

**PROGRAMAÇÃO MATEMÁTICA EM VARIEDADES RIEMANNIANAS:
ALGORITMOS SUBGRADIENTE E PONTO PROXIMAL**

Orizon Pereira Ferreira

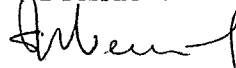
TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO EM ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS
PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS
EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO

Aprovada por:



Prof. Paulo Roberto Oliveira, Dr.Ing.

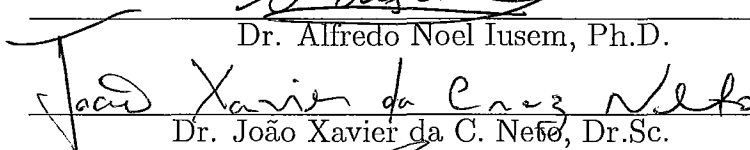
Presidente



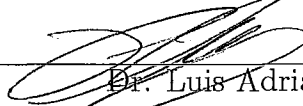
Profa. Susana S. de Makler, Dr.Sc.



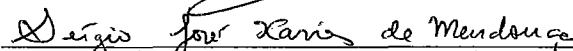
Dr. Alfredo Noel Iusem, Ph.D.



Dr. João Xavier da C. Neto, Dr.Sc.



Dr. Luis Adrián Florit, Ph.D.



Dr. Sérgio J. Xavier de Mendonça, Dr.Sc.

RIO DE JANEIRO, RJ - BRASIL
MARÇO DE 1997

FERREIRA, ORIZON PEREIRA

Programação Matemática em Variedades

Riemannianas: Algoritmos subgradiente

e ponto proximal [Rio de Janeiro] 1997.

VII, 81pp, 29,7 cm (COPPE/UFRJ, D.Sc.,

Engenharia de Sistemas e Computação, 1997)

Tese -- Universidade Federal do Rio de Janeiro, COPPE

1. Programação Matemática

I. COPPE/UFRJ II. Título (série).

À minha esposa Deller e ao
meu filho Alexandre

Agradecimentos

Primeiro agradeço meu orientador Paulo Roberto Oliveira pela sua orientação na escolha do tema desta tese e pelas palavras de estímulo durante a sua elaboração.

Agradeço os amigos Quirino, Xavier e Alfredo pelas suas companhias, importantes nas horas de incerteza.

Agradeço a oportunidade de participar do seminário de otimização do IMPA e os amigos feitos lá, especialmente Benar, Iusem, Luis, Mauricio e Regina pelas suas paciências em ouvir e pelas sugestões dadas as primeiras idéias deste trabalho.

Agradeço a banca examinadora por ler esta tese e pelas sugestões dadas que contribuíram bastante para melhoria dela.

Agradeço aos colegas do Departamento de Matemática da UFG que assumiram minha carga horária durante o tempo que estive realizando o doutorado.

Quero agradecer também a minha esposa Deller e ao meu filho Alexandre os maiores responsáveis pelo meu empenho em realizar este curso.

E por último agradeço ao Rogerio e ao Wilson por ter datilografado este trabalho.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências (D.Sc.)

PROGRAMAÇÃO MATEMÁTICA EM VARIEDADES RIEMANNIANAS: ALGORITMOS SUBGRADIENTE E PONTO PROXIMAL

Orizon Pereira Ferreira
Março, 1997

Orientador: Paulo Roberto Oliveira
Programa: Engenharia de Sistemas e Computação

Os algoritmos subgradiente e de ponto proximal para minimizar uma função convexa são generalizados para o contexto de variedades Riemannianas. Suas análises de convergências são feitas e obtém-se os mesmos resultados do \mathbb{R}^n . Também é generalizado o algoritmo de ponto proximal para encontrar zeros de operadores monótonos para o contexto de variedades Riemannianas mas, neste caso, a análise de convergência é feita apenas para o caso C^1 .

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

**MATHEMATICAL PROGRAMMING ON RIEMANNIAN MANIFOLDS:
ALGORITHMS SUBGRADIENT AND PROXIMAL POINT**

Orizon Pereira Ferreira
March, 1997

Thesis Supervisors: Paulo Roberto Oliveira
Department: Systems and Computation Engineering

The subgradient and proximal point algorithms to minimize convex functions are generalized to the context of Riemannian manifolds. Their convergence analysis are made and the same results as for \mathbb{R}^n are obtained. The proximal point algorithm to provide zeroes approximations of monotone operators is also generalized to the context of Riemannian manifolds, but in this case, the convergence analysis was made for the C^1 case.

Índice

I	Introdução	3
II	Variedades Riemannianas	5
II.1	Introdução	5
II.2	Variedades diferenciáveis	5
II.3	Métrica Riemanniana	6
II.4	Conexão Riemanniana	7
II.5	Geodésicas e aplicação exponencial	8
II.6	Curvatura	10
II.7	Campos de Jacobi e Fórmulas da Variação	10
II.8	Variedades de Hadamard	11
II.8.1	A função distância	12
II.9	O Teorema de Toponogov	14
III	Análise convexa	17
III.1	Introdução	17
III.2	Conjuntos convexos	18
III.3	Funções convexas	23

III.4 Derivada direcional de funções convexas	29
IV Campos monótonos	39
IV.1 Introdução	39
IV.2 Campos monótonos contínuos	40
IV.3 Campos monótonos ponto-conjunto	52
V Algoritmos para otimização	57
V.1 Introdução	57
V.2 Algoritmo subgradiente	58
V.2.1 Resultados preliminares	58
V.2.2 Convergência	62
V.2.3 Observações finais	63
V.3 Algoritmo de ponto proximal para otimização	64
V.3.1 Boa definição	64
V.3.2 Convergência	66
V.3.3 Observações finais	68
VI Algoritmo de ponto proximal	69
VI.1 O algoritmo para campos C^1	69
VI.1.1 Boa definição da sequência proximal	70
VI.1.2 Convergência	72

CAPÍTULO I

Introdução

É de certo modo natural a extensão dos conceitos e técnicas da Programação Matemática em \mathbb{R}^n para variedades Riemannianas. Isto tem sido feito com certa frequência nos últimos anos com objetivos teóricos e também com o de obter algoritmos efetivos; veja [14], [25], [29], [35]-[37], [45], [48], [50], [52] e [53]; daí vem a motivação de nosso trabalho, o qual está dividido em seis capítulos.

No segundo capítulo fazemos um breve resumo dos conceitos básicos de Geometria Riemanniana, necessários ao desenvolvimento dos capítulos seguintes, os quais se encontram basicamente em do Carmo [11].

No terceiro capítulo fazemos um resumo dos conceitos relativos a conjuntos e funções convexas em variedades Riemannianas necessárias aos três últimos capítulos. Estes conceitos serão discutidos com um certo detalhe, apesar deles já serem todos conhecidos (talvez nem todos). Fizemos esta opção devido ao fato que eles são pouco utilizados em Programação Matemática. As referências para este capítulo são: [1], [2], [5], [8], [9], [12], [15]-[23], [26], [27], [32], [35], [36], [38], [41]-[44] e [46]-[54].

A partir do quarto capítulo começa, de fato, nossa contribuição. Neste capítulo in-

Introduzimos o conceito de campos monótonos em Variedades Riemannianas. Mostramos que os campos monótonos estão de certa maneira intimamente ligados às funções convexas de classe C^1 , por exemplo, mostramos que uma função de classe C^1 é convexa se, e somente se, seu gradiente é um campo monótono. Introduzimos também o conceito de campo monótono ponto-conjunto em Variedades Riemannianas e mostramos que o subdiferencial de uma função convexa é um campo monótono ponto-conjunto.

No quinto capítulo propomos dois algoritmos para minimizar uma função convexa, não necessariamente diferenciável, numa variedade Riemanniana. Estes algoritmos são generalizações do algoritmo subgradiente clássico, introduzido por Shor [44], e do algoritmo de ponto proximal, introduzido por Martinet [31]. Fazemos suas análises de convergências e obtemos os mesmos resultados do \mathbb{R}^n , mais precisamente, provamos que a seqüência gerada pelo algoritmo subgradiente converge para o ínfimo da função se a variedade for completa e tiver curvatura seccional não-negativa, que o algoritmo de ponto proximal está bem definido e que a seqüência gerada por ele converge para um minimizador da função, caso exista algum, se a variedade for de Hadamard.

Finalmente, no último capítulo, propomos um algoritmo para encontrar singularidades de campos monótonos de classe C^1 . Este algoritmo é uma generalização do algoritmo de ponto proximal, introduzido por Rockafellar [40], para encontrar zeros de operadores monótonos. Mostramos que este algoritmo está bem definido e gera uma seqüência que converge para uma singularidade do campo, caso exista alguma, se a variedade for de Hadamard.

CAPÍTULO II

Variedades Riemannianas: um resumo dos conceitos básicos

II.1 Introdução

Neste capítulo serão fixadas algumas notações, e enumerados alguns resultados gerais das variedades Riemannianas, dando ênfase às variedades de Hadamard. Incluiremos as demonstrações dos resultados que desejarmos dar ênfase, em particular alguns resultados clássicos sobre a função distância. O conteúdo deste capítulo pode ser encontrado em do Carmo [11], as vezes com mudanças na notação.

II.2 Variedades diferenciáveis

Seja M uma variedade diferenciável e conexa (daqui para frente omitiremos a palavra conexa, pois só trabalharemos com variedades conexas). O *espaço tangente* a M em p é denotado por T_pM e $TM = \bigcup_{p \in M} T_pM$ denota o *fibrado tangente* de M . Um *campo de vetores* X em M

de classe C^r , $r \geq 0$, é uma aplicação $X: M \rightarrow TM$, dada por $p \mapsto X_p \in T_pM$, de classe C^r . Denotamos o espaço dos campos de vetores em M de classe C^r por $\mathfrak{X}^r(M)$.

II.3 Métrica Riemanniana

Seja M uma variedade diferenciável de dimensão n . Para todo ponto $p \in M$, denotaremos por g ou $\langle \cdot, \cdot \rangle$ uma *métrica Riemanniana* de M . Assim, para cada ponto $p \in M$, g_p ou $\langle \cdot, \cdot \rangle_p$ denotará um produto interno (as vezes omitiremos o ponto p , e isto não causará confusão) em T_pM que varia diferenciavelmente com p . Chamamos o par $(M, \langle \cdot, \cdot \rangle)$ de *variedade Riemanniana*.

Definição II.3.1. Sejam M uma variedade Riemanniana e $f: M \rightarrow \mathbb{R}$ uma função de classe C^1 . O *gradiente* de f é o único campo $\text{grad } f \in \mathfrak{X}^0(M)$ definido por

$$\langle \text{grad } f_p, v \rangle = df_p \cdot v$$

para todo $p \in M$, $v \in T_pM$.

Seja $c: [a, b] \rightarrow M$ uma curva de classe C^∞ por partes. O *comprimento da curva* c , denotado por $L(c)$, é

$$L(c) = \int_a^b \|c'(t)\| dt \quad (\text{II.3.1})$$

onde $\|c'(t)\| = \left(\left\langle \frac{dc}{dt}, \frac{dc}{dt} \right\rangle \right)^{1/2}$, e o *comprimento de arco* de c é

$$s(t) = \int_a^t \|c'(t)\| dt. \quad (\text{II.3.2})$$

Sejam p e $q \in M$, considere o conjunto $C_{pq} = \{c: [a, b] \rightarrow M/c \text{ é contínua e de classe } C^\infty \text{ por partes; } c(a) = p \text{ e } c(b) = q\}$, isto é, o conjunto de todas as curvas de classe C^∞ por partes ligando p a q .

Definição II.3.2. Seja M uma variedade Riemanniana. Dados p e $q \in M$, a *distância Riemanniana* de p a q , denotada por $d(p, q)$, é dada por

$$d(p, q) := \inf\{L(c) \mid c \in C_{pq}\}. \quad (\text{II.3.3})$$

A *função distância* $d: M \times M \rightarrow \mathbb{R}$ é contínua, o conjunto $B_r(p) = \{q \in M; d(p, q) < r\}$ é chamado *bola métrica* de centro p e raio $r > 0$, e seu fecho é dado por $\overline{B}_r(p) = \{q \in M \mid d(p, q) \leq r\}$.

Definição II.3.3. Sejam M e N variedades Riemannianas. Um difeomorfismo $\mu: M \rightarrow N$ de classe C^∞ é chamado uma *isometria* se

$$\langle u, v \rangle_p = \langle d\mu_p(u), d\mu_p(v) \rangle_{\mu(p)}$$

para todo $p \in M$, $u, v \in T_p M$.

II.4 Conexão Riemanniana

Seja M uma variedade Riemanniana. Iremos denotar por ∇ a *conexão de Levi-Civita* de M , e por $\nabla_Y X$ a *derivada covariante* de X por Y , onde $X \in \mathfrak{X}^1(M)$ e $Y \in \mathfrak{X}^0(M)$. Como $(\nabla_Y X)_p$ depende somente de Y_p e do valor de X ao longo de uma curva em M que é tangente a X_p , denotaremos este vetor por $\nabla_{Y_p} X$. Considere uma curva $c: [a, b] \rightarrow M$ de classe C^∞ e $Y: [a, b] \rightarrow TM$ um campo de classe C^r , $r \geq 1$, ao longo da curva, isto é, $Y(t) := Y(c(t)) \in T_{c(t)}M$, a *derivada covariante* de Y ao longo de c é denotada por $\frac{DY}{dt} := \nabla_{c'} Y$. O *transporte paralelo* ao longo da curva c é denotado por $P(c)_a^b$. As vezes, usaremos também a notação P_{pq} quando $c(a) = p$ e $c(b) = q$ e não for necessário explicitar a curva.

Definição II.4.1. Seja M uma variedade Riemanniana e $X \in \mathfrak{X}^1(M)$. A *diferencial* do campo X é o operador linear $A_X: \mathfrak{X}^0(M) \rightarrow \mathfrak{X}^0(M)$ dado por $A_X(Y) := \nabla_Y X$, e para cada ponto $p \in M$, temos definida uma aplicação linear

$$\begin{aligned} A_X(p): T_p M &\rightarrow T_p M & (II.4.1) \\ v &\mapsto A_X(p) \cdot v = \nabla_v X. \end{aligned}$$

Em particular, se $X = \text{grad } f$, onde $f: M \rightarrow \mathbb{R}$ é uma aplicação de classe C^2 , então $A_X(p) = \text{Hess } f_p$ é a *hessiana* de f .

II.5 Geodésicas e aplicação exponencial

Seja M uma variedade Riemanniana. Uma curva $\gamma: I \rightarrow M$ é chamada uma *geodésica* se

$$\gamma''(t) := \frac{D\gamma'(t)}{dt} = \nabla_{\gamma'(t)} \gamma'(t) = 0 \quad (II.5.1)$$

para todo $t \in I$. Pode-se provar que, se γ é uma geodésica, $\|\gamma'(t)\|^2 = \langle \gamma'(t), \gamma'(t) \rangle \equiv$ constante, isto é, γ tem velocidade constante. Assim de (II.3.2) segue-se que o comprimento do arco s de γ , a partir de $t = t_0$, é $s(t) = \|\gamma'(t)\|(t - t_0)$. Quando $\|\gamma'(t)\| = 1$ dizemos que γ está *parametrizada pelo comprimento de arco ou normalizada*. Dado $p \in M$ e $v \in T_p M$, a equação (II.5.1) tem uma única solução γ , definida num certo intervalo I , tal que $\gamma(0) = p$ e $\gamma'(0) = v$. Quando for conveniente, denotaremos tal geodésica por γ_v . *Iremos sempre considerar variedades Riemannianas completas*, ou seja, variedades Riemannianas cujas geodésicas estão definidas para todo t , isto é, $I = \mathbb{R}$. Assim, não é difícil mostrar que, para todo $a \in \mathbb{R}$, $a > 0$ a igualdade

$$\gamma_{av}(t) = \gamma_v(at) \quad (II.5.2)$$

é satisfeita para todo $t \in \mathbb{R}$.

Definição II.5.1. Seja M uma variedade Riemanniana completa. Para cada $p \in M$, a *aplicação exponencial* em p , denotada por \exp_p , é definida por

$$\begin{aligned} \exp: T_p M &\rightarrow M \\ v &\mapsto \exp_p v = \gamma_v(1) \end{aligned}$$

onde γ é a geodésica de M tal que $\gamma_v(0) = p$. Segue de (II.5.2) que $\gamma_v(t) = \exp_p tv$.

A aplicação exponencial \exp_p é uma função de classe C^∞ , e um difeomorfismo numa vizinhança Ω da origem em $T_p M$. O conjunto $\exp_p \Omega = \tilde{\Omega}$ é chamada uma *vizinhança normal* de p . Se $\tilde{\Omega}$ é uma vizinhança normal de cada um de seus pontos, então dizemos que $\tilde{\Omega}$ é uma *vizinhança totalmente normal*. Se $B_\varepsilon(0) = \{v \in T_p M \mid \|v\| < \varepsilon\}$ é tal que $\overline{B_\varepsilon(0)} \subset \Omega$, chamamos $\exp_p B_\varepsilon(0) = B_\varepsilon(p)$ a *bola normal*, ou *geodésica*, de centro p e raio ε que, neste caso, coincide com a bola métrica (veja Definição II.3.2).

Teorema II.5.2. (*Teorema de Hopf-Rinow*). *Seja M uma variedade Riemanniana. As seguintes afirmações são equivalentes:*

- i) *Para cada ponto $p \in M$, \exp_p está definida em todo $T_p M$, isto é, M é uma variedade Riemanniana completa.*
- ii) *(M, d) é completo como espaço métrico, onde d é distância Riemanniana.*
- iii) *Os subconjuntos limitados e fechados de M são compactos.*

Além disso, cada uma das afirmações acima implica que

- iv) *Para quaisquer dois pontos p e $q \in M$ existe um segmento de geodésica γ ligando p a q com $L(\gamma) = d(p, q)$. A geodésica γ com esta propriedade é chamada *minimizante*.*

Proposição II.5.3. *Sejam M e N variedades Riemannianas completas. Se $\mu: M \rightarrow N$ é uma isometria e γ é uma geodésica de M , então $\mu \circ \gamma$ é uma geodésica de N .*

Definição II.5.4. Seja M uma variedade Riemanniana completa. Um *triângulo geodésico* de M determinado pelos pontos $p_1, p_2, p_3 \in M$, denotado por $\Delta(p_1 p_2 p_3)$, é o conjunto formado pelos três pontos p_1, p_2 e p_3 chamados de *vértices* e três segmentos de geodésicas minimizantes γ_{i+1} ligando p_{i+1} a p_{i+2} , $i = 1, 2, 3 \pmod{3}$, chamados de *lados*.

II.6 Curvatura

O *tensor curvatura* R de uma variedade Riemanniana M é dado por $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z + \nabla_{[Y, X]}Z$, onde X, Y e $Z \in \mathfrak{X}^r(M)$, $r \geq 2$ e o *colchete* $[X, Y] = XY - YX$. Então a *curvatura seccional* $K(X, Y)$ segundo o espaço gerado por X, Y é definida por

$$K(X, Y) = \frac{\langle R(X, Y)Y, X \rangle}{\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2} \quad (\text{II.6.1})$$

onde $\|X\| = \langle X, X \rangle^{1/2}$. Se $K(X, Y) \leq 0$, respectivamente $K(X, Y) \geq 0$, para cada par X, Y , então M é dita uma *variedade Riemanniana de curvatura não-positiva*, respectivamente *não-negativa*; neste caso usaremos a notação $K \leq 0$, respectivamente $K \geq 0$.

II.7 Campos de Jacobi e Fórmulas da Variação

Sejam M variedade Riemanniana e γ uma geodésica de M . Um campo J ao longo de γ é chamado *campo de Jacobi* se ele satisfaz a equação diferencial

$$\nabla_{\gamma'} \nabla_{\gamma'} J + R(J, \gamma')\gamma' = 0 \quad (\text{II.7.1})$$

onde R é o tensor curvatura de M .

Definição II.7.1. Sejam M uma variedade Riemanniana e $\gamma: [a, b] \rightarrow M$ uma geodésica de M . Uma *variação* de γ é uma função $\alpha: [a, b] \times (-\varepsilon, \varepsilon) \rightarrow M$ de classe C^∞ , tal que $\alpha(t, 0) \equiv \gamma(t)$.

O campo de vetores ao longo de γ definido por $V(t) := \frac{\partial \alpha}{\partial s}(t, 0)$ é o *campo variacional* de α . Se a variação α é tal que, para cada s , a curva $\alpha(\cdot, s)$ é uma geodésica, então o campo $J(t) = \frac{\partial \alpha}{\partial s}(t, s)$ é um campo de Jacobi ao longo desta geodésica. A *fórmula da primeira variação de comprimento de arco* sobre a família de geodésicas $c_s: [a, b] \rightarrow M$, dadas por $c_s(t) = \alpha(t, s)$, onde $s \in (-\varepsilon, \varepsilon)$, é dada por

$$L'(\gamma) := \frac{d}{ds}L(c_s)|_{s=0} = \left\langle V, \frac{\gamma'}{\|\gamma'\|} \right\rangle_a^b \quad (\text{II.7.2})$$

e a *fórmula da segunda variação do comprimento de arco* é dada por

$$\begin{aligned} L''(\gamma) &:= \frac{d^2}{ds^2}L(c_s)|_{s=0} & (\text{II.7.3}) \\ &= \frac{1}{\|\gamma'\|} \int_a^b \{ \|\nabla_{\gamma'} V^\perp\|^2 - K(V, \gamma')(\|V\|^2 \|\gamma'\|^2 - \langle V, \gamma' \rangle^2) \} dt + \left\langle \nabla_V V, \frac{\gamma'}{\|\gamma'\|} \right\rangle_a^b \end{aligned}$$

onde $V^\perp = V - \langle V, \frac{\gamma'}{\|\gamma'\|} \rangle \frac{\gamma'}{\|\gamma'\|}$ denota a componente normal de V com relação a γ' .

