3.1, obtemos

$$\|\mathbf{x}'(t^+)\| = \|\partial^0 f(\mathbf{x}(t))\| = |\partial f(\mathbf{x}(t))|.$$

Pela proposição 2.21, segue que a função $t \mapsto \|\mathbf{x}'(t^+)\|$ é não crescente e por (3.11), temos

$$\lim_{t \rightarrow +\infty} \|\mathbf{x}'(t^+)\| = 0.$$

□

O próximo resultado foi provado em [13], mas apresentaremos uma prova alternativa.

Proposição 3.7. *Sejam M uma variedade de Hadamard, $f : M \rightarrow \mathbb{R}$ uma função convexa limitada inferiormente. Se $\mathbf{x}(t)$ é uma curva subgradiente de f e $\{t_k\} \subset [0, +\infty)$ é uma sequência tal que $t_k \rightarrow +\infty$ e $\lim_{k \rightarrow \infty} \mathbf{x}(t_k) = \mathbf{x}^*$, então $\mathbf{x}^* \in S = \arg \min f$.*

Demonstração. Como $-\mathbf{x}'(t_k^+) \in \partial f(\mathbf{x}(t_k))$ e $\lim_{k \rightarrow \infty} \mathbf{x}(t_k) = \mathbf{x}^*$, pela proposição 2.6, podemos supor sem perda de generalidade que

$$(\mathbf{x}(t_k), -\mathbf{x}'(t_k^+)) \rightarrow (\mathbf{x}^*, \mathbf{v}), \quad \mathbf{v} \in \partial f(\mathbf{x}^*).$$

Por outro lado, pela proposição 3.6 temos

$$\lim_{k \rightarrow \infty} \|\mathbf{x}'(t_k^+)\| = 0.$$

Logo $\mathbf{v} = 0$ e portanto $0 \in \partial f(\mathbf{x}^*)$, isto é, \mathbf{x}^* é ponto de mínimo. □

Proposição 3.8 ([13], lema 7). *Sejam M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa limitada inferiormente. Se $x(t)$ é solução de (3.3), então*

$$\lim_{t \rightarrow +\infty} x(t) = x^* \in \arg \min f.$$

Em particular, $\lim_{t \rightarrow +\infty} (f \circ x)(t) = \min f$.

Demonstração. Se $\{t_k\} \subset M$ é uma sequência tal que $t_k \rightarrow \infty$, segue da proposição 3.5 que a sequência $\{x(t_k)\}$ é limitada e a proposição 3.7 garante que seus pontos de acumulação são pontos de mínimo da função f , resta então mostrar que essa sequência possui apenas um ponto de acumulação.

De fato, se $\lim_{j \rightarrow +\infty} x(t_{k_j}) = x^*$ e $\lim_{s \rightarrow +\infty} x(t_{k_s}) = \bar{x}$, pela proposição 3.5 existe $\lim_{t \rightarrow +\infty} d(x(t), x^*)$, logo

$$\lim_{s \rightarrow +\infty} d(x(t_{k_s}), x^*) = \lim_{j \rightarrow +\infty} d(x(t_{k_j}), x^*).$$

Ou seja,

$$d(\bar{x}, x^*) = d(x^*, x^*) = 0.$$

Portanto

$$\lim_{t \rightarrow +\infty} x(t) = x^* \in \arg \min f.$$

□

Concluimos esta seção com o teorema abaixo, onde reunimos as principais propriedades da curva subgradiente obtidas neste capítulo.

Teorema 3.1. *Seja M uma variedade de Hadamard e $f : M \rightarrow \mathbb{R}$ uma função convexa. Para cada $x \in M$, existe uma única curva absolutamente contínua $x : [0, \infty) \rightarrow M$ tal que $x(0) = x$ e $-x'(t) \in \partial f(x(t))$ para quase todo $t > 0$. Além disso,*

1. $(f \circ x)'(t) = -\|x'(t)\|^2$ para quase todo $t > 0$;
2. $(f \circ x)'(t^+) = -\|x'(t^+)\|^2$ para todo $t \geq 0$;
3. $x'(t) = -\partial^0 f(x(t))$ para quase todo $t > 0$;
4. $x'(t^+) = -\partial^0 f(x(t^+))$ para todo $t > 0$;
5. A função $t \mapsto \|x'(t^+)\|$ é não crescente, e $\lim_{t \rightarrow +\infty} \|x'(t^+)\| = 0$;
6. Se $p \in S$, então a função $t \mapsto d(p, x(t))$ é decrescente;
7. Existe $\lim_{t \rightarrow +\infty} x(t) = x^*$ e $x^* \in \arg \min f$;
8. A função $t \mapsto f(x(t))$ é não crescente e $\lim_{t \rightarrow +\infty} f(x(t)) = \min f$.

Capítulo 4

Error bounds e desigualdade de Kurdyka - Łojasiewicz

4.1 A desigualdade de Kurdyka - Łojasiewicz

Nesta seção apresentaremos a desigualdade de Kurdyka - Łojasiewicz. Para simplificar a notação, escreveremos $[\alpha < f < \beta] := \{x \in M : \alpha < f(x) < \beta\}$. Dado $r_0 > 0$, considere o conjunto

$$\mathcal{K}(0, r_0) := \{\varphi \in C^0[0, r_0) \cap C^1(0, r_0); \varphi(0) = 0, \varphi \text{ é côncava e } \varphi' > 0\}.$$

A função f satisfaz a desigualdade de Kurdyka - Łojasiewicz (ou tem a propriedade KL) localmente em $\bar{x} \in M$ se existem $r_0 > 0$, $\varphi \in \mathcal{K}(0, r_0)$ e $\delta > 0$ tais que

$$\varphi'(f(x) - f(\bar{x})) \cdot \|\partial^0 f(x)\| \geq 1,$$

para todo $x \in B(\bar{x}, \delta) \cap [f(\bar{x}) < f < f(\bar{x}) + r_0]$. A função φ acima é chamada de função desingularizante de f em \bar{x} .

Se \bar{x} não é um ponto de mínimo de f , mostraremos que f tem a propriedade KL em \bar{x} . Para isso, precisaremos do seguinte lema:

Lema 4.1. *Sejam $f : M \rightarrow \mathbb{R}$ uma função convexa e $\bar{x} \in M$ tal que $0 \notin \partial f(\bar{x})$. Existe $\delta > 0$ tal que*

$$x \in M, d(x, \bar{x}) < \delta \Rightarrow \|\partial^0 f(x)\| \geq \delta.$$

Demonstração. Suponha por absurdo que o lema não seja verdadeiro, logo existem sequências

$(\delta_k)_{k \in \mathbb{N}}$ e $(x_k)_{k \in \mathbb{N}}$, com $\delta_k \rightarrow 0$, tais que

$$d(x_k, \bar{x}) < \delta_k \text{ e } \|\partial^0 f(x_k)\| < \delta_k. \quad (4.1)$$

Para cada $k \in \mathbb{N}$, tome $v_k = \partial^0 f(x_k)$, por (4.1), tem-se que $(x_k, v_k) \rightarrow (\bar{x}, 0)$ e pela proposição 2.6, segue que $0 \in \partial f(\bar{x})$, que é uma contradição.

□

Proposição 4.1. *Sejam $f : M \rightarrow \mathbb{R}$ uma função convexa e $\bar{x} \in M$. Se $0 \notin \partial f(\bar{x})$, então f satisfaz a desigualdade de Kurdyka - Łojasiewicz em \bar{x} .*

Demonstração. Pelo lema 4.1 existe $\delta > 0$ tal que

$$x \in M, d(x, \bar{x}) < \delta \Rightarrow \|\partial^0 f(x)\| \geq \delta.$$

Considere $r_0 > 0$ qualquer e note que a função $\varphi : [0, r_0] \rightarrow \mathbb{R}$ dada por $\varphi(t) = \frac{t}{\delta}$ pertence a $\mathcal{K}(0, r_0)$. Além disso, para cada $x \in B(\bar{x}, \delta) \cap [f(\bar{x}) < f < f(\bar{x}) + r_0]$, tem-se

$$\varphi'(f(x) - f(\bar{x})) \cdot \|\partial^0 f(x)\| = \frac{1}{\delta} \cdot \|\partial^0 f(x)\| \geq \frac{\delta}{\delta} = 1.$$

□

Assim, nosso foco é o caso em que $\bar{x} \in S$, onde $S = \operatorname{argmin} f$. Se $f(\bar{x}) = 0$, a desigualdade KL torna-se

$$\varphi'(f(x)) \cdot \|\partial^0 f(x)\| \geq 1$$

. para todo $x \in B(\bar{x}, \delta) \cap [0 < f < r_0]$.

