UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

STEFAN ALBERTO GÓMEZ GUEVARA

# Unificando a análise local do método de Newton em Variedades Riemannianas

Goiânia
2017

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

# AUTORIZAÇÃO PARA PUBLICAÇÃO DE DISSERTAÇÃO EM FORMATO ELETRÔNICO

Na qualidade de titular dos direitos de autor, **AUTORIZO** o Instituto de Matemática e Estatística da Universidade Federal de Goiás – UFG a reproduzir, inclusive em outro formato ou mídia e através de armazenamento permanente ou temporário, bem como a publicar na rede mundial de computadores (*Internet*) e na biblioteca virtual da UFG, entendendo-se os termos "reproduzir" e "publicar" conforme definições dos incisos VI e I, respectivamente, do artigo 5º da Lei nº 9610/98 de 10/02/1998, a obra abaixo especificada, sem que me seja devido pagamento a título de direitos autorais, desde que a reprodução e/ou publicação tenham a finalidade exclusiva de uso por quem a consulta, e a título de divulgação da produção acadêmica gerada pela Universidade, a partir desta data.

**Título:** Unificando a análise local do método de Newton em Variedades Riemannianas

**Autor(a):** Stefan Alberto Gómez Guevara

Goiânia, 07 de Março de 2017.

Stefan Alberto Gómez Guevara – Autor

Orizon Ferreira Pereira – Orientador

STEFAN ALBERTO GÓMEZ GUEVARA

# Unificando a análise local do método de Newton em Variedades Riemannianas

Dissertação apresentada ao Programa de Pós–Graduação do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Matemáticas.

**Área de concentração:** Otimização.

**Orientador:** Prof. Orizon Ferreira Pereira

Goiânia
2017

Stefan Alberto Gómez Guevara

# Unificando a análise local do método de Newton em Variedades Riemannianas

Dissertação defendida no Programa de Pós–Graduação do Instituto de Matemática e Estatística da Universidade Federal de Goiás como requisito parcial para obtenção do título de Mestre em Matemáticas, aprovada em 07 de Março de 2017, pela Banca Examinadora constituída pelos professores:

**Prof. Orizon Ferreira Pereira**
Instituto de Matemática e Estatística – UFG
Presidente da Banca

**Prof. Dr. Paulo Ségio Marques dos Santos**
Universidade Federal do Piauí, – UFPI

**Prof. Dr. Max Leandro Nobre Gonçalves**
Instituto de Matemática e Estatística – UFG

**Stefan Alberto Gómez Guevara**

Formado como Bacharel em Matemática pela UDFJC - Universidad Distrital Francisco José de Caldas. Durante sua graduação, foi monitor no departamento de Matemática da UDFJC. No Mestrado, na UFG - Universidade Federal de Goiás, foi bolsista da CAPES.

*A meus pais, que de muitas formas*
*me incentivaram e ajudaram para que fosse*
*possível a concretização deste trabalho.*

# Agradecimentos

A minha família, meus pais e meus irmãos, pelo amor, carinho, apoio e pela sólida formação que me foi dada.

A minha namorada, amiga e companheira, Diana Fajardo, pelo apoio incondicional, pela paciência e compreensão.

Aos meus amigos que fizeram parte desses momentos.

Ao meu orientador, professor Dr. Orizon Pereira Ferreira, pela orientação, confiança, disposição e paciência ao longo deste trabalho.

Aos professores do Instituto de Matemáticas e Estatística da UFG, que passaram pela minha trajetória durante o Mestrado em Matemática.

Por fim, agradeço ao apoio financeiro da CAPES que me permitiu executar este trabalho.

# Resumo

Gómez, Stefan. **Unificando a análise local do método de Newton em Variedades Riemannianas**. Goiânia, 2017. 79p. Dissertação de Mestrado. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

Neste trabalho consideramos o problema de encontrar uma singularidade de um campo de vetores diferenciável $X$ sobre uma variedade Riemanniana. Apresentamos uma análise local da convergência do método de Newton para encontrar uma singularidade do Campo $X$ sobre uma condição majorante. A análise mostra uma relação entre a função majorante e o campo de vetores $X$. Também apresentamos uma análise semi-local do tipo Kantorovich no contexto Riemanniana sob uma condição majorante, estabelecido em [3]. Os dois resultados permitem unificar alguns resultados nÃ£o previamente relacionados.

**Palavras–chave**

convergência local, convergência semi-local, função majorante, método de Newton, variedades Riemannianas.

# Abstract

Gómez, Stefan. **Unifying local analysis of Newton's method in Riemannian Manifolds**. Goiânia, 2017. 79p. MSc. Dissertation. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

In this work we consider the problem of finding a singularity of a field of differentiable vectors $X$ on a Riemannian manifold. We present a local analysis of the convergence of Newton's method to find a singularity of field $X$ on an increasing condition. The analysis shows a relationship between the major function and the vector field $X$. We also present a semi-local Kantorovich type analysis in the Riemannian context under a major condition, established in [3]. The two results allow to unify some previously unrelated results.

**Keywords**

Local convergence, Newton's method, majorant principle, Riemannian manifold, Semi-local convergence

# Sumário

# Lista de Figuras

# Lista de Tabelas

# Introdução

O método de Newton é uma ferramenta fundamental em análise numérica, pesquisa operacional e otimização. Tem inúmeras aplicações na ciência da administração, pesquisa industrial e financeira, mineração de dados.

O método de Newton é também uma poderosa ferramenta teórica com uma ampla gama de aplicações em matemática pura veja [4, 18, 21, 22, 36]. o método de Newton foi estendido a variedades Riemannianas com muitas finalidades diferentes (veja [1, 2, 31, 33]).

Nesta dissertação, estudaremos a análise da convergência local e semi-local do método de Newton para encontrar singularidades de um campo de vetores definido sobre uma variedade Riemanniana sob uma condição majorante apresentado por Ferreira e Silva em [12] e Alvarez et al. em [3] respectivamente. Este trabalho está divido em quatro capítulos:

No capítulo 1 apresentaremos o método de Newton em $\mathbb{R}^n$ e alguns resultados de convergência clássicos do método, após apresentaremos definições e resultados da Geometria Riemanniana que serão necessários para a compreensão dos próximos capítulos, por último apresentaremos a extensão do método de Newton para variedades Riemannianas.

No capítulo 2 estudaremos a análise local da convergência do método de Newton para encontrar uma singularidade de um campo de vetores definido sobre uma variedade Riemanniana sob uma condição majorante, apresentada por Ferreira em [12]. A análise apresentada fornece uma relação entre a função majorante e o campo de vetores. Também, permite-nos obter o maior raio para a singularidade e o raio ótimo de convergência para o método com relação à função majorante. As demostrações feitas nesta parte seguem as mesmas linhas apresentadas pelos autores em [12], acrescentando os cálculos dos passos intermédios. Além disso, agora no contexto Riemanniano são unificados vários resultados não previamente relacionados com o método (Veja [4, 5, 14, 23, 30, 35])

Concluímos este trabalho, estudando, no capítulo 3, a análise semi-local do método de Newton sob uma condição majorante apresentado por Alvarez et al. em [3]. na primeira parte apresentamos o conceito de função majorante o qual relaxa a clássica condição de Lipschitz. Também, apresentamos a demostração do teorema principal e por último mostraremos como vários resultados anteriores relacionados com o método de Newton são unificados por este (Veja [5, 9, 23, 30]). Como no capítulo 3 as demostrações feitas nesta parte seguem as mesmas linhas apresentadas pelos autores em [3], acrescentando os cálculos dos passos intermédios.

Capítulo 1

# Resultados preliminares

Neste capítulo apresentaremos alguns resultados básicos. Na primeira seção, apresentaremos o método de Newton em $\mathbb{R}^n$. Na segunda seção apresentaremos os conceitos básicos de Geometria Riemanniana para posteriormente estender o método de Newton a variedades Riemannianas.

## 1.1 Método de Newton em $\mathbb{R}^n$

O método de Newton é um algoritmo para encontrar um zero de uma função diferenciável não-linear. A ideia do método de Newton, consiste em aproximar o zero da função não linear, por zeros de aproximações lineares sucessivas em uma vizinhança apropriada. Isto é, queremos resolver a equação

$$F(x) = 0, \qquad x \in \Omega, \tag{1-1}$$

onde $\Omega \subset \mathbb{R}^n$ é um conjunto aberto e $F : \Omega \to \mathbb{R}^n$ é uma função continuamente diferenciável não-linear. Seja $x_0 \in \Omega$ o ponto inicial, e construímos a aproximação linear de $F$ neste ponto, isto é:

$$F(x) \approx F(x_0) + F'(x_0)(x - x_0).$$

Então, ao invés de resolver a equação não-linear $F(x) = 0$, resolvemos a equação linear

$$F(x_0) + F'(x_0)(x - x_0) = 0.$$

Se a $F'(x_0)$ for invertível, a equação linear acima tem uma única solução, digamos

$$x_1 = x_0 - F'(x_0)^{-1}F(x_0).$$

Sob condições adequados sobre $F$ e $x_0$, é possível mostrar que a sequência

$$x_{k+1} = x_k + v_k, \qquad k = 0, 1, \cdots, \tag{1-2}$$

está bem definida e converge para a solução de (1-1). Onde $v_k = -F'(x_k)^{-1}F(x_k)$ é chamado o passo de Newton.

### 1.1.1 Teorema de Convergência Local caso Lipschitz

A otimalidade do raio de convergência do Método de Newton, contidos no próximo teorema, foram expostos a partir dos artigos Rall [28] and Traub and Wozniakowski [32]. Uma prova dos resultados mencionados nesta subseção, podem ser encontrados no caso de espaços de Banach em [10].

**Teorema 1.1** *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto aberto e $F : \Omega \to \mathbb{R}^n$ e $x_* \in \Omega$. Suponha que $F$ é continuamente diferenciável em $\Omega$, $F'(x_*)$ não-singular, $F(x_*) = 0$ e existe $K > 0$ tal que*

$$\left\|F'(x_*)^{-1}[F'(x) - F'(y)]\right\| \leq K\|x - y\|, \qquad \forall\, x, y \in \Omega. \tag{1-3}$$

*Sejam $\kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}$ e $r := \min\{\kappa,\ 2/(3K)\}$. Então, a sequencia gerada pelo método de Newton para resolver $F(x) = 0$ com ponto inicial $x_0 \in B(x_*, r)$,*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \ldots, \tag{1-4}$$

*está bem definida, está contida em $B(x_*, r)$, converge para $x_*$ e vale*

$$\|x_* - x_{k+1}\| \leq \frac{K}{2}\,\frac{1}{1 - K\|x_0 - x_*\|}\,\|x_k - x_*\|^2, \qquad k = 0, 1, \ldots$$

*Além disso, o ponto $x_*$ é o único zero de $F$ em $B(x_*, 2/K)$ e se $2/(3K) < \kappa$ então, $r = 2/(3K)$ é o maior raio de convergência possível.*

**Teorema 1.2 (Teorema clássico)** *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto aberto e $F : \Omega \to \mathbb{R}^n$ e $x_* \in \Omega$. Suponha que $F$ é continuamente diferenciável em $\Omega$, $F'(x_*)$ não-singular, $F(x_*) = 0$ e existe $K > 0$ tal que*

$$\left\|F'(x) - F'(y)\right\| \leq L\|x - y\|, \qquad \forall\, x, y \in \Omega. \tag{1-5}$$

*Sejam*

$$\kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}, \qquad r := \min\left\{\kappa,\ \frac{2}{3L\|F'(x_*)^{-1}\|}\right\}.$$

*Então, a sequência* $\{x_k\}$ *gerada pelo método de Newton para resolver a equação* $F(x) = 0$*, com ponto inicial* $x_0 \in B(x_*, r)$*,*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \dots, \tag{1-6}$$

*está bem definida, está contida em* $B(x_*, r)$*, converge para* $x_*$ *e vale*

$$\|x_* - x_{k+1}\| \leq \frac{L\|F'(x_*)^{-1}\|}{2} \frac{1}{1 - L\|F'(x_*)^{-1}\| \, \|x_0 - x_*\|} \|x_k - x_*\|^2, \qquad k = 0, 1, \dots.$$

**Observação 1.3** *Utilizando a condição (1-3), ao invés da condição clássica de Lipschitz para a derivada*

$$\|F'(x) - F'(y)\| \leq L\|x - y\|, \qquad \forall\, x, y \in \Omega,$$

*onde* $L > 0$*, o Teorema 1.1 fica invariante sob transformações lineares inversíveis. De fato, suponhamos que* $F : \Omega \to \mathbb{R}^n$ *e* $x_* \in \Omega$ *satisfaz as hipótese do Teorema 1.1 e sejam* $A$ *uma matriz de ordem* $n \times n$ *e* $G : \Omega \to \mathbb{R}^n$ *definida por*

$$G(x) = AF(x).$$

*Note que* $G(x_*) = AF(x_*) = 0$*. Como* $G'(x) = AF'(x)$ *e* $G'(x)^{-1} = F'(x)^{-1}A^{-1}$ *temos que*

$$\|G'(x_0)^{-1}(G'(x) - G'(y))\| = \|F'(x_0)^{-1}A^{-1}(AF'(x) - AF'(y))\| \leq K\|x - y\|, \qquad x, y \in \Omega.$$

*Logo* $K$ *é invariante por transformações lineares inversíveis e consequentemente o raio* $r$ *também é invariante. Além disso, como*

$$G(x)^{-1}G(x) = F'(x)^{-1}A^{-1}AF'(x) = F'(x)^{-1}F'(x),$$

*as sequências geradas pelo método de Newton para resolver* $F(x) = 0$ *e* $G(x) = 0$ *com ponto inicial* $x_0 \in B(x_*, r)$ *são iguais, isto é, o método de Newton é invariante por transformações lineares inversíveis.*

Existem várias referências importantes sobre o método de Newton, veja por exemplo [6, 24, 25]. Para uma perspectiva histórica do uso do método de Newton em Otimização, veja Polyak [26]. Em Deuflhard [7] encontra-se um material mais recente e aplicações interessantes do método de Newton.

## 1.2 Variedades Riemannianas

Nesta seção, lembramos algumas definições e propriedades básicas de Geometria Riemanniana e listamos os resultados que serão utilizados nos capítulos posteriores.

Todos os conteúdos aqui contidos podem ser encontrados em [8, 19].

**Definição 1.4** *Uma variedade diferenciável de dimensão n, é um conjunto $\mathcal{M}$ e uma família de aplicações biunívocas $\mathbf{x}_\alpha : U_\alpha \subset \mathbb{R}^n \to \mathcal{M}$ de abertos $U_\alpha$ de $\mathbb{R}^n$ em $\mathcal{M}$ tais que:*

(1) $\bigcup\limits_{\alpha} \mathbf{x}_\alpha(U_\alpha) = \mathcal{M}$;

(2) *Para todo par $\alpha, \beta$, com $\mathbf{x}_\alpha(U_\alpha) \cap \mathbf{x}_\beta(U_\beta) = W \neq \emptyset$, os conjuntos $\mathbf{x}_\alpha^{-1}(W)$ e $\mathbf{x}_\beta^{-1}(W)$ são abertos em $\mathbb{R}^n$ e as aplicações $\mathbf{x}_\beta^{-1} \circ \mathbf{x}_\alpha$ são diferenciáveis;*

(3) *A família $\{(U_\alpha, \mathbf{x}_\alpha)\}$ é máxima relativa às condições (1) e (2).*

O par $(U_\alpha, \mathbf{x}_\alpha)$ (ou aplicação $\mathbf{x}_\alpha$) com $p \in \mathbf{x}_\alpha(U_\alpha)$ é chamado uma parametrização (ou sistema de coordenadas) de $\mathcal{M}$ em $p$; $\mathbf{x}_\alpha(U_\alpha)$ é chamada uma vizinhança coordenada em p.

**Observação 1.5** *A família $\{(U_\alpha, \mathbf{x}_\alpha)\}$ na Definição 1.4 induz de maneira natural uma topologia em $\mathcal{M}$. (Basta definir que $A \subset \mathcal{M}$ é aberto se $\mathbf{x}_\alpha^{-1}(A \cap \mathbf{x}_\alpha(U_\alpha))$ é um aberto de $\mathbb{R}^n$ para todo $\alpha$.)*
*Se $\mathcal{M}$ com a topologia induzida é conexa, dizemos que $\mathcal{M}$ é uma variedade diferenciável conexa .*
*Se para todo $\alpha$ e $\beta$ as aplicações $\mathbf{x}_\beta^{-1} \circ \mathbf{x}_\alpha$ na Definição 1.4 (2) são analíticas, então dizemos que a variedade $\mathcal{M}$ é analítica .*

Seja $\mathcal{M}$ uma variedade diferenciável conexa. Denotamos por $T_p\mathcal{M}$ ao espaço tangente de $\mathcal{M}$ em $p$ e o fibrado tangente sera denotado por

$$T\mathcal{M} := \{(p, v) : p \in \mathcal{M}, v \in T_p\mathcal{M}\}.$$

Um campo de vetores $X$ em uma variedade diferenciável $\mathcal{M}$ é uma correspondência que a cada ponto $p \in \mathcal{M}$, associa um vetor $X(p) \in T_p\mathcal{M}$. Em termos de aplicações, $X$ é uma aplicação de $M$ no fibrado tangente $T\mathcal{M}$. O campo é diferenciável se a aplicação $X : \mathcal{M} \to T\mathcal{M}$ é diferenciável.

Considerando uma parametrização $\mathbf{x} : U \subset \mathbb{R}^n \to \mathcal{M}$ é possível escrever

$$X(p) = \sum_{i=1}^{n} a_i(p) \frac{\partial}{\partial x_i},$$

onde cada $a_i : U \to \mathbb{R}$ é uma função em $U$ e $\left\{\frac{\partial}{\partial x_i}\right\}$ é a base associada a $\mathbf{x}$, $i = 1, \cdots, n$. É claro que $X$ é diferenciável se, e somente se, as funções $a_i$ são diferenciáveis para alguma (e, portanto, para qualquer) parametrização. Se denotamos por $\mathcal{D}$ ao conjunto de todas as funções reais diferenciáveis em $\mathcal{M}$ and $\mathcal{F}$ o conjunto de todas as funções reais em $\mathcal{M}$, então um campo vetorial pode ser pensado como uma aplicação $X : \mathcal{D} \to \mathcal{F}$, definida do seguinte modo

$$(Xf)(p) = \sum_i a_i(p)\frac{\partial f}{\partial x_i}(p), \quad f \in \mathcal{D},$$

onde $f$ indica, por um abuso de notação, a expressão de $f$ na parametrização $\mathbf{x}$, a função $Xf$ não depende da escolha da parametrização $\mathbf{x}$. Neste contexto, o campo vetorial $X$ é diferenciável se, e somente se $X : \mathcal{D} \to \mathcal{D}$, isto é $Xf \in \mathcal{D}$ para todo $f \in \mathcal{D}$.

**Observação 1.6** *Um campo de vetores é dito analítico em $p$ se existe um sistema de coordenadas local $(U, \mathbf{x})$ tal que as funções $a_i : U \to \mathbb{R}$ são analíticas. Então um campo de vetores $X$ é analítico em $\mathcal{M}$ se é analítico em cada ponto de $\mathcal{M}$.*

*Se $X, Y$ são campos diferenciáveis em $\mathcal{M}$ e $f : \mathcal{M} \to \mathbb{R}$ é uma função diferenciável, podemos considerar as funções $X(Yf), Y(Xf)$. Em geral, tais operações não conduzem a campos vetoriais, por envolverem derivadas de ordem superior à primeira. Entretanto, podemos afirmar que existe um único campo vetorial $Z$, tal que, para todo $f \in \mathcal{D}$, $Zf = (XY - YX)f$. O campo vetorial $Z$ é chamado o colchete de Lee e geralmente é denotado $[X,Y] = XY - YX$.*

**Definição 1.7** *Uma métrica Riemanniana em uma variedade diferenciável $\mathcal{M}$ é uma correspondência que associa a cada ponto $p$ de $\mathcal{M}$ um produto interno $\langle , \rangle_p$ (uma forma bilinear simétrica, positiva definida) no espaço tangente $T_p\mathcal{M}$, que varia diferenciavelmente no seguinte sentido: Se $\mathbf{x} : U \subset \mathbb{R}^n \to \mathcal{M}$ é um sistema de coordenadas locais em torno de $p$, com $\mathbf{x}(x_1, \cdots, x_n) = q \in \mathbf{x}(U)$ e $\frac{\partial}{\partial x_i}(q) = d\mathbf{x}(0, \cdots, 1, \cdots, 0)$, então $\left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle_q = g_{ij}(x_1, \cdots, x_n)$ é uma função diferenciável em $U$.*

Esta definição não depende da escolha do sistema de coordenadas. É usual excluir o índice $p$ em $\langle , \rangle_p$ sempre que não exista possibilidade de confusão. Uma variedade diferenciável com uma métrica Riemanniana é chamada variedade Riemanniana.

**Definição 1.8** *Sejam $\mathcal{M}$ e $\mathcal{N}$ variedades Riemannianas. Um difeomorfismo $f : \mathcal{M} \to \mathcal{N}$ (Uma bijeção diferenciável com inversa diferenciável) é chamado isometria se:*

$$\langle u, v \rangle_p = \langle df_p(u), df_p(v) \rangle_{f(p)}, \quad \textit{para todo } p \in \mathcal{M}, u, v \in T_p\mathcal{M}$$

Uma métrica Riemanniana pode ser usada para calcular comprimento de curvas. Nos referimos a uma curva diferenciável por partes a uma aplicação contínua $\zeta : [a,b] \to \mathcal{M}$ de um intervalo fechado $[a,b] \subset \mathbb{R}$ em uma variedade diferenciável $\mathcal{M}$, é tal que existe uma partição $a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b$ de $[a,b]$ tal que as restrições $\zeta|_{[t_i, t_{i+1}]}$, $i = 0, \cdots, k-1$ são diferenciáveis. Dizemos que $\zeta$ liga os ponto $\zeta(a)$ e $\zeta(b)$ e seu comprimento está dado por

$$\ell[\zeta,a,b] := \int_a^b \left\|\zeta'(t)\right\| dt.$$

Onde $\zeta'$ é o campo $dc\left(\frac{d}{dt}\right)$ e $\|\cdot\|$ é a norma induzida pela métrica Riemanniana.

**Definição 1.9** *Sejam $p,q$ dos pontos de uma variedade Riemanniana $\mathcal{M}$. Definimos a distância $d(p,q)$ por $d(p,q) =$ ínfimo dos comprimentos de todas as curvas $f_{p,q}$, onde $f_{p,q}$ é uma curva diferenciável por partes ligando $p$ a $q$.*

A distância $d$ acima está bem definida. De fato como $\mathcal{M}$ é conexa, dados dos pontos $p$ e $q$ em $\mathcal{M}$ existe uma curva diferenciável por partes ligando $p$ a $q$. As bolas abertas e fechadas de raio $r > 0$ centradas no ponto $p$ são definidas, respectivamente, como

$$B(p,r) := \{q \in \mathcal{M} : d(p,q) < r\}, \qquad B[p,r] := \{q \in \mathcal{M} : d(p,q) \leq r\}.$$

**Proposição 1.10** *Com a distância $d$, $\mathcal{M}$ é um espaço métrico. A topologia induzida por $d$ em $\mathcal{M}$ coincide com a topologia inicial de $\mathcal{M}$.*

**Demonstração**. Ver [8, Pág. 162]. □

**Corolário 1.11** *Seja $p_0 \in \mathcal{M}$ fixo, a função $d_{p_0}(p) = d(p_0, p)$ é contínua.*

### 1.2.1 Conexão Riemanniana e transporte paralelo

O conjunto dos campos de vetores em $\mathcal{M}$ de classe $\mathscr{C}^r$ será denotado por $\mathscr{X}^r(\mathcal{M})$ e indicaremos por $\mathscr{D}^r(\mathcal{M})$ o anel das funções reais de classe $\mathscr{C}^r$ definidas em $\mathcal{M}$.

