Riemannian Examples

Misha Gromov

September 14, 2026

Contents

1 (Quasi)divergences, Metrics λ -Lipschitz Maps Length Preserving Maps, Riemannian Spaces

A real valued function $D = D(x_1, x_2) \geq$ in two variables

$$D : X \times X \to \mathbb{R}$$

also written as $D(x_1||x_2)$ is called *quasidivergence* if it is positive,

$$D(x_1||x_2) \geq 0$$

and it vanishes on the diagonal $\{x_1 = x_2\} \in X \times X$,

$$x_1 = x_2 \Longrightarrow D(x_1||x_2) = 0$$

If D vanishes *if and only if* the points are equal,

$$D(x_1||x_2) = 0 \Longleftrightarrow x_1 = x_2$$

then D is called *divergence.*

Remark: Amari Regularity. In the original definition by Amari X is a smooth manifold and D must be twice or thrice continuously differentiable on the diagonal. `https://link.springer.com/book/10.1007/978-4-431-55978-8`

Example: Kullback–Leibler Divergence. Let $P = \{p_i\}, Q = \{q_i\}$ be strictly positive probability measures on a countable set I,

$$p_i, q_i > 0, i \in I, \sum_{i \in I} p_i = 1 \text{ and } \sum_{i \in I} q_i = 1.$$

Then

$$D_{QL}(P||Q) =_{def} \sum_{i\in I} p_i \log\frac{p_i}{q_i} = -\sum_{i\in I} p_i \log\frac{q_i}{p_i}$$

Since $\log x$ is a concave ($\frown$) function,

$$\sum_{i\in I} p_i \log\frac{q_i}{p_i} \le \log\sum_{i\in I} p_i\frac{q_i}{p_i} \le \log\sum_{i\in I} q_i = \log 1 = 0 \Longrightarrow D_{QL} \ge 0,$$

where the implication

$$D_{KL}(x_1||x_2) = 0 \Longrightarrow p_i = q_i, i \in I$$

follows from strict concavity of log.

Thus the Kullback–Leibler D_{KL} is indeed a divergence. .

Subexample: Shannon Entropy. The traditional definition of Shannon entropy of a probabilty measure on countable set reads

$$ent(P = \{p_i\}) =_{def} -\sum_{i\in I} p_i \log_2 p_i.$$

If I is a finite set of cardinality N, then "ent" is equal to $\log N$ minus the KL-divergence of P from *equidustributed* $Q_{uni\frac{1}{N}} = \{\frac{1}{N}, \frac{1}{N}, \frac{1}{N} ..., \frac{1}{N}\}$,

$$ent(P) = \log N - D_{KL}(P||Q_{uni\frac{1}{N}})$$

The basic property of the Shannon entropy is *multiplicative additivity*:

$$ent(P\times Q) = \{p_i\cdot q_j\}) = ent(P = \{p_i\}) + ent(Q = \{q_j\}).$$

In fact, the probability conditions $\sum_i p_i = 1$ and $\sum_j q_j = 1$ imply that

$$\sum_{(i,j)\in I\times J} p_i q_j \log p_i q_j = \left(\sum_j q_j\right)\sum_i p_i \log p_i + \left(\sum_i p_i\right)\sum_j q_j \log q_j =$$

$$= \sum_i p_i \log p_i + \sum_j q_j \log q_j$$

(This additivity property agrees with a physicist's definition of the entropy as log of the number of micro states, and suggests a more conceptual mathematical definition, see https://www.ihes.fr/ gromov/wp-content/uploads/2018/08/structre-serch-entropy-july5-2012.pdf)

The KL=divergence is also *multiplicatively additive*:

$$P = P_1 \otimes P_2, \qquad Q = Q_1 \otimes Q_2.$$

Then

$$D_{\mathrm{KL}}(P_1 \otimes P_2 \,\|\, Q_1 \otimes Q_2) = D_{\mathrm{KL}}(P_1\|Q_1) + D_{\mathrm{KL}}(P_2\|Q_2).$$

