Tutorial:
Geometric central limit theorems on singular spaces

Ezra Miller

Duke University, Department of Mathematics
and Department of Statistical Science

`ezra@math.duke.edu`

joint with Jonathan Mattingly (Duke)
Do Tran (Deutsche Bank (was: Göttingen))

`http://arxiv.org/abs/2311.09455`
`09454`
`09453`
`09451`

Interactions of Statistics and Geometry (ISAG) II

National University of Singapore

14 – 16 October 2024

Syllabus

1. overview (today)
2. geometry and measure on CAT(κ) spaces [`arXiv:...3, `§1-§2]
    - CAT(κ) spaces
    - angles and angular pairing
    - tangent cones
    - localized measures
    - escape and fluctuating cones
    - hulls
3. shadows and tangential collapse [`arXiv:...1`]
    - shadows [`arXiv:...3, `§3-§4]
    - radial transport
    - limit log
    - tangential collapse
4. convergence to Gaussian objects [`arXiv:...4`], [`arXiv:...5, `§6.1]
    - random tangent fields and their CLT
    - stratified Gaussians
5. central limit theorems via escape [`arXiv:...5, `§4-§6]
    - escape vectors
    - continuous mapping theorem

Outline

Linear Central Limit Theorem

Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

to population mean $\bar{\mu} = \int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem

Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

to population mean $\bar{\mu} = \displaystyle\int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem

Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

to population mean $\bar{\mu} = \int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem

Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

to population mean $\bar{\mu} = \int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem

Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

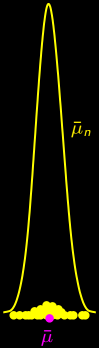
to population mean $\bar{\mu} = \int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem



Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$

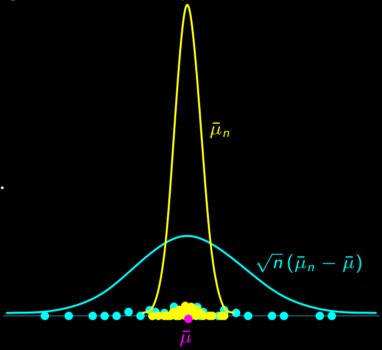
to population mean $\bar{\mu} = \displaystyle\int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem



Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

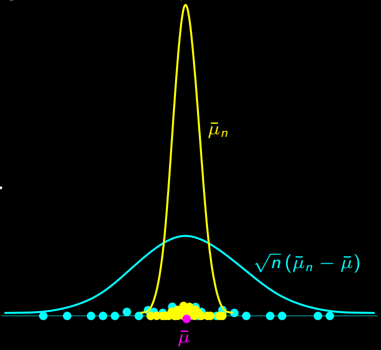
to population mean $\bar{\mu} = \int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Linear Central Limit Theorem



Input

- vector space $\mathbb{R}^d$
- independent random variables $X_1, X_2, \ldots$
- distributed according to μ

Compare empirical mean $\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

to population mean $\bar{\mu} = \int x\,\mu(dx)$

Law of Large Numbers (LLN): $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

Central Limit Theorem (CLT): $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution, for random variable $N(0,\Sigma)$

- Gaussian
- centered at 0
- same covariance Σ as μ.

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], ...
- products and mixtures of these: unions of subspaces, spheres, tori, ...
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

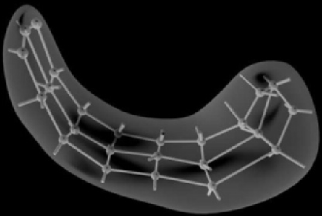
Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Hippocampus surfaces

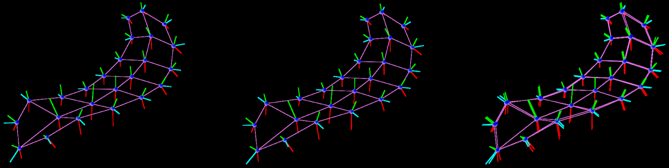
Skeletal representation



Fletcher, Pizer, and Joshi 2006

Dataset

276 skeleta of hippocampus surfaces:



courtesy S. Pizer

each datapoint $\in \mathbb{R}_+^{67} \times S^{68} \times (S^2)^{66}$, dim 267 in $\mathbb{R}^{334}$.

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: “Big Data” often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit “1”
- persistence diagrams: topological summaries of
  - datasets
  - data objects

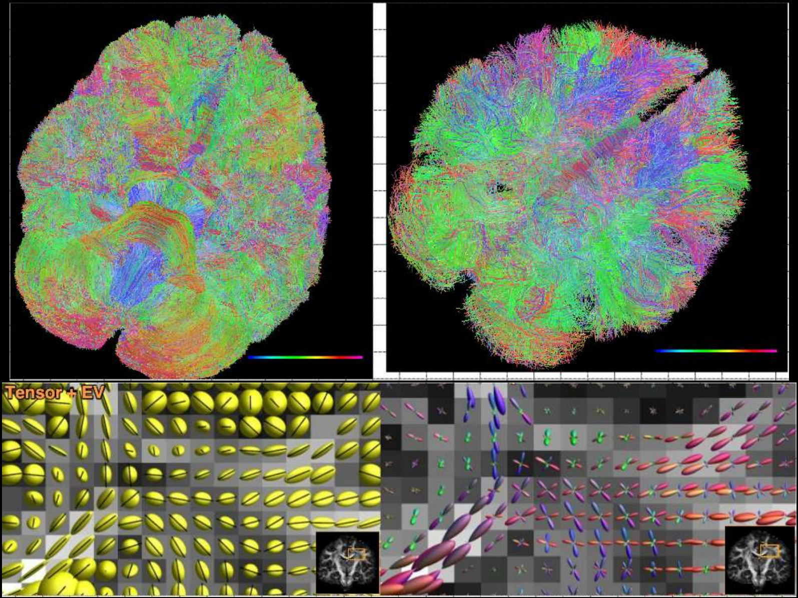
Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Streamlines from Diffusion Tensor Imaging



courtesy Zhengwu Zhang

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
    - wind direction
    - knee or elbow motion
- directions: points on a sphere
    - wrist or ankle motion
    - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
    - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
    - phylogenetic orange [Kim 2000],
    - tropical grassmannian [Speyer–Sturmfels 2004],
    - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
    - e.g., the digit "1"
- persistence diagrams: topological summaries of
    - datasets
    - data objects

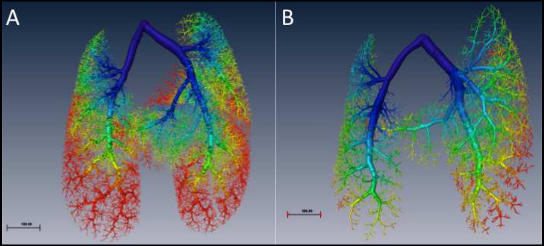
Nonlinear data

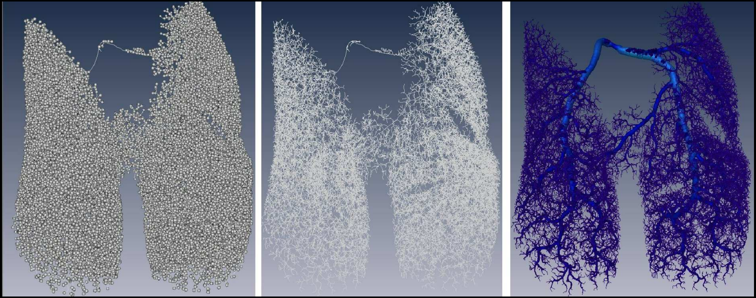
Initial rationale: “Big Data” often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit “1”
- persistence diagrams: topological summaries of
  - datasets
  - data objects

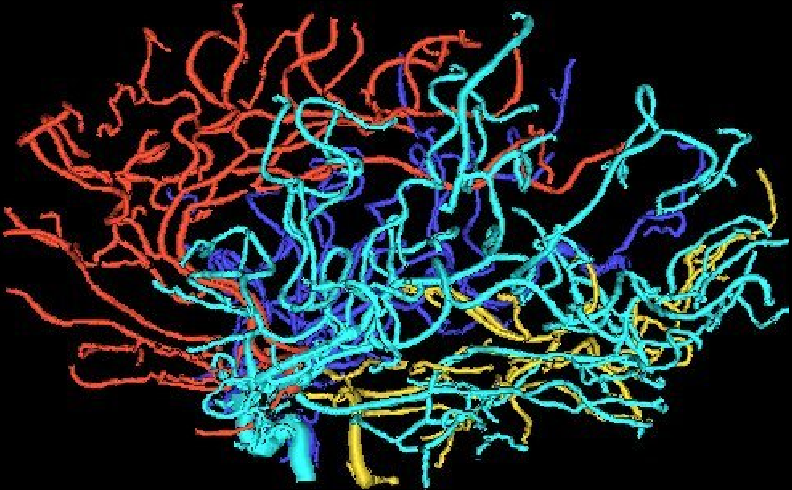
Lung vessels (CDH study)





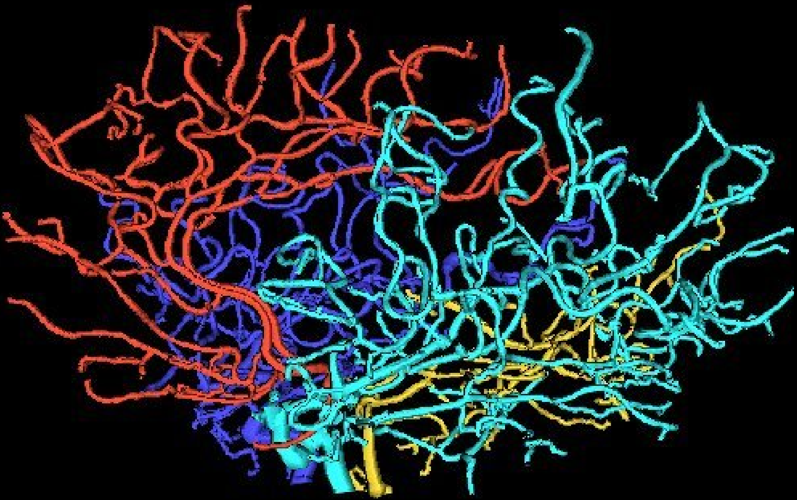
courtesy Sean McLean

Brain arteries



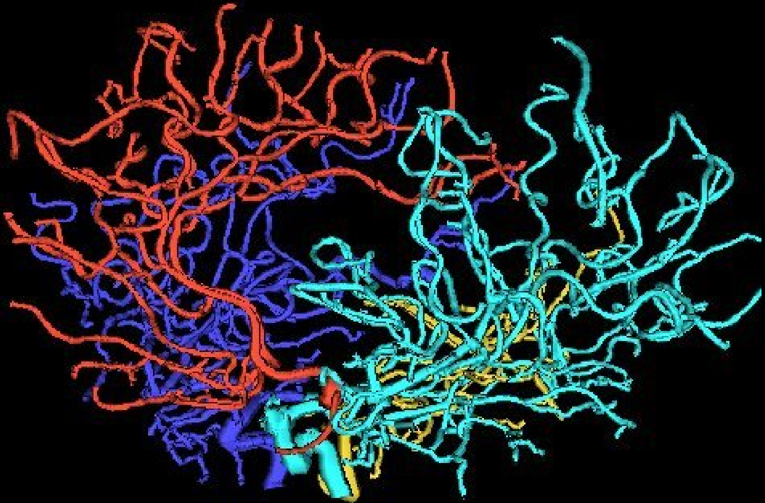
[Bullitt and Aylward, 2002]

