Introducción a las geodésicas en espacios métricos

Julián Fernández Bonder

Instituto de Cálculo, FCEN–CONICET · Departamento de Matemática, FCEN–UBA

Resumen. Estas notas introducen la noción de geodésica en un espacio métrico arbitrario. Se define la longitud de una curva mediante poligonales inscriptas y se prueban sus propiedades básicas: aditividad, continuidad de la longitud acumulada y semicontinuidad inferior. Luego se introducen los espacios de longitud y la métrica intrínseca d_L inducida por una métrica d, y se estudia la reparametrización por longitud de arco. El resultado central es el teorema de Hopf–Rinow para espacios métricos: todo espacio de longitud propio es geodésico. Para mostrar el papel de la compacidad, se analizan dos ejemplos sin compacidad local, uno geodésico (un espacio de Hilbert) y otro que no lo es (un grafo con infinitas aristas). La única herramienta externa es el teorema de Arzelà–Ascoli, que se enuncia en la forma que se necesita.

1. Longitud de curvas

El enfoque de estas notas (longitud de curvas definida por poligonales inscriptas, espacios de longitud, métrica intrínseca y existencia de geodésicas por compacidad) es el de la geometría métrica; las referencias generales son [1], Capítulo 2, [2], Capítulo I.1 y [4], Capítulo 1.

Sea (X, d) un espacio métrico y sea $\gamma\colon [a,b] \to X$ una curva continua. La *longitud* de γ se define por

$$L(\gamma) := \sup\left\{ \sum_{i=1}^{n} d\big(\gamma(t_i), \gamma(t_{i-1})\big) : a = t_0 < t_1 < \cdots < t_n = b \right\} \in [0, \infty].$$

Cuando haya más de una métrica en juego escribiremos $L_d(\gamma)$.

Definición 1.1. La curva γ es *rectificable* si $L(\gamma) < \infty$.

Cada suma de la definición es la longitud de la "poligonal" que une los puntos $\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_n)$, en la que cada tramo se mide con la distancia d (Figura 1). Por la desigualdad triangular, agregar puntos a la partición no disminuye la suma, así que la longitud es el límite de estas sumas a medida que la partición se refina.

Por la desigualdad triangular, toda curva que une x con y satisface $d(x,y) \le L(\gamma)$. Más en general, si $s \le t$, entonces $d(\gamma(s), \gamma(t)) \le L(\gamma|_{[s,t]})$.

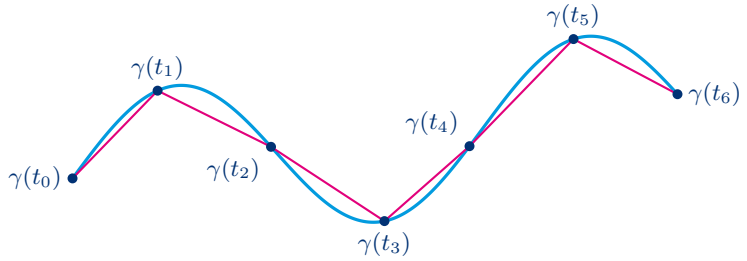


FIGURA 1. La curva γ (en cian) y la poligonal inscripta asociada a una partición $t_0 < \cdots < t_6$ (en magenta). La longitud de la poligonal es $\sum_i d(\gamma(t_{i-1}), \gamma(t_i)) \le L(\gamma)$.

Ejemplo 1.2 (Arcos de circunferencia)**.** Sea $\mathbb{S}^1 = \{z \in \mathbb{R}^2 : |z| = 1\}$ con la distancia euclídea $d(z, w) = |z - w|$, y sea $\gamma(t) = (\cos t, \sin t)$ para $t \in [0, \theta]$, con $0 < \theta \le 2\pi$. Veamos que $L(\gamma) = \theta$, es decir, que la definición de longitud da la longitud de arco usual.

Para $s \le t$, calculamos la cuerda entre $\gamma(s)$ y $\gamma(t)$:

$$\begin{aligned} |\gamma(t) - \gamma(s)|^2 &= (\cos t - \cos s)^2 + (\sin t - \sin s)^2 \\ &= \cos^2 t + \sin^2 t + \cos^2 s + \sin^2 s - 2(\cos t \cos s + \sin t \sin s) \\ &= 2 - 2\cos(t - s) = 4\sin^2\Big(\frac{t-s}{2}\Big), \end{aligned}$$

donde usamos $\cos(t - s) = \cos t \cos s + \sin t \sin s$ y $1 - \cos u = 2\sin^2(u/2)$. Como $0 \le \frac{t-s}{2} \le \pi$, se tiene $\sin\big(\frac{t-s}{2}\big) \ge 0$, y entonces

$$|\gamma(t) - \gamma(s)| = 2\sin\Big(\frac{t-s}{2}\Big) \le t - s,$$

porque $\sin u \le u$ para $u \ge 0$. Sumando sobre cualquier partición de $[0, \theta]$, cada suma es a lo sumo θ, y entonces $L(\gamma) \le \theta$. Por otro lado, para la partición en n partes iguales todas las cuerdas miden lo mismo (Figura 2), y

$$\sum_{i=1}^{n} d\big(\gamma(t_{i-1}), \gamma(t_i)\big) = 2n\sin\Big(\frac{\theta}{2n}\Big) \xrightarrow[n\to\infty]{} \theta,$$

porque $\sin u/u \to 1$ cuando $u \to 0$. Luego $L(\gamma) = \theta$.