**Definição 1.12** *Uma **conexão afim** $\nabla$ em uma variedades diferenciável $\mathcal{M}$ é uma aplicação*

$$\nabla : \mathscr{X}^r(\mathcal{M}) \times \mathscr{X}^r(\mathcal{M}) \to \mathscr{X}^r(\mathcal{M}),$$

*que se indica por $(X,Y) \xrightarrow{\nabla} \nabla_Y X$ e que satisfaz as seguintes propriedades:*

1. $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$,
2. $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$,
3. $\nabla_X(fY) = f\nabla_X Y + X(f)Y$,

*Onde $X,Y,Z \in \mathscr{X}(\mathcal{M})$ e $f,g \in \mathscr{D}^r(\mathcal{M})$.*

**Proposição 1.13** *Seja $\mathcal{M}$ uma variedade diferenciável com uma conexão afim $\nabla$. Então, existe uma única correspondência que associa a um campo vetorial $V$ ao longo da curva diferenciável $\zeta : I \to \mathcal{M}$ um outro campo vetorial $\frac{DV}{dt}$ ao longo de $\zeta$, denominado derivada covariante de $V$ ao longo de $\zeta$, tal que:*

1. $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$.
2. $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$, *Onde $W$ é um campo de vetores ao longo de $\zeta$ y $f$ é uma função diferenciável em $I$.*
3. *Se $V$ é induzido por um campo de vetores $Y \in \mathscr{X}(\mathcal{M})$, i.e., $V(t) = Y(\zeta(t))$, então $\dfrac{DV}{dt} = \nabla_{d\zeta/dt}Y$.*

**Demonstração**. Ver [8, Capítulo 2, Proposição 2.2]. □

**Definição 1.14** *Seja $\mathcal{M}$ uma variedade diferenciável com uma conexão afim $\nabla$. Um campo vetorial $V$ ao longo de uma curva $\zeta : I \to \mathcal{M}$ é chamado paralelo quando $\frac{DV}{dt} = 0$, para todo $t \in I$.*

**Proposição 1.15** *Seja $\mathcal{M}$ uma variedade diferenciável com uma conexão afim $\nabla$. Seja $\zeta : I \to \mathcal{M}$ uma curva diferenciável em $\mathcal{M}$ e $V_0$ um vetor tangente a $\mathcal{M}$ em $\zeta(t_0)$, $t_0 \in I$. Então existe um único campo de vetores paralelo $V$ ao longo de $\zeta$, tal que $V(t_0) = V_0$, ($V(t)$ é chamado o transporte paralelo de $V_0$ ao longo de $\zeta$).*

**Demonstração**. Ver [8, Capítulo 2, Proposição 2.6]. □

Se consideramos a aplicação $P = P_{\zeta,t_0,t} : T_{\zeta(t_0)}\mathcal{M} \to T_{\zeta(t)}\mathcal{M}$, definida por: $P_{\zeta,t_0,t}(v)$, $v \in T_{\zeta(t_0)}\mathcal{M}$, é o transporte paralelo do vetor $v$ ao longo da curva $\zeta$. Então $P$ é uma isometria. Note-se também que

$$P_{\zeta,b_1,b_2} \circ P_{\zeta,a,b_1} = P_{\zeta,a,b_2}, \qquad P_{\zeta,b,a} = P^{-1}_{\zeta,a,b}.$$

para $a, b, b_1, b_2 \in dom(\zeta)$.

**Definição 1.16** *Seja $\mathcal{M}$ uma variedade Riemanniana com uma conexão afim $\nabla$ e métrica $\langle , \rangle$. A conexão é dita compatível com a métrica $\langle , \rangle$, quando para toda curva diferenciável $\zeta$ e quaisquer pares de campos de vetores paralelos $P$ e $P'$ ao longo de $\zeta$, tivermos $\langle P, P' \rangle_{=}$constante.*

Uma conexão afim $\nabla$ em uma variedade diferenciável $\mathcal{M}$ é dita simétrica quando

$$\nabla_X Y - \nabla_Y X = [X,Y],$$

para todo $X, Y \in \mathscr{X}(\mathcal{M})$.

**Teorema 1.17 (Levi-Civita)** *Dada uma variedade Riemanniana $\mathcal{M}$, existe uma única conexão afim $\nabla$ em $\mathcal{M}$ satisfazendo as condições:*

1. *$\nabla$ é simétrica.*
2. *$\nabla$ é compatível com a métrica Riemanniana.*

**Demonstração**. Ver [8, Capítulo 2, Teorema 3.6]. □

A conexão dada pelo teorema acima é denominada conexão de Levi-Civita (ou Riemanniana) de $\mathcal{M}$.

### 1.2.2 Geodésicas

No que se segue, $\mathcal{M}$ será uma variedade Riemanniana munida de sua conexão Riemanniana.

**Definição 1.18** *Uma curva parametrizada $\xi : I \to \mathcal{M}$ é uma geodésica em $t_0 \in I$ se $\frac{D}{dt}\left(\frac{d\xi}{dt}\right) = 0$ no ponto $t_0$; se $\xi$ é geodésica em $t$, para todo $t \in I$, dizemos que $\xi$ é uma geodésica, a restrição de $\xi$ a $[a,b]$ é chamada (segmento de) geodésica ligando $\xi(a)$ a $\xi(b)$.*

Ás vezes, por abuso de linguagem, chamaremos de geodésica à imagem $\xi(I)$ de uma geodésica $\xi$.

Se $\xi : I \to \mathcal{M}$ é uma geodésica, então

$$\frac{d}{dt}\left\langle \frac{d\xi}{dt}, \frac{d\xi}{dt} \right\rangle = 2\left\langle \frac{D}{dt}\frac{d\xi}{dt}, \frac{d\xi}{dt} \right\rangle = 0,$$

isto é, o comprimento do vetor tangente $\frac{d\xi}{dt}$ é constante. O comprimento $\ell$ de $\xi$, a partir de uma origem fixa, digamos $t = t_0$, é então $\ell[\xi, t_0, t] = \left\|\frac{d\xi}{dt}\right\|(t - t_0)$. Quando o parâmetro é o próprio comprimento de arco, isto é, $\left\|\frac{d\xi}{dt}\right\| = 1$, diremos que a geodésica $\xi$ está normalizada.

**Proposição 1.19** *Dado $p \in \mathcal{M}$, existem uma vizinhança $V$ de $p$ em $\mathcal{M}$, um número $\varepsilon > 0$ e uma aplicação $\mathscr{C}^\infty$, $\xi : (-2,2) \times \mathscr{U} \to \mathcal{M}$, $\mathscr{U} = \{(q,w) \in T\mathcal{M}; q \in V, w \in T_q\mathcal{M}, |w| < \varepsilon\}$ tal que $t \to \xi(t,q,w)$, $t \in (-2,2)$, é a única geodésica de $\mathcal{M}$ que no instante $t = 0$ passa por $q$ com velocidade $w$, para cada $q \in V$ e cada $w \in T_q\mathcal{M}$, com $|\omega| < \varepsilon$.*

**Demonstração**. Ver [8, Capítulo 3, Proposição 2.7]. □

A proposição 1.19 nos permite introduzir o conceito de aplicação exponencial da maneira seguinte. Seja $p \in \mathcal{M}$ e $\mathcal{U} \subset T\mathcal{M}$ um aberto dado pela proposição 1.19. Então a aplicação $\exp : \mathcal{U} \to \mathcal{M}$ dada por

$$\exp(q,v) = \zeta(1,q,v) = \zeta\left(|v|, q, \frac{v}{|v|}\right), \qquad (q,v) \in \mathcal{U},$$

é chamada a aplicação exponencial em $\mathcal{U}$.

Na maior parte das aplicações, utilizaremos a restrição de exp a um aberto do espaço tangente $T_p\mathcal{M}$, isto é, definiremos

$$\exp_p : B_\varepsilon(0_p) \subset T_p\mathcal{M} \to \mathcal{M}$$

por $\exp_p(v) = \exp(p,v)$. Aqui, e de agora em diante, indicaremos por $B_\varepsilon(0_p)$ e $\overline{B}_\varepsilon(0_p)$ respectivamente as bolas aberta e fechada de raio $\varepsilon$ e centro na origem $0_p$ de $T_p\mathcal{M}$. Geometricamente, $\exp_p(v)$ é o ponto de $\mathcal{M}$ obtido percorrendo um comprimento igual a $|v|$, a partir de $p$, sobre a geodésica que passa por $p$ com velocidade igual a $\frac{v}{|v|}$.

**Proposição 1.20** *Dado $p \in \mathcal{M}$, existe um $\varepsilon > 0$ tal que $\exp_p : B_\varepsilon(0_p) \subset T_p\mathcal{M} \to \mathcal{M}$ é um difeomorfismo de $B_\varepsilon(0_p)$ sobre um aberto de $\mathcal{M}$.*

**Demonstração**. Ver [8, Capítulo 3, Proposição 2.9]. □

Para $p \in \mathcal{M}$, seja

$$r_p := \sup\left\{r > 0 : \exp_{p|B_r(0_p)} \text{ é um difeomorfismo}\right\}.$$

Pela Proposição 1.20 temos que $r_p$ está bem definido para todo $p \in \mathcal{M}$. O número $r_p$ é chamado o raio de injetividade de $\mathcal{M}$ em $p$. Note-se que se $0 < r < r_p$ então $\exp_p(B_r(0_p)) = B(p,r)$.

**Definição 1.21** *Uma variedade Riemanniana $\mathcal{M}$ é (geodesicamente) completa se para todo $p \in \mathcal{M}$, a aplicação exponencial, $\exp_p$, está definida para todo $v \in T_p\mathcal{M}$, i.e., se as geodésicas $\xi(t)$ que partem de $p$ estão definidas para todos os valores do parâmetro $t \in \mathbb{R}$.*

**Teorema 1.22 (Hopf-Rinow.)** . *Seja $\mathcal{M}$ uma variedade Riemanniana e seja $p \in \mathcal{M}$. As seguintes afirmações são equivalentes:*

*a)* $\exp_p$ *está definida em todo o* $T_p\mathcal{M}$.
*b) Os limitados e fechados de* $\mathcal{M}$ *são compactos.*
*c)* $\mathcal{M}$ *é completa como espaço métrico.*

*d)* *$\mathcal{M}$ é geodesicamente completa.*

*e)* *Existe uma sucessão de compactos $K_n \subset \mathcal{M}$, $K_n \subset int K_{n+1}$ e $\bigcup_n K_n = \mathcal{M}$, tais que se $q_n \neq K_n$ então $d(p,q_n) \to \infty$.*

*Além disso, cada uma das afirmações acima implica que*

*f)* *Para todo $q \in \mathcal{M}$ existe uma geodésica $\xi$ ligando $p$ a $q$ com $\ell[\xi] = d(p,q)$.*

**Demonstração**. Ver [8, Capítulo 7, Proposição 2.8]. □

**Observação 1.23** *Uma geodésica que satisfaz $f)$ é chamada geodésica minimizante ligando $p$ a $q$. Para todo $t \in \mathbb{R}$, temos $d(p,\xi(t)) \leq t|v|$, onde $\xi(t) = \exp_p(tv)$, a igualdade é válida se a geodésica $\xi$ restrita a $[0,1]$ é minimizante.*

*Seja $\zeta : [a,b] \to \mathcal{M}$ uma curva $\mathscr{C}^1$ por partes. Se existe $c \in (a,b)$ tal que a restrição $\zeta\,|_{[a,c]}$ é um segmento de geodésica minimizante, então, podemos escrever*

$$\ell[\zeta,a,b] = d(\zeta(a),\zeta(c)) + \ell[\zeta,c,b] \tag{1-7}$$

**Definição 1.24** *Seja $p \in \mathcal{M}$ e $r_p$ o raio de injectividade em $p$. Definimos a quantidade*

$$K_p := \sup\left\{ \frac{d(\exp_q u, \exp_q v)}{\|u-v\|} \,:\, q \in B(p,r_p),\ u,v \in T_q\mathcal{M},\ u \neq v,\ \|v\| \leq r_p,\ \|u-v\| \leq r_p \right\}.$$

**Observação 1.25** *A quantidade $K_p$ mede quão rapidamente as geodésicas se afastam em $\mathcal{M}$. Em particular, quando $u = 0$ ou mais geral $u$ e $v$ estão na mesma linha através $0_q$,*

$$d(\exp_q u, \exp_q v) = \|u-v\|.$$

*Assim, $K_p \geq 1$ para todo $p \in \mathcal{M}$. Agora, quando $\mathcal{M}$ tem curvatura seccional não-negativa [8, Capítulo 4], as geodésicas se afastam menos que os raios de $T_p\mathcal{M}$ [8, Capítulo 5], de modo que*

$$d(\exp_q u, \exp_q v) \leq \|u-v\|.$$

*Como consequência, $K_p = 1$ for all $p \in \mathcal{M}$.*

### 1.2.3 Teorema fundamental do Cálculo

Os resultados desta subseção são baseados em [9, Seção 2] e [20]. Seja $X$ um campo de vetores $\mathscr{C}^1$ em $\mathcal{M}$. A derivada covariante de $X$ determinada pela conexão de Levi-Civita $\nabla$ define para cada $p \in \mathcal{M}$ uma aplicação linear $\nabla X(p) : T_p\mathcal{M}$ em $T_p\mathcal{M}$, dada por

$$\nabla X(p)v := \nabla_Y X(p), \tag{1-8}$$

onde $Y$ é um campo de vetores, tal que $Y(p) = v$. Podemos chamar $\nabla X(p)$ a derivada covariante de $X$ em $p$. Seja $n$ um inteiro positivo. Vamos definir a noção de $n$-ésima derivada covariante.

**Definição 1.26** *Seja $\mathcal{M}$ uma variedade Riemanniana completa e $Y_1, \cdots, Y_n$ campos de vetores em $\mathcal{M}$. Então a n-ésima derivada covariante de $X$ em relação a $Y_1, \cdots, Y_n$ é definida indutivamente*

$$\nabla^2_{\{Y_1,Y_2\}} X := \nabla_{Y_2} \nabla_{Y_1} X, \qquad \nabla^n_{\{Y_i\}_{i=1}^n} X := \nabla_{Y_n}(\nabla_{Y_{n-1}} \cdots \nabla_{Y_1} X).$$

**Definição 1.27** *Seja $\mathcal{M}$ uma variedade Riemanniana completa, $p \in \mathcal{M}$. Então, a n-ésima derivada covariante de $X$ em $p$ é uma aplicação n-linear*

*$\nabla^n X(p) : T_p\mathcal{M} \times \ldots \times T_p\mathcal{M} \to T_p\mathcal{M}$ definida por*

$$\nabla^n X(p)(v_1, \ldots, v_n) := \nabla^n_{\{Y_i\}_{i=1}^n} X(p),$$

*onde $Y_1, \ldots, Y_n$ são campos de vetores em $\mathcal{M}$ tais que $Y_1(p) = v_1, \ldots, Y_n(p) = v_n$.*

Observamos que a Definiçao 1.27 só depende dos vetores $(v_1, \ldots, v_n)$, pois a derivada covariante é tensorial em cada campo de vetores $Y_i$. Conforme a definiçao 1.27, a $n$-ésima derivada covariante $\nabla^n X(p)$ num ponto fixo $p$ é a aplicaçao $n$-linear $(T_p\mathcal{M})^n \to T_p\mathcal{M}$.

**Definição 1.28** *Seja $\mathcal{M}$ uma variedade Riemanniana completa, $p \in \mathcal{M}$. A norma de uma aplicação n-linear $A : T_p\mathcal{M} \times \ldots \times T_p\mathcal{M} \to T_p\mathcal{M}$ é definida por*

$$\|A\| = \sup\left\{\|A(v_1, \ldots, v_n)\| \ : \ v_1, \ldots, v_n \in T_p\mathcal{M}, \|v_i\| = 1, i = 1, \ldots, n\right\}.$$

*Em particular, definimos a n-ésima derivada covariante de $X$ em $p$ por*

$$\|\nabla^n X(p)\| = \sup\left\{\|\nabla^n X(p)(v_1, \ldots, v_n)\| \ : \ v_1, \ldots, v_n \in T_p\mathcal{M}, \|v_i\| = 1, i = 1, \ldots, n\right\}.$$

Sejam $\zeta : \mathbb{R} \to \mathcal{M}$ uma curva $\mathscr{C}^1$ em $\mathcal{M}$, $X$ um campo de vetores $\mathscr{C}^1$, então

$$\begin{aligned} \nabla X(\zeta(s))\zeta'(s) &= \nabla_{\zeta'(s)} X(\zeta(s)) \\ &= \lim_{h \to 0} \frac{1}{h}(P_{\zeta,s+h,s} X(\zeta(s+h)) - X(\zeta(s))). \end{aligned} \tag{1-9}$$

A primeira igualdade é por (1-8) e a segunda é proposto como exercício em [8, Capítulo 2, Exercício 2]. Para provar isto, considere $W_1(t), \cdots, W_n(t)$ bases ortonormais transportadas paralelamente ao longo da curva $\zeta(t)$. Expressando $X(\zeta(t))$ nesta base, usando também [8, Capítulo 1, Proposição 2.2] e as propriedades da conexão obtemos (1-9). Isto mostra como a conexão pode ser reobtida da noção de paralelismo.

O Teorema Fundamental do cálculo para um campo de vetores $X$ é

**Lema 1.29** *Seja $\Omega$ um subconjunto aberto de $\mathcal{M}$, $X$ um campo de vetores $\mathscr{C}^1$ definido em $\Omega$ e $\zeta : [a,b] \to \Omega$ uma curva $\mathscr{C}^1$. Então,*

$$P_{\zeta,t,a} X(\zeta(t)) = X(\zeta(a)) + \int_a^t P_{\zeta,s,a} \nabla X(\zeta(s))\, \zeta'(s)\, \mathrm{ds}, \qquad \mathrm{t} \in [\mathrm{a}, \mathrm{b}].$$

**Demonstração**. Seja $\eta : [a,b] \to T_{\zeta(a)\mathcal{M}}$ uma curva dada por $\eta(s) = P_{\zeta,s,a} X(\zeta(s))$. Um cálculo direto dá

$$\begin{aligned}
\eta'(s) &= \lim_{h\to 0} \frac{\eta(s+h) - \eta(s)}{h} \\
&= \lim_{h\to 0} \frac{1}{h} \left[ P_{\zeta,s+h,a} X(\zeta(s+h)) - P_{\zeta,s,a} X(\zeta(s)) \right] \\
&= \lim_{h\to 0} \frac{1}{h} \left[ P_{\zeta,s,a} \circ P_{\zeta,s+h,s} X(\zeta(s+h)) - P_{\zeta,s,a} X(\zeta(s)) \right].
\end{aligned}$$

Como $P_{\zeta,s,a}$ é linear, temos

$$\begin{aligned}
\eta'(s) &= \lim_{h\to 0} \frac{1}{h} P_{\zeta,s,a} \left[ P_{\zeta,s+h,s} X(\zeta(s+h)) - X(\zeta(s)) \right] \\
&= \lim_{h\to 0} P_{\zeta,s,a} \left[ \frac{1}{h} (P_{\zeta,s+h,s} X(\zeta(s+h)) - X(\zeta(s))) \right] \\
&= P_{\zeta,s,a} \left[ \lim_{h\to 0} \frac{1}{h} (P_{\zeta,s+h,s} X(\zeta(s+h)) - X(\zeta(s))) \right] \\
&= P_{\zeta,s,a} \left[ \nabla_{\zeta'(s)} X(\zeta(s)) \right] = P_{\zeta,s,a} \nabla X(\zeta(s)) \zeta'(s).
\end{aligned}$$

Onde a última igualdade é por (1-9), daí

$$\begin{aligned}
\eta(t) - \eta(a) &= \int_a^t P_{\zeta,s,a} \nabla X(\zeta(s)) \zeta'(s) \\
P_{\zeta,t,a} X(\zeta(t)) &= X(\zeta(a)) + \int_a^t P_{\zeta,s,a} \nabla X(\zeta(s))\, \zeta'(s)\, \mathrm{ds}.
\end{aligned}$$

□

**Lema 1.30** *seja $\Omega$ um subconjunto aberto de $\mathcal{M}$, $X$ um campo de vetores $\mathscr{C}^2$ definido em $\Omega$ e $\zeta : [a,b] \to \Omega$ uma curva $\mathscr{C}^1$. Então para todo $Y \in \mathscr{X}(\mathcal{M})$ temos que*

$$P_{\zeta,t,a} \nabla X(\zeta(t)) Y(\zeta(t)) = \nabla X(\zeta(a)) Y(\zeta(a)) + \int_a^t P_{\zeta,s,a} \nabla^2 X(\zeta(s)) \big( Y(\zeta(s)), \zeta'(s) \big)\, \mathrm{ds},$$

*Onde $t \in [a,b]$.*

**Demonstração**. Seja $Z = \nabla_Y X$. Então $Z$ é um campo de vetores $\mathscr{C}^1$ em $\mathcal{M}$. Então pelo Lema 1.29.

$$P_{\zeta,t,a} Z(\zeta(t)) = Z(\zeta(a)) + \int_a^t P_{\zeta,s,a} \nabla Z(\zeta(s))\, \zeta'(s)\, \mathrm{ds}, \qquad \mathrm{t} \in [\mathrm{a}, \mathrm{b}].$$

Segue-se das Definições 1.26 e 1.27,

$$\nabla Z(\zeta(s))\zeta'(s) = \nabla\left[\nabla X(\zeta(s))Y(\zeta(s))\right]\zeta'(s) = \nabla^2 X(\zeta(s))\left[Y(\zeta(s))\zeta'(s)\right],$$

Assim, se tem o resultado. □

## 1.3 Método de Newton em variedades Riemannianas

Nesta seção estenderemos o método de Newton para variedades Riemannianas. Primeiramente, definiremos o conceito de condição de lipschitz a qual será substituída e generalizada pela condição majorante no Capítulo 2. Posteriormente seguindo a ideia do caso $\mathbb{R}^n$, apresentaremos o algoritmo que estende (1-2) para variedades Riemannianas. Todos os fatos apresentado nesta seção podem ser encontrados em [9].

**Definição 1.31** *Sejam $\Omega$ um subconjunto aberto de $\mathcal{M}$ e $X$ um campo de vetores $\mathscr{C}^1$ definido em $\Omega$. $\nabla X$ é Lipchitz com constante $L > 0$ se para toda curva geodésica $\xi : [a,b] \to \mathcal{M}$ contida em $\Omega$ satisfaz*

$$\left\| P_{\xi,b,a} \nabla X(\xi(b)) P_{\xi,a,b} - \nabla X(\xi(a)) \right\| \leq L \ell[\xi, a, b]$$

Por exemplo, se $\mathcal{M}$ é o espaço euclideano a definição acima coincide com (1-5).

Seja $p \in \mathcal{M}$ e $v \in T_p\mathcal{M}$. Sabemos que existe uma única geodésica $\xi$ tal que $\xi(0) = p$ e $\xi'(0) = v$ com $\exp_p(v) = \xi(1)$. Assim, a extensão do Método (1-2) para variedades Riemannianas é direto. Pois, consiste em encontrar uma singularidade do campo de vetores $X$ definido em uma variedade Riemanniana $\mathcal{M}$,

$$X(p) = 0, \qquad p \in \mathcal{M}.$$

De (1-8) temos que a sequência $\{p_k\} \subset \mathcal{M}$ gerada pelo método é tal que o passo de Newton é substituído por $v_k = -\nabla X(p_k)^{-1} X(p_k)$, sempre que faz sentido. Daí, o método de Newton em uma variedade Riemanniana fica.

**Algoritmo.** Tomando $p_0 \in \mathcal{M}$, para $k = 0, 1, \cdots$ definamos

$$\begin{aligned} v_k &= -\nabla X(p_k)^{-1} X(p_k) \\ p_{k+1} &= \exp_{p_k}(v_k) \end{aligned}$$

ou equivalentemente,

$$p_{k+1} = \exp_{p_k}(-\nabla X(p_k)^{-1} X(p_k)). \tag{1-10}$$

É possível estender os resultados mencionados na 1.1 em $\mathbb{R}^n$ para variedades Riemannianas. veja [3, 5, 9, 12, 16, 34, 35].
Para finalizar enunciaremos o seguinte lema, que é uma ferramenta que será usada mais adiante.

**Lema 1.32 (Lema de Banach)** *Sejam $B$ um operador linear e $I_p$ o operador identidade em $T_p\mathcal{M}$. Se $\|B - I_p\| < 1$, então $B$ é invertível e $\|B^{-1}\| \leq \frac{1}{1-\|B-I_p\|}$*

**Demonstração**. Fazendo um cálculo direto é fácil ver que $B^{-1} = \sum_{i=0}^{\infty}(B - I_p)^i$ portanto $\|B^{-1}\| \leq \sum_{i=0}^{\infty} \|B - I_p\|^i = 1/(1 - \|B - I_p\|)$. □

Capítulo 2

# Análise local do Método de Newton

Neste capítulo, nosso principal objetivo é enunciar e provar um teorema de convergência local para o método de Newton. Inicialmente, provaremos alguns resultados preliminares em relação à função majorante, o qual relaxa a Lipschitz-continuidade da derivada e depois vamos mostrar que o método de Newton está bem definido e converge. Provaremos também a unicidade da solução e estabeleceremos o raio ótimo de convergência. O teorema principal deste capítulo e os demais resultados e provas são baseados em [12].