Brain arteries



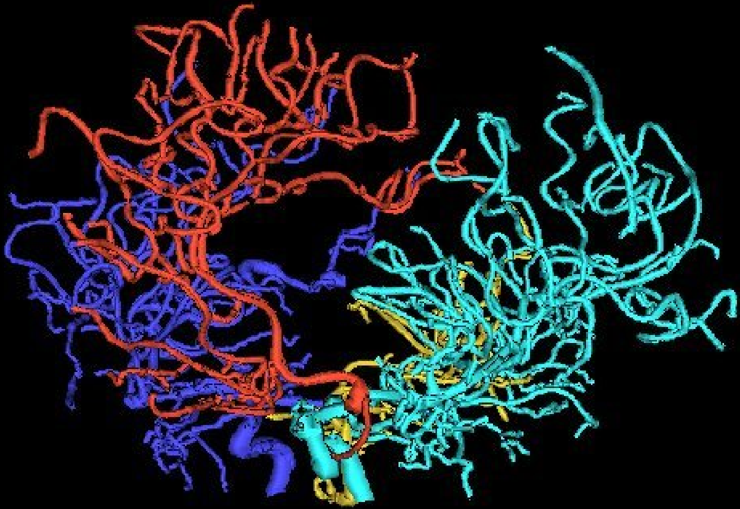
[Bullitt and Aylward, 2002]

Brain arteries



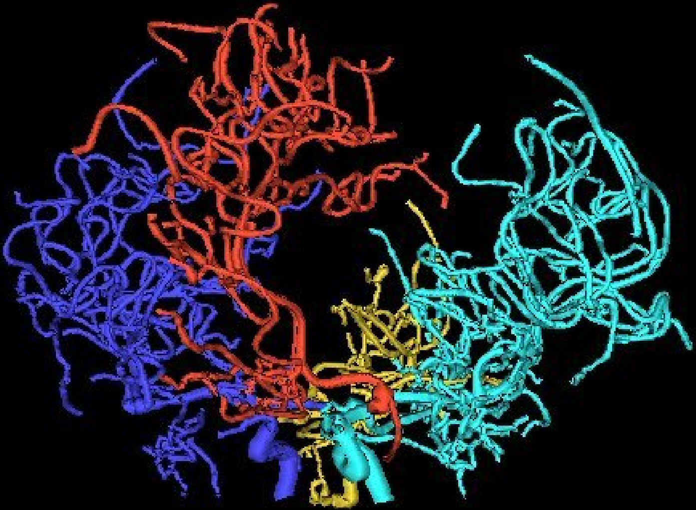
[Bullitt and Aylward, 2002]

Brain arteries



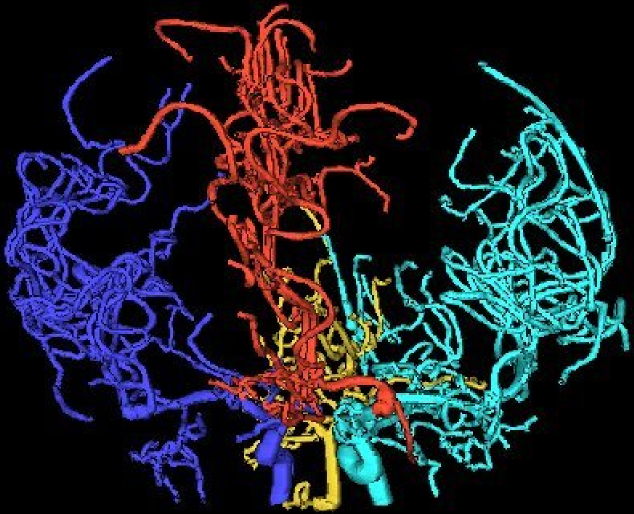
[Bullitt and Aylward, 2002]

Brain arteries



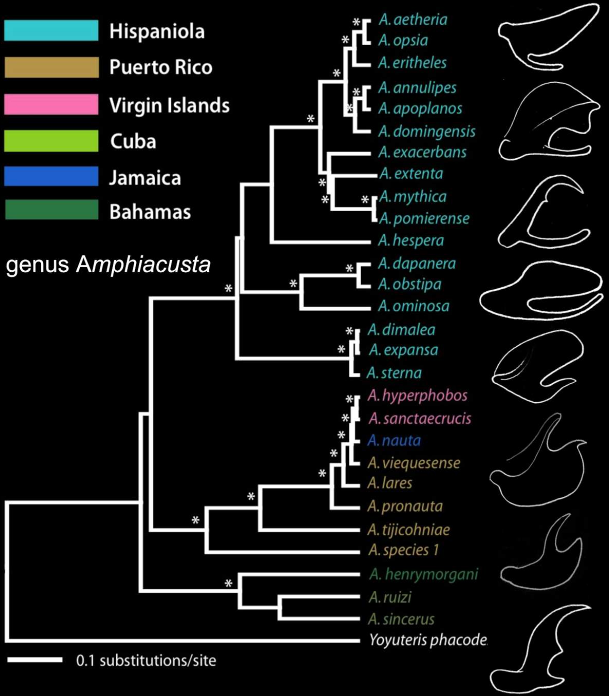
[Bullitt and Aylward, 2002]

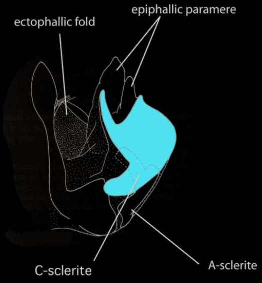
Brain arteries



[Bullitt and Aylward, 2002]

genus *Amphiacusta*





From Oneal, Otte & Knowles, 2010

Drawings by Dan Otte

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.: + phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Phylogenetic trees

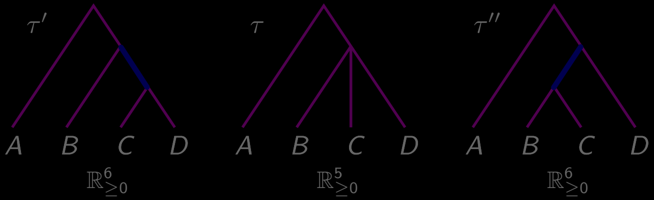
Def. A phylogenetic tree is a rooted metric tree with n labeled leaves

Distributions of trees come from

- tree reconstruction algorithms: LLN ⇒ sample mean → true tree
- evolutionary biology: “gene trees” from a “species tree”
- medical imaging: blood vessels, lungs, nerve cells, ...

Sample space $\mathcal{T}_n = \{\text{phylogenetic } n\text{-trees}\}$ is a union of polyhedral cones (orthants) [Billera–Holmes–Vogtmann 2001]

- $\mathcal{O}_\tau$ = trees with fixed topology $\tau \leftrightarrow \{\text{lists of edge lengths for } \tau\}$
  $= \text{orthant } \mathbb{R}^{E(\tau)}_{\geq 0}$



- $\mathcal{O}_\tau \subseteq \mathcal{O}_{\tau'} \Leftrightarrow \tau$ is a contraction of τ'
- $\mathcal{O}_\tau = \mathcal{O}_{\tau'} \cap \mathcal{O}_{\tau''} \Leftrightarrow \tau =$ biggest common contraction of τ' and τ''

Phylogenetic trees

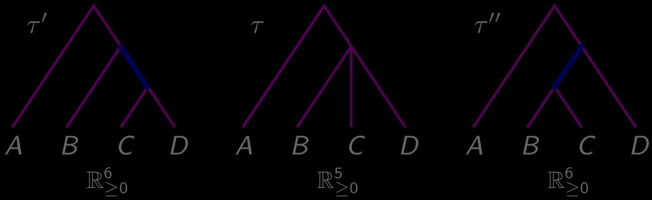
Def. A phylogenetic tree is a rooted metric tree with n labeled leaves

Distributions of trees come from

- tree reconstruction algorithms: LLN ⇒ sample mean → true tree
- evolutionary biology: "gene trees" from a "species tree"
- medical imaging: blood vessels, lungs, nerve cells, ...

Sample space $\mathcal{T}_n = \{$phylogenetic n-trees$\}$ is a union of polyhedral cones (orthants) [Billera–Holmes–Vogtmann 2001]

- $\mathcal{O}_\tau$ = trees with fixed topology $\tau \leftrightarrow \{$lists of edge lengths for $\tau\}$
  = orthant $\mathbb{R}_{\geq 0}^{E(\tau)}$



- $\mathcal{O}_\tau \subseteq \mathcal{O}_{\tau'} \Leftrightarrow \tau$ is a contraction of τ'
- $\mathcal{O}_\tau = \mathcal{O}_{\tau'} \cap \mathcal{O}_{\tau''} \Leftrightarrow \tau =$ biggest common contraction of τ' and τ''

Phylogenetic trees

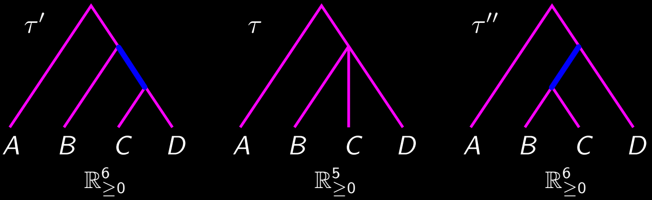
Def. A phylogenetic tree is a rooted metric tree with n labeled leaves

Distributions of trees come from

- tree reconstruction algorithms: LLN $\Rightarrow$ sample mean $\to$ true tree
- evolutionary biology: "gene trees" from a "species tree"
- medical imaging: blood vessels, lungs, nerve cells, ...

Sample space $\mathcal{T}_n = \{\text{phylogenetic } n\text{-trees}\}$ is a union of polyhedral cones (orthants) [Billera–Holmes–Vogtmann 2001]

- $\mathcal{O}_\tau$ = trees with fixed topology $\tau \leftrightarrow \{\text{lists of edge lengths for } \tau\}$ $= \text{orthant } \mathbb{R}_{\geq 0}^{E(\tau)}$



- $\mathcal{O}_\tau \subseteq \mathcal{O}_{\tau'} \Leftrightarrow \tau$ is a contraction of τ'
- $\mathcal{O}_\tau = \mathcal{O}_{\tau'} \cap \mathcal{O}_{\tau''} \Leftrightarrow \tau =$ biggest common contraction of τ' and τ''

Phylogenetic trees

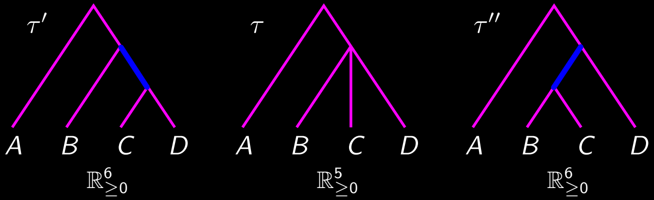
Def. A phylogenetic tree is a rooted metric tree with n labeled leaves

Distributions of trees come from

- tree reconstruction algorithms: LLN $\Rightarrow$ sample mean $\to$ true tree
- evolutionary biology: "gene trees" from a "species tree"
- medical imaging: blood vessels, lungs, nerve cells, ...