Lema 1.3 (Aditividad)**.** *Si $a \le c \le b$, entonces $L(\gamma) = L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]})$.*

Demostración. Si P es una partición de $[a, b]$, agregarle el punto c no disminuye la suma correspondiente (desigualdad triangular), y la partición resultante se separa en una de $[a, c]$ y una de $[c, b]$. Luego la suma de P es a lo sumo $L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]})$, y tomando supremo sale "$\le$". Recíprocamente, juntando una partición de $[a, c]$ con una de $[c, b]$ se obtiene una de $[a, b]$, cuya suma es la suma de ambas; tomando supremos por separado sale "$\ge$". □

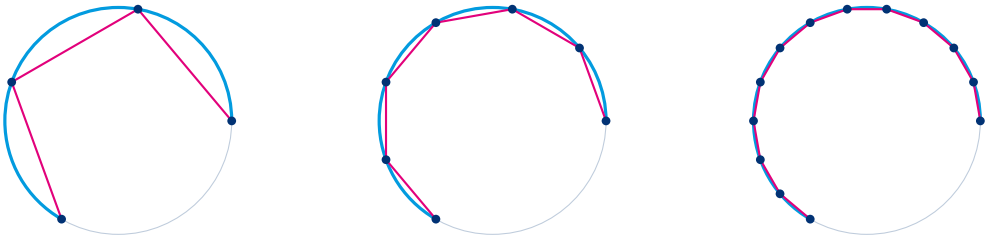


FIGURA 2. Poligonales inscriptas en un arco de ángulo $\theta = \frac{4\pi}{3}$ con particiones uniformes de n partes. Sus longitudes, $2n\sin(\theta/2n)$, crecen hacia θ.

Definición 1.4 (Longitud acumulada)**.** Sea $\gamma\colon [a,b] \to X$ una curva rectificable. La *longitud acumulada* de γ es la función $\ell\colon [a,b] \to [0, L(\gamma)]$ dada por

$$\ell(t) := L\big(\gamma|_{[a,t]}\big), \qquad \ell(a) = 0.$$

Por el Lema 1.3, para $a \le s \le t \le b$ se tiene $\ell(t) = \ell(s) + L(\gamma|_{[s,t]})$, es decir,

$$\ell(t) - \ell(s) = L\big(\gamma|_{[s,t]}\big) \ge 0. \tag{1}$$

En particular, ℓ es no decreciente, $\ell(b) = L(\gamma)$ y todos los valores de ℓ son finitos.

Lema 1.5 (Continuidad de la longitud acumulada)**.** *Si $\gamma\colon [a,b] \to X$ es rectificable, su longitud acumulada ℓ es continua.*

Demostración. Como ℓ es no decreciente, basta ver que es continua por derecha en cada $t_0 \in [a,b)$ y por izquierda en cada $t_0 \in (a,b]$. Sea $\varepsilon > 0$.

Continuidad por derecha en $t_0 \in [a,b)$. Por la definición de supremo existe una partición $t_0 = s_0 < s_1 < \cdots < s_n = b$ de $[t_0, b]$ tal que

$$\sum_{i=1}^{n} d\big(\gamma(s_{i-1}), \gamma(s_i)\big) > L\big(\gamma|_{[t_0,b]}\big) - \varepsilon.$$

Por continuidad de γ en t_0, existe $\delta \in (0, s_1 - t_0)$ tal que $d(\gamma(t_0), \gamma(t)) < \varepsilon$ para todo $t \in (t_0, t_0 + \delta)$. Fijemos un tal t. Por la desigualdad triangular, $d(\gamma(t_0), \gamma(s_1)) \le d(\gamma(t_0), \gamma(t)) + d(\gamma(t), \gamma(s_1))$. Como $t < s_1$, los puntos $t < s_1 < \cdots < s_n$ forman una partición de $[t, b]$, de modo que

$$L\big(\gamma|_{[t_0,b]}\big) - \varepsilon < d(\gamma(t_0), \gamma(t)) + \Big[d(\gamma(t), \gamma(s_1)) + \sum_{i=2}^{n} d(\gamma(s_{i-1}), \gamma(s_i))\Big] < \varepsilon + L\big(\gamma|_{[t,b]}\big).$$

Todas las longitudes son finitas, así que por aditividad $L(\gamma|_{[t_0,t]}) = L(\gamma|_{[t_0,b]}) - L(\gamma|_{[t,b]}) < 2\varepsilon$. Por (1), $0 \le \ell(t) - \ell(t_0) < 2\varepsilon$ para todo $t \in (t_0, t_0 + \delta)$.

Continuidad por izquierda en $t_0 \in (a,b]$. Existe una partición $a = s_0 < \cdots < s_n = t_0$ de $[a, t_0]$ con

$$\sum_{i=1}^{n} d\big(\gamma(s_{i-1}), \gamma(s_i)\big) > \ell(t_0) - \varepsilon.$$

Por continuidad de γ en t_0, existe $\delta \in (0, t_0 - s_{n-1})$ tal que $d(\gamma(t), \gamma(t_0)) < \varepsilon$ para todo $t \in (t_0 - \delta, t_0)$. Fijemos un tal t. Como $s_{n-1} < t$, los puntos $s_0 < \cdots < s_{n-1} < t$ forman una partición de $[a, t]$, y por la desigualdad triangular $d(\gamma(s_{n-1}), \gamma(t_0)) \leq d(\gamma(s_{n-1}), \gamma(t)) + d(\gamma(t), \gamma(t_0))$. Luego

$$\ell(t_0) - \varepsilon < \sum_{i=1}^{n-1} d\big(\gamma(s_{i-1}), \gamma(s_i)\big) + d(\gamma(s_{n-1}), \gamma(t)) + d(\gamma(t), \gamma(t_0)) < \ell(t) + \varepsilon.$$

Por la monotonía, $0 \leq \ell(t_0) - \ell(t) < 2\varepsilon$ para todo $t \in (t_0 - \delta, t_0)$. □

