VOS: A New Method for Visualizing Similarities between Objects

Nees Jan van Eck    Ludo Waltman

Econometric Institute, Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
Email: {nvaneck,lwaltman}@few.eur.nl

50 Years of Econometrics
June 10, 2006
A new method for visualizing similarities between objects is presented.

The method is called VOS, which is an abbreviation for *visualization of similarities*.

The objective of VOS is to provide a low-dimensional visualization in which objects are located in such a way that the distance between any pair of objects reflects their similarity as accurately as possible.

The relationship between VOS and multidimensional scaling (MDS) is established.
Mathematical notation

- $n$: number of objects
- $m$: number of dimensions of the visualization
- $S = (s_{ij})$: $n \times n$ similarity matrix satisfying $s_{ij} \geq 0$, $s_{ii} = 0$, and $s_{ij} = s_{ji}$ for all $i, j \in \{1, \ldots, n\}$
- $X$: $n \times m$ matrix containing the coordinates of the objects $1, \ldots, n$
- $x_i$: $i$th row of $X$, containing the coordinates of object $i$
A VOS solution is obtained by solving the following constrained optimization problem:

Minimize

\[ E(X; S) = \sum_{i<j} s_{ij} \|x_i - x_j\|^2 \]  \hspace{1cm} (1)

subject to

\[ \sum_{i<j} \|x_i - x_j\| = 1, \]  \hspace{1cm} (2)

where \(\| \cdot \|\) denotes the Euclidean norm.
Ideal coordinates

- The ideal coordinates of object $i$ are given by

$$c_i(X, S) = \frac{\sum_j s_{ij}x_j}{\sum_j s_{ij}}.$$

(3)

- Suppose the coordinates of all objects except object $i$ are fixed. The objective function of VOS then reduces to

$$E_i(x_i; X, S) = \sum_j s_{ij} \|x_i - x_j\|^2.$$

(4)

- Minimization of (4) results in

$$x_i = c_i(X, S).$$

(5)

- Apparently, VOS has the tendency to locate objects close to their ideal coordinates.
A simple example data set

Data set taken from Mardia, Kent, and Bibby (1979)\(^1\):

- \( n = 51 \)
- \( s_{ij} = \begin{cases} 
8 & \text{if } 1 \leq |i - j| \leq 3 \\
7 & \text{if } 4 \leq |i - j| \leq 6 \\
1 & \text{if } 22 \leq |i - j| \leq 24 \\
0 & \text{otherwise} 
\end{cases} \)

---

Results of VOS and MDS

The MDS solution demonstrates the horseshoe effect (Mardia et al., 1979).

VOS seems to take into account indirect similarities via third objects.
Sammon mapping

Objective

- To locate objects in a low-dimensional space in such a way that the distance between any pair of objects reflects their dissimilarity as accurately as possible.

Mathematical notation

- \( D = (d_{ij}) \): \( n \times n \) dissimilarity matrix satisfying \( d_{ij} > 0 \) and \( d_{ij} = d_{ji} \) for all \( i, j \in \{1, \ldots, n\} \)

Optimization problem

- A Sammon mapping solution is obtained by minimizing the following objective function

\[
\sigma(X; D) = \sum_{i<j} \frac{(d_{ij} - \|x_i - x_j\|)^2}{d_{ij}}.
\]
Relationship between VOS and MDS

Conditional equivalence of VOS and Sammon mapping

Theorem
Let \( s_{ij} > 0 \) for all \( i \) and \( j \) (\( i \neq j \)), and let similarities be transformed into dissimilarities using \( d_{ij} = s_{ij}^{-1} (i \neq j) \). VOS and Sammon mapping are then equivalent in the sense that VOS solutions and Sammon mapping solutions differ only by a multiplicative constant. A proof is provided in Van Eck and Waltman (2006)\(^2\).

Conditional equivalence of VOS and weighted MDS
Sammon mapping is equivalent to weighted MDS where for each pair of objects \( i \) and \( j \) the weight equals \( d_{ij}^{-1} \). It therefore follows from the above theorem that there also exists a conditional equivalence between VOS and weighted MDS.

---

Conclusions

VOS has the following properties:

- VOS has the tendency to locate objects close to their ideal coordinates.
- VOS seems to pay more attention to indirect similarities via third objects than MDS.
- VOS is, under certain conditions, equivalent to Sammon mapping and, as a consequence, to weighted MDS.
References

Description and analysis of VOS


Application of VOS to the visualization of associations between concepts based on co-occurrence data

For a practical application of VOS and an experimental comparison between VOS and MDS, please visit the poster presentation *A Comparison of Knowledge Domain Visualization Approaches* by Nees Jan van Eck.