**Teorema 2.1** *Seja $\mathcal{M}$ uma variedade Riemanniana, $\Omega \subseteq \mathcal{M}$ um subconjunto aberto e $X : \Omega \to T\mathcal{M}$ um campo de vetores continuamente diferenciável. Seja $p_* \in \Omega$, $R > 0$ e $\kappa = \sup\{t \in [0,R) : B_t(p_*) \subset \Omega\}$. Supor que $X(p_*) = 0$, $\nabla X(p_*)$ é invertível e existe uma função $f : [0,R) \to \mathbb{R}$ continuamente diferenciável tal que*

$$\left\|\nabla X(p_*)^{-1}\left[P_{\zeta,1,0}\nabla X(p) - P_{\zeta,\tau,0}\nabla X(\zeta(\tau))P_{\zeta,1,\tau}\right]\right\| \leq f'\left(d(p_*,p)\right) - f'\left(\tau d(p_*,p)\right). \tag{2-1}$$

*Para todo $\tau \in [0,1]$, $p \in B(p_*,\kappa)$, onde $\zeta : [0,1] \to M$ é uma geodésica minimizante ligando $p_*$ a $p$, e*

**(h1)** *$f(0) = 0$ e $f'(0) = -1$,*
**(h2)** *$f'$ é estritamente crescente.*

*Seja $\nu := \sup\{t \in [0,R) : f'(t) < 0\}$, $\rho := \sup\{\delta \in (0,\nu) : [f(t)/f'(t) - t]/t < 1/K_{p_*}, t \in (0,\delta)\}$ e $r := \min\left\{\kappa, \rho, r_{p_*}\right\}$, onde $K_{p_*}$ é a quantidade na Definição 1.24 e $r_{p_*}$ o raio de injectividade.*
*Então as sequências com pontos inicias $p_0 \in B(p_*,r)/\{p_*\}$ e $t_0 = d(p_0,p_*)/K_{p_*}^{1/\mu}$ com $0 \leq \mu \leq 1$, a saber*

$$p_{k+1} = \exp_{p_k}\left(-\nabla X(p_k)^{-1}X(p_k)\right), \quad e \quad t_{k+1} = \left|t_k - f(t_k)/f'(t_k)\right|, \qquad k = 0,1,\ldots \tag{2-2}$$

*respectivamente, estão bem definidas;* $\{t_k\}$ *é estritamente decrescente, está contida em* $(0,r)$ *e converge para* $0$*,* $\{p_k\}$ *está contida em* $B(p_*,r)$ *e converge para* $p_*$ *que é o único zero de* $X$ *em* $B(p_*,\sigma)$*, Onde* $\sigma := \sup\{t \in (0,\kappa) : f(t) < 0\}$ *e temos que*

$$\lim_{k\to\infty} \left[d(p_{k+1},\, p_*)/d(p_k,\, p_*)\right] = 0, \qquad \lim_{k\to\infty} [t_{k+1}/t_k] = 0. \tag{2-3}$$

*Além disso, se* $K_{p_*} = 1$*,* $f(\rho)/(\rho f'(\rho)) - 1 = 1$ *e* $\rho < \min\{\kappa, r_{p_*}\}$*, então* $r = \rho$ *é o melhor raio de convergência possível. Se, adicionalmente,*

**(h3)** *A função* $g : (0,\nu) \to \mathbb{R}$ *é estritamente crescente*

$$g(t) := [f(t)/f'(t) - t]/t^{\mu+1},$$

*então a sequência* $\{t_{k+1}/t_k^{\mu+1}\}$ *é estritamente decrescente e*

$$d(p_*, p_{k+1}) \le K_{p_*} \left[t_{k+1}/t_k^{\mu+1}\right] d(p_*,\, p_k)^{\mu+1}, \qquad k = 0,1,\ldots. \tag{2-4}$$

**Observação 2.2** *Se* $f$ *tem derivada convexa* $f'$*, então vale* ***(h3)*** *com* $\mu = 1$*. A demostração segue um argumento similar ao usado em [11, Proposição 2.6]. Neste caso, o método de Newton converge quadraticamente para* $p_*$*.*

Agora, daremos exemplos de funções que satisfazem **(h1)**, **(h2)** e **(h3)**.

**Exemplo 2.3** *As seguintes funções satisfazem* ***(h1),(h2)*** *e* ***(h3)****:*

*(i)* $f : [0,\infty) \to \mathbb{R}$ *tal que* $f(t) = t^{1+\mu} - t$*;*
*(ii)* $f : [0,\infty) \to \mathbb{R}$ *tal que* $f(t) = e^{-t} + t^2 - 1$*.*

*Fixando* $0 < \mu < 1$*, a derivada da primeira função não é convexa, nem a segunda.*

De agora em diante, assumimos que as hipóteses do teorema 2.1 são satisfeitas, com exceção de **(h3)**, o qual será valida somente quando explicitamente afirmada

## 2.1 Resultados preliminares

Nesta seção, provaremos os resultados do Teorema 2.1 que envolvem a sequência $\{t_k\}$ associada à função majorante. Além disso, serão estabelecidas as relações entre a função majorante e o campo de vetores.

### 2.1.1 A sequência escalar

Iniciaremos provando que as constantes $\kappa$, $\nu$, e $\sigma$ são positivas.

**Proposição 2.4** *As constantes* $\kappa, \nu$ *e* $\sigma$ *são positivas e* $t - f(t)/f'(t) < 0$ *para todo* $t \in (0, \nu)$.

**Demonstração**. Como $\Omega$ é aberto e $p_* \in \Omega$, existe $\varepsilon > 0$ tal que $B(p_*, \varepsilon) \subseteq \Omega$ daí $\kappa \geq \varepsilon > 0$. Como $f'$ é contínua em 0 e $f'(0) = -1$, existe $\delta > 0$ tal que $f'(t) < 0$ para todo $t \in (0, \delta)$. Assim, $\nu > 0$. Agora, como $f(0) = 0$ e novamente por ser $f'$ contínua em 0 com $f'(0) = -1$ temos que existe $\delta > 0$ tal que $f(t) < 0$ para todo $t \in (0, \delta)$, assim, $\sigma > 0$. Se assumimos **(h2)** isto implica que $f$ é estritamente convexa. Assim,

$$f(y) > f(x) + f'(x)(y - x). \tag{2-5}$$

Fazendo $y = 0$ e $x = t$ junto com **(h1)** temos, $f(t) - tf'(t) < f(0) = 0$ para todo $t \in (0, R)$, daí, $t - f(t)/f'(t) < 0$ para todo $t \in [0, \nu)$. $\square$

Seja $n_f$ a função iteração de Newton para $f$,

$$n_f : [0, \nu) \to (-\infty, 0] \tag{2-6}$$

$$t \to t - f(t)/f'(t). \tag{2-7}$$

Pela Proposição 2.4, $n_f$ é bem definida em $[0, \nu)$ e como $f'(t) \neq 0$ para todo $t \in [0, \nu)$ a função $n_f$ é uma função contínua.

**Proposição 2.5** $\lim\limits_{t \to 0} |n_f(t)|/t = 0$. *Como consequência,* $\rho > 0$ *e* $|n_f(t)| < t/K_{p_*}$ *para todo* $t \in (0, \rho)$.

**Demonstração**. Seja $t \in [0, \nu)$, pela Proposição 2.4 e o fato que $f(0) = 0$, temos

$$\frac{|n_f|}{t} = \frac{f(t)/f'(t) - t}{t} = \frac{1}{f'(t)} \frac{f(t) - f(0)}{t - 0} - 1, \tag{2-8}$$

pela continuidade de $f'$ em 0 e como $f'(0) \neq 0$, a primeira afirmação se cumpre tomando o limite em (2-8), quando $t$ tende a 0. Como $\lim\limits_{t \to 0} |n_f(t)|/t = 0$, para $\varepsilon = 1/K_{p_*}$, existe $\delta > 0$ tal que

$$0 < \frac{|n_f|}{t} = \frac{f(t)/f'(t) - t}{t} < \frac{1}{K_{p_*}}, \qquad t \in (0, \delta).$$

Assim, $\rho > 0$. Portanto $|n_f(t)| < t/K_{p_*}$ para todo $t \in (0, \rho)$. $\square$

Usando (2-6) temos que a sequência $\{t_k\}$ pode ser equivalentemente definida como

$$t_0 = d(p_*, p_0)/K_{p_*}^{1/\mu}, \qquad t_{k+1} = |n_f(t_k)|, \qquad k = 0, 1, \cdots \tag{2-9}$$

**Corolário 2.6** *A sequência $\{t_k\}$ está bem definida, é estritamente decrescente e contida em $(0,\rho)$. Além disso, $\{t_k\}$ converge superlinearmente para zero, i.e., $\lim_{k\to\infty} t_{k+1}/t_k = 0$. Se, adicionalmente, vale **(h)** então a sequência $\left\{t_{k+1}/t_k^{\mu+1}\right\}$ é estritamente decrescente.*

**Demonstração**. Pela observação 1.25 temos $K_{p_*} \geq 1$, então

$$0 < t_0 = d(p_*, p_0)/K_{p_*}^{1/\mu} \leq d(p_*, p_0) < r \leq \rho.$$

Daí, a primeira afirmação se tem por indução. De fato, para $k = 0$ pela proposição 2.5 temos

$$t_1 = |n_f(t_0)| < t_0/K_{p_*},$$

assim, $t_1 - t_0 < 0$. Agora, suponhamos que $t_k < t_{k-1}$, então novamente pela proposição 2.5

$$t_{k+1} = |n_f(t_k)| < t_k/K_{p_*}.$$

Como $\{t_k\} \subset (0,\rho)$ é estritamente decrescente então, $\{t_k\}$ é convergente. Assim, $\lim_{k\to\infty} t_k = t_*$ e $0 \leq t_* < \rho$, o qual pela continuidade tem

$$0 \leq t_* = \lim_{k\to\infty} |n_f(t_k)| = |n_f(t_*)|.$$

Afirmamos que $t_* = 0$ de fato, se $t_* \neq 0$ pela proposição 2.5 $t_* = |n_f(t_*)| < t_*/K_{p_*}$, o qual é verdade se, e somente se, $t_* = 0$. Assim, $t_k \to 0$. Fazendo $h = t_k$ então, novamente pela proposição 2.5, temos

$$\lim_{t_k\to 0} \frac{t_{k+1}}{t_k} = \lim_{t_k\to\infty} \frac{|n_f(t_k)|}{t_k} = \lim_{h\to 0} \frac{|n_f(h)|}{h} = 0,$$

pois $h = t_k \to 0$ quando $k \to \infty$. Daí, a segunda afirmação fica demostrada. Agora, supondo a validade de **(h3)** e como $\{t_k\}$ é estritamente decrescente então $t_{k+1} < t_k$ para todo $k$, assim, $g(t_{k+1}) < g(t_k)$. Isto é a sequência $\{t_{k+1}/t_k^{\mu+1}\}$ é estritamente decrescente. □

### 2.1.2 Relação entre a função majorante e o campo de vetores

No que segue, apresentaremos as principais relações entre a função majorante $f$ e o campo de vetores $X$.

**Lema 2.7** *Se $d(p_*,p) < \min\{\kappa,\nu\}$, então $\nabla X(p)$ é invertível e*

$$\left\|\nabla X(p)^{-1}P_{\zeta,0,1}\,\nabla X(p_*)\right\| \leq \frac{1}{|f'(d(p_*,p))|},$$

*onde $\zeta : [0,1] \to \mathcal{M}$ é a geodésica minimizante ligando $p_*$ a $p$. Em particular, $\nabla X(p)$ é invertível para todo $p \in B(p_*, r)$, onde $r$ é definido no Teorema 2.1.*

**Demonstração**. Seja $I_{p_*} : T_{p_*}\mathcal{M} \to T_{p_*}\mathcal{M}$ o operador identidade, $p \in B(p_*,\kappa)$ and $\zeta : [0,1] \to \mathcal{M}$ a geodésica minimizante ligando $p_*$ a $p$. Como $P_{\zeta,0,0} = I_{p_*}$ e $P_{\zeta,0,1}$ é uma isometria, temos

$$\begin{aligned}\left\|\nabla X(p_*)^{-1}P_{\zeta,1,0}\,\nabla X(p)P_{\zeta,0,1} - I_{p_*}\right\| &= \left\|\nabla X(p_*)^{-1}\left[P_{\zeta,1,0}\,\nabla X(p) - \nabla X(p_*)P_{\zeta,1,0}\right]P_{\zeta,0,1}\right\| \\ &= \left\|\nabla X(p_*)^{-1}\left[P_{\zeta,1,0}\nabla X(p) - P_{\zeta,0,0}\nabla X(p_*)P_{\zeta,1,0}\right]\right\|.\end{aligned}$$

Pela equação (2-1) com $\tau = 0$ e **(h1)** concluímos que

$$\|\nabla X(p_*)^{-1}P_{\zeta,1,0}\nabla X(p)P_{\zeta,0,1} - I_{p_*}\| \leq f'(d(p_*,p)) - f'(0) < -f'(0) = 1.$$

onde usamos o fato que $f'(d(p_*,\,p) < 0$ pois $d(p_*,\,p) < \nu$. Daí, pelo Lema de Banach 1.32, $\nabla X(p_*)^{-1}P_{\zeta,1,0}\nabla X(p)P_{\zeta,0,1}$ é invertível, e por tanto $\nabla X(p)$ também. Além disso, por ser $P_{\zeta,0,1}^{-1} = P_{\zeta,1,0}$ uma isometria temos

$$\begin{aligned}\left\|\left[\nabla X(p_*)^{-1}P_{\zeta,1,0}\,\nabla X(p)P_{\zeta,0,1}\right]^{-1}\right\| &= \left\|P_{\zeta,0,1}^{-1}\,\nabla X(p)^{-1}P_{\zeta,1,0}^{-1}\,\nabla X(p_*)\right\| \\ &= \left\|\nabla X(p)^{-1}P_{\zeta,0,1}\nabla X(p_*)\right\| \\ &\leq \frac{1}{1-(f'(d(p_*,p)) - f'(0))} \\ &= \frac{1}{|f'(d(p_*,p))|},\end{aligned}$$

Pois, $f'(0) = -1$ e $f' < 0$ em $[0,\,\nu)$ . Como $r \leq \nu$ temos que a última afirmação é provada. □

Seja $N_X$ definida por

$$\begin{aligned} N_X : B(p_*,\,r) &\to \mathcal{M} \\ p &\to \exp_p(-\nabla X(p)^{-1}X(p)). \end{aligned} \tag{2-10}$$

O Lema 2.7 assegura, em particular que $\nabla X(p)$ é invertível para todo $p \in B(p_*,r)$ e, consequentemente, a função iteração de Newton $N_X$ está bem definida. Não obstante, para $p \in B(p_*,r)$, $N_X(p)$ pode não pertencer a $B(p_*,\,r)$ e pode nem mesmo pertencer ao domínio de $X$. Assim, isto é suficiente para garantir que apenas a primeira iteração está bem definida. Para garantir que a iteração pode ser repetida indefinidamente, precisamos de algum resultados adicionais.

Seja $p \in B(p_*, r)$, sabemos que uma aproximação linear para o campo $X$ no ponto $p$ é $X(q) = P_{\xi,0,1}[X(p) + \nabla X(p)u]$ com $q \in B(p_*, r)$, onde $\xi : [0,1] \to \mathcal{M}$ é uma geodésica minimizante ligando $p$ a $q$ e $q = \exp_p(u)$. Se $q = p_*$ temos

$$\begin{aligned} 0_{p_*} &= P_{\xi,0,1}[X(p) + \nabla X(p)u] \\ \exp_p^{-1}(p_*) &= -\nabla X(p)^{-1} X(p) \\ p_* &= \exp_p(-\nabla X(p)^{-1} X(p)), \end{aligned}$$

daí, a iteração de Newton passa a ser um zero da linearização de $X$ em tal ponto. Agora estudaremos o erro $E_X(p, p_*)$ da linearização em um ponto $p$ de $B(p_*, \kappa)$, o qual é dado por

$$E_X(p, p_*) := X(p_*) - P_{\xi,0,1}[X(p) + \nabla X(p)u] \tag{2-11}$$

O seguinte Lema mostra como este erro está acotado pelo erro da linearização da função majorante $f$.

$$e_f(t,u) := f(u) - [f(t) + f'(t)(u-t)], \qquad t \in [0,R). \tag{2-12}$$

**Lema 2.8** *Se $d(p_*, p) \le \kappa$ então $\left\|\nabla X(p_*)^{-1} E_X(p, p_*)\right\| \le e_f(d(p_*, p), 0)$.*

**Demonstração**. Seja $p \in B(p_*, \kappa)$ e seja $\xi : [0,1] \to \mathcal{M}$ uma geodésica minimizante ligando $p_*$ a $p$. Seja $\zeta(u) = \xi(1-u)$. Como $P_{\zeta,0,1} = P_{\zeta,1,0}^{-1}$ pelo Lema 1.29 segue que

$$X(p_*) = P_{\zeta,0,1} X(p) + \int_0^1 P_{\zeta,u,1}[\nabla X(\zeta(u))\zeta'(u)]du,$$

como $\zeta$ é uma geodésica, $\zeta'$ é um campo paralelo e pela unicidade do transporte paralelo temos que $\zeta'(u) = P_{\zeta,0,u}\zeta'(0)$. Assim,

$$\begin{aligned} X(p_*) &= P_{\zeta,0,1} X(p) \pm P_{\zeta,0,1} \nabla X(p)\zeta'(0) + \int_0^1 P_{\zeta,u,1} \nabla X(\zeta(u)) P_{\zeta,0,u}\zeta'(0)du \\ E_X(p, p_*) &= \int_0^1 \left[P_{\zeta,u,1} \nabla X(\zeta(u)) P_{\zeta,0,u} - P_{\zeta,0,1} \nabla X(p)\right] \zeta'(0)du, \end{aligned}$$

A última igualdade acima é por (2-11). Agora, de $\zeta(u) = \xi(1-u)$, temos $P_{\zeta,u,1} = P_{\xi,1-u,0}$, $P_{\zeta,0,u} = P_{\xi,1,1-u}$ e $P_{\zeta,0,1} = P_{\xi,1,0}$. Como $\|\xi'(1)\| = d(p_*, p)$ pois $\xi$ é minimizante e usando

(2-1) com $\tau = 1-u$ segue que

$$\begin{aligned}
&\left\|\nabla X(p_*)^{-1}E_X(p,p_*)\right\| = \left\|\nabla X(p_*)^{-1}\left[\int_0^1 \left[P_{\zeta,u,1}\nabla X(\zeta(u))P_{\zeta,0,u} - P_{\zeta,0,1}\nabla X(p)\right]\zeta'(0)du\right]\right\| \\
&\le \int_0^1 \left\|\nabla X(p_*)^{-1}\left[P_{\xi,1,0}\nabla X(p) - P_{\xi,1-u,0}\nabla X(\xi(1-u))P_{\xi,1,1-u}\right]\xi'(1)\right\| du \\
&\le \int_0^1 \left[f'(d(p_*,p)) - f'((1-u)d(p_*,p))\right] d(p_*,p)du,
\end{aligned}$$

resolvendo a integral acima e lembrando que $f(0)=0$

$$\begin{aligned}
\left\|\nabla X(p_*)^{-1}E_X(p,p_*)\right\| &\le f'(d(p_*,p))d(p_*,p) + \int_0^1 f(\alpha(u))\alpha'(u)du \\
&= f'(d(p_*,p))d(p_*,p) + \left[\cancelto{0}{f(\alpha(1))} - f(\alpha(0))\right] \\
&= e_f(d(p_*,p),0),
\end{aligned}$$

onde $\alpha(u) = (1-u)d(p_*,p)$. Daí, o resultado se tem. $\square$

**Lema 2.9** *Se $d(p_*,p) < r$ então $d(p_*, N_X(p)) \le K_{p_*}|n_f(d(p_*,p))|$. Como consequência, $N_X(B((p_*,r)) \subset B(p_*,r)$*

**Demonstração**. Suponhamos que $0 < d(p_*,p) \le r$. Caso contrário $d(p_*,p) = 0$, isto implica $p = p_*$ e como $X(p_*) = 0$ o lema se tem. Pelo Lema 2.7 $X(p)$ é invertível. Seja $\xi : [0,1] \to \mathcal{M}$ uma geodésica minimizante ligando $p$ a $p_*$, aplicando o transporte paralelo de 1 a 0 ao longo de $\xi$ em (2-11) e multiplicando por $-\nabla X(p)^{-1}$ obtemos

$$-\nabla X(p)^{-1}P_{\xi,1,0}E_X(p,p_*) = \nabla X(p)^{-1}X(p) + \xi'(0).$$

Agora, os lemas 2.7 e 2.8 implicam

$$\begin{aligned}
\left\|\nabla X(p)^{-1}X(p) + \xi'(0)\right\| &\le \left\|-\nabla X(p)^{-1}P_{\xi,1,0}\nabla X(p_*)\right\| \left\|\nabla X(p_*)^{-1}E_X(p,p_*)\right\| \\
&\le \frac{1}{|f'(d(p_*,p))|}\cdot e_f(d(p_*,p),0) \\
&= \frac{f(0) - [f(d(p_*,p)) - f'(d(p_*,p))d(p_*,p)]}{-f'(d(p_*,p))} \\
&= \frac{f(d(p_*,p))}{f'(d(p_*,p))} - d(p_*,p) \\
&= -n_f(d(p_*,p)) = |n_f(d(p_*,p))|.
\end{aligned}$$

Como $d(p_*, p) < r \leq r_{p_*}$ e pela Observação temos 1.25 $K_{p_*} \geq 1$, segue pela Proposição 2.5 que

$$\left\|\nabla X(p)^{-1}X(p) + \xi'(0)\right\| \leq |n_f(d(p_*, p))| < \frac{d(p_*, p)}{K_{p_*}} < r_{p_*},$$

como $\|\xi'(0)\| = d(p_*, p) \leq r_{p_*}$, fazendo $p = p_*,\ q = p,\ v = \xi'(0),\ u = -\nabla X(p)^{-1}X(p)$ na definição 1.24

$$\frac{d(\exp_p(-\nabla X(p)^{-1}X(p)), \exp_p(\xi'(0))}{\|\nabla X(p)^{-1}X(p) + \xi'(0)\|} \leq K_{p_*},$$

segue se de (2-10)

$$d(p_*, N_X(p)) \leq K_{p_*}\left\|\nabla X(p)^{-1}X(p) + \xi'(0)\right\| \leq K_{p_*}|n_f(d(p_*, p))|.$$

Assim, a primeira afirmação fica demostrada. Para a inclusão, seja $p \in B(p_*, r)$ . Como $d(p_*, p) < r \leq \rho$, pelo feito acima e pela Proposição 2.5 temos

$$d(p_*, N_X(p)) \leq K_{p_*}|n_f(d(p_*, p))| < d(p_*, p) < r.$$

Isto mostra que para todo $p \in B(p_*, r),\ N_X(p) \in B(p_*, r)$. □

**Lema 2.10** *Se vale* **(h3)** *e* $d(p_*, p) \leq t < r$ *então* $d(p_*, N_X(p)) \leq \left[K_{p_*}|n_f(t)|/t^{\mu+1}\right] d(p_*, p)^{\mu+1}$.

**Demonstração**. Se $d(p_*, p) = 0$, isto implica que $p = p_*$ e vale o Lema. Suponhamos que $0 < d(p_*, p) \leq t$ por **(h3)** e (2-6), temos

$$\frac{|n_f(d(p_*, p))|}{d(p_*, p)^{\mu+1}} = g(d(p_*, p)) \leq g(t) = \frac{|n_f(t)|}{t^{\mu+1}}.$$

Assim, pelo Lema 2.9

$$d(p_*, N_X(p)) \leq K_{p_*}|n_f(d(p_*, p))| \leq \frac{K_{p_*}|n_f(t)|}{t^{\mu+1}} d(p_*, p)^{\mu+1}.$$

□

## 2.2 Unicidade e raio de convergência ótimo

Nesta seção mostraremos a unicidade da solução e o raio ótimo de convergência.