Lema 1.6 (Semicontinuidad inferior)**.** *Sean $\gamma_n, \gamma \colon [a, b] \to X$ curvas tales que $\gamma_n(t) \to \gamma(t)$ para todo $t \in [a, b]$. Entonces*

$$L(\gamma) \leq \liminf_{n\to\infty} L(\gamma_n).$$

Demostración. Para una partición $a = t_0 < \cdots < t_k = b$ fija, por continuidad de d,

$$\sum_{i=1}^{k} d(\gamma(t_{i-1}), \gamma(t_i)) = \lim_{n\to\infty} \sum_{i=1}^{k} d(\gamma_n(t_{i-1}), \gamma_n(t_i)) \leq \liminf_{n\to\infty} L(\gamma_n).$$

Tomando supremo sobre las particiones se obtiene el resultado. □

2. Espacios de longitud

Definición 2.1. El espacio métrico (X, d) es un *espacio de longitud* si, para todo $x, y \in X$ y todo $\varepsilon > 0$, existe una curva rectificable $\gamma \colon [0, 1] \to X$ con $\gamma(0) = x$, $\gamma(1) = y$ y

$$L(\gamma) < d(x, y) + \varepsilon.$$

Equivalentemente, la distancia entre dos puntos es el ínfimo de las longitudes de las curvas que los unen.

2.1. La métrica intrínseca.

Supongamos ahora que (X, d) es un espacio métrico cualquiera en el que todo par de puntos se puede unir con una curva rectificable. Definimos

$$d_L(x, y) := \inf\{L_d(\gamma) : \gamma \text{ es una curva que une } x \text{ con } y\}.$$

Proposición 2.2. *d_L es una métrica en X y $d \leq d_L$.*

Demostración. La finitud de d_L es exactamente la hipótesis de que todo par de puntos se une con una curva rectificable; sin ella, d_L sería una métrica extendida (con valores en $[0, \infty]$). Como $d(x, y) \leq L_d(\gamma)$ para toda curva que une x con y, tomando ínfimo se obtiene $d \leq d_L$; en particular, $d_L(x, y) = 0$ implica $x = y$. La simetría sale de que la curva recorrida al revés tiene la misma longitud. Para la desigualdad triangular, si γ_1 une x con y y γ_2 une y con z, su concatenación une x con z y, por el Lema 1.3, tiene longitud $L_d(\gamma_1) + L_d(\gamma_2)$. Tomando ínfimos, $d_L(x, z) \leq d_L(x, y) + d_L(y, z)$. □

Observación 2.3. Como $d \leq d_L$, si una sucesión converge para d_L, entonces converge para d y al mismo límite: de $d_L(x_n, x) \to 0$ se deduce $d(x_n, x) \leq d_L(x_n, x) \to 0$. La recíproca puede fallar: puede haber sucesiones que convergen para d y no convergen para d_L (Ejemplo 2.5).

Ejemplo 2.4 (La circunferencia y la esfera)**.** Volvamos a $\mathbb{S}^1 \subset \mathbb{R}^2$ con la distancia euclídea $d(z,w) = |z-w|$. Dados $z, w \in \mathbb{S}^1$, sea $\theta \in [0,\pi]$ el ángulo entre ellos, es decir, $\cos\theta = \langle z, w\rangle$. Entonces:

- La distancia euclídea es la longitud de la cuerda: $d(z,w) = 2\sin(\theta/2)$.
- La distancia intrínseca es la longitud del arco más corto: $d_L(z,w) = \theta$.

Para la segunda afirmación, el arco más corto tiene longitud θ por el Ejemplo 1.2, así que $d_L(z,w) \le \theta$. Recíprocamente, toda curva γ en $\mathbb{S}^1$ de z a w se puede escribir como $\gamma(t) = (\cos\varphi(t), \sin\varphi(t))$ con φ continua (el argumento continuo), y $|\varphi(1) - \varphi(0)| \ge \theta$. Tomando particiones finas, donde las cuerdas $2\sin(|\Delta\varphi_i|/2)$ son casi iguales a los incrementos $|\Delta\varphi_i|$, se obtiene $L(\gamma) \ge \sum_i |\Delta\varphi_i| \ge \theta$, salvo un error arbitrariamente chico. Luego $d_L(z,w) = \theta$.

Como $2\sin(\theta/2) < \theta$ para $\theta > 0$, el espacio $(\mathbb{S}^1, d)$ *no* es de longitud: la cuerda que realiza la distancia euclídea se sale de la circunferencia, y toda curva dentro de $\mathbb{S}^1$ es más larga (Figura 3, izquierda).

Sin embargo, las dos métricas son *equivalentes*. Como sin es cóncava en $[0,\pi/2]$, vale $\sin u \ge \frac{2}{\pi}u$ para $u \in [0, \frac{\pi}{2}]$; con $u = \theta/2$ se obtiene $\theta \le \frac{\pi}{2} \cdot 2\sin(\theta/2)$. Es decir,

$$d(z,w) \le d_L(z,w) \le \frac{\pi}{2}\, d(z,w) \qquad (z, w \in \mathbb{S}^1).$$

La constante $\frac{\pi}{2}$ es óptima: se alcanza en puntos antipodales, donde $d = 2$ y $d_L = \pi$. En particular, d y d_L tienen las mismas sucesiones convergentes y las mismas sucesiones de Cauchy.

Lo mismo ocurre en la esfera $\mathbb{S}^2 \subset \mathbb{R}^3$ (Figura 3, derecha). Si $\theta \in [0,\pi]$ es el ángulo entre z y w, la distancia euclídea es otra vez $2\sin(\theta/2)$, y la distancia intrínseca es θ: es la longitud del arco más corto del *círculo máximo* que pasa por z y w (la intersección de $\mathbb{S}^2$ con el plano que contiene a 0, z y w). Que ningún otro camino sobre la esfera sea más corto es un hecho clásico que aceptamos sin demostración. La desigualdad $d \le d_L \le \frac{\pi}{2}d$ vale con la misma cuenta.