O próximo resultado mostra uma importante classe de funções que possuem a propriedade KL em uma vizinhança de cada ponto $x \in M$. Esta classe de funções é formada pelas funções fortemente convexas e duas vezes diferenciáveis. Assim, pelo corolário 2.1, a função $\rho_p(q) = \frac{d(p, q)^2}{2}$ possui a propriedade KL localmente em cada $q \in M$.

Proposição 4.2. *Seja $f : M \rightarrow \mathbb{R}$ uma função fortemente convexa e duas vezes diferenciável. Então f satisfaz a desigualdade KL em cada ponto de M , com função desingularizante $\varphi(t) = c\sqrt{t}$ para algum $c > 0$.*

Demonstração. Consultar [12].

□

4.2 Error Bounds

Considere uma função não decrescente $\omega : [0, \infty) \rightarrow [0, \infty)$ com $\omega(0) = 0$. Dizemos que $f : M \rightarrow \mathbb{R}$ satisfaz um error bounds local com função residual ω se existir $r_0 > 0$ tal que

$$(\omega \circ f)(x) \geq \text{dist}(x, S), \quad (4.2)$$

para todo $x \in [0 \leq f \leq r_0]$ (Estamos considerando $\min f = 0$).

Um importante caso particular é quando a função residual é dada por $\omega(s) = \gamma^{-\frac{1}{p}} s^{\frac{1}{p}}$, nesse caso podemos rescrever a desigualdade (4.2) da seguinte forma:

$$f(x) \geq \gamma d(x, S)^p.$$

Definição 4.1. Dizemos que uma função $\varphi : [0, r) \rightarrow \mathbb{R}$ em $C^1(0, r) \cap C^0[0, r_0]$ e nula na origem tem comportamento moderado (próximo a origem) se ela satisfaz a desigualdade

$$s\varphi'(s) \geq c\varphi(s) \quad \forall s \in (0, r),$$

onde c é uma constante positiva.

Observação 4.1. Se φ for côncava, então $c \leq 1$. De fato, seja $s \in (0, r)$ tal que $\varphi(s) \neq 0$, pelo teorema do valor médio existe $t_s \in (0, s)$ tal que $\varphi(s) - \varphi(0) = s\varphi'(t_s)$, logo

$$\varphi(s) = \varphi(s) - \varphi(0) = s\varphi'(t_s) \geq s\varphi'(s) \geq c\varphi(s),$$

donde $c \leq 1$.

Exemplo 4.1. Considere a função $\varphi : [0, r) \rightarrow \mathbb{R}$ dada por $\varphi(s) = \gamma s^{\frac{1}{p}}$. Daí

$$\varphi'(s) = \frac{1}{p} \gamma s^{\frac{1}{p}-1}.$$

Fazendo $c = \frac{1}{p}$, obtemos

$$s\varphi'(s) \geq c\varphi(s), \quad \forall s \in (0, r),$$

ou seja, φ tem comportamento moderado próximo a origem. Além disso, se $\gamma > 0$ e $p \geq 1$, é fácil ver que φ é côncava e $c = \frac{1}{p} \leq 1$.

Quando f é uma função convexa definida em um espaço de Hilbert e satisfaz a desigualdade KL, o teorema 27 em [7] fornece uma estimativa para o comprimento das

trajetórias subgradientes. Utilizando a existência das curvas subgradientes em variedades de Hadamard e algumas propriedades que foram desenvolvidas no capítulo anterior, generalizaremos este resultado para variedades de Hadamard.

Relembramos aqui que estamos assumindo que $S = \arg \min f$ e que $\min f = 0$.

Proposição 4.3. *Sejam $\bar{x} \in S, \rho > 0$ e $\varphi \in \mathcal{K}(0, r_0)$. As seguintes afirmações são equivalentes:*

(i) *Para cada $x \in B(\bar{x}, \rho) \cap [0 < f < r_0]$, temos*

$$\varphi'(f(x)) \|\partial^0 f(x)\| \geq 1.$$

(ii) *Para cada $x \in B(\bar{x}, \rho) \cap [0 < f \leq r_0]$ e $0 \leq t < s$, temos*

$$\int_t^s \|\mathbf{x}'(\tau)\| d\tau \leq \varphi(f(\mathbf{x}(t))) - \varphi(f(\mathbf{x}(s))).$$

onde $\mathbf{x} : [0, \infty] \rightarrow M$ é a curva subgradiente, com $\mathbf{x}(0) = x$.

Demonstração. Primeiro mostraremos que (i) \Rightarrow (ii).

Dado $x \in B(\bar{x}, \rho) \cap [0 < f < r_0]$ e $0 \leq t < s$, considere a curva subgradiente partindo de x (ver proposição 3.1), isto é, uma curva $\mathbf{x} : [0, \infty) \rightarrow M$ absolutamente contínua satisfazendo

$$\begin{cases} -\mathbf{x}'(t) \in \partial f(\mathbf{x}(t)) \\ \mathbf{x}(0) = x, \end{cases}$$

para quase todo $t > 0$. Observe que

$$\begin{aligned} \varphi(f(\mathbf{x}(t))) - \varphi(f(\mathbf{x}(s))) &= \int_s^t \frac{d}{d\tau} \varphi(f(\mathbf{x}(\tau))) d\tau \\ &= \int_s^t \varphi'(f(\mathbf{x}(\tau))) \frac{d}{d\tau} (f(\mathbf{x}(\tau))) d\tau. \end{aligned}$$

Na proposição 3.2, provamos que $\frac{d}{d\tau} (f(\mathbf{x}(\tau))) = -\|\mathbf{x}'(\tau)\|^2$ para quase todo $\tau > 0$, e na proposição 3.3 mostramos que $\mathbf{x}'(\tau) = -\partial^0 f(\mathbf{x}(\tau))$ para quase todo $\tau > 0$, logo

$$\begin{aligned} \varphi(f(\mathbf{x}(t))) - \varphi(f(\mathbf{x}(s))) &= \int_s^t \varphi'(f(\mathbf{x}(\tau))) \cdot (-\|\mathbf{x}'(\tau)\|^2) d\tau \\ &= \int_t^s \varphi'(f(\mathbf{x}(\tau))) \cdot \|\mathbf{x}'(\tau)\|^2 d\tau \\ &= \int_t^s \varphi'(f(\mathbf{x}(\tau))) \cdot \|\partial^0 f(\mathbf{x}(\tau))\| \cdot \|\mathbf{x}'(\tau)\| d\tau. \end{aligned}$$

Note que $\mathbf{x}(t) \in B(\bar{\mathbf{x}}, \rho)$ para todo $t > 0$, pois $\mathbf{x}(0) = \mathbf{x} \in B(\bar{\mathbf{x}}, \rho)$ e a função $t \mapsto d(\mathbf{x}(t), \bar{\mathbf{x}})$ é decrescente (ver proposição 3.5). Além disso, pela proposição 3.2, a função $t \mapsto f(\mathbf{x}(t))$ é decrescente, portanto $\mathbf{x}(t) \in [0 < f < r_0]$ para todo $t > 0$. Assim,

$$\mathbf{x}(t) \in B(\bar{\mathbf{x}}, \rho) \cap [0 < f < r_0] \quad \forall t > 0,$$

dessa forma, por hipótese, tem-se que $\varphi'(f(\mathbf{x}(\tau))) \|\partial^0 f(\mathbf{x}(\tau))\| \geq 1$ para todo $\tau > 0$. Logo,

$$\varphi(f(\mathbf{x}(t))) - \varphi(f(\mathbf{x}(s))) = \int_t^s \varphi'(f(\mathbf{x}(\tau))) \cdot \partial^0 f(\mathbf{x}(\tau)) \cdot \|\mathbf{x}'(\tau)\| d\tau \geq \int_t^s \|\mathbf{x}'(\tau)\| d\tau.$$

Mostraremos agora que (ii) \Rightarrow (i).

