Iniciação Científica

Relatório Final

Aprendizado de máquina e geometria hiperbólica complexa

Universidade de São Paulo

Instituto de Ciências Matemáticas e Computação

Aluno:	**Lucas Giraldi Almeida Coimbra**
Orientador:	**Carlos Henrique Grossi Ferreira**

Junho de 2023

São Carlos

Conteúdo

1 Introdução

2 Geometria hiperbólica

2.1 Variedades e métricas riemannianas

Uma *variedade topológica de dimensão* n é um espaço topológico M Hausdorff com base enumerável que é *localmente euclidiano de dimensão* n, isso é, para cada $p \in M$ existe um aberto U e um homeomorfismo $\phi\colon U \to V \subset\circ \mathbb{R}^n$. O par (U, ϕ) será comumente chamado de *carta sobre* p. Se (V, ψ) é uma outra carta em M tal que $U \cap V \neq \varnothing$, chamamos de *mapas de transição* as funções

$$\phi \circ \psi^{-1}\colon \psi(U \cap V) \to \mathbb{R}^n \quad \text{e} \quad \psi \circ \phi^{-1}\colon \phi(U \cap V) \to \mathbb{R}^n. \tag{1}$$

Se os mapas de transição forem suaves, diremos que (U, ϕ) e (V, ψ) são *compatíveis*. Uma *estrutura diferenciável* em M é uma cobertura de M por cartas que são duas a duas compatíveis. Dizemos que M é *suave* ou *diferenciável* se possuir uma estrutura diferenciável.

A partir de agora, toda carta estará em uma estrutura diferenciável previamente fixada, e portanto toda variedade será suave. Se $p \in M$ dizemos que $F\colon M \to N$ é *suave em* p se existirem (U, ϕ) carta sobre p e (V, ψ) carta sobre $F(p)$ tais que $\psi \circ F \circ \phi^{-1}$ é suave. A função F é *suave em* $U \subset\circ M$ se for suave em todo ponto de U, e é apenas *suave* se for suave em todo ponto de M.

Uma *curva* em M é um mapa suave $c\colon I \to M$ onde I é um intervalo de $\mathbb{R}$. Se $p \in M$, definimos por C_p^∞ como o conjunto dos mapas $f\colon U \subset\circ M \to \mathbb{R}$ suaves, onde U é uma vizinhança qualquer de p. Esse espaço é uma álgebra com as três operações:

- se $f\colon U \to \mathbb{R}$ e $g\colon V \to \mathbb{R}$, definimos $f + g\colon U \cap V \to \mathbb{R}$ por $(f + g)(p) = f(p) + g(p)$;
- se $f\colon U \to \mathbb{R}$ e $\lambda \in \mathbb{R}$, definimos $\lambda f\colon U \to \mathbb{R}$ por $(\lambda f)(p) = \lambda f(p)$;
- se $f\colon U \to \mathbb{R}$ e $g\colon V \to \mathbb{R}$, definimos $fg\colon U \cap V \to \mathbb{R}$ por $(fg)(p) = f(p)g(p)$.

Dada uma curva $c\colon]-\varepsilon, \varepsilon[\to M$, definimos $c'(0)$ como sendo um mapa $c'(0)\colon C_p^\infty \to \mathbb{R}$ dado por

$$c'(0)f = \left.\frac{d}{dt}\right|_{t=0} (f \circ c)(t). \tag{2}$$

Esse mapa é linear e satisfaz a *regra de Leibniz*, isso é,

$$c'(0)(fg) = f(c(0)) \cdot c'(0)g + c'(0)f \cdot g(c(0)). \tag{3}$$

Se $p \in M$, o *espaço tangente a* M *em* $p \in M$ como o conjunto

$$T_pM = \{c'(0) \mid c\colon]-\varepsilon, \varepsilon[\to \mathbb{R} \text{ e } c(0) = p\}. \tag{4}$$

Se M tem dimensão n, então T_pM é um espaço vetorial de dimensão n. Seus elementos são chamados de *vetores tangentes*. Uma *métrica riemanniana* em M é a associação de um produto interno $\mathfrak{g}_p(-,-)$ em T_pM para cada $p \in M$. Mais do que isso, pedimos que essa associação seja suave. Entenderemos o que isso significa a seguir.

