MULTIDIMENSIONAL SCALING

Euclidean Embedding

A dissimilarity matrix is a square matrix D with entries d_{ij} satisfying

- $d_{ij} \geq 0$
- $d_{ij} = d_{ji}$
- $d_{ii} = 0$

Conceptually, an $n \times n$ dissimilarity matrix represents the pairwise dissimilarities between n patterns which may or may not be Euclidean vectors.

MDS refers to a family of algorithms for solving the Euclidean embedding problem:

Given an $n \times n$ dissimilarity matrix $D = [d_{ij}]$, find a dimension p and $x_1, \dots, x_n \in \mathbb{R}^p$

such that $d_{ij} = \|x_i - x_j\|$.

Applications of Euclidean embedding include:

1. Dimensionality reduction
2. Visualization
3. Extend algorithms to non-Euclidean data

Euclidean embeddings don't always exist (dissimilarities need not satisfy the triangle inequality), or the minimum p might be larger than desired. Therefore, approximate solutions are also of interest.

$\boxed{\text{Euclidean Distance Matrices}}$

An $n \times n$ matrix D is called a Euclidean distance matrix if there exist p and $x_1, \dots, x_n \in \mathbb{R}^p$ s.t.

$$d_{ij} = \|x_i - x_j\| \quad \forall i,j.$$

Theorem Let D be an $n \times n$ dissimilarity matrix. Set $B = HAH$ where

$$A = [a_{ij}], \qquad a_{ij} = -\tfrac{1}{2} d_{ij}^2$$

$$H = I - \frac{1}{n} \underline{1}\, \underline{1}^T$$

Then D is a Euclidean distance matrix iff B is positive semi-definite. If B is PSD with positive eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n$ and corresponding eigenvectors

$$u_1 = \begin{bmatrix} u_{11} \\ \vdots \\ u_{1n} \end{bmatrix}, \quad \ldots, \quad u_p = \begin{bmatrix} u_{p1} \\ \vdots \\ u_{pn} \end{bmatrix}$$

normalized such that

$$u_k^T u_k = \lambda_k,$$

then the vectors

$$x_i = (u_{1i}, \ldots, u_{pi})$$

satisfy $d_{ij} = \| x_i - x_j \|$. In addition $\bar{x} = \frac{1}{n} \sum x_i = 0$.

Proof: See Mardia, Kent, and Bibby, Multivariate Analysis, 1979

Classical MDS

Even if a Euclidean distance matrix does not exist, the previous result suggests an approximate algorithm known as classical MDS.

Input: D, desired $p \le n$

1. Form B as in the theorem
2. Compute the eigenvalue decomposition

$$B = V \Lambda V^T$$

where $V = [v_1 \cdots v_n]$, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$

3. Set $u_k = \sqrt{\lambda_k}\, v_k$, $U = [u_1 \cdots u_p]$

Return: $x_i = i^{th}$ row of U

Note that the algorithm can only be applied if $\lambda_p > 0$.

If the dissimilarities are themselves Euclidean distances between points in $\mathbb{R}^q$, $q > p$, then classical MDS is equivalent to applying PCA to the q dimensional points. See Mardia, Kent, and Bibby for details.

Stress Criterion

A second approach to MDS is to minimize the stress objective function

$$\sum_{i,j=1}^{n} w_{ij} \left(d_{ij} - \| x_i - x_j \| \right)^2$$

wrt $x_1, \dots, x_n \in \mathbb{R}^p$. Examples for w_{ij} are $w_{ij} = 1$ or $w_{ij} = d_{ij}^{-\alpha}$, $\alpha > 0$. This problem is nonconvex, but a local minimizer can be obtained efficiently with a majorize-minimize algorithm.

Final Thoughts

The two MDS algorithms are called metric methods because they try to preserve the dissimilarities. There are also nonmetric methods that just try to preserve the rank ordering of the interpoint distances.