Dado $\mathbf{x} \in B(\bar{\mathbf{x}}, \rho) \cap [0 < f < r_0]$, considere a curva subgradiente partindo de \mathbf{x} , segue da hipótese que

$$\frac{1}{h} \int_0^h \|\mathbf{x}'(\tau)\| d\tau \leq -\frac{\varphi(f(\mathbf{x}(h))) - \varphi(f(\mathbf{x}(0)))}{h}, \quad \forall h > 0.$$

Fazendo $h \rightarrow 0^+$, obtemos

$$\|\mathbf{x}'(0^+)\| \leq -\varphi'(f(\mathbf{x}(0))) \cdot (f \circ \mathbf{x})'(0^+).$$

Lembre-se que $(f \circ \mathbf{x})'(t^+) = -\|\mathbf{x}'(t^+)\|^2$ e $\mathbf{x}'(t^+) = -\partial^0 f(\mathbf{x}(t))$ para todo $t \geq 0$ (ver proposição 3.4 e o corolário 3.1), daí

$$\begin{aligned} \|\mathbf{x}'(0^+)\| \leq -\varphi'(f(\mathbf{x})) \cdot (f \circ \mathbf{x})'(0^+) &\Rightarrow \|\partial^0 f(\mathbf{x})\| \leq \varphi'(f(\mathbf{x})) \cdot \|\partial^0 f(\mathbf{x})\|^2 \\ &\Rightarrow 1 \leq \varphi'(f(\mathbf{x})) \cdot \|\partial^0 f(\mathbf{x})\|. \end{aligned}$$

□

O teorema a seguir é o principal resultado desse capítulo. Ele garante que se uma função convexa $f : M \rightarrow \mathbb{R}$ possui a propriedade KL em $\bar{\mathbf{x}}$, então f satisfaz um error bounds em $\bar{\mathbf{x}}$, a recíproca não é verdade, mas se f satisfaz um error bound em $\bar{\mathbf{x}}$ e sua função residual tem comportamento moderado, então f satisfaz a desigualdade KL em $\bar{\mathbf{x}}$. Além disso, no primeiro caso a função desingularizante e a função residual são iguais e no segundo caso diferem apenas por uma constante multiplicativa. Este resultado encontra-se feito em [7], no caso em que f está definida em um espaço de Hilbert. Usando a existência da curva subgradiente, que desenvolvemos no capítulo anterior, também iremos generalizar este resultado para Variedades de Hadamard.

Teorema 4.1 (Caracterização da desigualdade KL para funções convexas).

Sejam $f : M \rightarrow \mathbb{R}$ uma função convexa, com $\min f = 0$, $r_0 > 0$, $c > 0$, $\varphi \in \mathcal{K}(0, r_0)$, $\rho > 0$ e $\bar{x} \in S$. Então

(i) (Desigualdade KL implica error bounds.)

Se $\|\partial^0 f(x)\| \varphi'(f(x)) \geq 1$ para todo $x \in [0 < f < r_0] \cap B(\bar{x}, \rho)$, então

$$\text{dist}(x, S) \leq \varphi(f(x)), \quad \forall x \in [0 < f < r_0] \cap B(\bar{x}, \rho).$$

(ii) (Error bounds e comportamento moderado implica desigualdade KL.)

Se $s\varphi'(s) \geq c\varphi(s)$ para todo $s \in (0, r_0)$ e $\text{dist}(x, S) \leq \varphi(f(x))$ para todo $x \in [0 < f < r_0] \cap B(\bar{x}, \rho)$, então

$$\varphi'(f(x)) \|\partial^0 f(x)\| \geq c, \quad \forall x \in [0 < f < r_0] \cap B(\bar{x}, \rho).$$

Demonstração. (i) Como f tem a propriedade KL em \bar{x} , a proposição 4.3 garante que para cada $x \in [0 < f < r_0] \cap B(\bar{x}, \rho)$ e $0 \leq t < s$, tem-se

$$d(x(s), x(t)) \leq \int_t^s \|\dot{x}(\tau)\| d\tau \leq \varphi(f(x(t))) - \varphi(f(x(s))), \quad (4.3)$$

onde $x : [0, \infty] \rightarrow M$ é a curva subgradiente, com $x(0) = x$. Lembre-se que na proposição 3.8 provamos que $\lim_{t \rightarrow +\infty} x(t) = x^* \in \arg \min f$ e $\lim_{t \rightarrow +\infty} (f \circ x)(t) = \min f = 0$. Assim, tomando $t = 0$ e fazendo $s \rightarrow \infty$ em (4.3), obtemos

$$d(x^*, x) \leq \varphi(f(x)).$$

Portanto,

$$\text{dist}(x, S) \leq \varphi(f(x)), \quad \forall x \in [0 < f < r_0] \cap B(\bar{x}, \rho).$$

(ii) Tome $x \in [0 < f < r_0] \cap B(\bar{x}, \rho)$ e seja $y = P_S(x)$. Por convexidade, temos

$$0 = f(y) \geq f(x) + \langle \partial^0 f(x), \exp_x^{-1} y \rangle.$$

Assim,

$$\begin{aligned} f(x) &\leq \|\langle \partial^0 f(x), \exp_x^{-1} y \rangle\| \leq \|\partial^0 f(x)\| \cdot \|\exp_x^{-1} y\| \\ &= \|\partial^0 f(x)\| \cdot d(x, y) \\ &= \|\partial^0 f(x)\| \cdot d(x, S) \\ &\leq \|\partial^0 f(x)\| \cdot \varphi(f(x)) \\ &= \frac{1}{c} \cdot c\varphi(f(x)) \cdot \|\partial^0 f(x)\| \\ &\leq \frac{1}{c} \cdot f(x)\varphi'(f(x)) \cdot \|\partial^0 f(x)\|. \end{aligned}$$

Observe que na ultima desigualdade usamos que φ tem comportamento moderado. Além disso, note que $f(x) > 0$, pois $x \in [0 < f < r_0]$, portanto

$$\varphi'(f(x))\|\partial^0 f(x)\| \geq c.$$

□

A seguir caracterizaremos a existência global da propriedade KL. A demonstração do caso global é completamente análoga à do teorema anterior.

Corolário 4.1 (Caracterização da desigualdade KL para funções convexas: caso global).

Sejam $f : M \rightarrow \mathbb{R}$ uma função convexa, com $\min f = 0$, $r_0 > 0$, $c > 0$ e $\varphi \in \mathcal{K}(0, r_0)$.

(i) (Desigualdade KL implica error bounds.) Se $\|\partial^0 f(x)\|\varphi'(f(x)) \geq 1$ para todo $x \in [0 < f]$, então $\text{dist}(x, S) \leq \varphi(f(x))$, $\forall x \in [0 < f]$.

(ii) (Error bounds e comportamento moderado implica desigualdade KL.)

Se $s\varphi'(s) \geq c\varphi(s)$ para todo $s \in (0, r_0)$ e $\text{dist}(x, S) \leq \varphi(f(x))$ para todo $x \in [0 < f]$, então $\varphi'(f(x))\|\partial^0 f(x)\| \geq c$, $\forall x \in [0 < f]$.

Observação 4.2. Observe que há uma certa assimetria entre as conclusões de (i) e (ii): Em (i) a função residual é exatamente a função desingularizante, enquanto em (ii), a função residual e a função desingularizante diferem por uma constante multiplicativa.