Um *campo vetorial* em M é uma associação X de um vetor $X_p \in T_pM$ para cada $p \in M$. Se $\phi = (x^1, \dots, x^n)$ é uma carta sobre $p \in M$ e $r = (r^1, \dots, r^n)$ são as coordenadas em $\mathbb{R}^n$, definimos as derivadas parciais de $f \in C_p^\infty$ por

$$\left.\frac{\partial f}{\partial x^i}\right|_p = \left.\frac{\partial}{\partial r^i}\right|_{\phi(p)} (f \circ \phi^{-1})(r). \tag{5}$$

Cada derivada parcial em p pode ser vista como um elemento de T_pM, afinal, se $e^1, \dots, e^n$ é a base canônica de $\mathbb{R}^n$, então dadas as curvas $c^i(t) = te^i$ temos

$$\left.\frac{\partial}{\partial x^i}\right|_p = (\phi^{-1} \circ c^i)'(0). \tag{6}$$

Esses vetores tangentes formam uma base para T_pM.

Se (U, ϕ) é uma em M e X é um campo vetorial em M, então para cada $p \in M$ podemos escrever, de maneira única,

$$X_p = \sum_{k=1}^{n} a^i(p) \left. \frac{\partial}{\partial x^i} \right|_p . \tag{7}$$

Dizemos que o campo vetorial X é *suave* se existir uma cobertura de M por cartas tais que os mapas a^i são sempre suaves. Ao dizermos que a métrica riemanniana tem que ser suave, queremos dizer que, para quaisquer X, Y campos suaves em M, o mapa $p \mapsto \mathfrak{g}_p(X_p, Y_p)$ tem que ser suave. Uma *variedade riemanniana* é uma variedade suave equipada com uma métrica riemanniana.

2.2 Conexões e derivada covariante

Denotamos o conjunto de todos os campos suaves em M por $\mathfrak{X}(M)$. Se $M = \mathbb{R}^n$, vamos entender quem é a derivada direcional. Se $X = (v^1, \dots, v^n) \in \mathbb{R}^n$ e X_p é o vetor tangente a p na direção X, então dada $f\colon \mathbb{R}^n \to \mathbb{R}$ definimos a *derivada direcional de f na direção X_p*

$$D_{X_p} f = \lim_{t \to 0} \frac{f(p + tX) - f(p)}{t} = \sum_{k=1}^{n} v^k \left. \frac{\partial f}{\partial x^i} \right|_p = X_p f. \tag{8}$$

Podemos então trocar f por um campo vetorial suave $Y = \sum b^i \partial/\partial x^i$ e obtermos a *derivada direcional de Y na direção X_p*

$$D_{X_p} Y = \sum_{k=1}^{n} D_{X_p} b^i \left. \frac{\partial}{\partial x^i} \right|_p . \tag{9}$$

Note que a derivada $D_{X_p} Y$ é um vetor tangente em p. Dessa forma, se X é um campo vetorial em $\mathbb{R}^n$ podemos definir $D_X Y$ como o campo vetorial que, em p, vale $D_{X_p} Y$. Esse mapa é a *derivada direcional de Y na direção X*.

Agora vamos generalizar a derivada direcional em $\mathbb{R}^n$ para uma variedade riemanniana qualquer. Uma *conexão afim* em M é um mapa

$$\begin{aligned} \nabla\colon \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathfrak{X}(M) \\ (X, Y) &\mapsto \nabla_X Y \end{aligned}$$

que satisfaz as seguintes propriedades:

- se $C^\infty(M)$ é o conjunto dos mapas suaves $M \to \mathbb{R}$, então ∇ é $C^\infty(M)$-linear na primeira coordenada;
- ∇ satisfaz a regra de Leibniz na segunda coordenada, isso é, se $f \in C^\infty(M)$, então

  $$\nabla_X(fY) = (Xf)Y + f\nabla_X Y, \tag{10}$$

  onde Xf é o mapa suave dado por $(Xf)(p) = X_p f$.