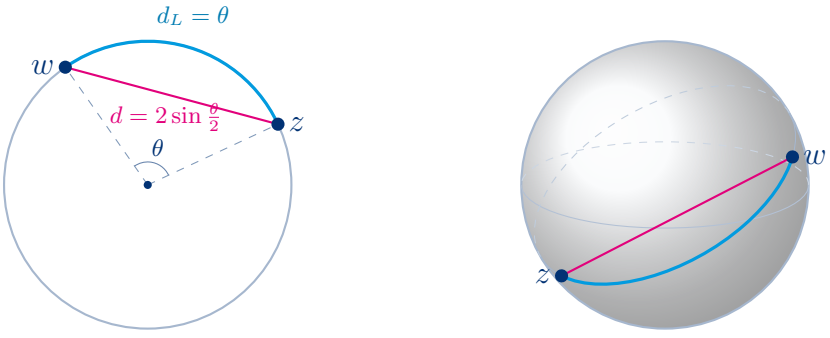


FIGURA 3. Izquierda: en $\mathbb{S}^1$, la distancia euclídea es la longitud de la cuerda (magenta) y la distancia intrínseca es la del arco más corto (cian). Derecha: en $\mathbb{S}^2$, el arco más corto es un arco de círculo máximo.

El ejemplo siguiente muestra que, en cambio, d y d_L pueden no ser equivalentes, e incluso definir topologías distintas.

Ejemplo 2.5 (El peine)**.** Sea $X \subset \mathbb{R}^2$ la unión de $[0,1] \times \{0\}$, de $\{0\} \times [0,1]$ y de los segmentos $\{1/n\} \times [0,1]$, $n \geq 1$, con la métrica euclídea d. Sean $p_n = (1/n, 1)$ y $p = (0,1)$. Entonces $d(p_n, p) = 1/n \to 0$, pero toda curva en X que une p_n con p tiene que bajar hasta el eje horizontal y volver a subir (Figura 4), así que $d_L(p_n, p) \geq 2$. Luego $p_n \to p$ en (X, d) pero no en (X, d_L).

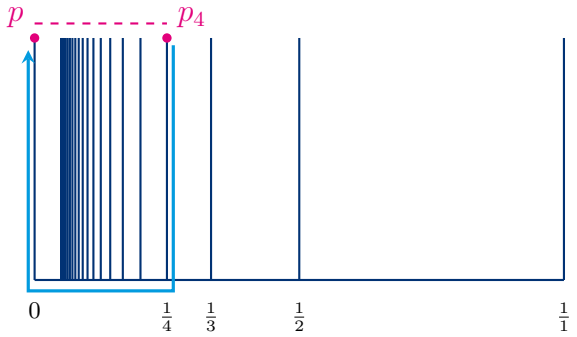


FIGURA 4. El peine. La distancia euclídea entre p_n y p es $1/n$ (en magenta), pero toda curva dentro del peine tiene que bajar y volver a subir (en cian), así que $d_L(p_n, p) \geq 2$.

Proposición 2.6. *Sea $\gamma\colon [a,b] \to X$ una curva.*

(a) Si γ es continua para d_L, entonces es continua para d.
(b) Si γ es d-rectificable, entonces es continua para d_L.
(c) Si γ es continua para ambas métricas, entonces $L_{d_L}(\gamma) = L_d(\gamma)$ (como elementos de $[0,\infty]$).

Demostración. (a) Es inmediato de $d \leq d_L$.

(b) Para $s \leq t$, la restricción $\gamma|_{[s,t]}$ une $\gamma(s)$ con $\gamma(t)$, así que, con ℓ la longitud acumulada respecto de d,

$$d_L(\gamma(s), \gamma(t)) \leq L_d(\gamma|_{[s,t]}) = \ell(t) - \ell(s).$$

Como ℓ es continua (Lema 1.5), γ es continua para d_L.

(c) Como $d \leq d_L$, cada suma sobre una partición calculada con d es menor o igual que la calculada con d_L; tomando supremo, $L_d(\gamma) \leq L_{d_L}(\gamma)$. Si $L_d(\gamma) = \infty$ no hay nada más que probar. Si $L_d(\gamma) < \infty$, para una partición $a = t_0 < \cdots < t_n = b$, cada tramo $\gamma|_{[t_{i-1},t_i]}$ une $\gamma(t_{i-1})$ con $\gamma(t_i)$, de modo que, por el Lema 1.3,

$$\sum_{i=1}^{n} d_L(\gamma(t_{i-1}), \gamma(t_i)) \leq \sum_{i=1}^{n} L_d\big(\gamma|_{[t_{i-1},t_i]}\big) = L_d(\gamma).$$

Tomando supremo, $L_{d_L}(\gamma) \leq L_d(\gamma)$. □

Corolario 2.7. *$(d_L)_L = d_L$; es decir, (X, d_L) es un espacio de longitud.*

Demostración. Por la Proposición 2.6, las curvas d_L-rectificables son exactamente las curvas d-rectificables, y en ambas métricas tienen la misma longitud. Por lo tanto los ínfimos que definen $(d_L)_L$ y d_L se toman sobre el mismo conjunto de números. □

Esto justifica llamar a d_L la *métrica intrínseca* inducida por d: aplicar la construcción una segunda vez ya no cambia nada. Además, $d = d_L$ si y sólo si (X, d) es un espacio de longitud.