Capítulo 5

Problema de viabilidade convexa

Um problema muito comum em diversas áreas de matemática e ciências físicas consiste em tentar encontrar um ponto na interseção de conjuntos convexos, mas precisamente, se C_1, C_2, \dots, C_m são conjuntos convexos e fechados tais que $C := \bigcap_{i=1}^m C_i \neq \emptyset$, estamos interessados em encontrar algum ponto $x \in C$, esse problema é conhecido na literatura como problema de viabilidade convexa.

5.1 Método do gradiente

Enunciaremos a seguir o método do gradiente, que também é conhecido na literatura como método de máxima descida. Para mais detalhes sobre este método, consulte o apêndice A, desta dissertação.

Algoritmo 5.1 (Método do gradiente).

Passo 1: Escolha um ponto inicial $x^0 \in M$;

Passo 2: Dado x_k , se x_k é um ponto crítico de f , então faça $x_{k+p} = x_k$ para todo $p \in \mathbb{N}$;

Passo 3: Caso contrário, tome como a próxima iterada $x_{k+1} \in M$ tal que

$$x_{k+1} = \exp_{x_k}(-t_k \text{grad}f(x_k)), \quad (5.1)$$

onde o comprimento de passo t_k é dado pela busca de Armijo's, isto é,

$$t_k := \max\{2^{-j} : j \in \mathbb{N}, f(-\exp_{x_k}(2^{-j} \text{grad}f(x_k))) \leq f(x_k) - \alpha 2^{-j} \|\text{grad}f(x_k)\|^2\},$$

com $\alpha \in (0, 1)$.

Observação 5.1. *A boa definição da igualdade (5.1) segue do Teorema de Cartan-Hadamard, mais precisamente, pelo fato de, em uma variedade de Hadamard, a aplicação exponencial é um difeomorfismo (global). Além disso, segue de (5.1) que*

$$t_k \|\text{grad} f(x_k)\| = d(x_{k+1}, x_k), \quad \forall k \geq 0. \quad (5.2)$$

Se $x_{k+1} = x_k$, segue de (5.1) que $\text{grad} f(x_k) = 0$, uma vez que $t_k > 0$, ou seja, x_k é ponto crítico de f . Com isso, temos um critério de parada prático para o Algoritmo 5.1.

Considere a função $f : M \rightarrow \mathbb{R}$ dada por

$$f(x) = \frac{1}{2} \sum_{i=1}^m \alpha_i \text{dist}^2(x, C_i), \quad (5.3)$$

onde $\alpha_i > 0$ para todo $i = 1, 2, \dots, m$ e $\sum_{i=1}^m \alpha_i = 1$. Note que $x \in C$ se, e somente se, $x \in \arg \min f$, isto é, resolver este problema de viabilidade convexa é equivalente à minimizar a função f .

Para minimizar f , aplicaremos o método do gradiente para variedades de Hadamard, que pode ser encontrado com mais detalhes em [6].

Observe que pela definição da busca de Armijo's, a sequência $\{x_k\}$ gerada pelo algoritmo de máxima descida satisfaz a desigualdade

$$f(x_{k+1}) + \alpha t_k \|\text{grad} f(x_k)\|^2 \leq f(x_k).$$

Em particular, a sequência $f(x_k)$ é estritamente decrescente.

A partir da definição do comprimento de passo t_k , é possível mostrar que se a função objetivo f tem gradiente Lipschitz, então zero não é um ponto de acumulação da sequência $\{t_k\}$.

Quando $\{x_k\}$ tem um ponto de acumulação \bar{x} e f satisfaz a desigualdade KL em \bar{x} , a sequência gerada pelo método do gradiente converge para um ponto crítico de f .

Teorema 5.1. *Seja $\{x_k\}$ a sequencia gerada pelo algoritmo 5.1. Se $\{x_k\}$ tem um ponto de acumulação $\bar{x} \in M$ e f satisfaz a desigualdade de Kurdyka-Lojasiewicz em \bar{x} , então \bar{x} é um ponto crítico de f . Além disso,*

$$\lim_{k \rightarrow +\infty} f(x_k) = f(\bar{x}) \text{ e } \lim_{k \rightarrow +\infty} x_k = \bar{x}.$$

Demonstração. Ver teorema A.1, no apêndice A. □

Finalmente resolveremos o problema de viabilidade convexa, mas para isso precisaremos acrescentar uma hipótese adicional:

Definição 5.1. *Considere os conjuntos $A_1, A_2, \dots, A_m \subset M$ e $A = \bigcap_{i=1}^m A_i$.*

(i) *Dizemos que os conjuntos A_1, A_2, \dots, A_m são linearmente limitadamente regular se para cada conjunto limitado existir uma constante $\kappa > 0$ tal que*

$$\text{dist}(x, A) \leq \kappa \max\{\text{dist}(x, A_i), i = 1, \dots, m\}, \forall x \in S.$$

(ii) *Dizemos que os conjuntos A_1, A_2, \dots, A_m são linearmente regular se existir uma constante $\kappa > 0$ tal que*

$$\text{dist}(x, A) \leq \kappa \max\{\text{dist}(x, A_i), i = 1, \dots, m\}, \forall x \in M.$$

Observação 5.2. *Segue da definição que todo conjunto linearmente limitadamente regular é linearmente regular.*

Teorema 5.2. *Sejam C_1, C_2, \dots, C_m conjuntos convexos e fechados tais que $C := \bigcap_{i=1}^m C_i$ é compacto e não vazio. Considere a sequência $\{x^k\}$ obtida ao aplicar o método do gradiente na função*

$$f(x) = \frac{1}{2} \sum_{i=1}^m \alpha_i \text{dist}^2(x, C_i),$$

onde $\alpha_i > 0$ para todo $i = 1, 2, \dots, m$ e $\sum_{i=1}^m \alpha_i = 1$.

Se os conjuntos C_1, C_2, \dots, C_m são linearmente limitadamente regular, então $\lim_{k \rightarrow +\infty} f(x_k) = 0$ e $\{x_k\}$ converge para algum $\bar{x} \in C$.

Demonstração. Inicialmente mostraremos que f possui um error bounds local. De fato, como que o conjunto $[0 \leq f \leq r_0] = f^{-1}([0, r_0])$ é limitado, temos

$$\text{dist}(x, C) \leq \kappa \max\{\text{dist}(x, C_i), i = 1, \dots, m\}, \forall x \in [0 \leq f \leq r_0].$$

Seja $\alpha = \left(\frac{\min_{i=1, \dots, m} \{\alpha_i\}}{2} \right)^{\frac{1}{2}}$ e $x \in [0 \leq f \leq r_0]$, assim

$$\begin{aligned} \text{dist}(x, C) &\leq \kappa \max\{\text{dist}(x, C_i), i = 1, \dots, m\} \\ &\leq \kappa \sqrt{\text{dist}^2(x, C_1) + \dots + \text{dist}^2(x, C_m)} \\ &= \frac{\kappa}{\alpha} \sqrt{\frac{\min_{i=1, \dots, m} \{\alpha_i\}}{2} (\text{dist}^2(x, C_1) + \dots + \text{dist}^2(x, C_m))} \\ &\leq \frac{\kappa}{\alpha} \sqrt{\frac{\alpha_1 \text{dist}^2(x, C_1) + \dots + \alpha_m \text{dist}^2(x, C_m)}{2}} \\ &= \frac{\kappa}{\alpha} \sqrt{f(x)}. \end{aligned}$$

Assim, f admite um error bound com função residual dada por $\omega(s) = \frac{\kappa}{\alpha} \sqrt{s}$, como vimos no exemplo 4.1, esta função tem comportamento moderado em $(0, \infty)$ com $c = \frac{1}{2}$. Segue do teorema 4.1 que f satisfaz a desigualdade de Kurdyka-Lojasiewicz em cada ponto $x \in [0 < f < r_0]$, com função desingularizante dada por $\varphi(s) = \left(\frac{1}{2}\right)^{-1} \omega(s) = \frac{2\kappa}{\alpha} \sqrt{s}$.