**Lema 2.11** *O ponto* $p_*$ *é o único zero de* $X$ *em* $B(p_*, \sigma)$.

**Demonstração**. Sejam $q \in B(p_*, \sigma)$ com $X(q) = 0$, $\xi : [0,1] \to \mathcal{M}$ uma geodésica minimizante ligando $p_*$ a $q$. Como $X(q) = X(p_*) = 0$, segue do lema 1.29

$$
\begin{aligned}
0 &= \int_0^1 P_{\xi,u,0} \nabla X(\xi(u))\xi'(u)du \\
\xi'(0) &= -\int_0^1 \nabla X(p_*)^{-1} P_{\xi,u,0} \nabla X(\xi(u))\xi'(u)du + \int_0^1 \xi'(0)du \\
&= -\int_0^1 \left[\nabla X(p_*)^{-1} P_{\xi,u,0} \nabla X(\xi(u)) - P_{\xi,u,0}\right] P_{\xi,0,u}\xi'(0)du \\
&= -\int_0^1 \nabla X(p_*)^{-1} \left[P_{\xi,u,0} \nabla X(\xi(u)) - P_{\xi,0,0} \nabla X(\xi(0)) P_{\xi,u,0}\right] P_{\xi,0,u}\xi'(0)du. \quad (2\text{-}13)
\end{aligned}
$$

Seja $\zeta(t) = \xi(tu)$. Como $P_{\zeta,1,0} = P_{\xi,u,0}$ temos que

$$P_{\xi,u,0} \nabla X(\xi(u)) - P_{\xi,0,0} \nabla X(\xi(0)) P_{\xi,u,0} = P_{\zeta,1,0} \nabla X(\zeta(1)) - P_{\zeta,0,0} \nabla X(\zeta(0)) P_{\zeta,1,0}.$$

Fazendo uso da igualdade acima y por (2-1) com $q = \zeta(1) = \xi(u)$ e $\tau = 0$, segue-se

$$\left\| \nabla X(p_*)^{-1} \left[P_{\xi,u,0} \nabla X(\xi(u)) - P_{\xi,0,0} \nabla X(\xi(0)) P_{\xi,u,0}\right] \right\| \leq f'(d(p_*, \xi(u))) + 1.$$

Pois, $f'(0) = -1$. Por ser $\xi$ minimizante temos que $\|\xi'(0)\| = d(p_*, q)$ e pela Observação 1.23 segue $d(p_*, \xi(u)) = ud(p_*, q) \geq 0$. Assim,

$$
\begin{aligned}
d(p_*, q) &\leq \int_0^1 \left\| \nabla X(p_*)^{-1} \left[P_{\xi,u,0} \nabla X(\xi(u)) - P_{\xi,0,0} \nabla X(\xi(0)) P_{\xi,u,0}\right] P_{\xi,0,u}\xi'(0) \right\| du \\
&\leq \int_0^1 \left[f'(ud(p_*, q)) + 1\right] d(p_*, q)du \\
&= \int_0^1 \frac{d}{du} f(ud(p_*, q))du + d(p_*, q) \\
&= f(d(p_*, q)) + d(p_*, q)
\end{aligned}
$$

Como $p_* \neq q$ temos $f(d(p_*, q)) \geq 0$, mas $d(p_*, q) < \sigma$, isto implica que $d(p_*, q) = 0$ assim $p_* = q$, pois 0 é o único zero de $f$ em $[0, \sigma)$. De fato, como $f$ é estritamente convexa , $f(t) < 0$ para todo $t \in (0, \sigma)$. $\square$

**Lema 2.12** *Se $K_{p_*} = 1$, $f(\rho)/(\rho f'(\rho)) - 1 = 1$ e $\rho < \min\{\kappa, r_{p_*}\}$, então $r = \rho$ é o raio de convergência ótimo.*

**Demonstração**. Seja $\mathcal{M} = \mathbb{R}$. Temos que a curvatura de $\mathcal{M}$ é igual a zero, $r_p = \infty$ e $K_t = 1$ para todo $t \in \mathcal{M}$. Definamos a função $h : (-\kappa, \kappa) \to \mathbb{R}$ por

$$h(t) = \begin{cases} -f(-t), & t \in (-\kappa, 0] \\ f(t), & t \in [0, \kappa). \end{cases} \tag{2-14}$$

Mediante um cálculo direto podemos mostrar que $X = h$, $\Omega = (-\kappa, \kappa)$ e $p_* = 0$ satisfaz todos as hipóteses do Teorema 2.1. De fato, como $h(0) = 0$ e $h'(0) = -1$, por **(h2)** temos

$$\left\| h'(0)^{-1}[h'(t) - h'(\tau t)] \right\| \leq f'(|t|) - f'(\tau|t|)$$

pois $h'(t) = f'(|t|)$. Se $K_{p_*} = 1$, $\rho < \kappa$ e $f(\rho)/(\rho f'(\rho)) - 1 = 1$, pela definição de $h$ em (2-14) segue

$$t_1 = -\rho - h(-\rho)/h'(-\rho) = -\rho + f(\rho)/f'(\rho) = [f(\rho)/(\rho f'(\rho)) - 1]\rho = \rho.$$

Novamente por (2-14) temos

$$t_2 = \rho - h(\rho)/h'(\rho) = \rho - f(\rho)/f'(\rho) = -[f(\rho)/(\rho f'(\rho)) - 1]\rho = -\rho.$$

Então aplicando o método de Newton para resolver $h(t) = 0$, com ponto inicial $t_0 = \rho < \kappa$, produze um ciclo $\{(-\rho)^{\kappa}\}$. Isto é, o método de Newton está bem definido e não é convergente. Portanto, o Lema fica provado. □

## 2.3 Sequência de Newton

Nesta seção, provaremos as afirmações no Teorema 2.1 que envolvem a sequência $\{p_k\}$. Por (1-10) e (2-10) a sequência pode ser escrita como

$$p_{k+1} = N_X(p_k), \qquad k = 0, 1, \cdots \tag{2-15}$$

**Proposição 2.13** *A sequência $\{p_k\}$ está bem definida, contida em $B(p_*, r)$ e converge ao ponto $p_*$, que é o único zero de $X$ em $B(p_*, \sigma)$, e*

$$\lim_{k \to \infty} \left[ \frac{d(p_*, p_{k+1})}{d(p_*, p_k)} \right] = 0. \tag{2-16}$$

*Se adicionalmente, vale **(h3)**, então as sequências* $\{p_k\}$ *e* $\{t_k\}$ *satisfaz*

$$d(p_*, p_{k+1}) \leq K_{p_*} \left[ \frac{t_{k+1}}{t_k^{\mu+1}} \right] d(p_*, p_k)^{\mu+1}, \qquad k = 0, 1, \cdots \tag{2-17}$$

**Demonstração**. Seja $p_0 \in B(p_*, r)$ é fácil ver por indução que a sequência $\{p_k\}$ está bem definida e contida em $B(p_*, r)$. De fato, $N_X(p_0) \in B(p_*, r)$ pelo Lema 2.9. Supondo que $p_i \in B(p_*, r)$ para todo $i = 0, \cdots, k$. Como $r \leq \nu$ pelo lema 2.7 $\nabla X(p_K)$ é invertível daí novamente pelo lema 2.9 $p_{k+1} = N_X(p_x) \in B(p_*, r)$.
Agora, para provar que $\{p_k\}$ converge a $p_*$, provaremos que $d(p_*, p_k)$ converge a 0. Como para todo $k = 0, \cdots,$ $d(p_*, p_k) < r \leq \rho$. Então pelo Lema 2.9 e a proposição 2.5 obtemos

$$d(p_*, p_{k+1}) = d(p_*, N_X(p_k)) \leq K_{p_*} |n_f(d(p_*, p_k))| < d(p_*, p_k), \qquad k = 0, \cdots \tag{2-18}$$

Assim, a sequência $\{d(p_*, p_k)\}$ é estritamente decrescente e portanto convergente. Seja $\ell_* = \lim_{k\to\infty} d(p_*, p_k)$, como $\{d(p_*, p_k)\}$ é estritamente decrescente e contida em $(0, \rho)$, temos que $0 \leq \ell_* < \rho$. Por (2-18) e pela continuidade de $n_f$ em $[0, \rho)$ segue-se $0 \leq \ell_* = K_{p_*} |n_f(\ell_*)|$. Mas, pela proposição 2.5 se $\ell_* > 0$, então $K_{p_*} |n_f(\ell*)| < \ell_*$ o que é absurdo. Assim, devemos ter $\ell_* = 0$ que implica a convergência de $p_k$ a $p_*$. A unicidade foi provada no Lema 2.11. Para provar (2-16) note que por (2-18) e pela Proposição 2.5 temos

$$\lim_{k\to\infty} \left[ \frac{d(p_*, p_{k+1})}{d(p_*, p_k)} \right] \leq K_{p_*} \lim_{k\to\infty} \left[ \frac{|n_f(d(p_*, p_k))|}{d(p_*, p_k)} \right] = K_{p_*} \lim_{t\to 0} \frac{|n_f(t)|}{t} \nearrow 0$$

Onde $t = d(p_*, p_k) \to 0$, quando $k \to \infty$.

para mostrar (2-17). Mostraremos por indução que as sequências $\{p_k\}$ e $\{t_k\}$, satisfazem

$$d(p_*, p_k) \leq \frac{t_k}{K_{p_*}^{1/\mu}}, \tag{2-19}$$

para todo $k = 0, 1, \cdots$. Como $t_0 = d(p_*, p_0)/K_{p_*}^{1/\mu}$, então (2-19) vale para $k = 0$. Agora, suponhamos por indução que vale (2-19) para $k$. Em particular como $K_{p_*} \geq 1$, temos que

$d(p_*, p_k) \leq t_k < r$, Assim

$$\begin{aligned} d(p_*, p_{k+1}) &= d(p_*, N_X(p_k)) \\ &\leq K_{p_*} \frac{|n_f(t_k)|}{t_k^{\mu+1}} d(p_*, p_k)^{\mu+1} \\ &\leq K_{p_*} \frac{|n_f(t_k)|}{t_k^{\mu+1}} \left[ \frac{t_k}{K_{p_*}^{1/\mu}} \right]^{\mu+1} \\ &= \frac{t_{k+1}}{K_{p_*}^{1/\mu}}, \end{aligned}$$

pois, $|n_f(t_k)| = t_{k+1}$. A primeira desigualdade acima se tem pelo Lema 2.10 e a segunda pela hipótese de indução. Assim, (2-19) vale para todo $k$. Em particular temos novamente por ser $K_{p_*} \geq 1$ que $d(p_*, p_{k+1}) \leq t_{k+1}$. Daí, pelo Lema 2.10 obtemos (2-17), que era o desejado. □

**Demonstração do Teorema 2.1**. A demonstração segue combinando os Lemas 2.9 e 2.12, Corolário 2.6 e Proposição 2.5. □

## 2.4 Casos especiais

Nesta seção apresentaremos 3 casos especiais do Teorema 2.1.

### 2.4.1 Convergência sob a condição Hölder-like

Apresentaremos o teorema de convergência para o método de Newton sob a condição de Holder-like, o qual estende o resultado que aparece em [14] e [35].

**Teorema 2.14** *Seja $\mathcal{M}$ uma variedade diferenciável, $\Omega \subset \mathcal{M}$ um conjunto aberto e $X :\to T\mathcal{M}$ um campo de vetores continuamente diferenciável. Sejam $p \in \Omega$ e $\kappa :=$ $\sup\{t \in [0,R) : B(p_*,t) \subset \Omega\}$. Suponhamos que $X(p_*) = 0$, $\nabla X(p_*)$ é invertível e existem constantes $L > 0$ e $0 \leq \mu \leq 1$ tal que*

$$\left\| \nabla X(p_*)^{-1} [P_{\xi,1,0} \nabla X(p) - P_{\xi,\tau,0} \nabla X(\xi(\tau)) P_{\xi,1,\tau}] \right\| \leq L(1-\tau^{\mu}) d(p_*, p)^{\mu} \qquad \text{(2-20)}$$

*para todo $\tau \in [0,1]$, $p \in B(p_*, \kappa)$, onde $\xi : [0,1] \to \mathcal{M}$ é uma geodésica minimizante ligando $p_*$ a $p$. Seja*

$$r = \min \left\{ \kappa, \left[ \frac{\mu+1}{L((1+K_{p_*})\mu+1)} \right]^{1/\mu}, r_{p_*} \right\}.$$

*Então, as sequências* $\{p_k\}$ *e* $\{t_k\}$ *a saber*

$$p_{k+1} = \exp_{p_k} -\nabla X(p_k)^{-1} X(p_k) \qquad e \qquad t_{k+1} = \frac{L\,\mu\, t_k^{\mu+1}}{(\mu+1)[1-Lt_k^{\mu}]}, \quad k = 0, \cdots,$$

*com pontos iniciais* $p_0 \in B(p_*, r)\backslash\{p_*\}$ *e* $t_0 = d(p_*, p_0)/K_{p_*}^{1/\mu}$*, respectivamente, estão bem definidas;* $\{t_k\}$ *é estritamente decrescente, contida em* $(0,r)$ *e converge a 0;* $\{p_k\}$ *está contida em* $B(p_*, r)$ *e converge a* $p_*$ *que é o único zero de* $X$ *em* $B(p_*, [\frac{\mu+1}{L}]^{1/\mu})$ *e vale a seguinte desigualdade*

$$d(p_*, p_{k+1}) \leq \frac{K_{p_*} L\mu}{(\mu+1)[1-Lt_k^{\mu}]} d(p_*, p_k)^{\mu+1}, \qquad k = 0, \cdots.$$

*mais ainda, se* $K_{p_*} = 1$ *e* $[(\mu+1)/(L(2\mu+1))]^{1/\mu} < \min\{\kappa, r_{p_*}\}$*, então*

$$r = [(\mu+1)/(L(2\mu+1))]^{1/\mu}$$

*é o melhor raio de convergência possível.*

**Demonstração**. Temos que $X, p_*$ e $f : [0,\kappa) \to \mathbb{R}$, definido por $f(t) = \frac{Lt^{\mu+1}}{(\mu+1)} - t$, satisfaz (2-1) e as condições **(h1)**, **(h2)** e **(h3)** no Teorema 2.1. Seja $p \in B(p_*, \kappa)$ então $f'(d(p_*, p)) - f'(\tau d(p_*, p)) = Ld(p_*, p)^{\mu}(1-\tau)$, assim por (2-20) temos que $X$ satisfaz (2-1). Agora, $f(t) < 0$ se, e somente se $t^{\mu} < (\mu+1)/L$ e $f'(t) < 0$ se, e somente se $t < 1/L^{1/\mu}$. Assim pelas definições segue que $\sigma = [(\mu+1)/L]^{1/\mu}$ e $\nu = \frac{1}{L^{1\mu}}$. daí, $\rho = \sup\left\{\delta \in \left(0, 1/L^{1/\mu}\right) : [f(t)/f'(t) - t]/t < 1/K_{p_*} : t \in (0,\delta)\right\}$. Assim,

$$\begin{aligned}
\frac{f(t)}{tf'(t)} - 1 < \frac{1}{K_{p_*}} &\iff \frac{f(t)}{f'(t)} \leq \frac{t(K_{p_*}+1)}{K_{p_*}} \\
&\iff t^{\mu}\left[1-(\mu+1)\left[\frac{K_{p_*}+1}{K_{p_*}}\right]\right] < -\frac{\mu+1}{L}\left[\frac{1}{K_{p_*}}\right] \\
&\iff t^{\mu} < \frac{\mu+1}{L[\mu(K_{p_*}+1)+1]}
\end{aligned}$$

portanto

$$\rho = \left[\frac{\mu+1}{L[\mu(K_{p_*}+1)+1]}\right]^{1/\mu} \leq \nu = \frac{1}{L^{1/\mu}}.$$

Além disso, para ρ definido acima temos,

$$
\begin{aligned}
\frac{f(\rho)}{\rho f'(\rho)} &= \frac{L\rho^{\mu+1}/(\mu+1)-1}{L\rho^{\mu}-1} \\
&= \frac{1/((1+K_{p_*})\mu+1)-1}{(\mu+1)/((1+K_{p_*})\mu+1)-1} \\
&= \frac{1+K_{p_*}}{K_{p_*}}
\end{aligned}
$$

Asim, se $K_{p_*} = 1$ temos $\dfrac{f(\rho)}{\rho f'(\rho)} - 1 = 1$. O resultado segue do Teorema 2.1 com $p_{k+1} = \exp_{p_k}(-\nabla X(p_k)^{-1}X(p_k))$ e

$$
t_{k+1} = \left| t_k - \frac{f(t_k)}{f'(t_k)} \right| = \left| t_k + \frac{Lt_k^{\mu+1} - t_k(\mu+1)}{(\mu+1)(1-Lt_k^{\mu})} \right| = \frac{\mu L t_k^{\mu}}{(\mu+1)[1-Lt_k^{\mu}]}
$$

□

### 2.4.2 Convergência sob a condição de Smale

A seguir, apresentaremos um teorema de convergência local para o método de Newton sob a condição de Smale. Veja [5, Teorema 1.1] (Veja também [34, Teorema 3.1]) o qual é uma generalização no contexto de variedades Riemannianas de [30, Pág. 195] (Veja também [4, Pág. 158] ).

No que resta da seção, consideraremos tanto a variedade Riemanniana completa $\mathcal{M}$ e o campo de vetores $X$ analíticos (Veja Observações 1.5 e 1.6). Como foi definido em [5, Definição 1.6], seja

$$
\gamma(p) = \sup_{k\geq 2} \left\| \frac{\nabla X(p)^{-1}\nabla^k X(p)}{k!} \right\|_p^{1/(k-1)} < \infty \tag{2-21}
$$

para cada ponto $p \in \mathcal{M}$ tal que $\nabla X(p)$ é invertível, caso contrario $\gamma(p) = \infty$. Pela analiticidade de $X$, $\gamma(p) < \infty$ sempre que $\nabla X(p)^{-1}$ seja invertível.

**Teorema 2.15** *Seja $\mathcal{M}$ uma variedade Riemanniana analítica, $\Omega \subset \mathcal{M}$ um conjunto aberto e $X : \Omega \to T\mathcal{M}$ um campo de vetores analítico. Seja $p_* \in \Omega$ e seja $\kappa := \sup\{t \in [0,R) : B(p_*,t) \subset \Omega\}$. Suponhamos que $X(p_*) = 0$, $\nabla X(p_*)$ é invertível e $\gamma = \gamma(p_*)$ como* *foi definido em* (2-21). *Seja*

$$r := \min\left\{\kappa, \frac{K_{p_*} + 4 - \sqrt{K_{p_*}^2 + 8K_{p_*} + 8}}{4\gamma}, r_{p_*}\right\}.$$

*Então, as sequências* $\{p_k\}$ *e* $\{t_k\}$ *a saber*

$$p_{k+1} = \exp_{p_k} -\nabla X(p_k)^{-1} X(p_k) \qquad e \qquad t_{k+1} = \frac{\gamma t_k^2}{2(1-\gamma t_k)^2 - 1}, \quad k = 0, \cdots,$$

*com pontos inicias* $p_0 \in B(p_*, r)\backslash\{p_*\}$ *e* $t_0 = d(p_*, p_0)/K_{p_*}$, *respectivamente, estão bem definidas;* $\{t_k\}$ *é estritamente decrescente, contida em* $(0, r)$ *e converge a 0;* $\{p_k\}$ *está contida em* $B(p_*, r)$ *e converge a* $p_*$ *que é o único zero de* $X$ *em* $B(p_*, \frac{1}{2\gamma})$. *Além disso,* $\left\{\frac{t_{k+1}}{t_k^2}\right\}$ *é estritamente decrescente, e*

$$\frac{t_{k+1}}{t_k} < \frac{\gamma}{2(1-\gamma d(p_*, p_0)/K_{p_*})^2 - 1},$$

*e*

$$d(p_*, p_{k+1}) \leq \frac{\gamma}{2(1-\gamma t_k)^2 - 1} d(p_*, p_k)^2 \leq \frac{\gamma}{2(1-\gamma d(p_*, p_*)/K_{p_*})^2 - 1} d(p_*, p_k)^2,$$

*para todo k. Se, adicionalmente,* $K_{p_*} = 1$, $\frac{5-\sqrt{17}}{4\gamma} < \min\{\kappa, r_{p_*}\}$, *então* $t = (5-\sqrt{17})/(4\gamma)$ *é o melhor raio de convergência possível.*

Para fazer a demostração do teorema acima, vamos precisar o seguinte Lema. Este lema é um caso particular do Lema 3.17 do próximo capítulo.

**Lema 2.16** *Seja* $\mathcal{M}$ *uma variedade Riemanniana analítica,* $\Omega \subset \mathcal{M}$ *um conjunto aberto e* $X : \Omega \to T\mathcal{M}$ *um campo de vetores analítico. Suponhamos que* $p_* \in \Omega$, $\nabla X(p_*)$ *é invertível e que* $B(p_*, \frac{1}{\gamma}) \subset \Omega$, *onde* $\gamma$ *é definida em* (2-21). *Então, para todo* $p \in B(p_*, \frac{1}{\gamma})$,

$$\left\|\nabla X(p_*)^{-1} P_{\zeta,1,0} \nabla^2 X(p)\right\| \leq \frac{2\gamma}{(1-\gamma d(p_*, p))^3},$$

*onde* $\zeta : [0,1] \to \mathcal{M}$ *é uma geodésica minimizante ligando* $p_*$ *a* $p$.

O próximo resultado dá uma condição alternativa para verificação de condição (2-1) Sempre que o campo vetorial em consideração seja duas vezes continuamente diferenciável.

**Lema 2.17** *Seja* $\mathcal{M}$ *uma variedade Riemanniana analítica,* $\Omega \subset \mathcal{M}$ *um conjunto aberto e* $X : \Omega \to T\mathcal{M}$ *um campo de vetores analítico. Suponha que* $p_* \in \Omega$ *e* $\nabla X(p_*)$ *é invertível.*

*Se existe uma função $f : [0.R) \to \mathbb{R}$ duas vezes diferenciável tal que*

$$\left\|\nabla X(p_*)^{-1} P_{\xi,1,0} \nabla^2 X(q)\right\| \le f''(d(p_*,q)) \qquad \textit{para todo } q \in B(p_*,\kappa), \tag{2-22}$$

*onde $\xi : [0,1] \to \mathcal{M}$ e uma geodésica minimizante ligando $p_*$ a $q$, então $X$ e $f$ satisfaz* (2-1).

**Demonstração**. Seja $\tau \in [0,1]$ e $p \in B(p_*,\kappa)$ e $\zeta : [0,1] \to \mathcal{M}$ uma geodésica minimizante ligando $p_*$ a $p$. Sejam $v \in T_p\mathcal{M}$ e $Y$ um campo de vetores paralelo ao longo de $\zeta$ com $Y(p) = v$. Então pelo Lema 1.30 temos

$$P_{\zeta,1,\tau} \nabla X(p) Y(p) = \nabla X(\zeta(\tau)) Y(\zeta(\tau)) + \int_\tau^1 P_{\zeta,s,\tau} \nabla^2 X(\zeta(s)) \left(Y(\zeta(s)), \zeta'(s)\right) ds.$$

De $Y(p) = v$ e $Y(\zeta(\tau)) = P_{\zeta,1,\tau} v$, fazendo alguns manipulações algébricas obtemos

$$\nabla X(p_*)^{-1} \left[P_{\zeta,1,0} \nabla X(p) - P_{\zeta,\tau,0} \nabla X(\zeta(\tau)) P_{\zeta,1,\tau}\right] v = \int_\tau^1 \nabla X(p_*)^{-1} P_{\zeta,s,0} \nabla^2 X(\zeta(s)) \left(Y(\zeta(s)), \zeta'(s)\right) ds.$$

Como $\|\zeta'(s)\| = \|v\|$ para todo $s \in [0,1]$ e $v$ arbitrário, pela Definição 1.28 concluímos que

$$\left\|\nabla X(p_*)^{-1} \left[P_{\zeta,1,0} \nabla X(p) - P_{\zeta,\tau,0} \nabla X(\zeta(\tau)) P_{\zeta,1,\tau}\right]\right\| \le \int_\tau^1 \left\|\nabla X(p_*)^{-1} P_{\zeta,s,0} \nabla^2 X(\zeta(s))\right\| \left\|\zeta'(s)\right\| ds.$$

Agora, como $\|\zeta'(s)\| = d(p_*,p)$ e $s\,d(p_*,p) = d(p_*,\zeta(s)) < \kappa$ para todo $s \in [0,1]$ e $f$ satisfaz (2-22) com $\xi(t) = \zeta(ts)$ e $q = \zeta(s)$, obtemos da desigualdade acima que

$$\left\|\nabla X(p_*)^{-1} \left[P_{\zeta,1,0} \nabla X(p) - P_{\zeta,\tau,0} \nabla X(\zeta(\tau)) P_{\zeta,1,\tau}\right]\right\| \le \int_\tau^1 f''(sd(p_*,p)) d(p_*,p) ds.$$

Avaliando a integral acima, a afirmação segue. □

**Corolário 2.18** *Seja $\mathcal{M}$ una variedade Riemanniana analítica, $\Omega \subset \mathcal{M}$ um conjunto aberto e $X : \Omega \to T\mathcal{M}$ um campo de vetores analítico. Seja $p_* \in \Omega$ e $\kappa := \sup\{t \in [0,R) B(p_*,t) \subset \Omega\}$ e $\gamma$ como foi definida em* (2-21). *Suponhamos que $X(p_*) = 0$ e $\nabla X(p_*)$ é invertível. Então*

$$\left\|\nabla X(p_*)^{-1} \left[P_{\zeta,1,0} \nabla X(p) - P_{\zeta,\tau,0} \nabla X(\zeta(\tau)) P_{\zeta,1,\tau}\right]\right\| \le \frac{1}{(1-\gamma d(p_*,p))^2} - \frac{1}{(1-\tau\gamma d(p_*,p))^2}$$

*para todo $\tau \in [0,1]$, $p \in B(p_*, \frac{1}{\gamma})$, onde $\zeta : [0,1] \to \mathcal{M}$ é uma geodésica minimizante ligando $p_*$ a $p$.*

**Demonstração**. A demonstração é uma imediata consequência dos Lemas 2.16 e 2.17. □

**Demonstração do Teorema 2.15**. Suponhamos que todas as hipóteses do teorema são válidas. Consideremos a função $f : [0, 1/\gamma) \to \mathbb{R}$ definida por

$$f(t) = \frac{t}{1-\gamma t} - 2t.$$

Temos que $f$ é analítica. De fato como $0 \leq \gamma t < 1$, $h(t) = 1/(1-\gamma t)$ é limite de uma serie geométrica, daí $f(t) = th(t) - 2t$. Agora, como

$$f(0) = 0, \quad f'(t) = 1/(1-\gamma t)^2 - 2, \quad f'(0) = -1, \quad f''(t) = (2\gamma)/(1-\gamma t)^3,$$

temos que $f$ satisfaz **(h1)** e **(h2)**. Além disso, para $\mu = 1$

$$\begin{aligned} g(t) &= \frac{1}{t}\left[\frac{f(t)}{tf'(t)} - 1\right] \\ &= \frac{1}{t}\left[\frac{(1-2(1-\gamma t))(1-\gamma t)}{1-2(1-\gamma t)^2} - 1\right] \\ &= \frac{\gamma}{2(1-\gamma t)^2 - 1}. \end{aligned}$$

Segue que vale **(h3)** com $\mu = 1$, pois $g'(t) > 0$. Ademais, para $p \in B(p_*, 1/\gamma)$ vale o Corolário 2.18, concluímos que $X$ e $f$ satisfaz (2-1) com $R = 1/\gamma$. Neste caso as constantes $\nu$, $\rho$, e $\sigma$ definidas no Teorema 2.1, satisfaz

$$\rho = \frac{K_{p_*} + 4 - \sqrt{K_{p_*}^2 + 8K_{p_*} + 8}}{4\gamma} < \nu = \frac{\sqrt{2}-1}{\sqrt{2}\gamma} < \frac{1}{\gamma}, \quad , \sigma = 1/(2\gamma), \quad r := \min\{\kappa, \rho, r_{p_*}\}.$$

De fato, $f(t) < 0$, se, e somente se $2\gamma t < 1$ e $f'(t) < 0$, se, e somente se $1/\sqrt{2} < 1 - \gamma t$. Assim, o resultado segue pelas definições de $\sigma$ e $\nu$. Além disso, temos

$$\frac{f(t)}{tf'(t)} - 1 < \frac{1}{K_{p_*}} \iff \frac{\gamma t}{2(1-\gamma t)^2 - 1} \leq \frac{1}{K_{p_*}} \iff 0 \leq q(t),$$

onde $q(t) = 2\gamma^2 t^2 - t\gamma(K_{p_*} + 4) + 1$, mas

$$\begin{aligned} q(t) = 0 &\iff t = \frac{1}{\gamma} + \frac{K_{p_*} \pm \sqrt{K_{p_*}^2 + 8K_{p_*} + 8}}{4\gamma} \\ &\iff t = \frac{K_{p_*} + 4 - \sqrt{K_{p_*}^2 + 8K_{p_*} + 8}}{4\gamma} \end{aligned}$$

Pois $t \in [0, 1/\gamma)$. Como $q(0) = 1$, e $q(1/\gamma) < 0$, o resultado segue da continuidade de $q$ e pela definição de $\rho$. Se $K_{p_*} = 1$, então

$$\rho = \frac{5-\sqrt{17}}{4\gamma} \qquad \text{e} \qquad \frac{f(\rho)}{\rho f'(\rho)} - 1 = \frac{\gamma\rho}{2(1-\gamma\rho)^2-1} = 1.$$

O resultado segue do Teorema 2.1 com $p_{k+1} = \exp_{p_k}(-\nabla X(p_k)^{-1} X(p_k))$ e

$$t_{k+1} = \left| t_k - \frac{f(t_k)}{f'(t_k)} \right| = \frac{\gamma t_k^2}{2(1-\gamma t_k)^2 - 1}$$

Assim, para todo $k$

$$\frac{t_{k+1}}{t_k^2} = \frac{\gamma}{2(1-\gamma t_k)^2-1} < \frac{\gamma}{2(1-\gamma t_0)^2-1} = \frac{\gamma}{2(1-\gamma d(p_0, p_*)/K_{p_*})^2-1}.$$

Assim, o resultado segue pela aplicação do Teorema 2.1. □

### 2.4.3 Convergência sob a condição de Nesterov-Nemirovskii

Nesta seção, mostraremos um teorema correspondente ao Teorema 2.1, bajo a condição de Nesterov-Nemirovskii (veja [23]).