II.8 Variedades de Hadamard

O primeiro resultado importante em variedades Riemannianas completas de curvatura não-positiva é o *teorema de Cartan-Hadamard* a seguir.

Teorema II.8.1. *Seja M uma variedade Riemanniana completa (e conexa), simplesmente conexa, com curvatura $K \leq 0$. Então M é difeomorfa ao espaço euclidiano \mathbb{R}^n $n = \dim M$; mais precisamente, $\exp_p: T_p M \rightarrow M$ é um difeomorfismo de classe C^∞ para cada $p \in M$.*

Uma variedade Riemanniana que satisfaça as hipóteses do Teorema II.8.1, isto é, completa (e conexa), simplesmente conexa, com curvatura $K \leq 0$ é chamada uma *variedade de Hadamard*. O Teorema II.8.1 assegura que se M é uma variedade de Hadamard, então M tem a mesma topologia e estrutura diferenciável do espaço Euclidiano \mathbb{R}^n . Além destas

propriedades topológicas e diferenciais são conhecidas algumas propriedades geométricas análogas às do espaço Euclidiano, por exemplo o teorema a seguir, o qual usaremos bastante.

Teorema II.8.2. (*Teorema de Topogonov*). *Seja M uma variedade de Hadamard e $\Delta(p_1p_2p_3)$ um triângulo geodésico. Denote por $\gamma_{i+1}: [0, \ell_{i+1}] \rightarrow M$ o segmento de geodésica, ligando p_{i+1} a p_{i+2} , $\ell_{i+1} := L(\gamma_{i+1})$, e seja $\theta_{i+1} = \sphericalangle(\gamma'_{i+1}(0), -\gamma'_i(\ell_i))$, onde $i \equiv 1, 2, 3 \pmod{3}$.*

Então

$$\theta_1 + \theta_2 + \theta_3 \leq \pi. \quad (\text{II.8.1})$$

$$\ell_{i+1}^2 + \ell_{i+2}^2 - 2\ell_{i+1}\ell_{i+2} \cos \theta_{i+2} \leq \ell_i^2. \quad (\text{II.8.2})$$

$$\ell_{i+1} \cos \theta_{i+2} + \ell_i \cos \theta_i \geq \ell_{i+2} \quad (\text{II.8.3})$$

Observamos ainda que se $K < 0$ as desigualdades (II.8.1), (II.8.2) e (II.8.3) são estritas.

II.8.1 A função distância

Seja M uma variedade de Hadamard e seja $p' \in M$. Pelo Teorema II.8.1 podemos definir a inversa da aplicação exponencial $\exp_{p'}^{-1}: M \rightarrow T_{p'}M$ e deste modo temos a seguinte relação entre a distância Riemanniana e a aplicação exponencial

$$d(p, p') = \|\exp_{p'}^{-1} p\|. \quad (\text{II.8.4})$$

Sendo $\exp_{p'}^{-1}$ uma função de classe C^∞ segue-se de (II.8.4) que a função

$$\rho_{p'}: M \rightarrow \mathbb{R} \quad (\text{II.8.5})$$

$$p \mapsto \rho_{p'}(p) := \frac{1}{2}d^2(p, p')$$

é também de classe C^∞ .

Proposição II.8.3. *O gradiente da aplicação $\rho_{p'}$, definida em (II.8.5), no ponto $p \in M$, é $\text{grad } \rho_{p'}(p) = -\exp_p^{-1} p'$.*

Demonstração. São dados $p \in M$ e $u \in T_p M$. Seja $\ell = d(p, p')$ e $\gamma: [0, \ell] \rightarrow M$ a geodésica tal que $\gamma(0) = p'$ e $\gamma(\ell) = p$. Deste modo, temos que $\gamma'(\ell) = -(\exp_p^{-1} p')/d(p, p')$. Considere a variação $\alpha: [0, \ell] \times (-\varepsilon, \varepsilon) \rightarrow M$ da geodésica γ , definida por $\alpha(t, s) = \exp_{p'} tv(s)$, onde $v(s) = v + (s/\ell)w$, $v = \gamma'(0)$ e w é dado pela equação $d(\exp_{p'}^{-1})_{\ell v} \cdot w = u$. Observe que o campo variacional $V(t) = \frac{\partial \alpha}{\partial s}(t, 0)$ de α satisfaz $V(0) = 0$ e $V(\ell) = u$. Assim da definição de $\rho_{p'}$, regra da cadeia e (II.7.2) temos

$$\begin{aligned} d(\rho_{p'})_p \cdot u &= \frac{d}{ds} \rho_{p'}(\alpha(\ell, s))|_{s=0} \\ &= d(p, p') L'(\gamma) \\ &= d(p, p') \langle \gamma', V \rangle|_0^\ell \\ &= d(p, p') \langle \gamma', u \rangle \\ &= \langle -\exp_p^{-1} p', u \rangle. \end{aligned}$$

Portanto, da Definição II.3.1, temos $\text{grad } \rho_{p'}(p) = -\exp_p^{-1} p'$. □

Lema II.8.4. *Seja M uma variedade de Hadamard e sejam γ_1, γ_2 duas geodésicas de M . Então a função $f: \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(s) = d(\gamma_1(s), \gamma_2(s))$ é convexa.*

Demonstração. Se $\gamma_1 \equiv \gamma_2$, então $f \equiv 0$ e portanto convexa. Vamos assumir agora que $\gamma_1 \not\equiv \gamma_2$. Fixe $s_0 \in \mathbb{R}$ tal que $\gamma_1(s_0) \neq \gamma_2(s_0)$ e seja $\gamma: [0, 1] \rightarrow M$ a geodésica que liga $\gamma_1(s_0)$ a $\gamma_2(s_0)$, isto é, $\gamma(0) = \gamma_1(s_0)$ e $\gamma(1) = \gamma_2(s_0)$. Considere a variação $\alpha: [0, 1] \times (-\varepsilon, \varepsilon) \rightarrow M$ de γ dada por $\alpha(t, s) = \exp_{\gamma_1(s_0+s)} t(\exp_{\gamma_1(s_0+s)}^{-1} \gamma_2(s_0+s))$, onde $\varepsilon > 0$ é tal que $\gamma_1(s_0+s) \neq \gamma_2(s_0+s)$ para todo $s \in (-\varepsilon, \varepsilon)$. Observamos que $\alpha(0, s) = \gamma_1(s_0+s)$, $\alpha(1, s) = \gamma_2(s_0+s)$ e que para cada s a curva $\alpha_s := \alpha(\cdot, s): [0, 1] \rightarrow M$ é uma geodésica que liga $\gamma_1(s_0+s)$ a $\gamma_2(s_0+s)$. Desde modo, $L(\alpha_s) = d(\gamma_1(s_0+s), \gamma_2(s_0+s)) := f(s_0+s)$. Do fato que

$\gamma_1(s_0 + s) \neq \gamma_2(s_0 + s)$ para todo $s \in (-\varepsilon, \varepsilon)$, segue-se que a função $g: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ dada por $g(s) = f(s_0 + s) := L(\alpha_s)$ é C^∞ ; então pela fórmula da segunda variação do comprimento de arco (II.7.3) temos

$$\begin{aligned} f''(s_0) &= g''(0) \\ &= \frac{d^2}{ds^2} L(\alpha_s)|_{s=0} \\ &= \frac{1}{\|\gamma'\|} \int_0^1 \{ \|\nabla_{\gamma'} V^\perp\|^2 - K(V, \gamma') (\|V\|^2 \|\gamma'\|^2 - \langle V, \gamma' \rangle^2) \} dt + \langle \nabla_V V, \frac{\gamma'}{\|\gamma'\|} \rangle|_0^1 \end{aligned} \quad (\text{II.8.6})$$

onde $V(t) = \frac{\partial \alpha}{\partial s}(t, 0)$, isto é, o campo de Jacobi ao longo de γ . Como M é de Hadamard segue-se que $K(V, \gamma') \leq 0$. Assim o integrando de (II.8.6) é não-negativo, e como γ_1, γ_2 são geodésicas temos $\langle \nabla_V V, \frac{\gamma'}{\|\gamma'\|} \rangle|_0^1 = 0$, e isto implica que $f''(s_0) \geq 0$. Então $f''(s_0) \geq 0$ para todo $s_0 \in \mathbb{R}$ tal que $\gamma_1(s_0) \neq \gamma_2(s_0)$, e se $\gamma_1(s_0) = \gamma_2(s_0)$, e esta igualdade só pode ocorrer no máximo em um ponto pelo Teorema II.8.1. Então f assume o valor mínimo 0 em s_0 . Portanto f é convexa. \square

II.9 O Teorema de Toponogov

Os resultados desta seção são usados apenas na prova do Lema V.2.3.

Teorema II.9.1. (*Teorema de Toponogov*). *Seja M uma variedade Riemanniana completa com curvatura seccional $K \geq H$. Seja γ_1 e γ_2 segmentos de geodésicas normalizadas em M com $\gamma_1(0) = \gamma_2(0)$. Indiquemos por $M^2(H)$ uma variedade de dimensão dois com curvatura constante H . Admitamos que a geodésica γ_1 é minimizante e que, se $H > 0$, $L(\gamma_2) \leq \frac{\pi}{\sqrt{H}}$. Considere em $M^2(H)$ duas geodésicas normalizadas $\bar{\gamma}_1, \bar{\gamma}_2$, tais que $\bar{\gamma}_1(0) = \bar{\gamma}_2(0)$, $L(\bar{\gamma}_1) = L(\bar{\gamma}_2) = \ell_i$, $i = 1, 2$ e $\sphericalangle(\bar{\gamma}'_1(0), \bar{\gamma}'_2(0)) = \sphericalangle(\gamma'_1(0), \gamma'_2(0))$. Então*

$$d(\gamma_1(\ell_1), \gamma_2(\ell_2)) \leq d(\bar{\gamma}_1(\ell_1), \bar{\gamma}_2(\ell_2)).$$

Corolário II.9.2. *Seja M uma variedade Riemanniana completa com curvatura seccional $K \geq 0$. Se γ_{v_1} e γ_{v_2} são geodésicas normalizadas tal que $\gamma_{v_1}(0) = \gamma_{v_2}(0)$ e se γ_{v_1} é minimizante, então*

$$d(\gamma_{v_1}(t_1), \gamma_{v_2}(t_2)) \leq \|t_2 v_2 - t_1 v_1\|$$

Demonstração. Com a notação do Teorema II.9.1, $M^2(H)$ é o subespaço gerado pelos vetores v_1, v_2 , $H = 0$, $\gamma_1 = \gamma_{v_1}$, $\gamma_2 = \gamma_{v_2}$, $\bar{\gamma}_1(t) = tv_1$ e $\bar{\gamma}_2(t) = tv_2$. Observe também que neste caso não temos nenhuma hipótese sobre $\gamma_2 = \gamma_{v_2}$. Feita esta identificação, a prova é uma consequência imediata do Teorema II.9.1. \square

CAPÍTULO III

Análise convexa: Um resumo dos conceitos básicos

III.1 Introdução

Neste capítulo vamos tratar de alguns conceitos básicos a respeito de conjuntos convexos e funções convexas em variedades Riemannianas os quais serão utilizados nos capítulos seguintes. Do mesmo modo que no capítulo anterior, daremos ênfase a estes conceitos em variedades de Hadamard que é o ambiente onde desenvolveremos alguns algoritmos. Apesar dos resultados deste capítulo serem todos conhecidos iremos demonstrar a grande maioria deles. Fizemos esta opção porque ainda são pouco utilizados os resultados da Geometria Riemanniana na Programação Matemática, e também porque desejamos uniformizar as notações e, na medida do possível, simplificar as demonstrações com argumentos comuns à análise convexa em \mathbb{R}^n (as referências de análise convexa em \mathbb{R}^n que utilizamos foram [28], [40] e [46]). Os resultados deste capítulo se encontram distribuídos em [1], [5], [8], [17]-[25], [29], [37], [45] e [48]-[54]. À medida em que eles forem aparecendo citaremos as referências específicas.

III.2 Conjuntos convexos

Em uma variedade Riemanniana M dada pode existir mais de uma geodésica ligando dois pontos; este fato torna o conceito de convexidade bem mais complicado que o da geometria Euclidiana. Uma situação mais simples é restringir o estudo a variedades Riemannianas onde dados dois pontos existe uma única geodésica que os liga, por exemplo, variedades de Hadamard: esta é a situação que mais nos interessa no momento. Os resultados desta seção serão usados apenas para mostrar que o subdiferencial de uma função convexa (estes conceitos serão definidos na próxima seção) definido em uma variedade de Hadamard é não vazio, Teorema III.3.19, da próxima seção). As variedades de Hadamard possuem propriedades geométricas muito parecidas com às Euclidianas; veremos que isto se transfere ao estudo de convexidade. Os resultados desta seção estão em Bishop e O'Neill [5].

Definição III.2.1. Seja M uma variedade de Hadamard. Um subconjunto $C \subset M$ é dito *convexo*, se para quaisquer pontos p e $p' \in C$ a única geodésica ligando p a p' em M está contida em C , isto é, $\gamma: [a, b] \rightarrow M$ tal que $\gamma(a) = p$, $\gamma(b) = p'$ e $\gamma([a, b]) \subset C$.

Observação III.2.2. Seja M uma variedade Riemanniana. Algumas noções particulares de conjuntos convexos em M são:

i) Um subconjunto $C \subset M$ é dito *fortemente convexo*, se para quaisquer pontos p e $p' \in C$ existir uma única geodésica minimal γ ligando p a p' em M , e γ está contida em C , isto é, $\gamma: [a, b] \rightarrow M$ tal que $\gamma(a) = p'$, $\gamma(b) = p$ e $\gamma([a, b]) \subset C$.

ii) Um subconjunto $C \subset M$ é dito *totalmente convexo*, se para quaisquer pontos p e $p' \in C$, toda geodésica γ ligando p e p' em M está contida em C , isto é se $\gamma: [a, b] \rightarrow M$ é tal que $\gamma(a) = p$, $\gamma(b) = p'$ então $\gamma([a, b]) \subset C$.

Em variedades de Hadamard estas duas definições coincidem com a Definição III.2.1. Não é nosso interesse explorar as definições dadas na Observação III.2.2 mas existem trabalhos

nesta direção, por exemplo, relacionado a noção de totalmente convexo com propriedade global da variedade; veja [8].

Sejam M uma variedade de Hadamard e $C \subset M$ um subconjunto fechado e convexo, fixe $p' \in M$, e considere o seguinte problema¹:

$$\inf \left\{ d(p', p) \mid p \in C \right\}. \quad (\text{III.2.1})$$

Dado $p_0 \in C$, tome o conjunto sub-nível $A_{p_0} = \{p \in M \mid d(p', p) \leq d(p', p_0)\}$ que é compacto, pois a aplicação $p \mapsto d(p', p)$ é contínua e M é completa. Assim o problema:

$$\inf \left\{ d(p', p) \mid p \in C \cap A_{p_0} \right\} \quad (\text{III.2.2})$$

tem solução, pois $C \cap A_{p_0}$ é compacto. Sendo (III.2.1) equivalente a (III.2.2) segue-se que (III.2.1) tem solução. Portanto deduzimos a *existência* de um ponto em C que minimiza a distância a p' e desta forma (III.2.1) é de fato um mínimo. Veremos a seguir que existe um único ponto que minimiza a distância de p' ao convexo C . O próximo resultado nos ajudará nesta direção.

Proposição III.2.3. *Seja C um conjunto convexo fechado de M e seja $p' \in M$. Se $q_{p'} \in C$ é tal que $d(p', q_{p'}) \leq d(p', p)$ para todo $p \in C$, então $\langle \exp_{q_{p'}}^{-1} p', \exp_{q_{p'}}^{-1} p \rangle \leq 0$, para todo $p \in C$.*

Demonstração. Se $p' \in C$ o resultado vale. Suponhamos que $p' \notin C$. Sejam $\ell = d(p', q_{p'})$, $\gamma: [0, \ell] \rightarrow M$ a geodésica tal que $\gamma(0) = p'$ e $\gamma(\ell) = q_{p'}$. Observe que $\gamma'(\ell) = -(\exp_{q_{p'}}^{-1} p') / (\|\exp_{q_{p'}}^{-1} p'\|)$. Suponha por absurdo que exista $\bar{p} \in C$ tal que $\langle \exp_{q_{p'}}^{-1} p', \exp_{q_{p'}}^{-1} \bar{p} \rangle > 0$. Considere a variação $\alpha: [0, \ell] \times (-\varepsilon, \varepsilon) \rightarrow M$ da geodésica γ definida por $\alpha(t, s) = \exp_{p'} t(\exp_{p'}^{-1} \beta(s))$, onde $\beta: (-\varepsilon, \ell + \varepsilon) \rightarrow M$ é a geodésica tal que $\beta(0) = q_{p'}$ e $\beta(\ell) = \bar{p}$.

¹Aqui estamos cometendo um abuso na linguagem.

Deste modo o campo variacional $V(t) = \frac{\partial \alpha}{\partial s}(t, 0)$ de α satisfaz $V(0) = 0$ e $V(\ell) = \beta'(0) = \exp_{q_{p'}}^{-1} \bar{p}$. Assim de (III.2.2) segue-se que

$$L'(\gamma) = \frac{d}{ds} L(c_s) \Big|_{s=0} = \langle -\exp_{q_{p'}}^{-1} p' / \|\exp_{q_{p'}}^{-1} p'\|, \exp_{q_{p'}}^{-1} \bar{p} \rangle < 0$$

onde $c_s(\cdot) = \alpha(\cdot, s)$ e $s \in (-\varepsilon, \ell + \varepsilon)$. Portanto pela última desigualdade existe $\delta > 0$ tal que $L(c_s) < L(\gamma)$ para $0 < s < \delta$, isto é, $d(p', \beta(s)) < d(p', q_{p'})$ para $0 < s < \delta$. Sendo C convexo segue-se que $\beta(s) \in C$ para $0 < s < \delta$, e deste modo temos um absurdo. \square

Proposição III.2.4. *Seja C um conjunto convexo fechado de M . Então para cada ponto $p' \in M$ existe um único ponto $q_{p'} \in C$ tal que $d(p', q_{p'}) \leq d(p', p)$ para todo $p \in C$.*

Demonstração. Suponhamos que $p' \notin C$. Suponha por absurdo que existam $q_{p'}, \bar{q}_{p'} \in C$ tais que $q_{p'} \neq \bar{q}_{p'}$ e $d(q_{p'}, p') = d(q_{p'}, \bar{q}_{p'}) = d(p', C)$. Pela Proposição III.2.3 segue-se que

$$\theta := \sphericalangle \left(\exp_{q_{p'}}^{-1} p', \exp_{q_{p'}}^{-1} \bar{q}_{p'} \right) \geq \frac{\pi}{2}, \quad \bar{\theta} = \sphericalangle \left(\exp_{\bar{q}_{p'}}^{-1} p', \exp_{\bar{q}_{p'}}^{-1} q_{p'} \right) \geq \frac{\pi}{2}$$

como $q_{p'} \neq \bar{q}_{p'}$, $d(p', q_{p'}) = d(q_{p'}, \bar{q}_{p'})$ e C é convexo temos que $\theta' := \sphericalangle \left(\exp_{p'}^{-1} q_{p'}, \exp_{p'}^{-1} \bar{q}_{p'} \right) > 0$, deste modo a soma dos ângulos internos do triângulo geodésico $\Delta(p' q_{p'} \bar{q}_{p'})$ definido pelos pontos p' , $q_{p'}$ e $\bar{q}_{p'}$ é maior que π , isto é, $\theta + \bar{\theta} + \theta' > \pi$, e isto contradiz Teorema II.8.2. \square

Denotemos por $p_C(p')$ o único ponto dado pelo Teorema III.2.4, que é chamado de *projeção* de p' sobre o convexo C . Vamos então resumir os dois últimos resultados no seguinte teorema.

Proposição III.2.5. *Seja M uma variedade de Hadamard e seja C um subconjunto convexo fechado de M . Então para cada $p' \in M$ existe uma única projeção $p_C(p') \in C$. Além disso, vale a desigualdade*

$$\langle \exp_{p_C(p')}^{-1} p', \exp_{p_C(p')}^{-1} p \rangle \leq 0,$$

para todo $p \in C$.

Demonstração. Teorema III.2.3 e Teorema III.2.4 □

Observação III.2.6. Veja que na demonstração do Teorema III.2.4 usamos fortemente o fato que M é de Hadamard. Na verdade, este teorema vale em variedades onde temos a unicidade de geodésica ligando dois pontos quaisquer. Em geral este fato não é verdade, por exemplo na esfera euclidiana.

