Computing medians and means of phylogenetic trees

Miroslav Bacak Philipp Benner

Max Planck Institute, Leipzig

Mathematical and Computational Evolutionary Biology, Hameau de l'Etoile, May 27–31, 2013

Contents of the talk

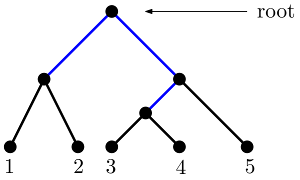
n-trees

Definition

A *metric n-tree* is a tree (connected graph with no circuit) with

- a distinguished vertex called *root*,
- n vertices called *leaves* that are labeled $1, \ldots, n$,
- leaf and inner edges of positive length.

5-tree

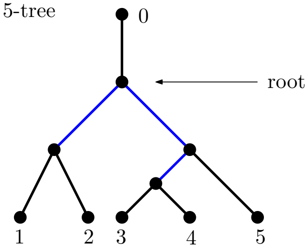


n-trees

Definition

A *metric n-tree* is a tree (connected graph with no circuit) with

- a distinguished vertex called *root*,
- n vertices called *leaves* that are labeled $1, \ldots, n$,
- leaf and inner edges of positive length.



Need for a space of trees

We would like to

- measure distances between a given pair of trees,
- compute medians and means of a given set of trees.

We hence need a space of trees.

Construction due to **Billera, Holmes, and Vogtmann** in 2001:

BHV Tree space: a metric space whose points are trees.

(Metric space means we can measure distances.)

Moreover, tree space is an **Hadamard space** (i.e. it is nice).

Need for a space of trees

We would like to

- measure distances between a given pair of trees,
- compute medians and means of a given set of trees.

We hence need a space of trees.

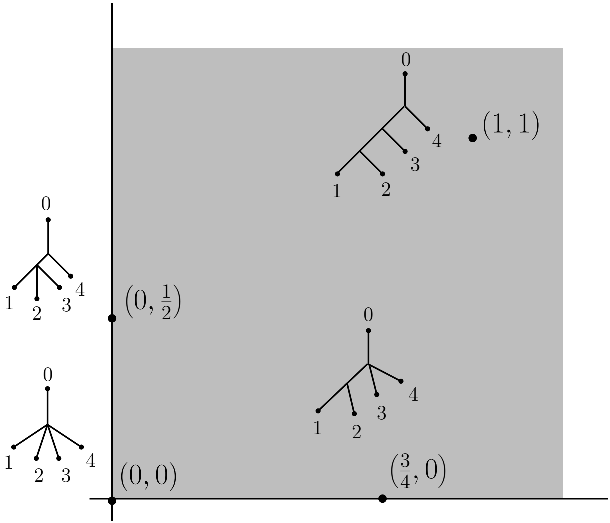
Construction due to **Billera, Holmes, and Vogtmann** in 2001:

BHV Tree space: a metric space whose points are trees.

(Metric space means we can measure distances.)

Moreover, tree space is an **Hadamard space** (i.e. it is nice).

Orthant representation of a 4-tree



A piece of tree space $\mathcal{T}_4$

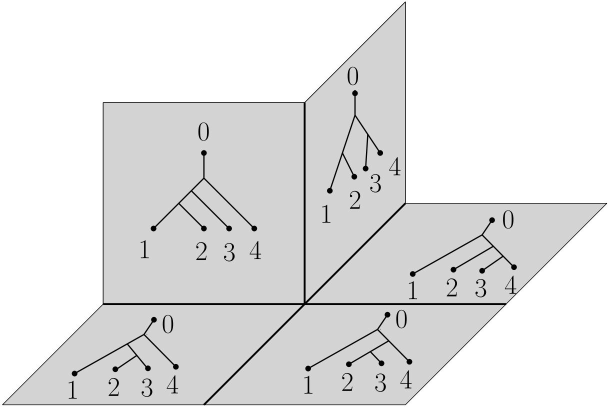


Figure : 5 out of 15 orthants of $\mathcal{T}_4$

Tree space

It is easy to define a metric in $\mathcal{T}_n$ – induced by Euclidean distances.

(Hence we are able to measure distances.)

Geodesics are piecewise linear (broken line segments).

(Geodesic = shortest path between a given pair of points.)

Theorem (Billera, Holmes, Vogtmann)

Tree space $\mathcal{T}_n$ is an Hadamard space.

(Hadamard space = geodesic space of non-positive curvature.)

Tree space

It is easy to define a metric in $\mathcal{T}_n$ – induced by Euclidean distances.

(Hence we are able to measure distances.)

Geodesics are piecewise linear (broken line segments).

(Geodesic = shortest path between a given pair of points.)

Theorem (Billera, Holmes, Vogtmann)

Tree space $\mathcal{T}_n$ is an Hadamard space.

(Hadamard space = geodesic space of non-positive curvature.)

Tree space

It is easy to define a metric in $\mathcal{T}_n$ – induced by Euclidean distances.

(Hence we are able to measure distances.)

Geodesics are piecewise linear (broken line segments).