Suponhamos que $\mathcal{M} = \mathbb{R}^n$ com a identificação usual $T_x\mathbb{R}^n \cong \mathbb{R}^n$. Denotaremos por $\langle \cdot, \cdot \rangle$ o produto Euclideano em $\mathbb{R}^n$. Lembrando que uma função convexa $f \in C^3(\Omega, \mathbb{R})$ definida em um conjunto não-vazio, aberto e convexo $\Omega \subset \mathbb{R}^n$ é dito $a$-auto-concordante para $a > 0$ se para todo $x \in \Omega$, para todo $h \in \mathbb{R}^n$,

$$|f'''(x)[h,h,h]| \leq 2a^{-1/2}(f''(x)[h,h])^{3/2}. \tag{2-23}$$

Agora, defina

$$\langle h_1, h_2 \rangle_{H_x} = a^{-1} f''(x)[h_1, h_2] = a^{-1} \langle f''(x) h_1, h_2 \rangle$$

e

$$\|h\|_{H_x} = \sqrt{\langle h, h \rangle_{H_x}}.$$

No que resta da seção, $\mathbb{R}^n$ é dotado com a estrutura métrica induzida pelo produto interno $\langle \cdot, \cdot \rangle_{H_x}$. O elipsoide de Dikin com centro em $x \in \Omega$ e raio $r > 0$ é definido por

$$D_r(x) = \{ y \in \mathbb{R}^n | \, \|y - x\|_{H_x} \leq r \}.$$

É correspondentemente a bola fechada centrada em $x$ e raio $r$.

Assim o gradiente e a Hessiana da função $h$ em $\mathcal{M}$ são dados como segue

$$grad\,h(p) := h''(p_*)^{-1}h'(p) \qquad \text{e} \qquad Hess\,g(p) := g''(p_*)^{-1}g''(p), \tag{2-24}$$

respectivamente. Dado que a curvatura seccional de $\mathcal{M}$ é zero, temos que as geodésicas de $\mathcal{M}$ são linhas retas, o transporte paralelo é a identidade e a aplicação exponencial é dada por

$$\exp_p v = p + v \qquad \text{e} \qquad p \in \mathcal{M}, \quad v \in T_p\mathcal{M}, \tag{2-25}$$

$r_p = \infty$ e $K_{p_*} = 1$.

**Teorema 2.19** *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto aberto convexo e $h : \Omega \to \mathbb{R}$ uma função estritamente convexa, três vezes continuamente diferenciável em $\Omega$. Seja $p_* \in \Omega$ com $h''(p_*)$ não singular. Suponha que $h$ é a-auto-concordante, isto é vale* (2-23)*. Seja*

$$r := \min\{\kappa, (5-\sqrt{17})/4\},$$

*Então, a sequências com ponto inicial $p_0 \in D_r(p_*) \setminus \{p_*\}$ e $t_0 = d(p_0, p_*)$, a saber,*

$$p_{k+1} = p_k - h''(p_k)^{-1}h'(p_k) \qquad e \qquad t_{k+1} = t_k^2/[2(1-t_k)^2-1], \qquad k = 0, 1, \cdots,$$

*respectivamente, estão bem definidas; $\{t_k\}$ é estritamente decrescente, contida em $(0, r)$ e converge para $0$; e $\{p_k\}$ está contida em $D_r(p_*)$ e convergente para o ponto $p_*$ que é o único zero de $h'$ em $D_{1/2}(p_*)$. Além disso, $\{t_{k+1}/t_k^2\}$ é estritamente decrescente, $t_{k+1}/t_k^2 < 1/[2(1-d(p_*, p_0))^2-1]$ para $k = 0, 1, \cdots$, e*

$$d(p_*, p_{k+1}) \leq \frac{1}{2(1-t_k)^2-1} d(p_*, p_k)^2 \leq \frac{1}{2(1-d(p_*, p_0))^2-1} d(p_*, p_k)^2.$$

Para poder fazer a demonstração do teorema acima, vamos precisar o seguinte lema. Este lema é um caso particular do Lema 3.13 do próximo capítulo.

**Lema 2.20** *Se $r < 1$ temos que para todo $x \in D_r(p_*)$*

$$\left\| g''(p_*)^{-1} g'''(p) \right\|_{p_*} \leq \frac{2}{(1-d(p_*, p))^3},$$

**Demonstração do Teorema 2.19**. Suponhamos que valem todas as hipóteses do Teorema. Seja $f : [0, 1) \to \mathbb{R}$ uma função definida como segue

$$f(t) = \frac{t}{1-t} - 2t$$

mediante um cálculo direto temos que

$$f(0)=0, \quad f'(t)=1/(1-t)^2-2, \quad f'(0)=-1, \quad f''(t)=2/(1-t)^3,$$

temos que $f$ satisfaz **(h1)** e **(h2)**. Além disso, para $\mu=1$

$$\begin{aligned} g(t) &= \frac{1}{t}\left[\frac{f(t)}{tf'(t)}-1\right] \\ &= \frac{1}{t}\left[\frac{(1-2(1-t))(1-t)}{1-2(1-t)^2}-1\right] \\ &= \frac{1}{2(1-t)^2-1}. \end{aligned}$$

Segue que vale **(h3)** com $\mu=1$, pois $g'(t)>0$. Agora, fazendo $\mathcal{M}:=(\mathbb{R}^n, \langle\cdot,\cdot\rangle_{p_*})$ e $X=h'$ concluímos que $X$ e $f$ satisfaz (2-1) com $R=1$. Neste caso como $K_{p_*}=1$ e $r_{p_*}=\infty$, as constantes $\nu$, $\rho$, e $\sigma$ definidas no Teorema 2.1, satisfaz

$$\rho=\frac{5-\sqrt{17}}{4}<\nu=\frac{\sqrt{2}-1}{\sqrt{2}}<1, \quad ,\sigma=1/2, \quad r:=\min\{\kappa,\rho\}.$$

De fato, $f(t)<0,$ se, e somente se $2t<1$ e $f'(t)<0,$ se, e somente se $1/\sqrt{2}<1-t$. Assim, o resultado segue pelas definições de $\sigma$ e $\nu$. Além disso, temos

$$\frac{f(t)}{tf'(t)}-1<1 \Longleftrightarrow \frac{t}{2(1-t)^2-1}\le 1 \Longleftrightarrow 0\le q(t).$$

Onde $q(t)=2t^2-5t+1$, mas

$$\begin{aligned} q(t)=0 &\Longleftrightarrow t=1+\frac{1\pm\sqrt{17}}{4} \\ &\Longleftrightarrow t=\frac{5-\sqrt{17}}{4} \end{aligned}$$

Pois $t\in[0,1)$. Como $q(0)=1,$ e $q(1)<0$, o resultado segue da continuidade de $q$ e pela definição de $\rho$. Agora, $p_{k+1}=\exp_{p_k}(-\nabla X(p_k)^{-1}X(p_k))$ e

$$t_{k+1}=\left|t_k-\frac{f(t_k)}{f'(t_k)}\right|=\frac{t_k^2}{2(1-t_k)^2-1}$$

Assim, para todo $k$

$$\frac{t_{k+1}}{t_k^2}=\frac{1}{2(1-t_k)^2-1}<\frac{1}{2(1-t_0)^2-1}=\frac{1}{2(1-d(p_0,p_*))^2-1}.$$

Como $X = h'$, usando (2-24) e (2-25), a sequência de Newton fica

$$p_{k+1} = p_k - h''(p_k)^{-1} g'(p_k).$$

Assim, o resultado segue pela aplicação do Teorema 2.1. □

Capítulo 3

# Análise semi-local do Método de Newton

Neste capítulo, nosso principal objetivo é enunciar e provar um teorema de convergência semi-local para o método de Newton. O teorema principal e os demais resultados e provas enunciados neste capítulo são baseados na referencia [3].

## 3.1 Método de Newton e função majorante

Nesta seção definiremos função majorante que estabelece uma condição sobre a derivada de um campo de vetores $X$ sobre uma variedade Riemanniana. Esta condição é chamada condição majorante e é uma generalização da condição de Lipschitz na Definição 1.31. Além disso, provaremos alguns resultados preliminares em relação à função majorante e enunciaremos e provaremos nosso Teorema principal.

**Definição 3.1** *Seja $\mathcal{M}$ uma variedade Riemanniana e $X : \mathcal{M} \to T\mathcal{M}$ um campo de vetores continuamente diferenciável. Então para $p_0 \in \mathcal{M}$ e $T > 0$, definiremos a classe $\mathcal{G}_2(p_0, r)$ como sendo: Todas as curvas geodésicas por partes $\zeta : [0,T] \to \mathcal{M}$ que satisfaz as seguentes condições:*

*(1) $\zeta(0) = p_0$ e o comprimento de $\zeta$ não é mais grande que $r$.*

*(2) Existe $\tau \in (0,T]$ tal que $\zeta|_{[0,\tau]}$ é uma geodésica minimizante e $\zeta|_{[\tau,T]}$ é uma geodésica.*

Suponhamos que para algum $R > 0$ existe uma função continua e não decrescente $\mathcal{L} : [0,R] \to [0,\infty)$ satisfazendo as seguentes propriedades: para todo $r \in [0,R]$ e $\zeta \in \mathcal{G}_2(p_0, r)$,

$$\left\| \nabla X(p_0)^{-1}[P_{\zeta,b,0}\nabla X(\zeta(b))P_{\zeta,a,b} - P_{\zeta,a,0}\nabla X(\zeta(a))] \right\| \le \mathcal{L}(r)\ell[\zeta,a,b], \tag{3-1}$$

onde $0 \le a \le b$. A função $\mathcal{L}$ é uma parametrização radial do tipo de Lipschitz-continua para $\nabla X$ em torno $p_0$. Sem perda de generalidade podemos assumir que $\mathcal{L}(r) > 0$ para todo $r \in [0,R]$. Seja $f:[0,R] \to \mathbb{R}$ definida por

$$\begin{cases} f''(r) = \mathcal{L}(r), & \text{para todo } \; r \in (0,R), \\ f'(0) = -1, \\ f(0) = \beta, \end{cases}$$

onde $\beta \geq 0$. É fácil ver que a função

$$f(r) = \beta - r + \int_0^r (r-s)\mathcal{L}(s)ds, \tag{3-2}$$

é uma sulução da equação acima e, além disso, temos

$$\begin{aligned} \int_0^r (r-s)\mathcal{L}(s)ds &= r\int_0^r \mathcal{L}(s)ds - \int_0^r s\mathcal{L}(s)ds \\ &= rf'(r) - rf'(0) - \left[ sf'(s)|_0^r - \int_0^r f'(s)ds \right] \\ &= r + f(r) - \beta \end{aligned}$$

de modo que

$$f'(r) = -1 + \int_0^r \mathcal{L}(s)ds,$$

para todo $r \in [0,R]$. De agora em diante vamos também supor que que $X$ satisfaz

$$\left\| \nabla X(p_0)^{-1} X(p_0) \right\| \leq \beta. \tag{3-3}$$

**Definição 3.2** *A função continuamente diferenciável $f$ definida acima é dita função majorante no ponto $p_0$ para o campo de vetores $X$ com respeito a $\mathcal{G}_2(p_0,r)$.*

De agora em diante vamos a assumir que a função majorante $f$ tem um único zero $r_*$ em $[0,R]$ com $f(R) \leq 0$. Assim, por (3-2) temos que $r_* \leq \beta$, assim podemos supor $r_* > 0$, caso contrário $r_* = 0$ implica $p_0$ é uma singularidade de $X$ e não há nada a fazer. Também pode acontecer $r_* = R$.

**Observação 3.3** *sob condições de regularidade apropriados no campo de vetores $X$, (3-1) sempre se mantém em uma vizinhança de $p_0$. Por exemplo, se assumimos que $X$ é de classe $\mathscr{C}^2$ então podemos tomar $\bar{\ell} := \sup_{\zeta \in \mathcal{G}_2(p_0,r)} \{ \left\| \nabla X(p_0)^{-1} [P_{\zeta,t,0} \nabla^2(\zeta(t))] \right\|_0 \}$. De fato, mediante um calculo direito é possível mostrar que o Lema 1.29 implica que se tomamos $\ell = \bar{\ell}$ então (3-1) se mantém. (veja Lema 2.17)*

**Lema 3.4** *Seja $f$ a função dado por (3-2). Então:*

$(i)$ *$f$ é estritamente convexa por isso $f'$ é crescente em $[0,R]$, e $f'(r)<0$ para todo $r\in[0,r_*)$.*

$(ii)$ *Se $r\in[0,r_*)$ e $r_+=r-f'(r)^{-1}f(r)$, então $r_+\in(r,r_*)$.*

$(iii)$ *Para todo $r\in[0,R]$, $\beta-r\le f(r)\le\beta+r(f'(r)-1)/2$. Consequentemente, $\beta\le r_*\le 2\beta$.*

$(iv)$ *Para cada $a\in(0,R)$, a função $f_a(r)=f'(r+a)-f'(r)=\int_r^{r+a}L(s)ds>0$ é não decrescente em $[0,R-a]$.*

**Demonstração**.

$(i)$ Como $f''(r)=L(r)>0$ para todo $r\in(0,R)$, a primeira parte é imediata. Para provar que $f'(r)<0$ para todo $r\in[0,r_*)$. Suponhamos por contradição que existe $\tilde{r}\in[0,r_*)$, tal que $f'(\tilde{r})=0$. Pela convexidade de $f$, $\tilde{r}$ é um minimo em $[0,R]$, então $0\le f(\tilde{r})\le f(R)\le 0$. Portanto, $\tilde{r}$ é um zero de $f$, o que contradiz a unicidade de $r_*$. Como $f'(0)=-1$ e pela continuidade de $f'$, temos que $f'<0$ em $[0,r*)$.

$(ii)$ Se $r\in[0,r_*)$, por $(i)$ temos $r_+>r$, por outro lado como $f$ é estritamente convexa vale (2-5) com $y=r_*$ e $x=r$, como $f(r_*)=0$ temos $0>f'(r)(r_*-r_+)$. Por $(i)$ segue que $r_+<r_*$.

$(iii)$ Por (3-2) temos $\beta-r\le f(r)$. Por outro lado

$$\begin{aligned}2f(r)&=2\beta-2r+2r\int_0^r sL(s)ds-2\int_0^r sL(s)ds\\&=2\beta+r\left(-1+\int_0^r L(s)ds\right)-r+\int_0^r(r-2s)L(s)ds\\&=2\beta+r(f'(r)-1)+\omega(r),\end{aligned}$$

pois $f'(r)=-1+\int_0^r L(s)ds$. Onde $\omega(r)=\int_0^r(r-2s)L(s)ds$, dai como $L$ é estritamente crescente segue que

$$\omega'(r)=\int_0^r L(s)ds-rL(r)=\int_0^r[L(s)-L(r)]ds\le 0.$$

Portanto $\omega(r)\le\omega(0)=0$, e assim $f(r)\le\beta+r(f'(r)-1)/2$. para todo $r\in[0,R]$. Em particular para $r_*$, por ser $f'(r_*)<0$, temos $\beta-r_*\le 0\le\beta-r_*/2$ Assim, $\beta\le r_*\le 2\beta$.

$(iv)$ Como $L$ é crescente em $[0,R]$, $f_a'(r)=L(r+a)-L(r)\ge 0$, para todo $r\in[0,R-a]$.

□

**Corolário 3.5** *A sequencia escalar $\{r_k\}$ definida por*

$$\begin{cases} r_0 = 0, \\ r_{k+1} = r_k - f'(r_k)^{-1} f(r_k). \end{cases} \tag{3-4}$$

*esta bem definida com $f'(r_k) < 0$ para todo $k \in \mathbb{N}$, consequentemente monótona crescente e converge a $r_*$.*

**Demonstração**. A demostração é uma imediata consequência do Lema 3.4 $(i)$, $(ii)$. □

O seguinte lema mostra que a condição majorante (3-1) pode ser definida equivalentemente como (2-1).

**Lema 3.6** *Se $\zeta \in \mathcal{G}_2(p_0, r)$ então para quaisquer $a \leq b$ temos que*

$$\left\| \nabla X(p_0)^{-1} [P_{\zeta,b,0} \nabla X(\zeta(b)) P_{\zeta,a,b} - P_{\zeta,a,0} \nabla X(\zeta(a))] \right\| \leq f'(\ell[\zeta,0,b]) - f'(\ell[\zeta,0,a]), \tag{3-5}$$

*como consequência, se $f'(\ell[\zeta,0,t]) < 0$ então $\nabla X(\zeta(t))$ é invertível e, Além disso,*

$$\left\| [P_{\zeta,t,0} \nabla X(\zeta(t))]^{-1} \nabla X(p_0) \right\| \leq \frac{1}{|f'(\ell[\zeta,0,t])|}. \tag{3-6}$$

**Demonstração**. Seja $\triangle_0 = \left\| \nabla X(p_0)^{-1} [P_{\zeta,b,0} \nabla X(\zeta(b)) P_{\zeta,a,b} - P_{\zeta,a,0} \nabla X(\zeta(a))] \right\|$.
Para $N \in \mathbb{N}$ com $N \geq 1$, defina $t_i = a + (i/N)(b-a)$ e $y_i = c(t_i)$ para todo $i \in \{1, \cdots, N\}$.
Por (3-1) e como o transporte paralelo é uma isometria temos

$$\begin{aligned}
\triangle_0 &= \left\| \nabla X(p_0)^{-1} \left[ P_{\zeta,b,0} \nabla X(\zeta(b)) P_{\zeta,a,b} - P_{\zeta,a,0} \nabla X(\zeta(a)) \pm \sum_{i=1}^{N-1} P_{\zeta,t_i,0} \nabla X(y_i) P_{\zeta,t_0,t_i} \right] \right\| \\
&\leq \sum_{i=0}^{N-1} \left\| \nabla X(p_0)^{-1} \left[ P_{\zeta,t_{i+1},0} \nabla X(y_{i+1}) P_{\zeta,t_0,t_{i+1}} - P_{\zeta,t_i,0} \nabla X(y_i) P_{\zeta,t_0,t_i} \right] \right\|_{p_0} \\
&\leq \sum_{i=0}^{N-1} \left\| \nabla X(p_0)^{-1} \left[ P_{\zeta,t_{i+1},0} \nabla X(y_{i+1}) P_{\zeta,t_i,t_{i+1}} - P_{\zeta,t_i,0} \nabla X(y_i) \right] \right\|_{p_0} \\
&\leq \sum_{i=0}^{N-1} \mathcal{L}(r_i) \ell[\zeta, t_i, t_{i+1}]
\end{aligned} \tag{3-7}$$

Para todo $r_i \in [0,R]$. Escolhendo $N$ de forma que $\zeta|_{[t_i,t_{i+1}]}$ seja uma geodésica, temos que $\ell[\zeta,t_i,t_{i+1}] = \|\zeta'(t_i)\| (t_{i+1} - t_i) = \|\zeta'(t_i)\| (b-a)/N$ e como $\ell[\zeta,0,t_i] \in [0,R]$ para todo $i$.
Substituindo em (3-7) y tomando o limite quando $N \to \infty$, temos

$$\triangle_0 \leq \lim_{N\to\infty} \frac{b-a}{N} \sum_{i=1}^{N} \mathcal{L}(\ell[\zeta,0,t_i]) \left\| \zeta'(t_i) \right\| = \int_a^b \mathcal{L}(\ell[\zeta,0,\tau]) \left\| \zeta'(\tau) \right\| d\tau = \int_{\ell[\zeta,0,a]}^{\ell[\zeta,0,b]} \mathcal{L}(s) ds$$

que prova (3-5).

Agora, fazendo $a=0$ e $b=t$ em (3-5) obtemos que

$$\left\|\nabla X(p_0)^{-1}P_{\zeta,t,0}\nabla X(\zeta(t))P_{\zeta,0,t}-I\right\| \leq f'(\ell[\zeta,0,t])+1,$$

onde $I$ é a identidade em $T_{p_0}\mathcal{M}$. Se $f'(\ell[\zeta,0,t])<0$, então esta estimativa é estritamente menor que 1, então pelo Lema de Banach 1.32 $\nabla X(p_0)^{-1}P_{\zeta,t,0}\nabla X(\zeta(t))P_{\zeta,0,t}$ é invertível, e portanto $\nabla X(\zeta(t))$ também. Além disso vale

$$\begin{aligned}\left\|\left[P_{\zeta,t,0}\nabla X(\zeta(t))\right]^{-1}\nabla X(p_0)\right\| &= \left\|\left[\nabla X(p_0)^{-1}P_{\zeta,t,0}\nabla X(\zeta(t))P_{\zeta,0,t}\right]^{-1}\right\| \\ &\leq -f'(\ell[\zeta,0,t])^{-1}.\end{aligned}$$

pois a $P_{\zeta,t,0}$ é uma isometria. □

**Teorema 3.7** *Seja $\mathcal{M}$ uma variedade Riemanniana completa, $\Omega\subset\mathcal{M}$ um subconjunto aberto de $\mathcal{M}$ e $X:\Omega\to T\mathcal{M}$ um campo de vetores continuamente diferenciável. Seja $p_0\in\mathcal{M}$ tal que $\nabla X(p_0)$ é invertível e $f:[0,R]\to\mathcal{M}$ uma função majorante para $X$ em $p_0$ com respeito a $\mathcal{G}_2(p_0,r)$, então temos os seguentes resultados:*

- $(i)$ *O campo de vetores $X$ admite uma única singularidade $p_*\in B[p_0,R]$ com $d(p_*,p_0)\leq r_*$. Se $f'(r_*)<0$ Então $\nabla X(p_*)$ é invertível.*
- $(ii)$ *A sequencia $\{p_k\}$ definida em (1-10) com ponto inicial $p_0$, esta bem definida, isto é, $\nabla X(p_k)$ é invertível para todo $k\in\mathbb{N}$.*
- $(iii)$ *Para todo $k\in\mathbb{N}$, $d(p_*,p_k)\leq r_k$ y a seguente estimativa se tem*

$$d(p_{k+1},p_k)\leq\left\|\nabla X(p_k)^{-1}x(p_k)\right\|\leq r_{k+1}-r_k \tag{3-8}$$

*onde $\{r_k\}$ é dada por (3-4). Como consequência, $p_k$ converge a $p_*$, quando $k\to\infty$, e além disso, para cada $k\in\mathbb{N}$,*

$$d(p_*,p_k)\leq r_*-r_k. \tag{3-9}$$

- $(iv)$ *Para todo $k\in\mathbb{N}$,*

$$\frac{\left\|\nabla X(p_{k+1})^{-1}X(p_{k+1})\right\|}{r_{k+2}-r_{k+1}}\leq\left(\frac{\left\|\nabla X(p_k)^{-1}X(p_k)\right\|}{r_{k+1}-r_k}\right)^2. \tag{3-10}$$

*Consequentemente, para todo $k\geq k_0\geq 0$,*

$$d(p_*,p_k)\leq(r_*-r_k)\left(\frac{\left\|\nabla X(p_{k_0})^{-1}X(p_{k_0})\right\|}{r_{k_0+1}-r_{k_0}}\right)^{2^{k-k_0}}. \tag{3-11}$$

$(v)$ *Seja $q \in (0,1]$, como sendo*

$$q = \frac{1-\sqrt{1-2\lambda}}{1+\sqrt{1-2\lambda}}, \tag{3-12}$$

*com*

$$\lambda = \frac{r_*^2 f''(r_*)^2 - 2r_* f''(r_*) f'(r_*)}{2[r_* f''(r_*) - f'(r_*)]^2} \tag{3-13}$$

*Então, para todo $k \in \mathbb{N}$, $r_k \geq r_*(1-q^{2^k-1})(1-q^{2^k})^{-1}$ se $q < 1$, e $r_k \geq r_*(1-2^{-k})$ se $q = 1$. Em particular, se $\lambda < 1/2$ então, para todo $k \in \mathbb{N}$,*

$$r_* - r_k \leq r_* \frac{1-q}{1-q^{2^k}} q^{2^k-1}. \tag{3-14}$$

### 3.1.1 Existência, convergência e Unicidade

A seguir, apresentaremos e demostraremos três lemas, dos quais precisaremos na hora de fazer a prova do teorema principal.