Sample space $\mathcal{T}_n = \{$phylogenetic n-trees$\}$ is a union of polyhedral cones (orthants) [Billera–Holmes–Vogtmann 2001]

- $\mathcal{O}_\tau =$ trees with fixed topology $\tau \leftrightarrow \{$lists of edge lengths for $\tau\}$
  $= \text{orthant } \mathbb{R}_{\geq 0}^{E(\tau)}$



- $\mathcal{O}_\tau \subseteq \mathcal{O}_{\tau'} \Leftrightarrow \tau$ is a contraction of τ'
- $\mathcal{O}_\tau = \mathcal{O}_{\tau'} \cap \mathcal{O}_{\tau''} \Leftrightarrow \tau =$ biggest common contraction of τ' and τ''

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1"
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: "Big Data" often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit "1" e.g.  or  or 
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Nonlinear data

Initial rationale: “Big Data” often sampled from nonlinear spaces.

Examples

- angles: points on a circle
  - wind direction
  - knee or elbow motion
- directions: points on a sphere
  - wrist or ankle motion
  - surface unit normal (e.g., medical imaging)
- shapes: points on a quotient of $d \times n$ matrices
- diffusion tensors: positive semidefinite matrices
- trees, e.g.:
  - phylogenetic tree space [Billera–Holmes–Vogtmann 2001],
  - phylogenetic orange [Kim 2000],
  - tropical grassmannian [Speyer–Sturmfels 2004],
  - wald space [Lueg–Garba–Nye–Huckemann 2021], . . .
- products and mixtures of these: unions of subspaces, spheres, tori, . . .
  - e.g., the digit “1” e.g.  or  or 
- persistence diagrams: topological summaries of
  - datasets
  - data objects

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals
- Central Limit Theorem (CLT)

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - "omnibus CLT" [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals
- Central Limit Theorem (CLT)

History

- for smooth M
    - CLT [Bhattacharya and Patrangenaru 2003, 2005]
    - "omnibus CLT" [Bhattacharya and Lin 2017]
    - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
    - open books [SAMSI Working Group 2013]
    - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
    - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT)

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - "omnibus CLT" [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT) ← focus on this for tutorial

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - "omnibus CLT" [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT) ← focus on this for tutorial

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - “omnibus CLT” [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT) ← focus on this for tutorial

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - "omnibus CLT" [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT) ← focus on this for tutorial

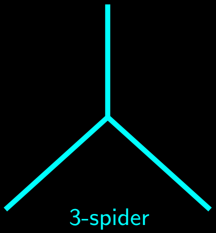
History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - “omnibus CLT” [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

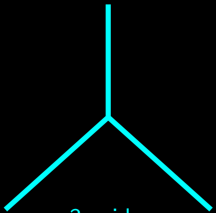
Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

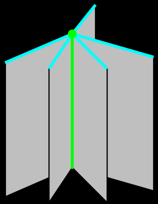
Motivation and history



Motivation and history



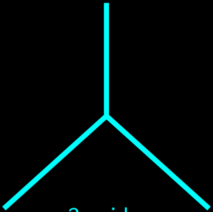
3-spider



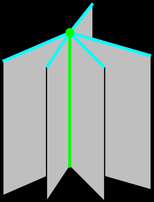
open book

$\mathbb{R}^d \times$ spider

Motivation and history

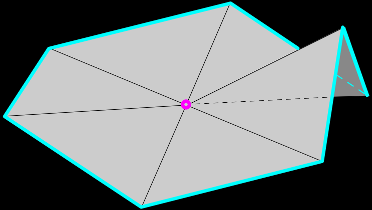


3-spider



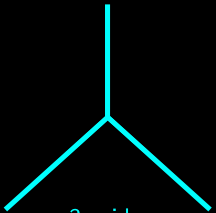
open book

$\mathbb{R}^d \times$ spider

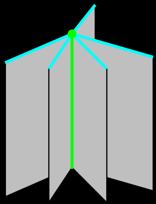


isolated hyperbolic planar singularity

Motivation and history

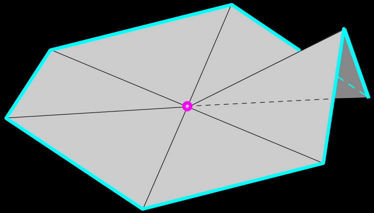


3-spider

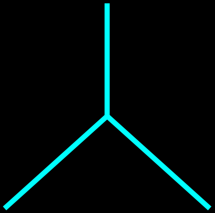


open book

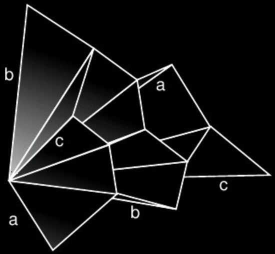
$\mathbb{R}^d$ × spider



isolated hyperbolic planar singularity



$\mathcal{T}_3$



$\mathcal{T}_4$

from [BHV 2001]

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT) ← focus on this for tutorial

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - "omnibus CLT" [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Motivation and history

Mimic ordinary statistics: assume nonlinear M given; want

- averages
- variance, PCA
- Law of Large Numbers (LLN), confidence intervals (regions)
- Central Limit Theorem (CLT) ← focus on this for tutorial

History

- for smooth M
  - CLT [Bhattacharya and Patrangenaru 2003, 2005]
  - "omnibus CLT" [Bhattacharya and Lin 2017]
  - smeary phenomena [Eltzner and Huckemann 2018]
- for singular M
  - open books [SAMSI Working Group 2013]
  - isolated planar singularity [Huckemann, Mattingly, M–, Nolen 2015]
  - phylogenetic tree spaces [Barden, Le 2018, w/Owen 2013, 2014]

Goals for minicourse

- What is an appropriate notion of stratified space?
- What is a Gaussian in the stratified setting?
- What geometry governs CLT there?

Fréchet means

Sample space: Riemannian manifold M

What fails

1. sum or average of
   - $+$ points in M
   - $+$ random variables in M
2. “Gaussian” on M

Workarounds

1. **Def.** Probability distribution μ on any metric space M has Fréchet function

$$F_\mu(y) = \frac{1}{2}\int_M \underset{\substack{\uparrow\\ \text{square}\\ \text{distance}}}{d(x,y)^2}\,\underset{\substack{\uparrow\\ \text{measure}\\ \text{induced}\\ \text{by } \mu}}{\mu(dx)}$$

   and Fréchet mean $\bar{\mu} = \underset{y\in M}{\operatorname{argmin}}\, F_\mu(y)$.
   - “least squares approximation”
   - empirical mean $\bar{\mu}_n$ from empirical measure $\mu_n = \frac{1}{n}(\delta_{X_1} + \cdots + \delta_{X_n})$
   - LLN unaffected: $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.
2. Reduce to linear case

Fréchet means

Sample space: Riemannian manifold M

What fails

1. sum or average of
   - points in M
   - random variables in M
2. "Gaussian" on M

Workarounds

1. **Def.** Probability distribution μ on any metric space M has Fréchet function

$$F_\mu(y) = \frac{1}{2}\int_M \underset{\substack{\uparrow\\ \text{square}\\ \text{distance}}}{d(x,y)^2}\,\underset{\substack{\uparrow\\ \text{measure}\\ \text{induced}\\ \text{by } \mu}}{\mu(dx)}$$

   and Fréchet mean $\bar{\mu} = \operatorname*{argmin}_{y\in M} F_\mu(y)$.

   - "least squares approximation"
   - empirical mean $\bar{\mu}_n$ from empirical measure $\mu_n = \frac{1}{n}(\delta_{X_1} + \cdots + \delta_{X_n})$
   - LLN unaffected: $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.
2. Reduce to linear case

Fréchet means

Sample space: Riemannian manifold M

What fails

1. sum or average of
    + points in M
    + random variables in M
2. "Gaussian" on M

Workarounds

1. **Def.** Probability distribution μ on any metric space M has Fréchet function

$$F_\mu(y) = \frac{1}{2}\int_M \underset{\substack{\uparrow\\ \text{square}\\ \text{distance}}}{d(x,y)^2}\,\underset{\substack{\uparrow\\ \text{measure}\\ \text{induced}\\ \text{by } \mu}}{\mu(dx)}$$

   and Fréchet mean $\bar{\mu} = \operatorname*{argmin}_{y\in M} F_\mu(y)$.

   - "least squares approximation"
   - empirical mean $\bar{\mu}_n$ from empirical measure $\mu_n = \frac{1}{n}(\delta_{X_1} + \cdots + \delta_{X_n})$
   - LLN unaffected: $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.

2. Reduce to linear case

Fréchet means

Sample space: Riemannian manifold M

What fails

1. sum or average of
   - points in M
   - random variables in M
2. "Gaussian" on M

Workarounds

1. **Def.** Probability distribution μ on any metric space M has Fréchet function

$$F_\mu(y) = \frac{1}{2}\int_M \underset{\substack{\uparrow\\ \text{square}\\ \text{distance}}}{d(x,y)^2}\, \underset{\substack{\uparrow\\ \text{measure}\\ \text{induced}\\ \text{by } \mu}}{\mu(dx)}$$

   and Fréchet mean $\bar{\mu} = \operatorname*{argmin}_{y\in M} F_\mu(y)$.