3. Reparametrización por longitud de arco

Proposición 3.1. *Sea $\gamma\colon [0,1] \to X$ una curva rectificable de longitud $L > 0$ y sea ℓ su longitud acumulada (Definición 1.4). Existe una curva $\tilde\gamma\colon [0,1] \to X$, con los mismos extremos y la misma imagen que γ, tal que $\gamma = \tilde\gamma \circ (\ell/L)$ y*

$$d\big(\tilde\gamma(s), \tilde\gamma(t)\big) \le L|t-s|, \qquad L\big(\tilde\gamma|_{[s,t]}\big) = L(t-s) \qquad (0 \le s \le t \le 1).$$

Demostración. Por el Lema 1.5, $\ell/L\colon [0,1] \to [0,1]$ es continua, no decreciente y suryectiva. Dado $s \in [0,1]$, definimos $\tilde\gamma(s) := \gamma(t)$ para cualquier t con $\ell(t) = sL$. Está bien definida: si $\ell(t) = \ell(t')$ con $t \le t'$, entonces $d(\gamma(t), \gamma(t')) \le L(\gamma|_{[t,t']}) = \ell(t') - \ell(t) = 0$. Por construcción $\gamma = \tilde\gamma \circ (\ell/L)$, y $\tilde\gamma$ tiene los mismos extremos y la misma imagen que γ.

Si $\ell(t) = sL$ y $\ell(t') = s'L$ con $s \le s'$ (y entonces podemos tomar $t \le t'$),

$$d(\tilde\gamma(s), \tilde\gamma(s')) = d(\gamma(t), \gamma(t')) \le \ell(t') - \ell(t) = L(s'-s).$$

En particular, $\tilde\gamma$ es continua. Para la longitud, las particiones de $[t,t']$ se transforman por ℓ/L en particiones de $[s,s']$ (eventualmente con puntos repetidos, que no aportan a la suma), y toda partición de $[s,s']$ proviene de este modo de una de $[t,t']$, porque ℓ/L es suryectiva y monótona. Como $\gamma = \tilde\gamma \circ (\ell/L)$, las sumas correspondientes coinciden, y entonces $L(\tilde\gamma|_{[s,s']}) = L(\gamma|_{[t,t']}) = \ell(t') - \ell(t) = L(s'-s)$. □

En particular, $\tilde\gamma$ tiene velocidad métrica constante L. En general esto *no* implica que $d(\tilde\gamma(s), \tilde\gamma(t)) = L|t-s|$: esa igualdad caracteriza a las geodésicas (Definición 4.1). Si $L = 0$, la curva es constante y no hace falta reparametrizarla.

4. Geodésicas y el teorema de Hopf–Rinow

Definición 4.1. Una *geodésica* (minimizante, parametrizada a velocidad constante) entre x e y es una curva $\gamma\colon [0,1] \to X$ con $\gamma(0) = x$, $\gamma(1) = y$ y

$$d\big(\gamma(s), \gamma(t)\big) = |t-s|\, d(x,y) \qquad (s,t \in [0,1]).$$

El espacio (X,d) es un *espacio geodésico* si todo par de puntos está unido por una geodésica.

Toda geodésica entre x e y tiene longitud $d(x,y)$. Recíprocamente, si una curva rectificable γ une x con y y $L(\gamma) = d(x,y)$, su reparametrización por longitud de arco es una geodésica. En particular, todo espacio geodésico es un espacio de longitud, en el que además el ínfimo de la definición se alcanza.

Por ejemplo, $(\mathbb{S}^1, d_L)$ y $(\mathbb{S}^2, d_L)$ son espacios geodésicos: las geodésicas son los arcos más cortos de círculo máximo, recorridos a velocidad constante. En cambio, con la distancia euclídea, $\mathbb{S}^1$ y $\mathbb{S}^2$ no tienen geodésicas entre puntos distintos (Ejemplo 2.4).

Ejemplo 4.2 (Las geodésicas no tienen por qué ser únicas)**.** En $\mathbb{R}^2$ con la distancia euclídea, la única geodésica entre dos puntos es el segmento que los une. Pero si consideramos la distancia

$$d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|,$$

hay muchas más. Sean $x = (0, 0)$ e $y = (1, 1)$, de modo que $d_1(x, y) = 2$. El segmento $\gamma(t) = (t, t)$ es una geodésica, pero también lo es el camino que primero va hacia la derecha y después sube,

$$\gamma(t) = \begin{cases} (2t, 0) & \text{si } 0 \le t \le \frac{1}{2}, \\ (1, 2t - 1) & \text{si } \frac{1}{2} \le t \le 1. \end{cases}$$

En efecto, si $s \le \frac{1}{2} \le t$, entonces $d_1(\gamma(s), \gamma(t)) = (1 - 2s) + (2t - 1) = 2(t - s)$, y si s y t están del mismo lado de $\frac{1}{2}$ la cuenta es directa. Con el mismo argumento, cualquier "escalera" en la que ambas coordenadas son monótonas, recorrida a velocidad constante, es una geodésica (Figura 5). Esta falta de unicidad no es una patología: va a aparecer también en los espacios de Wasserstein.

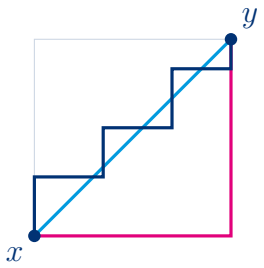


FIGURA 5. Tres geodésicas distintas entre $x = (0, 0)$ e $y = (1, 1)$ para la distancia d_1: el segmento (cian), el camino en "L" (magenta) y una escalera (azul). Todas tienen longitud $d_1(x, y) = 2$.

Un espacio métrico es *propio* si sus bolas cerradas son compactas o, equivalentemente, si sus subconjuntos cerrados y acotados son compactos. Todo espacio propio es completo y localmente compacto.