**Lema 3.8** *Seja $r \in [0, r_*)$ e $p \in B[p_0, r]$ tal que, $\left\|\nabla X(p)^{-1} X(p)\right\| \leq -f'(r)^{-1} f(r)$. Sejam $\tau(\theta) = r - \theta f'(r)^{-1} f(r)$ e $\xi(\theta) = \exp_p[-\theta \nabla X(p)^{-1} X(p)]$ para todo $\theta \in [0,1]$. Então, para todo $\theta \in (0,1]$,*

$(i)$ $\tau(\theta) \in (r, r_*)$ *e* $\xi(\theta) \in B[p_0, \tau(\theta)]$;

$(ii)$ $\nabla X(\xi(\theta))$ *é invertível;*

$(iii)$ $\left\|\nabla X(\xi(\theta))^{-1} X(\xi(\theta))\right\| \leq \dfrac{f(\tau(\theta))}{|f'(\tau(\theta))|}$.

**Demonstração**. $(i)$ Fixando $\theta \in (0,1]$. Pelo Lema 3.4 $(i)$ $\tau$ é crescente, assim $r = \tau(0) < \tau(\theta) < \tau(1) = r_+$. Agora, pela definição, temos que $\xi$ é uma geodésica com $\xi(0) = p$, $\xi'(0) = -\nabla X(p)^{-1} X(p)$ e $d(\xi(\theta), p) \leq \theta \left\|\nabla X(p)^{-1} X(p)\right\|$ daí

$$d(p_0, \xi(\theta)) \leq d(p_0, p) + d(p, \xi(\theta)) \leq r + \theta \left\|\nabla X(p)^{-1} X(p)\right\| \leq \tau(\theta).$$

Assim, $\xi(\theta) \in B[p_0, r]$ para todo $\theta \in [0,1]$.

$(ii)$ Seja $\zeta : [0,2] \to \mathcal{M}$ a curva obtida pela concatenação da geodésica minimizante ligando $p_0$ a $p$, definida em $[0,1]$, e $\xi$. Isto é $\zeta$ é uma curva geodésica por partes com $\zeta(0) = p_0$, $\zeta(1+\theta) = \xi(\theta)$ para todo $\theta \in [0,1]$. Por $(i)$ temos que $\zeta \in \mathcal{G}_2(p_0, r_*)$. Para mostrar que $X(\xi(\theta))$ é invertível consideremos

$$
\begin{aligned}
&\left\|\nabla X(p)^{-1}P_{\xi,\theta,0}\nabla X(\xi(\theta))P_{\xi,0,\theta}-I\right\| \\
&=\left\|\nabla X(p)^{-1}\left[P_{\xi,\theta,0}\nabla X(\xi(\theta))P_{\xi,0,\theta}-\nabla X(p)\right]\right\| \\
&\leq\left\|\nabla X(p)^{-1}P_{\zeta,0,1}\nabla X(p_0)\right\|\left\|\nabla X(p_0)^{-1}P_{\zeta,1,0}\left[P_{\xi,\theta,0}\nabla X(\xi(\theta))P_{\xi,0,\theta}-\nabla X(p)\right]\right\| \\
&=\left\|[P_{\zeta,1,0}\nabla X(p)]^{-1}\nabla X(p_0)\right\|\left\|\nabla X(p_0)^{-1}\left[P_{\zeta,1+\theta,0}\nabla X(\zeta(1+\theta))P_{\zeta,1,1+\theta}-P_{\zeta,1,0}\nabla X(p)\right]\right\|
\end{aligned}
$$

Segue dos Lemas 3.6 e 3.4 $(i)$.

$$
\left\|[P_{\zeta,1,0}\nabla X(p)]^{-1}\nabla X(p_0)\right\|\leq\frac{1}{|f'(d(p_0,p))|}\leq\frac{1}{|f'(r)|},
$$

e dado que

$$
\ell[\zeta,0,1+\theta]=d(p_0,p)+\ell[\xi,0,\theta]=d(p_0,p)+\theta\left\|\zeta'(0)\right\|\leq\tau(\theta) \tag{3-15}
$$

$$
\left\|\nabla X(p_0)^{-1}\left[P_{\zeta,1+\theta,0}\nabla X(\zeta(1+\theta))P_{\zeta,1,1+\theta}-P_{\zeta,1,0}\nabla X(p)\right]\right\|\leq f'(\tau(\theta))-f'(r)
$$

consequentemente,

$$
\left\|\nabla X(p)^{-1}P_{\xi,\theta,0}\nabla X(\xi(\theta))P_{\xi,0,\theta}-I\right\|\leq 1-f'(r)^{-1}f'(\tau(\theta))<1.
$$

Segue pelo Lema de Banach 1.32 que $\nabla X(\xi(\theta))$ é invertível. e vale

$$
\left\|\left[P_{\xi,\theta,0}\nabla X(\xi(\theta))\right]^{-1}\nabla X(p)\right\|\leq\frac{1}{1-(1-f'(r)^{-1}f'(\tau(\theta)))}\leq\frac{f'(r)}{f'(\tau(\theta))}. \tag{3-16}
$$

Agora, só resta provar $(iii)$. Para isto seja $\Theta=\left\|\nabla X(\xi(\theta))^{-1}X(\xi(\theta))\right\|$, como $\xi$ é geodésica, $\xi'$ é um campo paralelo, assim $\xi'(s)=P_{\xi,0,s}\xi'(0)=-P_{\xi,0,s}\nabla X(p)^{-1}X(p)$ segue do Lema 1.29

$$
\begin{aligned}
\Theta&=\left\|\nabla x(\xi(\theta))^{-1}P_{\xi,0,\theta}\left[X(p)\pm\theta X(p)+\int_0^\theta P_{\xi,s,0}\nabla X(\xi(s))\xi'(s)ds\right]\right\| \\
&=\left\|\nabla X(\xi(\theta))^{-1}P_{\xi,0,\theta}\left[(1-\theta)X(p)+\int_0^\theta\left[\nabla X(p)+P_{\xi,s,0}\nabla X(\xi(s))P_{\xi,0,s}ds\right]\xi'(0)\right]\right\| \\
&\leq(1-\theta)\left\|\nabla X(\xi(\theta))^{-1}P_{\xi,0,\theta}X(p)\right\|+R(\theta) \qquad (3\text{-}17)
\end{aligned}
$$

onde

$$
\begin{aligned}
R(\theta) &= \left\| \nabla X(\xi(\theta))^{-1} P_{\xi,0,\theta} \int_0^\theta \left[ \nabla X(p) + P_{\xi,s,0} \nabla X(\xi(s)) P_{\xi,0,s} ds \right] \xi'(0) \right\| \\
&= \left\| \nabla X(\zeta(1+\theta))^{-1} P_{\zeta,1,1+\theta} \int_0^\theta \left[ \nabla X(p) + P_{\xi,s,0} \nabla X(\xi(s)) P_{\xi,0,s} ds \right] \xi'(0) \right\| \\
&= \left\| [P_{\zeta,1+\theta,0} \nabla X(\zeta(1+\theta))]^{-1} P_{\zeta,1,0} \int_0^\theta \left[ \nabla X(p) + P_{\xi,s,0} \nabla X(\xi(s)) P_{\xi,0,s} ds \right] \xi'(0) \right\|
\end{aligned}
$$

Por (3-16)

$$
\begin{aligned}
\left\| \nabla X(\xi(\theta))^{-1} P_{\xi,0,\theta} X(p) \right\| &= \left\| \left[ P_{\xi,\theta,0} \nabla X(\xi(\theta)) \right]^{-1} X(p) \right\| \\
&\leq \left\| \left[ P_{\xi,\theta,0} \nabla X(\xi(\theta)) \right]^{-1} \nabla X(p) \right\| \left\| \nabla X(p)^{-1} X(p) \right\| \\
&\leq \frac{f'(r)}{f'(\tau(\theta))} [-f'(r)^{-1} f(r)] \\
&\leq \frac{f(r)}{|f'(\tau(\theta))|}.
\end{aligned} \tag{3-18}
$$

agora lembrado que $f'$ é crescente, por (3-6) e (3-15) segue

$$
\left\| [P_{\zeta,1+\theta,0} \nabla X(\zeta(1+\theta))]^{-1} \nabla X(p_0) \right\| \leq \frac{1}{|f'(\ell[\zeta,0,1+\theta])|} \leq \frac{1}{|f'(\tau(\theta))|},
$$

daí, pela equação acima, por (3-15) e (3-5)

$$
\begin{aligned}
& R(\theta) \\
& \leq \frac{1}{|f'(\tau(\theta))|} \left\| \nabla X(p_0)^{-1} P_{\zeta,1,0} \int_0^\theta \left[ \nabla X(p) + P_{\xi,s,0} \nabla X(\xi(s)) P_{\xi,0,s} ds \right] \xi'(0) \right\| \\
& \leq \frac{1}{|f'(\tau(\theta))|} \int_0^\theta \left\| \nabla X(p_0)^{-1} \left[ P_{\zeta,1,0} \nabla X(p) + P_{\zeta,s+1,0} \nabla X(\zeta(s+1)) P_{\zeta,1,s+1} \right] \xi'(0) \right\| ds \\
& \leq \frac{\left\| \nabla X(p)^{-1} X(p) \right\|}{|f'(\tau(\theta))|} \int_0^\theta \left\| \nabla X(p_0)^{-1} \left[ P_{\zeta,s+1,0} \nabla X(\zeta(s+1)) P_{\zeta,1,s+1} - P_{\zeta,1,0} \nabla X(p) \right] \right\| ds \\
& \leq \frac{f'(r)^{-1} f(r)}{f'(\tau(\theta))} \int_0^\theta \left[ f'(\ell[\zeta,0,1+s]) - f'(\ell[\zeta,0,1]) \right] ds \\
& = -\frac{1}{f'(\tau(\theta))} \int_0^\theta \left[ f'(\tau(s)) \tau'(s) - f(r) \right] ds \\
& = \frac{1}{|f'(\tau(\theta))|} [f(\tau(\theta)) - f(\tau(0)) - \theta f(r)] \\
& = \frac{1}{|f'(\tau(\theta))|} [f(\tau(\theta)) - (1-\theta) f(r)]
\end{aligned}
$$

pela equação acima junto com (3-17) e (3-18) obtemos,

$$\Theta \leq (1-\theta)\frac{f(r)}{|f'(\tau(\theta))|} + \frac{1}{|f'(\tau(\theta))|}\left[f(\tau(\theta)) - (1-\theta)f(r)\right] \leq \frac{f(\tau(\theta))}{f'(\tau(\theta))}$$

□

**Lema 3.9** *Seja* $0 \leq r < r_*$ *e* $q \in B[p_0, r]$ *tal que* $\left\|\nabla X(q)^{-1}X(q)\right\| \leq -f'(r)^{-1}f(r)$. *Suponha que* $q_* \in B[p_0, r_*]$ *satisfaz* $X(q_*) = 0$ *e* $r + d(q, q_*) = r_*$. *Então* $d(p_0, q) = r$. *Além disso, para* $r_+ = r - f'(r)^{-1}f(r)$ *e* $q_+ = \exp_q(-\nabla X(q)^{-1}X(q))$, *temos*

$$r_+ + d(q_+, q_*) = r_*.$$

**Demonstração**. Seja $\xi : [0, 1] \to \mathcal{M}$ uma geodésica minimizante ligando $q$ a $q_*$. Seja $v = \xi'(0)$, então $\|v\| = d(q, q_*)$, pois $\xi$ é minimizante. Assim,

$$\|v\| = d(q, q_*) \leq \left\|\nabla X(q)^{-1}X(q) + v\right\| + \left\|\nabla X(q)^{-1}X(q)\right\| \tag{3-19}$$

$$\leq \left\|\nabla X(q)^{-1}X(q) + v\right\| - f'(r)^{-1}f(r). \tag{3-20}$$

Por outro lado, como $X(q_*) = 0$ pelo Lema (1.29) e lembrando que $\xi'(s) = P_{\xi,0,s}v$

$$X(q) = -\int_0^1 P_{\xi,s,0}\nabla X(\xi(s))P_{\xi,0,s}v\, ds,$$

daí

$$\left\|\nabla X(q)^{-1}X(q) + v\right\| = \left\|\nabla X(q)^{-1}\int_0^1 \left[P_{\xi,s,0}\nabla X(\xi(s))P_{\xi,0,s} - \nabla X(q)\right] v\, ds\right\|.$$

Seja $\zeta : [0, 2] \to \mathcal{M}$ a concatenação de uma geodésica minimizante ligando $p_0$ a $q$, com $\xi$. Pelo Lema 3.4, $f'$ é crescente e $f'(d(p_0, q)) \leq f'(r) \leq 0$. Segue do Lema 3.6

$$\begin{aligned}
&\left\|\nabla X(q)^{-1}X(q) + v\right\| \\
&\leq \left\|[P_{\zeta,1,0}\nabla X(q)]^{-1}\nabla X(p_0)\right\| \left\|\nabla X(p_0)^{-1}P_{\zeta,1,0}\int_0^1 \left[P_{\xi,s,0}\nabla X(\xi(s))P_{\xi,0,s} - \nabla X(q)\right] v\, ds\right\| \\
&\leq \frac{1}{|f'(\ell[\zeta, 0, 1])|}\int_0^1 \left\|\nabla X(p_0)^{-1}\left[P_{\zeta,1+s,0}\nabla X(\xi(1+s))P_{\zeta,1,1+s} - P_{\zeta,1,0}\nabla X(q)\right] v\, ds\right\|
\end{aligned}$$

$$
\begin{aligned}
\left\|\nabla X(q)^{-1}X(q)+v\right\| &\leq \frac{\|v\|}{|f'(d(p_0,q)|}\int_0^1 \left[f'(\ell[\zeta,0,1+s])-f'(\ell[\zeta,0,1])\right] ds && \text{(3-21)}\\
&\leq \frac{\|v\|}{|f'(r)|}\int_0^1 \left[f'(r+s\|v\|)-f'(r)\right] ds && \text{(3-22)}\\
&\leq \|v\|+\frac{1}{|f'(r)|}\int_0^1 f'(r+s\|v\|)\,\|v\|\,ds\\
&= \|v\|+\cancelto{0}{\frac{f(r+\|v\|)}{|f'(r)|}}+\frac{f(r)}{f'(r)}
\end{aligned}
$$

Onde o termo da expressão acima é zero pelo fato de que $r+\|v\|=r_*$. Assim, por (3-20) temos a igualdade $\left\|\nabla X(q)^{-1}X(q)+v\right\|=\|v\|+\dfrac{f(r)}{f'(r)}$. Em particular (3-21) e (3-22) são iguais, além do mais (3-19) e (3-20) também são iguais. Como $\|v\|\neq 0$ temos

$$
\begin{aligned}
f'(d(p_0,q)) &= f'(r), && \text{(3-23)}\\
\left\|\nabla X(q)^{-1}X(q)+v\right\| &= \|v\|-\left\|\nabla X(q)^{-1}X(q)\right\|, && \text{(3-24)}\\
\left\|\nabla X(q)^{-1}X(q)\right\| &= -f'(r)^{-1}f(r). && \text{(3-25)}
\end{aligned}
$$

Como $f'$ é injetiva em $[0,r_*)$, segue de (3-23) que $d(p_0,q)=r$. □

**Lema 3.10** *Seja $q_*\in B[p_0,r_*]$ com $X(q_*)=0$. Se existem $\tilde{r},\tilde{q}$ tais que $0\leq\tilde{r}<r_*$, $\tilde{q}\in B[p_0,\tilde{r}]$, $\left\|\nabla X(\tilde{q})^{-1}X(\tilde{q})\right\|\leq -f'(\tilde{r})^{-1}f(\tilde{r})$, $\tilde{r}+d(\tilde{q},q_*)=r_*$, então $d(p_0,q_*)=r_*$.*

**Demonstração**. Sejam $\{\tau_k\}$ e $\{q_k\}$ definidas por $\tau_0=\tilde{r},\ \tau_{k+1}=\tau_k-f'(\tau_k)^{-1}f(\tau_k)$ e $q_0=\tilde{q},\ q_{k+1}=\exp_{q_k}(-\nabla X(q_k))^{-1}X(q_k)$, respectivamente. Assim, $\{\tau_k\}$ é crescente e converge para $r_*$, $\{q_k\}$ converge para algum $\tilde{q}_*\in B[p_0,r_*]$, e para todo $k$, $\left\|\nabla X(q_k)^{-1}X(q_k)\right\|\leq -f'(\tau_k)^{-1}f(\tau_k)$. Além disso, pelo Lema 3.9, temos que para todo $k$, $\tau_k+d(q_k,q_*)=r_*$ e $d(p_0,q_k)=\tau_k$. Pelo Corolário 1.11 segue que

$$
d(\tilde{q}_*,q_*)=\lim_{k\to\infty}d(q_k,q_*)=0 \qquad d(p_0,q_*)=\lim_{k\to\infty}d(p_0,q_k)=r_*
$$

□

**Demonstração do Teorema 3.7**. $(i)-(iii)$ Primeiramente, provaremos por indução que para todo $k\geq 0$, $p_k\in B[p_0,r_k]$ e $\nabla X(p_k)$ é invertível com

$$
\left\|\nabla X(p_k)^{-1}X(p_k)\right\|\leq r_{k+1}-r_k.
$$

Se $k=0$, pelas hipóteses do teorema temos que $\nabla X(p_0)$ é invertível e vale (3-3). Assim, para $k \geq 1$ assumamos que o resultado é válido para todo $i \in \{1, \cdots, k-1\}$. Isto é, como $r_{k-1} \in [0, r_*)$ e $p_{k-1} \in B[p_0, r_k]$ e

$$\left\| \nabla X(p_{k-1})^{-1} X(p_{k-1}) \right\| \leq r_k - r_{k-1} = -f'(r_{k-1}) f(r_{k-1}),$$

então vale o Lema 3.8 com $r = r_{k-1}, p = p_{k-1}$ e $\theta = 1$. Isto é $r_k \in B[p_0, r_k]$, $\nabla X(p_k)$ é invertível e $\left\| \nabla X(p_k)^{-1} X(p_{k)} \right\| \leq r_{k+1} - r_k$, pois $\tau(1) = r_k$ e $\xi(1) = p_k$.
Dai, $\{p_k\}$ esta bem definida. A estimativa (3-8) se tem de $d(p_{k+1}, p_k) \leq \left\| \nabla X(p_k)^{-1} X(p_k) \right\|$, pois $p_{k+1} = \exp_{p_k}\left(-\nabla X(p_k)^{-1} X(p_k)\right)$. Como $\{r_k\}$ converge para $r_*$ pelo Corolario 3.5, temos que $\{p_k\}$ é uma sequencia de Cauchy e como $\mathcal{M}$ é completa, então $\{p_k\}$ converge para $p_* \in \mathcal{M}$, que é uma singularidade de $X$. De fato,

$$\|X(p_k)\| \leq \|\nabla X(p_k)\| \left\| \nabla X(p_k)^{-1} X(p_k) \right\| \leq \|\nabla X(p_k)\| (r_{k+1} - r_k),$$

se $k \to \infty$ então $X(p_*) = 0$. Seja $k \in \mathbb{N}$ fixo, Pelo Corolário 1.11 temos que $d(\cdot, p_k)$ é continua. Assim, sem perda de generalidade supondo $k < j$

$$d(p_*, p_k) = \lim_{j \to \infty} d(p_j, p_k) \leq \lim_{j \to \infty} \sum_{i=k}^{j-1} d(p_i, p_{i+1}) \leq \lim_{j \to \infty} \sum_{i=k}^{j-1} \left\| \nabla X(p_i)^{-1} X(p_i) \right\| \leq r_* - r_k$$

Agora, temos $p_* \in B[p_0, r_*]$. De fato

$$d(p_*, p_0) \leq d(p_*, p_k) + d(p_k, p_0) \leq r_* - r_k + r_k.$$

Por último, Se $f'(r_*) < 0$, então pelo Lema 3.6 $\nabla X(p_*)$ é invertível. Agora, provaremos a unicidade de $p_*$ na bola $B[p_0, R]$. Começamos por estabelecer a unicidade de $p_*$ na bola $B[p_0, r_*]$. Seja $q_* \in B[p_0, r_*]$ tal que $X(q_*) = 0$. Para provar que necessariamente $q_* = p_*$, consideramos dois casos.

**Caso 1.** Se $d(p_0, q_*) < r_*$, Então temos por indução que para todo $k$, $d(p_k, q_*) + r_k < r_*$. De fato, como $q_* \in B[p_0, r_*]$ o resultado se tem para $k = 0$. Suponhamos que para $k$ fixo vale a afirmação e para $\theta \in [0, 1]$, seja $\xi(\theta) = \exp_{p_k}(-\theta \nabla X(p_k)^{-1} X(p_k))$ e $\psi(\theta) = d(\xi(\theta), q_*) + \tau(\theta)$, onde $\tau$ é a função do Lema 3.8 com $r = r_k$. Assim, temos que $d(p_k, q_*) + r_k = \psi(0) < r_*$, se existe $\tilde{\theta} \in [0, 1]$ tal que $\psi(\tilde{\theta}) = r_*$, temos

$$r_* = d(\tilde{q}, q_*) + \tilde{r}, \tag{3-26}$$

onde, $\tilde{q} = \xi(\tilde{\theta})$ e $\tilde{r} = \tau(\tilde{\theta})$. Além disso vale vale $\left\| \nabla X(p_k)^{-1} X(p_k) \right\| \leq -f'(r_k)^{-1} f(r_k)$. Dai, podemos aplicar o Lema 3.8 com $r = r_k$, $p = p_k$, e $\theta = \tilde{\theta}$. Portanto $\tilde{r} \in (r_k, r_*)$,

$\tilde{q} \in B[p_0, \tilde{r}]$, $\nabla X(\tilde{q})$ é invertível e vale $\left\|\nabla X(\tilde{q})^{-1}X(\tilde{q})\right\| \leq -f'(\tilde{r})^{-1}f(\tilde{r})$. Assim junto com (3-26) pelo Lema 3.10 segue que $d(p_0, q_*) = r_*$ o que é uma contradição. O que implica pela continuidade de $\psi$, que para todo $\theta \in [0,1]$ vale $\psi(\theta) < r_*$. Em particular $\psi(1) < r_*$, isto é $d(p_{k+1}, q_*) + r_{k+1} < r_*$.

**Caso 2.** Se $d(p_=, q_*) = r_*$, então por um argumento indutivo, os lemas 3.8 e 3.9 mostram que para todo $k \geq 0$, $d(p_k, q_*) + r_k = r_*$. Em quaisquer dos casos temos $d(p_k, q_*) + r_k \leq r_*$, Como $r_k \to r_*$ e $p_k \to p_*$, temos que $p_* = q_*$. Mostramos que $p_*$ é a única singularidade de $X$ em $B[p_0, r_*]$.

Para completar a prova se $R = r_*$ o resultado é obvio, assim $f(R) < 0$. Seja $q_* \in B[p_0, R]$, tal que $X(q_*) = 0$. Seja $\xi : [0,1] \to \mathcal{M}$ uma geodésica minimizante ligando $p_0$ a $q_*$, e $v = \xi'(0)$. Como $P_{\xi,1,0}X(q_*) = 0$, temos pelo Lema 1.29 e por (3-5)

$$
\begin{aligned}
\left\|\nabla X(p_0)^{-1}X(p_0) + v\right\| &= \left\|\nabla X(p_0)^{-1}\left[-X(p_0) - \nabla X(p_0)v\right]\right\| \\
&= \left\|\nabla X(p_0)^{-1}\int_0^1 \left[P_{\xi,s,0}X(\xi(s))P_{\xi,0,s} - \nabla X(p_0)\right] v\, ds\right\| \\
&\leq \int_0^1 \left[f'(\ell[\xi,0,s]) - f'(0)\right] \|v\|\, ds \\
&\leq f(\|v\|) - f(0) + \|v\|\,.
\end{aligned}
$$

Além do mais, por (3-19) com $q = p_0$, $\|v\| - f(0) \leq \left\|\nabla X(p_0)^{-1}X(p_0) + v\right\|$. Usando a desigualdade acima, segue que $0 \leq f(\|v\|)$. Como $\|v\| \leq R$ e $f(R) < 0$, pela continuidade de $f$ necessariamente $\|v\| \leq r_*$ o que implica que $q_* \in B[p_0, r_*]$ e então $q_* = p_*$.

