**SAMMON EMBEDDING**

This is an embedding algorithm that preserves the distances between all points. How is it achieved?

Assume there are high dimensional data described by d-dimensional vectors, X_i where i=1, 2, \ldots, N. And they will be mapped into vectors Y_i, with dimensions 2 or 3. Denote the distances to be d_{ij}^{*} = \sqrt{| X_i - X_j|^2} and d_{ij} = \sqrt{| Y_i - Y_j|^2}. In this problem, Y_i are the variables to be learned. The cost function to minimize is

E = \frac{1}{c} \sum_{i<j} \frac{(d_{ij}^{*} - d_{ij})^2}{d_{ij}^{*}},

where c = \sum_{i<j} d_{ij}^{*}. To minimize this, use Newton's method by

Y_{pq} (m+1) = Y_{pq} (m) - \alpha \Delta_{pq} (m),

where , and \alpha is the learning rate.

**LINKS:**

[1]. https://datawarrior.wordpress.com/2016/10/23/sammon-embedding/

[2].