

Applications of Tropical Geometry

Karthik Balakrishnan

July 11, 2017

1 Phylogenetics

One important tool in computational biology is creating phylogenetic tree from distance data involving n leaves. The leaves may be genes or organism and biologists use DNA sequences to measure the evolutionary distance between the leaves. Using the Hamming distance, which is the proportion of different between two strings of DNA, biologists create a measure of evolutionary distance between two leaves. Below is a possible phylogenetic table.

	<i>H</i>	<i>M</i>	<i>R</i>	<i>C</i>
<i>H</i>	0	1.1	1.0	1.4
<i>M</i>	1.1	0	0.3	1.3
<i>R</i>	1.0	0.3	0	1.2
<i>C</i>	1.4	1.3	1.2	0

More generally an $n \times n$ -symmetric matrix of distances is created to represent genetic distances.

$$D = \begin{pmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{12} & 0 & d_{23} & \cdots & d_{2n} \\ d_{13} & d_{23} & 0 & \cdots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{1n} & d_{2n} & d_{3n} & \cdots & 0 \end{pmatrix}$$

D is a metric if and only if, which means that triangle inequality holds true between terms, $d_{(ik)}$, $d_{(ij)}$, and $d_{(jk)}$ for all i , j , and k . From this we can extrapolate that D is only a metric if and only if $D \odot D = D$.

Sources

<https://math.berkeley.edu/~bernd/mathmag.pdf>