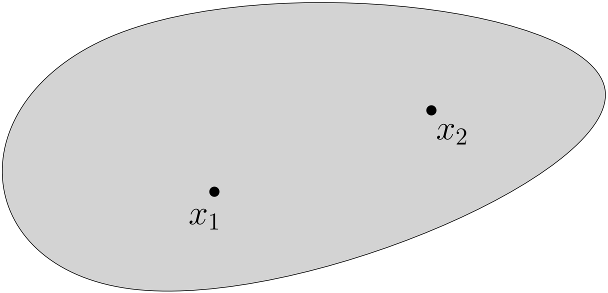
(Geodesic $=$ shortest path between a given pair of points.)

Theorem (Billera, Holmes, Vogtmann)

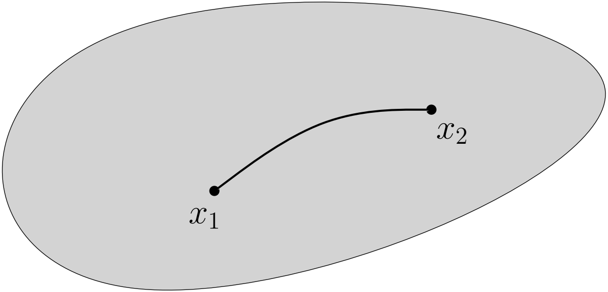
Tree space $\mathcal{T}_n$ is an Hadamard space.

(Hadamard space $=$ geodesic space of non-positive curvature.)

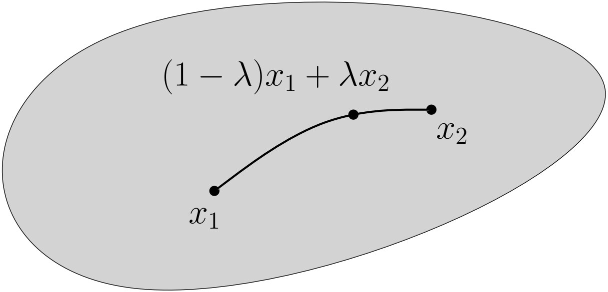
Geodesic space



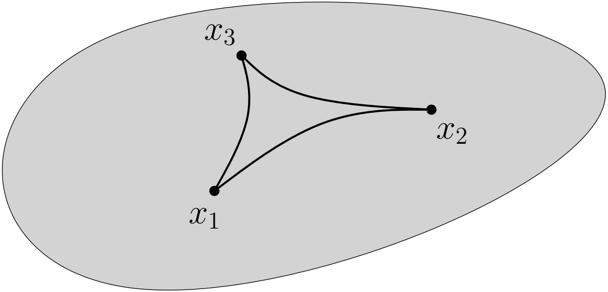
Geodesic space



Geodesic space



Definition of nonpositive curvature



The Fréchet mean

The arithmetic mean of $x_1, \ldots, x_K \in \mathbb{R}^m$ is defined as

$$\Xi\left(x_1, \ldots, x_K\right) := \frac{x_1 + \cdots + x_K}{K} = \frac{K-1}{K} \Xi\left(x_1, \ldots, x_{K-1}\right) + \frac{1}{K} x_K .$$

We cannot directly extend this into tree space.

Theorem (Methode der kleinsten Quadrate, Gauss, 1809)

The arithmetic mean is the unique vector in $\mathbb{R}^m$ such that

$$\sum_{k=1}^{K} d\left(\Xi, x_k\right)^2 = \min_{y \in \mathbb{R}^m} \sum_{k=1}^{K} d\left(y, x_k\right)^2 .$$

The Fréchet mean

The arithmetic mean of $x_1, \ldots, x_K \in \mathbb{R}^m$ is defined as

$$\Xi\left(x_1, \ldots, x_K\right) := \frac{x_1 + \cdots + x_K}{K} = \frac{K-1}{K} \Xi\left(x_1, \ldots, x_{K-1}\right) + \frac{1}{K} x_K .$$

We cannot directly extend this into tree space.

Theorem (Methode der kleinsten Quadrate, Gauss, 1809)

The arithmetic mean is the unique vector in $\mathbb{R}^m$ such that

$$\sum_{k=1}^{K} d\left(\Xi, x_k\right)^2 = \min_{y \in \mathbb{R}^m} \sum_{k=1}^{K} d\left(y, x_k\right)^2 .$$

The Fréchet mean

The arithmetic mean of $x_1, \dots, x_K \in \mathbb{R}^m$ is defined as

$$\Xi\left(x_1, \dots, x_K\right) := \frac{x_1 + \cdots + x_K}{K} = \frac{K-1}{K} \Xi\left(x_1, \dots, x_{K-1}\right) + \frac{1}{K} x_K.$$

We cannot directly extend this into tree space.

Theorem (Methode der kleinsten Quadrate, Gauss, 1809)

The arithmetic mean is the unique vector in $\mathbb{R}^m$ such that

$$\sum_{k=1}^{K} d\left(\Xi, x_k\right)^2 = \min_{y \in \mathbb{R}^m} \sum_{k=1}^{K} d\left(y, x_k\right)^2.$$

The Fréchet mean

The arithmetic mean of $x_1, \ldots, x_K \in \mathbb{R}^m$ is defined as

$$\Xi(x_1, \ldots, x_K) := \frac{x_1 + \cdots + x_K}{K} = \frac{K-1}{K} \Xi(x_1, \ldots, x_{K-1}) + \frac{1}{K} x_K.$$

We cannot directly extend this into tree space.