Mostraremos agora que a sequência $\{x_k\}$ é limitada. Com efeito,

$$d(x_k, P_C(x_k)) = \text{dist}(x_k, C) \leq \frac{\kappa}{\alpha} \sqrt{f(x_k)}.$$

Como $\{f(x_k)\}$ é limitada (pois é decrescente), segue que a sequência $\{d(x_k, P_C(x_k))\}$ também é limitada. Além disso, fixado $y_0 \in C$, note que $\{d(P_C(x_k), y_0)\}$ também é limitada, pois C é compacto. Finalmente, pela desigualdade triangular obtemos

$$d(x_k, y_0) \leq d(x_k, P_C(x_k)) + d(P_C(x_k), y_0).$$

Logo a sequência $\{x_k\}$ é limitada e portanto admite algum ponto de acumulação \bar{x} . Pelo teorema 5.1, segue que \bar{x} é um ponto crítico de f e

$$\lim_{k \rightarrow +\infty} f(x_k) = f(\bar{x}) = 0 \text{ e } \lim_{k \rightarrow +\infty} x_k = \bar{x}.$$

□

5.2 Complexidade do método

Na seção anterior, aplicamos o método do gradiente e obtemos uma sequência $\{x_k\}$ que converge para a interseção dos conjuntos C_1, C_2, \dots, C_m , nesta seção estamos interessados em obter uma taxa de convergência para esta sequência, para isso, usaremos um resultado sobre a complexidade do método do gradiente, que pode ser encontrado em [6].

Teorema 5.3. *Assuma que f satisfaz a desigualdade de Kurdyka - Lojasiewicz em um ponto de acumulação \bar{x} de $\{x_k\}$ com função desingularizante $\varphi(t) = \frac{c}{\theta} t^\theta$ com $c > 0$ e $\theta \in (0, 1]$. Então, vale a seguinte estimativa:*

1. Se $\theta = 1$, a sequência $\{x_k\}$ converge em um número finito de passos.
2. Se $\theta \in [\frac{1}{2}, 1)$, então existem constantes $c_1 > 0$ e $c_2 > 0$ tais que

$$f(x_k) - f(\bar{x}) \leq c_2 e^{-c_1 k}.$$

3. Se $\theta \in (0, \frac{1}{2})$, então existe uma constante $c > 0$ e $k_0 \in \mathbb{R}$ tais que

$$f(x_k) - f(\bar{x}) \leq c(k - k_0 - 1)^{\frac{-1}{1-2\theta}}.$$

Demonstração. Veja Teorema 3, em [6]. □

Sob as hipóteses do teorema 5.2, temos a seguinte taxa de convergência para a sequência $\{x_k\}$.

Teorema 5.4 (Taxa de convergência). *Existem constantes $\mu_1 > 0$, $\mu_2 > 0$ e $k_0 \in \mathbb{N}$ tais que*

$$d(x_k, C) \leq \mu_2 e^{-\mu_1 k},$$

para todo $k > k_0$.

Demonstração. Lembre-se que a sequência $\{x_k\}$ é gerada aplicando o método do gradiente na função $f(x) = \frac{1}{2} \sum_{i=1}^m \alpha_i \text{dist}^2(x, C_i)$. Na demonstração do teorema 5.2, vimos que a função desingularizante de f é dada por $\varphi(s) = \frac{2k}{\alpha} \sqrt{s}$, ou seja, $\theta = \frac{1}{2} \in [\frac{1}{2}, 1)$. Pelo teorema 5.3, existem constantes positivas c_1 e c_2 tais que

$$f(x_k) - f(\bar{x}) \leq c_2 e^{-c_1 k}.$$

Como $f(\bar{x}) = \min f = 0$ e $d(x, C) \leq \frac{\kappa}{\alpha} \sqrt{f(x)}$, tem-se

$$\begin{aligned} f(x_k) - f(\bar{x}) \leq c_2 e^{-c_1 k} &\Rightarrow f(x_k) \leq c_2 e^{-c_1 k} \\ &\Rightarrow \sqrt{f(x_k)} \leq \sqrt{c_2} \cdot e^{\frac{-c_1 k}{2}} \\ &\Rightarrow \frac{\kappa}{\alpha} \sqrt{f(x)} \leq \frac{\kappa}{\alpha} \sqrt{c_2} \cdot e^{\frac{-c_1 k}{2}} \\ &\Rightarrow d(x, C) \leq \frac{\kappa}{\alpha} \sqrt{c_2} \cdot e^{\frac{-c_1 k}{2}}. \end{aligned}$$

Tomando $\mu_1 = \frac{c_1}{2}$ e $\mu_2 = \frac{\kappa}{2} \sqrt{c_2}$, temos

$$d(x_k, C) \leq \mu_2 e^{-\mu_1 k}.$$

□

Apêndice A

Análise de convergência do método do gradiente

A.1 Método do gradiente

Nesta seção abordaremos o método do gradiente também presente na literatura como método de máxima descida. Os resultados da primeira parte deste capítulo seguem a mesma ideia de [1, 4]. Foi em 1972 com Luenberger [18] que o método de máxima descida ao longo de geodésicas foi primeiramente estudado. Mas, foi somente em 1993 que Smith [20] obteve os primeiros resultados de convergência desse método no contexto das variedades Riemannianas. Em 2008, Cruz Neto et al.[9] observa a influência da curvatura da variedade para obter a convergência global da sequência gerada pelo método de máxima descida (com o passo de Armijo e passo fixo) para funções continuamente diferenciáveis e convexas definidas em uma variedade Riemanniana completa de dimensão finita com curvatura não-negativa. Esse resultado foi estendido para funções quase-convexas (continuamente diferenciáveis) por Papa Quiroz et al.[19] em 2008. Dessa forma, tem sido comum considerar o contexto das variedades Riemannianas com curvatura não-negativa como o ambiente apropriado para se estudar a convergência global de métodos como gradiente e subgradiente, veja por exemplo [9, 10, 15, 19]. No decorrer dessa seção mostraremos que o método de máxima descida pode ser estendido para as variedades de Hadamard usando a propriedade de Kurdyka-Łojasiewicz.

Considere o problema de minimização irrestrito

$$\min_{x \in M} f(x) \tag{A.1}$$

onde M é uma variedade de Hadamard de dimensão finita e $f : M \rightarrow \mathbb{R}$ é uma função continuamente diferenciável com gradiente Lipschitz de constante $L > 0$. Assuma que o conjunto solução de (A.1) é não-vazio.

O método de máxima descida gera uma sequência da seguinte forma:

Algoritmo A.1. *Passo 1: Escolha $x^0 \in M$ e $\delta_1, \delta_2 > 0$ tal que $L\delta_1 + \delta_2 < 1$;*

Passo 2: Dado $x_k \in M$, se x_k é um ponto crítico de f , então faça $x_{k+p} = x_k$ para todo $p \in \mathbb{N}$;

Passo 3: Caso contrário, tome como próxima iterada $x_{k+1} \in M$ tal que

$$x_{k+1} = \exp_{x_k}(-t_k \text{grad}f(x_k)), \quad (\text{A.2})$$

onde

$$t_k \in \left(\delta_1, \frac{2}{L}(1 - \delta_2) \right). \quad (\text{A.3})$$

Observação A.1. *Lembre-se que a boa definição da igualdade (A.2) segue do Teorema de Cartan-Hadamard, mais precisamente, pelo fato de, em uma variedade de Hadamard, a aplicação exponencial é um difeomorfismo (global). Além disso, segue de (A.2) que*

$$t_k \|\text{grad}f(x_k)\| = d(x_{k+1}, x_k), \quad \forall k \geq 0. \quad (\text{A.4})$$

Se $x_{k+1} = x_k$, segue de (A.4) que $\text{grad}f(x_k) = 0$, uma vez que $t_k > 0$, ou seja, x_k é ponto crítico de f . Com isso, temos um critério de parada prático para o Algoritmo A.1.