En la demostración del teorema de Hopf–Rinow (Teorema 4.4) vamos a usar el teorema de Arzelà–Ascoli, pero no en la versión habitual de los cursos de Análisis, que es para funciones a valores reales. Como acá las curvas toman valores en un espacio métrico cualquiera, enunciamos a continuación la versión que vamos a usar, para tenerla a mano y poder referirnos a ella más adelante.

Teorema 4.3 (Arzelà–Ascoli)**.** *Sean (K, d_K) un espacio métrico compacto, (Y, d_Y) un espacio métrico y $f_n \colon K \to Y$ una sucesión de funciones continuas que cumple:*

*(i) (*equicontinuidad*) para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, si $d_K(x, x') < \delta$, entonces $d_Y(f_n(x), f_n(x')) < \varepsilon$ para todo n;*

*(ii) (*imágenes en un compacto*) existe un subconjunto compacto $C \subset Y$ tal que $f_n(K) \subset C$ para todo n.*

Entonces existe una subsucesión (f_{n_k}) que converge uniformemente a una función continua $f\colon K \to Y$.

La versión que suele verse en los cursos de Análisis es el caso $Y = \mathbb{R}$, donde la condición (ii) se pide como equiacotación: $|f_n(x)| \leq M$ para todo n y todo x, es decir, $C = [-M, M]$. Como en $\mathbb{R}$ los compactos son los cerrados y acotados, ambas formulaciones coinciden. Cuando el espacio de llegada es un espacio métrico general, la acotación ya no alcanza (por ejemplo, en un espacio de Hilbert de dimensión infinita la bola unidad cerrada no es compacta, Ejemplo 4.6), y hay que pedir que las imágenes caigan en un mismo compacto. Una referencia es [1], Sección 2.5, donde el teorema se enuncia en una forma más general (alcanza con que, para cada $x \in K$, la clausura de $\{f_n(x)\}$ sea compacta); también puede consultarse [6], Sección 47.

El siguiente resultado es la versión métrica del teorema de Hopf y Rinow [5], que lo probaron para superficies (completitud implica existencia de geodésicas minimizantes), y de su extensión por Cohn-Vossen [3] a espacios de longitud localmente compactos. El enunciado en espacios métricos puede verse en [1], Capítulo 2, y [2], Capítulo I.3.

Teorema 4.4 (Hopf–Rinow para espacios métricos)**.** *Todo espacio de longitud propio es un espacio geodésico. Es decir, para todo $x, y \in X$ existe una geodésica $\gamma\colon [0,1] \to X$ entre x e y; en particular, $L(\gamma) = d(x,y)$.*

Demostración. Pongamos $D = d(x,y)$. Si $D = 0$, entonces $x = y$ y la curva constante sirve. Supongamos $D > 0$.

Por ser X un espacio de longitud, existen curvas $\gamma_n\colon [0,1] \to X$ que unen x con y y cumplen $D \leq L(\gamma_n) \leq D + 1/n$. Como $L(\gamma_n) \geq D > 0$, podemos reparametrizarlas por longitud de arco (Proposición 3.1); llamamos $\tilde{\gamma}_n$ a las curvas reparametrizadas, que tienen los mismos extremos y la misma longitud $L_n := L(\gamma_n)$, y satisfacen

$$d\big(\tilde{\gamma}_n(s), \tilde{\gamma}_n(t)\big) \leq L_n|t-s| \leq (D+1)|t-s|.$$

Así, la familia $\{\tilde{\gamma}_n\}$ es equicontinua (de hecho, uniformemente Lipschitz). Además, para todo t,

$$d(x, \tilde{\gamma}_n(t)) \leq L\big(\tilde{\gamma}_n|_{[0,t]}\big) \leq L_n \leq D + 1,$$

de modo que todas las imágenes están contenidas en la bola cerrada $\overline{B}(x, D+1)$, que es compacta por hipótesis.

Por el Teorema 4.3, aplicado con $K = [0,1]$, $Y = X$ y $C = \overline{B}(x, D+1)$, existe una subsucesión $\tilde{\gamma}_{n_k}$ que converge uniformemente a una curva continua $\gamma\colon [0,1] \to X$, con $\gamma(0) = x$ y $\gamma(1) = y$ (Figura 6). Como la convergencia uniforme implica la convergencia puntual, y el Lema 1.6 sólo pide convergencia puntual, resulta

$$L(\gamma) \leq \liminf_{k\to\infty} L(\tilde{\gamma}_{n_k}) = \lim_{k\to\infty} L_{n_k} = D.$$

Como siempre $D = d(x,y) \leq L(\gamma)$, concluimos que $L(\gamma) = D$.

Para la última afirmación, pasando al límite en la desigualdad de Lipschitz se obtiene $d(\gamma(s), \gamma(t)) \leq D|t-s|$. Por otro lado, si $s \leq t$, por la desigualdad triangular,

$$D = d(x,y) \leq d(x,\gamma(s)) + d(\gamma(s),\gamma(t)) + d(\gamma(t),y) \leq Ds + d(\gamma(s),\gamma(t)) + D(1-t),$$

de donde $d(\gamma(s),\gamma(t)) \geq D(t-s)$. Luego $d(\gamma(s),\gamma(t)) = D|t-s|$. □