Theorem (Methode der kleinsten Quadrate, Gauss, 1809)

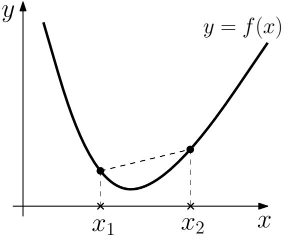
The arithmetic mean is the unique vector in $\mathbb{R}^m$ such that

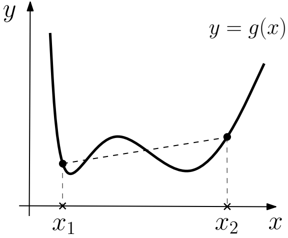
$$\sum_{k=1}^{K} d(\Xi, x_k)^2 = \min_{y \in \mathbb{R}^m} \sum_{k=1}^{K} d(y, x_k)^2.$$

Convexity in tree space

Definition (Convex function)

A function $f : \mathcal{T}_n \to (-\infty, \infty]$ is *convex* if $f \circ \gamma$ is a convex function for any geodesic $\gamma : [0, 1] \to \mathcal{T}_n$.

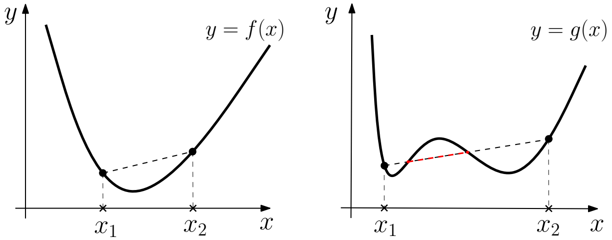




Convexity in tree space

Definition (Convex function)

A function $f : \mathcal{T}_n \to (-\infty, \infty]$ is *convex* if $f \circ \gamma$ is a convex function for any geodesic $\gamma : [0, 1] \to \mathcal{T}_n$.



Definition of the Fréchet mean

Let $T_1, \ldots, T_K \in \mathcal{T}_n$. The function

$$\xi(S) := \sum_{k=1}^{K} d\left(S, T_k\right)^2$$

is (strongly) convex and continuous. (**By nonpositive curvature.**)

Theorem

1. *There exists a unique minimizer $\Xi \in \mathcal{T}_n$ of the function ξ.*
2. *The function $\Xi = \Xi\left(T_1, \ldots, T_K\right)$ is Lipschitz.*

This Ξ is called the *Fréchet mean* of $\{T_1, \ldots, T_K\}$.

Definition of the Fréchet mean

Let $T_1, \dots, T_K \in \mathcal{T}_n$. The function

$$\xi(S) := \sum_{k=1}^{K} d\left(S, T_k\right)^2$$

is (strongly) convex and continuous. (**By nonpositive curvature.**)

Theorem

1. *There exists a unique minimizer* $\Xi \in \mathcal{T}_n$ *of the function* ξ.
2. *The function* $\Xi = \Xi\left(T_1, \dots, T_K\right)$ *is Lipschitz.*

This Ξ is called the *Fréchet mean* of $\left\{T_1, \dots, T_K\right\}$.

Probabilistic interpretation of the mean

Let $T_1, \ldots, T_K \in \mathcal{T}_n$. Denote the probability measure

$$\pi := \frac{1}{K} \sum_{k=1}^{K} \delta_{T_k}.$$

We can consider a random variable $Y : \Omega \to \mathcal{T}_n$ with distr. π.

If each of the values $T_1, \ldots, T_K$ occurs with probability $\frac{1}{K}$, then

$$\mathbb{E}Y := \operatorname*{arg\,min}_{S \in \mathcal{T}_n} \frac{1}{K} \sum_{k=1}^{K} d\left(S, T_k\right)^2 = \Xi\left(T_1, \ldots, T_K\right)$$

is the expectation of Y. (Also called the barycenter of π.)

The law of large numbers

Given a sequence of random variables Y_i with values in $\mathcal{T}_n$, we define $S_1 := Y_1$, and

$$S_{i+1} := \frac{i}{i+1} S_i + \frac{1}{i+1} Y_{i+1},$$

Theorem (The law of large numbers, Sturm 2003)

Let (Y_i) be a sequence i.i.d. according to π. Then

$$S_i \to \Xi\left(T_1, \ldots, T_K\right), \quad \text{as } i \to \infty,$$

almost everywhere.

The law of large numbers

Given a sequence of random variables Y_i with values in $\mathcal{T}_n$, we define $S_1 := Y_1$, and

$$S_{i+1} := \frac{i}{i+1} S_i + \frac{1}{i+1} Y_{i+1},$$

Theorem (The law of large numbers, Sturm 2003)

Let (Y_i) be a sequence i.i.d. according to π. Then

$$S_i \to \Xi\left(T_1, \ldots, T_K\right), \quad \text{as } i \to \infty,$$

almost everywhere.