   - "least squares approximation"
   - empirical mean $\bar{\mu}_n$ from empirical measure $\mu_n = \frac{1}{n}(\delta_{X_1} + \cdots + \delta_{X_n})$
   - LLN unaffected: $\bar{\mu}_n \xrightarrow{n\to\infty} \bar{\mu}$ almost surely.
2. Reduce to linear case

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

⇒ limit is a random tangent vector
in tangent space $T_{\bar{\mu}}$

Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}}M \\ x &\mapsto d(\bar{\mu}, x)V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}}M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

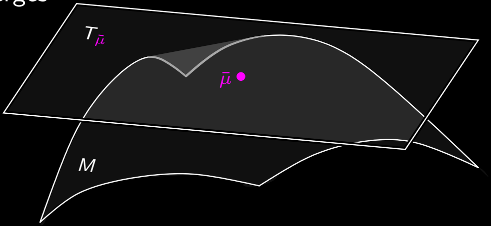
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

⇒ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}}M \\ x &\mapsto d(\bar{\mu}, x)V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}}M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

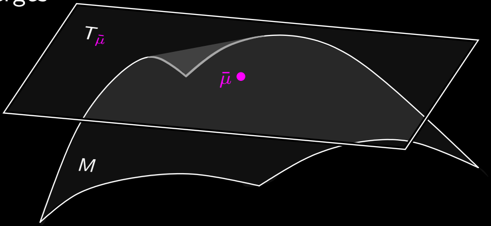
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

⇒ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned}\log_{\bar{\mu}} : M &\to T_{\bar{\mu}}M \\ x &\mapsto d(\bar{\mu}, x)V,\end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}}M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

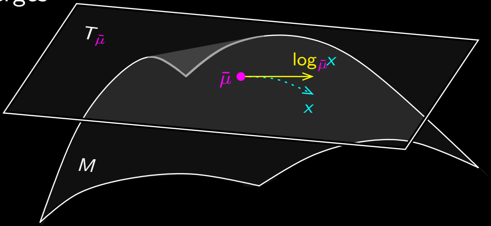
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

$\Rightarrow$ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}} M \\ x &\mapsto d(\bar{\mu}, x)\, V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}} M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0, \Sigma)$ in distribution

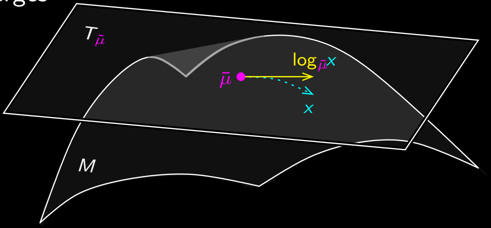
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

⇒ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}}M \\ x &\mapsto d(\bar{\mu}, x)V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}}M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0, \Sigma)$ in distribution

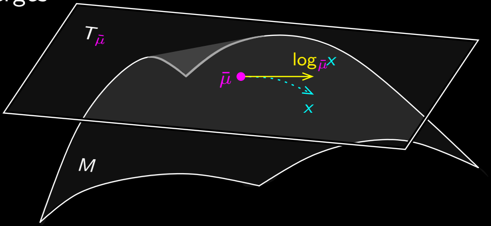
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

$\Rightarrow$ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}}M \\ x &\mapsto d(\bar{\mu}, x)V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}}M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$ Def. pushforward of μ to $T_{\bar{\mu}}M$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

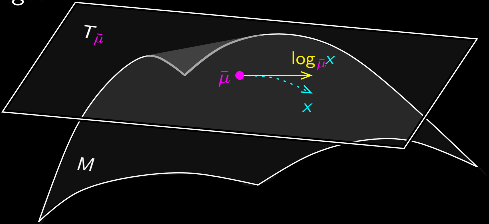
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

⇒ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned}\log_{\bar{\mu}} : M &\to T_{\bar{\mu}} M\\ x &\mapsto d(\bar{\mu}, x)V,\end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}} M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$ Def. pushforward of μ to $T_{\bar{\mu}} M$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

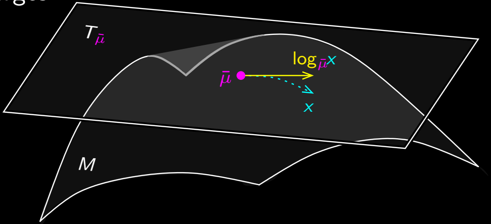
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

$\Rightarrow$ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}} M \\ x &\mapsto d(\bar{\mu}, x) V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}} M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$ Def. pushforward of μ to $T_{\bar{\mu}} M$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n \to \infty} N(0, \Sigma)$ in distribution

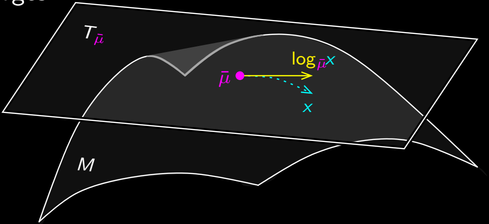
Question: Is this the manifold CLT?

Logarithm maps

Recast CLT on manifold M

- variation of rescaled differences $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu})$
- as (moving) empirical mean converges
- to (fixed) population mean

$\Rightarrow$ limit is a random tangent vector in tangent space $T_{\bar{\mu}}$



Def. The logarithm map is

$$\begin{aligned} \log_{\bar{\mu}} : M &\to T_{\bar{\mu}} M \\ x &\mapsto d(\bar{\mu}, x)\,V, \end{aligned}$$

where V = unit tangent to geodesic from $\bar{\mu}$ to x.

Back to linear setting

- μ on $M \rightsquigarrow \nu$ on $T_{\bar{\mu}} M$ for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$ Def. pushforward of μ to $T_{\bar{\mu}} M$
- linear CLT: $\sqrt{n}\,\bar{\nu}_n \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

Question: Is this the manifold CLT? Not quite. . . .

Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar\mu}M \to T_{\bar\mu}M$$

is the inverse of the Hessian at $\bar\mu$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar\mu} F_\mu)^{-1}.$$

- $(M, \mu) \rightsquigarrow (T_{\bar\mu}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar\mu$ tug covariance of μ toward $\bar\mu$ as compared with ν around $0 \in T_{\bar\mu}M$

Def. Any $T_{\bar\mu}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar\mu}\bar\mu_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0,\Sigma)$ in distribution,
where $N(0,\Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar\mu}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar\mu_n - \bar\mu) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

Differences:
- LHS $\log_{\bar\mu}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar\mu}M$

Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar\mu}M \to T_{\bar\mu}M$$

is the inverse of the Hessian at $\bar\mu$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar\mu} F_\mu)^{-1}.$$

- $(M, \mu) \rightsquigarrow (T_{\bar\mu}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar\mu$ tug covariance of μ toward $\bar\mu$ as compared with ν around $0 \in T_{\bar\mu}M$

Def. Any $T_{\bar\mu}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar\mu}\bar\mu_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0,\Sigma)$ in distribution, where $N(0,\Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar\mu}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar\mu_n - \bar\mu) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

Differences:
- LHS $\log_{\bar\mu}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar\mu}M$

Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar{\mu}}M \to T_{\bar{\mu}}M$$

is the inverse of the Hessian at $\bar{\mu}$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar{\mu}} F_\mu)^{-1}.$$

- $(M, \mu) \rightsquigarrow (T_{\bar{\mu}}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar{\mu}$ tug covariance of μ toward $\bar{\mu}$ as compared with ν around $0 \in T_{\bar{\mu}}M$

Def. Any $T_{\bar{\mu}}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution, where $N(0, \Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0, \Sigma)$ in distribution

Differences:
- LHS $\log_{\bar{\mu}}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar{\mu}}M$

Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar{\mu}}M \to T_{\bar{\mu}}M$$

is the inverse of the Hessian at $\bar{\mu}$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar{\mu}} F_\mu)^{-1}.$$

- $(M, \mu) \rightsquigarrow (T_{\bar{\mu}}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar{\mu}$ tug covariance of μ toward $\bar{\mu}$ as compared with ν around $0 \in T_{\bar{\mu}}M$

Def. Any $T_{\bar{\mu}}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution, where $N(0, \Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0, \Sigma)$ in distribution

Differences:
- LHS $\log_{\bar{\mu}}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar{\mu}}M$

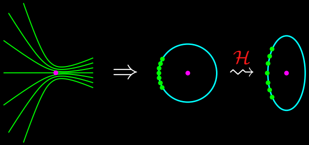
Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar{\mu}}M \to T_{\bar{\mu}}M$$

is the inverse of the Hessian at $\bar{\mu}$ of the Fréchet function F_{μ}:

$$\mathcal{H} = (\nabla\nabla_{\bar{\mu}} F_{\mu})^{-1}.$$



- $(M, \mu) \rightsquigarrow (T_{\bar{\mu}}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar{\mu}$ tug covariance of μ toward $\bar{\mu}$ as compared with ν around $0 \in T_{\bar{\mu}}M$

Def. Any $T_{\bar{\mu}}M$-valued random variable N has pushforward $\mathcal{H}_{\sharp}N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_{\sharp}N(0,\Sigma)$ in distribution, where $N(0,\Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0,\Sigma)$ in distribution

Differences:
- LHS $\log_{\bar{\mu}}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar{\mu}}M$

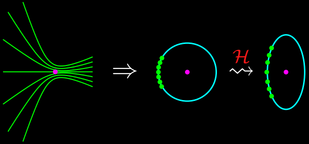
Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar{\mu}}M \to T_{\bar{\mu}}M$$

is the inverse of the Hessian at $\bar{\mu}$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar{\mu}} F_\mu)^{-1}.$$



- $(M, \mu) \rightsquigarrow (T_{\bar{\mu}}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar{\mu}$ tug covariance of μ toward $\bar{\mu}$ as compared with ν around $0 \in T_{\bar{\mu}}M$

Def. Any $T_{\bar{\mu}}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution, where $N(0, \Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0, \Sigma)$ in distribution

Differences:
- LHS $\log_{\bar{\mu}}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar{\mu}}M$

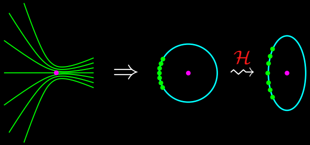
Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar{\mu}}M \to T_{\bar{\mu}}M$$

is the inverse of the Hessian at $\bar{\mu}$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar{\mu}} F_\mu)^{-1}.$$



- $(M, \mu) \rightsquigarrow (T_{\bar{\mu}}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar{\mu}$
  tug covariance of μ toward $\bar{\mu}$ as compared with ν around $0 \in T_{\bar{\mu}}M$

Def. Any $T_{\bar{\mu}}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution,
where $N(0, \Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n\to\infty} N(0, \Sigma)$ in distribution

Differences:
- LHS $\log_{\bar{\mu}}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar{\mu}}M$

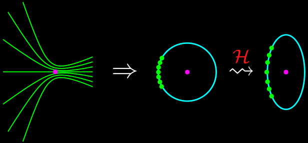
Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar{\mu}}M \to T_{\bar{\mu}}M$$

is the inverse of the Hessian at $\bar{\mu}$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar{\mu}} F_\mu)^{-1}.$$



- $(M, \mu) \rightsquigarrow (T_{\bar{\mu}}M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar{\mu}$ tug covariance of μ toward $\bar{\mu}$ as compared with ν around $0 \in T_{\bar{\mu}}M$

Def. Any $T_{\bar{\mu}}M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n} \log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n \to \infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution,
where $N(0, \Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar{\mu}}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar{\mu}_n - \bar{\mu}) \xrightarrow{n \to \infty} N(0, \Sigma)$ in distribution

Differences:
- LHS $\log_{\bar{\mu}}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar{\mu}}M$

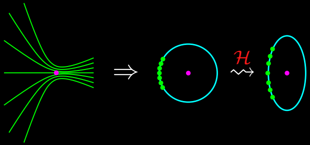
Smooth manifold CLT

Def. The distortion map

$$\mathcal{H} : T_{\bar\mu} M \to T_{\bar\mu} M$$

is the inverse of the Hessian at $\bar\mu$ of the Fréchet function F_μ:

$$\mathcal{H} = (\nabla\nabla_{\bar\mu} F_\mu)^{-1}.$$



- $(M, \mu) \rightsquigarrow (T_{\bar\mu} M, \nu)$ forgets curvature of M
- e.g., rapidly spreading geodesics exiting $\bar\mu$ tug covariance of μ toward $\bar\mu$ as compared with ν around $0 \in T_{\bar\mu} M$

Def. Any $T_{\bar\mu} M$-valued random variable N has pushforward $\mathcal{H}_\sharp N = \mathcal{H} \circ N$.