Conexões e métricas riemannianas não estão sempre conectadas. Porém, se M é uma variedade riemanniana e ∇ uma conexão afim em M, então podemos falar sobre alguns aspectos geométricos de ∇. Definimos o *tensor torção* de ∇ como sendo o mapa $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$, onde $[X, Y]_p f = X_p(Y f) - Y_p(X f)$ é o *bracket de Lie*. Do mesmo modo, definimos o *tensor curvatura* de ∇ como sendo o mapa $R(X, Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}$, isso é, para um campo vetorial suave Z, temos

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{11}$$

Dizemos que uma conexão ∇ em uma variedade riemanniana M é *compatível com a métrica* se $Z\mathfrak{g}(X, Y) = \mathfrak{g}(\nabla_Z X, Y) + \mathfrak{g}(X, \nabla_Z Y)$. Uma *conexão de Levi-Civita* é uma conexão compatível com a métrica e que satisfaz $T(X, Y) = 0$ para todos X, Y campos suaves em M.

Proposição 2.1. *Toda variedade riemanniana possui uma, e apenas uma, conexão de Levi-Civita.*

Um campo vetorial *ao longo* de uma curva $c\colon I \to M$ é a associação V de um vetor $V(t) \in T_{c(t)}M$ para cada $t \in I$. Dizemos que V é *suave* se, para cada $f\colon M \to \mathbb{R}$, $(Vf)(t) = V(t)f$ é suave.

Se $c\colon I \to \mathbb{R}^n$ é uma curva e V é um campo ao longo de c, temos

$$V(t) = \sum_{k=1}^{n} v^i(t) \left.\frac{\partial}{\partial x^i}\right|_{c(t)}, \tag{12}$$

portanto podemos definir a *derivada de V com respeito a t* como sendo o campo

$$\frac{dV}{dt} = \sum_{k=1}^{n} \frac{dv^i}{dt}\frac{\partial}{\partial x^i}. \tag{13}$$

Essa derivada satisfaz algumas propriedades importantes:

- ela é linear com respeito a V, isso é, se $\lambda \in \mathbb{R}$ e U é outro campo ao longo de c, então

$$\frac{d(\lambda V + U)}{dt} = \lambda\frac{dV}{dt} + \frac{dU}{dt}; \tag{14}$$

- ela satisfaz a regra de Leibniz, isso é, se $f\colon I \to \mathbb{R}$ (lembrando aqui que I é o domínio de c) é suave, então

$$\frac{d(fV)}{dt} = \frac{df}{dt}V + f\frac{dV}{dt}; \tag{15}$$

- ela é compatível com a derivada direcional em $\mathbb{R}^n$, isso é, se V se estende para um campo $\tilde{V}$ em $\mathbb{R}^n$, então

$$\frac{dV}{dt} = D_{c'(t)}\tilde{V}. \tag{16}$$

Vamos agora generalizar o conceito da derivada de V para uma variedade M qualquer, utilizando de conexões afins. Se ∇ é uma conexão afim em M e $c\colon I \to \mathbb{R}$ é uma curva, então definimos uma *derivada covariante* como um operador D/dt que, para cada campo V ao londo de c associa um outro campo DV/dt ao longo de c. Pedimos que essa associação satisfaça as três propriedades que a derivada definida acima satisfaz:

- D/dt é linear, isso é, se V e U são campos ao longo de c e $\lambda \in \mathbb{R}$ então

$$\frac{D(\lambda V + U)}{dt} = \lambda\frac{DV}{dt} + \frac{DU}{dt}; \tag{17}$$

- D/dt satisfaz a regra de Leibniz, isso é, se $f\colon I \to \mathbb{R}$ é suave, então

$$\frac{D(fV)}{dt} = \frac{df}{dt}V + f\frac{DV}{dt}; \tag{18}$$

- D/dt é compatível com a conexão afim, isso é: se $\tilde{V}$ é um campo em M que estende V, então

$$\frac{DV}{dt} = \nabla_{c'(t)}V. \tag{19}$$

Definimos acima o que seria **uma** derivada covariante, mas acontece que, fixadas uma conexão e uma curva, sempre existe uma e apenas uma derivada covariante, portanto podemos falar **da** derivada covariante.