Dados $C \subset M$ um subconjunto convexo com fronteira $\partial C \neq \emptyset$, $p' \in \partial C$, $s \in T_{p'}M$ e $s \neq 0$. Considere o subespaço definido por

$$S_{s,p'} = \{v \in T_{p'}M \mid \langle s, v \rangle = 0\}. \quad (\text{III.2.3})$$

O subespaço $S_{s,p'}$ é *suporte* a C em p' se vale a desigualdade

$$\langle s, \exp_{p'}^{-1} p \rangle \leq 0 \quad (\text{III.2.4})$$

para cada $p \in C$.

Proposição III.2.7. *Seja $\bar{p} \notin C$, onde $C \neq M$ é um subconjunto não vazio, convexo e fechado de M . Então existe $s \in T_{\bar{p}}M$ tal que*

$$\sup\{\langle -s, \exp_{\bar{p}}^{-1} p \rangle \mid p \in C\} < 0. \quad (\text{III.2.5})$$

Demonstração. Considere o triângulo geodésico $\Delta(\bar{p}, p, p_C(\bar{p}))$ e sejam $\beta = \sphericalangle(\exp_{\bar{p}}^{-1} p_C(\bar{p}), \exp_{\bar{p}}^{-1} p)$, $\theta = \sphericalangle(\exp_{p_C(\bar{p})}^{-1} \bar{p}, \exp_{p_C(\bar{p})}^{-1} p)$. Segue-se do Teorema II.8.2 que

$$d(\bar{p}, p) \cos \beta + d(p, p_C(\bar{p})) \cos \theta \geq d(\bar{p}, p_C(\bar{p})) \quad (\text{III.2.6})$$

Agora no plano tangente $T_{\bar{p}}M$ considere o triângulo $\Delta(0, \exp_{\bar{p}}^{-1} p, \exp_{\bar{p}}^{-1} p_C(\bar{p}))$ e sejam $s := \exp_{\bar{p}}^{-1} p_C(\bar{p})$ e $\alpha = \sphericalangle(-s, \exp_{\bar{p}}^{-1} p - s)$ também do Teorema II.8.2, neste caso com igualdade, temos

$$\|\exp_{\bar{p}}^{-1} p\| \cos \beta + \|\exp_{\bar{p}}^{-1} p - s\| \cos \alpha = \|s\|. \quad (\text{III.2.7})$$

Como $d(\bar{p}, p) = \|\exp_{\bar{p}}^{-1} p\|$ e $d(\bar{p}, p_C(\bar{p})) = \|s\|$ segue-se de (III.2.6) e (III.2.7) que

$$d(p, p_C(\bar{p})) \cos \theta \geq \|\exp_{\bar{p}}^{-1} p - s\| \cos \alpha$$

mas $d(p, p_C(\bar{p})) = \|\exp_{p_C(\bar{p})}^{-1} p\|$ e $\|s\| = \|\exp_{\bar{p}}^{-1} p_C(\bar{p})\|$, assim

$$\langle -s, \exp_{\bar{p}}^{-1} p - s \rangle \leq \langle \exp_{p_C(\bar{p})}^{-1} \bar{p}, \exp_{p_C(\bar{p})}^{-1} p \rangle. \quad (\text{III.2.8})$$

Da Proposição III.2.5 $\langle \exp_{p_C(p')}^{-1} p', \exp_{p_C(p')}^{-1} p \rangle \leq 0$ então da equação (III.2.8) temos $\langle -s, \exp_{\bar{p}}^{-1} p - s \rangle \leq 0$ que é equivalente a

$$\langle -s, \exp_{\bar{p}}^{-1} p \rangle \leq -\|s\|^2$$

e isto implica (III.2.5). □

Teorema III.2.8. *Seja $C \subset M$ um subconjunto convexo fechado ($C \neq M$) e seja $p' \in \partial C$. Então existe um subespaço suporte a C em p' .*

Demonstração. Seja $n = \dim M$. Tome $\delta > 0$ tal que $\bigcup_{p \in B} T_p M = B \times \mathbb{R}^n$, onde $B := B_\delta(p')$ é a bola métrica. Isto é sempre possível pois o fibrado tangente TM é localmente um produto. Tome também uma seqüência $\{p'_k\} \subset B$ tal que $p'_k \notin C$ para $k = 1, 2, \dots$ e $\lim_{k \rightarrow +\infty} p'_k = p'$. Pelo Teorema III.2.7 existe $s_k \in T_{p'_k} M$, tal que

$$\langle -s_k, \exp_{p'_k}^{-1} p \rangle < 0 \quad (\text{III.2.9})$$

para todo $p \in C$. Sem perda de generalidade podemos tomar $\|s_k\| = 1$. Seja \bar{B} o fecho de B . Como a seqüência $\{(p'_k, s_k)\}$ está contida no compacto $\{(p, s) \mid p \in \bar{B}, s \in T_p M \text{ e } \|s\| = 1\}$ de TM podemos extrair uma subsequência convergente $\{(p'_{k_j}, s_{k_j})\}$. Mas como já temos que $\lim_{j \rightarrow +\infty} p'_{k_j} = p'$, então existe $\bar{s} \in T_{p'}$ tal que $\lim_{j \rightarrow +\infty} s'_{k_j} = \bar{s}$. Assim de (III.2.9) e do fato que a métrica varia continuamente com o ponto, obtemos

$$\langle -\bar{s}, \exp_{p'}^{-1} p \rangle \leq 0$$

para todo $p \in C$. Portanto $S_{s,p'}$, definido em (III.2.3), é suporte a C em p' , onde $s = -\bar{s}$.

□

Observação III.2.9. Podemos observar que os resultados são análogos aos do espaço euclidiano. Isto se deve ao fato, como já observamos, de que as variedades de Hadamard possuem propriedades geométricas muito semelhantes às euclidianas. É possível obter os mesmos resultados em variedades onde temos a unicidade de geodésicas ligando dois pontos.

III.3 Funções convexas

O conceito de funções convexas em variedades Riemannianas desempenha um papel de destaque no estudo de propriedades topológicas das variedades Riemannianas não compactas (veja as referências [5], [8], [14] e [45]). Nesta seção estudaremos algumas propriedades das funções convexas e daremos alguns exemplos de tais funções. Em particular daremos uma prova alternativa, usando a noção de subespaço suporte, de que o subdiferencial de uma função convexa definida em uma variedade de Hadamard é não vazio. Este fato foi demonstrado por Udriste [50], para uma variedade Riemanniana completa, usando a noção de derivada direcional (definida na próxima seção).

Definição III.3.1. Seja M uma variedade Riemanniana completa. Uma função $f: M \rightarrow \mathbb{R}$ é dita *convexa* se para toda geodésica $\gamma: \mathbb{R} \rightarrow M$ a composição $f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$ é uma função convexa, ou seja que $f \circ \gamma(ta + (1-t)b) \leq tf(\gamma(a)) + (1-t)f(\gamma(b))$ para qualquer $a, b \in \mathbb{R}$ e $0 \leq t \leq 1$.

Observação III.3.2. Seja $f: M \rightarrow \mathbb{R}$ uma função convexa não constante. A convexidade da função f impõe certas restrições topológicas a M . Intuitivamente isto pode ser visto observando que $S_r(f) := \{p \in M \mid f(p) \leq r\}$ é um subconjunto totalmente convexo. Outra

consequência topológica é que M é não compacta; veja [5]. Pelo menos uma consequência métrica sobre M é conhecida: M tem volume infinito; veja Yau [54]. O estudo de convexidade com objetivos de compreender a estrutura topológica e métrica das variedades Riemannianas tem sido bastante explorado; veja [6], [8], [17], [18], [19], [29], [45], [53], e suas referências.

Proposição III.3.3. *Seja M uma variedade Riemanniana completa. Se $f_i: M \rightarrow \mathbb{R}$, $i = 1, \dots, n$ são funções convexas e $\alpha_i \geq 0$, $i = 1, \dots, n$, então $f := \sum_{i=1}^n \alpha_i f_i$ é uma função convexa*

Demonstração. Segue imediatamente da Definição III.3.1. □

Proposição III.3.4. *Sejam M uma variedade Riemanniana completa e $\{f_k\}_{k \in \mathbb{N}}$ uma sequência de funções convexas em M . Se $\{f_k\}$ converge ponto a ponto para $f: M \rightarrow \mathbb{R}$, isto é, $\lim_{k \rightarrow \infty} f_k(p) = f(p)$ para todo ponto $p \in M$, então f é uma função convexa.*

Demonstração. Segue imediatamente da desigualdade na Definição III.3.1. □

Proposição III.3.5. *Sejam M e N variedades Riemannianas completas e $\mu: M \rightarrow N$ uma isometria. Se $f: M \rightarrow \mathbb{R}$ é uma função convexa, então $f \circ \mu: M \rightarrow \mathbb{R}$ é uma função convexa.*

Demonstração. Segue imediatamente da definição e do fato que isometria leva geodésica em geodésica, isto é, se γ é uma geodésica de M então $\mu \circ \gamma$ é uma geodésica de N . □

Exemplo III.3.6. Seja M uma variedade de Hadamard. Fixe $p' \in M$ e considere a aplicação $\rho_{p'}$ definida em (II.8.5) Seja γ uma geodésica de M , então no Lema II.8.4. Tomando $\gamma_1 \equiv \gamma$ e $\gamma_2 \equiv p'$ temos que $f(s) = d(\gamma(s), p')$ é convexa. Deste modo $\rho_{p'}(\gamma(s)) = \frac{1}{2}f^2(s)$ é convexa. Portanto, pela Definição III.3.1, $\rho_{p'}$ é convexa.

Exemplo III.3.7. Sejam M uma variedade de Hadamard e $d: M \times M \rightarrow \mathbb{R}$ a função distância. Então d é uma função convexa com respeito a métrica produto. Isto segue-se

imediatamente do Lema II.8.4, observando que toda geodésica γ de $M \times M$ pode ser escrita como $\gamma = (\gamma_1, \gamma_2)$, onde γ_1 e γ_2 são geodésicas de M .

Exemplo III.3.8. Sejam M uma variedade de Hadamard e $\mu: M \rightarrow M$ uma isometria. Então a função $f: M \rightarrow \mathbb{R}$ dada por $f(p) = d(p, \mu(p))$ é uma função convexa. Isto segue-se do Lema II.8.4 e da Proposição II.5.3.

Observação III.3.9. A função do exemplo anterior desempenha importante papel no estudo das isometrias de uma variedade de Hadamard, veja [9] e [44].

Exemplo III.3.10. Seja M uma variedade Riemanniana completa e não-compacta. Uma geodésica $\gamma: [0, +\infty] \rightarrow M$ partindo de p' parametrizada pelo comprimento de arco é chamada um *raio* se $d(\gamma(t), \gamma(s)) = |t - s|$, para todo $t, s \geq 0$. Seja γ um raio partindo de algum ponto de M (sempre existe um raio para qualquer ponto de M , pois M é não-compacta). A função de Busemann b_γ é definida por

$$b_\gamma: M \rightarrow \mathbb{R} \tag{III.3.1}$$

$$p \mapsto b_\gamma(p) = \lim_{t \rightarrow \infty} (d(p, \gamma(t)) - t)$$

onde d é a distância Riemanniana. Se M é uma variedade de Hadamard então b_γ é uma função convexa, pois $d(\cdot, \gamma(t)) - t$ é convexa para cada t , isto segue do Exemplo III.3.7 e Proposição III.3.4. Também é verdade que $-b_\gamma$ é convexa se M tem curvatura não-negativa. A prova, neste caso, é um pouco mais técnica, e pode ser encontrada em Cheeger-Gromoll [8].

Observação III.3.11. A função de Busemann desempenha papel importante no estudo da estrutura das variedades Riemannianas completas e não compactas cuja curvatura tenha um sinal fixado, veja [8], [14], [44], [45] e [53].

Definição III.3.12. Seja M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. O subconjunto $\text{epi}(f) := \{(p, r) \in M \times \mathbb{R} / f(p) \leq r\}$ da variedade produto $M \times \mathbb{R}$ é chamado de *epigrafo* de f .

Proposição III.3.13. *Sejam M variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$. Então f é convexa se, e somente se $\text{epi}(f)$ é um subconjunto totalmente convexo da variedade produto $M \times \mathbb{R}$.*

Demonstração. Análoga ao caso $M = \mathbb{R}^n$; veja [28]. Basta observar que $\alpha = (\gamma, \beta)$ é uma geodésica de $M \times \mathbb{R}$ se, e somente se γ e β são geodésicas de M e \mathbb{R} respectivamente. \square

Definição III.3.14. Seja M uma variedade Riemanniana completa. Uma função $f: M \rightarrow \mathbb{R}$ é chamada *Lipschitziana* se existe $L \geq 0$ tal que

$$|f(p) - f(p')| \leq Ld(p, p') \quad (\text{III.3.2})$$

para todo p e $p' \in M$. Dizemos ainda que f é *localmente lipschitziana* se para cada $q \in M$ existe $L(q) \geq 0$ e $\delta = \delta(q) \geq 0$ tal que a desigualdade (III.3.2) ocorre, com $L = L(q)$, para todo $p, p' \in B_\delta(q) = \{p \in M / d(p, q) \leq \delta\}$.

Observação III.3.15. Segue imediatamente da desigualdade triangular que $|d(p, q) - d(p', q)| \leq d(p, p')$ para todo p, p' e $q \in M$, e de (III.3.1) temos

$$|b_\gamma(p) - b_\gamma(p')| \leq \lim_{t \rightarrow \infty} |d(p, \gamma(t)) - d(p', \gamma(t))| \leq d(p, p').$$

Então da Definição III.3.14 segue-se que tanto a função distância Riemanniana a um ponto fixo $d(\cdot, q)$, quanto, a função de Busemann b_γ , são Lipschitzianas e portanto localmente Lipschitzianas. De fato, é sabido que toda função convexa é localmente Lipschitziana e conseqüentemente contínua, veja [21]. Na verdade f é diferenciável. Em “quase todos os pontos” de M .

Definição III.3.16. Seja M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Dado $p' \in M$ um vetor $s \in T_{p'}M$ é um *subgradiente* de f em p' se para toda geodésica $\gamma: \mathbb{R} \rightarrow M$ com $\gamma(0) = p'$

$$f(\gamma(t)) \geq f(p') + t\langle s, \gamma'(0) \rangle \quad (\text{III.3.3})$$

para todo $t \geq 0$. O conjunto de todos os subgradientes de f em p' denotado por $\partial f(p')$, é chamado de *subdiferencial* de f em p' .

No caso particular em que M é uma variedade de Hadamard a Definição III.3.16 é equivalente a

Definição III.3.17. Sejam M uma variedade de Hadamard e $f: M \rightarrow \mathbb{R}$ uma função convexa. Dado $p' \in M$, um vetor $s \in T_{p'}M$ é um *subgradiente* de f em p' se

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

para todo $p \in M$.

Observação III.3.18. Se na Definição III.3.16 $f: M \rightarrow \mathbb{R}$ é diferenciável, em $p' \in M$, então $\partial f(p') = \{\text{grad } f(p')\}$. Neste caso, uma função $f: M \rightarrow \mathbb{R}$ de classe C^1 é convexa se, e somente se, para todo ponto $p' \in M$ e toda geodésica γ de M com $\gamma(0) = p'$, $f(\gamma(t)) \geq f(p') + t\langle \text{grad } f(p'), \gamma'(0) \rangle$, para todo $t \geq 0$. Veja [52] e [13].

Teorema III.3.19. Sejam M uma variedade de Hadamard e $f: M \rightarrow \mathbb{R}$ uma função convexa. Então para todo $p' \in M$ existe $s \in T_{p'}M$ tal que

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

para todo $p \in M$. Isto é, $\partial f(p') \neq \emptyset$ para todo $p' \in M$.

Demonstração. Sendo f contínua, segue-se da Proposição III.3.13 que $\text{epi}(f)$ é um subconjunto convexo e fechado da variedade produto $M \times \mathbb{R}$. Observe que a fronteira de $\text{epi}(f)$ é $\partial(\text{epi}(f)) = \{(p, f(p))/p \in M\}$ e que a inversa da aplicação exponencial de $M \times \mathbb{R}$ é igual a $\exp_{(p', f(p'))}^{-1}(p, r) = (\exp_{p'}^{-1} p, r - f(p'))$. Então seja $S_{((s, \alpha), (p', f(p')))$ o subespaço suporte, dado pelo Teorema III.2.8, a $\text{epi}(f)$ em $(p', f(p'))$, assim

$$\langle s, \exp_{p'}^{-1} p \rangle + \alpha(r - f(p')) := \langle (s, \alpha), (\exp_{p'}^{-1} p, r - f(p')) \rangle \leq 0 \quad (\text{III.3.4})$$

para todo $(p, r) \in \text{epi}(f)$. Agora sejam $\tilde{p} = \exp_{p'} s$ e $\tilde{r} > f(\tilde{p})$, que substituídos em (III.3.4) implicam $\|s\|^2 + \alpha(\tilde{r} - f(\tilde{p})) \leq 0$. Deste modo $0 \leq \|s\|^2 \leq \alpha(f(\tilde{p}) - \tilde{r})$. Assim $\alpha \neq 0$, pois caso contrário $s = 0$. Sem perda de generalidade tomamos $\alpha = -1$ em (III.3.4), que fazendo $r = f(p)$ nos dá

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

para todo $p \in M$, que é o desejado. \square

Observação III.3.20. É ainda verdade que $\partial f(p) \neq \emptyset$ para todo $p \in M$, onde M é uma variedade Riemanniana completa. Este resultado foi obtido por Udriste [50], cuja prova é baseada no conceito de derivada direcional (a qual iremos repeti-la na próxima seção). Aqui na prova do Teorema III.3.19 substituímos este conceito pelo conceito de subespaço suporte, obtendo assim uma versão geométrica para o caso específico das variedades de Hadamard. A prova como está aqui pode ser estendida, com um pouco mais de cuidado, a qualquer variedade Riemanniana completa, pois sendo o conceito de subgradiente local basta apenas mudar a definição de subespaço suporte por um conceito local. Isto se faz necessário pois devemos levar em conta a presença do “cut locus”. Veja a definição e alguns resultados sobre subespaço suporte em Cheeger-Gromoll [8].

Proposição III.3.21. *Sejam M uma variedade Riemanniana e $f: M \rightarrow \mathbb{R}$ uma função convexa. Um ponto $p_* \in M$ é um minimizador da função f se, e somente se, $0 \in \partial f(p_*)$.*

Demonstração. Segue imediatamente da Definição III.3.16. \square

III.4 Derivada direcional de funções convexas

Nesta seção estudaremos algumas propriedades da derivada direcional de uma função convexa definida em uma variedade Riemanniana completa. Alguns dos resultados aqui estão enunciados, mas sem demonstração, em [1], e outros, com demonstração, em [49], [52] e alguns ainda desconhecidos no contexto de variedades Riemannianas, mas que são resultados clássicos da análise convexa do \mathbb{R}^n (veja [26], [45] e [46]), cujas provas em alguns casos, são análogas às do \mathbb{R}^n . Omitiremos a prova daqueles que não iremos utilizar.

Sejam M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Dados $p \in M$ e $v \in T_pM$, seja $c: (-\varepsilon, \varepsilon) \rightarrow M$ uma curva tal que $c(0) = p$ e $c'(0) = v$, e considere o quociente

$$q_c(t) := \frac{f(c(t)) - f(p)}{t}. \quad (\text{III.4.1})$$

Se $\gamma_v: \mathbb{R} \rightarrow M$ é uma geodésica tal que $\gamma_v(0) = p$, então $f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$ é uma função convexa. Deste modo $q_{\gamma_v}: \mathbb{R} \rightarrow \mathbb{R}$ é não-decrescente, e do fato que f é localmente Lipschitziana, segue-se também que q_{γ_v} é limitada próxima de zero. Assim a definição a seguir faz sentido.

Definição III.4.1. Sejam M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. A *derivada direcional* de f em p na direção $v \in T_pM$ é

$$f'(p, v) := \lim_{t \rightarrow 0^+} q_{\gamma_v}(t) = \inf_{t > 0} q_{\gamma_v}(t) \quad (\text{III.4.2})$$

onde $\gamma_v: \mathbb{R} \rightarrow M$ é a geodésica tal que $\gamma_v(0) = p$ e $\gamma_v'(0) = v$.

A seguir vamos demonstrar que a derivada direcional de f em p na direção $v \in T_pM$, dada pela Definição III.4.2, depende apenas da direção e não da curva; isto é, no limite

(III.4.2) podemos tomar qualquer curva c tal que $c(0) = p$ e $c'(0) = v$, e ainda teremos $\lim_{t \rightarrow 0^+} q_c(t) = f'(p, v)$. Antes de demonstrar isto, necessitamos de alguns resultados.