$(iv)$ Dado $k \geq 0$, Seja $\zeta : [0,2] \to \mathcal{M}$ a concatenação de uma geodésica minimizante ligando $p_0$ a $p_k$ com $\xi(\theta) = \exp_{p_k}(-\theta\nabla X(p_k)^{-1}X(p_k))$, $t \in [0,1]$. Por tanto, seguindo os mesmos argumentos na demostração do Lema 3.8, temos

$$
\begin{aligned}
&\left\|\nabla X(p_0)^{-1}P_{\zeta,2,0}X(p_{k+1})\right\| \\
&= \left\|\nabla X(p_0)^{-1}P_{\zeta,1,0}\left[P_{\xi,1,0}X(p_{k+1})\right]\right\| \\
&= \left\|\nabla X(p_0)^{-1}P_{\zeta,1,0}\left[X(p_k) - \int_0^1 P_{\xi,s,0}\nabla X(\xi(s))P_{\xi,0,s}\nabla X(p_k)^{-1}X(p_k)\, ds\right]\right\| \\
&= \left\|\nabla X(p_0)^{-1}\int_0^1 \left[P_{\zeta,1,0}\nabla X(p_k) - P_{\xi,1+s,0}\nabla X(\xi(s))P_{\xi,1,1+s}\right]\nabla X(p_k)^{-1}X(p_k)\, ds\right\| \\
&\leq \beta_k \int_0^1 \left\|\nabla X(p_0)^{-1}\left[P_{\zeta,1,0}\nabla X(p_k) - P_{\xi,1+s,0}\nabla X(\xi(s))P_{\xi,1,1+s}\right]\right\| ds \\
&\leq \beta_k \int_0^1 \left[f'(\ell[\zeta,0,1+s]) - f'(\ell[\zeta,0,1])\right] ds,
\end{aligned}
$$

onde $\beta_k = \left\|\nabla X(p_k)^{-1}X(p_k)\right\|$. Agora, como $\zeta|_{[0,1]}$ é geodésica minimizante temos que $\ell[\zeta,0,1] = d(p_0,p_k)$ e $\ell[\zeta,0,1+s] = d(p_0,p_k)+s\beta_k$. Portanto pelo Lema 3.4 (iv)

$$\begin{aligned}
\left\|\nabla X(p_0)^{-1}P_{\zeta,2,0}X(p_{k+1})\right\| &\le \beta_k \int_0^1 \int_{d(p_0,p_k)}^{d(p_0,p_k)+s\beta_k} \mathcal{L}(r)\,dr\,ds\\
&= \beta_k \int_0^1 \int_0^{s\beta_k} \mathcal{L}(d(p_0,p_k)+r)\,dr\,ds\\
&= \beta_k \int_0^{\beta_k} \int_{\frac{r}{\beta_k}}^1 \mathcal{L}(d(p_0,p_k)+r)\,ds\,dr\\
&= \int_0^{\beta_k} \mathcal{L}(d(p_0,p_k)+r)(\beta_k - r)\,dr
\end{aligned}$$

Considerando a função $\psi_a(s) = (1/s^2)\int_0^s \mathcal{L}(a+r)(s-r)dr$ para $s\in[0,R-a]$. é possível verificar que $\psi_a$ é não decrescente. Por (3-8) temos $\beta_k \le r_{k+1}-r_k$ e $d(p_k,p_0)\le r_k$, dai

$$\left\|\nabla X(p_0)^{-1}P_{\zeta,2,0}X(p_{k+1})\right\| \le \beta_k^2\,\psi_{d(p_0,p_k)}(r_{k+1}-r_k) \le \beta_k^2\,\psi_{r_k}(r_{k+1}-r_k),$$

mas, por (3-2)

$$\begin{aligned}
f(r_{k+1}) &= \beta - r_{k+1} + \int_0^{r_{k+1}} \mathcal{L}(s)(r_{k+1}-s)ds\\
&= -r_{k+1} + f(r_k) + r_k - f'(r_k)^{-1}f(r_k)(f'(r_k)+1) + \int_{r_k}^{r_{k+1}} \mathcal{L}(s)(r_{k+1}-s)\,ds\\
&= \int_{r_k}^{r_{k+1}} \mathcal{L}(s)(r_{k+1}-s)\,ds\\
&= \int_0^{r_{k+1}-r_k} \mathcal{L}(s+r_k)(r_{k+1}-s-r_k)\,ds\\
&= (r_{k+1}-r-k)^2\,\psi_{r_k}(r_{k+1}-r_k)
\end{aligned}$$

dai,

$$\left\|\nabla X(p_0)^{-1}P_{\zeta,2,0}X(p_{k+1})\right\| \le f(r_{k+1})\frac{\beta_k^2}{(r_{k+1}-r_k)^2}.$$

Pelo Lema 3.6 obtemos

$$\begin{aligned}
\left\|\nabla X(p_{k+1})^{-1}X(p_{k+1})\right\| &\le \left\|[P_{\zeta,2,0}\nabla X(p_{k+1})]^{-1}\nabla X(p_0)\right\|\left\|\nabla X(p_0)^{-1}P_{c,2,0}X(p_{k+1})\right\|\\
&\le \frac{f(r_{k+1})}{|f'(\ell[\zeta,0,2])|}\frac{\beta_k^2}{(r_{k+1}-r_k)^2}\\
&\le \frac{f(r_{k+1})}{|f'(r_{k+1})|}\frac{\beta_k^2}{(r_{k+1}-r_k)^2}\\
&\le (r_{k+2}-r_{k+1})\frac{\beta_k^2}{(r_{k+1}-r_k)^2}
\end{aligned}$$

Assim, para todo $k \in \mathbb{N}$ vale (3-10). Se $k \geq k_0 \geq 0$ e $j \geq 0$, fazendo uso da desigualdade acima $j+k-k_0$-vesses implica,

$$\begin{aligned} d(p_{k+j+1}, p_{k+j}) &\leq \left\| \nabla X(p_{k+j})^{-1} X(p_{k+j}) \right\| \\ &\leq (r_{k+n+1} - r_{k_j}) \left[ \frac{\beta_{k_0}}{r_{k_0+1} - r_{k_0}} \right]^{2^{j+k-k_0}} \\ &\leq (r_{k+n+1} - r_{k_j}) \left[ \frac{\beta_{k_0}}{r_{k_0+1} - r_{k_0}} \right]^{2^{k-k_0}}. \end{aligned}$$

Pois $\beta_{k_0} \leq r_{k_0+1} - r_{k_0}$. Somando na última desigualdade para todo $j \geq 0$. Obtemos (3-11). $(v)$ A continuação provaremos a última parte do Teorema 3.7. Definamos $Q : [0, R] \to \mathbb{R}$ como o seguente polinômio quadrático

$$Q(r) = f'(r_*)(r - r_*) + \frac{1}{2} f''(r_*)(r - r_*)^2, \tag{3-27}$$

Cuja menor raiz é $r_*$, porque $f''(r_*) < 0$ e $f'(r_*) \leq 0$. O método de Newton para resolver a equação $Q(r) = 0$, começando no ponto $\rho_0 := 0$, gera a sequencia

$$\rho_{k+1} = \rho_k - Q(\rho_k)^{-1} Q(\rho_k). \tag{3-28}$$

É fácil verificar que $\{\rho_k\}$ esta bem definida, é monotonicamente crescente e convergente para $r_*$. Além disso

$$\rho_k = \begin{cases} r_*(1 - q^{2^k - 1})(1 - q^{2^k})^{-1} & q \in (0, 1), \\ r_*(1 - 2^{-k}) & q = 1, \end{cases}$$

onde $q$ é dado em (3-12); veja por exemplo [13, 27]. Por lo tanto o Teorema 3.7-$(v)$ fica demostrado mediante uma aplicação direta do seguinte lema. □

**Lema 3.11** *sejam $\{r_k\}$ e $\{\rho_k\}$ como sendo, respetivamente definidas por* (3-28) *e* (3-28), *então para todo $k \geq 0$, $r_k \geq r_k$.*

**Demonstração**. A prova sera feita por indução. A propriedade é imediata para $k=0$ pois $r_0 = \rho_0$. Agora, suponhamos que $\rho_k \leq r_k$ para algum $k \geq 0$. usando (3-28) temos

$$\begin{aligned} r_{k+1} - \rho_{k+1} &= r_{k+1} - r_k + \rho_k - \rho_{k+1} + r_k - \rho_k \\ &= -f'(r_k)^{-1} f(r_k) + Q'(\rho_k)^{-1} Q(\rho_k) r_k - \rho_k \\ &= Q'(\rho_k)^{-1} \left[ Q(\rho_k) + Q'(\rho_k)(r_k - \rho_k) \right] - f'(r_k)^{-1} f(r_k). \end{aligned}$$

Pela convexidade, $Q(\rho_k) + Q'(\rho_k)(r_k - \rho_k) \leq Q(r_k)$, e $Q(\rho_k) \leq Q(r_k)$ pois $\rho_k \leq r_k \leq r_*$. Como $Q'$ é negativa em $[0, r_*]$, deduzimos que

$$Q'(\rho_k)^{-1} \left[ Q(\rho_k) + Q'(\rho_k)(r_k - \rho_k) \right] \geq Q'(\rho_k)^{-1} Q(r_k) \geq Q'(r_k)^{-1} Q(r_k)$$

segue que

$$r_{k+1} - \rho_{k+1} \geq Q'(r_k)^{-1} Q(r_k) - f'(r_k)^{-1} f(r_k).$$

Como $f''$ é não decrescente $f'$ é convexa e então

$$f(r_k) \geq -\frac{f'(r_k) + f'(r_*)}{2} (r_* - r_k).$$

Segue-se que

$$-f'(r_k)^{-1} f(r_k) \geq \frac{1}{2}(r_* - r_k) \left[ 1 + f'(r_k)^{-1} f'(r_*) \right].$$

Por outro lado $Q'(r_k)^{-1} Q(r_k) = \frac{1}{2}(r_* - r_k) \left[ -1 - Q'(r_k)^{-1} f'(r_*) \right]$. De fato,

$$\begin{aligned} &2Q(r_k) + Q'(r_k)(r_* - r_k) \\ &= 2\left[ f'(r_*)(r_k - r_*) - \frac{1}{2} f''(r_*)(r_- r_*)^2 \right] + Q'(r_k)[r_* - r_k] \\ &= 2f'(r_*)(r_k - r_*) + f''(r_*)(r_k - r_*)^2 + [f'(r_*) + f''(r_*)(r_k - r_*)](r_* - r_k) \\ &= f'(r_k)(r_k - r_*) \end{aligned}$$

assim, $Q(r_k) = -\frac{1}{2}[f'(r_*) + Q'(r_k)](r_* - r_k)$. Mas pela convexidade de $f'$ temos $Q'(r_k) = f'(r_*) + f''(r_*)(r_k - r_*) \leq f'(r_k) < 0$. Como $f'(r_*) \leq 0$.

$$Q'(r_k)^{-1} Q(r_k) \geq -\frac{1}{2}(r_* - r_k)[1 + f'(r_k)^{-1} f'(r_*)]$$

Portanto, obtemos

$$r_{k+1} - \rho_{k+1} = Q'(r_k)^{-1} Q(r_k) - f'(r_k)^{-1} f(r_k) \geq 0.$$

O que prova o resultado. □

## 3.2 Casos especiais

Nesta seção apresentaremos 3 casos especiais do Teorema 3.7.

### 3.2.1 Teorema de Kantorovich

O Teorema de Kantorovich para o método de Newton em espaços de Banach veja [17], fornece um critério para (quadrática) convergência que é verificável no ponto inicial, para isto é necessário conhecer uma constante de Lipschitz local. Uma extensão do Teorema de Kantorovich para variedades Riemannianas completas de dimensão finita foi obtida por Ferreira e Svaiter em [9]. Neste primeiro caso vamos ver que a extensão do Teorema de Kantorovich [9, Teorema 3.2] pode ser visto como um caso especial do Teorema 3.7.

**Teorema 3.12** *Dado $p_0 \in \mathcal{M}$ tal que $\nabla X(p_0)$ é invertível, seja $a = \left\|\nabla X(p_0)^{-1}\right\|$ e $\beta \geq \left\|\nabla X(p_0)^{-1}X(p_0)\right\|$. Assumamos que $2a\beta L \leq 1$, onde a constante $L$ é tal que $\nabla X$ é L-Lipschitz continuo na bola fechada $B[p_0, R_0]$ para algum $R_0 \geq r_*$ onde*

$$r_* = \frac{1-\sqrt{1-2a\beta L}}{aL}.$$

*Então a sequencia $\{p_k\}$ gerado pelo método de Newton com ponto inicial $p_0$ esta bem definida e converge à singularidade $p_*$ de $X$.*
*Se $2a\beta L = 1$, isto é, $r_* = 2\beta$, então $p_*$ é a única singularidade de $X$ em $B[p_0, 2\beta]$.*
*Se $2a\beta L < 1$ então $\nabla X(p_*)$ é invertível e $p_*$ é a única singularidade de $X$ em $B[p_0, R]$ para qualquer $R \in [r_*, R_0]$ tal que $R < \frac{1+\sqrt{1-2a\beta L}}{aL}$. Em quaisquer dos casos, $d(p_*, p_0) \leq r_* \leq 2\beta$ e, para todo $k \in \mathbb{N}$, temos*

$$d(p_*, p_k) \leq r_* - r_k = r_* \left[\frac{1-q}{1-q^{2^k}}\right] q^{2^k-1},$$

*onde $q$ é dado por (3-12) para $\lambda = a\beta L$, e $\{r_k\}$ é a sequencia gerada por o método de Newton, com ponto inicial $r_0 = 0$, aplicado à função escalar*

$$f(r) = \beta - r + \frac{aL}{2}r^2. \tag{3-29}$$

**Demonstração**. Sem perda de generalidade suponhamos $\beta > 0$. Seja $\zeta \in \mathcal{G}_2(p_0, r)$ e sejam $(t, y)$ e $(t', y')$ em $\zeta$ com $t \leq t'$. Assim, pela Definição 1.31 e como o transporte paralelo é uma isometria, temos

$$
\begin{aligned}
&\left\|\nabla X(p_0)^{-1}[P_{\zeta,t',0}\nabla X(y')P_{\zeta,t',t} - P_{\zeta,t,0}\nabla X(y)]\right\|_{p_0} \\
&\qquad\qquad \leq \left\|\nabla X(p_0)^{-1}\right\|_{p_0} \left\|P_{\zeta,t,0}[P_{\zeta,t',t}\nabla X(y')P_{\zeta,t,t'} - \nabla X(y)]\right\|_y \\
&\qquad\qquad = a\left\|P_{\zeta,t',t}\nabla X(\zeta(t'))P_{\zeta,t,t'} - \nabla X(\zeta(t))\right\|_y \\
&\qquad\qquad \leq aL\ell[\zeta,t',t].
\end{aligned}
$$

Definamos $\mathcal{L}(r) = aL$ para todo $t \in [0,R]$ dai a função majorante $f$ fica da forma (3-29). Pela continuidade de $f$ temos que $f(R) \leq 0$, pois a raízes de $f$ são $r^{\pm} = (1 \pm \sqrt{1-2a\beta L})/aL$ e $R < r^{+}$. De $2a\beta L \leq 1$ segue $2a\beta L(2a\beta L - 1) \leq 0$ daí $(1-2a\beta L)^2 \leq 1 - 2a\beta L$ implica $r^{-} \leq 2\beta \leq \frac{1}{aL} \leq r^{+}$. Onde a igualdade é valida se, e somente se $2a\beta L = 1$. neste casso, $r^{-} = r^{+} = 2\beta$ é a única raiz de $f$ em $[0,R]$. Com $f'(r_*) = -1 + aLr_*$, o qual é negativo quando $2a\beta L < 1$. Finalmente, um calculo direito em (3-13) produz $\lambda = a\beta L$. O resultado segue aplicando o Teorema 3.7. □

### 3.2.2 Convergência sob a condição de Nesterov-Nemirovskii

Nesterov e Nemirovskii desenvolveram em [23] uma teoria geral da complexidade computacional do método de ponto-interior para otimização convexa, baseado na noção de funções auto-concordantes. Para uma simplificada presentação desta teoria veja [29]. Alvarez et at. em [3], mostraram que o Teorema 3.7 Produz um resultado de convergência local do método de Newton para a minimização de funções auto-concordantes que é uma ligeira variante de um resultado em [23].

A igual que na Subseção 2.4.3, $\mathbb{R}^n$ é dotado com a estrutura métrica induzida pelo produto interno $\langle\cdot,\cdot\rangle_{H_x}$. Denotaremos por $D_r(x)$ ao elipsoide de Dikin com centro em $x \in \Omega$ e raio $r > 0$. Agora, seguindo [23], assumimos que $f : \Omega \to \mathbb{R}$ é uma função fortemente $a$-auto-concordante não degenerada, isto é

1. $f(y_k) \to \infty$ sempre que $\{y_k\}$ converge a um ponto na fronteira $\partial\Omega$.
2. $H_x := f''(x)$ é definida positiva para todo $x \in \Omega$.
3. $f$ satisfaz (2-23).

Em particular, $f$ é estritamente convexa em $\Omega$ e $x_* \in \Omega$ é o único minimizador de $f$ em $\Omega$ se, e somente se $f'(x_*) = 0$. Por [23, Teorema 2.1.1] (Veja também [29, Teorema 2.5.18]), se $r < 1$ então $D_r(x) \subset \Omega$ e para todo $y \in D_r(x)$ temos

$$\text{Para todo } h \in \mathbb{R}^n \qquad \|h\|_{H_y} \leq \frac{1}{1-r}\|h\|_{H_x}. \tag{3-30}$$

**Lema 3.13** *Seja $x_0 \in \Omega$ fixo. Então, para todo $R \in (0,1)$, $X = f'$ satisfaz (3-1) com $\mathcal{L} : [0,R] \to [0,\infty)$ dada como segue*

$$\mathcal{L}(r) = \frac{2}{(1-r)^3}.$$

**Demonstração.**Pela Observação 3.3 é suficiente mostrar que

$$\left\|f''(x_0)^{-1}f'''(y)\right\|_{H_{x_0}} \leq \mathcal{L}(r), \qquad \text{para } r \in [0,R] \qquad \text{e} \qquad y \in B[x_0,r] = D_r(x_0).$$

Usando a notação $\langle f'''(y)h_1h_2,h_3,\rangle = f'''(y)[h_1,h_2,h_3]$ e pela Definição 1.28 temos

$$\begin{aligned}
\left\|f''(x_0)^{-1}f'''(y)\right\|_{H_{x_0}} &= \sup\left\{\left\|f''(x_0)^{-1}f'''(y)h_1h_2h_3\right\|_{H_{x_0}} : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\} \\
&= \sup\left\{\left\langle f''(x_0)^{-1}f'''(y)h_1h_2,h_3\right\rangle_{H_{x_0}} : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\} \\
&= \sup\left\{a^{-1}f''(x_0)[f''(x_0)^{-1}f'''(y)h_1h_2,h_3] : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\} \\
&= a^{-1}\sup\left\{\left\langle f'''(y)h_1h_2,h_3\right\rangle : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\} \\
&= a^{-1}\sup\left\{f'''(y)[h_1,h_2,h_3] : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\}
\end{aligned}$$

Por outro lado, por (2-23) e por uma propriedade geral das formas simétricas trilineares (veja [23, Proposição 9.1.1]) segue que $f'''(y)[h_1,h_2,h_3] \leq 2a\|h_1\|_{H_y}\|h_2\|_{H_y}\|h_3\|_{H_y}$. (Para uma prova alternativa desse Lema veja [15]) Então

$$\begin{aligned}
\left\|f''(x_0)^{-1}f'''(y)\right\|_{H_{x_0}} &\leq 2\sup\left\{\|h_1\|_{H_y}\|h_2\|_{H_y}\|h_3\|_{H_y} : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\} \\
&\leq 2\sup\left\{\frac{1}{(1-r)^3}\|h_1\|_{H_{x_0}}\|h_2\|_{H_{x_0}}\|h_3\|_{H_{x_0}} : \|h_i\|_{H_{x_0}} = 1, i = 1,2,3\right\} \\
&\leq \frac{2}{(1-r)^3},
\end{aligned}$$

Onde na ultima desigualdade usamos (3-30). Daí o Lema fica demonstrado. □

**Teorema 3.14** *Seja $x_0 \in \Omega$ e $\beta \geq \left\|f''(x_0)^{-1}f'(x_0)\right\|_{H_{x_0}}$.*

*Se $\beta \leq \beta_0 := (\sqrt{2}-1)^2 = 3-2\sqrt{2}$, então $f$ admite um único minimizador $x_*$ o qual pertence a $D_{r_*}(x_0)$ para*

$$r_* = \frac{1}{4}\left[\beta + 1 - \sqrt{\beta^2 - 6\beta + 1}\right]$$

*e, em particular, $\|x_* - x_0\|_{H_{x_0}} \leq r_* \leq 2\beta$. A sequencia gerada pelo método de Newton*

$x_{k+1} = x_k - f''(x_k)^{-1} f'(x_k)$ *com ponto inicial* $x_0$ *esta bem definida, contida em* $D_{r_*}(x_0)$, *e convergente a* $X_*$.

*Para todo* $k \leq 0$, $\|x_* - x_k\|_{H_{x_0}} \leq r_* - r_k$ *onde* $\{r_k\}$ *é a sequencia gerada pelo método de Newton, com ponto inicial em* $r_0 = 0$, *aplicado à função escalar*

$$f(r) = \beta - 2r + \frac{r}{1-r} \tag{3-31}$$

*a sequencia* $\{r_k\}$ *converge para* $r_*$ *que é o menor zero de* $f$ *em* $[0,1)$. *Além disso,*

$$r_k = \frac{1 - \nu^{2^k - 1}}{1 - \nu^{2^k - 1}\eta} r_*, \tag{3-32}$$

*onde*

$$\nu = \frac{1 - \beta - \sqrt{\beta^2 - 6\beta + 1}}{1 - \beta + \sqrt{\beta^2 - 6\beta + 1}}, \qquad \eta = \frac{1 + \beta - \sqrt{\beta^2 - 6\beta + 1}}{1 + \beta + \sqrt{\beta^2 - 6\beta + 1}}.$$

**Demonstração**. Sem perda de generalidade suponhamos $\beta > 0$. Seja $R < 1$ então pelo Lema 3.13, $X = f'$ satisfaz (3-1) com $\mathcal{L}(r) = 2/(1-r)^3$, $r \in [0, R]$. Assim, mediante um calculo direito a função majorante $f$ fica da forma (3-31). o problema de encontar um zero de $f$ é equivalente a resolver $2r^2 - (\beta+1)r + \beta = 0$. Esta equação tem uma raiz reais, e somente se, $\beta^2 - 6\beta + 1 = (\beta - 3 + 2\sqrt{2})(\beta - 3 - 2\sqrt{2}) \geq 0$, que é o caso pois $\beta \leq \beta_0$. As raízes são então $r^{\pm} = [1 + \beta \pm \sqrt{\beta^2 - 6\beta + 1}]/4$. Temos que

$$r^- \leq \frac{1+\beta_0}{4} = 1 - \frac{1}{\sqrt{2}} \leq r^+,$$

pois $\beta_0$ é raiz de $\beta^2 - 6\beta + 1$. Pela continuidade de $f$ temos que $f(R) \leq 0$. é possível escolher $R \in [0, 1 - 1/\sqrt{2}]$ tal que as premissas do Teorema 3.7 valem. Agora,

$$\begin{aligned} f'(r) = -2 + \frac{1}{(1-r)^2} \leq 0 &\Leftrightarrow 2r^2 - 4r + 1 \geq 0 \\ &\Leftrightarrow 2\left[r - \frac{1+\beta_0}{4}\right]\left[r + \frac{-7+\beta_0}{4}\right] \geq 0. \end{aligned}$$

Daí, se $r_* = r^-$, $f'(r_*) < 0$ pois $r_* < (1 - 1/\sqrt{2})$. O resultado segue aplicando o Teorema 3.7. □

### 3.2.3 Convergência sob a condição de Smale

Outro resultado conhecido da convergência local do método de Newton é [30] para funções analíticas, que é uma ferramenta muito útil para a construção e analise de complexidade computacional de algoritmos de homotopia para resolver equações não lineares [4]. Uma extensão do Teorema para variedades Riemanniana foi estabelecido por Dedieu et at. em [5] (Veja também [20]).

A igual que na Subseção 2.4.2 tanto a variedade Riemanniana completa $\mathcal{M}$ e o campo de vetores $X$ serão considerados analíticos. Agora, para todo $p \in \mathcal{M}$, seja $\gamma(p)$ como foi definido em (2-21) e consideremos a seguente formula de Taylor que pode ser encontrada em [5, Teorema 2.1 e Corolário 2.1].