Indeed, set

$$P_1 = (p_i), \qquad Q_1 = (q_i), \qquad P_2 = (r_j), \qquad Q_2 = (s_j),$$

then

$$\begin{aligned} D_{\mathrm{KL}}(P_1 \otimes P_2 \,\|\, Q_1 \otimes Q_2) &= \sum_{i,j} p_i r_j \log \frac{p_i r_j}{q_i s_j} \\ &= \sum_{i,j} p_i r_j \left(\log \frac{p_i}{q_i} + \log \frac{r_j}{s_j} \right) \\ &= \sum_i p_i \log \frac{p_i}{q_i} + \sum_j r_j \log \frac{r_j}{s_j} \\ &= D_{\mathrm{KL}}(P_1 \| Q_1) + D_{\mathrm{KL}}(P_2 \| Q_2). \end{aligned}$$

Lipschitz Maps and Lipschitz Constants. A map between divergence or quasi-divergence spaces

$$f : (X, D_X) \to (Y, D_Y)$$

is *λ -Lipschitz,*[1] $\lambda \geq 0$, if

$$D_Y(f(x_1)||f(x_2)) \leq \lambda \cdot D_X(x_1||x_2) \text{ for all } x_1, x_2 \in X.$$

The minimal such $\lambda \geq 0$ is called Lipschitz constant and denoted $Lip(f)$. Clearly

$$Lip(f_1 \circ f_2) \leq Lip(f_1) \cdot Lip(f_2)$$

A map f is *λ-bi-Lipschitz* if

$$\lambda^{-1} D_X(x_1||x_2) \leq \lambda \cdot D_Y(f(x_1)||f(x_2)) \leq D_X(x_1||x_2) \text{ for all } x_1, x_2 \in X.$$

If a λ-bi-Lipschitz map f is bijective (one-to-one and onto) then the inverse map is also λ-bi-Lipschitz.

Metric Spaces. Divergence D is upgraded to (square of) *metric* if
$(\triangle)$ it (or $\sqrt{D}$) satisfies the *triangle inequality* $D(x_1, x_2) + D(x_2, x_3) \geq D(x_1, x_3)$
and
$(\cdot \leftrightarrow \cdot)$ it is *symmetric* $D(x_1, x_2) = D(x_2, x_1)$.

On Amari Regularity and non-Regularity. The triangle inequality is incompatible with Amari regularity (Prove it). For instance the metric $|x_1 - x_2|$ is not twice differential at $x_1, = x_2$. Yet the square of it is.

$\triangle$-Remark The $\triangle$-inequality is hard to verify for a given divergence function D; in most examples of such a D this inequality is enforced by the very definition of D

Edge-Metric Example. Let X be the vertex set of a connected graph and $d = d(e) > 0$ be a strictly positive positive function on the edges of X called *edge length*

The corresponding edge metric $D(x_1, x_2)$ is the infimum of the length of edge paths between x_1 and x_2. Equivalently it is the supremum of the functions,
which are equal to d on pairs of vertices joint by an edge of length d and
which satisfy the triangle inequality.

Euclidean/Pythagorean Example The triangle inequality

$$\sqrt{\sum_i (x_i - y_i)^2} + \sqrt{\sum_i (y_i - z_i)^2} \geq \sqrt{\sum_i (x_i - z_i)^2}$$

[1] I might be more logical to call it $\sqrt{\lambda}$-Lipschitz, since, at least locally, D behaves as $dist^2$.

however simple is by no means trivial.

Isometric and Contracting Maps. A map between metric spaces

$$f : (X, dist_X) \to (Y, dist_Y)$$

is *isometric* if the distance is preserved,

$$dist_Y(f(x_1), f(x_2)) = dist_X(x_1, x_2) \text{ for all } x_1, x_2 \in X,$$

In the Lipschitz terms this is the same as ($\lambda = 1$)-bi-Lipschitz.

Such maps are rare. Yet, these are abundant in the Euclidean world according to the following:

Main Euclidean theorem[2]: Partially defined isometric maps

$$\mathbb{R}^m \supset X \to \mathbb{R}^n$$

extends to isometric maps to all space

$$\mathbb{R}^m \to \mathbb{R}^n.$$

Distance decreasing, or 1-Lipschitz maps are called contracting.

Remarkably, these satisfy the extension property as well.