Seja M uma variedade Riemanniana completa. Sejam c_1 e c_2 curvas diferenciáveis em M , tais que $c_1(0) = c_2(0) = p$ e $c_1'(0) = v$ e $c_2'(0) = w$. Considere a variação por geodésicas dada por

$$\begin{aligned} \alpha: [0, 1] \times (-\varepsilon, \varepsilon) &\rightarrow M & (III.4.3) \\ (t, s) &\mapsto V(t, s) = \exp_{c_1(s)}(t \exp_{c_1(s)}^{-1} c_2(s)) \end{aligned}$$

onde $\varepsilon > 0$ é tal que $B_\varepsilon(p)$ seja uma vizinhança totalmente normal. Observe que $\alpha(0, s) = c_1(s)$, $\alpha(1, s) = c_2(s)$, e que para cada s , a curva $\alpha_s: [0, 1] \rightarrow M$ dada por $\alpha_s(t) = \alpha(t, s)$ é uma geodésica. Em particular, para $s = 0$ temos a geodésica constante $\alpha_0(t) = \alpha(t, 0) = p$. Considere ainda, para cada s , os campos $T(\cdot, s) := \frac{\partial \alpha}{\partial t}(\cdot, s)$ tangente à geodésica α_s e $J(\cdot, s) = \frac{\partial \alpha}{\partial s}(\cdot, s)$ o campo de Jacobi ao longo de α_s . Deste modo $J(\cdot, s)$ satisfaz a equação

$$\frac{D^2 J}{\partial t^2}(t, s) + R(J(t, s), T(t, s))T(t, s) = 0 \quad (III.4.4)$$

Lema III.4.2. *Sejam c_1 e c_2 curvas diferenciáveis em M , tais que $c_1(0) = c_2(0) = p$, $c_1'(0) = v$ e $c_2'(0) = w$. Se $T(\cdot, s)$ e $J(\cdot, s)$ são os campos definidos acima, então*

i) $J(t, 0) = v + t(w - v)$ é o campo de Jacobi ao longo da geodésica constante $\alpha_0(t) = \alpha(t, 0) = p$.

Além disso, por simetria, temos,

$$ii) \frac{DT}{\partial s}(t, 0) = \frac{DJ}{\partial t}(t, 0) = w - v.$$

Demonstração. Substituindo $s = 0$ em (III.4.4) temos

$$\frac{D^2 J}{\partial t^2}(t, 0) = 0$$

pois $\alpha_0(s) = p'$ e $T(t, 0) = 0$. Como $J(0, 0) = v$ e $J(1, 0) = w$ resolvendo a equação

$$\begin{cases} \frac{D^2}{dt^2} J(t, 0) & = 0 \\ J(0, 0) & = v \\ J(1, 0) & = w \end{cases}$$

temos $J(t, 0) = v + t(w - v)$ que é o desejado em i). Imediatamente do Lema de simetria (Lema 3.4 pág. 68 de [11]) temos

$$\frac{DT}{\partial s}(t, 0) = \frac{D}{\partial s} \frac{\partial}{\partial t} V(t, 0) = \frac{D}{\partial t} \frac{\partial}{\partial s} V(t, 0) = \frac{D}{\partial t} J(t, 0) = w - v.$$

□

Lema III.4.3. *Sejam c_1 e c_2 curvas diferenciáveis em M , tais que $c_1(0) = c_2(0) = p$, $c_1'(0) = v$ e $c_2'(0) = w$. Se $\psi(s) = d(c_1(s), c_2(s))$, então*

$$i) \frac{d}{ds}(\psi^2(s))|_{s=0} = 0$$

$$ii) \frac{d^2}{ds^2}(\psi^2(s))|_{s=0} = 2\|w - v\|^2.$$

Assim a fórmula de Taylor para ψ^2 numa vizinhança de $s = 0$ é dada por

$$\psi^2(s) = \|w - v\|^2 s^2 + \mathcal{O}(s^2) \quad (\text{III.4.5})$$

onde $\lim_{s \rightarrow 0^+} \frac{\mathcal{O}(s^2)}{s^2} = 0$.

Demonstração. Como $\alpha_s(t) = \alpha(t, s)$, onde α é definida em (III.4.3), temos $\psi(s) = \|\alpha'_s(t)\|^2 = \|T(t, s)\|^2$, assim

$$\frac{d}{ds}(\psi^2(s))|_{s=0} = 2 \left\langle \frac{DT}{\partial s}(t, 0), T(t, 0) \right\rangle = 0$$

pois $T(t, 0) = 0$. Agora

$$\frac{d^2}{ds^2}(\psi^2(s))|_{s=0} = 2 \left[\left\langle \frac{D^2 J}{\partial s^2}(t, 0), T(t, 0) \right\rangle + \left\langle \frac{DT}{\partial s}(t, 0), \frac{DT}{\partial s}(t, 0) \right\rangle \right]$$

e o resultado segue-se do fato que $T(t, 0) = 0$ e do Lema III.4.2, item ii). □

Corolário III.4.4. *Sejam c_1 e c_2 curvas diferenciáveis em M , tais que $c_1(0) = c_2(0) = p$, $c_1'(0) = v$ e $c_2'(0) = w$. Então*

$$\lim_{s \rightarrow 0^+} \frac{d(c_1(s), c_2(s))}{s} = \|w - v\|$$

onde d é a distância Riemanniana.

Demonstração. Segue imediatamente de (III.4.5), onde $\psi(s) = d(c_1(s), c_2(s))$. \square

Teorema III.4.5. *Sejam M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Se $c: (-\varepsilon, \varepsilon) \rightarrow M$ é uma curva diferenciável tal que $c(0) = p$ e $c'(0) = v$, então*

$$f'(p, v) = \lim_{s \rightarrow 0^+} q_c(s)$$

onde q_c é definida em (III.4.1).

Demonstração. Seja γ_v a geodésica tal que $\gamma_v(0) = p$ e $\gamma_v'(0) = v$, então por definição temos $f'(p, v) = \lim_{s \rightarrow 0^+} q_{\gamma_v}(s)$. Como f é localmente Lipschitziana existe $L(p) \geq 0$ tal que

$$\begin{aligned} |f'(p, v) - \lim_{s \rightarrow 0^+} q_c(s)| &= \lim_{s \rightarrow 0^+} |q_{\gamma_v}(s) - q_c(s)| \\ &= \lim_{s \rightarrow 0^+} \frac{f(\gamma_v(s)) - f(c(s))}{s} \\ &\leq L(p) \lim_{s \rightarrow 0^+} \frac{d(\gamma_v(s), c(s))}{s}. \end{aligned} \quad (\text{III.4.6})$$

Como $\gamma_v'(0) = c'(0) = v$, segue do Corolário III.4.4, com $c_1 = \gamma_v$ e $c_2 = c$, que $\lim_{s \rightarrow 0^+} \frac{d(\gamma_v(s), c(s))}{s} = 0$. Portanto de (III.4.6) temos que $f'(p, v) = \lim_{s \rightarrow 0^+} q_c(s)$. \square

Teorema III.4.6. *Seja M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Então a aplicação derivada direcional de f no ponto p definida por*

$$\begin{aligned} f'(p, \cdot): T_p M &\rightarrow \mathbb{R} \\ v &\mapsto f'(p, v) \end{aligned} \quad (\text{III.4.7})$$

é uma função convexa.

Demonstração. Sejam $v, v' \in T_p M$ e $t \in [0, 1]$. Considere a variação por geodésicas definida em (III.4.3), onde $c_1 = \gamma_v$ e $c_2 = \gamma_{v'}$. Pelo Lema III.4.2 temos $\frac{\partial \alpha}{\partial s}(t, 0) = J(t, 0) = v + t(v' - v)$, que é o vetor tangente a curva $s \mapsto \alpha_t(s) := \alpha(t, s)$ (que em geral não é uma geodésica) em $s = 0$, isto é, $\alpha'_t(0) = v + t(v' - v)$. Pelo Teorema III.4.5, segue-se que

$$f'(p, (1-t)v + tv') = \lim_{s \rightarrow 0^+} \frac{f(\alpha_t(s)) - f(p)}{s}. \quad (\text{III.4.8})$$

Para cada s , a curva $t \mapsto \alpha_s(t) = \alpha(t, s)$ é uma geodésica com $\alpha_s(0) = \gamma_v(s)$ e $\alpha_s(1) = \gamma_{v'}(s)$, então da definição de α_t e da convexidade de f temos

$$\begin{aligned} f(\alpha_t(s)) &= f(\alpha_s(t)) \leq (1-t)f(\alpha_s(0)) + tf(\alpha_s(1)) \\ &= (1-t)f(\gamma_v(s)) + tf(\gamma_{v'}(s)) \end{aligned}$$

substituindo a última desigualdade em (III.4.8), segue da Definição III.4.1 que

$$\begin{aligned} f'(p, (1-t)v + tv') &\leq (1-t) \lim_{s \rightarrow 0^+} \frac{f(\gamma_v(s)) - f(p)}{s} + t \lim_{s \rightarrow 0^+} \frac{f(\gamma_{v'}(s)) - f(p)}{s} \\ &= (1-t)f'(p, v) + tf'(p, v') \end{aligned}$$

e isto implica que $f'(p, \cdot)$ é convexa. □

Observação III.4.7. Este resultado foi obtido por Udriste [49], mas em sua demonstração, que é idêntica a que está aqui, a igualdade (III.4.8) a nosso ver não é clara. Aqui ela é justificada pelo Teorema III.4.5.

Proposição III.4.8. *Sejam M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Então para cada $p \in M$, temos*

i) $f'(p, \lambda d) = \lambda f'(p, v)$ para todo $\lambda > 0$ e $v \in T_p M$, isto é, $f'(p, \cdot)$ é positivamente homogênea.

ii) $-f'(p, -v) \leq f'(p, v)$ para todo $v \in T_p M$.

iii) $|f'(p, v)| \leq L(p)\|v\|$ para todo $v \in T_pM$, onde $L(p) \geq 0$ é a constante de Lipschitz de f em p .

Demonstração. Mesma demonstração de \mathbb{R}^n ; veja [28]. □

Teorema III.4.9. *Sejam M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Então para todo $p \in M$ o subdiferencial $\partial f(p)$ é não-vazio.*

Demonstração. Dado $p \in M$. Seja γ uma geodésica de M com $\gamma(0) = p$. De (III.4.2) segue-se que

$$f'(p, \gamma'(0)) \leq \frac{f(\gamma(t)) - f(p)}{t} \quad (\text{III.4.9})$$

para todo $t > 0$. Como $f'(p, \cdot): T_pM \rightarrow \mathbb{R}$ é convexa e $f'(p, 0) = 0$, segue do Teorema III.3.19 (ou da Proposição IV-1.2.1, pag. 147, de [28]) que existe $\bar{s} \in T_pM$ tal que

$$f'(p, v) \geq f'(p, 0) + \langle \bar{s}, v \rangle$$

para todo $v \in T_pM$. Substituindo esta última desigualdade, com $v = \gamma'(0)$, em (III.4.9) temos

$$f(\gamma(t)) \geq f(p) + t\langle \bar{s}, \gamma'(0) \rangle$$

para todo $t > 0$. Deste modo, segue da Definição III.3.16 que $\bar{s} \in \partial f(p)$. □

Observação III.4.10. Do mesmo modo que no \mathbb{R}^n , podemos mostrar que $\partial f(p)$ é convexo e compacto e $\partial f(p) \subset B(0, L)$, onde $L = L(p)$ é a constante de Lipschitz de f em p .

Proposição III.4.11. *Sejam M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Então para cada $p \in M$ temos*

i) $f'(p, v) = \max_{s \in \partial f(p)} \langle s, v \rangle$ para cada $v \in T_pM$.

ii) $\partial f(p) = \{s \in T_pM / f'(p, v) \geq \langle s, v \rangle, \forall v \in T_pM\}$.

Demonstração. i) Sejam $v \in T_p M$. Seja γ_v a geodésica tal que $\gamma_v(0) = p$; da definição de subgradiente, Definição III.3.16, temos

$$\frac{f(\gamma_v(t)) - f(p)}{t} \geq \langle s, v \rangle \quad (\text{III.4.10})$$

para todo $t > 0$ e todo $s \in \partial f(p)$. Da última desigualdade e (III.4.2) segue-se que $f'(p, v) \geq \max_{s \in \partial f(p)} \langle s, v \rangle$. Agora suponha por absurdo que exista $v_1 \in T_p M$ tal que $f'(p, v_1) > \max_{s \in \partial f(p)} \langle s, v_1 \rangle$. Pelo teorema de Hahn-Banach existe $\bar{s} \in T_p M$ tal que, para todo $v \in T_p M$, $f'(p, v) \geq \langle \bar{s}, v \rangle$ e $f'(p, v_1) = \langle \bar{s}, v_1 \rangle$. Deste modo, para todo $v \in T_p M$, segue-se de (III.4.2) que $f(\gamma_v(t)) - f(p) \geq t f'(p, v) \geq t \langle \bar{s}, v \rangle$ para todo $t \geq 0$ e isto implica que $\bar{s} \in \partial f(p)$. Agora temos

$$\begin{aligned} f'(p, v_1) &> \max_{s \in \partial f(p)} \langle s, v_1 \rangle \\ &\geq \langle \bar{s}, v_1 \rangle \\ &= f'(p, v_1) \end{aligned}$$

o que é absurdo. Isto prova i).

ii) Seja $\Gamma := \{s \in T_p M / f'(p, v) \geq \langle s, v \rangle, \forall v \in T_p M\}$. Tome $s \in \Gamma$. Para todo $v \in T_p M$ e $t > 0$ temos

$$\begin{aligned} t \langle s, v \rangle &\leq f'(p, tv) \\ &= \lim_{\lambda \rightarrow 0^+} \frac{f(\gamma_{tv}(\lambda)) - f(p)}{\lambda} \\ &\leq \lim_{\lambda \rightarrow 0^+} \frac{(1 - \lambda)f(\gamma_{tv}(0)) + \lambda f(\gamma_{tv}(1)) - f(p)}{\lambda} \\ &= \lim_{\lambda \rightarrow 0^+} \frac{(1 - \lambda)f(p) + \lambda f(\gamma_v(t)) - f(p)}{\lambda} \\ &= f(\gamma_v(t)) - f(p) \end{aligned}$$

onde γ_v é a geodésica tal que $\gamma_v(0) = p$ e $\gamma'_v(0) = v$. Isto implica que $s \in \partial f(p)$. Então mostramos que $\Gamma \subset \partial f(p)$. Agora tome $s \in \partial f(p)$. Para todo $v \in T_p M$ temos

$$\begin{aligned} f'(p, v) &= \lim_{t \rightarrow 0^+} \frac{f(\gamma_v(t)) - f(p)}{t} \\ &\geq \lim_{t \rightarrow 0^+} \frac{t \langle s, v \rangle}{t} \\ &= \langle s, v \rangle \end{aligned}$$

onde γ_v é a geodésica tal que $\gamma_v(0) = p$. Isto implica que $\partial f(p) \subset \Gamma$. Portanto $\Gamma = \partial f(p)$. \square

Seja M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função convexa. Dada uma geodésica $\gamma: \mathbb{R} \rightarrow M$ considere a composição

$$\begin{aligned} \varphi &= f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R} \\ t &\mapsto \varphi(t) = f(\gamma(t)) \end{aligned} \tag{III.4.11}$$

Desejamos calcular $\partial \varphi$.

Lema III.4.12. (*Regra da cadeia*). *O subdiferencial da função φ definida em (III.4.11) é*

$$\begin{aligned} \partial \varphi(t) &= \{ \langle s, \gamma'(t) \rangle / s \in \partial f(\gamma(t)) \} \\ &:= \langle \partial f(\gamma(t)), \gamma'(t) \rangle \end{aligned}$$

Demonstração. Pela Definição III.4.1 temos

$$\begin{aligned} \varphi'(t, 1) &= \lim_{\lambda \rightarrow 0^+} \frac{\varphi(t + \lambda) - \varphi(t)}{\lambda} \\ &= \lim_{\lambda \rightarrow 0^+} \frac{f(\gamma(t + \lambda)) - f(\gamma(t))}{\lambda} \\ &= f'(\gamma(t), \gamma'(t)) \end{aligned} \tag{III.4.12}$$

$$\begin{aligned}
\varphi'(t, -1) &= \lim_{\lambda \rightarrow 0^+} \frac{\varphi(t - \lambda) - \varphi(t)}{\lambda} \\
&= \lim_{\lambda \rightarrow 0^+} \frac{f(\gamma(t - \lambda)) - f(\gamma(t))}{\lambda} \\
&= f'(\gamma(t), -\gamma'(t)).
\end{aligned} \tag{III.4.13}$$

Sabemos que $\partial\varphi(t) = [-\varphi'(t, -1), \varphi'(t, 1)]$ (veja capítulo I de [28]), e da Proposição III.4.11 temos que

$$f'(\gamma(t), \gamma'(t)) = \max_{s \in \partial f(\gamma(t))} \langle s, \gamma'(t) \rangle$$

e

$$-f'(\gamma(t), -\gamma'(t)) = \min_{s \in \partial f(\gamma(t))} \langle s, \gamma'(t) \rangle.$$

Então de (III.4.12) e (III.4.13) obtemos que

$$\begin{aligned}
\varphi(t) &= \left[\min_{s \in \partial f(\gamma(t))} \langle s, \gamma'(t) \rangle, \max_{s \in \partial f(\gamma(t))} \langle s, \gamma'(t) \rangle \right] \\
&= \{ \langle s, \gamma'(t) \rangle / s \in \partial f(\gamma(t)) \}.
\end{aligned}$$

A última igualdade se deve à convexidade do conjunto $\partial f(\gamma(t))$. □

CAPÍTULO IV

Campos monótonos

IV.1 Introdução

Neste capítulo introduziremos o conceito de campos monótonos, estritamente monótonos e fortemente monótonos, em variedades Riemannianas, que generaliza o conceito de operadores monótonos, estritamente monótonos e fortemente monótonos, respectivamente, em \mathbb{R}^n (veja Iusem [30] e Ortega-Rheinboldt [35] para a definição de operadores monótonos). Daremos uma caracterização destes campos onde explicitaremos seu significado geométrico e estudaremos suas relações com as funções convexas. Campos monótonos existem em abundância. Alguns exemplos são os campos gradientes de funções convexas e os campos de Killing (observamos que estes últimos não são gradientes). Mostraremos que o campo gradiente do quadrado da função distância, em uma variedade de Hadamard, é um campo fortemente monótono. Este fato será fundamental, no Capítulo V, para definir o método de ponto proximal para encontrar singularidades de campos monótonos. Introduziremos também o conceito de campos monótonos ponto-conjunto em variedades Riemannianas que generaliza o conceito de operadores monótonos ponto-conjunto em \mathbb{R}^n , veja Iusem [30]. Como no caso

contínuo, daremos uma caracterização geométrica dos campos monótonos e mostraremos que os subdiferenciais de funções convexas são exemplos de campos monótonos ponto-conjunto.

IV.2 Campos monótonos contínuos

Definiremos, nesta seção, campos monótonos, estritamente monótonos e fortemente monótonos em variedades Riemannianas que, via o transporte paralelo, generaliza de modo natural o conceito de operadores monótonos, estritamente monótonos e fortemente monótonos, respectivamente, em \mathbb{R}^n , veja Iusem [30] e Rheinboldt [35]. A definição algébrica, via o transporte paralelo, deixa oculto seu significado geométrico o qual será explicitado pela Proposição IV.2.6. Veremos que o conceito de convexidade está relacionado com o conceito de monotonicidade, isto é, veremos que uma função de classe C^1 é convexa, estritamente convexa e fortemente convexa se, e somente se, o seu campo gradiente é monótono, estritamente monótono e fortemente monótono, respectivamente. Daremos alguns exemplos de campos monótonos. Em particular, mostraremos que os campos de Killing são monótonos. Também mostraremos que o campo gradiente do quadrado da função distância é fortemente monótono. Finalmente mostraremos um modo de construir campos monótonos.

Definição IV.2.1. Seja M uma variedade Riemanniana completa e seja $X \in \mathfrak{X}^0(M)$. Dizemos que o campo X é *monótono* se

$$\langle \gamma'(0), P_{p'}^{-1}X(p) - X(p') \rangle \geq 0 \quad (\text{IV.2.1})$$

para todo p e $p' \in M$ e toda geodésica γ ligando p a p' tal que $\gamma(0) = p$, onde $P_{pp'}$ é o transporte paralelo ao longo de γ . Dizemos ainda que X é *estritamente monótono* se a desigualdade (IV.2.1) é estrita.

No caso particular em que M é uma variedade de Hadamard a Definição IV.2.1 se reduz a

Definição IV.2.2. Seja M uma variedade de Hadamard e seja $X \in \mathcal{X}^0(M)$. Dizemos que o campo X é *monótono*, se

$$\langle \exp_{p'}^{-1} p, P_{p'p}^{-1} X(p) - X(p') \rangle \geq 0 \quad (\text{IV.2.2})$$

para todo p e $p' \in M$, onde $P_{p'p}^{-1}$ é o transporte paralelo ao longo da geodésica que liga p' e p . Dizemos ainda que X é *estritamente monótono* se a desigualdade (IV.2.2) é estrita.

Observação IV.2.3. No caso particular que $M = \mathbb{R}^n$ com a métrica usual, a desigualdade (IV.2.1), e consequentemente (IV.2.2), se reduz a

$$\langle p - p', X(p) - X(p') \rangle \geq 0$$

pois, $\exp_{p'}^{-1} p = p - p'$ e $P_{p'p}^{-1} = I$. Portanto a Definição IV.2.1 coincide com a definição usual de operador monótono no caso que $M = \mathbb{R}^n$ e $X: \mathbb{R}^n \rightarrow \mathbb{R}^n$, veja Iusem [30] e Ortega-Rheinboldt [35].