Manifold CLT: $\sqrt{n} \log_{\bar\mu} \bar\mu_n \xrightarrow{n \to \infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution, where $N(0, \Sigma)$ is the limit law for $\nu = \mu \circ \log_{\bar\mu}{}^{-1}$.

Linear CLT: $\sqrt{n}\,(\bar\mu_n - \bar\mu) \xrightarrow{n \to \infty} N(0, \Sigma)$ in distribution

Differences:
- LHS $\log_{\bar\mu}$ pushes to linear setting
- RHS $\mathcal{H}$ accounts for curvature lost by $M \rightsquigarrow T_{\bar\mu} M$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$

     so $$\begin{tikzcd} & \text{tangent cone: singular!} \arrow[d] & \\ M \arrow[r, "\log_{\bar{\mu}}"] & T_{\bar{\mu}}M \arrow[r, "\mathcal{L}"] & \mathbb{R}^m \end{tikzcd}$$
   - $N(0,\Sigma)$ = Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n}\,\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0,\Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse

     $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$

     so $M \xrightarrow{\log_{\bar{\mu}}} \overset{\substack{\text{tangent cone: singular!} \\ \downarrow}}{T_{\bar{\mu}}M} \xrightarrow{\mathcal{L}} \mathbb{R}^m$
   - $N(0, \Sigma) =$ Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse

   $$\mathcal{L} : T_{\bar\mu}M \to \mathbb{R}^m$$

   tangent cone: singular!

   so $M \xrightarrow{\log_{\bar\mu}} \overset{\downarrow}{T}_{\bar\mu}M \xrightarrow{\mathcal{L}} \mathbb{R}^m$

   - $N(0,\Sigma) =$ Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar\mu}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar\mu}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n}\log_{\bar\mu}\bar\mu_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0,\Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse

     $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$

     tangent cone: singular!

     so $M \xrightarrow{\log_{\bar{\mu}}} \underset{\uparrow}{T_{\bar{\mu}}M} \xrightarrow{\mathcal{L}} \mathbb{R}^m$
   - $N(0, \Sigma)$ = Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse

     $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$

     tangent cone: singular!

     ↓

     so $M \xrightarrow{\log_{\bar{\mu}}} T_{\bar{\mu}}M \xrightarrow{\mathcal{L}} \mathbb{R}^m$
   - $N(0, \Sigma) =$ Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1.
   - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2.
   - reduce to linear case using extra step: tangential collapse

     $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$

     tangent cone: singular!
     ↓
     so $M \xrightarrow{\log_{\bar{\mu}}} T_{\bar{\mu}}M \xrightarrow{\mathcal{L}} \mathbb{R}^m$
   - $N(0, \Sigma) =$ Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n} \log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse

     tangent cone: singular! $\downarrow$ $\qquad$ $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$

     so $M \xrightarrow{\log_{\bar{\mu}}} T_{\bar{\mu}}M \xrightarrow{\mathcal{L}} \mathbb{R}^m$
   - $N(0, \Sigma) =$ Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n} \log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1. - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2. - reduce to linear case using extra step: tangential collapse

     tangent cone: singular! $\qquad \mathcal{L} : T_{\bar{\mu}} M \to \mathbb{R}^m$

     $$\text{so } M \xrightarrow{\log_{\bar{\mu}}} \overset{\downarrow}{T_{\bar{\mu}} M} \xrightarrow{\mathcal{L}} \mathbb{R}^m$$

   - $N(0, \Sigma)$ = Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar{\mu}}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar{\mu}} M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n} \log_{\bar{\mu}} \bar{\mu}_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Singular CLT

Problems: need appropriate

1. classes of spaces M and measures μ
2. analogues of Gaussian random variables as limiting distributions N
3. reflection of geometry ("curvature") of M in N

Solutions

1.
   - smoothly stratified metric space M of curvature bounded above
   - localized immured amenable probability measure μ on M
2.
   - reduce to linear case using extra step: tangential collapse

     $$\mathcal{L} : T_{\bar\mu}M \to \mathbb{R}^m$$

     $$\text{so } M \xrightarrow{\log_{\bar\mu}} \overset{\substack{\text{tangent cone: singular!}\\ \downarrow}}{T_{\bar\mu}M} \xrightarrow{\mathcal{L}} \mathbb{R}^m$$
   - $N(0, \Sigma)$ = Gaussian random vector supported on $\mathbb{R}^\ell \subseteq \mathbb{R}^m$ with same covariance Σ as pushforward of μ under the composite $\mathcal{L} \circ \log_{\bar\mu}$
3. distortion map $\mathcal{H} : \mathbb{R}^\ell \to T_{\bar\mu}M$

Singluar CLT [Mattingly, M–, Tran 2023]: $\sqrt{n}\log_{\bar\mu}\bar\mu_n \xrightarrow{n\to\infty} \mathcal{H}_\sharp N(0, \Sigma)$ in distribution.

Note: same expression as Manifold CLT!
$\mathcal{L}$ is hiding in N and $\mathcal{H}$

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - only really needed at $\bar{\mu}$
  - morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - standard, mild, analytic sort of hypothesis on μ
  - allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - avoids cut locus and smeary issues
  - ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is $\mathrm{CAT}(\kappa)$

Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata ↔ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces

Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces

Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces

Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - + is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - + has closure $\overline{M^j} = \bigcup_{k \le j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces

Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

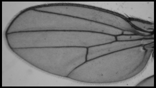
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata ↔ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces



Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

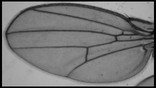
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata ↔ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces



Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

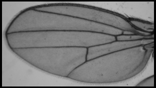
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - \+ is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - \+ has closure $\overline{M^j} = \bigcup_{k \le j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

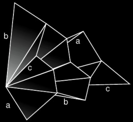
Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces





Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

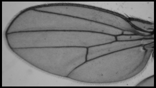
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

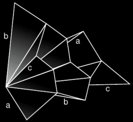
Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces





Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

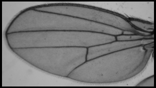
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - + is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - + has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is $\mathrm{CAT}(\kappa)$

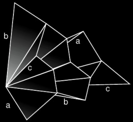
Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces





Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

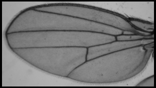
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - has closure $\overline{M^j} = \bigcup_{k \leq j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is $\mathrm{CAT}(\kappa)$

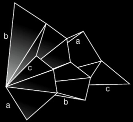
Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata ↔ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces





Stratified spaces

Def [Mattingly, M–, Tran 2023]. M is smoothly stratified with distance $\mathbf{d}$ if

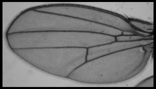
- M is a complete, locally compact, geodesic space
- $M = \bigsqcup_{j=0}^{d} M^j$ has disjoint locally closed strata M^j
- each stratum M^j
  - + is a manifold with geodesic distance $\mathbf{d}|_{M^j}$
  - + has closure $\overline{M^j} = \bigcup_{k \le j} M^k$
- locally well defined exponential maps that are local homeomorphisms
- curvature bounded above by κ: M is CAT(κ)

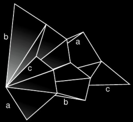
Examples

- graph (or network): strata are vertices and edges
- polyhedron: strata are (relatively open) faces
- real (semi)algebraic variety: strata $\leftrightarrow$ equisingular loci

Actual examples

- fruit fly wings
- tree spaces
- shape spaces





Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - only really needed at $\bar{\mu}$
  - morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - standard, mild, analytic sort of hypothesis on μ
  - allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - avoids cut locus and smeary issues
  - ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - + essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is $\mathrm{CAT}(\kappa)$
  - + only really needed at $\bar{\mu}$
  - + morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - + standard, mild, analytic sort of hypothesis on μ
  - + allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - + always if M is $\mathrm{CAT}(0)$ (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - + always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - + maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - + always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - + avoids cut locus and smeary issues
  - + ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - + essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - + only really needed at $\bar{\mu}$
  - + morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - + standard, mild, analytic sort of hypothesis on μ
  - + allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp} \mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - + always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - + always if $\operatorname{supp} \mu$ contains a neighborhood of $\bar{\mu}$
  - + maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - + always if $\operatorname{supp} \mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - + avoids cut locus and smeary issues
  - + ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - + essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - + only really needed at $\bar{\mu}$
  - + morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - + standard, mild, analytic sort of hypothesis on μ
  - + allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - + always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - + always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - + maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - + always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - + avoids cut locus and smeary issues
  - + ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - \+ essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - \+ only really needed at $\bar{\mu}$
  - \+ morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - \+ standard, mild, analytic sort of hypothesis on μ
  - \+ allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp} \mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - \+ always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - \+ always if $\operatorname{supp} \mu$ contains a neighborhood of $\bar{\mu}$
  - \+ maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - \+ always if $\operatorname{supp} \mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - \+ avoids cut locus and smeary issues
  - \+ ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - only really needed at $\bar{\mu}$
  - morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - standard, mild, analytic sort of hypothesis on μ
  - allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - avoids cut locus and smeary issues
  - ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - + essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - + only really needed at $\bar{\mu}$
  - + morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - + standard, mild, analytic sort of hypothesis on μ
  - + allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp} \mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - + always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - + always if $\operatorname{supp} \mu$ contains a neighborhood of $\bar{\mu}$
  - + maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - + always if $\operatorname{supp} \mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - + avoids cut locus and smeary issues
  - + ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - + essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - + only really needed at $\bar{\mu}$
  - + morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x, -)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - + standard, mild, analytic sort of hypothesis on μ
  - + allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - + always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - + always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - + maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - + always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - + avoids cut locus and smeary issues
  - + ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Hypotheses

Hypotheses on M

- "nice" union of finitely many manifolds (strata)
- locally well defined exponential maps that are local homeomorphisms
  - essential for bringing asymptotics of sampling to $T_{\bar{\mu}}M$ and back to M
- curvature bounded above by κ: M is CAT(κ)
  - only really needed at $\bar{\mu}$
  - morally won't be infinitely curved anyway: Fréchet means would flee

Hypotheses on μ

- amenable: $d(x,-)^2$ has finite μ-expectation directional 2nd derivatives at $\bar{\mu}$
  - standard, mild, analytic sort of hypothesis on μ
  - allows differential and Taylor expansion techniques for optimization
- immured: $\log_{\bar{\mu}} \bar{\mu}_n \in$ convex cone generated by $\operatorname{supp}\mu$ when $\bar{\mu}_n$ is near $\bar{\mu}$
  - always if M is CAT(0) (e.g., $M = T_{\bar{\mu}}M$, such as $M = \mathcal{T}_n$ or open book)
  - always if $\operatorname{supp}\mu$ contains a neighborhood of $\bar{\mu}$
  - maybe always no matter what
- localized: unique Fréchet mean, locally convex F_μ, and $\log_{\bar{\mu}}$ defined a.s.
  - always if $\operatorname{supp}\mu \subseteq B(\bar{\mu}, \pi/\sqrt{\kappa})$
  - avoids cut locus and smeary issues
  - ensures pushforward of μ to $T_{\bar{\mu}}M$ is defined