2.3 Geodésicas e transporte paralelo

Se $c\colon I \to M$ é uma curva, então dizemos que c é uma *geodésica* se a derivada covariante DT/dt do seu campo velocidade $T(t) = c'(t)$ é nula. Note que a existência de uma conexão, e portanto de uma derivada covariante, não depende da existência de uma métrica riemanniana. Porém, caso a variedade M possua uma métrica, vamos sempre assumir que a conexão considerada é a conexão de Levi-Civita em M.

Proposição 2.2. *Geodésicas em variedades riemannianas possuem velocidade constante, isso é, se $c\colon I \to M$ é uma geodésica, então $||c'(t)||$ é constante para cada $t \in I$.*

Seja M uma variedade suave com uma conexão ∇. Se $(U, x^1, \ldots, x^n)$ é uma carta em M, então temos os campos vetoriais $\partial_i = \partial/\partial x^i$. Sabemos que todo campo vetorial em U se escreve como combinação linear destes, e portanto temos

$$\nabla_{\partial_i}\partial_j = \sum_{k=1}^{n} \Gamma_{ij}^{k}\partial_k. \tag{20}$$

Os coeficientes Γ_{ij}^k são chamados de *símbolos de Christoffel de* ∇ *em* $(U, x^1, \ldots, x^n)$.

Sejam M uma variedade com uma conexão ∇, $(U, \phi) = (U, x^1, \ldots, x^n)$ uma carta em M e Γ_{ij}^k os seus símbolos de Christoffel. Note que, se $c\colon I \to M$ é uma curva e $y = \phi \circ c$, então temos

$$T = c'(t) = \sum_{k=1}^{n} \frac{dy^k}{dt}\partial_k. \tag{21}$$

Dessa maneira, segue que

$$\frac{DT}{dt} = \sum_{j=1}^{n} \frac{d^2y^j}{dt^2}\partial_j + \sum_{j=1}^{n} \frac{dy^j}{dt}\frac{D\partial_j}{dt} = \sum_{j=1}^{n} \frac{d^2y^j}{dt^2}\partial_j + \sum_{j=1}^{n} \frac{dy^j}{dt}\nabla_{c'(t)}\partial_j \tag{22}$$

$$= \sum_{j=1}^{n} \frac{d^2y^j}{dt^2}\partial_j + \sum_{i,j=1}^{n} \frac{dy^j}{dt}\nabla_{\frac{dy^i}{dt}\partial_i}\partial_j = \sum_{j=1}^{n} \frac{d^2y^j}{dt^2}\partial_j + \sum_{i,j=1}^{n} \frac{dy^j}{dt}\frac{dy^i}{dt}\nabla_{\partial_i}\partial_j \tag{23}$$

$$= \sum_{k=1}^{n} \frac{d^2y^k}{dt^2}\partial_k + \sum_{i,j,k=1}^{n} \frac{dy^j}{dt}\frac{dy^i}{dt}\Gamma_{ij}^k\partial_k = \sum_{k=1}^{n}\left(\frac{d^2y^k}{dt^2} + \sum_{i,j=1}^{n}\frac{dy^i}{dt}\frac{dy^j}{dt}\Gamma_{ij}^k\right)\partial_k. \tag{24}$$

Portanto, temos o seguinte resultado.

