Geometry of Gaussian Measures

Joseph Anderson

The Ohio State University

andejose@cse.ohio-state.edu

24 March 2014

Curves

Let (X, d) be a metric space.

Definition

Let $\gamma : [0,1] \to X$ be continuous. We say that γ is a *curve*, and the *length* of γ is

$$L_d(\gamma) = \sup_{0=t_1<t_2<\cdots<t_N=1} \sum_{i=1}^{N} d(\gamma(t_i), \gamma(t_{i+1})).$$

The curve γ is said to be *rectifiable* if its length is finite.

Length Spaces

Definition

Let $x, x' \in X$ and let $\Gamma(x, x')$ be the family of curves joining x and x'. The *intrinsic metric* on X is defined as

$$d^*(x, x') = \inf(L_d(\gamma) \mid \gamma \in \Gamma(x, x')).$$

If $\Gamma(x, x') = \emptyset$, we say that $d^*(x, x') = \infty$.

- If $d = d^*$, then we say that d is intrinsic and we call (X, d^*) a *path metric space* or *length space*.
- One calls (X, d^*) *geodesic* if for any pair of points x, x' there exist $\gamma \in \Gamma(x, x')$ so that $L_d(\gamma) = d(x, x')$.

Space of Measures

Let $\mathcal{P}_2^{ac}(\mathbb{R}^d)$ be the set of all absolutely continuous measures on $\mathbb{R}^d$ with finite second moment.

- Absolutely continuous: $\mu << \lambda$ if $\lambda(A) = 0 \implies \mu(A) = 0$
- Finite second moment: $\int_{\mathbb{R}^d} d(x, x_0) d\mu(x) < \infty$ for all x_0.

Gaussian Measures

Recall the Gaussian measure on $\mathbb{R}^d$, $\phi_{\mu,\Sigma}$, where

$$\phi_{\mu,\Sigma}(A) = \frac{1}{\sqrt{\det(2\pi\Sigma)}^d} \int_A \exp\left(\frac{-1}{2}\langle x-\mu, \Sigma^{-1}(x-\mu)\rangle\right) d\lambda_n(x).$$

- Let $(\mathcal{N}^d, d_{\mathcal{W},2}^{\mathbb{R}^d})$ be the space of all Gaussian measures on $\mathbb{R}^d$ with the (restriction of the) 2-Wasserstein distance.
- Note: Gaussian measures are square integrable and absolutely continuous with respect to the Lebesgue measure, so $\mathcal{N}^d \subset \mathcal{P}_2^{ac}$

Some Basic Questions about $\mathcal{N}$

Let $\phi_1, \phi_2 \in \mathcal{N}^d$.

- Is it true that we can find a measure $\mu \in \mathcal{M}(\phi_1, \phi_2)$?
- How about $\mathcal{N}^{d^2} \cap \mathcal{M}(\phi_1, \phi_2)$?
- Does $\mathcal{N}^d$ have recognizable structure?

One-Dimensional Warm Up

Let $\phi_1, \phi_2 \in \mathcal{N}_1$ with mean and variance μ_1, σ_1^2 and μ_2, σ_2^2, respectively.

- Couplings exist, and are Gaussian on $\mathbb{R}^2$
- We can assume that $\mu_1 = \mu_2 = 0$
- We can compute

$$\begin{aligned} d_{\mathcal{W},2}(\phi_1, \phi_2) &= \left(\inf_{\mu \in \mathcal{M}(\phi_1, \phi_2)} \iint |x - y|^2 d\mu(x, y) \right)^{1/2} \\ &= |\sigma_1 - \sigma_2|. \end{aligned}$$

- Furthermore, this even forms a length space

Known Results

Theorem ([GM96])