Geometric median

Let $T_1, \ldots, T_K \in \mathcal{T}_n$. Then

$$\psi(S) := \sum_{k=1}^{K} d\left(S, T_k\right)$$

is convex and continuous on $\mathcal{T}_n$.

($=$ the *Fermat-Weber problem* for optimal facility location)

Theorem

1. *There exists a minimizer $\Psi \in \mathcal{T}_n$ of the function ψ.*
2. *The minimizer is unique unless all the points lie on a geodesic.*

This Ψ is called the *geometric median* of $\{T_1, \ldots, T_K\}$.

Geometric median

Let $T_1, \ldots, T_K \in \mathcal{T}_n$. Then

$$\psi(S) := \sum_{k=1}^{K} d\left(S, T_k\right)$$

is convex and continuous on $\mathcal{T}_n$.

(= the *Fermat-Weber problem* for optimal facility location)

Theorem

1. *There exists a minimizer $\Psi \in \mathcal{T}_n$ of the function ψ.*
2. *The minimizer is unique unless all the points lie on a geodesic.*

This Ψ is called the *geometric median* of $\{T_1, \ldots, T_K\}$.

Geometric median in $\mathbb{R}^m$

Let $x_1, \ldots, x_K \in \mathbb{R}^m$ and

$$\psi(y) := \sum_{k=1}^{K} d\left(y, x_k\right).$$

No explicit formula for a minimizer, only approximation algorithms, e.g. Weiszfeld's algorithm. (Compare with means.)

In $\mathbb{R}$ it coincides with the usual definition of a median:

$$\Pr(Y \leq \mu) \geq \frac{1}{2} \quad \text{and} \quad \Pr(Y \geq \mu) \geq \frac{1}{2},$$

where $Y : \mathbb{R} \to \mathbb{R}$ is a random variable. Then μ is a median of Y.

Geometric median in $\mathbb{R}^m$

Let $x_1, \ldots, x_K \in \mathbb{R}^m$ and

$$\psi(y) := \sum_{k=1}^{K} d\left(y, x_k\right).$$

No explicit formula for a minimizer, only approximation algorithms, e.g. Weiszfeld's algorithm. (Compare with means.)

In $\mathbb{R}$ it coincides with the usual definition of a median:

$$\Pr(Y \leq \mu) \geq \frac{1}{2} \quad \text{and} \quad \Pr(Y \geq \mu) \geq \frac{1}{2},$$

where $Y : \mathbb{R} \to \mathbb{R}$ is a random variable. Then μ is a median of Y.

Let $f : \mathcal{T}_n \to \mathbb{R}$ be convex continuous function.

Assume f attains its minimum. How to **compute** a minimizer?

Algorithm (Proximal point algorithm)

Choose $S_0 \in \mathcal{T}_n$ and set

$$S_{i+1} := \arg\min_{T \in \mathcal{T}_n} \left[f(T) + \frac{1}{2\lambda_i} d(T, S_i)^2 \right],$$

for $i \in \mathbb{N}$.

The sequence S_i converges to a minimizer of f.

This is a classical optimization method in $\mathbb{R}^m$.

Works also in Hadamard spaces (M.B. 2011)

Let $f : \mathcal{T}_n \to \mathbb{R}$ be convex continuous function.

Assume f attains its minimum. How to **compute** a minimizer?

Algorithm (Proximal point algorithm)

Choose $S_0 \in \mathcal{T}_n$ *and set*

$$S_{i+1} := \underset{T \in \mathcal{T}_n}{\arg\min} \left[f(T) + \frac{1}{2\lambda_i} d(T, S_i)^2 \right],$$

for $i \in \mathbb{N}$.

The sequence S_i converges to a minimizer of f.

This is a classical optimization method in $\mathbb{R}^m$.

Works also in Hadamard spaces (M.B. 2011)

Let $f : \mathcal{T}_n \to \mathbb{R}$ be convex continuous function.

Assume f attains its minimum. How to **compute** a minimizer?

Algorithm (Proximal point algorithm)

Choose $S_0 \in \mathcal{T}_n$ *and set*

$$S_{i+1} := \operatorname*{arg\,min}_{T \in \mathcal{T}_n} \left[f(T) + \frac{1}{2\lambda_i} d(T, S_i)^2 \right],$$

for $i \in \mathbb{N}$.

The sequence S_i converges to a minimizer of f.

This is a classical optimization method in $\mathbb{R}^m$.

Works also in Hadamard spaces (M.B. 2011)

Splitting proximal point algorithm

Let $f_1, \ldots, f_K$ be convex continuous and consider

$$f(T) := \sum_{k=1}^{K} f_k(T), \qquad T \in \mathcal{T}_n.$$

Example (Median and mean)

$$\psi(T) := \sum_{k=1}^{K} d\left(T, T_k\right), \qquad \xi(T) := \sum_{k=1}^{K} d\left(T, T_k\right)^2.$$

Key idea: use the PPA for $f_1, \ldots, f_K$ in a cyclic or random order.

