Metric spaces

lecture 4: Hopf-Rinow theorem

Misha Verbitsky

IMPA, sala 23

January 6, 2022, 17:00

Local metrics and weakly intrinsic metrics (reminder)

CLAIM: Let d_i be a family of metrics (possibly infinite), and $d(x,y) := \sup_i d_i(x,y)$. **Then d is also a metric.**

Proof: We need to check only that $d(x,y) \leqslant d(x,z) + d(z,y)$. This is clear, because

$$d(x,y) = \sup_i d_i(x,y) \leqslant \sup_i (d_i(x,z) + d_i(z,y)) \leqslant$$
$$\leqslant \sup_i d_i(x,z) + \sup_i d_i(z,y) = d(x,z) + d(z,y).$$

■

DEFINITION: Let $\{U_i\}$ be an open covering of s metric space $\{M,d\}$. Denote by $d_{\{U_i\}}$ the metric $\sup_\alpha d_\alpha$, where the supremum is taken over all metrics d_α which satisfy $d_\alpha\big|_{U_i} = d$ for all open sets U_i in the cover. A metric d is called **$\{U_i\}$-local** if $d_{\{U_i\}} = d$. It is called **ε-local**, if it is $\{U_i\}$-local with respect to the covering $\{U_i\}$ consisting of all ε-balls, and **local** if it is ε-local for all $\varepsilon > 0$.

DEFINITION: For any two points x,y in a metric space (M,d), **an ε-chain**, connecting x to y is a collection of points $x = z_0, z_1, ..., z_{n-1}, z_n = y$ such that $d(z_i, z_{i+1}) \leqslant \varepsilon$. Its **defect** is the number $\sum_{i=0}^{n-1} d(z_i, z_{i+1}) - d(x,y)$. The space (M,d) is **weakly intrinsic** if for any two points $x,y \in M$ such that $d(x,y) < \infty$ and any $\varepsilon > 0$, $\delta > 0$, there exists an ε-chain connecting x to y with defect $\leqslant \delta$.

The distance between metric balls (reminder)

DEFINITION: For any two subsets $A, B \subset M$, we denote by $d(A,B)$ the number $\inf_{a\in A, b\in B} d(a,b)$.

DEFINITION: We say that a metric space (M,d) **admits ε-midpoints** if any $x, y \in M$ we have $d(B_x(r/2), B_y(r/2)) = 0$.

THEOREM: Let (M,d) be a metric space. **Then the following conditions are equivalent.**
(1). (M,d) is weakly intrinsic.
(2). (M,d) is local.
(3). For any $x, y \in M$, and any $r_1, r_2 > 0$ such that $d(x,y) = r_1 + r_2$, we have $d(B_x(r_1), B_y(r_2)) = 0$.
(4). (M,d) admits ε-midpoints.

Proof: Lecture 3. ■

Compact metric spaces

DEFINITION: For a subset $Z \subset M$ in a metric space, we denote by $Z(\varepsilon)$ the union of all ε-balls with centers in Z. **An ε-net** in a metric space M is a subset $Z \subset M$ such that $Z(\varepsilon) = M$. The space M is called **totally bounded** if for any $\varepsilon > 0$ it contains a finite ε-net.

PROPOSITION: A metric space M **is compact if and only if it is complete and totally bounded.**

Proof. Step 1: If M is compact, any sequence in M has a limit point; therefore, any Cauchy sequence converges. Also, any open covering has a finite subcovering, taking a covering by ε-balls, we **obtain M as a union of finitely many ε-balls,** this gives a finite ε-net.

Step 2: Conversely, let $\{a_i\}$ be an infinite sequence. Then infinitely many elements of a_i are contained in $B_{x_1}(1/2)$, where x_1 is a point on a finite 1/2-net. Take as b_1 the first of these. Again, infinitely many elements of $\{a_i\} \cap B_{x_1}(1/2)$ are contained in $B_{x_2}(1/4)$. Takie b_2 the first of these after b_1. Repeat, by taking as b_n the first element of $\{a_i\} \cap \bigcap_{i=1}^n B_{x_i}(1/2^i)$ after b_{n-1}. Almost all elements of $\{b_i\}$ are contained in $B_{x_n}(1/2^n)$, for all $n > 0$, hence it is a Cauchy sequence, and it converges. **We proved that M is sequentially compact.** ■

Hopf-Rinow theorem

DEFINITION: A topological space M is **locally compact** if every point $x \in M$ is contained in an open subset $U \subset M$ such that the closure of U is compact.

THEOREM: (Hopf-Rinow)
Let M be a complete, locally compact space with a weakly intrinsic metric. **Then every closed metric ball $B^{cl}_x(r)$ in M is compact.**

Proof. Step 1: For any $y \in M \backslash B_x(r)$, one has $d(y, B_x(r)) = d(x,y) - r$. Indeed, $d(B_x(r), B_y(d(x,y)-r)) = 0$, hence for each $\delta > 0$ there exists a point $z \in B_x(r)$ such that $d(z,y) > d(x,y) - r - \delta$, which gives $d(y, B_x(r)) \geqslant d(x,y) - r - \delta$; passing to the limit $\delta \to 0$, obtain $d(y, B_x(r)) \geqslant d(x,y) - r$, and the opposite inequality follows from the triange inequality.