Let μ, ν be Borel probability measures on $\mathbb{R}^d$. Then

1. *there exists a convex function ψ on $\mathbb{R}^d$ whose gradient $\nabla\psi$ pushes μ forward to ν*
2. *the gradient of ψ is determined up to μ-measure 0*
3. *the measure $\pi = (\mathrm{id} \times \nabla\psi)_{\#}\mu$ is optimal*
4. *π is the only optimal measure in $\mathcal{M}(\mu, \nu)$ unless $d_{\mathcal{W},2}(\mu, \nu) = +\infty$*

Theorem ([GS84])

For $\phi_{m,V}, \phi_{n,U} \in \mathcal{N}^d$, we have

$$d_{\mathcal{W},2}^2 = \|m - n\|^2 + \mathrm{tr}\,(V) + \mathrm{tr}\,(U) - 2\mathrm{tr}\left(U^{\frac{1}{2}} V U^{\frac{1}{2}}\right)^{\frac{1}{2}}$$

Geometry of $\mathcal{N}^d$ and Applications

- Proved by [GM96] that $\mathcal{N}_0^d$ (mean 0 Gaussian measures) is geometrically convex as a subspace of $\mathcal{P}_2^{ac}(\mathbb{R}^d)$.
- The sectional curvature of $\mathcal{N}_0^d$ can be computed, c.f. [Tak08]
- Known curvature can help learning algorithms which rely on regularization
- Distance between "soft" shapes with Gaussian-like geodesic distances becomes easily computable and interpolated

References

Wilfrid Gangbo and Robert J McCann.
The geometry of optimal transportation.
Acta Mathematica, 177(2):113–161, 1996.

Clark R Givens and Rae Michael Shortt.
A class of wasserstein metrics for probability distributions.
The Michigan Mathematical Journal, 31(2):231–240, 1984.

Asuka Takatsu.
On wasserstein geometry of the space of gaussian measures.
2008.

Harmonic Soft Maps and the Laplace-Beltrami Operator

Overview Presentation

Alfred Rossi

The Ohio State University

March 23, 2014

Introduction

- Project goal is to discover and understand the landscape of ideas surrounding harmonic maps and the Laplace-Beltrami operator while exploring related applications.
- In particular
  - The variational formulation of the Laplace-Beltrami operator
  - The generalization of Harmonic Maps from functions between manifolds to soft maps
  - The interplay between these ideas

Laplace-Beltrami Operator

Throughout this presentation M_S and M_T will denote two smooth, compact Riemannian manifolds with gradient operators ∇_S, ∇_T, respectively.

Definition

Let g_S denote the metric for M_S, then the Laplace-Beltrami Operator at a point $x \in M_S$ is given by $\Delta_{g_S} u \equiv \mathrm{div} \cdot \mathrm{grad} u$.

- The Laplace-Beltrami Operator on a Riemannian Manifold is determined by the metric.
- The converse direction holds too, by intermediate construction of the heat kernel

The Heat Equation

Consider a heat diffusion process on a manifold. The goal is to find the amount of heat at each location of the manifold $u(x, t) : M_S \times \mathbb{R}^+ \to \mathbb{R}$ for all $t > 0$, given an initial distribution $u(x, 0)$.

Definition

$$\Delta_{g_S} u(x, t) = -\frac{\partial u(x, t)}{\partial t}$$

Solving the Heat Equation

Intuitively, suppose we had a function $K : M_S \times M_S \times \mathbb{R}^+ \to \mathbb{R}$ which given two points $x, y \in M_S$ encodes the proportion of heat at position y from position x after a time Δt. Then,

Fact

$$u(x,t) = \int_{M_S} K(x,y,t)\, u(y,0)\, dy$$

What would such a K look like?

The Heat Kernel

Definition

Let λ_n, ϕ_n be the eigenvalues and eigenfunctions of Δ_{g_S}. The heat kernel $K(x, y, t) \in C^\infty(M_S \times M_S \times \mathbb{R}^+)$ is given by

$$K(x, y, t) = \sum_{n=0}^{\infty} e^{-\lambda_n t} \phi_n(x) \phi_n(y)$$

- Analogous to Fourier analysis on the manifold
- Contains all of the information regarding the structure of the manifold
  - can be used to determine the metric

Soft Maps Between Surfaces [SNBBcG]

Definition

A *Soft Map* from M_S to M_T is a function assigning to each point $x \in M_S$ a probability distribution μ_x over M_T.