Proposição A.1. *Seja $\{x_k\}$ uma sequência gerada pelo Algoritmo A.1. Se $x_{k+1} = x_k$, então o algoritmo para em x_k e x_k é um ponto crítico de f .*

Observação A.2. *Note que em (A.3) o passo pode variar desde que permaneça no intervalo determinado. Mesmo assim, esse passo é denominado na literatura como passo fixo. Nos resultados desta seção iremos considerar o método de máxima descida com o passo fixo como em (A.3), mas ressaltamos que resultados similares podem ser obtidos com o passo clássico conhecido como “busca de Armijo” e definido da seguinte forma:*

$$t_k := \arg \max_t \left\{ f(\exp_{x_k}(-t \text{grad}f(x_k))) \leq f(x_k) - \alpha t \|\text{grad}f(x_k)\|^2, t = 2^{-j} : j \in \mathbb{N} \right\},$$

com $\alpha \in (0, 1)$. Dessa forma, quando a função objetivo f tem gradiente Lipschitz, temos que zero não é um ponto de acumulação da sequência $\{t_k\}$. Além disso, pela definição da

busca de Armijo, a sequência gerada pelo método de máxima descida satisfaz

$$f(x_{k+1}) + \alpha t_k^2 \|\text{grad} f(x_k)\|^2 \leq f(x_k), \quad \forall k \geq 0,$$

que é uma desigualdade importante para a convergência do método, conforme veremos na análise de convergência.

A seguir, apresentamos um resultado que mostra que o Algoritmo A.1 satisfaz a desigualdade acima sem a necessidade de nenhum procedimento adicional, como por exemplo a busca de Armijo. A proposição seguinte é a motivação da escolha do passo como em (A.3) ao invés da busca de Armijo.

Proposição A.2. *Seja $\{x_k\}$ uma sequência gerada pelo Algoritmo A.1. Então, existe uma constante $\beta > 0$ tal que*

$$f(x_{k+1}) + \beta t_k^2 \|\text{grad} f(x_k)\|^2 \leq f(x_k), \quad \forall k \geq 0. \quad (\text{A.5})$$

Em particular, vale um dos seguintes itens:

1. $\{x_k\}$ para em um ponto crítico;
2. a sequência $\{f(x_k)\}$ é estritamente decrescente.

Demonstração. Veja [9, Teorema 5.1]. □

Observação A.3. *De uma forma mais geral, a proposição anterior foi provada em [9] apenas supondo f continuamente diferenciável com gradiente Lipschitz e sem nenhuma hipótese na curvatura da variedade. Porém, para obter os resultados de convergência parcial (no sentido de que pontos de acumulação de $\{x_k\}$ são pontos críticos de f) os autores assumem que os conjuntos de níveis de f são compactos. Substituindo essa hipótese pela convexidade de f juntamente com a não negatividade da curvatura da variedade, eles provam a convergência da sequência para um minimizador de f .*

A seguir, apresentamos um resultado técnico que será usado na análise de convergência do método.

Lema A.1. *Seja $\{a_k\}$ uma sequência de números reais positivos tal que*

$$\sum_{k=1}^{+\infty} a_k^2 / a_{k-1} < +\infty.$$

Então, $\sum_{k=1}^{+\infty} a_k < +\infty$.

Demonstração. Veja [5, Lema 4.1]. □

A.2 Análise de convergência

Recentemente, Bento et al. [5] provaram a convergência do método do ponto proximal para variedades de Riemann sem nenhuma hipótese sobre a curvatura desde que a função objetivo satisfaça a propriedade de Kurdyka-Lojasiewicz. Nesta seção substituiremos a hipótese de convexidade da função objetivo analisada em [9] para convergência do método de máxima descida em uma variedade de Riemann com curvatura não negativa pela propriedade de Kurdyka-Lojasiewicz para analisar a convergência do Algoritmo A.1 em variedades de Hadamard.

No decorrer desta seção os termos φ , η e \mathbf{U} são os mesmo que aparecem na definição da desigualdade KL. Além disso, nos resultados que seguem iremos supor que a sequência $\{\mathbf{x}_k\}$ gerada pelo Algoritmo A.1 não satisfaz o critério de parada, ou seja, $\{\mathbf{x}_k\}$ é uma sequência de infinitos termos tais que $\mathbf{x}_{k+1} \neq \mathbf{x}_k$ e $\text{grad } f(\mathbf{x}_k) \neq 0$ para todo $k \geq 0$, pois caso contrário os resultados são triviais.

Proposição A.3. *Seja $\{\mathbf{x}_k\}$ uma sequência gerada pelo Algoritmo A.1. Suponha que $\{\mathbf{x}_k\}$ tenha um ponto de acumulação $\tilde{\mathbf{x}} \in \mathbf{M}$ e f satisfaz a propriedade de Kurdyka-Lojasiewicz no ponto $\tilde{\mathbf{x}}$. Então, dada uma constante $\mathbf{c} > 0$, para todo $\rho > 0$ existe $k_0 \in \mathbb{N}$ tal que*

$$f(\tilde{\mathbf{x}}) < f(\mathbf{x}_k) < f(\tilde{\mathbf{x}}) + \eta, \quad \forall k \geq k_0 \quad (\text{A.6})$$

e

$$\mathbf{d}(\mathbf{x}_{k_0}, \tilde{\mathbf{x}}) + 2\mathbf{d}(\mathbf{x}_{k_0+1}, \mathbf{x}_{k_0}) + \mathbf{c}\varphi(f(\mathbf{x}_{k_0}) - f(\tilde{\mathbf{x}})) < \rho. \quad (\text{A.7})$$

Demonstração. Seja $\{\mathbf{x}_{k_j}\}$ uma subsequência de $\{\mathbf{x}_k\}$ convergindo para $\tilde{\mathbf{x}}$. Segue da continuidade de f que $\{f(\mathbf{x}_{k_j})\}$ converge para $f(\tilde{\mathbf{x}})$. Além disso, segue de (A.5) que

$$f(\mathbf{x}_{k+1}) + \beta\delta_1^2 \|\text{grad } f(\mathbf{x}_k)\|^2 \leq f(\mathbf{x}_k), \quad \forall k \geq 0, \quad (\text{A.8})$$

onde usamos que $t_k > \delta_1$. Daí, segue que $\{f(\mathbf{x}_k)\}$ é uma sequência estritamente decrescente e sendo o conjunto solução de (A.1) não vazio, temos que $\{f(\mathbf{x}_k)\}$ é convergente e, como $f(\mathbf{x}_{k_j}) \rightarrow f(\tilde{\mathbf{x}})$ quando $j \rightarrow +\infty$, temos que $f(\mathbf{x}_k) \rightarrow f(\tilde{\mathbf{x}})$ quando $k \rightarrow +\infty$. Logo, para todo $\eta > 0$, existe $k_0 \in \mathbb{N}$ tal que

$$f(\tilde{\mathbf{x}}) < f(\mathbf{x}_k) < f(\tilde{\mathbf{x}}) + \eta, \quad \forall k \geq k_0.$$

Agora, combinando (A.2) com (A.8), obtemos

$$2\mathbf{d}(\mathbf{x}_{k+1}, \mathbf{x}_k) < \frac{2}{\delta_1\sqrt{\beta}} \sqrt{f(\mathbf{x}_k) - f(\mathbf{x}_{k+1})}, \quad \forall k \geq 0.$$