Teorema 2.3. *Se M é uma variedade suave com uma conexão ∇ e $c\colon I \to M$ é uma curva, então c é uma geodésica se, com respeito a qualquer carta $(U, \phi) = (U, x^1, \ldots, x^n)$, as componentes de $y = \phi \circ c$ satisfazem o sistema de EDOs*

$$\frac{d^2y^k}{dt^2} + \sum_{i,j=1}^{n}\frac{dy^i}{dt}\frac{dy^j}{dt}\Gamma_{ij}^k = 0 \tag{25}$$

As equações do sistema acima são chamadas de *equações geodésicas*. Pelo teorema de existência e unicidade de solução para EDOs temos a existência e unicidade de geodésicas.

Teorema 2.4. *Seja M uma variedade suave com uma conexão ∇. Dado $p \in M$ e $X_p \in T_pM$, existe uma geodésica $c\colon I \to M$ tal que $c(0) = p$ e $c'(0) = X_p$. Mais do que isso, essa geodésica é única no sentido de que qualquer outra geodésica satisfazendo essas propriedades deve coincidir com c na intersecção de seus domínios.*

Um *difeomorfismo* entre variedades suaves M e N é um mapa $F\colon M \to N$ suave, bijetor e com inversa suave. Se M e N forem riemannianas, dizemos que F é uma *isometria* se, para todos $p \in M$ e $X_p, Y_p \in T_pM$, temos

$$\mathfrak{g}_p(X_p, Y_p) = \mathfrak{g}_{F(p)}(D_pF(X_p), D_pF(Y_p)). \tag{26}$$

Proposição 2.5. *Isometrias preservam conexões de Levi-Civita. Mais ainda, mapas que preservam conexões, preservam geodésicas. Como corolário, isometrias preservam geodésicas.*

Se $c\colon I \to M$ é uma curva e V é um campo ao longo de c, então dizemos que V é *paralelo* se $DV/Dt = 0$. Dessa forma, uma geodésica é uma curva cujo campo velocidade é paralelo. Fixado $X_p \in T_{c(t_0)}M$, existe um único campo V ao longo de c, paralelo, tal que $V(t_0) = X_p$. Se $c\colon [a, b] \to M$ é uma curva e V é um campo paralelo ao longo de c, dizemos que $V(b)$ é obtido a partir de $V(a)$ por *translação paralela*. Dizemos que $V(b)$ é o *transporte paralelo* de $V(a)$ ao longo de c.

Proposição 2.6. *Se V e W são paralelos ao longo de c em uma variedade riemanniana M, então $||V||$ e $\mathfrak{g}(V, W)$ são constantes.*

2.4 Mapa exponencial e mapa logarítmico

Uma geodésica $c\colon I \to M$ é *maximal* se não podemos estender c para um intervalo maior do que I sem que a curva deixe de ser uma geodésica. Do Teorema 2.4 temos que, dado $p \in M$ e $X_p \in T_pM$ existe uma única geodésica maximal c com $c(0) = p$ e $c'(0) = X_p$. Vamos denotar essa geodésica por γ_{X_p}.

O *mapa exponencial* em um ponto $p \in M$ é a função dada por $\operatorname{Exp}_p(X_p) = \gamma_{X_p}(1)$. Note que podemos assumir que γ_{X_p} sempre está definida em 1, pois caso não esteja, basta reparametrizá-la por algum mapa da forma $\phi(t) = kt$.

2.5 Modelos para a geometria hiperbólica

3 Redes neurais

4 Misturando tudo

Referências

[1] Hugo Cattarucci Botós. «Geometrias Clássicas». 2020.

[2] Loring W. Tu. *Differential Geometry: Connections, Curvature and Characteristic Classes*. Springer-Verlag New York Inc, 2017. ISBN: 978-3-319-55082-4.

[3] Wei Peng e Tuomas Varanka e Abdelrahman Mostafa e Henglin Shi e Guoying Zhao. «Hyperbolic Deep Neural Networks: A Survey». Em: *JOURNAL OF LATEX CLASS FILES* 14.8 (2015).