- Let $\mathcal{U} \subseteq M_T$ then $\mu_x(\mathcal{U})$ encodes the probability that x maps into $\mathcal{U}$. In particular it is required that:
  - $\mu_x(\mathcal{U}) \in [0, 1]$
  - $\mu_x(M_T) = 1$
- Soft maps generalize point to point maps.

Harmonic Maps [SGB]

Definition

The *Dirichlet Energy* of a map $\phi : M_S \to M_T$ is given by

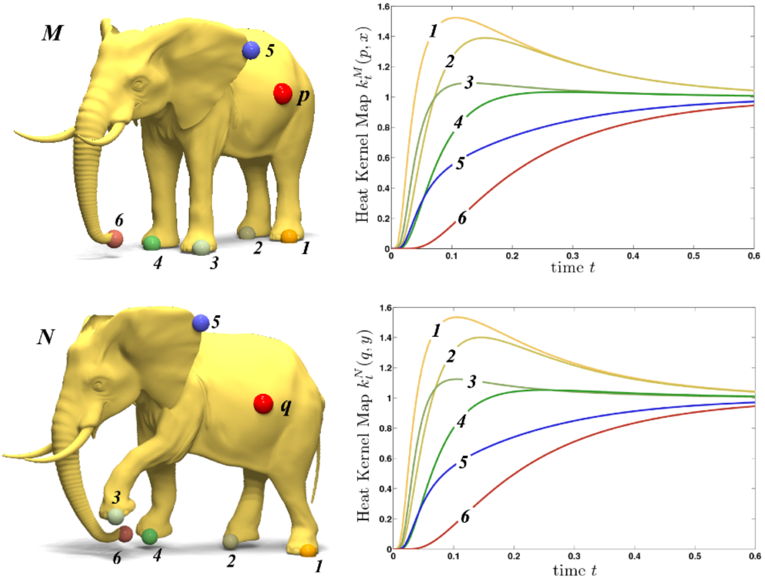
$$\mathcal{E}_D(\phi) \equiv \int_{M_S} \|\nabla_S \phi(x)\|^2 \, dx$$

- Can think of this as a measure of the "intrinsic stretching" of ϕ.
- Maps which extremize this functional are known as *harmonic maps.*
- Project goal: Investigate and understand how to generalize this notion to soft maps.

One Point Isometric Matching by Heat Kernel [OMMG]

- Gives a method to recover an isometry when given only a single pair of corresponding points between two geodesically identical meshes
- Works by exploiting the invariance of the heat kernel under isometry
  - Given a pair of isometric manifolds M, N and a pair of points $p \in M$, $q \in M$ such that for some isometry $\psi : M \to N$, $\psi(p) = q$. One can associate to every point on $x \in M$ a signature $k_t^M(p, x)$ and likewise for any $y \in N$ a signature $k_t^N(q, y)$.
  - Related points can be identified by matching signatures
  - Both the initial correspondence and isometry can be found efficiently in practice
- Very general. That is, it does not depend genus of the surface, or dimension

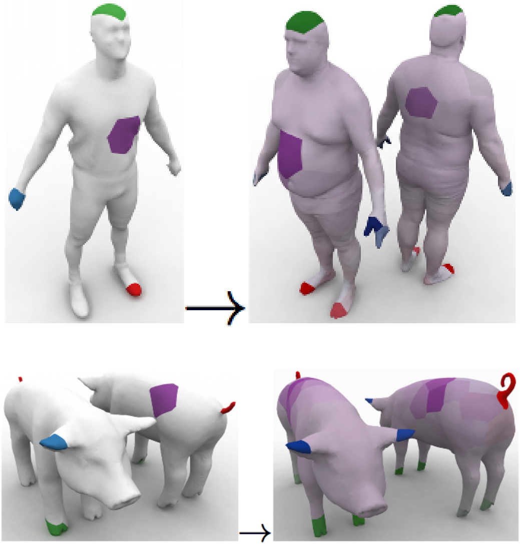
One Point Isometric Matching by Heat Kernel [OMMG]