Step 2: This implies that **an ε-neighbourhood of an open ball $B_x(r)$ contains the corresponding closed ball $B^{cl}_x(r)$.** In particular, $B^{cl}_x(r) = \overline{B}_x(r)$ **(the closed ball is the closure of the open ball of the same radius)**,

REMARK: This is false, for instance, when M is discrete.

Hopf-Rinow theorem (2)

THEOREM: (Hopf-Rinow)
Let M be a complete, locally compact space with a weakly intrinsic metric. **Then every closed metric ball $B_x^{cl}(r)$ in M is compact.**

Step 1-2: In a weakly intrinsic metric space, a closed ball is a closure of an open ball: $B_x^{cl}(r) = \overline{B}_x(r)$. Also, $d(y, B_x(r)) = d(x,y) - r$.

Step 3: Let $m \in M$ be a point such that the ball $B_m(r-\varepsilon)$ is totally bounded for all $\varepsilon > 0$. **Then $B_m(r)$ is also totally bounded.** Indeed, an $\frac{1}{2}\varepsilon$-net in $B_{r-\frac{1}{2}\varepsilon}(m)$ is an ε-net in $B_r(m)$, because any $y \in B_r(m)$ satisfies $d(y, B_{r-\frac{1}{2}\varepsilon}(m)) < \frac{1}{2}\varepsilon$ (Step 1), hence the distance from y and a point in an $\frac{1}{2}\varepsilon$-net in $B_{r-\frac{1}{2}\varepsilon}(m)$ is $< \varepsilon$.

Step 4: Define the following function on a metric space M: $\rho(m) := \sup_r\{r \in \mathbb{R} \mid \overline{B}_r(m) \text{ compact }\}$. Clearly, an $r - d(x,y)$-ball with center in y is contained in $B_x(r)$. If $\rho(x) > \rho(y)$, the $\rho(x) - \varepsilon$-ball with center in x is contained in the $\rho(x) - d(x,y) - \varepsilon$-ball with center in y. Therefore, $\rho(y) \geqslant \rho(x) - d(x,y)$, hence $|\rho(x) - \rho(x)| \leqslant d(x,y)$. In other words, **the function ρ is 1-Lipschitz**.

Hopf-Rinow theorem (3)

Step 1-2: In a weakly intrinsic metric space, **a closed ball is a closure of an open ball: $B_x^{cl}(r) = \overline{B}_x(r)$.** Also, $d(y, B_x(r)) = d(x,y) - r$.

Step 3-4: A closed ball $\overline{B}_x(r)$ **is compact** if $\overline{B}_x(r-\varepsilon)$ is compact for all $\varepsilon > 0$. Also, **the function $\rho(m) := \sup_r\{r \in \mathbb{R} \ \mid \ \overline{B}_r(m)$ compact $\}$ is 1-Lipschitz.**

Step 5: Let $\overline{B}_m(r)$ be a compact closed ball in a locally compact space. Then there exists $\varepsilon > 0$ such that **every closed ε-ball with center in $\overline{B}_m(r)$ is compact.** Indeed, ρ is continuous, hence reaches its minimum somewhere on $\overline{B}_m(r)$.

Step 6: Let $\overline{B}_m(r)$ be a compact closed ball, $\rho\Big|_{B_m(r)} > \varepsilon$, and $Z \subset \overline{B}_m(r)$ a finite $\frac{1}{4}\varepsilon$-net. Then $\overline{Z(\varepsilon)}$ is compact, because it is a finite union of compact sets. However, any $u \in \overline{B}_m(r + \frac{1}{2}\varepsilon)$ satisfies $d(u, \overline{B}_m(r) \leqslant \frac{1}{2}\varepsilon$, hence $d(u, Z) \leqslant \frac{3}{4}\varepsilon < \varepsilon$, **Therefore, $\overline{B}_m(r + \frac{1}{2}\varepsilon) \subset \overline{Z(\varepsilon)}$ is also compact.**

Step 7: Step 3 implies that $\overline{B}_x(\rho(x))$ is compact, Step 6 implies that in this case $\overline{B}_x(\rho(x) + \frac{1}{2}\varepsilon)$ is also compact. This is impossible, unless $\rho(x) = \infty$. **Then $\rho(x) = \infty$, and all closed balls in M are compact.** ■

Hopf, Rinow, Cohn-Vossen

H. Hopf & W. Rinow, Ueber den Begriff der vollständigen differentialgeometrischen Fläche, *Commentarii Mathematici Helvetici volume 3, pages 209-225 (1931).*

Stefan Cohn-Vossen, Sur la courbure totales des surfaces ouvertes, *C. R. Acad. Sci. Paris 197 (1933), 1165-1167.*

Stefan Cohn-Vossen, Kürzeste Wege und Totalkrümmung auf Flächen, *Compositio Math. 2 (1935), 69-133.*

Stefan Cohn-Vossen, Totalkrum̈mung und geodätische Linien auf einfachzusammenhängenden offenen vollständigen Flächenstücken, *Mat. Sbornik N. Ser. 1 (43):2 (1936), 139-164.*

Cohn-Vossen significantly improved the statement and the proof of Hopf-Rinow theorem; he also proved existence of geodesics in a very general setting.