Sendo $\{f(\mathbf{x}_k)\}$ convergente, segue da desigualdade acima que $d(\mathbf{x}_{k+1}, \mathbf{x}_k) \rightarrow 0$ quando $k \rightarrow +\infty$. Com isso, levando em consideração que φ é contínua, temos que

$$d(\mathbf{x}_{k_j}, \tilde{\mathbf{x}}) + 2d(\mathbf{x}_{k_{j+1}}, \mathbf{x}_{k_j}) + c\varphi(f(\mathbf{x}_{k_j}) - f(\tilde{\mathbf{x}})) \rightarrow 0,$$

quando $j \rightarrow +\infty$. Portanto, podemos assumir (tomando o máximo dos índices se necessário) que

$$d(\mathbf{x}_{k_0}, \tilde{\mathbf{x}}) + 2d(\mathbf{x}_{k_0+1}, \mathbf{x}_{k_0}) + c\varphi(f(\mathbf{x}_{k_0}) - f(\tilde{\mathbf{x}})) < \rho,$$

e isso conclui a demonstração. \square

Proposição A.4. *Seja $\{\mathbf{x}_k\}$ uma sequência gerada pelo Algoritmo A.1. Se as hipóteses da Proposição A.3 são satisfeitas e existem $k_0 \in \mathbb{N}$ e $\epsilon > 0$ tais que $\mathbf{x}_{k_0} \in \mathcal{B}(\tilde{\mathbf{x}}, \epsilon)$, então existem $\alpha, \beta > 0$ tais que*

$$\frac{d^2(\mathbf{x}_{k_0+1}, \mathbf{x}_{k_0})}{d(\mathbf{x}_{k_0}, \mathbf{x}_{k_0-1})} \leq \frac{\alpha}{\beta} [\varphi(f(\mathbf{x}_{k_0}) - f(\tilde{\mathbf{x}})) - \varphi(f(\mathbf{x}_{k_0+1}) - f(\tilde{\mathbf{x}}))]. \quad (\text{A.9})$$

Demonstração. Assuma, sem perda de generalidade, que $k_0 \in \mathbb{N}$ é tal que (A.6) se verifica e $\epsilon > 0$ é suficientemente pequeno tal que $\mathcal{B}(\tilde{\mathbf{x}}, \epsilon) \subset \mathcal{U}$. Assim,

$$\mathbf{x}_{k_0} \in \mathcal{U} \cap [f(\tilde{\mathbf{x}}) < f < f(\tilde{\mathbf{x}}) + \eta].$$

Como f satisfaz a propriedade de Kurdyka-Lojasiewicz em $\tilde{\mathbf{x}}$, temos que

$$\varphi'(f(\mathbf{x}_{k_0}) - f(\tilde{\mathbf{x}})) \geq \frac{1}{\|\text{grad}f(\mathbf{x}_{k_0})\|}. \quad (\text{A.10})$$

Como φ é côncava e $\varphi' > 0$, lembrando que $f(\mathbf{x}_{k_0+1}) \leq f(\mathbf{x}_{k_0})$, obtemos

$$\begin{aligned} \varphi(f(\mathbf{x}_{k_0}) - f(\tilde{\mathbf{x}})) - \varphi(f(\mathbf{x}_{k_0+1}) - f(\tilde{\mathbf{x}})) &\geq \varphi'(f(\mathbf{x}_{k_0}) - f(\tilde{\mathbf{x}}))(f(\mathbf{x}_{k_0}) - f(\mathbf{x}_{k_0+1})) \\ &\geq \beta t_{k_0}^2 \frac{\|\text{grad}f(\mathbf{x}_{k_0})\|^2}{\|\text{grad}f(\mathbf{x}_{k_0})\|} \\ &\geq \beta \frac{d^2(\mathbf{x}_{k_0+1}, \mathbf{x}_{k_0})}{\|\text{grad}f(\mathbf{x}_{k_0})\|}, \end{aligned} \quad (\text{A.11})$$

onde a segunda desigualdade segue de (A.10) e (A.5), e a terceira vem de (A.2). Além disso, sendo a aplicação $\mathbf{x} \mapsto \text{grad}f(\mathbf{x})$ Lipschitz contínua e o transporte paralelo uma isometria, segue da desigualdade triangular que

$$\begin{aligned} \|\text{grad}f(\mathbf{x}_{k_0})\| &\leq \|\text{grad}f(\mathbf{x}_{k_0}) - \mathbf{P}_{\mathbf{x}_{k_0-1}, \mathbf{x}_{k_0}} \text{grad}f(\mathbf{x}_{k_0-1})\| \\ &\quad + \|\mathbf{P}_{\mathbf{x}_{k_0-1}, \mathbf{x}_{k_0}} \text{grad}f(\mathbf{x}_{k_0-1})\| \\ &\leq \left(L + \frac{1}{\delta_1}\right) d(\mathbf{x}_{k_0}, \mathbf{x}_{k_0-1}). \end{aligned} \quad (\text{A.12})$$

Portanto, o resultado desejado segue combinando as inequações (A.12) e (A.11), com $\alpha = (L + \frac{1}{\delta_1})$. \square

Proposição A.5. *Seja $\{x_k\}$ uma sequência gerada pelo algoritmo A.1. Se as hipóteses da proposição A.3 são satisfeitas. Então, existem $k_0 \in \mathbb{N}$ e $\epsilon > 0$ tais que*

$$x_k \in B(\tilde{x}, \epsilon), \quad \forall k \geq k_0. \quad (\text{A.13})$$

Demonstração. Seja $\{x_{k_j}\}$ uma subsequência de $\{x_k\}$ convergindo para \tilde{x} e tome $\epsilon > 0$ suficientemente pequeno tal que $B(\tilde{x}, \epsilon) \subset \mathcal{U}$. Segue da Proposição A.3, para $\rho = \epsilon$, que existe $k_0 \in \mathbb{N}$ tal que (A.7) se verifica. Provaremos (A.13) por indução sobre k . Para $k = k_0$, a afirmação segue diretamente de (A.7). Agora, suponha que $x_k \in B(\tilde{x}, \epsilon)$, para todo $k = k_0 + 1, \dots, k_0 + j - 1$. Então, pela proposição A.4, obtemos

$$\sqrt{d(x_k, x_{k-1})(\alpha/\beta)[\varphi(f(x_k) - f(\tilde{x})) - \varphi(f(x_{k+1}) - f(\tilde{x}))]} \geq d(x_{k+1}, x_k), \quad (\text{A.14})$$

para $k = k_0 + 1, \dots, k_0 + j - 1$. Note que, para quaisquer números reais $r, s \geq 0$, temos que $r + s \geq 2\sqrt{rs}$. Assim, para $k = k_0 + 1, \dots, k_0 + j - 1$, tomando

$$r = d(x_k, x_{k-1}) \text{ e } s = (\alpha/\beta)[\varphi(f(x_k) - f(\tilde{x})) - \varphi(f(x_{k+1}) - f(\tilde{x}))],$$

pela desigualdade (A.14), segue que

$$2d(x_{k+1}, x_k) \leq d(x_k, x_{k-1}) + \frac{\alpha}{\beta}[\varphi(f(x_k) - f(\tilde{x})) - \varphi(f(x_{k+1}) - f(\tilde{x}))],$$

para $k = k_0 + 1, \dots, k_0 + j - 1$. Somando a última desigualdade de $k = k_0 + 1$ até $k = k_0 + j - 1$, obtemos

$$\begin{aligned} \sum_{i=k_0+1}^{k_0+j-1} d(x_{i+1}, x_i) + d(x_{k_0+j}, x_{k_0+j-1}) &\leq d(x_{k_0+1}, x_{k_0}) \\ &+ \frac{\alpha}{\beta}[\varphi(f(x_{k_0+1}) - f(\tilde{x}))] \\ &- \frac{\alpha}{\beta}[\varphi(f(x_{k_0+j}) - f(\tilde{x}))], \end{aligned}$$

que nos leva a concluir que

$$\sum_{i=k_0+1}^{k_0+j-1} d(x_{i+1}, x_i) \leq d(x_{k_0+1}, x_{k_0}) + \frac{\alpha}{\beta}[\varphi(f(x_{k_0}) - f(\tilde{x}))] \quad (\text{A.15})$$

porque φ é crescente, $f(x_{k+1}) \leq f(x_k)$, para todo $k \geq 0$, e $d(x, y) \geq 0$ para todo $x, y \in M$. Agora, usando a desigualdade triangular, temos que

$$\begin{aligned} d(x_{k_0+j}, \tilde{x}) &\leq d(x_{k_0+j}, x_{k_0}) + d(x_{k_0}, \tilde{x}) \\ &\leq d(x_{k_0}, \tilde{x}) + d(x_{k_0+1}, x_{k_0}) + \sum_{i=k_0+1}^{k_0+j-1} d(x_{i+1}, x_i), \end{aligned}$$

que, combinado com (A.15), obtemos

$$d(x_{k_0+j}, \tilde{x}) \leq d(x_{k_0}, \tilde{x}) + 2d(x_{k_0+1}, x_{k_0}) + \frac{\alpha}{\beta} \varphi(f(x_{k_0}) - f(\tilde{x})) < \epsilon,$$

onde a desigualdade do lado direito vem da Proposição A.3 para $\rho = \epsilon$. Portanto, concluímos que $x_{k_0+j} \in B(\tilde{x}, \epsilon)$, e a indução está completa. \square