Applications of Soft Maps [SNBBcG]

- Models M_1, M_2 are descritized into patches corresponding to geodesic voroni cells
- A descriptor $\phi_s(U) \in \mathbb{R}^k$ is associated with each patch $U \subset M_s$, $s \in \{1, 2\}$ that encodes some information about its geometry.
- A solver numerically finds a soft correspondence A which minimizes the energy $E(A) = E_\phi(A) + \lambda E_{cont}(A) + \beta E_s(A)$ where
  - E_ϕ ensures the corresponding patches have similar geometry
  - E_{cont} attempts to keep local patches local in the correspondence by applying a cost based on *GEMD*
  - E_s is a sharpness term that tries to mediate tradeoff between E_ϕ and E_{cont} when there is an exact symmetry.

Applications of Soft Maps [SNBBcG]



Bibliography

- Ovansjanikov, Merigot, Mémoli, and Gubias. *One Point Isometric Matching with the Heat Kernel*
- Solomon, Gubias, Butscher. *Dirichlet Energy for Analysis and Synthesis of Soft Maps*
- Zeng, Guo, Luo, Gu. *Discrete Heat Kernel Determines Discrete Riemannian Metric*
- Crane, Weischedel, and Wardetzky. *Geodesics in Heat*
- Eells and Sampson. *Harmonic Mappings of Riemannian Manifolds*
- Mémoli. *Spectral Gromov-Wasserstein Distances for Shape Matching*
- Solomon, Nguyen, Butscher, Ben-Chen, Gubias. *Soft Maps Between Surfaces*