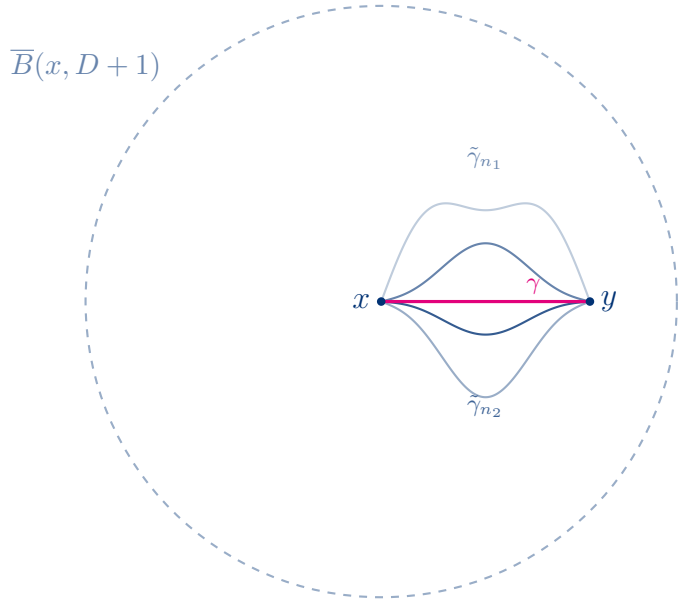


Figura 6. Idea de la demostración del Teorema 4.4. Las curvas casi minimizantes $\tilde{\gamma}_{n_k}$ quedan dentro de la bola compacta $\overline{B}(x, D+1)$ y, por Arzelà–Ascoli, convergen a una curva γ de longitud $d(x, y)$.

Observación 4.5. El enunciado clásico de Hopf–Rinow–Cohn-Vossen [5, 3] pide que X sea un espacio de longitud completo y localmente compacto, y prueba que entonces X es propio; en ese contexto, el Teorema 4.4 es la parte que asegura la existencia de geodésicas. Acá tomamos directamente la propiedad como hipótesis, que es lo único que usa la demostración.

4.1. **Dos ejemplos sin compacidad local.** El Teorema 4.4 da una condición *suficiente* para que un espacio sea geodésico. Los dos ejemplos que siguen muestran que, cuando falta la compacidad, puede pasar cualquier cosa: en ambos el espacio es completo, es de longitud y *no* es localmente compacto, pero el primero es geodésico y el segundo no.

Ejemplo 4.6 (Un espacio geodésico que no es localmente compacto)**.** Sea H un espacio de Hilbert de dimensión infinita (por ejemplo, ℓ^2), con la distancia $d(x, y) = \|x - y\|$. Es completo por definición.

Es geodésico. Dados $x, y \in H$, el segmento $\gamma(t) = (1-t)x + ty$ cumple

$$\|\gamma(s) - \gamma(t)\| = \|(t-s)(x-y)\| = |t-s|\,\|x-y\|,$$

así que es una geodésica entre x e y en el sentido de la Definición 4.1. En particular, H también es un espacio de longitud.

No es localmente compacto. Sea $\{u_k\}_{k\geq 1}$ una sucesión ortonormal y sea $r > 0$. Los puntos $x + r\,u_k$ están en la bola cerrada $\overline{B}(x, r)$, y para $k \neq j$,

$$\|r\,u_k - r\,u_j\| = r\sqrt{\|u_k\|^2 + \|u_j\|^2} = r\sqrt{2}.$$

Como todos los puntos están a distancia $r\sqrt{2}$ entre sí, la sucesión no tiene ninguna subsucesión de Cauchy, y por lo tanto ninguna subsucesión convergente. Así, ninguna bola cerrada centrada en x es compacta, por chica que sea.

Ejemplo 4.7 (Un espacio de longitud completo que no es geodésico)**.** Tomamos dos puntos x e y y los unimos con infinitas aristas $e_1, e_2, e_3, \ldots$, donde la arista e_n es un segmento de longitud

$$\ell_n := 1 + \tfrac{1}{n}.$$

Las aristas sólo se tocan en sus extremos x e y (Figura 7). Cada punto de e_n queda determinado por su distancia $s \in [0, \ell_n]$ a x medida *a lo largo de la arista*. Llamamos *camino* a una curva que se mueve por las aristas, y su longitud es la suma de los tramos recorridos en cada arista. Definimos

$$d(p, q) := \inf\{\text{longitud de un camino de } p \text{ a } q\}.$$

Llamemos X al espacio resultante.

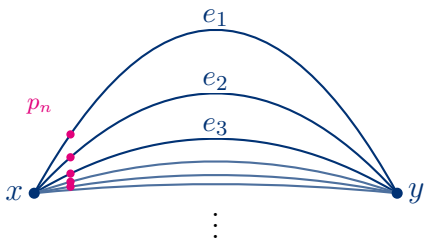


Figura 7. El grafo del Ejemplo 4.7. Las aristas se van acortando: e_n mide $1 + \frac{1}{n}$. Los puntos p_n (en magenta) están todos cerca de x, pero a distancia al menos $2r$ entre sí.

La observación clave es la siguiente. Si sacamos los puntos x e y, el espacio se parte en las aristas abiertas, que son disjuntas entre sí. Por lo tanto:

(∗) *Una curva que sale de un punto interior de e_n no puede entrar en otra arista sin pasar antes por x o por y.*

Esto se justifica por conexión: el tramo de la curva anterior a la primera visita a $\{x, y\}$ es un conjunto conexo que no contiene a x ni a y, así que queda dentro de una única arista abierta. De (∗) salen dos consecuencias:

(i) Si $p \in e_n$ está a distancia s de x a lo largo de la arista, entonces $d(p, \{x, y\}) = \min\{s,\ \ell_n - s\}$: para llegar a x o a y hay que recorrer la arista hasta uno de sus extremos.

(ii) Si p y q están en aristas distintas, entonces $d(p, q) \geq d(p, \{x, y\}) + d(q, \{x, y\})$: todo camino de p a q tiene que llegar primero a x o a y y, al final, salir de x o de y hacia q.

Veamos ahora las propiedades de X.