Singular distortion

What can limiting distribution $\mathcal{H}_\sharp N(0, \Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]

- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)

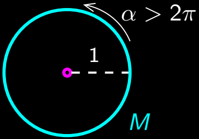
embedded in $\mathbb{R}^3$:

- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0, \Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at 0, $\pm\frac{2\pi}{5}$, $\pm\frac{4\pi}{5}$

Singular distortion

What can limiting distribution $\mathcal{H}_\sharp N(0,\Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]



- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)

embedded in $\mathbb{R}^3$:

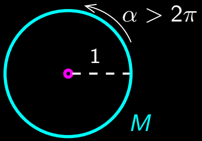
- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0,\Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at 0, $\pm\frac{2\pi}{5}$, $\pm\frac{4\pi}{5}$

Singular distortion

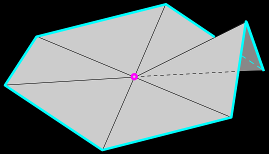
What can limiting distribution $\mathcal{H}_\sharp N(0, \Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]

- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)



embedded in $\mathbb{R}^3$:



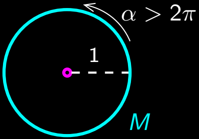
- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0, \Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at $0, \pm\frac{2\pi}{5}, \pm\frac{4\pi}{5}$

Singular distortion

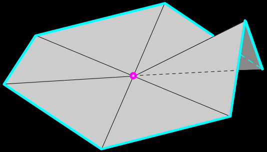
What can limiting distribution $\mathcal{H}_\sharp N(0,\Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]

- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)



embedded in $\mathbb{R}^3$:



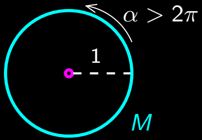
- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0,\Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at $0, \pm\frac{2\pi}{5}, \pm\frac{4\pi}{5}$

Singular distortion

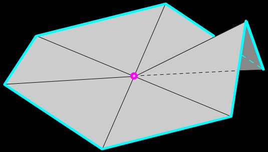
What can limiting distribution $\mathcal{H}_\sharp N(0, \Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]

- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)



embedded in $\mathbb{R}^3$:

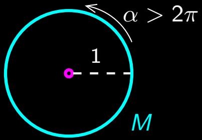


- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0, \Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at $0, \pm\frac{2\pi}{5}, \pm\frac{4\pi}{5}$

Singular distortion

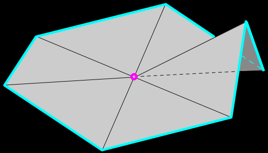
What can limiting distribution $\mathcal{H}_\sharp N(0,\Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]

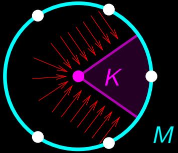


- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)

embedded in $\mathbb{R}^3$:



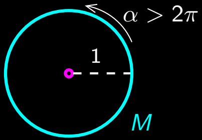
- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0,\Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at $0, \pm\frac{2\pi}{5}, \pm\frac{4\pi}{5}$



Singular distortion

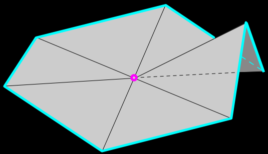
What can limiting distribution $\mathcal{H}_\sharp N(0,\Sigma)$ look like?

Example [Huckemann, Mattingly, M–, Nolen 2015]

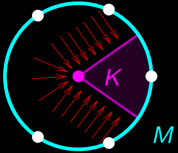


- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)

embedded in $\mathbb{R}^3$:



- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0,\Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at 0, $\pm\frac{2\pi}{5}$, $\pm\frac{4\pi}{5}$

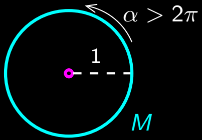


Singular distortion

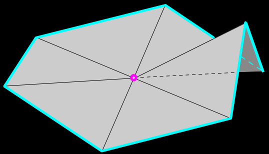
What can limiting distribution $\mathcal{H}_\sharp N(0, \Sigma)$ look like?

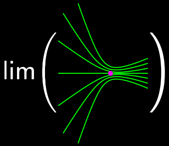
Example [Huckemann, Mattingly, M–, Nolen 2015]

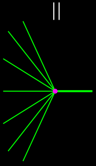
- Isolated hyperbolic planar singularity: angle sum at apex is $\alpha > 2\pi$ (that is, circumference at radius 1 is α)



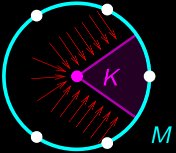
embedded in $\mathbb{R}^3$:







- Note: singularity of M is geometric, not topological
- Pushforward under distortion map $\mathcal{H}_\sharp$ is convex projection from tangent cone $T_{\bar{\mu}}M$ to fluctuating cone K.
- limiting measure $\mathcal{H}_\sharp N(0, \Sigma)$ on isolated hyperbolic planar singularity for $\mu = \sum$ atoms at $0, \pm\frac{2\pi}{5}, \pm\frac{4\pi}{5}$



New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty} \sqrt{n}\log_{\bar\mu} \bar\mu_n \overset{d}{=} \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{V\in T_{\bar\mu}M} \big(F_\mu(\exp_{\bar\mu} V) - tG(V)\big)$$

- $G =$ real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar\mu}M$
- spatial variation in $T_{\bar\mu}M \rightsquigarrow$ radial variation in $S_{\bar\mu}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar\mu}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim_{n\to\infty} \sqrt{n}\log_{\bar\mu} \bar\mu_n = \nabla_\mu \mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b} : \mathcal{P}_2M \to M$ sending $\mu \mapsto \bar\mu$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim_{n\to\infty} \sqrt{n}\log_{\bar\mu} \bar\mu_n \overset{d}{=} \nabla_{F_\mu\circ\exp_{\bar\mu}} \mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar\mu}M, \mathbb{R})$,
- of the minimizer map $\mathfrak{B} : \mathcal{C}(T_{\bar\mu}M, \mathbb{R}) \to T_{\bar\mu}M$ that sends $f \mapsto \operatorname*{argmin}_{X\in C_\mu} f(X)$
- at $F_\mu \circ \exp_{\bar\mu}$
- along the Gaussian tangent field $G = G(\,\cdot\,) = \langle \Gamma_\mu, \cdot\,\rangle_{\bar\mu}$ induced by μ

New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty}\sqrt{n}\log_{\bar\mu}\bar\mu_n \overset{d}{=} \lim_{t\to 0}\frac{1}{t}\operatorname*{argmin}_{V\in T_{\bar\mu}M}\big(F_\mu(\exp_{\bar\mu}V)-tG(V)\big)$$

- G = real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar\mu}M$
- spatial variation in $T_{\bar\mu}M \rightsquigarrow$ radial variation in $S_{\bar\mu}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar\mu}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\log_{\bar\mu}\bar\mu_n = \nabla_\mu\mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b}:\mathcal{P}_2M\to M$ sending $\mu\mapsto\bar\mu$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\log_{\bar\mu}\bar\mu_n \overset{d}{=} \nabla_{F_\mu\circ\exp_{\bar\mu}}\mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar\mu}M,\mathbb{R})$,
- of the minimizer map $\mathfrak{B}:\mathcal{C}(T_{\bar\mu}M,\mathbb{R})\to T_{\bar\mu}M$ that sends $f\mapsto\operatorname*{argmin}_{X\in C_\mu}f(X)$
- at $F_\mu\circ\exp_{\bar\mu}$
- along the Gaussian tangent field $G = G(\cdot) = \langle\Gamma_\mu,\cdot\rangle_{\bar\mu}$ induced by μ

New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty}\sqrt{n}\,\log_{\bar\mu}\bar\mu_n \overset{d}{=} \lim_{t\to 0}\frac{1}{t}\operatorname*{argmin}_{V\in T_{\bar\mu}M}\big(F_\mu(\exp_{\bar\mu}V) - t\,G(V)\big)$$

- G = real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar\mu}M$
- spatial variation in $T_{\bar\mu}M \rightsquigarrow$ radial variation in $S_{\bar\mu}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar\mu}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\,\log_{\bar\mu}\bar\mu_n = \nabla_\mu\mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b} : \mathcal{P}_2M \to M$ sending $\mu \mapsto \bar\mu$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\,\log_{\bar\mu}\bar\mu_n \overset{d}{=} \nabla_{F_\mu\circ\exp_{\bar\mu}}\mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar\mu}M,\mathbb{R})$,
- of the minimizer map $\mathfrak{B} : \mathcal{C}(T_{\bar\mu}M,\mathbb{R}) \to T_{\bar\mu}M$ that sends $f \mapsto \operatorname*{argmin}\limits_{X\in C_\mu} f(X)$
- at $F_\mu \circ \exp_{\bar\mu}$
- along the Gaussian tangent field $G = G(\cdot) = \langle\Gamma_\mu, \cdot\rangle_{\bar\mu}$ induced by μ

New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty} \sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \overset{d}{=} \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{V\in T_{\bar{\mu}}M} \big(F_\mu(\exp_{\bar{\mu}} V) - tG(V)\big)$$

- G = real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar{\mu}}M$
- spatial variation in $T_{\bar{\mu}}M \rightsquigarrow$ radial variation in $S_{\bar{\mu}}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar{\mu}}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim_{n\to\infty} \sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n = \nabla_\mu \mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2 M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b} : \mathcal{P}_2 M \to M$ sending $\mu \mapsto \bar{\mu}$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim_{n\to\infty} \sqrt{n}\log_{\bar{\mu}} \bar{\mu}_n \overset{d}{=} \nabla_{F_\mu \circ \exp_{\bar{\mu}}} \mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar{\mu}}M, \mathbb{R})$,
- of the minimizer map $\mathfrak{B} : \mathcal{C}(T_{\bar{\mu}}M, \mathbb{R}) \to T_{\bar{\mu}}M$ that sends $f \mapsto \operatorname*{argmin}_{X\in C_\mu} f(X)$
- at $F_\mu \circ \exp_{\bar{\mu}}$
- along the Gaussian tangent field $G = G(\cdot) = \langle \Gamma_\mu, \cdot \rangle_{\bar{\mu}}$ induced by μ

New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty}\sqrt{n}\,\log_{\bar{\mu}}\bar{\mu}_n \overset{d}{=} \lim_{t\to 0}\frac{1}{t}\operatorname*{argmin}_{V\in T_{\bar{\mu}}M}\big(F_\mu(\exp_{\bar{\mu}}V)-tG(V)\big)$$