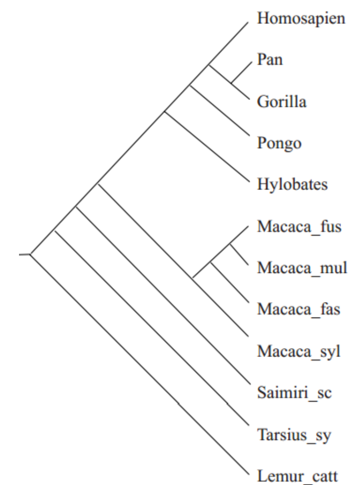
Averaging Phylogenetic Trees

Presenter: Suyi Wang

Phylogenetic Tree Space

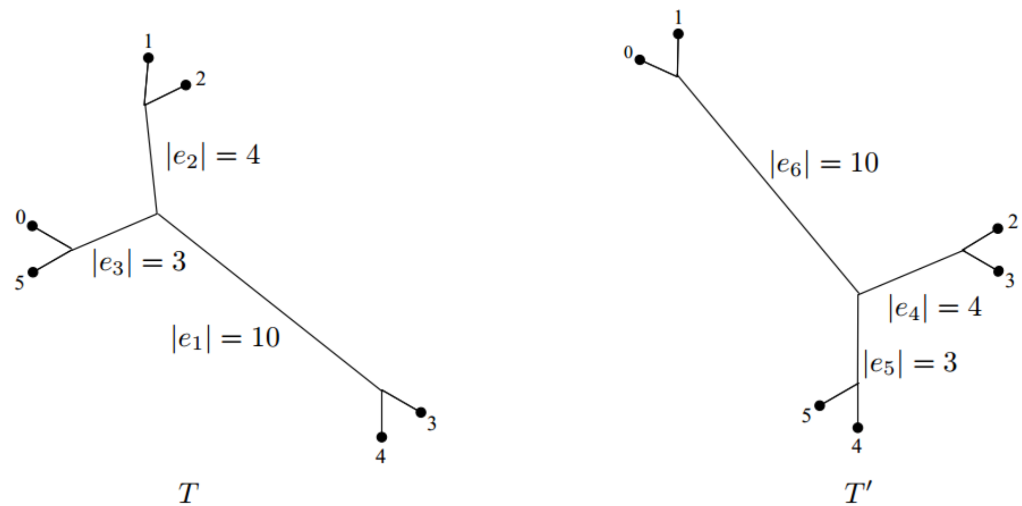
- A phylogenetic n-tree is:
  - A Tree with n leaves
  - Leaves: represent distinguishable species
  - Interior vertices: #degree >= 3
  - Edges: weighted
    - Max: 2*n – 1 edges

```
'Lemur_catta'      AAGCTTCATAGG
'Tarsius_syrichta' AAGTTTCATTGG
'Saimiri_sciureus' AAGCTTCACCGG
'Macaca_sylvanus'  AAGCTTCTCCGG
'Macaca_fascicul.' AAGCTTCTCCGG
'Macaca_mulatta'   AAGCTTTTCTGG
'Macaca_fuscata'   AAGCTTTTCCGG
'Hylobates'        AAGCTTTACAGG
'Pongo'            AAGCTTCACCGG
'Gorilla'          AAGCTTCACCGG
'Pan'              AAGCTTCACCGG
'Homo_sapiens'     AAGCTTCACCGG
```



Phylogenetic Tree Space

- Splits:
  - Associates with an edge e
  - A partition (Xe, Xe*) of leaves

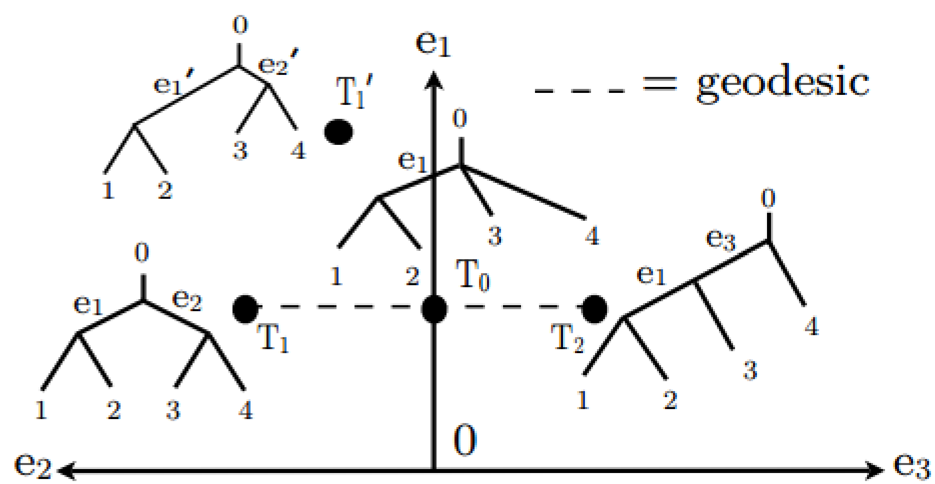


splits
$e_1 : \{0,1,2,5\} \,|\, \{3,4\}$
$e_2 : \{0,3,4,5\} \,|\, \{1,2\}$
$e_3 : \{0,5\} \,|\, \{1,2,3,4\}$

splits
$e_4 : \{0,1,4,5\} \,|\, \{2,3\}$
$e_5 : \{0,1,2,3\} \,|\, \{4,5\}$
$e_6 : \{0,1\} \,|\, \{2,3,4,5\}$

Phylogenetic Tree Space

- Trees are uniquely defined by compatible set of splits [Semple 03]
  – Suggests a tree space
  – Axis – length of edges
  – 2n – 1 orthants