Es un espacio de longitud. Si un camino recorre un tramo de arista de largo s, sus extremos están a distancia d a lo sumo s. Por eso la longitud de un camino calculada

con d (como en la Sección 1) es menor o igual que la suma de los tramos recorridos. Tomando ínfimo sobre los caminos de p a q, el ínfimo de las longitudes es a lo sumo $d(p,q)$; como siempre es al menos $d(p,q)$, coinciden.

Es completo. Sea (q_k) una sucesión de Cauchy. Hay dos casos.

- Si infinitos términos están en una misma arista e_n (incluyendo sus extremos), esos términos forman una sucesión en e_n, que es compacta por ser imagen continua del intervalo $[0,\ell_n]$. Entonces tienen una subsucesión convergente.
- Si no, podemos elegir una subsucesión (q_{k_j}) cuyos términos están en aristas distintas y no son x ni y. Por (ii), $d(q_{k_j},q_{k_{j+1}}) \geq d(q_{k_j},\{x,y\})$, y como la sucesión es de Cauchy, el lado izquierdo tiende a 0. Luego $d(q_{k_j},\{x,y\}) \to 0$, y alguna subsucesión converge a x o a y.

En ambos casos, (q_k) es una sucesión de Cauchy con una subsucesión convergente, así que converge.

No es localmente compacto. Fijemos $0 < r \leq \frac{1}{2}$ y sea p_n el punto de e_n a distancia r de x a lo largo de la arista. Recorriendo la arista vemos que $d(x,p_n) \leq r$, de modo que todos los p_n están en la bola cerrada $\overline{B}(x,r)$. Por (i), $d(p_n,\{x,y\}) = \min\{r,\ \ell_n - r\} = r$, ya que $\ell_n - r > 1 - r \geq r$. Por (ii), si $n \neq m$,

$$d(p_n,p_m) \geq 2r.$$

Como en el Ejemplo 4.6, la sucesión (p_n) no tiene subsucesiones convergentes, así que $\overline{B}(x,r)$ no es compacta. Esto vale para todo $r \leq \frac{1}{2}$, de modo que x no tiene ningún entorno compacto.

No es geodésico. Por un lado, recorriendo la arista e_n se obtiene $d(x,y) \leq \ell_n$ para todo n, así que $d(x,y) \leq 1$. Por otro lado, sea γ cualquier curva de x a y. Consideremos el último instante t_0 en que γ pasa por x antes de su primera visita a y, y el instante $t_1 > t_0$ de esa primera visita a y. Entre t_0 y t_1 la curva no pasa ni por x ni por y, así que, por $(*)$, se queda dentro de una sola arista e_n, que recorre de punta a punta. Como para puntos cercanos de una misma arista la distancia d coincide con la distancia medida a lo largo de la arista, tomando particiones finas se obtiene

$$L(\gamma) \geq \ell_n = 1 + \tfrac{1}{n} > 1.$$

Esto muestra que $d(x,y) = 1$ (el ínfimo de los ℓ_n) y que *ninguna* curva de x a y tiene longitud $d(x,y)$. Por lo tanto, X no es geodésico.

Observación 4.8 (¿Dónde falla la demostración de Hopf–Rinow?). En el Ejemplo 4.7, las curvas γ_n que recorren e_n forman una sucesión minimizante: $L(\gamma_n) = 1 + \frac{1}{n} \to d(x,y)$. Reparametrizadas por longitud de arco, son uniformemente Lipschitz, igual que en la demostración del Teorema 4.4. Lo que falla es el paso de Arzelà–Ascoli: los puntos medios $\tilde{\gamma}_n(\frac{1}{2})$ están en aristas distintas y lejos de x y de y, así que, por (ii), están a distancia mayor que 1 entre sí, y no tienen subsucesiones convergentes. La sucesión minimizante existe, pero "se escapa" por infinitas direcciones distintas y no tiene límite (Figura 8).

La diferencia con el Ejemplo 4.6 no es que allí haya más compacidad en la topología de la métrica: en ninguno de los dos casos la hay. En el espacio de Hilbert las geodésicas

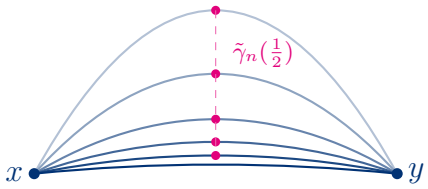


Figura 8. Los puntos medios $\tilde{\gamma}_n(\frac{1}{2})$ de las curvas minimizantes. Parecen acercarse a algo, pero están a distancia mayor que 1 entre sí: no hay ningún punto límite, porque la "arista de longitud 1" no existe.

existen por una razón distinta (allí se pueden escribir explícitamente). Más en general, para encontrar geodésicas en espacios que no son localmente compactos, una estrategia habitual es extraer límites de sucesiones minimizantes usando compacidad en una topología *más débil* que la de la métrica, para la cual la longitud (o el costo que se quiera minimizar) siga siendo semicontinua inferiormente. Esta idea va a reaparecer en los espacios de Wasserstein.

Referencias

[1] D. Burago, Yu. Burago, S. Ivanov, *A Course in Metric Geometry*, Graduate Studies in Mathematics **33**, American Mathematical Society, Providence, RI, 2001.
[2] M. R. Bridson, A. Haefliger, *Metric Spaces of Non-Positive Curvature*, Grundlehren der mathematischen Wissenschaften **319**, Springer-Verlag, Berlin, 1999.
[3] S. Cohn-Vossen, Kürzeste Wege und Totalkrümmung auf Flächen, *Compositio Math.* **2** (1935), 69–133.
[4] M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Progress in Mathematics **152**, Birkhäuser, Boston, 1999.
[5] H. Hopf, W. Rinow, Über den Begriff der vollständigen differentialgeometrischen Fläche, *Comment. Math. Helv.* **3** (1931), 209–225.
[6] J. R. Munkres, *Topology*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2000.