- G = real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar{\mu}}M$
- spatial variation in $T_{\bar{\mu}}M \rightsquigarrow$ radial variation in $S_{\bar{\mu}}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar{\mu}}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\,\log_{\bar{\mu}}\bar{\mu}_n = \nabla_\mu\mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b}:\mathcal{P}_2M\to M$ sending $\mu\mapsto\bar{\mu}$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\,\log_{\bar{\mu}}\bar{\mu}_n \overset{d}{=} \nabla_{F_\mu\circ\exp_{\bar{\mu}}}\mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar{\mu}}M,\mathbb{R})$,
- of the minimizer map $\mathfrak{B}:\mathcal{C}(T_{\bar{\mu}}M,\mathbb{R})\to T_{\bar{\mu}}M$ that sends $f\mapsto\operatorname*{argmin}_{X\in C_\mu}f(X)$
- at $F_\mu\circ\exp_{\bar{\mu}}$
- along the Gaussian tangent field $G=G(\,\cdot\,)=\langle\Gamma_\mu,\cdot\,\rangle_{\bar{\mu}}$ induced by μ

New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty}\sqrt{n}\,\log_{\bar{\mu}}\bar{\mu}_n \overset{d}{=} \lim_{t\to 0}\frac{1}{t}\operatorname*{argmin}_{V\in T_{\bar{\mu}}M}\big(F_\mu(\exp_{\bar{\mu}}V) - tG(V)\big)$$

- G = real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar{\mu}}M$
- spatial variation in $T_{\bar{\mu}}M \rightsquigarrow$ radial variation in $S_{\bar{\mu}}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar{\mu}}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\,\log_{\bar{\mu}}\bar{\mu}_n = \nabla_\mu\mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b}:\mathcal{P}_2M\to M$ sending $\mu\mapsto\bar{\mu}$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty}\sqrt{n}\,\log_{\bar{\mu}}\bar{\mu}_n \overset{d}{=} \nabla_{F_\mu\circ\exp_{\bar{\mu}}}\mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar{\mu}}M,\mathbb{R})$,
- of the minimizer map $\mathfrak{B}:\mathcal{C}(T_{\bar{\mu}}M,\mathbb{R})\to T_{\bar{\mu}}M$ that sends $f\mapsto\operatorname*{argmin}_{X\in C_\mu}f(X)$
- at $F_\mu\circ\exp_{\bar{\mu}}$
- along the Gaussian tangent field $G = G(\cdot) = \langle\Gamma_\mu,\cdot\rangle_{\bar{\mu}}$ induced by μ

New interpretations of CLTs

Fundamental shifts in perspective via random fields or directional derivatives

CLT 2 [Mattingly, M–, Tran 2023]. Intrinsic, with Gaussian random field as limit:

$$\lim_{n\to\infty} \sqrt{n}\,\log_{\bar\mu} \bar\mu_n \overset{d}{=} \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{V\in T_{\bar\mu}M} \big(F_\mu(\exp_{\bar\mu} V) - t\,G(V)\big)$$

- G = real-valued Gaussian random field indexed by unit tangent sphere $S_{\bar\mu}M$
- spatial variation in $T_{\bar\mu}M \rightsquigarrow$ radial variation in $S_{\bar\mu}M$
- CLT proved via convergence of random fields: sidesteps nonlinearity of $T_{\bar\mu}M$!

CLT 3 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty} \sqrt{n}\,\log_{\bar\mu} \bar\mu_n = \nabla_\mu \mathfrak{b}(\Gamma_\mu)$,

- the directional derivative at μ, in the space $\mathcal{P}_2M$ of L^2 measures on M,
- of the barycenter map $\mathfrak{b} : \mathcal{P}_2M \to M$ sending $\mu \mapsto \bar\mu$
- along any Gaussian tangent mass Γ_μ

CLT 4 [Mattingly, M–, Tran 2023]. $\lim\limits_{n\to\infty} \sqrt{n}\,\log_{\bar\mu} \bar\mu_n \overset{d}{=} \nabla_{F_\mu\circ\exp_{\bar\mu}} \mathfrak{B}(G)$,

- the directional derivative, in the space of continuous maps $\mathcal{C}(T_{\bar\mu}M, \mathbb{R})$,
- of the minimizer map $\mathfrak{B} : \mathcal{C}(T_{\bar\mu}M, \mathbb{R}) \to T_{\bar\mu}M$ that sends $f \mapsto \operatorname*{argmin}\limits_{X\in C_\mu} f(X)$
- at $F_\mu \circ \exp_{\bar\mu}$
- along the Gaussian tangent field $G = G(\,\cdot\,) = \langle \Gamma_\mu, \cdot\, \rangle_{\bar\mu}$ induced by μ

Summary of singular setting

1. shadow geometry at singular points of CAT(κ) spaces `[arXiv:2311.09451]`
   - CLT $\leftrightarrow$ variation of Fréchet means Lec 1
   - $\bar{\mu}$ can only escape in certain directions Z Lec 2
   - "shadow" directions beyond opposite Z, at angle $\geq \pi$, collapse to a ray Lec 3
   - directions within π of Z remain metrically intact ⋮
2. geometry of measures on smoothly stratified spaces `[arXiv:2311.09453]`
   - smooth setting: reduce to linear via $\log_{\bar{\mu}}$ Lec 1
   - singular setting: take log from infinitesimally nearby $\bar{\mu}$ Lec 3
   - miracle: this "limit log" isometrically preserves all escape directions ⋮
   - carefully iterating yields tangential collapse $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$ ⋮
3. CLT for random tangent fields `[arXiv:2311.09454]`
   - spatial variation of $\bar{\mu}_n$ around $\bar{\mu}$ is too much to handle Lec 4
   - ask about variation along each ray separately ⋮
   - $\to$ Gaussian on that ray ⋮
   - multivariate Gaussian on any finite set of rays ⋮
4. escape vectors and stratified Gaussians `[arXiv:2311.09455]`
   - singular analogue of $N(0, \Sigma)$: Gaussian random mass Γ_μ Lec 4
   - $\mathscr{E}(x)$: where does $\bar{\mu}$ go when mass is added at x? Lec 5
   - $\mathscr{E}$ linearizes #3 ⋮
   - CLT $\leftrightarrow$ where does $\bar{\mu}$ go when Gaussian random mass Γ_μ is added? ⋮

Summary of singular setting

1. shadow geometry at singular points of CAT(κ) spaces [arXiv:2311.09451]
   - CLT ↔ variation of Fréchet means Lec 1
   - $\bar{\mu}$ can only escape in certain directions Z Lec 2
   - "shadow" directions beyond opposite Z, at angle $\geq \pi$, collapse to a ray Lec 3
   - directions within π of Z remain metrically intact ⋮
2. geometry of measures on smoothly stratified spaces [arXiv:2311.09453]
   - smooth setting: reduce to linear via $\log_{\bar{\mu}}$ Lec 1
   - singular setting: take log from infinitesimally nearby $\bar{\mu}$ Lec 3
   - miracle: this "limit log" isometrically preserves all escape directions ⋮
   - carefully iterating yields tangential collapse $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$ ⋮
3. CLT for random tangent fields [arXiv:2311.09454]
   - spatial variation of $\bar{\mu}_n$ around $\bar{\mu}$ is too much to handle Lec 4
   - ask about variation along each ray separately ⋮
   - → Gaussian on that ray ⋮
   - multivariate Gaussian on any finite set of rays ⋮
4. escape vectors and stratified Gaussians [arXiv:2311.09455]
   - singular analogue of $N(0, \Sigma)$: Gaussian random mass Γ_μ Lec 4
   - $\mathscr{E}(x)$: where does $\bar{\mu}$ go when mass is added at x? Lec 5
   - $\mathscr{E}$ linearizes #3 ⋮
   - CLT ↔ where does $\bar{\mu}$ go when Gaussian random mass Γ_μ is added? ⋮

Summary of singular setting

1. shadow geometry at singular points of CAT(κ) spaces [arXiv:2311.09451]
    - CLT $\leftrightarrow$ variation of Fréchet means Lec 1
    - $\bar{\mu}$ can only escape in certain directions Z Lec 2
    - "shadow" directions beyond opposite Z, at angle $\geq \pi$, collapse to a ray Lec 3
    - directions within π of Z remain metrically intact
2. geometry of measures on smoothly stratified spaces [arXiv:2311.09453]
    - smooth setting: reduce to linear via $\log_{\bar{\mu}}$ Lec 1
    - singular setting: take log from infinitesimally nearby $\bar{\mu}$ Lec 3
    - miracle: this "limit log" isometrically preserves all escape directions
    - carefully iterating yields tangential collapse $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$
3. CLT for random tangent fields [arXiv:2311.09454]
    - spatial variation of $\bar{\mu}_n$ around $\bar{\mu}$ is too much to handle Lec 4
    - ask about variation along each ray separately
    - $\to$ Gaussian on that ray
    - multivariate Gaussian on any finite set of rays
4. escape vectors and stratified Gaussians [arXiv:2311.09455]
    - singular analogue of $N(0, \Sigma)$: Gaussian random mass Γ_μ Lec 4
    - $\mathscr{E}(x)$: where does $\bar{\mu}$ go when mass is added at x? Lec 5
    - $\mathscr{E}$ linearizes #3
    - CLT $\leftrightarrow$ where does $\bar{\mu}$ go when Gaussian random mass Γ_μ is added?

Summary of singular setting

1. shadow geometry at singular points of CAT(κ) spaces `[arXiv:2311.09451]`
   - CLT ↔ variation of Fréchet means Lec 1
   - $\bar{\mu}$ can only escape in certain directions Z Lec 2
   - "shadow" directions beyond opposite Z, at angle $\geq \pi$, collapse to a ray Lec 3
   - directions within π of Z remain metrically intact ⋮
2. geometry of measures on smoothly stratified spaces `[arXiv:2311.09453]`
   - smooth setting: reduce to linear via $\log_{\bar{\mu}}$ Lec 1
   - singular setting: take log from infinitesimally nearby $\bar{\mu}$ Lec 3
   - miracle: this "limit log" isometrically preserves all escape directions ⋮
   - carefully iterating yields tangential collapse $\mathcal{L} : T_{\bar{\mu}}M \to \mathbb{R}^m$ ⋮
3. CLT for random tangent fields `[arXiv:2311.09454]`
   - spatial variation of $\bar{\mu}_n$ around $\bar{\mu}$ is too much to handle Lec 4
   - ask about variation along each ray separately ⋮
   - $\to$ Gaussian on that ray ⋮
   - multivariate Gaussian on any finite set of rays ⋮
4. escape vectors and stratified Gaussians `[arXiv:2311.09455]`
   - singular analogue of $N(0, \Sigma)$: Gaussian random mass Γ_μ Lec 4
   - $\mathscr{E}(x)$: where does $\bar{\mu}$ go when mass is added at x? Lec 5
   - $\mathscr{E}$ linearizes #3 ⋮
   - CLT ↔ where does $\bar{\mu}$ go when Gaussian random mass Γ_μ is added? ⋮

Looking forward

Interpretations of Gaussian objects on singular spaces

- heat dissipation
- random walks
- infinite divisibility of probability distributions

Statistical developments

- convergence rates
- confidence regions
- geometric PCA, e.g., in the sense of [Marron, et al. since 2010s]
- smoothness/singularity testing
- learning stratified spaces
- singular influence functions

Infinite-dimensional singular settings

- persistence diagrams [Mileyko, Mukherjee, Harer 2011]
- spaces of measures [Lott 2006]

Looking forward

Interpretations of Gaussian objects on singular spaces

- heat dissipation
- random walks
- infinite divisibility of probability distributions

Statistical developments

- convergence rates
- confidence regions
- geometric PCA, e.g., in the sense of [Marron, et al. since 2010s]
- smoothness/singularity testing
- learning stratified spaces
- singular influence functions

Infinite-dimensional singular settings

- persistence diagrams [Mileyko, Mukherjee, Harer 2011]
- spaces of measures [Lott 2006]

Looking forward

Interpretations of Gaussian objects on singular spaces

- heat dissipation
- random walks
- infinite divisibility of probability distributions

Statistical developments

- convergence rates
- confidence regions
- geometric PCA, e.g., in the sense of [Marron, et al. since 2010s]
- smoothness/singularity testing
- learning stratified spaces
- singular influence functions

Infinite-dimensional singular settings

- persistence diagrams [Mileyko, Mukherjee, Harer 2011]
- spaces of measures [Lott 2006]

Looking forward

Curvature invariants from distortion maps and tangential collapse

- generalize 2D angle deficit
- variation from point to point in M
- integrate to reflect topology of singular spaces?
- compare with singular homology or intersection cohomology
- how to construct measures with given Fréchet mean?