Heinz Hopf and Willi Rinow

Heinz Hopf (1894-1971), photographed in 1936

Willi Rinow (1907-1979), 1960, Greifswald

Stefan Cohn-Vossen, 1902 - 1936

Stefan Cohn-Vossen with his son Richard

Intrinsic metrics (reminder)

Let (M, d) be a metric space, and $\gamma : [a, b] \mapsto M$ a continuous path (here $[a, b]$ denotes the closed interval). Let $x_0 = a < x_1 < ... < x_{n-1} < b = x_n$ be the partition of the interval, and $L_\gamma(x_1, ...x_{n-1}) := \sum_{i=0}^{n-1} d(\gamma(x_i), \gamma(x_{i+1}))$ the length of the corresponding polygonal chain.

DEFINITION: We define **the arc-length** (or **the length**) of the path γ as

$$L_d(\gamma) := \sup_{a<x_1<...<x_{n-1}<b} L_\gamma(x_1, ...x_{n-1}),$$

where supremum is taken over all partitions of the interval $[a, b]$. A path is called **rectifiable** if its arc-length is finite.

DEFINITION: A metric d on M us called **intrinsic metric** if $d(x, y) = \inf_\gamma L_d(\gamma)$, where the infimum is taken over all rectifiable paths γ connecting x to y.

Normal parametrization

DEFINITION: A continuous path $\gamma : [a,b] \longrightarrow M$ is called **minimizing**, if its arc-length is equal to $d(\gamma(a),\gamma(b))$.

DEFINITION: Given a homeomorphism $\varphi : [a,b] \longrightarrow [a,b]$, and a path $\gamma : [a,b] \longrightarrow M$, the composition $\varphi \circ \gamma$ is also a path from $x := \gamma(a)$ to $y := \gamma(b)$. Such a path is called **a reparametrization of γ**.

DEFINITION: Let $\gamma : [a,b] \longrightarrow M$ be a path, and γ_1 takes $t \in \mathbb{R}^{\geqslant 0}$ to $\gamma(t)$, where t is the minimal of all $x \in [a,b]$ such that $L_d(\gamma\big|_{[a,x]}) = t$ (here, as everywhere, L_d denotes the arc length). When $t_i \in \mathbb{R}^{\geqslant 0}$ is the sequence converging to t, the sequence $\gamma(t_i)$ converges to t, because $L_d(\gamma_1\big|_{[t_i,t]}) \geqslant d(\gamma_1(t),\gamma_1(t_i))$. Therefore, γ_1 is continuous. Such a path γ_1 is called **the normal parametrization** of γ.

Normal parametrization and minimizing geodesics

CLAIM: Let $\gamma: [a,b] \longrightarrow M$ be minimizing, and $\gamma_1: [0,\alpha] \longrightarrow M$ its normal parametrization. **Then γ_1 is an isometry.**

Proof. Step 1: Any interval of a minimizing path is minimizing. Indeed, consider a partition $[a,c] \cup [b,c] = [a,b]$. Then

$$d(\gamma(a),\gamma(b)) \leqslant L_\gamma(c) \leqslant L_d(\gamma) = d(\gamma(a),\gamma(b)).$$

This gives $L_\gamma(c) = d(\gamma(a),\gamma(c)) + d(\gamma(c),\gamma(b)) = d(\gamma(a),\gamma(b))$. Then

$$L_d(\gamma) = L_d(\gamma\big|_{[a,c]}) + L_d(\gamma\big|_{[c,b]}) \geqslant d(\gamma(a),\gamma(c)) + d(\gamma(c),\gamma(b)) = d(\gamma(a),\gamma(b))$$

hence **the inequalities $L_d(\gamma\big|_{[a,c]}) \geqslant d(\gamma(a),\gamma(c))$ and $L_d(\gamma\big|_{[c,b]}) \geqslant d(\gamma(c),\gamma(b))$ are equalities.**

Step 2: Let $v \in [0,\alpha]$. The same argument applied to a partition $[0,v] = [0,u] \cup [u,v]$ would imply that

$$L_d(\gamma_1\big|_{[0,u]}) + L_d(\gamma_1\big|_{[u,v]}) = u + L_d(\gamma_1\big|_{[u,v]}) = L_d(\gamma_1\big|_{[0,v]}) = v,$$

giving $L_d(\gamma_1\big|_{[u,v]}) = v-u$. Then Step 1 gives $v-u = L_d(\gamma_1\big|_{[u,v]}) = d(\gamma_1(u),\gamma_1(v))$, hence γ_1 is an isometry. ■