Proposição A.6. *Seja $\{x_k\}$ uma sequência gerada pelo Algoritmo A.1. Se as hipóteses da Proposição A.3 são satisfeitas, então $\sum_{k=0}^{+\infty} d(x_{k+1}, x_k) < +\infty$. Em particular,*

$$\lim_{k \rightarrow +\infty} d(x_{k+1}, x_k) = 0.$$

Demonstração. Tome $k_0, N \in \mathbb{N}$ números naturais tais que $N > k_0$ e (A.13) se verifica. Então, combinando a Proposição 3.2 com a Proposição A.5, temos que

$$\frac{d^2(x_{k+1}, x_k)}{d(x_k, x_{k-1})} \leq \frac{\alpha}{\beta} [\varphi(f(x_k) - f(\tilde{x})) - \varphi(f(x_{k+1}) - f(\tilde{x}))], \quad \forall k \geq k_0, \quad (\text{A.16})$$

que implica em

$$\sum_{i=k_0+1}^N \frac{d^2(x_{i+1}, x_i)}{d(x_i, x_{i-1})} \leq \frac{\alpha}{\beta} [\varphi(f(x_{k_0+1}) - f(\tilde{x}))], \quad (\text{A.17})$$

porque $\varphi' > 0$ e $f(x_k) \leq f(\tilde{x})$, para todo $k \geq 0$. Portanto, o resultado desejado segue fazendo $k \rightarrow +\infty$ em (A.17) e aplicando o Lema A.1. \square

A maioria dos resultados de convergência de métodos numéricos em minimização convexa são obtidos supondo a existência de um minimizador local (e com isso global) da função objetivo. No presente cenário não-convexo, a existência de um minimizador local não implica a existência de um minimizador global mesmo se a função em consideração for limitada inferiormente. Dessa forma, não é esperado se obter limitação da sequência das iteradas.

Teorema A.1. *Seja $\{x_k\}$ uma sequência gerada pelo Algoritmo A.1. Assuma que $\{x_k\}$ tem um ponto de acumulação $\tilde{x} \in M$ e f satisfaz a propriedade de Kurdyka-Lojasiewicz em \tilde{x} . Então, $\lim_{k \rightarrow +\infty} f(x_k) = f(\tilde{x})$ e $\{x_k\}$ converge para \tilde{x} que é um ponto crítico de f .*

Demonstração. Seja $\{x_{k_j}\}$ uma subsequência de $\{x_k\}$ convergindo para \tilde{x} . Da Proposição A.6 segue que $\{x_k\}$ é uma sequência de Cauchy, e com isso, $\{x_k\}$ converge para \tilde{x} quando $k \rightarrow +\infty$. Sendo f contínua, temos que $f(x_k)$ converge para $f(\tilde{x})$ quando $k \rightarrow +\infty$. Além disso, de (A.2), temos que

$$t_k \|\text{grad}f(x_k)\| = d(x_{k+1}, x_k), \quad \forall k \geq 0.$$

Aplicando o limite com $k \rightarrow +\infty$ na igualdade acima obtemos

$$\|\text{grad}f(\tilde{x})\| = 0,$$

uma vez que $\liminf_{k \rightarrow +\infty} t_k > 0$ e $\lim_{k \rightarrow +\infty} d(x_{k+1}, x_k) = 0$. Portanto, \tilde{x} é um ponto crítico de f e a demonstração está concluída. \square

Referências Bibliográficas

- [1] Attouch, H., Bolte, J., Svaiter, B.F. - *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*. Math. Program., 137(1-2), 91–129, 2013.
- [2] Bacák, M. - *Convex analysis and optimization in Hadamard spaces*, in: De Gruyter Series in Nonlinear Analysis and Applications 22 - Walter de Gruyter, Berlin/Boston, 2014.
- [3] Bauschke, H. H., Borwein, J. M. - *On projection algorithms for solving convex feasibility problems* . Siam Review, 38 (3), 367–426, 1996.
- [4] Bento, G.C., Cruz Neto, J.X., Oliveira, P.R. - *Convergence of inexact descent methods for nonconvex optimization on Riemannian manifolds* . Preprint, 2011.
- [5] Bento, G.C., Cruz Neto, J.X., Oliveira, P.R. - *A new approach to the proximal point method: convergence on general Riemannian manifolds* . J. Optim. Theory Appl., 168, 743–755, 2016.
- [6] Souza, J. C., Bento, G.C., Cruz Neto, J.X., Oliveira, P. R., *Convergence of the steepest descent method for computing Riemannian center of mass on Hadamard manifolds, Preprint*, 2016.
- [7] Bolte, J., Nguyen, T.P., Peypouquet, J., Suter, B.W. - *From error bounds to the complexity of first-order descent methods for convex functions*. Math. Program., 1–37, 2016. doi:10.1007/s10107-016-1091-6
- [8] Burtscher, A.Y. - *Length structures on manifolds with continuous Riemannian metrics*. New York J. Math., 21, 273–296, 2015.

- [9] Cruz Neto, J.X., Lima, L.L., Oliveira, P.R. - *Geodesic algorithms in Riemannian geometry*. Balkan J. Geom. Appl., 3(2), 89–100, 1998.
- [10] Cruz Neto, J.X., Ferreira, O.P., Lucambio Perez, L.R. - *A proximal regularization of the steepest descent method in Riemannian manifold*. Balkan J. Geom. Appl., 4(2), 1–8, 1999.
- [11] Cruz Neto, J.X., Ferreira, O.P., Lucâmbio Pérez, L.R. - *Contribution to the study of monotone vector fields*. Acta Math. Hungar., 94, 307–320, 2002.
- [12] Cruz Neto, J.X., Oliveira, P.R., Soares Jr., P.A., Soubeyran, A. - *Learning how to play Nash, potential games and alternating minimization method for structured nonconvex problems on Riemannian manifolds*. J. Convex Anal., 20, 395–438, 2013.
- [13] Cruz Neto, J.X., Melo, I.D., Sousa, P.A., Souza, J.C.O. - *Solving Convex Feasibility Problems in Hadamard spaces*, Preprint, 2019.
- [14] do Carmo, M. P. - *Geometria Riemanniana*. Projeto Euclides - IMPA, 2005.
- [15] Ferreira, O.P., Oliveira, P.R. - *Subgradient algorithm on Riemannian manifolds*. J. Optim. Theory Appl., 97, 93–104, 1998.
- [16] Lang, S. - *Fundamentals of differential geometry*. Graduate Texts in Mathematics - Springer-Verlag, New York, 191, 1999.
- [17] Li, C., López, G., Martín-Márquez, V. - *Monotone vector fields and the proximal point algorithm on Hadamard manifolds*. J. Lond. Math. Soc. (2), 79(3):663-683, 2009.
- [18] Luenberger, D.G. - *The gradient projection method along geodesics*. Management Science, 18(11), 620–631, 1972.
- [19] Papa Quiroz, E.A., Quispe, E.M., Oliveira, P.R. - *Steepest descent method with a generalized Armijo search for quasiconvex functions on Riemannian manifolds*. J. Math. Anal. Appl., 341, 467–477, 2008.
- [20] Smith, S.T. - *Optimization techniques on Riemannian manifolds*. Fields institute communications, 3(3), 113–135, 1994.
- [21] Udriste, C. - *Convex functions and optimization algorithms on Riemannian manifolds*. Mathematics and its applications - Kluwer Academic, Dordrecht, 297, 1994.