Kirszbraun Theorem. A partially defined contracting map

$$\mathbb{R}^m \supset X \to \mathbb{R}^n$$

always extends to a contracting map from all of the space

$$\mathbb{R}^m \to \mathbb{R}^n.$$

Corollary/Example/Lemma. Let $\{x_i\} \subset S^{n-1} \subset \mathbb{R}^n$, $i \in I$, be a subset in the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n\}_{\|x\|=1}$ and let $x \mapsto x' \in S^{n-1}$ be a contracting map $\{x_i\} \to \{x'_i\} \subset \mathbb{S}^{n-1}$. If the convex hull of $\{x'_i\}$ contains the origin $0 \in \mathbb{R}^n \supset S^{n-1}$, then the map $\{x'_i\} \to \{x'_i\}$ is an isometry.

Ask gpt: What are metric spaces X, where all partial contracting maps $X \supset Y \to X$ extend to such maps $X \to X$?

When is this extension property equivalent to that for isometric maps?

Beside the Euclidean, both properties are enjoyed by *hyperbolic spaces of constant negative curvature* and with minor restrictions by spherical spaces. (We prove that along with the Kirszbraun Theorem in lecture ??.

When is a Metric Space Riemannian? Let $(X, dist)$ be a metric space. The r-ball at $x_0 \in X$ is

$$B_{x_0}(r)\{x \in X\}_{dist(x,x_0)\leq r}.$$

[2]For $m, n = 3$, this must have been known for more than 23 centuries but I don't know where it was first formulated in the the present terms

The metric dist called infinitesimally n-Euclidean at x_0 if all sufficiently small balls $B_{x_0}(r), r > 0$, admit $\lambda-$bi-Lpschitz maps onto Euclidean domains (e.g Euclidaen r-ball)

$$f_r : B_{x_0}(r) \to U_r \subset \mathbb{R}^n$$

where

$$\lambda \to 1 \text{ for } r \to 0.$$

In other words small balls are are approximately isometric to the Euclidaen ones.

A good way to visualise the definition is to multiply *dist* by a large constant A (e.g. $A = 1/r$) and express the above by saying that the (scaled) pointed metric spaces $(A \times X, x_0) = (X, x_0, A \cdot dist)$ *Lipschitz converge* to $(\mathbb{R}^n, y_0 = 0, dist_{Eucl})$

Exercises. Work out a definition of Lipschitz convergence.

2. Define a distance between metric spaces using $\log \lambda$.

3. Suggests other vague ideas of mathematical concepts and ask chat gpt to make these precise.

Now define Riemannin in the following way:

(a) A metric is of "Riemannian origin" if it is infinitesimally Euclidean at all points.

(b) A metric function $dist(x_1, x_2)$ is of "Riemannian origin" if its square $dist(x_1, x_2)^2$ is twice differential on the diagonal $\{x_1 = x_2\}$

From Divergence (e.g a metric) to Riemannin Tensor (Quadratic form) g

The definition of a *Riemannin metric tensor* is usually preceded by the definition of an abstract smooth manifold X which corresponds in physics to an empty space. This makes it awkward.

Alternatively, (quasi)following Riemann start with $X \subset \mathbb{R}^n$ – an open subset and

$$g = \{g_{ij}\} = \{g_{ij}(x_1, ..., x_n\}$$

is a function with values in square symmetric matrices representing positive definite quadratic n-forms written as

$$g = \sum_{i,j} g_{ij} dx_i dx_j.$$

Here the coordinates x_i are regarded as functions on X and their differentials dx_i are liner forms.

The key word is still missing: these are *differential* forms both quadratic nd linear ones : they are not functions on some liner space but the form $g_{ij}(x_1, ..., x_n)$ is a bilinear function on the tangent space $T_{x_1,...,x,n}(X)$.

In the coordinate terms this "differential" corresponds to the transformation rule under changes of coordinates.

$$g = g_{ij}(x(y)) \frac{\partial x_i}{\partial y_\alpha} \frac{\partial x_j}{\partial y_\beta} dy_\alpha dy_\beta.$$

In the modern language, g is defined invariantly to start with as a family $g(x)$ $x \in X$ of quadratic forms on the *tangent* spaces $T_x(X)$, where "tangent" makes the forms "differential".

