Notes 7: Tree-metric theorem

MATH 833 - Fall 2012

Lecturer: Sebastien Roch

References: [SS03, Chapter 7]

1 Equidistant Representation

We begin with a result about ultrametrics.

DEF 7.1 (Equidistant Representation) Let $\delta$ be a dissimilarity map on $X$. An equidistant representation of $\delta$ is a rooted phylogenetic tree $T = (T, \phi)$ with $T = (V, E)$ and root $\rho$, and an edge weight function $w : E \to \mathbb{R}$ such that:

1. For all $x, y \in X$
   $$d_{T,w}(\rho, \phi(x)) = d_{T,w}(\rho, \phi(y)).$$
   (By definition of a path metric, the equality then holds with $\rho$ replaced with $u \in V$ as long as $u \leq_T \phi(x), \phi(y)$, that is, $u$ is a common ancestor of $\phi(x)$ and $\phi(y)$.)

2. If $u \leq_T v \leq_T \phi(x)$ for $u, v \in \hat{V}$ and $x \in X$ then
   $$d_{T,w}(\phi(x), v) \leq d_{T,w}(\phi(x), u).$$
   (In particular, all interior edge weights are non-negative.)

It is straightforward to check that a dissimilarity map admitting an equidistant representation is an ultrametric. There is also a converse:

THM 7.2 If $\delta$ is an ultrametric on $X$, then it has an equidistant representation.

Proof: The proof is based on a simple reconstruction algorithm.

DEF 7.3 (Cherry) A cherry is a pair of leaves $(u, v)$ with a common neighbour.

Let $\delta$ be an ultrametric on $X$. Consider the following recursive procedure:

function EQUIDISTANT

Input: Dissimilarity map $\delta$ on $X$
Output: Equidistant representation \((T, w)\) of \(\delta\)
- If \(|X| = 2\), return a cherry with edge weights \(\frac{1}{2} \delta(a, b)\).
- Otherwise:
  * Find \(a, b \in X\) minimizing \(\delta(a, b)\).
  * Set \(\delta^{(ab)}\) to be \(\delta\) restricted to \(X \setminus \{b\}\).
  * Compute \((T^{(ab)}, w^{(ab)}) = \text{EQUIDISTANT}(\delta^{(ab)})\).
  * Let \(l^{(ab)} = \phi^{(ab)}(a)\). Let \(T\) be \(T^{(ab)}\) where \(l^{(ab)}\) is replaced with a new cherry \((l_a, l_b)\) with \(l_a = \phi(a)\) and \(l_b = \phi(b)\) and edge weights \(\frac{1}{2} \delta(a, b)\). Let \(e^{(ab)}\) be the edge adjacent to \(\phi^{(ab)}(a)\) in \(T^{(ab)}\). Let \(\hat{e}\) be interior edge of \(T\) adjacent to the common neighbour of \(l_a\) and \(l_b\). Set \(w_{\hat{e}} = w_{e^{(ab)}} - \frac{1}{2} \delta(a, b)\).
  * Return \((T, w)\).

The correctness of this procedure follows by induction on \(|X| \geq 2\). The case \(|X| = 2\) is trivial. Assume the reconstruction is correct for \(|X| - 1\). The choice of \(a, b\) above guarantees that

\[
\delta(a, b) \leq \delta(a, x) = \delta(b, x)
\]

for all \(x \in X \setminus \{a, b\}\) and \(w_{\hat{e}} \geq 0\). \(\blacksquare\)

2 Proof of Tree-Metric Theorem

Proof: (of Theorem ??) Choose \(r \in X\). Since \(\delta\) satisfies the 4PC, \(\delta_r\) is an ultrametric on \(X' = X \setminus \{r\}\) and there exists an equidistant representation \((T', w')\) of \(\delta_r\) with \(T' = (T', \phi')\) and root \(\rho'\). Define

\[
p = -d_{T', w'}(\rho', \phi'(x)),
\]

which is independent of \(x \in X'\).

To obtain a tree metric representation \((T, w)\) of \(\delta\), we add a leaf edge \(e_r\) to \(\rho'\) with a new leaf \(r\). Guided by the formula

\[
\delta(x, y) = \delta(r, x) + \delta(r, y) + 2\delta_r(x, y),
\]

for \(r \notin \{x, y\}\), we set

\[
w_e = \begin{cases} 
2w'_e & \text{if } e \in E(T') \\
2w'_e + \delta(r, x) & \text{if } e = \{\phi'(x), u\} \text{ for some } u \in V(T') \text{ and } x \in X' \\
2p & \text{if } e = e_r.
\end{cases}
\]
The choice of \( w_e \) in (1) is justified by

\[
d_{T,w}(x,r) = 2p + 2d_{T',w'}(\phi'(x),\rho') + \delta(r,x) = \delta(r,x).
\]

To see that the weights in (1) are non-negative, note that, since \( \delta \) satisfies the 4PC, it satisfies the triangle inequality so

\[
\delta_r(x,y) = \frac{1}{2} (\delta(x,y) - \delta(r,x) - \delta(r,y)) \leq 0.
\]

Hence, taking \( x, y \in X' \) such that \( \phi'(x) \) and \( \phi'(y) \) go through \( \rho' \) in \( T' \), we have

\[
2p = -\delta_r(x,y) \geq 0.
\]

Similarly, since \( \delta \) satisfies the triangle inequality, leaf edges in any tree metric representation of \( \delta \) must have non-negative weight (see the proof of the Uniqueness of Tree Representation Theorem). Finally, by definition of an equidistant representation, \( w'_e \geq 0 \) for \( e \in \tilde{E}(T') \).

Contracting zero-weight edges, we obtain a tree metric representation with positive edge weights.

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