Splitting proximal point algorithm

Let $f_1, \dots, f_K$ be convex continuous and consider

$$f(T) := \sum_{k=1}^{K} f_k(T), \qquad T \in \mathcal{T}_n.$$

Example (Median and mean)

$$\psi(T) := \sum_{k=1}^{K} d\left(T, T_k\right), \qquad \xi(T) := \sum_{k=1}^{K} d\left(T, T_k\right)^2 .$$

Key idea: use the PPA for $f_1, \dots, f_K$ in a cyclic or random order.

Splitting proximal point algorithm

Let $f_1, \ldots, f_K$ be convex continuous and consider

$$f(T) := \sum_{k=1}^{K} f_k(T), \qquad T \in \mathcal{T}_n.$$

Example (Median and mean)

$$\psi(T) := \sum_{k=1}^{K} d\left(T, T_k\right), \qquad \xi(T) := \sum_{k=1}^{K} d\left(T, T_k\right)^2.$$

Key idea: use the PPA for $f_1, \ldots, f_K$ in a cyclic or random order.

Splitting proximal point algorithm (for mean)

Hence instead of computing (the usual PPA)

$$S_{i+1} := \operatorname*{arg\,min}_{T \in \mathcal{T}_n} \left[\sum_{k=1}^{K} d\left(T, T_k\right)^2 + \frac{1}{2\lambda_i} d\left(T, S_i\right)^2 \right],$$

we are to minimize the function

$$S_{i+1} := \operatorname*{arg\,min}_{T \in \mathcal{T}_n} \left[d\left(T, T_k\right)^2 + \frac{1}{2\lambda_i} d\left(T, S_i\right)^2 \right],$$

where T_k chosen by a selection rule (cyclic/random).

This is a **one-dimensional** problem!

$\implies$ S_{i+1} is a convex combination of T_k and S_i.

Splitting proximal point algorithm (for mean)

Hence instead of computing (the usual PPA)

$$S_{i+1} := \operatorname*{arg\,min}_{T \in \mathcal{T}_n} \left[\sum_{k=1}^{K} d\left(T, T_k\right)^2 + \frac{1}{2\lambda_i} d\left(T, S_i\right)^2 \right],$$

we are to minimize the function

$$S_{i+1} := \operatorname*{arg\,min}_{T \in \mathcal{T}_n} \left[d\left(T, T_k\right)^2 + \frac{1}{2\lambda_i} d\left(T, S_i\right)^2 \right],$$

where T_k chosen by a selection rule (cyclic/random).

This is a **one-dimensional** problem!

$\Longrightarrow \quad S_{i+1}$ is a convex combination of T_k and S_i.

Computing the mean: Cyclic order version

Algorithm (M.B. 2012)

Input: $T_1, \ldots, T_K \in \mathcal{T}_n$

Step 1: $S_1 := T_1$ *and* $i := 1$

Step 2: $q := \lceil \frac{i}{K} \rceil$ *and* $p := i \mod K$

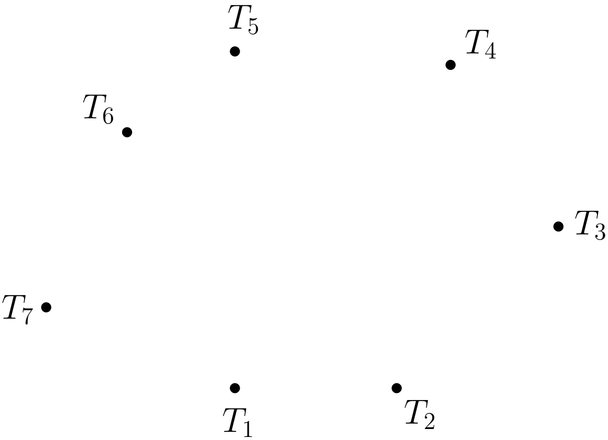
Step 3: $S_{i+1} := \frac{q}{q+1} S_i + \frac{1}{q+1} T_p$

Step 4: $i := i + 1$

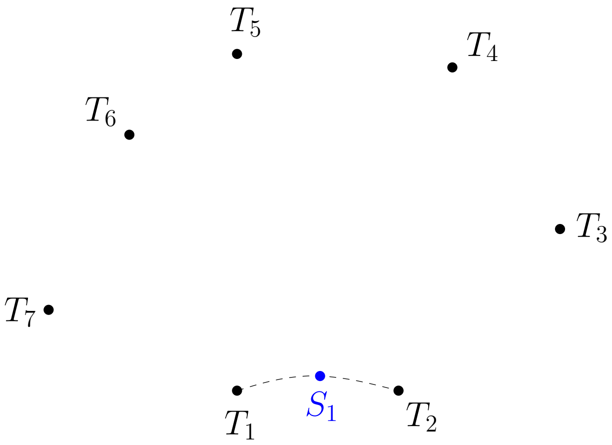
Step 5: *go to Step 2*

The sequence S_i converges to the mean of $T_1, \ldots, T_K$.