Intuitively, however, if $X = \mathbb{R}^n$ or if it is an open subset in $\mathbb{R}^n$ it is hard not to confuse the tangent spaces from $\mathbb{R}^n \supset X$ itself since these are "almost naturally" isomorphic. This is much better seen if $X \subset \mathbb{R}^{n+1}$ is a smooth hypersurface.

From Riemannin metric tensor $g(x)$ to the distance function $dist(x_1, x_2)$
(A) Riemannian g assigns *length* to all smooth curves in X as follows.

$$g = \sum_{i,j} g_{ij}(x)\, dx^i\, dx^j,$$

$$x(t) = \big(x^1(t), \ldots, x^n(t)\big), \qquad a \leq t \leq b,$$

The length of $x(t)$ is $\int_a^b \left(\sum_{i,j} g_{ij}(x(t)) \frac{dx^i}{dt} \frac{dx^j}{dt}\right)^{1/2} dt$.

Then $dist(x_1, x_2)$ is defined as the infimum of lengths of the curves from x_1 to x_2

Hessian Geometries Let $f(x) \geq 0$ be a *positive* C^2-smooth (twice continuously differential) function on a smooth manifold X, which vanishes at the point $x_0 \in$.

Then the second differential of f at x_0 defines a positive semi-definite quadratic form at x_0, that is on the tangent space $T_{x_0}(X)$: the second derivative of f restricted to a smooth curve $[0, 1] \ni t \mapsto x \in X$ issuing from x_0 with speed $\tau \in T_{x_0}$ depends on τ but not on a particular curve,

$$(d^2 f(x = x_0))(\tau, \tau) = \frac{d^2 f(x(t = 0))}{dt^2}.$$

Corollary/Example. If $X = \mathbb{R}^n$ and $x_0 = 0$, then $|f(x) - f(-x)| = o(\|x\|^2)$ for $\|x\| \to 0$.

Next let $D(x_1||x_2$ be a C^2-smooth quasi-divergence. Then the second differential is defined as a quadratic form, on the tangent bundle $T(X \times X)$ restricted to the diagonal $X_\Delta \subset X \times X$, call it $g = g_D$.

Since D vanishes on X_Δ the forms factors through

$$T(X_\Delta) \subset T(X \times X)|X_\Delta$$

and defines a form on the quotient bundle

$$T_\Delta = (T(X \times X)|X_\Delta)/T(X_\Delta).$$

Since the the bundle $T_\Delta \to X_\Delta$ is equal (canonically isomorphic/equivalent) to $T(X) \to X$ the form $g = g_\Delta$ becomes a quadratic form on X.

Since $D \geq 0$ this g is positive semidefinite. If X is a domain in a Euclidean space and the function $D(x_1||x_2$ is strictly (locally) convex with respect to x_1 near the diagonal

KL-Example

For two distributions p, q,

$$D_{\mathrm{KL}}(p\|q) = \sum_i p_i \log \frac{p_i}{q_i}.$$

If $q = p + dp$, then

$$D_{\mathrm{KL}}(p\|p + dp) = \frac{1}{2} \sum_i \frac{(dp_i)^2}{p_i} + O(|dp|^3).$$

and

$$g_{KL} = \sum_i \frac{(dp_i)^2}{p_i} = \frac{\sum_i \partial^2(p_i \log p_i)}{\partial^2 p_i}.$$

Thus g_{KL} equal to the *Fisher Metric*, that is the Hessian of the minus ($\log_e$) entropy

$$\Phi = \Phi(p_i) = \sum_i \partial^2(p_i \log p_i).$$

(This Hessian explicitly depends on the affine structure in the simplex, while in the definition of g_{KL} one uses only the smooth structure.)