Definição IV.2.4. uma função $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ é chamada *monótona* (respectivamente, *estritamente monótona*) se, $(t_2 - t_1)(\varphi(t_2) - \varphi(t_1)) \geq 0$ (respectivamente, $(t_1 - t_2)(\varphi(t_1) - \varphi(t_2)) > 0$) para todo $t_1, t_2 \in \mathbb{R}$, $t_1 \neq t_2$.

Observação IV.2.5. Na Definição IV.2.4 não usamos a terminologia monótona crescente porque não trabalharemos com função monótona decrescente.

Proposição IV.2.6. *Seja M uma variedade Riemanniana completa e $X \in \mathcal{X}^0(M)$. O campo X é monótono (respectivamente, estritamente monótono) se, e somente se, a função*

$$\begin{aligned} \varphi_{(X,\gamma)}: \mathbb{R} &\rightarrow \mathbb{R} & (\text{IV.2.3}) \\ t &\mapsto \varphi_{(X,\gamma)}(t) = \langle \gamma'(t), X(\gamma(t)) \rangle \end{aligned}$$

é monótona (respectivamente, estritamente monótona) para toda geodésica γ de M .

Demonstração. (\Rightarrow) Seja γ uma geodésica de M . Sejam $t_1 \neq t_2$ e considere a geodésica $\alpha(t) = \gamma(t_1 + t(t_2 - t_1))$, que é uma reparametrização de γ . Seja $p' = \alpha(0) = \gamma(t_1)$, $p = \alpha(1) = \gamma(t_2)$ e $P_{p'p}$ o transporte paralelo ao longo de α . Como $P_{p'p}$ é uma isometria temos

$$\begin{aligned} (t_2 - t_1)(\varphi_{(X,\gamma)}(t_2) - \varphi_{(X,\gamma)}(t_1)) &= (t_2 - t_1)(\langle \gamma'(t_2), X(\gamma(t_2)) \rangle - \langle \gamma'(t_1), X(\gamma(t_1)) \rangle) \\ &= \langle \alpha'(1), X(p) \rangle - \langle \alpha'(0), X(p') \rangle \\ &= \langle P_{p'p}^{-1} \alpha'(1), P_{p'p}^{-1} X(p) \rangle - \langle \alpha'(0), X(p') \rangle \\ &= \langle \alpha'(0), P_{p'p}^{-1} X(p) - X(p') \rangle \end{aligned}$$

Deste modo, $(t_2 - t_1)(\varphi_{(X,\gamma)}(t_2) - \varphi_{(X,\gamma)}(t_1)) = \langle \alpha'(0), P_{p'p}^{-1} X(p) - X(p') \rangle \geq 0$, pois X é monótono. Portanto, $\varphi_{(X,\gamma)}$ é monótona para toda geodésica γ de M . A mesma prova é feita no caso de monótona estrita.

(\Leftarrow) Dados $p, p' \in M$, $p \neq p'$, seja γ uma geodésica ligando p a p' com $\gamma(0) = p'$ e $\gamma(1) = p$. Seja $P_{p'p}$ o transporte paralelo ao longo de γ . Assim

$$\begin{aligned} \langle \gamma'(0), P_{p'p}^{-1} X(p) - X(p') \rangle &= \langle P_{p'p}^{-1} \gamma'(1), P_{p'p}^{-1} X(p) \rangle - \langle \gamma'(0), X(p') \rangle \\ &= \langle \gamma'(1), X(\gamma(1)) \rangle - \langle \gamma'(0), X(\gamma(0)) \rangle \\ &= (1 - 0)(\varphi_{(X,\gamma)}(1) - \varphi_{(X,\gamma)}(0)) \end{aligned}$$

Deste modo $\langle \gamma'(0), P_{p'p}^{-1} X(p) - X(p') \rangle = \varphi_{(X,\gamma)}(1) - \varphi_{(X,\gamma)}(0) \geq 0$, pois $\varphi_{(X,\gamma)}$ é monótona. Portanto, X é monótono. A mesma prova é feita no caso de monótona estrita. \square

A equivalência dada pelo Teorema IV.2.6 irá nos deixar livres de certas manipulações algébricas comuns quando se trabalha com a Definição IV.2.1, além de explicitar o significado geométrico oculto nesta definição. A demonstração da próxima proposição exemplificará isto. Compare com a demonstração usual do \mathbb{R}^n , veja Ortega-Rheinboldt [35].

Proposição IV.2.7. *Seja M uma variedade Riemanniana completa e seja $f: M \rightarrow \mathbb{R}$ uma função de classe C^1 . A função f é convexa se, e somente se o campo $\text{grad } f$ é monótono.*

Demonstração. Considere a função $\psi = f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$, onde γ é uma geodésica de M . A função ψ é convexa se, e somente se ψ' é monótona. Mas $\psi' = \langle \gamma', \text{grad } f \circ \gamma \rangle = \varphi_{(\text{grad } f, \gamma)}$, então $\psi = f \circ \gamma$ é convexa se, e somente se $\varphi_{(\text{grad } f, \gamma)}$ é monótona. Deste modo, f é convexa se, e somente se $\varphi_{(\text{grad } f, \gamma)}$ é monótona para toda geodésica γ . Portanto, pelo Teorema IV.2.6, f é convexa se, e somente se o campo $\text{grad } f$ é monótono. \square

Proposição IV.2.8. *Seja M uma variedade Riemanniana completa e seja $f: M \rightarrow \mathbb{R}$ uma função de classe C^1 . A função f é estritamente convexa se, e somente se o campo $\text{grad } f$ é estritamente monótono.*

Demonstração. Análoga a demonstração da Proposição IV.2.7. \blacksquare

Observação IV.2.9. As proposições acima nos fornecem uma classe de exemplos de campos monótonos, a saber, aqueles campos que são gradientes de funções convexas. Existem outros exemplos de campos monótonos que não são gradientes de funções convexas. Nosso primeiro exemplo será na esfera, onde sabemos que as únicas funções convexas existentes são as constantes e portanto o único campo gradiente monótono é o campo identicamente nulo.

Exemplo IV.2.10. (Campo monótono na esfera). Considere a esfera $S^n = \{p \in \mathbb{R}^{n+1} / \langle p, p \rangle = 1\}$ e seja A uma matriz anti-simétrica $(n+1) \times (n+1)$. O campo $X(p) = Ap$ é um campo monótono em S^n . De fato, seja γ uma geodésica de S^n , então

$$\varphi_{(X, \gamma)}(t) = \langle \gamma'(t), X(\gamma(t)) \rangle$$

A derivada covariante de X na direção $v \in T_p M$ é igual a $\nabla_v X(p) = (I - pp^T)Av$, onde p^T denota o transposto do vetor $p \in \mathbb{R}^{n+1}$. Deste modo,

$$\frac{d}{dt} \varphi_{(X, \gamma)}(t) = \langle \gamma''(t), X(\gamma(t)) \rangle + \langle \gamma'(t), \nabla_{\gamma'(t)} X \rangle = \langle \gamma'(t), \nabla_{\gamma'(t)} X \rangle \quad (\text{IV.2.4})$$

Como $\langle v, \nabla_v X \rangle = \langle v, (I - pp^T)Av \rangle = 0$, pois A é anti-simétrica e $p^T v = 0$, segue-se de (IV.2.4) que $\frac{d}{dt}\varphi_{(X,\gamma)}(t) = 0$. Portanto, $\varphi_{(X,\gamma)}(t) \equiv \text{cte}$ para toda geodésica γ , então segue do Teorema IV.2.6 que X é monótono.

Na verdade o exemplo IV.2.10 é um caso particular do exemplo mais geral a seguir.

$$Y \mapsto A_X(Y) := \nabla_Y X$$

Exemplo IV.2.11. Seja M uma variedade Riemanniana completa e seja $X \in \mathfrak{X}^1(M)$. Dizemos que X é *campo de Killing* se

$$\langle \nabla_Y X, Z \rangle + \langle \nabla_Z X, Y \rangle = 0 \tag{IV.2.5}$$

para todo $Y, Z \in \mathfrak{X}^0(M)$, isto é, se A_X é um operador anti-simétrico (veja Definição II.4.1). Seja γ uma geodésica de M , da equação (IV.2.5) segue-se que $\frac{d}{dt}\varphi_{(X,\gamma)}(t) = \langle \gamma'(t), \nabla_{\gamma'(t)} X \rangle \equiv 0$ e isto implica que $\varphi_{(X,\gamma)}(t) \equiv \text{cte}$. Portanto do Teorema IV.2.6 segue-se que X é monótono. No exemplo anterior $A_X(p) = (I - pp^T)A$ é uma matriz anti-simétrica.

Exemplo IV.2.12. Seja M uma variedade Riemanniana completa e seja $X \in \mathfrak{X}^1(M)$. Dizemos que X é um *campo concorrente* se

$$\nabla_Y X = Y$$

para todos $Y \in \mathfrak{X}^0(M)$, isto é, A_X é o operador identidade; veja [52] (veja Definição II.4.1). Com o mesmo argumento usado no exemplo IV.2.11 podemos mostrar que todo campo concorrente é monótono.

Observação IV.2.13. i) Seja $X \in \mathfrak{X}^0(M)$ um campo monótono. Suponhamos que γ seja uma geodésica fechada em M . Então

$$\varphi_{(X,\gamma)}(t) = \langle \gamma'(t), X(\gamma(t)) \rangle \equiv \text{cte}.$$

Disso segue-se que se M tem uma geodésica fechada, e portanto não podemos definir um campo $X \in \mathfrak{X}^0(M)$ estritamente monótono. Em particular, não podemos definir um campo estritamente monótono em variedades compactas. Intuitivamente podemos esperar que a existência de um campo estritamente monótono numa variedade Riemanniana impõe restrições sobre a mesma como ocorre com a existência de funções convexas, veja [5], [45] e [54]. ii)¹ Se M^n é completa e não-compacta com $K \geq 0$ e admite um campo estritamente monótono, então M é difeomorfa a \mathbb{R}^n . Caso contrário, dimensão (alma) ≥ 1 e existe geodésica fechada em M . Já sabemos que as únicas funções convexas em variedades completas de volume finito são as funções constantes; veja [5]. Assim, da Proposição IV.2.7 segue-se que o único campo gradiente monótono em variedades de volume finito é o campo identicamente nulo. Em particular o único campo gradiente monótono na esfera é o campo identicamente nulo.

Definição IV.2.14. Seja M uma variedade Riemanniana completa e seja $X \in \mathfrak{X}^0(M)$. Dizemos que o campo X é *fortemente monótono com módulo* $\lambda > 0$ se,

$$\langle \hat{t}\gamma'(0), P_{p'p}^{-1}X(p) - X(p') \rangle \geq \lambda \|\gamma'(0)\|^2 \hat{t}^2 \quad (\text{IV.2.6})$$

para todo ponto p e $p' \in M$, $p \neq p'$ e toda geodésica γ ligando p a p' com $\gamma(0) = p$ e $\gamma(\hat{t}) = p'$, onde $P_{p'p}$ é o transporte paralelo ao longo de γ .

Observe que $\|\gamma'(0)\|\hat{t}$ é igual ao comprimento da geodésica que liga p a p' , assim a Definição IV.2.14 se simplifica no caso que M é uma variedade de Hadamard.

Definição IV.2.15. Seja M uma variedade de Hadamard e seja $X \in \mathfrak{X}^0(M)$. Dizemos que o campo X é *fortemente monótono com módulo* $\lambda \geq 0$ se

$$\langle \exp_{p'}^{-1} p, P_{p'p}^{-1}X(p) - X(p') \rangle \geq \lambda d^2(p, p') \quad (\text{IV.2.7})$$

¹Esta observação foi feita pelo professor Sérgio J. X. de Mendonça.

para todo p e $p' \in M$, onde $P_{p'p}$ é o transporte paralelo ao longo da única geodésica γ que liga p a p' e d é a distância Riemanniana.

Observação IV.2.16. No caso particular que $M = \mathbb{R}^n$, com a métrica usual, a desigualdade (IV.2.6), e conseqüentemente (IV.2.7), se reduzem a

$$\langle p - p', X(p) - X(p') \rangle \geq \lambda \|p - p'\|^2$$

pois $\gamma'(0)\hat{t} = \exp_{p'}^{-1} p = p - p'$ e $P_{p'p}^{-1} = I$. Portanto a Definição IV.2.15 coincide com a definição usual de operador fortemente monótono no caso que $M = \mathbb{R}^n$ e $X: \mathbb{R}^n \rightarrow \mathbb{R}^n$, veja [35].

Proposição IV.2.17. *Seja M uma variedade Riemanniana completa e seja $X \in \mathfrak{X}^0(M)$. O campo X é fortemente monótono com módulo $\lambda > 0$ se, e somente se, a função $\varphi_{(X,\gamma)}(t) - \lambda \|\gamma'\|^2 t$ é monótona para toda geodésica γ , onde $\varphi_{(X,\gamma)}$ é definida em (IV.2.3).*

Demonstração. (\Rightarrow) Seja γ uma geodésica de M . Seja $t_1 \neq t_2$ e considere a geodésica $\alpha(t) = \gamma(t_1 + t(t_2 - t_1))$ reparametrização de γ . Seja $p' = \alpha(0) = \gamma(t_1)$, $p = \alpha(1) = \gamma(t_2)$ e $P_{p'p}$ o transporte paralelo ao longo de α . Como $P_{p'p}$ é isometria temos

$$\begin{aligned} & (t_2 - t_1)[(\varphi_{(X,\gamma)}(t_2) - \lambda \|\gamma'(0)\|^2 t_2) - (\varphi_{(X,\gamma)}(t_1) - \lambda \|\gamma'(0)\|^2 t_1)] \\ &= (t_2 - t_1)(\langle \gamma'(t_2), X(\gamma(t_2)) \rangle - \langle \gamma'(t_1), X(\gamma(t_1)) \rangle - \lambda \|\gamma'(0)\|^2 (t_2 - t_1)) \\ &= \langle \alpha'(1), X(p) \rangle - \langle \alpha'(0), X(p') \rangle - \lambda \|\alpha'(0)\|^2 \\ &= \langle P_{p'p}^{-1} \alpha'(1), P_{p'p}^{-1} X(p) \rangle - \langle \alpha'(0), X(p') \rangle - \lambda \|\alpha'(0)\|^2 \\ &= \langle \alpha'(0), P_{p'p}^{-1} X(p) - X(p') \rangle - \lambda \|\alpha'(0)\|^2. \end{aligned}$$

Sendo X fortemente monótono segue-se da igualdade anterior que $\varphi_{(X,\gamma)}(t) - \lambda \|\gamma'(0)\|^2 t$ é monótona para toda geodésica γ .

(\Leftarrow) Dados $p, p' \in M$, $p \neq p'$ e γ uma geodésica ligando p a p' com $\gamma(0) = p$ e $\gamma(\hat{t}) = p'$. Considere a geodésica $\alpha(t) = \gamma(\hat{t}t)$, reparametrização de γ e $P_{p'p}$ o transporte paralelo ao longo de γ . Como $P_{p'p}$ é isometria temos

$$\begin{aligned} & \langle \hat{t}\gamma'(0), P_{p'p}^{-1}X(p) - X(p') \rangle - \lambda \|\gamma'(0)\|^2 \hat{t}^2 \\ &= \langle P_{p'p}^{-1}\hat{t}\gamma'(\hat{t}), P_{p'p}^{-1}X(p) \rangle - \lambda \|\gamma'(0)\|^2 \hat{t}^2 - \langle \hat{t}\gamma'(0), X(p') \rangle \\ &= (\langle \alpha'(1), X(p) \rangle - \lambda \|\alpha'(0)\|^2) - \langle \alpha'(0), X(p') \rangle \\ &= (1 - 0)[(\varphi_{(X,\alpha)}(1) - \lambda \|\alpha'(0)\|^2) - (\varphi_{(X,\alpha)}(0))]. \end{aligned}$$

Sendo $\varphi_{(X,\alpha)}(t) - \lambda \|\alpha'(0)\|^2 t$ monótona segue-se da desigualdade anterior que X é fortemente monótono com módulo $\lambda \geq 0$. \square

Definição IV.2.18. Seja M uma variedade Riemanniana completa. Uma função $f: M \rightarrow \mathbb{R}$ é dita fortemente convexa com módulo $\lambda > 0$ se para toda geodésica $\gamma: \mathbb{R} \rightarrow M$ a composição $f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$ é uma função fortemente convexa com módulo $\lambda \|\gamma'(0)\|^2 > 0$.

Proposição IV.2.19. Uma função $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ é fortemente convexa com módulo $\tilde{\lambda} > 0$ se, e somente se, a função $\psi: \mathbb{R} \rightarrow \mathbb{R}$ dada por $\psi(t) = \varphi(t) - \frac{1}{2}\tilde{\lambda}t^2$ é convexa.

Demonstração. Imediata. \square

Proposição IV.2.20. Seja M uma variedade Riemanniana completa e seja $f: M \rightarrow \mathbb{R}$ uma função de classe C^1 . A função f é fortemente convexa se, e somente se o campo $\text{grad } f$ é fortemente monótono.

Demonstração. Usando o Teorema IV.2.17 e a Proposição IV.2.19 a demonstração é análoga à Proposição IV.2.7. \square

Proposição IV.2.21. Sejam M uma variedade de Hadamard e $p' \in M$. Então o campo $\text{grad } \rho_{p'}$ é fortemente monótono com módulo $\lambda = 1$ para todo $p' \in M$, onde $\rho_{p'}$ é definida em (II.8.5)

Demonstração. Seja γ uma geodésica de M , e seja $\varphi_\gamma: \mathbb{R} \rightarrow \mathbb{R}$ dada por $\varphi_\gamma(t) = \varphi_{(X,\gamma)}(t) - \|\gamma'(0)\|^2 t$, onde $X = \text{grad } \rho_{p'}$ e $\varphi_{(X,\gamma)}$ é definida em (IV.2.3). Então de acordo com o Teorema IV.2.17 basta mostrar que φ_γ é monótona, isto é, $(t_2 - t_1)(\varphi_\gamma(t_2) - \varphi_\gamma(t_1)) \geq 0$ para todo $t_1, t_2 \in \mathbb{R}$. Primeiro suponhamos que γ passa por p' . Deste modo existe $t_0 \in \mathbb{R}$ tal que $\gamma(t_0) = p'$. Sem perda de generalidade vamos supor que $t \geq t_0$. Assim

$$\begin{aligned} \varphi_\gamma(t) &= \langle \gamma'(t), \text{grad } \rho_{p'}(\gamma(t)) \rangle - \|\gamma'(0)\|^2 t \\ &= \langle \gamma'(t), -\exp_{\gamma(t)}^{-1} p' \rangle - \|\gamma'(0)\|^2 t \\ &= \|\gamma'(0)\|^2 (t - t_0) - \|\gamma'(0)\|^2 t \\ &= -\|\gamma'(0)\|^2 t_0. \end{aligned}$$

Portanto, φ_γ é monótona. Agora suponhamos que γ não passa pelo ponto p' . Dados $t_1, t_2 \in \mathbb{R}$, considere o triângulo geodésico $\Delta(p_1 p_2 p_3)$, onde $p_1 = \gamma(t_1)$, $p_2 = \gamma(t_2)$ e $p_3 = p'$. Denote por $\gamma_{i+1}: [0, \ell_{i+1}] \rightarrow M$ o segmento de geodésica ligando p_{i+1} a p_{i+2} , $\ell_{i+1} = L(\gamma_{i+1}) = d(p_{i+1}, p_{i+2})$ e $\theta_{i+1} = \sphericalangle(\gamma'_{i+1}(0), -\gamma'_{i+3}(\ell_{i+3}))$, onde $i = 1, 2, 3 \pmod{3}$ e $|\gamma'_+(0)| = 1$; $i = 2, 3$. Com esta notação segue-se, que $\gamma = \gamma_1$, e da Proposição (II.8.3) segue-se que $\gamma'_2(0) = \exp_{p_2}^{-1} p' = -\text{grad } \rho_{p'}(p_2)$ e $\gamma'_2(\ell_2) = -\exp_{p_1}^{-1} p' = \text{grad } \rho_{p'}(p_1)$. Suponhamos, sem perda de generalidade, que $t_1 < t_2$. Como $\gamma = \gamma_1$ temos $\ell_1 = (t_2 - t_1)\|\gamma'_1(0)\| = (t_2 - t_1)\|\gamma'(0)\|$.