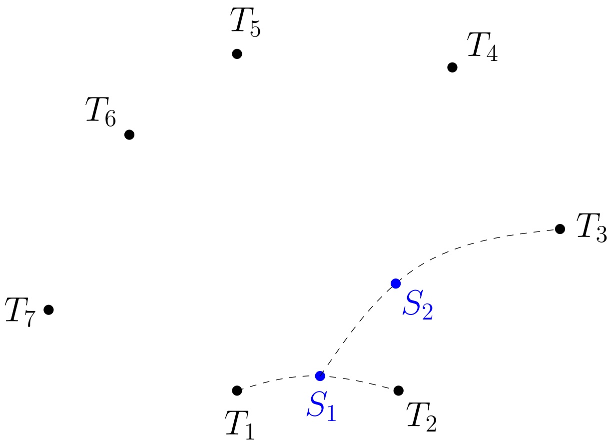
Computing the mean: Cyclic order version



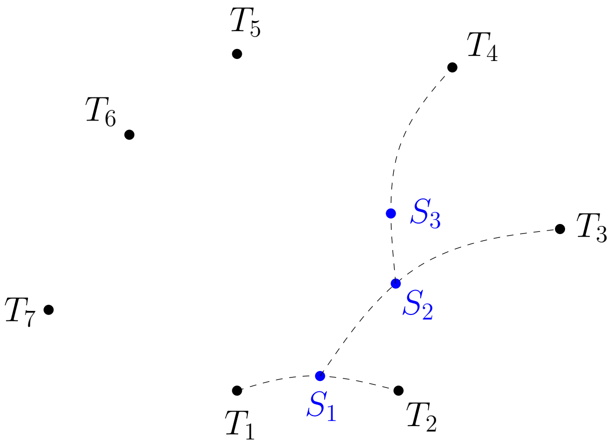
Computing the mean: Cyclic order version



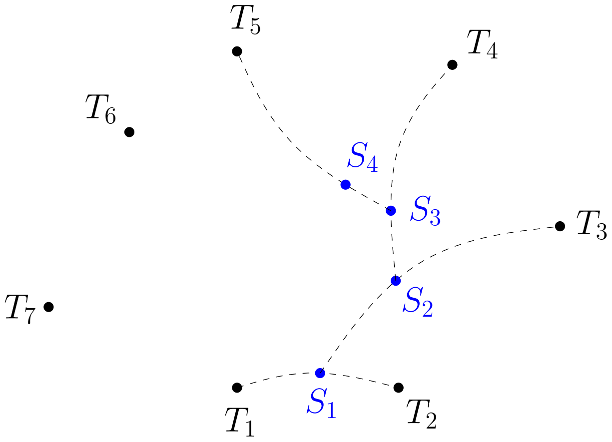
Computing the mean: Cyclic order version



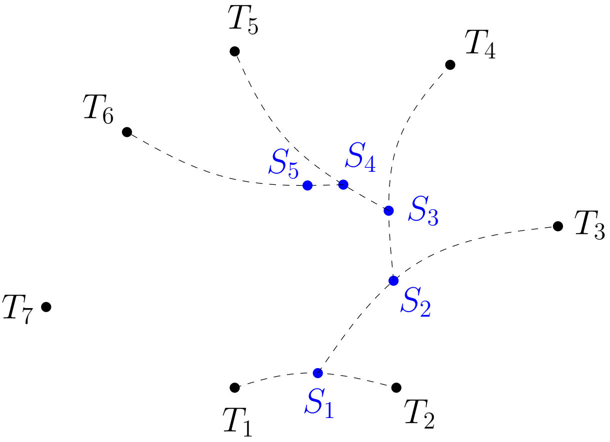
Computing the mean: Cyclic order version



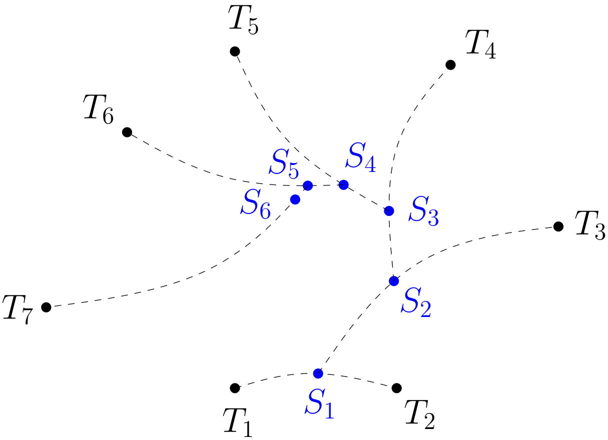
Computing the mean: Cyclic order version



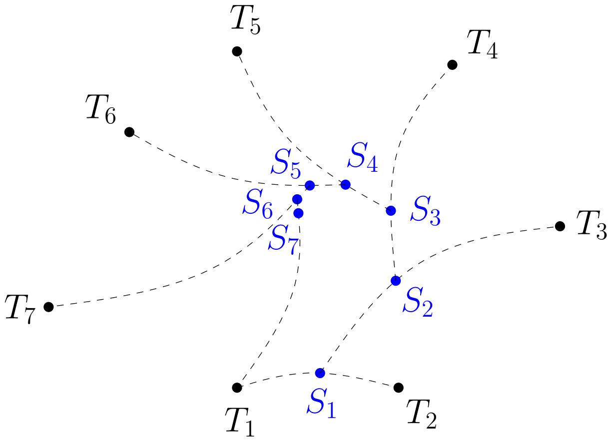
Computing the mean: Cyclic order version



