Some additions: **Multidimensional Scaling**

(F. Chiaromonte)

Given a matrix $D = \{d_{ij}\}$ containing the distances (dissimilarities) between each pair of n objects in a set, and a chosen number of dimensions, k, MDS places each object into a k-dimensional Euclidian space as to retain as much as possible the distances between objects. If $k = 1, 2$ or 3, the resulting data cloud can be visualized.

Stress function $$S_k(\boldsymbol{X}) = \sqrt{\frac{\sum_{i<j}(d(x_i;x_j) - d_{ij})^2}{\sum_{i<j} d(x_i;x_j)^2}} \longrightarrow \boldsymbol{X}_k^{\ *} = argmin\, S_k(\boldsymbol{X})$$

The minimization can be performed numerically under a variety of specifications of the problem.

In the simplest, where D is a true Euclidian distance matrix, this reduces to an Eigen decomposition problem like in PCA. Form the scores in the Eigen-space spanned by the eigenvectors of the k largest eigenvalues of

$$A = \left\{-\frac{1}{2}\, d^2{}_{ij}\right\}$$

Centered Matrix

$$B = \{a_{ij} - a_{i.} - a_{.j} + a_{..}\}$$

Can pick a reasonable k using the eigenvalues.

Notes:

- The stress can be brought to 0 for k $\geq$ n-1.
- Non-uniqueness, e.g., any translation or rotation of the solution produces an equally good solution.