Lecture 7 : Uniqueness of tree representation

MATH285K - Spring 2010 *Lecturer: Sebastien Roch*

References: [SS03, Chapter 7]

Previous class

Recall:

THM 7.1 (Quartet theorem) *Let $\mathcal{T}, \mathcal{T}'$ be X-trees. Then, $\mathcal{T} \cong \mathcal{T}'$ if and only if $\mathcal{T}|S \cong \mathcal{T}'|S$ for all $S \subseteq X$ with $|S| \leq 4$.*

1 Tree metrics

DEF 7.2 (Path metric) *Let $T = (V, E)$ be a tree with edge weights $\{w_e\}_{e \in E}$. For two vertices $u, v \in V$, we let $\mathrm{Path}(u, v)$ be the set of edges on the unique path between u and v. The* path metric *corresponding to (T, w) is then defined as*

$$d_{T,w}(u, v) = \sum_{e \in \mathrm{Path}(u,v)} w_e.$$

If $\mathcal{T} = (T, \phi)$ is an X-tree, then for $x, y \in X$ we define

$$d_{\mathcal{T},w}(x, y) = d_{T,w}(\phi(x), \phi(y)).$$

DEF 7.3 (Dissimilarity map) *A* dissimilarity map *on X is a function $\delta : X \times X \to \mathbb{R}$ such that $\delta(x, x) = 0$ and $\delta(x, y) = \delta(y, x)$ for all $x, y \in X$.*

DEF 7.4 (Tree metric) *A dissimilarity map δ is a* tree metric *if there exists an X-tree $\mathcal{T} = (T, \phi)$ and edge weight function $w : E \to \mathbb{R}_{++}$ such that for all $x, y \in X$*

$$\delta = d_{\mathcal{T},w}(x, y).$$

We say that $(\mathcal{T}, w)$ is a tree metric representation *of δ.*

2 Uniqueness of tree metric representation

THM 7.5 (Uniqueness of tree metric representation) *Let δ be a tree metric on X. Up to isomorphism, there is an unique tree metric representation of δ.*

Proof: *Tree structure.* Let $(\mathcal{T}, w)$ be a tree metric representation of δ with $\mathcal{T} = (T, \phi)$ and $T = (V, E)$. By the Quartet Theorem, it suffices to check that for all $X' \subseteq X$ of size four, δ determines $\mathcal{T}|X'$. Note that:

- the expression
$$\frac{1}{2}(\delta(x,y) + \delta(x,z) - \delta(y,z)) \tag{1}$$
is the (positive) weight of the leaf edge at x if such an edge is present in $\mathcal{T}|X'$, that is, if $\{x\}|X' - \{x\} \in \Sigma(\mathcal{T}|X')$ (and 0 otherwise).

- the expression
$$\frac{1}{2}(\delta(x,w) + \delta(y,z) - \delta(x,y) - \delta(w,z)) \tag{2}$$
is:
  - the (positive) weight of the interior edge of $\mathcal{T}|X'$ if $\{x,y\}|\{w,z\} \in \Sigma(\mathcal{T}|X')$,
  - minus the weight of the interior edge of $\mathcal{T}|X'$ if $\{x,w\}|\{y,z\} \in \Sigma(\mathcal{T}|X')$,
  - 0 otherwise.

Since the positivity of the expressions in (1) and (2) is independent of the tree metric representation chosen, any such representation must have the same set of splits.

Edge weights. The proof that δ also determines the edge weights of its tree metric representation follows similarly. Let $(\mathcal{T}, w)$ be as above. Let $e = \{u, v\} \in E$. There are two cases:

- e *is pendant.* W.l.o.g. u is a leaf. Let $x \in X$ such that $\phi(x) = u$. Choose two elements y, z of X such that the path in $T \backslash u$ between $\phi(y)$ and $\phi(z)$ passes through v. Apply formula (1) to obtain w_e.

- e *is interior.* Choose two elements x, y (resp. w, z) of X such that the path in $T \backslash u$ (resp. $T \backslash v$) between $\phi(x)$ and $\phi(y)$ (resp. $\phi(w)$ and $\phi(z)$) passes through v (resp. u). Apply formula (2) to obtain w_e.

■

3 Four-point condition

The proof of Theorem 7.5 indicates the central role played by the expression (2).

DEF 7.6 (Four-point condition) *A dissimilarity map δ satisifies that* four-point condition (4PC) *if for all $x, y, w, z \in X$ (not necessarily distinct)*

$$\delta(w,x) + \delta(y,z) \leq \max\{\delta(w,y) + \delta(x,z), \delta(w,z) + \delta(x,y)\}. \tag{3}$$

Note that if δ satisfies the 4PC then one of the sums in (3) must be less or equal than the other which must be equal. (Argue by contradiction that one of the sums is strictly larger than the other two.) Also, δ satisfies the triangle inequality since taking $y = z$ in (3) gives

$$\delta(w,x) \leq \delta(w,y) + \delta(x,y),$$

and δ is non-negative since taking $w = x$ and $y = z$ gives

$$0 \leq 2\delta(w,y).$$

The 4PC leads to the following characterization of tree metrics.

THM 7.7 (Tree-Metric Theorem) *Let δ be a nonnegative dissimilarity map. Then, δ is a tree metric if and only if δ saitisfies the 4PC.*

One direction follows directly from the proof of Theorem 7.5. We will prove the other direction in the next lecture. We first give a different characterization of the 4PC in terms of ultrametrics which will be useful in the proof of the Tree-Metric Theorem.

4 Ultrametrics

DEF 7.8 (Ultrametric) *A dissimilarity map δ on X is an* ultrametric *if for every three distinct $x, y, z \in X$,*

$$\delta(x,y) \leq \max\{\delta(x,z), \delta(y,z)\}. \tag{4}$$

Clearly, a non-negative ultrametric satisfies the triangle inequality. Also, a case analysis shows that a non-negative ultrametric satisfies the 4PC. In fact, we will show in the next lecture that ultrametrics are very special cases of tree metrics—corresponding essentially to rooted trees.

DEF 7.9 (Gromov Product) *Let δ be a dissimilarity map on X and fix $r \in X$. The* Gromov product *of x and y in $X \setminus \{r\}$ is*

$$\delta_r(x,y) = \begin{cases} \frac{1}{2}(\delta(x,y) - \delta(r,x) - \delta(r,y)) & x \neq y \\ 0 & \textit{o.w.} \end{cases} \tag{5}$$

In the case of tree metrics, the Gromov product is minus the length of the shared subpath between the path connecting r and x and the path connecting r and y.

LEM 7.10 *A nonnegative dissmilarity map δ satisfies the 4PC if and only if δ_r is an ultrametric for all $r \in X$.*

Proof: To prove (3), take $r = w$, write down the ultrametric condition for $\delta_r(y,z)$ and add $\delta(r,x)+\delta(r,y)+\delta(r,z)$ to both sides. Reversing the previous computation gives the other direction. ■

Further reading

The definitions and results discussed here were taken from Chapter 7 of [SS03]. Much more on the subject can be found in that excellent monograph. See also [SS03] for the relevant bibliographic references.

References

[SS03] Charles Semple and Mike Steel. *Phylogenetics*, volume 24 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2003.