NUMBER 39
SEPTEMBER 24, 1979

This Week's Citation Classic

Kruskal J B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29**:1-27, 1964.
[Bell Telephone Laboratories, Murray Hill, NJ]

This paper presented the first widely-used method and computer program to implement the statistical model known as multidimensional scaling (MDS). It introduced 'monotonic regression' as a tool to perfect a method for nonmetric MDS. [The *Social Sciences Citation Index*™ (*SSCI*™) and the *Science Citation Index*® (*SCI*®) indicate that this paper has been cited over 635 times since 1964.]

Joseph B. Kruskal
Bell Telephone Laboratories
Murray Hill, NJ 07974

August 29, 1979

"It is especially gratifying that a paper in statistics should receive the recognition of being frequently cited, since workers in other subjects often find statistics forbidding. The reason for this recognition is that this paper, in concert with papers by Roger Shepard and Warren Torgerson, helped introduce something which is quite unusual in statistics: a widely usable new statistical model called multidimensional scaling.[1,2]

"In the early 50s Torgerson emphasized the value of generalizing ordinary (psycho-physical) scaling to the multidimensional case, and called attention to a forgotten method by which the procedure could be carried out numerically. In the early 60s Shepard made substantial advances and provided some strikingly interesting applications which attracted wide interest. A few years later, I provided further improvements in rationale and methodology, and made available the first computer program for MDS which was to receive wide use.

"The idea of MDS is basically quite simple. Suppose we have selected several objects belonging to a single domain, such as colors, facial expressions, brands of soap, political parties in Holland, or Slavic languages, to cite a few applications. By measuring experimentally how different each selected object is from every other, we obtain a dissimilarity for each *pair* of objects. Now we want to treat these dissimilarities as if they were spatial distances, though of course we must allow for experimental error, often quite substantial (and perhaps also for systematic distortion). Take some fixed low-dimensional space, such as a line, a plane, three-dimensional or even four-dimensional space. In the MDS model each object in the domain is represented by a point in the space. The central concept is that the experimental dissimilarities should 'agree with' the spatial distances among the points. In the simplest kind of metric MDS, 'agreement' means that each dissimilarity should be approximately equal to the corresponding spatial distance. More generally, 'agreement' means an (approximate) curvilinear relationship between dissimilarities and spatial distances. In nonmetric MDS we only assume that dissimilarities and distances should match approximately up to an arbitrary order-preserving relationship.

"MDS participates in the new emphasis on methods of data analysis which are *exploratory.* Its value is not helping to measure something accurately, nor in determining how accurate a measurement is. Instead, it helps provide insight into the relationships among the objects of the domain.

"MDS has found greatest use in psychology and such uses of psychology as marketing and candidate choice. The reason for this is that the concept of similarity (and dissimilarity) appears to be an integral part of people's thinking. It is easy to elicit reliable judgments of similarity by a variety of techniques. This approach greatly enriches the older approach to cognition and perception which relied heavily on judgments of single objects on various scales.

"Outside of psychology, MDS has found diverse applications, including the arrangement of the macromolecules which make up a ribosome, and the relationship among species based on the serum-antiserum reaction.

"One of the most important recent advances in MDS was the invention of the INDSCAL method by Carroll and Chang in 1970.[3] We now appear to be in the midst of a revived surge of interest centering on the analysis of three-way and many-way tables."

1. **Torgerson W S.** Multidimensional scaling: 1. Theory and method. *Psychometrika* **17**:401-19, 1952.
2. **Shepard R N.** Analysis of proximities: multidimensional scaling with an unknown distance function.I & II. *Psychometrika* **27**:125-39, 219-46, 1962.
3. **Carroll J D & Chang J J.** Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika* **35**:283-319, 1970.