Deste modo

$$\begin{aligned}
& (t_2 - t_1)(\varphi_\gamma(t_2) - \varphi_\gamma(t_1)) \\
&= (t_2 - t_1)[(\langle \gamma'(t_2), \text{grad } \rho_{p'}(\gamma(t_2)) \rangle - \|\gamma'(0)\|^2 t_2) - (\langle \gamma'(t_1), \text{grad } \rho_{p'}(\gamma(t_1)) \rangle - \|\gamma'(0)\|^2 t_1)] \\
&= (t_2 - t_1)[(\langle \gamma'(t_2), -\gamma'_2(0) \rangle - \|\gamma'(0)\|^2 t_2) - (\langle \gamma'(t_1), \gamma'_3(\ell_3) \rangle - \|\gamma'(0)\|^2 t_1)] \\
&= (t_2 - t_1)(\langle -\gamma'_1(\ell_1), \gamma'_2(0) \rangle + \langle \gamma'_1(0), -\gamma'_3(\ell_3) \rangle - \|\gamma'_1(0)\|^2 (t_2 - t_1)) \\
&= (t_2 - t_1)\|\gamma'_3(\ell_3)\| \|\gamma'_1(0)\| \cos \theta_2 + (t_2 - t_1)\|\gamma'_3(0)\| \|\gamma'_2(\ell_2)\| \cos \theta_1 - (t_2 - t_1)^2 \|\gamma'_3(0)\|^2 \\
&= \ell_1 \ell_2 \cos \theta_2 + \ell_1 \ell_3 \cos \theta_1 - \ell_1^2 \\
&= \ell_1(\ell_2 \cos \theta_2 + \ell_3 \cos \theta_1 - \ell_1).
\end{aligned}$$

Assim $(t_2 - t_1)(\varphi_\gamma(t_2) - \varphi_\gamma(t_1)) = \ell_1(\ell_2 \cos \theta_2 - \ell_3 \cos \theta_1 - \ell_1) \geq 0$ pela desigualdade (II.8.3), fazendo $i = 2$. O caso $t_1 > t_2$ se faz de modo análogo. Portanto φ_γ é monótono para toda geodésica γ e isto implica que $\text{grad } \rho_{p'}$ é fortemente monótono com módulo $\lambda = 1$. \square

Veremos agora um teorema de caracterização para campos monótonos em termos do operador A_X , Definição II.4.1.

Teorema IV.2.22. *Seja M uma variedade Riemanniana completa e seja $X \in \mathfrak{X}^1(M)$.*

Então:

- i) O campo X é monótono se, e somente se, $\langle A_X(p)v, v \rangle \geq 0$ para todo $p \in M$ e todo $v \in T_p M$.*
- ii) Se $\langle A_X(p)v, v \rangle > 0$ para todo $p \in M$ e todo $v \in T_p M \setminus \{0\}$, então o campo X é estritamente monótono.*
- iii) O campo X é fortemente monótono com módulo $\lambda > 0$ se, e somente se, $\langle A_X(p)v, v \rangle \geq \lambda \|v\|^2$ para todo $p \in M$ e todo $v \in T_p M$.*

Demonstração. Faremos só a demonstração de iii); os outros itens são análogos. Seja γ uma geodésica de M e $\varphi_\gamma(t) = \varphi_{(X,\gamma)}(t) - \lambda\|\gamma'(0)\|^2 t$, onde $\varphi_{(X,\gamma)}$ é definida em (IV.2.3). Do Teorema IV.2.17 o campo X é fortemente monótono com módulo $\lambda \geq 0$ se, e somente se, φ_γ é monótono para toda geodésica γ . A aplicação φ_γ é monótona se, e somente se,

$$\varphi'_\gamma(t) = \langle \gamma'(t), \nabla_{\gamma'(t)} X \rangle - \lambda\|\gamma'(0)\|^2 = \langle \gamma'(t), A_X(\gamma(t)) \cdot \gamma'(t) \rangle - \lambda\|\gamma'(0)\|^2 \geq 0$$

para todo $t \in \mathbb{R}$. Finalmente, $\langle \gamma'(t), A_X(\gamma(t)) \cdot \gamma'(t) \rangle - \lambda\|\gamma'(0)\|^2 \geq 0$ para toda geodésica γ se, e somente se, $\langle v, A_X(p) \cdot v \rangle \geq \lambda\|v\|^2$ para todo ponto $p \in M$ e todo $v \in T_p M$. Portanto, X é fortemente monótono com módulo $\lambda \geq 0$ se, e somente se, $\langle A_X(p) \cdot v, v \rangle \geq \lambda\|v\|^2$ para todo $p \in M$ e todo $v \in T_p M$. \square

Corolário IV.2.23. *Seja M uma variedade Riemanniana completa e $f: M \rightarrow \mathbb{R}$ uma função de classe C^2 . Então*

- i) f é convexa se, e somente se, $\text{Hess } f_p$ é positiva semidefinida para todo $p \in M$.*
- ii) Se $\text{Hess } f_p$ é positiva definida para todo $p \in M$, então f é estritamente convexa.*
- iii) f é fortemente convexa com módulo $\lambda > 0$ se, e somente se, $\langle \text{Hess } f_p \cdot v, v \rangle \geq \lambda\|v\|^2$ para todo $p \in M$ e todo $v \in T_p M$.*

Demonstração. Para demonstrar i), ii) e iii) use o Teorema IV.2.22 e Proposição IV.2.7, Proposição IV.2.8 e Proposição IV.2.20 respectivamente. \square

Observação IV.2.24. Veja também este resultado em [50] e [53].

Corolário IV.2.25. *Sejam M uma variedade de Hadamard e $p' \in M$. Então a aplicação $\rho_{p'}$, definida em (II.8.5) é fortemente convexa. Além disso, $\langle (\text{Hess } \rho_{p'})_p \cdot v, v \rangle \geq \|v\|^2$ para todo $p \in M$ e $v \in T_p M$.*

Demonstração. Segue-se da Proposição IV.2.20, Proposição IV.2.21 e Corolário IV.2.25. \square

Finalmente mostraremos como construir campos monótonos a partir de um grupo a 1-parâmetro de difeomorfismos.

Definição IV.2.26. Seja M uma variedade Riemanniana completa. Um difeomorfismo $\psi: M \rightarrow M$ de classe C^1 é *expansivo* se

$$\begin{aligned} \psi^*(p)\langle u, v \rangle &:= \langle d\psi_p(u), d\psi_p(v) \rangle_{\psi(p)} \\ &\geq \langle u, v \rangle_p \end{aligned} \quad (\text{IV.2.8})$$

para todo $p \in M$, $u, v \in T_p M$ tal que $\angle(u, v) < \pi/2$. Usaremos a notação $\psi^*g \geq g$, onde $g := \langle \cdot, \cdot \rangle$ é o tensor métrico.

Seja M uma variedade riemanniana completa e $\{\psi_t\}_{t \in \mathbb{R}}$ um grupo a 1-parâmetro de difeomorfismos de classe C^1 , isto é,

- i) $\psi_0(\cdot) = \text{id}$,
- ii) $\psi_s \circ \psi_t = \psi_{s+t}$ para todo $s, t \in \mathbb{R}$,
- iii) $\psi_t: M \rightarrow M$ é um difeomorfismo de classe C^1 para cada $t \in \mathbb{R}$.

Proposição IV.2.27. ² Se $\{\psi_t\}_{t \in \mathbb{R}}$ é um grupo a 1-parâmetro de difeomorfismos expansivos de classe C^1 , então o campo $X \in \mathfrak{X}^1(M)$ dado por:

$$X(p) := \left. \frac{d}{dt} \psi_t(p) \right|_{t=0} \quad (\text{IV.2.9})$$

é monótono.

²Esta maneira de construir campos monótonos foi sugerida pelo Professor Luis A. Florit.

Para demonstrar esta proposição necessitamos de alguns resultados.

Proposição IV.2.28. *Se $X \in \mathfrak{X}^1(M)$, então*

$$L_X g = \lim_{t \rightarrow 0} \frac{1}{t} (\psi_t^* g - g)$$

onde $\{\psi_t\}_{t \in \mathbb{R}}$ é o fluxo de X e L_X é a derivada de Lie com respeito a X .

Demonstração. Caso particular da proposição 21 do Capítulo 9 de O'Neill [34].

Proposição IV.2.29. *Um campo $X \in \mathfrak{X}^1(M)$ é tal que $L_X g \geq 0$ se, e somente se, seu fluxo $\{\psi_t\}_{t \in \mathbb{R}}$ é expansivo.*

Demonstração. Análoga à demonstração da proposição 23 do Capítulo 9 de O'Neill [34].

□

Proposição IV.2.30. *Um campo $X \in \mathfrak{X}^1(M)$ é tal que $L_X g \geq 0$ se, e somente se, $\langle \nabla_Y X, Z \rangle + \langle \nabla_Z X, Y \rangle \geq 0$ para todo $Y, Z \in \mathfrak{X}^0(M)$.*

Demonstração. Análoga à demonstração da proposição 24 do Capítulo 9 de O'Neill [34].

□

Demonstração da Proposição IV.2.27.

Como o fluxo $\{\psi_t\}_{t \in \mathbb{R}}$ de X é expansivo segue-se da Proposição IV.2.29 e Proposição IV.2.30 que $\langle \nabla_Y X, Z \rangle + \langle \nabla_Z X, Y \rangle \geq 0$ para todo $Y, Z \in \mathfrak{X}^0(M)$. Seja γ uma geodésica de M . Segue-se da desigualdade anterior que $\frac{d}{dt} \varphi_{(X, \gamma)} = \langle \nabla_{\gamma'} X, \gamma' \rangle \geq 0$, onde $\varphi_{(X, \gamma)}(t) = \langle X(\gamma(t)), \gamma'(t) \rangle$. Isto implica que $\varphi_{(X, \gamma)}$ é monótona. Portanto X é monótono. □

IV.3 Campos monótonos ponto-conjunto

Nesta seção definiremos campos monótonos ponto-conjunto em variedades Riemannianas que, via o transporte paralelo, generaliza de modo natural o conceito de operadores monótonos

ponto-conjunto em \mathbb{R}^n ; veja Iusem [30]. O primeiro teorema desta seção fornece uma caracterização destes campos e no final da seção mostraremos que o subdiferencial de uma função convexa é um campo monótono ponto-conjunto.

Definição IV.3.1. Seja M uma variedade Riemanniana. Um *campo ponto-conjunto* é uma aplicação X que a cada $p \in M$ associa um conjunto $X(p) \subset T_p M$, isto é,

$$M \ni p \overset{X}{\mapsto} X(p) \subset T_p M.$$

Observação IV.3.2. Esta definição generaliza a definição usual de campos. Exêmplos de campos ponto-conjunto são o subdiferencial de funções convexas, Definição III.3.16

Definição IV.3.3. Sejam M uma variedade Riemanniana completa e X um campo ponto-conjunto. Dizemos que X é *monótono* se,

$$\langle \gamma'(0), P_{p'}^{-1}u - v \rangle \geq 0 \quad (\text{IV.3.1})$$

para todo p e $p' \in M$, $v \in X(p')$, $u \in X(p)$ e toda geodésica γ ligando p' a p tal que $\gamma(0) = p'$, onde $P_{p'p}$ é o transporte paralelo ao longo de γ .

Se M é uma variedade de Hadamard a Definição IV.3.3 se reduz a

Definição IV.3.4. Seja M uma variedade de Hadamard e X um campo ponto-conjunto. Dizemos que o campo X é monótono se,

$$\langle \exp_{p'}^{-1} p, P_{p'p}^{-1}u - v \rangle \geq 0 \quad (\text{IV.3.2})$$

para todo p e $p' \in M$, $v \in X(p')$ e $u \in X(p)$, onde $P_{p'p}$ é o transporte paralelo ao longo da única geodésica γ que liga p' a p .

Observação IV.3.5. No caso que $M = \mathbb{R}^n$ com a métrica usual, a desigualdade (IV.3.1), e consequentemente (IV.3.2), se reduz a

$$\langle p - p', u - v \rangle \geq 0.$$

pois $\exp_{p'}^{-1} p = p - p'$ e $P_{p'p} = I$. Portanto a Definição IV.3.3 coincide com a definição usual de operador ponto-conjunto monótono, no caso que $M = \mathbb{R}^n$ e $X : \mathbb{R}^n \rightarrow \mathbb{P}(\mathbb{R}^n)$, veja Iusem [30] e Rockafellar [42].

Definição IV.3.6. Uma aplicação $\varphi : \mathbb{R} \rightarrow \mathbb{P}(\mathbb{R})$ é chamada monótona (respectivamente, estritamente monótona) se, $(t_2 - t_1)(r_2 - r_1) \geq 0$ (respectivamente, $(t_2 - t_1)(r_2 - r_1) > 0$) para todo $t_1, t_2 \in \mathbb{R}$, $t_1 \neq t_2$, $r_1 \in \varphi(t_1)$ e $r_2 \in \varphi(t_2)$.

Proposição IV.3.7. *Seja M uma variedade Riemanniana completa e seja X um campo ponto-conjunto. O campo X é monótono (respectivamente, estritamente monótono) se, e somente se, a função*

$$\begin{aligned} \varphi_{(X,\gamma)} : \mathbb{R} &\rightarrow \mathbb{P}(\mathbb{R}) \\ T &\mapsto \varphi_{(X,\gamma)}(t) = \{ \langle \gamma'(t), v \rangle / v \in X(\gamma(t)) \} \end{aligned}$$

é monótona (respectivamente, estritamente monótona) para toda geodésica γ de M .

Demonstração. (\Rightarrow) Seja γ uma geodésica. Sejam $t_1, t_2 \in \mathbb{R}$, $t_1 \neq t_2$, $r_1 \in \varphi(t_1)$ e $r_2 \in \varphi(t_2)$. Considere a geodésica $\alpha(t) = \gamma(t_1 + t(t_2 - t_1))$ reparametrização de γ . Seja $p' = \alpha(0) = \gamma(t_1)$, $p = \alpha(1) = \gamma(t_2)$ e $P_{p'p}$ o transporte paralelo ao longo de α . Como $P_{p'p}$ é uma isometria temos

$$\begin{aligned} (t_2 - t_1)(r_2 - r_1) &= (t_2 - t_1)(\langle \gamma'(t_2), v_2 \rangle - \langle \gamma'(t_1), v_1 \rangle) \\ &= \langle \alpha'(1), v_2 \rangle - \langle \alpha'(0), v_1 \rangle \\ &= \langle P_{p'p}^{-1} \alpha'(1), P_{p'p}^{-1} v_2 \rangle - \langle \alpha'(0), v_1 \rangle \\ &= \langle \alpha'(0), P_{p'p}^{-1} v_2 - v_1 \rangle \end{aligned}$$

para algum $v_1 \in X(\gamma(t_1))$ e $v_2 \in X(\gamma(t_2))$. Deste modo $(t_2 - t_1)(r_2 - r_1) = \langle \alpha'(0), P_{p'p}^{-1} v_2 - v_1 \rangle \geq 0$, pois X é monótono. Portanto, $\varphi_{(X,\gamma)}$ é monótona para toda geodésica γ de M . A mesma prova é feita no caso de monótona estrita.

(\Leftarrow) Dados $p, p' \in M$, $v \in X(p')$, $u \in X(p)$ e γ uma geodésica ligando p a p' com $\gamma(0) = p'$. Reparametrizando γ de modo que $\gamma(0) = p'$ e $\gamma(1) = p$. Seja $P_{p',p}$ o transporte paralelo ao longo de γ . Assim

$$\begin{aligned} \langle \gamma'(0), P_{p',p}^{-1}u - v \rangle &= \langle P_{p',p}^{-1}\gamma'(1), P_{p',p}^{-1}u \rangle - \langle \gamma'(0), v \rangle \\ &= \langle \gamma'(1), u \rangle - \langle \gamma'(0), v \rangle \\ &= (1 - 0)(\langle \gamma'(1), u \rangle - \langle \gamma'(0), v \rangle) \end{aligned}$$

Deste modo $\langle \gamma'(0), P_{p',p}^{-1}u - v \rangle = (1 - 0)(\langle \gamma'(1), u \rangle - \langle \gamma'(0), v \rangle) \geq 0$, pois $\varphi_{(X,\gamma)}$ é monótona. Portanto X é monótono. A mesma prova é feita no caso de convexidade estrita. \square

Teorema IV.3.8. *Seja M uma variedade Riemanniana completa. Se $f: M \rightarrow \mathbb{R}$ é uma função convexa, então o subdiferencial ∂f da função f é um campo ponto-conjunto monótono.*

Demonstração. Que ∂f é um campo ponto-conjunto segue da Definição III.3.16. Do Teorema IV.3.7 basta mostrar que $\varphi_{(\partial f, \gamma)}(t) = \{\langle \gamma'(t), v \rangle / v \in \partial f(\gamma(t))\}$ é monótona para toda geodésica γ . Seja γ uma geodésica de M . Como f é convexa segue-se que $f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$ é convexa e assim $\partial(f \circ \gamma): \mathbb{R} \rightarrow \mathbb{P}(\mathbb{R})$ é monótona. Do Lema III.4.12 temos

$$\partial(f \circ \gamma)(t) = \{\langle \gamma'(t), v \rangle / v \in \partial f(\gamma(t))\} = \varphi_{(\partial f, \gamma)}(t).$$

Portanto $\varphi_{(\partial f, \gamma)}$ é monótona para toda geodésica γ de M . \square

Definição IV.3.9. Seja M uma variedade Riemanniana. Dizemos que $p_* \in M$ é uma *singularidade* do campo ponto-conjunto X se,

$$0 \in X(p_*).$$

Observação IV.3.10. i) Seja X um campo ponto-conjunto monótono em uma variedade Riemanniana M . Suponhamos que γ seja uma geodésica fechada de M . Então

$$\varphi_{(X,\gamma)} = \{\langle \gamma'(t), v \rangle / v \in X(\gamma(t))\} \equiv \{\langle \gamma'(0), v \rangle / v \in X(\gamma(0))\}.$$

Disso segue-se que se M tem uma geodésica fechada, então não podemos definir um campo ponto-conjunto estritamente monótono em M . Em particular, não podemos definir um campo estritamente monótono em variedades compactas. Então intuitivamente, podemos esperar que a existência de um campo ponto-conjunto estritamente monótono numa variedade Riemanniana impõe restrições sobre a mesma como ocorre com a existência de funções convexas veja [5], [45] e [54]. Vale a mesma observação feita na página 45.

Como na seção anterior, pode-se definir campos ponto-conjunto fortemente monótonos e estudar suas relações com as funções fortemente convexas.

CAPÍTULO V

Algoritmos para otimização

V.1 Introdução

Neste capítulo proporemos dois algoritmos para minimizar uma função convexa em uma variedade Riemanniana completa. Estes algoritmos são generalizações, para o contexto de variedades Riemannianas, do algoritmo subgradiente clássico, veja Shor [46] e Hiriart-Urruty e Lemaréchal [28], e do algoritmo de ponto proximal, para minimizar uma função convexa em \mathbb{R}^n , veja Iusem [30]. Para ambos os algoritmos iremos provar suas convergências, mas hipóteses sobre a curvatura da variedade serão necessárias. Provaremos a convergência do algoritmo do subgradiente se a curvatura da variedade for não negativa. E também provaremos a convergência, e boa definição, do algoritmo de ponto proximal se a variedade for Hadamard.

V.2 Algoritmo subgradiente

Seja M uma variedade Riemanniana completa e seja $f: M \rightarrow \mathbb{R}$ uma função convexa. O subconjunto \mathcal{O}^* de M denotará o conjunto dos minimizadores de f e $f^* = \inf_{p \in M} f(p)$ denotará o valor infimo de f . O nosso problema é estimar f^* e também calcular um ponto de \mathcal{O}^* caso exista algum. Isto será feito de modo iterativo pelo *algoritmo subgradiente* que gera uma seqüência $\{p_k\} \subset M$ do seguinte modo: Tome $p_0 \in M$ e defina

$$p_{k+1} = \exp_{p_k} \left(-t_k \frac{s_k}{\|s_k\|} \right) \quad (\text{V.2.1})$$

onde $t_k > 0$, $s_k \in \partial f(p_k)$ e $s_k \neq 0$.

Observação V.2.1. No caso particular que $M = \mathbb{R}^n$, $\exp_{p'} s = p' + s$ e assim, a iteração (V.2.1) se reduz a

$$p_{k+1} = p_k - t_k \frac{s_k}{\|s_k\|}.$$

Logo o algoritmo subgradiente em variedades Riemannianas é uma extensão natural do algoritmo subgradiente introduzido por Shor e generaliza o algoritmo gradiente no caso de ser de f seu diferenciável, estudado por da Cruz Neto-Oliveira [13]

V.2.1 Resultados preliminares

Como é sabido a seqüência obtida pelo algoritmo (V.2.1) não decresce a função. Deve-se portanto escolher uma seqüência $\{t_k\}$ de modo que a respectiva seqüência $\{p_k\}$ se aproxime do conjunto \mathcal{O}^* , como é usual no caso em que $M = \mathbb{R}^n$. Como em [28], para \mathbb{R}^n , obteremos algumas propriedades preliminares, em particular uma cota superior para os t_k 's.