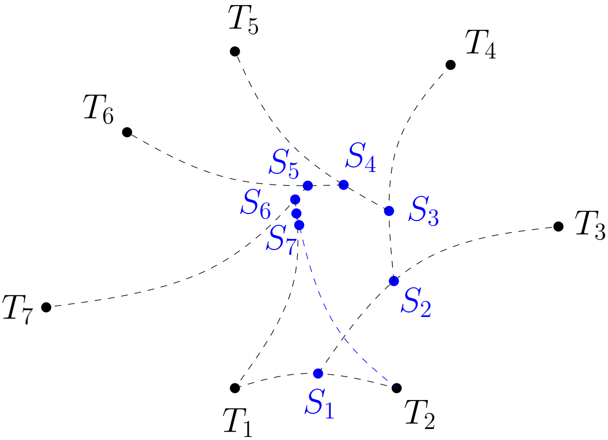
Computing the mean: Cyclic order version



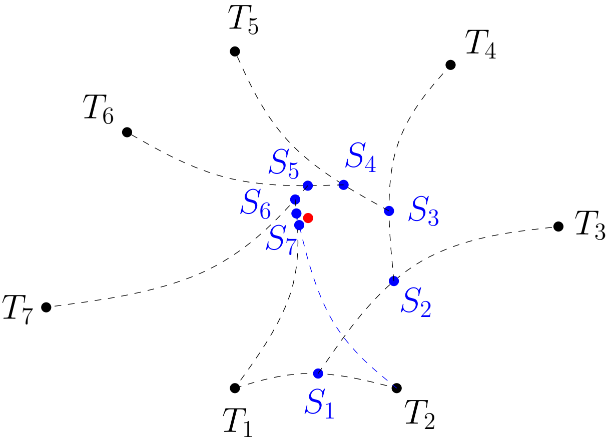
Computing the mean: Cyclic order version



Computing the mean: Cyclic order version



Computing the mean: Cyclic order version



Computing the mean: Cyclic order version

Algorithm (revisited)

Input: $T_1, \ldots, T_K \in \mathcal{T}_n$

Step 1: $S_1 := T_1$ *and* $i := 1$

Step 2: $q := \lceil \frac{i}{K} \rceil$ *and* $p := i \mod K$

Step 3: $S_{i+1} := \frac{q}{q+1} S_i + \frac{1}{q+1} T_p$

Step 4: $i := i + 1$

Step 5: *go to Step 2*

Geodesics can be computed in **polynomial** time:

The Owen-Provan algorithm (2011)

Computing the mean: Cyclic order version

Algorithm (revisited)

Input: $T_1, \dots, T_K \in \mathcal{T}_n$

Step 1: $S_1 := T_1$ *and* $i := 1$

Step 2: $q := \lceil \frac{i}{K} \rceil$ *and* $p := i \mod K$

Step 3: $S_{i+1} := \frac{q}{q+1} S_i + \frac{1}{q+1} T_p$

Step 4: $i := i + 1$

Step 5: *go to Step 2*

Geodesics can be computed in **polynomial** time:

The Owen-Provan algorithm (2011)

Computing the mean: Cyclic order version

Algorithm (revisited)

Input: $T_1, \ldots, T_K \in \mathcal{T}_n$

Step 1: $S_1 := T_1$ *and* $i := 1$

Step 2: $q := \lceil \frac{i}{K} \rceil$ *and* $p := i \mod K$

Step 3: $S_{i+1} := \frac{q}{q+1} S_i + \frac{1}{q+1} T_p$

Step 4: $i := i + 1$

Step 5: *go to Step 2*

Geodesics can be computed in **polynomial** time:

The Owen-Provan algorithm (2011)

Statistical model (see Philipp Benner's poster for details)

We start with multiple sequence alignments $\rightsquigarrow$

Posterior distribution is defined:

- first on each orthant $\mathcal{O}_i$ of tree space (fixed tree topology) $\implies \mu_i$
- posterior distribution on the whole tree space $\mathcal{T}_n$:

$$\mu := \sum_{i=1}^{(2n-3)!!} w_i \mu_i.$$

Difficulties:

- the weights w_i require to compute a complicated integral
- the number of orthants (tree topologies) is **big:** $(2n-3)!!$

Statistical model (see Philipp Benner's poster for details)

We start with multiple sequence alignments $\rightsquigarrow$

Posterior distribution is defined:

- first on each orthant $\mathcal{O}_i$ of tree space (fixed tree topology) $\Longrightarrow \mu_i$
- posterior distribution on the whole tree space $\mathcal{T}_n$:

$$\mu := \sum_{i=1}^{(2n-3)!!} w_i \mu_i .$$

Difficulties:

- the weights w_i require to compute a complicated integral
- the number of orthants (tree topologies) is **big:** $(2n-3)!!$