The self map of the positive cone $\mathbb{R}^n_+ = \{x_i > 0\} \subset \mathbb{R}^n$ for $x_i \mapsto p_i = x_i^2$

$$\left(\{x_i > 0\}, 4\sum_i dx_i^2\right) \to \left(\{p_i > 0\}, \sum_i \frac{(dp_i)^2}{p_i}\right)$$

is an isometry, since $dp_{=}2x_idx_i$, and

$$\frac{(dp_i)^2}{p_i} = \frac{4x_i^2(dx_i)^2}{x_i^2} = 4(dx_i)^2.$$

$dp_{=}2x_idx_i$,

Riemanni Families Let g_t, $t \in [a,b]\mathbb{R}$ also written as dX_t^2 be a family of Riemannin metrics on X. Then $g^{\rtimes} = dt^2 + gX_t^2$ is a Riemannin metric on $X \times [a,b]$

Warped Product Example. Given a metric $g = dX^2$ on X and a positive function $c(t) > 0$ one gets a metric $dt^2 + c^2(t)dX^2$ on $X \times [a,b]$

Euclidean, Spherical, Hyperbolic

Let $\mathbf{S}_{n-1}$ be the unit sphere with its spherical Riemannin metric $d\mathbf{S}^2_{n-1}$.

Then the Euclidean metric on $\mathbb{R}^n \setminus \mathbf{0} = (\mathbb{R}_+ \setminus 0) \times \mathbf{S}_{n-1}$ is $dr^2 + r^2\, d\mathbf{S}^2_{n-1}$.

Space	Metric	Radial factor $c(r)$
Euclidean $\mathbb{R}^n$	$dr^2 + r^2\, d\mathbf{S}^2_{n-1}$	r
Sphere S^n (radius 1)	$dr^2 + \sin^2 r\, d\mathbf{S}^2_{n-1}$	$\sin r$
Hyperbolic $\mathbb{H}^n$	$dr^2 + \sinh^2 r\, d\mathbf{S}^2_{n-1}$	$\sinh r$

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

Exercise. Let $X = \mathbb{R}^{n-1}$ with the Euclidean $dX^2 = \sum_{i=1,\ldots n-1} dx_i^2$

Show that the metric $dt^2 + e^{2t}dX^2$ is the hyperbolic one.

SOLUTION BY CHAT GPT

Let

$$x = (x_1, \ldots, x_n), \qquad |x|^2 = \sum_{i=1}^n x_i^2.$$

Consider the hyperboloid

$$-X_0^2 + X_1^2 + \cdots + X_{n+1}^2 = -1, \qquad X_0 > 0,$$

with the metric induced from

$$-dX_0^2 + \sum_{i=1}^{n+1} dX_i^2.$$

Define

$$X_i = e^t x_i, \qquad i = 1, \ldots, n,$$

and

$$X_0 = \frac{1}{2}\left(e^t(1+|x|^2) + e^{-t}\right),$$

$$X_{n+1} = \frac{1}{2}\left(e^t(1-|x|^2) - e^{-t}\right).$$

Then

$$X_0 + X_{n+1} = e^t,$$

and

$$X_0 - X_{n+1} = e^{-t} + e^t|x|^2.$$

Hence

$$-X_0^2 + X_{n+1}^2 = -(X_0 + X_{n+1})(X_0 - X_{n+1}) = -1 - e^{2t}|x|^2.$$

Since

$$\sum_{i=1}^{n} X_i^2 = e^{2t}|x|^2,$$

we obtain

$$-X_0^2 + \sum_{i=1}^{n+1} X_i^2 = -1.$$

Now put

$$A = X_0 + X_{n+1} = e^t, \qquad B = X_0 - X_{n+1} = e^{-t} + e^t|x|^2.$$

Since

$$X_0 = \frac{A+B}{2}, \qquad X_{n+1} = \frac{A-B}{2},$$

we have

$$-dX_0^2 + dX_{n+1}^2 = -dA\,dB.$$

Furthermore,

$$dA = e^t\,dt,$$

and

$$dB = \left(-e^{-t} + e^t|x|^2\right)dt + 2e^t\sum_{i=1}^{n} x_i\,dx_i.$$

Therefore

$$-dA\,dB = \left(1 - e^{2t}|x|^2\right)dt^2 - 2e^{2t}\sum_{i=1}^{n} x_i\,dx_i\,dt.$$

Also,

$$dX_i = e^t(dx_i + x_i\, dt),$$

so

$$\sum_{i=1}^{n} dX_i^2 = e^{2t}\sum_{i=1}^{n} dx_i^2 + 2e^{2t}\sum_{i=1}^{n} x_i\, dx_i\, dt + e^{2t}|x|^2 dt^2.$$