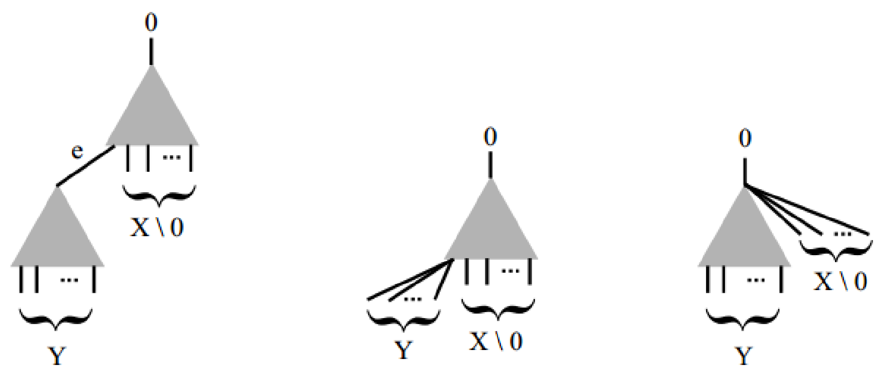
Geodesic distance

- Computing geodesic distances in tree space [Owen 07]
  - Shortest path from T1 to T2
  - Sum (sub path in each orthant)
  - Unique [Billera 01]

- Same orthant
  - Euclidean distance

Geodesic distance

- Trees share a common edge
  - [Vogtmann 07] [Owen 08]

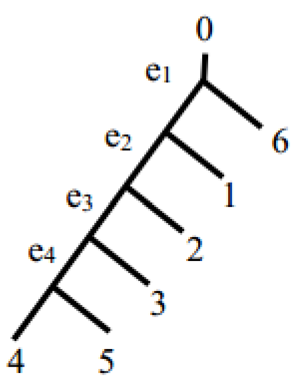


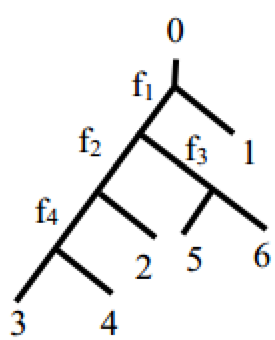
(a) Tree T_i. (b) Tree T_i^X. (c) Tree T_i^Y.

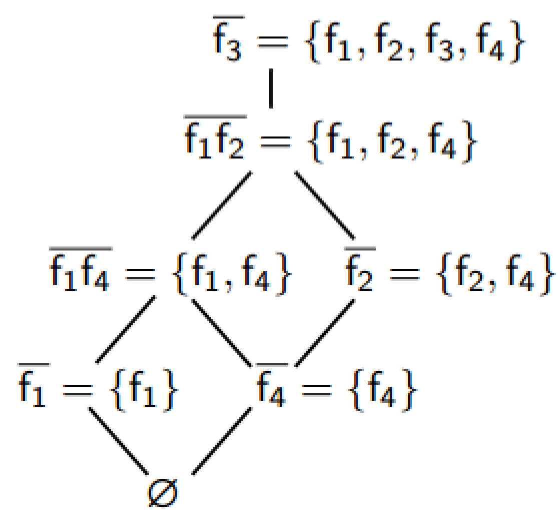
$$d(T_1, T_2) = \sqrt{d(T_1^X, T_2^X)^2 + d(T_1^Y, T_2^Y)^2 + (|e|_{T_1} - |e|_{T_2})^2}$$

Geodesic distance

- Trees share no common edge
  – A path space: $K(\Sigma_1, \Sigma_2)$
  – Try all possible orthants series.
  – A dynamic programming







Averaging in tree space

- Tree space T={T1, T2, ..., Tr}
- Variance S(X, T)
  - Sum of squares distances from X to each tree in T
- Mean
  - X that minimize the variance

Averaging in tree space

- Iterative approach [Strum 03]
  - Random sample on T for k times – {T1…Tk}
  - Current result u_i, i=[1, k]
  - u_i+1 = 1/(k+1) from u_i to Ti
  - Converge to mean

- A descent method [Millier 12]
  - To be continued…