Proposição V.2.2. *Seja $p_* \in \mathcal{O}^* \neq \emptyset$ e seja $B_\varepsilon(p_*)$ uma bola normal em M . Se $p_k \notin \mathcal{O}^*$, $p_k \in B_\varepsilon(p_*)$, e então existe $\delta_k > 0$ tal que escolhendo $0 < t_k < \delta_k$ no algoritmo (V.2.1),*

temos

$$d(p_{k+1}, p^*) < d(p_k, p^*).$$

Demonstração. Seja γ_v a geodésica parametrizada pelo comprimento de arco, tal que $\gamma_v(0) = p_k$ e $\gamma_v(t^*) = p_*$, com $d(p_k, p_*) = t^*$. Considere a esfera normal $S = S_{t^*}(p_*)$, fronteira da bola normal $B_{t^*}(p_*)$, que pelo Lema de Gauss, veja do Carmo [11], é uma subvariedade de codimensão 1 com $T_{p_k}S = \{u \in T_{p_k}M / \langle u, v \rangle = 0\}$. Seja $s_k \in \partial f(p_k)$. Por definição,

$$f \circ \gamma_v(t) \geq f(p_k) + t \langle s_k, v \rangle.$$

Fazendo $t = t^*$ obtemos

$$f(p_*) - f(p_k) \geq t^* \langle s_k, v \rangle.$$

Como $p_k \notin \mathcal{O}^*$, segue-se que $f(p_*) < f(p_k)$ o que implica $\langle s_k, v \rangle < 0$. Assim existe $\delta_k > 0$ tal que $\gamma_{d_k}(t) \in B_{t^*}(p_*)$ para todo $0 < t < \delta_k$, onde $d_k = -\frac{s_k}{\|s_k\|}$. Portanto para todo $0 < t_k < \delta_k$

$$d(\gamma_{d_k}(t_k), p_*) < d(p_k, p_*),$$

e isto demonstra o resultado. □

É consequência do teorema de Hadamard que se M tem curvatura seccional $K \leq 0$ então $\exp_{p_*} : T_{p_*}M \rightarrow M$ é difeomorfismo global e isto implica que podemos tomar $B_\varepsilon(p_*) = M$. Mas o Teorema V.2.2 não dá uma estimativa para δ_k em função de p_k e p_* , como ocorre quando $M = \mathbb{R}^n$. Esta estimativa pode no entanto ser obtida se $K \geq 0$, é o que virá a seguir.

Intuitivamente já se pode ter uma idéia da influência da curvatura K no comportamento da seqüência definida pelo algoritmo (V.2.1). Primeiro observe que se $K > 0$ as geodésicas tendem a se aproximar uma das outras, o contrário ocorre se $K < 0$. Isto sugere que podemos “andar mais” ao longo da geodésica, sem nos afastar de p_* , em curvatura não-negativa do que em negativa.

Lema V.2.3. *Seja $\{p_k\}$ a seqüência gerada pelo algoritmo (V.2.1). Se M tem curvatura $K \geq 0$, então para todo $q \in M$*

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) + t_k^2 + 2 \frac{t_k}{\|s_k\|} (f(q) - f(p_k))$$

para todo $k \in \mathbb{N}$.

Demonstração. Seja γ_{d_1} a geodésica minimizante, isto é, para $\|d_1\| = 1$, se tem $\gamma_{d_1}(0) = p_k$, $\gamma_{d_1}(t_1) = q$ com $\bar{t}_1 = d(p_k, q)$ e seja γ_{d_2} a geodésica tal que $d_2 = d_k = -\frac{s_k}{\|s_k\|}$, $\gamma_{d_2}(0) = p_k$ e $\gamma_{d_2}(t_k) = p_{k+1}$ com $t_k = \bar{t}_2$. Então segue do Corolário II.9.2 e da Definição III.3.16 que

$$\begin{aligned} d^2(p_{k+1}, q) &\leq \left\| -t_k \frac{s_k}{\|s_k\|} - t_1 d_1 \right\|^2 \\ &= t_1^2 + t_k^2 + 2 \frac{t_k}{\|s_k\|} \langle s_k, t_1 d_1 \rangle \\ &\leq d^2(p_k, q) + t_k^2 + 2 \frac{t_k}{\|s_k\|} (f(q) - f(p_k)). \end{aligned}$$

□

Observação V.2.4. A demonstração do Lema IV.2.3 foi feita por da Cruz Neto-Oliveira [13], foi incluída aqui para comodidade do leitor.

Considere o conjunto $\mathcal{O} = \{q \in M / f(q) \leq \inf_k f(p^k)\}$.

Corolário V.2.5. *Seja $\{p_k\}$ a seqüência gerada pelo algoritmo (V.2.1) e seja $K \geq 0$ a curvatura de M . Para todo $q \in \mathcal{O}$, temos*

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) + t_k^2$$

para todo $k \in \mathbb{N}$.

Demonstração. Segue-se imediatamente da definição de \mathcal{O} e do Lema V.2.3

□

Proposição V.2.6. *Seja $p_* \in \mathcal{O}^* \neq \emptyset$ e seja $\{p_k\}$ a seqüência gerada pelo algoritmo (V.2.1). Se M tem curvatura $K \geq 0$ e $p_k \notin \mathcal{O}^*$ então*

$$d(p_{k+1}, p_*) < d(p_k, p_*)$$

para todo

$$0 < t_k < \frac{2}{\|s_k\|} (f(p_k) - f(p_*)). \quad (\text{V.2.2})$$

Demonstração. No Lema V.2.3 tome $q = p_*$. Então

$$d^2(p_{k+1}, p_*) \leq d^2(p_k, p_*) + t_k^2 + 2 \frac{t_k}{\|s_k\|} (f(p_*) - f(p_k)).$$

Sendo $p_* \neq p_k$, segue-se que para todo $0 < t_k < \frac{2}{\|s_k\|} (f(p_k) - f(p_*))$ tem-se $t_k^2 + 2 \frac{t_k}{\|s_k\|} (f(p_*) - f(p_k)) < 0$ e isto conclui a demonstração. \square

Definição V.2.7. A seqüência $\{p_k\}$ é quasi Fejér convergente ao conjunto W se, para cada $q \in W$ existir uma seqüência $\{\varepsilon_k\} \subset \mathbb{R}$ tal que $\varepsilon_k \geq 0$, $\sum_{k=0}^{\infty} \varepsilon_k < +\infty$,

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) + \varepsilon_k$$

para todo k .

Proposição V.2.8. *Seja $\{p_k\}$ uma seqüência no espaço métrico completo (M, d) . Se $\{p_k\}$ é quasi Fejér convergente ao conjunto $W \subset M$, então $\{p_k\}$ é limitada. Se, além disso, um ponto de acumulação p de $\{p_k\}$ pertence a W então $\lim_{k \rightarrow \infty} p_k = p$.*

Demonstração. Mesma demonstração que em Burachik et al [7] substituindo $\|\cdot\|$ por d . \square

V.2.2 Convergência

Teorema V.2.9. *Seja $\{p_k\}$ a seqüência gerada pelo algoritmo (V.2.1) e seja $K \geq 0$ a curvatura seccional de M , onde M é uma variedade Riemanniana completa. Se a seqüência $\{t_k\}$ é escolhida satisfazendo*

$$\sum_{k=0}^{\infty} t_k = +\infty; \quad (\text{V.2.3})$$

$$\sum_{k=0}^{\infty} t_k^2 < +\infty. \quad (\text{V.2.4})$$

Então $\liminf_{k \rightarrow +\infty} f(p_k) = f^$; além disso, a seqüência $\{p_k\}$ converge para um ponto $p_* \in \mathcal{O}^*$ se $\mathcal{O}^* \neq \emptyset$.*

Demonstração. Por absurdo suponha que $\liminf_k f(p_k) > f^*$. Por um lado implica que $\mathcal{O} \neq \emptyset$. Assim de (V.2.4) e do Corolário V.2.5 $\{p_k\}$ é quasi Fejér convergente a \mathcal{O} , onde $\varepsilon_k = t_k^2$. Então $\{p_k\}$ é limitada e conseqüentemente $\{s_k\}$, também é limitada onde $s_k \in \partial(p_k)$; veja Observação III.4.10. Digamos que $\|s_k\| < C_0$ para todo $k \in \mathbb{N}$, onde $C_0 > 0$.

Por outro lado, existem $q \in \mathcal{O}$, $C_1 > 0$ e $k_0 \in \mathbb{N}$ tais que $f(q) < f(p_k) - C_1$ para todo $k > k_0$. Para este q segue-se do Lema V.2.3 que

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) + t_k \left(t_k - \frac{2C_1}{C_0} \right). \quad (\text{V.2.5})$$

Da hipótese (V.2.4), $t_k \rightarrow 0$ e assim podemos supor que k_0 é tal que $t_k < (C_1/C_0)$ para todo $k > k_0$. Substituindo em (V.2.5) tem-se

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) - \frac{C_1}{C_0} t_k \quad (\text{V.2.6})$$

para todo $k \geq k_0$. Somando (V.2.6) obtemos

$$\sum_{j=k_0}^{\ell+k_0} t_j \leq \frac{C_0}{C_1} (d^2(p_{k_0}, q) - d^2(p_{\ell+k_0}, q))$$

para todo ℓ . Isto contraria (V.2.3), pois $d(p_{\ell+k_0}, q)$ é limitada.

Segue da primeira parte que $\{f(p_k)\}$ possui uma subsequência monótona decrescente $\{f(p_{k_j})\}$ tal que

$$\lim_{j \rightarrow \infty} f(p_{k_j}) = f^*.$$

Sem perda de generalidade vamos supor que a própria seqüência $\{f(p_k)\}$ é monótona decrescente e converge para f^* . A seqüência $\{p_k\}$, sendo limitada, possui uma subsequência $\{p_{k_j}\}$ convergente. Digamos $\lim_{j \rightarrow \infty} p_{k_j} = p_*$, que pela continuidade de f implica $f(p_*) = \lim_{j \rightarrow \infty} f(p_{k_j}) = f^*$ e assim $p_* \in \mathcal{O}$. Portanto existe p_* ponto de acumulação de $\{p_k\}$ com $p_* \in \mathcal{O}$. Como $\{p_k\}$ é quasi Fejér convergente a \mathcal{O} segue-se do Teorema V.2.8 que a própria seqüência $\{p_k\}$ converge a p_* . \square

V.2.3 Observações finais

No parágrafo 1.1 pode-se notar a necessidade de fazermos hipótese sobre a curvatura seccional da variedade, pois é central o papel do Lema V.2.3, na prova apresentada, que só vale em curvatura não-negativa. Como as geodésicas, na presença de curvatura positiva, tendem a se aproximar, é intuitivo pensar que é possível melhorar a estimativa (V.2.2), em função da curvatura. Também fica em aberto a prova de convergência sem hipótese sobre a curvatura.

Como último comentário, observamos que o algoritmo (V.2.1) resolve o problema com restrição

$$\min_{p \in M} f(p)$$

onde M é uma variedade Riemanniana conexa, completa de curvatura seccional não-negativa, e as geodésicas de M estão disponíveis ou são facilmente aproximadas.

V.3 Algoritmo de ponto proximal para otimização

Seja M uma variedade de Hadamard e seja $f: M \rightarrow \mathbb{R}$ uma função convexa. O subconjunto \mathcal{O}^* de M denotará o conjunto dos minimizadores de f e $f^* = \inf_{p \in M} f(p)$ denotará o valor ínfimo de f . O nosso problema é estimar f^* e também calcular um ponto de \mathcal{O}^* , caso exista algum. Isto será feito de modo iterativo pelo *algoritmo de ponto proximal* que, formalmente, gera uma sequência $\{p_k\} \subset M$ do seguinte modo: tome $p_0 \in M$ e defina

$$p_{k+1} = \arg \min_{p \in M} \{f(p) + \lambda_k \rho_{p_k}(p)\} \quad (\text{V.3.1})$$

com $\lambda_k > 0$, onde $\rho_{p'}$ é definida em (II.8.5).

Observação V.3.1. No caso particular que $M = \mathbb{R}^n$, $\rho_{p'}(p) = \frac{1}{2} \|p - p'\|^2$ e assim, a iteração (V.3.1) se reduz a

$$p_{k+1} = \arg \min_{p \in \mathbb{R}^n} \left\{ f(p) + \frac{\lambda_k}{2} \|p - p_k\|^2 \right\}.$$

Assim o algoritmo de ponto proximal em variedades Riemannianas é uma extensão natural do algoritmo do ponto proximal introduzido por Martinet [33] e Rockafellar [40].

V.3.1 Boa definição

Uma aplicação $h: M \rightarrow \mathbb{R}$ é dita *1-coerciva* com respeito a $p' \in M$ se $\lim_{d(p', p) \rightarrow +\infty} \frac{h(p)}{d(p', p)} = +\infty$. Observe que o conjunto dos minimizadores de uma função 1-coerciva é não-vazio.

Lema V.3.2. A aplicação $f_k = f + \lambda_k \rho_{p_k}$ é 1-coerciva com respeito a $p_k \in M$ para todo k .

Demonstração. Pelo Teorema III.3.19 e do fato de que f é convexa temos que existe $s \in T_{p_k} M$ tal que

$$\begin{aligned} \frac{f_k(p)}{d(p_k, p)} &\geq \frac{f(p_k)}{d(p_k, p)} + \left\langle s, \frac{\exp_{p_k}^{-1} p}{d(p_k, p)} \right\rangle + \lambda_k \frac{\rho_{p_k}(p)}{d(p_k, p)} \\ &\geq \frac{f(p_k)}{d(p_k, p)} + \left\langle s, \frac{\exp_{p_k}^{-1} p}{d(p_k, p)} \right\rangle + \frac{\lambda_k}{2} d(p_k, p). \end{aligned}$$

Sendo $\|\exp_{p_k}^{-1} p\| = d(p_k, p)$ e $\lambda_k > 0$ para todo k , temos

$$\lim_{d(p_k, p) \rightarrow +\infty} \frac{f_k(p)}{d(p_k, p)} = +\infty$$

para todo k . □

Lema V.3.3. *A aplicação $f_k = f + \lambda_k \rho_{p_k}$ possui um único minimizador denotado por p_{k+1} , que é caracterizado como sendo o único ponto satisfazendo a relação*

$$0 \in \partial f(p_{k+1}) - \{\lambda_k \exp_{p_{k+1}}^{-1} p_k\}.$$

Demonstração. Do Corolário IV.2.25 segue-se que ρ_{p_k} é estritamente convexa e isto implica em que f_k é estritamente convexa. Isto, juntamente com o Lema V.3.2, implica a existência de um único minimizador de f_k . Da Proposição III.3.21 segue-se que p_{k+1} é minimizador de f_k se, e somente se

$$0 \in \partial f_k(p_{k+1}) = \partial f(p_{k+1}) - \{\lambda_k \exp_{p_{k+1}}^{-1} p_k\},$$

pois $\rho_{p_k}: M \rightarrow \mathbb{R}$ é convexa e diferenciável e a Proposição III.8.3 implica em que $\partial \rho_{p_{k+1}}(p_k) = \{-\exp_{p_{k+1}}^{-1} p_k\}$. □

Teorema V.3.4. *A seqüência acima está bem definida, e é caracterizada pela relação*

$$0 \in \partial f(p_{k+1}) - \{\lambda_k \exp_{p_{k+1}}^{-1} p_k\}. \quad (\text{V.3.2})$$

Isto é

$$p_{k+1} = \operatorname{argmin}_{p \in M} \{f(p) + \lambda_k \rho_{p_k}(p)\} \Leftrightarrow \lambda_k \exp_{p_{k+1}}^{-1} p_k \in \partial f(p_{k+1}). \quad (\text{V.3.3})$$

Demonstração. Imediata a partir do Lema V.3.3. □

Observação V.3.5. Os fatos fundamentais para a boa definição da seqüência proximal foram o Teorema III.3.19, que vale em qualquer variedade Riemanniana completa (veja Teorema III.4.9), e convexidade estrita de ρ_{p_k} que vale em variedades Riemannianas completas,

simplesmente conexa sem ponto focal (veja Innami [29]); portanto a definição do método de Ponto Proximal pode ser estendida a variedades Riemannianas completas, simplesmente conexas sem ponto focal¹.

V.3.2 Convergência

Uma seqüência $\{p_k\} \subset M$, onde (M, d) é um espaço métrico completo com distância d , é dita *Féjer convergente* ao conjunto $U \neq \emptyset$, $U \subset M$ se,

$$d(p_{k+1}, q) \leq d(p_k, q) \quad (\text{V.3.4})$$

para todo $k \geq 0$ e todo $q \in U$.

Lema V.3.6. *Se $\{p_k\}$ é Féjer convergente a $U \neq \emptyset$, então $\{p_k\}$ é limitada. Se além disso um ponto de acumulação p de $\{p_k\}$ pertence a U , então $p = \lim_{k \rightarrow +\infty} p_k$.*

Demonstração. (Veja [46] ou [30]). Dado $q \in U$, a desigualdade (V.3.4) implica que $d(p_k, q) \leq d(p_0, q)$ para todo k , logo $\{p_k\}$ é limitada. Agora sejam $p \in U$ e $\{p_{k_j}\}$ uma subsequência de $\{p_k\}$ tal que $\lim_{j \rightarrow +\infty} p_{k_j} = p$. Como $p \in U$ segue de (V.3.4) que a seqüência de números reais positivos $\{d(p, p_k)\}$ é monótona decrescente e possui uma subsequência, a saber $\{d(p, p_{k_j})\}$, convergindo para zero. Então toda a seqüência converge para zero, isto é, $0 = \lim_{k \rightarrow 0} d(p, p_k)$ o que significa $\lim_{k \rightarrow 0} p_k = p$. \square

Proposição V.3.7. *Se $\{p_k\}$ é a seqüência gerada por (V.3.1), então vale a desigualdade*

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) - d^2(p_k, p_{k+1}) + \frac{2}{\lambda_k}(f(q) - f(p_{k+1})) \quad (\text{V.3.5})$$

para todo $q \in M$.

¹Dizemos que uma variedade Riemanniana não tem ponto focal se, nenhuma geodésica de M , considerada como sendo uma subvariedade, tem ponto focal.

Demonstração. Seja $q \in M$. Considere o triângulo geodésico $\Delta(p_k p_{k+1} q)$. Segue-se do Teorema II.8.2 que

$$d^2(p_k, p_{k+1}) + d^2(p_{k+1}, q) - 2d(p_k, p_{k+1})d(p_{k+1}, q) \cos \theta \leq d^2(p_k, q)$$

onde $\theta = \sphericalangle \left(\exp_{p_{k+1}}^{-1} p_k, \exp_{p_{k+1}}^{-1} q \right)$. Isto implica que

$$d^2(p_k, p_{k+1}) + d^2(p_{k+1}, q) - 2 \langle \exp_{p_{k+1}}^{-1} p_k, \exp_{p_{k+1}}^{-1} q \rangle \leq d^2(p_k, q). \quad (\text{V.3.6})$$

Agora de (V.3.3) e da Definição III.3.17, com $p = q$ e $p' = p_{k+1}$, obtemos

$$\frac{1}{\lambda_k} (f(q) - f(p_{k+1})) \geq \langle \exp_{p_{k+1}}^{-1} p_k, \exp_{p_{k+1}}^{-1} q \rangle$$

que juntamente com (V.3.6) implica (V.3.5). □

Corolário V.3.8. *Se $\{p_k\}$ é a seqüência gerada por (V.3.1), então vale a desigualdade*

$$d^2(p_{k+1}, q) \leq d^2(p_k, q) + \frac{2}{\lambda_k} (f(q) - f(p_{k+1})) \quad (\text{V.3.7})$$

para todo $q \in M$.

Demonstração. Segue da Proposição V.3.7. □

Teorema V.3.9. *Seja a seqüência $\{p_k\}$ gerada por (V.3.1) e tome uma seqüência $\{\lambda_k\}$ tal que $\sum_{k=0}^{\infty} \frac{1}{\lambda_k} = +\infty$. Então $\lim_{k \rightarrow +\infty} f(p_k) = f^*$, onde $f^* = \inf_{p \in M} f(p)$. Se, além disso, o conjunto U^* dos minimizadores de f é não vazio, então $\lim_{k \rightarrow +\infty} p_k = p^*$, com $p^* \in U^*$.*

Demonstração. Substituindo q por p_k em (V.3.7) concluímos que a seqüência $\{f(p_k)\}$ é estritamente decrescente se $p_k \notin U^*$. Suponha por absurdo que $\lim_{k \rightarrow +\infty} f(p_k) > f^*$, então existe $q \in M$ é $\delta > 0$ tal que

$$f(q) < f(p_k) - \delta \quad (\text{V.3.8})$$

para todo $k > 0$. A desigualdade (V.3.8) juntamente com o Corolário V.3.8 implica

$$d^2(p_{k+1}, q) < d^2(p_k, q) - \frac{2\delta}{\lambda_k}$$

para todo k , mas isto implica que

$$\sum_{k=0}^j \frac{1}{\lambda_k} \leq \frac{1}{2\delta}(d^2(p_0, q) - d^2(p_{j+1}, q)) \leq \frac{1}{2\delta}d^2(p_0, q)$$

para todo j , o que contraria $\sum_{k=0}^{\infty} \frac{1}{\lambda_k} = +\infty$; logo $\lim_{k \rightarrow +\infty} f(p_k) = f^*$. Se $U^* \neq \emptyset$ tome $\bar{p} \in U^*$. Deste modo, $f(\bar{p}) \leq f(p_k)$ para todo k , substituindo q por \bar{p} em (V.3.7) obtemos $d(p_{k+1}, \bar{p}) \leq d(p_k, \bar{p})$ e isto implica que $\{p_k\}$ é Féjer convergente a U^* . Pelo Lema V.3.6 $\{p_k\}$ é limitada e, como M é completa, $\{p_k\}$ possui subsequência convergente. Seja $\{p_{k_j}\}$ uma subsequência convergente de $\{p_k\}$, digamos que $\lim_{j \rightarrow +\infty} p_{k_j} = p_*$. Da primeira parte $\lim_{j \rightarrow +\infty} f(p_{k_j}) = f^*$, sendo f continua $f(p_*) = f^*$, o que implica $p_* \in U^*$. Portanto novamente do Lema V.3.6 concluímos que a seqüência $\{p_k\}$ converge a p^* , isto é, $\lim p_k = p^*$. \square

V.3.3 Observações finais

Alguns métodos clássicos de otimização em variedades são estudados em Botsaris [6], Gabay [16], Rapcsák [37], Smith [47], Udriste [52] e da Cruz Neto–Oliveira [13]. A metodologia usada em suas provas de convergência as vezes impõe restrições sobre a variedade, como ocorre em da Cruz Neto–Oliveira [13]. Aqui observamos a necessidade de impor restrições sobre a variedade para a própria definição do método. Estas considerações dão margem a estudos futuros, por exemplo, verificar para quais variedades ainda podemos estender o Método de Ponto Proximal. Também fica em aberto o estudo de sua taxa de convergência.