Como $X$ é um campo analítico em $\mathcal{M}$, para todo $p \in \mathcal{M}$ existe uma certa vizinhança $B(p,r)$ de $p$ $(r > 0)$, tal que para todo $q \in B(p,r)$ vale

$$X(\zeta(1)) = P_{\zeta,0,1}\left[\sum_{j=0}^{\infty}\frac{1}{j!}\nabla^j X(p)[u]^j\right]P_{\zeta,1,0},$$

onde $\zeta(1) = \exp_p(u) = q$. Aplicando a $k-$ésima derivada covariante acima obtemos

$$P_{\zeta,1,0}\nabla^k X(\zeta(1)) = \left[\sum_{j=0}^{\infty}\frac{1}{j!}\nabla^{j+k} X(p)[u]^j\right]P_{\zeta,1,0}. \tag{3-33}$$

**Observação 3.15**

$(i)$ *A serie de Taylor em $p \in \mathcal{M}$ para $X$ e $\nabla^k X$ converge na bola centrada em $p$ e raio $\gamma(p)^{-1}$ isto é (3-33) vale para todo $q \in \mathcal{M}$ com $d(p,q) < \min\{r_p, \gamma(p)^{-1}\}$ (Veja [5, Proposição 2.1]).*

$(ii)$ *Para $|r| < 1$ e para todo inteiro $k \geq 0$*

$$\sum_{j=0}\frac{(j+k)!}{j!}r^j = \frac{k!}{(1-r)^{k+1}}$$

**Lema 3.16** *Seja $p_0 \in \mathcal{M}$ tal que $\nabla X(p_0)$ é invertível e $p \in B[p_0, r]$, com $r > 0$.*

$(i)$ *Se $r < \gamma(p_0)^{-1}$ entao, para todo $k \geq 2$,*

$$\left\|\nabla^{-1}X(p_0)[P_{\zeta,1,0}\nabla^k X(p)]\right\|_{p_0} \leq \frac{k!\gamma(p_0)^{k-1}}{(1-\gamma(p_0)r)^{k+1}}, \tag{3-34}$$

*onde $\zeta : [0,1] \to \mathcal{M}$ é uma geodésica minimizante ligando $p_0$ a $p$.*

$(ii)$ *Se $r < (1 - 1/\sqrt{2})\gamma(p_0)^{-1}$ entao $\nabla^{-1}X(p)$ é invertível e*

$$\gamma(p) \leq \frac{1}{(1-\gamma(p_0)r)\psi(\gamma(p_0)r)}\gamma(p_0), \tag{3-35}$$

*para $\psi(\alpha) = 1 - 4\alpha + 2\alpha^2 = (\alpha - 1 + 1/\sqrt{2})(\alpha - 1 - 1/\sqrt{2})$.*

**Demonstração**. $(i)$ Suponhamos que $r < \gamma(p_0)^{-1}$. Seja $\zeta(t) = \exp_{p_0}(tu)$, $t \in [0,1]$, com $u \in B_r(0_{p_0})$ e $p = \exp_{p_0}(u)$. Temos de (3-33) que

$$\begin{aligned}
\left\|\nabla X(p_0)^{-1}[P_{\zeta,1,0}\nabla^k X(p)]\right\| &\leq \sum_{j=0}^{\infty} \frac{1}{j!}\left\|\nabla^{-1} X(p_0)\nabla^{j+k} X(p_0)[u]^j\right\| \\
&\leq \sum_{j=0}^{\infty} \frac{(j+k)!}{j!}\gamma^{j+k-1}\|u\|^j \\
&\leq \gamma(p_0)^{k-1}\sum_{j=0}^{\infty}\frac{(j+k)!}{j!}(\gamma(p_0)r)^j \\
&= \frac{k!\gamma(p_0)^{k-1}}{(1-\gamma(p_0)r)^{k+1}}
\end{aligned}$$

como $\gamma(p_0)r < 1$ a serie na última desigualdade é convergente. Veja Observação 3.15 $(ii)$.

$(ii)$ Agora, se $r < (1-1/\sqrt{2})\gamma(p_0)^{-1}$. Seja $I : T_{p_0}\mathcal{M} \to T_{p_0}\mathcal{M}$ o operador identidade em $T_{p_0}\mathcal{M}$, então por (2-21) e (3-33)

$$\begin{aligned}
\left\|\nabla X(p_0)^{-1}[P_{\zeta,1,0}\nabla X(p)] - I\right\| &= \left\|\nabla X(p_0)^{-1}\left[\sum_{j=0}^{\infty}\frac{1}{j!}\nabla^{j+1}X(p)[u]^j\right]P_{\zeta,1,0} - I\right\| \\
&= \left\|\sum_{j=1}^{\infty}\nabla X(p_0)^{-1}\frac{1}{j!}\nabla^{j+1}X(p)[u]^j\right\| \\
&\leq \sum_{j=1}^{\infty}\left\|\nabla X(p_0)^{-1}\frac{1}{j!}\nabla^{j+1}X(p)[u]^j\right\| \\
&\leq \sum_{j=0}^{\infty}\left\|\nabla X(p_0)^{-1}\frac{1}{j!}\nabla^{j+1}X(p)\right\| r^j - 1 \\
&\leq \sum_{j=0}^{\infty}(j+1)(\gamma(p_0)r)^j - 1 \\
&= \frac{1}{(1-\gamma(p_0)r)^2} - 1.
\end{aligned}$$

A última igualdade acima se tem pela Observação 3.15 $(ii)$ com $k = 1$. Como $\gamma(p_0)\,r < 1 - 1/\sqrt{2}$, temos que $1/(1-\gamma(p_0)r)^2 - 1 < 1$. Assim pelo Lema de Banach 1.32 deduzimos que $\nabla X(p_0)^{-1}P_{\zeta,1,0}\nabla X(p)$ é invertível e portanto $\nabla X(p)$ também, além disso

$$\begin{aligned}\left\|[\nabla X(p_0)^{-1}P_{\zeta,1,0}\nabla X(p)]^{-1}\right\| &= \left\|\nabla X(p)^{-1}P_{\zeta,0,1}\nabla X(p_0)\right\| \\ &\le \frac{(1-\gamma(p_0)r)^2}{2(1-\gamma(p_0)r)^2-1} \\ &= \frac{(1-\gamma(p_0)r)^2}{\psi(\gamma(p_0)r)}.\end{aligned}$$

Portanto,

$$\begin{aligned}\left\|\nabla X(p)^{-1}\nabla^k X(p)\right\| &= \left\|\nabla X(p)^{-1}P_{\zeta,0,1}\nabla X(p_0)\nabla X(p_0)^{-1}P_{\zeta,1,0}\nabla^k X(p)\right\| \\ &\le \left\|\nabla X(p)^{-1}P_{\zeta,0,1}\nabla X(p_0)\right\| \left\|\nabla X(p_0)^{-1}P_{\zeta,1,0}\nabla^k X(p)\right\| \\ &\le \frac{k!\gamma(p_0)^{k-1}}{(1-\gamma(p_0)r)^{k-1}\psi(\gamma(p_0)r)}.\end{aligned}$$

Como $\psi(0) = 1$ e $\psi(1-\sqrt{2}) = 0$, temos pela continuidade de $\psi$ que $0 < \psi(\alpha) \le 1$ quando $0 \le \alpha < 1 - 1/\sqrt{2}$. Assim

$$\gamma(p) \le \sup_{k\ge 2}\frac{\gamma(p_0)}{(1-\gamma(p_0)r)\psi(\gamma(p_0)r)^{1/k-1}} = \frac{\gamma(p_0)}{(1-\gamma(p_0)r)\psi(\gamma(p_0)r)},$$

o que prova (3-35). □

**Lema 3.17** *Seja $p_0 \in \mathcal{M}$ tal que $\nabla X(p_0)$ é invertível. Para todo $R \le (1-1/\sqrt{2})\gamma(p_0)^{-1}$. Então $X$ satisfaz (3-1) com*

$$\mathcal{L}(r) = \frac{2\gamma(p_0)}{(1-\gamma(p_0)r)^3}, \qquad r \in [0,R].$$

**Demonstração**. Pela Observação 3.3 é suficiente mostrar que para todo $r \in [0,R]$ e $\zeta \in \mathcal{G}_2(p_0,r)$ temos

$$\left\|\nabla X(p_0)^{-1}[P_{\zeta,a,0}\nabla^2 X(\zeta(a))]\right\| \le \mathcal{L}(r), \qquad \text{Para todo } a \in [0,2]. \tag{3-36}$$

Pela continuidade é suficiente estabelecer a propriedade para $r \in [0,R)$.

Fixando $\zeta \in \mathcal{G}_2(p_0,r)$ com $r < R$. Pela definição de $\mathcal{G}_2(p_0,r)$, $\zeta$ é a concatenação de dois geodésicas $\zeta_i : [0,1] \to \mathcal{M}$ $(i = 0,1)$ com $\zeta_0$ minimizante. Assim escreveremos $\zeta(t) = \zeta_0(t)$ se $t \in [0,1]$ e $\zeta(t) = \zeta_1(t-1)$ se $t \in [1,2]$. Como $\zeta(0) = p_0$, aplicando o Lema 3.16 $(i)$ a $\zeta_0$ com $k = 2$, obtemos que vale (3-36) para todo $a \in [0,1]$.

Para provar (3-36) para $a \in (1,2]$, se procede de la seguinte forma. Primeiro, vamos a considerar $a \in (1,2]$ tal que $d(\zeta(1),\zeta(a)) < \gamma(p_0)^{-1}$. Como $\zeta|_{[1,2]}$ é geodésia, existe $u_1 \in B_{\gamma(\zeta(1))^{-1}}(\zeta(1))$, com $\|u_1\|_{\zeta(1)} = d(\zeta(1),\zeta(a))$ tal que $\zeta(a) = \exp_{\zeta(1)} u_1$. Pela formula de Taylor (3-33), temos $P_{\zeta,a,1}\nabla^2 X(\zeta(a)) = \left[\sum_{j=0}^{\infty} \frac{1}{j!}\nabla^{j+2}X(\zeta(1))[u_1]^j\right] P_{\zeta,a,1}$, então

$$P_{\zeta,a,0}\nabla^2 X(\zeta(a)) = \left[\sum_{j=0}^{\infty} \frac{1}{j!} P_{\zeta,1,0}\left[\nabla^{j+2}X(\zeta(1))[u_1]^j\right]\right] P_{\zeta,a,1}.$$

Agora, como $\zeta|_{[0,1]}$ é minimizante, existe $u_0 \in T_{p_0}\mathcal{M}$, com $\|u_0\|_{p_0} = d(p_0,\zeta(1)) \leq r < \gamma(p_0)^{-1}$, tal que $\zeta(1) = \exp_{p_0} u_0$. Assim, podemos utilizar (3-33) de novo para obter

$$\begin{aligned}
P_{\zeta,a,0}\nabla^2 X(\zeta(a)) &= \left[\sum_{j_1=0}^{\infty} \frac{1}{j_1!}\left[\sum_{j_0=0}^{\infty} \frac{1}{j_0!}\nabla^{j_0+j_1+2}X(p_0)[u_0]^{j_0}\right] P_{\zeta,1,0}[u_1]^{j_1}\right] P_{\zeta,a,1} \\
&= \left[\sum_{j_0,j_1=0}^{\infty} \frac{1}{j_0!j_1!}\nabla^{j_0+j_1+2}X(p_0)([u_0]^{j_0}, P_{\zeta,1,0}[u_1]^{j_1})\right] P_{\zeta,a,1}.
\end{aligned}$$

Portanto,

$$\begin{aligned}
&\left\|\nabla X(p_0)^{-1}\left[P_{\zeta,a,0}\nabla^2 X(\zeta(a))\right]\right\|_{p_0} \\
&\qquad \leq \sum_{j_0,j_1=0}^{\infty} \frac{1}{j_0!j_1!}\left\|\nabla X(p_0)^{-1}\nabla^{j_0+j_1+2}X(p_0)\right\|_{p_0} \|u_0\|_{p_0}^{j_0} \left\|P_{\zeta,1,0}[u_1]\right\|_{p_0}^{j_1} \\
&\qquad \leq \sum_{j_0,j_1=0}^{\infty} \frac{1}{j_0!j_1!}(j_0+j_1+2)!\gamma(p_0)^{j_0+j_1+1}\|u_0\|_{p_0}^{j_0}\|u_1\|_{p_1}^{j_1}
\end{aligned}$$

Se $j = j_0 + j_1$ temos $0 \leq j_1 = j - j_0$, então $j_0 \leq j$ dai, se $j_0 = i$ por (2-21) e (3-33) segue

$$\begin{aligned}
\left\|\nabla X(p_0)^{-1}\left[P_{\zeta,a,0}\nabla^2 X(\zeta(a))\right]\right\|_{p_0} &\leq \sum_{j=0}^{\infty}\left[\sum_{i=0}^{j}\frac{(j+2)!}{i!(j-i)!}\|u_0\|_{p_0}^{i}\|u_1\|_{p_1}^{j-i}\right]\gamma(p_0)^{j+1} \\
&= \sum_{j=0}^{\infty}(j+2)(j+1)\left(\|u_0\|_{p_0}+\|u_1\|_{p_1}\right)^j\gamma(p_0)^{j+1} \\
&= \gamma(p_0)\sum_{j=0}^{\infty}(j+2)(j+1)(\gamma(p_0)\,r)^j \\
&= \frac{2\gamma(p_0)}{(1-\gamma(p_0)\,r)^3}
\end{aligned}$$

Onde usamos o fato que $\|u_0\|_{p_0} + \|u_1\|_{p_1} = d(p_0,\zeta(1)) + d(\zeta(1),\zeta(a)) < r$ e na última igualdade acima a Observação 3.15 $(ii)$ com $k = 2$.

O resultado anterior mostra que (3-36) vale para todo $a \in (1,2]$ com $d(\zeta(1),\zeta(a)) < \gamma(\zeta(1))^{-1}$. Se $d(\zeta(1),\zeta(a)) \geq \gamma(\zeta(1))^{-1}$ então não é possível garantir que a expansão de Taylor em $\zeta(1)$ vale. No entanto, como $r < R \leq (1-1/\sqrt{2})\gamma(p_0)^{-1}$, por (3-35), existe uma constante positiva $K$ dependendo de $r$ e $\gamma(p_0)$ tal que $\gamma(p) \leq K$ para todo $p \in B[p_0, r]$. Assim, para $n$ suficientemente grande, existe uma partição $1 = t_1 < \cdots < t_{n+1} = a$ de $[1,a]$ e correspondentemente vectores tangentes $u_i \in T_{\zeta(t_i)}\mathcal{M}$ tal que $\zeta(t_{i+1}) = \exp_{\zeta(t_i)}[u_i]$ com $\|u_i\|_{\zeta(t_i)} = d(\zeta(t_{i+1}), \zeta(t_i)) \leq K^{-1} \leq \gamma(c(t_i))^{-1}$ para $i = 1, \cdots, n$. Neste caso, mediante aplicações sucessivas de transportes paralelos e formulas de Taylor obtemos

$$P_{\zeta,a,0}\nabla^2 X(\zeta(a)) = \sum_{j_0,\cdots,j_n=0}^{\infty} \frac{1}{j_0!\cdots j_n!}\nabla^{j_0+\cdots+j_n+2}X(p_0)[u_0^{j_0}, P_{\zeta,t_1,0}u_1^{j_1}, \cdots, P_{\zeta,t_n,0}u_n^{j_n}].$$

Então como $\|u_0\|_{p_0} + \sum_{i=1}^{n}\|u_i\|_{c(t_i)} \leq r$, é posivel mostrar mediante argumentos similares que vale (3-36). $\square$

**Teorema 3.18** *Seja $p_0 \in \mathcal{M}$ tal que $\nabla X(p_0)$ é invertível e seja $\alpha = \beta\gamma$ para $\beta \leq \left\|\nabla X(p_0)^{-1}X(p_0)\right\|_{p_0}$ e $\gamma = \gamma(p_0)$.*

*Se $\alpha \leq \alpha_0 := (\sqrt{2}-1)^2 = 3-2\sqrt{2}$ então a sequencia $\{p_k\}$ gerada pelo método de Newton iniciando no ponto $p_0$ esta bem definida, contida em $B[p_0, r_*]$ onde*

$$r_* = \frac{1}{4\gamma}[1+\alpha-\sqrt{\alpha^2-6\alpha+1}],$$

*e convergente para $p_*$, o qual é a única singularidade de $X$ em $B[p_0, (1-\sqrt{2})\gamma^{-1}]$. Em particular, $d(p_*, p_0) \leq r_* \leq 2\beta$.*

*Para todo $k \leq 0$, $d(p_*, p_k) \leq r_* - r_k$ onde $\{r_k\}$ é a sequencia gerada pelo método de Newton, iniciando em $r_0 = 0$, aplicado à função escalar*

$$f(r) = \beta - 2r + \frac{r}{1-\gamma r}. \tag{3-37}$$

*A sequencia $\{r_k\}$ converge a $r_*$ que é menor zero de $f$ em $[0, \gamma^{-1})$. Além disso, $r_k$ tem forma fechada*

$$r_k = \frac{1-\nu^{2^k-1}}{1-\nu^{2^k-1}\eta}r_*,$$

*para*

$$\nu = \frac{1-\alpha-\sqrt{\alpha^2-6\alpha+1}}{1-\alpha+\sqrt{\alpha^2-6\alpha+1}}, \qquad \eta = \frac{1+\alpha-\sqrt{\alpha^2-6\alpha+1}}{1+\alpha+\sqrt{\alpha^2-6\alpha+1}}.$$

**Demonstração**. Sem perda de generalidade suponhamos $\beta > 0$. Se $R \leq (1-1/\sqrt{2})\gamma^{-1}$ temos pelo Lema 3.17 que $X$ satisfaz (3-1) com $\mathcal{L}(r) = 2\gamma/(1-\gamma r)^3$, $r \in [0,R]$, assim, a função majorante $f$ fica da forma (3-37), cujas raízes são $r^{\pm} = [1+\alpha\pm\sqrt{\alpha^2-6\alpha+1}]/4\gamma$. De fato, o problema é equivalente a resolver $2\gamma r^2 - r(1+\beta\gamma)+\beta = 0$. Esta equação tem raiz real se, e somente se, $\alpha^2-6\alpha+1 = (\alpha-3+2\sqrt{2})(\alpha-3-2\sqrt{2}) \leq 0$. O qual é o caso, pois $\alpha \leq \alpha_0$. Temos que

$$r^- \leq \frac{1+\alpha_0}{4\gamma} = (1-1/\sqrt{2})\gamma^{-1} \leq r^+,$$

pois $\alpha_0$ é raiz de $\alpha^2 - 6\alpha + 1$. Pela continuidade de $f$ temos que $f(R) \leq 0$. Agora,

$$\begin{aligned} f'(r) = -2 + \frac{1}{(1-\gamma r)^2} \leq 0 &\iff 2\gamma^2 r^2 - 4\gamma r + 1 \geq 0 \\ &\iff 2\gamma^2 \left[r - \frac{1+\alpha_0}{4\gamma}\right]\left[r + \frac{-7+\alpha_0}{4\gamma}\right] \geq 0. \end{aligned}$$

Daí, se $r_* = r^{-1}$ $f'(r_*) < 0$ pois $r_* < (1 - 1\sqrt{2})\gamma^{-1}$ o resultado segue aplicando o Teorema 3.7. $\square$

Capítulo 4

# Considerações finais

Nesta dissertação estudamos análises local e semi-local do método de Newton sob a condição majorante. Como foi mostrado esta duas análises tem objetivos bem distintos. Na análise local supondo que a derivada covariante na solução, que em geral não é conhecida, é não singular, é mostrado que existe uma vizinhança desta solução tal que, se o ponto inicial é tomado nesta vizinhança o método está bem definido e converge. Enquanto a análise semi-local a hipótese é feita sob o ponto inicial e, neste caso, é mostrado existência de solução, boa-definição do método e convergência. Em ambos os caso, a codição majorante é usada para unificar alguns resultados sobre o método de Newton que aparentemente não tinham relação prévia evidente. Seria interessante estender o estudo feito aqui neste trabalho a situações mais gerais, por exemplo, inclusões do tipo

$$X(p)+Y(p)\ni 0$$

onde $X$ é um campo continuamente diferenciável e $Y$ um campo ponto-conjunto.

# Referências Bibliográficas


[1] ABSIL, P.-A.; MAHONY, R.; SEPULCHRE, R. **Optimization algorithms on matrix manifolds**. Princeton University Press, Princeton, NJ, 2008. With a foreword by Paul Van Dooren.

[2] ADLER, R. L.; DEDIEU, J.-P.; MARGULIES, J. Y.; MARTENS, M.; SHUB, M. **Newton's method on Riemannian manifolds and a geometric model for the human spine**. *IMA J. Numer. Anal.*, 22(3):359–390, 2002.

[3] ALVAREZ, F.; BOLTE, J.; MUNIER, J. **A unifying local convergence result for Newton's method in Riemannian manifolds**. *Found. Comput. Math.*, 8(2):197–226, 2008.

[4] BLUM, L.; CUCKER, F.; SHUB, M.; SMALE, S. **Complexity and real computation**. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.

[5] DEDIEU, J. P.; PRIOURET, P.; MALAJOVICH, G. **Newton's method on Riemannian manifolds: convariant alpha theory**. *IMA J. Numer. Anal.*, 23(3):395–419, 2003.

[6] DENNIS, JR., J. E.; SCHNABEL, R. B. **Numerical methods for unconstrained optimization and nonlinear equations**, volume 16 de **Classics in Applied Mathematics**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. Corrected reprint of the 1983 original.

[7] DEUFLHARD, P. **Newton methods for nonlinear problems**, volume 35 de **Springer Series in Computational Mathematics**. Springer, Heidelberg, 2011. Affine invariance and adaptive algorithms, First softcover printing of the 2006 corrected printing.

[8] DO CARMO, M. P. **Riemannian geometry**. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[9] FERREIRA, O. P.; SVAITER, B. F. **Kantorovich's theorem on Newton's method in Riemannian manifolds**. *J. Complexity*, 18(1):304–329, 2002.

[10] FERREIRA, O. P.; SVAITER, B. F. **Kantorovich's majorants principle for Newton's method**. *Comput. Optim. Appl.*, 42(2):213–229, 2009.

[11] FERREIRA, O. P. **Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle**. *IMA J. Numer. Anal.*, 29(3):746–759, 2009.

[12] FERREIRA, O. P.; SILVA, R. C. M. **Local convergence of Newton's method under a majorant condition in Riemannian manifolds**. *IMA J. Numer. Anal.*, 32(4):1696–1713, 2012.

[13] GRAGG, W. B.; TAPIA, R. A. **Optimal error bounds for the Newton-Kantorovich theorem**. *SIAM J. Numer. Anal.*, 11:10–13, 1974.

[14] HUANG, Z. **The convergence ball of Newton's method and the uniqueness ball of equations under Hölder-type continuous derivatives**. *Comput. Math. Appl.*, 47(2-3):247–251, 2004.

[15] JARRE, F. **Interior-point methods for convex programming**. *Appl. Math. Optim.*, 26(3):287–311, 1992.

[16] JIANG, D.; MOORE, J. B.; JI, H. **Self-concordant functions for optimization on smooth manifolds**. *J. Global Optim.*, 38(3):437–457, 2007.

[17] KANTOROVICH, L. V.; AKILOV, G. P. **Functional analysis in normed spaces**. Translated from the Russian by D. E. Brown. Edited by A. P. Robertson. International Series of Monographs in Pure and Applied Mathematics, Vol. 46. The Macmillan Co., New York, 1964.

[18] KRANTZ, S. G.; PARKS, H. R. **The implicit function theorem**. Birkhäuser Boston, Inc., Boston, MA, 2002. History, theory, and applications.

[19] LANG, S. **Differential and Riemannian manifolds**, volume 160 de **Graduate Texts in Mathematics**. Springer-Verlag, New York, third edition, 1995.

[20] LI, C.; WANG, J. **Newton's method on Riemannian manifolds: Smale's point estimate theory under the $\gamma$-condition**. *IMA J. Numer. Anal.*, 26(2):228–251, 2006.

[21] MOSER, J. **A new technique for the construction of solutions of nonlinear differential equations**. *Proc. Nat. Acad. Sci. U.S.A.*, 47:1824–1831, 1961.

[22] NASH, J. **The imbedding problem for Riemannian manifolds**. *Ann. of Math. (2)*, 63:20–63, 1956.

[23] NESTEROV, Y.; NEMIROVSKII, A. **Interior-point polynomial algorithms in convex programming**, volume 13 de **SIAM Studies in Applied Mathematics**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[24] ORTEGA, J. M.; RHEINBOLDT, W. C. **Iterative solution of nonlinear equations in several variables**, volume 30 de **Classics in Applied Mathematics**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.

[25] ORTEGA, J. M. **Numerical analysis. A second course**. Academic Press, New York-London, 1972. Computer Science and Applied Mathematics.

[26] POLYAK, B. T. **Newton's method and its use in optimization**. *European J. Oper. Res.*, 181(3):1086–1096, 2007.

[27] PTÁK, V. **The rate of convergence of Newton's process**. *Numer. Math.*, 25(3):279–285, 1975/76.

[28] RALL, L. B. **A note on the convergence of Newton's method**. *SIAM J. Numer. Anal.*, 11:34–36, 1974.

[29] RENEGAR, J. **A mathematical view of interior-point methods in convex optimization**. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2001.

[30] SMALE, S. **Newton's method estimates from data at one point**. In: *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, p. 185–196. Springer, New York, 1986.

[31] SMITH, S. T. **Optimization techniques on Riemannian manifolds**. In: *Hamiltonian and gradient flows, algorithms and control*, volume 3 de **Fields Inst. Commun.**, p. 113–136. Amer. Math. Soc., Providence, RI, 1994.

[32] TRAUB, J. F.; WOŹNIAKOWSKI, H. **Convergence and complexity of Newton iteration for operator equations**. *J. Assoc. Comput. Mach.*, 26(2):250–258, 1979.

[33] UDRI¸STE, C. **Convex functions and optimization methods on Riemannian manifolds**, volume 297 de **Mathematics and its Applications**. Kluwer Academic Publishers Group, Dordrecht, 1994.

[34] WANG, J.-H.; LI, C. **Uniqueness of the singular points of vector fields on Riemannian manifolds under the $\gamma$-condition**. *J. Complexity*, 22(4):533–548, 2006.

[35] WANG, X. H.; LI, C. **Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces. II**. *Acta Math. Sin. (Engl. Ser.)*, 19(2):405–412, 2003.

[36] WAYNE, C. E. **An introduction to KAM theory**. In: *Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994)*, volume 31 de **Lectures in Appl. Math.**, p. 3–29. Amer. Math. Soc., Providence, RI, 1996.

# Índice Remissivo