Statistical model - continued

We'll therefore give point estimates of posterior distribution μ:

- median:

$$\operatorname*{arg\,min}_{S\in\mathcal{T}_n} \int_{\mathcal{T}_n} d(S,T)\ \mathrm{d}\mu(T)$$

- mean:

$$\operatorname*{arg\,min}_{S\in\mathcal{T}_n} \int_{\mathcal{T}_n} d(S,T)^2\ \mathrm{d}\mu(T)$$

Markov chain Monte Carlo (MCMC) methods yield samples of posterior distribution:

$\rightsquigarrow T_1,\ldots,T_K \in \mathcal{T}_n$ $\qquad \rightsquigarrow \pi := \frac{1}{K}\sum_{k=1}^{K}\delta_{T_k}$ $\qquad (\pi \approx \mu)$

Median and mean of π are computed with the above algorithms.

Statistical model - continued

We'll therefore give point estimates of posterior distribution μ:

- median:
$$\arg\min_{S\in\mathcal{T}_n} \int_{\mathcal{T}_n} d(S,T)\ \mathrm{d}\mu(T)$$
- mean:
$$\arg\min_{S\in\mathcal{T}_n} \int_{\mathcal{T}_n} d(S,T)^2\ \mathrm{d}\mu(T)$$

Markov chain Monte Carlo (MCMC) methods yield samples of posterior distribution:

$\rightsquigarrow T_1,\ldots,T_K \in \mathcal{T}_n \qquad \rightsquigarrow \pi := \frac{1}{K}\sum_{k=1}^{K} \delta_{T_k} \qquad (\pi \approx \mu)$

Median and mean of π are computed with the above algorithms.

Statistical model - continued

We'll therefore give point estimates of posterior distribution μ:

- median:
$$\operatorname*{arg\,min}_{S\in\mathcal{T}_n} \int_{\mathcal{T}_n} d(S,T)\ \mathrm{d}\mu(T)$$
- mean:
$$\operatorname*{arg\,min}_{S\in\mathcal{T}_n} \int_{\mathcal{T}_n} d(S,T)^2\ \mathrm{d}\mu(T)$$

Markov chain Monte Carlo (MCMC) methods yield samples of posterior distribution:

$\rightsquigarrow T_1,\ldots,T_K \in \mathcal{T}_n \qquad \rightsquigarrow \pi := \frac{1}{K}\sum_{k=1}^{K} \delta_{T_k} \qquad (\pi \approx \mu)$

Median and mean of π are computed with the above algorithms.

Real data experiments (see Philipp Benner's poster)

We used ribosomal subunit rRNA sequence alignment:

- number of species: 12
- number of trees: $20,000$
- number of iterations: 10^7

Conclusion:

- computations took less than 5 minutes
- very good speed of convergence (no theory though)
- random-order versions seem to be better

More computational studies certainly **needed** in the future!

Real data experiments (see Philipp Benner's poster)

We used ribosomal subunit rRNA sequence alignment:

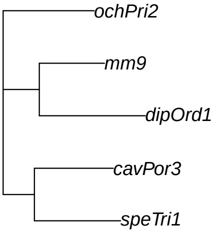
- number of species: 12
- number of trees: $20,000$
- number of iterations: 10^7

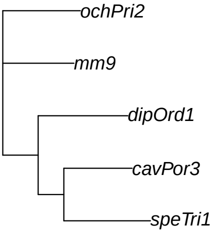
Conclusion:

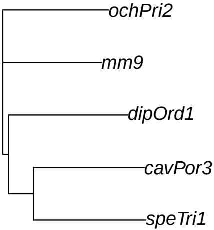
- computations took less than 5 minutes
- very good speed of convergence (no theory though)
- random-order versions seem to be better

More computational studies certainly **needed** in the future!

Real data experiments - continued







Summary:

- The BHV Tree space has nice geometrical properties.
- ... it is rather "big", but that doesn't seem to be an issue.
- The median and mean are well-defined and behave nicely.
- One can compute distances in polynomial time.
- There are rigorous approximation algorithms for medians and means.
- We used all that in phylogenetic inference and **would like to hear your opinion!**

References

- **M. Bacak:** *Computing medians and means in Hadamard spaces.* Preprint, arXiv:1210.2145.
- **P. Benner, M. Bacak:** *Computing the posterior expectation of phylogenetic trees.* Preprint, arXiv:1305.3692.
- **L. Billera, S. Holmes, K. Vogtmann:** *Geometry of the space of phylogenetic trees.* Adv. in Appl. Math., 2001.
- **E. Miller, M. Owen, S. Provan:** *Averaging metric phylogenetic trees.* Preprint, arXiv:1211.7046v1.
- **M. Owen, S. Provan:** *A fast algorithm for computing geodesic distances in tree space.* IEEE/ACM Trans. Computational Biology and Bioinformatics, 2011.

Invitation to Leipzig

International Symposium on Discrete Mathematics and Mathematical Biology

August 26–27

Summary School on Phylogenetic Combinatorics

August 28–30

Max Planck Institute for Mathematics in the Sciences

Leipzig

www.mis.mpg.de