CAPÍTULO VI

Algoritmo de ponto proximal para campos monótonos

VI.1 O algoritmo para campos C^1

Seja M uma variedade de Hadamard e seja $X \in \mathfrak{X}^1(M)$ um campo monótono em M de classe C^1 . O subconjunto S^* de M denotará o conjunto das singularidades de X . O nosso problema é calcular um ponto de S^* caso exista algum. Isto será feito de modo iterativo pelo *algoritmo de ponto proximal* que formalmente gera uma sequência $\{p_k\} \subset M$ a qual chamamos de *sequência proximal*, do seguinte modo: tome $p_0 \in M$ e defina p_{k+1} como sendo um ponto satisfazendo a equação

$$-\lambda_k \operatorname{grad} \rho_{p_k}(p_{k+1}) = X(p_{k+1}) \quad (\text{VI.1.1})$$

com $0 < \lambda_k < \tilde{\lambda}$, onde $\rho_{p'}$ é definida em (II.8.5).

Observação VI.1.1. No caso particular que $M = \mathbb{R}^n$, $\text{grad } \rho_{p_k}(p_{k+1}) = p_{k+1} - p_k$ e assim, a iteração (VI.1.1) se reduz a

$$-\lambda_k(p_{k+1} - p_k) = X(p_{k+1}).$$

O campo monótono X pode ser visto como sendo uma aplicação de \mathbb{R}^n em \mathbb{R}^n e as vezes é chamado de operador monótono, assim o algoritmo de ponto proximal para campos monótonos é uma extensão natural do algoritmo de ponto proximal para operadores monótonos introduzido por Rockafellar [42].

VI.1.1 Boa definição da sequência proximal

Seja $X \in \mathfrak{X}^1(M)$ um campo monótono de M de classe C^1 . Fixado $p' \in M$ e $\lambda > 0$ considere o campo regularizado $X + \lambda \text{grad } \rho_{p'} \in \mathfrak{X}^1(M)$. Mostraremos que este campo possui uma única singularidade ficando assim estabelecida a boa definição da sequência proximal.

Definição VI.1.2. Seja M uma variedade de Hadamard. Um campo $X \in \mathfrak{X}^1(M)$ é dito *coercivo* se

$$\lim_{d(p', p) \rightarrow +\infty} \frac{\langle -\exp_p^{-1} p', X(p) \rangle}{d(p', p)} = +\infty$$

para algum $p' \in M$.

Proposição VI.1.3. *Seja M uma variedade de Hadamard. Se $X \in \mathfrak{X}^1(M)$ é um campo monótono, então para todo $p' \in M$ e $\lambda > 0$ o campo regularizado $X + \lambda \text{grad } \rho_{p'}$ é coercivo.*

Demonstração. Sejam $p' \in M$ e $\lambda > 0$. Segue da Proposição IV.2.21 que o campo $X + \lambda \text{grad } \rho_{p'}$ é fortemente monótono com módulo $\lambda > 0$. Assim para todo ponto $p \in M$ temos

$$\lambda d^2(p, p') + \langle \exp_{p'}^{-1} p, (X + \lambda \text{grad } \rho_{p'})(p') \rangle \leq \langle \exp_{p'}^{-1} p, (X + \lambda \text{grad } \rho_{p'})(p) \rangle$$

esta desigualdade é equivalente a

$$\lambda d(p, p') + (\langle \exp_p^{-1} p, (X + \lambda \operatorname{grad} \rho_{p'}) (p') \rangle / d(p, p')) \leq (\langle -\exp_p^{-1} p', (X + \lambda \operatorname{grad} \rho_{p'}) (p) \rangle / d(p, p')).$$

Desta desigualdade e do fato que $|\langle -\exp_p^{-1} p', (X + \lambda \operatorname{grad} \rho_{p'}) (p') \rangle / d(p, p')| \leq \|(X + \lambda \operatorname{grad} \rho_{p'}) (p')\|$ segue-se que

$$\lim_{d(p, p') \rightarrow +\infty} \frac{\langle -\exp_p^{-1} p', (X + \lambda \operatorname{grad} \rho_{p'}) (p) \rangle}{d(p, p')} = +\infty$$

e isto implica que $X + \lambda \operatorname{grad} \rho_{p'}$ é coercivo. \square

Observação VI.1.4. De modo análogo à demonstração da Proposição VI.1.3 podemos mostrar que todo campo fortemente monótono é coercivo.

Definição VI.1.5. Seja M uma variedade Riemanniana. Um campo $X \in \mathfrak{X}^1(M)$ é dito *coercivo em norma* se

$$\lim_{d(p, p') \rightarrow +\infty} \|X(p)\| = +\infty$$

para algum $p' \in M$.

Proposição VI.1.6. *Seja M uma variedade de Hadamard. Se $X \in \mathfrak{X}^1(M)$ é um campo monótono, então para todo $p' \in M$ e $\lambda > 0$ o campo regularizado $X + \lambda \operatorname{grad} \rho_{p'}$ é coercivo em norma*

Demonstração. Basta observar que $(\langle -\exp_p^{-1} p', (X + \lambda \operatorname{grad} \rho_{p'}) (p) \rangle / d(p, p')) \leq \|(X + \lambda \operatorname{grad} \rho_{p'}) (p)\|$ e o resultado segue da Proposição VI.1.3. \square

Observação VI.1.7. De modo análogo à demonstração da Proposição VI.1.6 podemos mostrar que todo campo coercivo é coercivo em norma.

Teorema VI.1.8. *Seja M uma variedade de Hadamard. Se $X \in \mathfrak{X}^1(M)$ é um campo monótono, então para todo $p' \in M$ e $\lambda > 0$ o campo regularizado $X + \lambda \operatorname{grad} \rho_{p'}$ possui uma única singularidade.*

Demonstração. Considere a aplicação $N: M \rightarrow \mathbb{R}$ definida por $N(p) = \|(X + \lambda \operatorname{grad} \rho_{p'}) (p)\|^2$. Segue da Proposição VI.1.6 que N é uma aplicação coerciva e isto implica a existência de um ponto de mínimo global. Seja p_* este ponto. Como $X + \lambda \operatorname{grad} \rho_{p'} \in \mathfrak{X}^1(M)$ e p_* é um ponto de mínimo de N temos

$$\begin{aligned} 0 = dN_{p_*} \cdot v &= 2 \langle \nabla_v (X + \lambda \operatorname{grad} \rho_{p'}), (X + \lambda \operatorname{grad} \rho_{p'}) (p_*) \rangle \\ &= 2 \langle A_{(X + \lambda \operatorname{grad} \rho_{p'}) (p_*)} \cdot v, (X + \lambda \operatorname{grad} \rho_{p'}) (p_*) \rangle \end{aligned} \quad (\text{VI.1.2})$$

para todo $v \in T_{p_*} M$, onde $A_{(X + \lambda \operatorname{grad} \rho_{p'}) (p_*)}$ é definida em (II.4.1). Mas da Proposição IV.2.21 temos que o campo $X + \lambda \operatorname{grad} \rho_{p'}$ é fortemente monótono. Assim do Teorema IV.2.22, tomando $v = (X + \lambda \operatorname{grad} \rho_{p'}) (p_*)$ em (VI.1.2), concluimos que $(X + \lambda \operatorname{grad} \rho_{p'}) (p_*) = 0$. A unicidade segue do fato que todo campo fortemente monótono possui no máximo uma singularidade. \square

Observação VI.1.9. Com a mesma demonstração do Teorema VI.1.8 podemos mostrar que todo campo fortemente monótono possui uma única singularidade.

Teorema VI.1.10. *Seja M uma variedade de Hadamard e seja $X \in \mathfrak{X}^1(M)$ um campo monótono. Então dados $p_k \in M$ e $\lambda_k > 0$ existe um único ponto $p_{k+1} \in M$ satisfazendo a equação (VI.1.1), isto é, a sequência proximal $\{p_k\}$ está bem definida.*

Demonstração. Segue imediatamente do Teorema VI.1.8. \square

VI.1.2 Convergência

Lema VI.1.11. *Seja M uma variedade de Hadamard e seja $X \in \mathfrak{X}^1(M)$ um campo monótono. Se $\{p_k\} \subset M$ é a sequência proximal, definida em (VI.1.1), então para todo $q \in M$ temos*

$$d^2(p_k, p_{k+1}) + d^2(p_{k+1}, q) - \frac{2}{\lambda_k} \langle X(p_{k+1}), \exp_{p_{k+1}}^{-1} q \rangle \leq d^2(p_k, q) \quad (\text{VI.1.3})$$

Demonstração. Seja $q \in M$. Considere o triângulo geodésico $\Delta(qp_k p_{k+1})$. Do Teorema II.8.2, temos $d^2(p_k, p_{k+1}) + d^2(p_{k+1}, q) - 2d(p_k, p_{k+1})d(p_{k+1}, q) \cos \theta \leq d^2(p_k, q)$, onde $\theta = \sphericalangle(\exp_{p_{k+1}}^{-1} p_k, \exp_{p_{k+1}}^{-1} q)$. Esta última desigualdade é equivalente a $d^2(p_k, p_{k+1}) + d^2(p_{k+1}, q) - 2\langle \exp_{p_{k+1}}^{-1} p_k, \exp_{p_{k+1}}^{-1} q \rangle \leq d^2(p_k, q)$, que juntamente com (VI.1.1) e Proposição II.8.3 implica (VI.1.3). \square

Teorema VI.1.12. *Seja M uma variedade de Hadamard e seja $X \in \mathcal{X}^1(M)$ um campo monótono. Se $\{p_k\} \subset M$ é a sequência proximal, definida em (VI.1.1), e $\mathcal{O}^* \neq \emptyset$, então $\lim_{k \rightarrow +\infty} p_k = p_*$ para algum $p_* \in \mathcal{O}^*$.*

Demonstração. Dado $\tilde{p} \in \mathcal{O}^*$. Da monotonicidade de X temos que $\langle \exp_{p_{k+1}}^{-1} \tilde{p}, X(p_{k+1}) \rangle \leq \langle \exp_{\tilde{p}}^{-1} p_{k+1}, X(\tilde{p}) \rangle$ e como $X(\tilde{p}) = 0$ segue-se que $\langle \exp_{p_{k+1}}^{-1} \tilde{p}, X(p_{k+1}) \rangle \leq 0$. Tomando $q = \tilde{p}$ em (VI.1.3) temos da última desigualdade que

$$0 \leq d^2(p_{k+1}, p_k) \leq d^2(p_k, \tilde{p}) - d^2(p_{k+1}, \tilde{p}) \quad (\text{VI.1.4})$$

para todo $\tilde{p} \in \mathcal{O}^*$. A desigualdade (VI.1.4) implica que a sequência $\{p_k\}$ é Fejér convergente ao conjunto \mathcal{O}^* , veja Lema V.3.6, e sendo (M, d) completo, segue-se que existe uma subsequência $\{p_{k_j}\}$ de $\{p_k\}$ tal que $\lim_{j \rightarrow \infty} p_{k_j} = p_* \in M$. De (II.8.4) Proposição II.8.3 e (VI.1.1) temos $\lambda_{k_j} d(p_{k_j+1}, p_{k_j}) = \lambda_{k_j} \|\exp_{p_{k_j+1}}^{-1} p_{k_j}\| = \|X(p_{k_j+1})\|$, isto juntamente com $\lim_{j \rightarrow +\infty} p_{k_j} = p_*$, continuidade de X , (VI.1.4) e $0 < \lambda_k < \tilde{\lambda}$ implica que

$$\|X(p_*)\| = \lim_{j \rightarrow +\infty} \|X(p_{k_j})\| = \lambda_{k_j} d(p_{k_j+1}, p_{k_j}) = 0,$$

deste modo $p_* \in \mathcal{O}^*$. Portanto do Lema V.3.2 temos que $\lim_{k \rightarrow \infty} p_k = p_*$. \square

Observação VI.1.13. A técnica que utilizamos na prova de convergência se baseia fortemente no Lema VI.1.11 que por sua vez é uma consequência, é claro, da definição da sequência e da lei dos cossenos, que só vale, com a desigualdade no sentido que desejamos, em variedades com curvatura não positiva.

Bibliografia

- [1] S. Alexander, R.L. Bishop. Convex-supporting domains on spheres, Illinois Journal of Mathematics **18** (1974), 37–47.
- [2] W. Ballmann. Lectures on spaces of nonpositive curvature, DMV seminar Band **25**, Birkhäuser (1995).
- [3] D.A. Bayer and J.C. Lagarias. The nonlinear geometry of linear programming I, affine and projective scaling trajectories, Trans. Amer. Math. Soc. **314** (1989), 499–526.
- [4] D.A. Bayer and J.C. Lagarias. The nonlinear geometry of linear programming II, Legendre transform coordinates and central trajectories, Trans. Amer. Math. Soc. **314** (1989), 527–581.
- [5] R.L. Bishop and B. O’Neill, Manifolds of negative curvature, Trans. Amer. Math. Soc. **145** (1969), 1–49.
- [6] C.A. Botsaris, Constrained optimization along Geodesic, J. Math. Anal. Appl., **79** (1981), 295–306.

- [7] R. Burachik, L.M. Graña Drummond, A.N. Iusem and B.F. Svaiter, Full convergence of the steepest descent method with inexact line searches, *Optimization* **32** (1995), pp. 137–146.
- [8] J. Cheeger, D. Gromoll. On the structure of complete manifolds of nonnegative curvature, *Ann. of Math. (2)* **96** (1972) 413–443.
- [9] J. Cheeger and D. Ebin, *Comparison Theorems in Riemannian Geometry*, North-Holland, 1975.
- [10] R. Correa and C. Lemaréchal, Convergence of some algorithms for convex minimization, *Mathematical Programming* **62** (1993) 267–275.
- [11] M.P. do Carmo, *Riemannian Geometry*, Birkhauser (1992).
- [12] J.X. da Cruz Neto Métodos geodésicos na programação matemática, Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, (1995).
- [13] J.X. da Cruz Neto and P.R. Oliveira, *Geodesic Methods in Riemannian Manifolds*, Universidade Federal do Rio de Janeiro, Research Report 95–10 PESC/COPPE (1995).
- [14] P. Eberlein, B.O’Neill. Visibility manifolds. *Pacific. Math. J.* **46** (1973), 45–109.
- [15] O.P. Ferreira and P.R. Oliveira, *Subgradient Algorithm on Riemannian Manifolds*, Universidade Federal do Rio de Janeiro, Research Report ES-354/95 (1995).
- [16] D. Gabay, Minimizing a differentiable function over a differentiable manifolds, *JOTA* **37**, (1982), 177–219.

- [17] R.E. Greene, K. Shiohama. Riemannian manifolds having a nowhere constant convex function, *Notices Amer. Math. Soc.* **26**, (1979), 223.
- [18] R.E. Greene, K. Shiohama. Convex functions on complete noncompact manifolds: Differential structure, *Ann. Scient. Ec. Norm. Sup. 4e Serie*, **14** (1981) 357–367.
- [19] R.E. Greene and K. Shiohama, Convex functions on complete noncompact manifolds: topological structure, *Invent. Math.* **63** 129–157 (1981).
- [20] R.E. Greene, K. Shiohama. The isometry groups of manifolds admitting non-constant convex functions, *J. Math. Soc. Japan*, **39**, 1 (1987).
- [21] R.E. Greene, H. Wu. On the subharmonicity and plurisubharmonicity of geodesically convex functions, *Ind. Univ. Math. J.* **22** (1973), 641–654.
- [22] R.E. Greene, H. Wu. Integrals of subharmonic functions on manifolds of non-negative, *Invent. Math.* **27** (1974), 265–298.
- [23] R.E. Greene, H. Wu. Approximation Theorems, C^∞ convex exhaustion and manifolds of positive curvature, *Bull. Amer. Math. Soc.* **81** (1975), 101–104.
- [24] R.E. Greene, H. Wu. C^∞ convex functions and manifolds of positive curvature. *Acta Math.* **137** (1976), 209–245.
- [25] R.E. Greene, H. Wu. C^∞ approximations of convex, subharmonic and plurisubharmonic functions, *Ann. Scient. Ec. Norm. Sup. 4e Serie*, **12** (1979), 47–84.
- [26] O. Guler, On the Convergence of the Proximal Point Algorithm for Convex Minimization, *SIAM J. Cont. and Opt.*, **29**, (1991), 403–419.

- [27] U. Helmke and J.B. Moore, Optimization and Dynamical Systems, Springer-Verlag (1994).
- [28] J.-B. Hiriart-Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms I and II, Springer-Verlag (1993).
- [29] N. Innami, convexity in riemannian manifolds without Focal Points, Advanced Studies in Pure Mathematics 3, Geometry of Geodesics and Related Topics (1984), 311–332.
- [30] A.N. Iusem, Proximal Point Methods in Optimization, 20^o Colóquio Brasileiro de Matemática, IMPA, (1995).
- [31] N. Karmarkar, Riemannian geometry underlying interior point methods for linear programming, Contemporary Mathematics, **114** (1990), 51–75.
- [32] K.C. Kiwiel, An aggregate subgradient method for nonsmooth convex minimization, Mathematical Programming **27** (1983) 320–341.
- [33] B. Martinet, Régularisation, d'inéquations variationnelles par approximations successives, Revue Française d'Informatique et de Recherche Operationelle, (1970), 154–159.
- [34] G.D. Mostow, Strong rigidity of locally symmetric spaces, Ann. of Math. Studies, **78** (1973).
- [35] J.M. Ortega, W.C. Rheinboldt. Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, Inc, New York (1970).
- [36] B. O'Neill, Semi-Riemannian Geometry, Academic Press, Inc., (1983).

- [37] T. Rapcsák. Geodesic convexity in nonlinear optimization, *JOTA*, **69** (1991), 169–183.
- [38] T. Rapcsák and T.T. Thang. Nonlinear coordinate representations of smooth optimization problems, *JOTA* **86** (1995), 459–489.
- [39] T. Rapcsák and T.T. Thang. A class of polynomial variable metric algorithms for linear optimization, *Math. Prog.*, **74** (1996), 319–331.
- [40] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
- [41] R.T. Rockafellar, Augmented Lagrangian and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.*, **1** (1976), 97–116.
- [42] R.T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. cont. optim.* **14** (1976), 877-898.
- [43] T. Sakai. *Riemannian Geometry*, Translations of Mathematical Monographs, vol. **149**, Amer. Math. Soc., Providence, R.I. (1996).
- [44] K. Shiga, Hadamard manifolds, *Advanced Studies in Pure Mathematics 3, Geometry of Geodesics and Related Topic* (1984), 239–281.
- [45] K. Shiohama, *Topology of complete noncompact manifolds, Geometry of Geodesics and Related Topics, Advanced Studies in Pure Mathematics 3*, Edited by K. Shiohama, (1984), 423–450.
- [46] N.Z. Shor, *Minimization Methods for Non-Differentiable Function*, Springer-Verlag 1985.

- [47] S.T. Smith, Optimization techniques on Riemannian Manifolds, VOL. 3 of Fields Institute Communications, Am. Math. Soc., Providence, (1994), 113–146.
- [48] C. Udriste, Continuity of convex functions on Riemannian manifolds, Bull. math. Soc. Sci. Math. Roumanie, **21**, (1977), 215–218.
- [49] C. Udriste. Directional derivatives of convex functions on Riemannian manifolds. Rev. Roum. Math. Pures Appl. **24**, (1979), 1385–1388.
- [50] C. Udriste. Subgradients of convex functions on Riemannian manifolds, An. St. Univ. Al. I. Cuza, Iasi, **26**, 2 (1980), 303–305.
- [51] C. Udriste, O. Sandru, C. Nitescu. Convex programming on the Poincare plane, Tensor, N.S. **51**, 2 (1992), 103–116.
- [52] C. Udriste. Convex Functions and Optimization Methods on Riemannian Manifolds, Kluwer Academic Publishers (1994).
- [53] H. Wu. An elementary method in the study of nonnegative curvature, Acta math. **142** (1979), 57–78.
- [54] S.T. Yau. Non-existence of continuous convex functions on certain Riemannian manifolds, Math. Ann. **207** (1974), 269–270.