Functoriality and moduli

- distortion $\leftrightarrow$ how CLT transforms under morphism
- proposal for real or complex variety X:
  - take resolution of singularities $\widetilde{X} \to X$
  - push CLT on $\widetilde{X}$ forward to X
  - correction terms should involve local sheaf-theoretic data around $\bar{\mu}$
  - conj: results in well defined CLT on X
  - e.g.: compare pushforward CLT with singular CLT in smoothly stratified case
  - analogy: multiplier ideals
- asymptotics of sampling from moduli spaces
  - statistical invariants $\leftrightarrow$ typical or expected variation of algebraic structures
  - in neighborhoods of a fixed degeneration

Looking forward

Curvature invariants from distortion maps and tangential collapse

- generalize 2D angle deficit
- variation from point to point in M
- integrate to reflect topology of singular spaces?
- compare with singular homology or intersection cohomology
- how to construct measures with given Fréchet mean?

Functoriality and moduli

- distortion ↔ how CLT transforms under morphism
- proposal for real or complex variety X:
  - take resolution of singularities $\widetilde{X} \to X$
  - push CLT on $\widetilde{X}$ forward to X
  - correction terms should involve local sheaf-theoretic data around $\bar{\mu}$
  - conj: results in well defined CLT on X
  - e.g.: compare pushforward CLT with singular CLT in smoothly stratified case
  - analogy: multiplier ideals
- asymptotics of sampling from moduli spaces
  - statistical invariants ↔ typical or expected variation of algebraic structures
  - in neighborhoods of a fixed degeneration

Looking forward

Curvature invariants from distortion maps and tangential collapse

- generalize 2D angle deficit
- variation from point to point in M
- integrate to reflect topology of singular spaces?
- compare with singular homology or intersection cohomology
- how to construct measures with given Fréchet mean?

Functoriality and moduli

- distortion ↔ how CLT transforms under morphism
- proposal for real or complex variety X:
  - + take resolution of singularities $\widetilde{X} \to X$
  - + push CLT on $\widetilde{X}$ forward to X
  - + correction terms should involve local sheaf-theoretic data around $\bar{\mu}$
  - + conj: results in well defined CLT on X
  - + e.g.: compare pushforward CLT with singular CLT in smoothly stratified case
  - + analogy: multiplier ideals
- asymptotics of sampling from moduli spaces
  - + statistical invariants ↔ typical or expected variation of algebraic structures
  - + in neighborhoods of a fixed degeneration

References

- Stephen Aylward and Elizabeth Bullitt, *Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction*, IEEE Trans. on Medical Imaging **21**, no. 2 (2002), 61–75.
- Dennis Barden and Huiling Le, *The logarithm map, its limits and Fréchet means in orthant spaces*, Proc. of the London Mathematical Society (3) **117** (2018), no. 4, 751–789.
- Dennis Barden, Huiling Le, and Megan Owen, *Central limit theorems for Fréchet means in the space of phylogenetic trees*, Electronic J. of Probability **18** (2013), no. 25, 25 pp.
- Dennis Barden, Huiling Le, and Megan Owen, *Limiting behaviour of Fréchet means in the space of phylogenetic trees*, Annals of the Institute of Statistical Mathematics **70** (2013), no. 1, 99–129.
- Rabi Bhattacharya and Lizhen Lin, *Omnibus CLTs for Fréchet means and nonparametric inference on non-Euclidean spaces*, Proc. AMS **145** (2017), no. 1, 413–428.
- Rabi Bhattacharya and Vic Patrangenaru, *Large sample theory of intrinsic and extrinsic sample means on manifolds: I*, Annals of Statistics **31** (2003), no. 1, 1–29.
- Rabi Bhattacharya and Vic Patrangenaru, *Large sample theory of intrinsic and extrinsic sample means on manifolds: II*, Annals of Statistics **33** (2005), no. 3, 1225–1259.
- Benjamin Eltzner and Stephan Huckemann, *A smeary central limit theorem for manifolds with application to high-dimensional spheres*, Ann. Stat. **47** (2019), no. 6, 3360–3381.
- Thomas Hotz, Stephan Huckemann, Huiling Le, J.S. Marron, Jonathan C. Mattingly, Ezra Miller, James Nolen, Megan Owen, Vic Patrangenaru, and Sean Skwerer, *Sticky central limit theorems on open books*, Annals of Applied Probability **23** (2013), no. 6, 2238–2258.
- Stephan Huckemann, Jonathan Mattingly, Ezra Miller, and James Nolen, *Sticky central limit theorems at isolated hyperbolic planar singularities*, Electronic Journal of Probability **20** (2015), 1–34.
- J. Kim, *Slicing hyperdimensional oranges: The geometry of phylogenetic estimation*, Mol. Phylogenet. Evol. **17**, no. 1 (2000), 58–75.
- Jonas Lueg, Maryam Garba, Tom Nye, and Stephan Huckemann, *Wald space for phylogenetic trees*, in Geometric Sci. of Inform., Lect. Notes in Comp. Sci. **12829** (2021), 710–717.
- Jonathan Mattingly, Ezra Miller, and Do Tran, *Shadow geometry at singular points of* CAT(κ) *spaces*, preprint, 2023. arXiv:math.MG/2311.09451
- Jonathan Mattingly, Ezra Miller, and Do Tran, *Geometry of measures on smoothly stratified metric spaces*, preprint, 2023. arXiv:math.MG/2311.09453
- Jonathan Mattingly, Ezra Miller, and Do Tran, *A central limit theorem for random tangent fields on stratified spaces*, preprint, 2023. arXiv:math.PR/2311.09454
- Jonathan Mattingly, Ezra Miller, and Do Tran, *Central limit theorems for Fréchet means on stratified spaces*, preprint, 2023. arXiv:math.PR/2311.09455
- Yuriy Mileyko, Sayan Mukherjee, and John Harer, *Probability measures on the space of persistence diagrams*, Inverse Problems **27.12** (2011): 124007.
- David Speyer and Bernd Sturmfels, *The tropical Grassmannian*, Adv. Geom. **4** (2004), no. 3, 389–411.

References

- Stephen Aylward and Elizabeth Bullitt, *Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction*, IEEE Trans. on Medical Imaging **21**, no. 2 (2002), 61–75.
- Dennis Barden and Huiling Le, *The logarithm map, its limits and Fréchet means in orthant spaces*, Proc. of the London Mathematical Society (3) **117** (2018), no. 4, 751–789.
- Dennis Barden, Huiling Le, and Megan Owen, *Central limit theorems for Fréchet means in the space of phylogenetic trees*, Electronic J. of Probability **18** (2013), no. 25, 25 pp.
- Dennis Barden, Huiling Le, and Megan Owen, *Limiting behaviour of Fréchet means in the space of phylogenetic trees*, Annals of the Institute of Statistical Mathematics **70** (2013), no. 1, 99–129.
- Rabi Bhattacharya and Lizhen Lin, *Omnibus CLTs for Fréchet means and nonparametric inference on non-Euclidean spaces*, Proc. AMS **145** (2017), no. 1, 413–428.
- Rabi Bhattacharya and Vic Patrangenaru, *Large sample theory of intrinsic and extrinsic sample means on manifolds: I*, Annals of Statistics **31** (2003), no. 1, 1–29.
- Rabi Bhattacharya and Vic Patrangenaru, *Large sample theory of intrinsic and extrinsic sample means on manifolds: II*, Annals of Statistics **33** (2005), no. 3, 1225–1259.
- Benjamin Eltzner and Stephan Huckemann, *A smeary central limit theorem for manifolds with application to high-dimensional spheres*, Ann. Stat. **47** (2019), no. 6, 3360–3381.
- Thomas Hotz, Stephan Huckemann, Huiling Le, J.S. Marron, Jonathan C. Mattingly, Ezra Miller, James Nolen, Megan Owen, Vic Patrangenaru, and Sean Skwerer, *Sticky central limit theorems on open books*, Annals of Applied Probability **23** (2013), no. 6, 2238–2258.
- Stephan Huckemann, Jonathan Mattingly, Ezra Miller, and James Nolen, *Sticky central limit theorems at isolated hyperbolic planar singularities*, Electronic Journal of Probability **20** (2015), 1–34.
- J. Kim, *Slicing hyperdimensional oranges: The geometry of phylogenetic estimation*, Mol. Phylogenet. Evol. **17**, no. 1 (2000), 58–75.
- Jonas Lueg, Maryam Garba, Tom Nye, and Stephan Huckemann, *Wald space for phylogenetic trees*, in Geometric Sci. of Inform., Lect. Notes in Comp. Sci. **12829** (2021), 710–717.
- Jonathan Mattingly, Ezra Miller, and Do Tran, *Shadow geometry at singular points of* CAT(κ) *spaces*, preprint, 2023. arXiv:math.MG/2311.09451
- Jonathan Mattingly, Ezra Miller, and Do Tran, *Geometry of measures on smoothly stratified metric spaces*, preprint, 2023. arXiv:math.MG/2311.09453
- Jonathan Mattingly, Ezra Miller, and Do Tran, *A central limit theorem for random tangent fields on stratified spaces*, preprint, 2023. arXiv:math.PR/2311.09454
- Jonathan Mattingly, Ezra Miller, and Do Tran, *Central limit theorems for Fréchet means on stratified spaces*, preprint, 2023. arXiv:math.PR/2311.09455
- Yuriy Mileyko, Sayan Mukherjee, and John Harer, *Probability measures on the space of persistence diagrams*, Inverse Problems **27.12** (2011): 124007.
- David Speyer and Bernd Sturmfels, *The tropical Grassmannian*, Adv. Geom. **4** (2004), no. 3, 389–411.

Next lecture: geometry and measure on CAT(κ) spaces