Adding,

$$-dX_0^2 + \sum_{i=1}^{n+1} dX_i^2 = dt^2 + e^{2t}\sum_{i=1}^{n} dx_i^2.$$

On the same hyperboloid, introduce geodesic polar coordinates

$$X_0 = \cosh r,$$

and

$$(X_1, \ldots, X_{n+1}) = \sinh r\, \omega, \qquad \omega \in S^n.$$

Then

$$dX_0 = \sinh r\, dr,$$

and

$$\sum_{i=1}^{n+1} dX_i^2 = \cosh^2 r\, dr^2 + \sinh^2 r\, dS^2.$$

Therefore

$$-dX_0^2 + \sum_{i=1}^{n+1} dX_i^2 = -\sinh^2 r\, dr^2 + \cosh^2 r\, dr^2 + \sinh^2 r\, dS^2,$$

and hence

$$-dX_0^2 + \sum_{i=1}^{n+1} dX_i^2 = dr^2 + \sinh^2 r\, dS^2.$$

Thus

$$\boxed{dt^2 + e^{2t}\sum_{i=1}^{n} dx_i^2 = dr^2 + \sinh^2 r\, dS^2.}$$

The coordinate transformation is given by

$$\cosh r = \frac{1}{2}\left(e^t(1+|x|^2) + e^{-t}\right),$$

$$\sinh r\, \omega_i = e^t x_i, \qquad i = 1, \ldots, n,$$

and

$$\sinh r\, \omega_{n+1} = \frac{1}{2}\left(e^t(1-|x|^2) - e^{-t}\right).$$

2 Riemannin Isometric

Let

$$f = (f^1, \dots, f^n) : \mathbb{R}^m \longrightarrow \mathbb{R}^n, \qquad m \leq n.$$

We equip both spaces with their standard Euclidean metrics.

The map f is an isometric immersion if, for every $X, Y \in T_x\mathbb{R}^m$,

$$\langle df_x(X), df_x(Y) \rangle_{\mathbb{R}^n} = \langle X, Y \rangle_{\mathbb{R}^m}.$$

Using the standard coordinates

$$x = (x^1, \dots, x^m),$$

this condition is equivalent to

$$\left\langle \frac{\partial f}{\partial x^i}, \frac{\partial f}{\partial x^j} \right\rangle_{\mathbb{R}^n} = \delta_{ij}, \qquad 1 \leq i, j \leq m.$$

Since

$$f = (f^1, \dots, f^n),$$

we have

$$\frac{\partial f}{\partial x^i} = \left(\frac{\partial f^1}{\partial x^i}, \dots, \frac{\partial f^n}{\partial x^i} \right).$$

Therefore the isometric immersion equations are

$$\boxed{\sum_{\alpha=1}^{n} \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\alpha}{\partial x^j} = \delta_{ij}, \qquad 1 \leq i, j \leq m.}$$

Equivalently,

$$\boxed{\sum_{\alpha=1}^{n} \partial_i f^\alpha \, \partial_j f^\alpha = \delta_{ij}.}$$

If Df denotes the $n \times m$ Jacobian matrix

$$Df = \left(\frac{\partial f^\alpha}{\partial x^i} \right),$$

then the same equations can be written as

$$\boxed{(Df)^T Df = I_m.}$$

Thus the m vectors

$$\frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^m}$$

form an orthonormal system in $\mathbb{R}^n$.

In particular,

$$\left|\frac{\partial f}{\partial x^i}\right|^2 = 1$$

for every i, and

$$\left\langle \frac{\partial f}{\partial x^i}, \frac{\partial f}{\partial x^j} \right\rangle = 0, \qquad i \neq j.$$

Question to chat gpt What is the minimal N such that all *flat Riemannian* manifolds X, i.e locally isometric to $\mathbb{R}^n$ admit Riemannin isometric immersions to $\mathbb{R}^N$?

More specifically are there such X homeomorphic to $n-tori$ which admit no Riemannin isometric immersions to $\mathbb{R}^2n$?