Galileo and Multidimensional Scaling[1]

Joseph Woelfel

Carolyn Evans

University at Buffalo

State University of New York

December 2008

[1] Thanks to HyunJoo Lee for providing the impetus for this essay.

Since the early 1970's there has been some confusion about the relationship between Galileo and Multidimensional Scaling (MDS). Galileo, or The Galileo System as it is sometimes called, was initially produced by sociologists and then later by communication scientists; both research groups were attempting to model cognitive processes as motions in space. Work began at the University of Wisconsin in Madison as part of a project initiated and overseen by Archie Haller, designed to study the social psychological determinants of the educational and occupational aspirations of adolescent youth (Woelfel & Haller, 1971; Haller & Woelfel, 1972; Woelfel, 1972). The principle workers on the Galileo aspects of the project were Edward L. Fink and myself.

We had discovered that the educational and occupational aspirations of adolescent students were very well predicted by the average expectations of their significant others. The algebra that devolved from the average showed that the mean was the point of balance of forces (long known in physics, but new to us in sociology). Moreover, equations for attitude change represented a vector space. This implied that attitude changes and other cognitive processes might be modeled as motions of points in a mathematical space.

Later, at the University of Illinois, I spent considerable time trying to find the coordinates of points in space from their interpoint distances, with no success. Statisticians at Illinois' SOUPAC (Statistically Oriented Users Package) suggested a variant of factor analysis as a method of determining the coordinates of points in such a space, and the first Galileo spaces were made using this model in Gail Wisan's dissertation (Wisan, 1972). She referred to the method as *unstandardized factor analysis*, which is a good description of the procedure. Later work, such as "*Standardized versus unstandardized data matrices: which type is best for factor analysis*?" (Woelfel, et. al., 1980) reprised this terminology.

An early study designed to test the precision of the Galileo procedures and their correspondence to conventional measurement systems used a questionnaire that referred to the units of measures as *galileos*[2] in recognition of Galileo's use of comparative measurement procedures. Since the research design required respondents – Illinois sociology faculty and graduate students – to respond to the questionnaire three times, they began informally referring to the questionnaire and the technique as "galileo" – often in a phrase like, "Oh, no! Not another Galileo!"

When I moved to Michigan State University in 1973, SOUPAC was no longer available, and so a new method of computing the Galileo spatial coordinates was required. We were

---

[2] Specifically, the instructions said "Unlike physical distance, which is measured in feet or inches, psychological distance is measured in galileos." It then asked respondents to estimate how different or far apart a set of concepts were in the form "How different or far apart are (x) and (y) ? _______ gal." Since, for the N concepts, the question was repeated N(N-1)/2 times in each administration, respondents read the term "galileo" many times over the three administrations.

aware by then that a solution to the problem was known to some psychometricians, and was referred to as *multidimensional scaling*. The problem was first raised – long before we were aware of it, and in fact before I was born -- by L. L. Thurstone, who referred to it as the "box problem."

Thurstone had spent many years trying to discover the factors underlying human intelligence, and had developed a technique known as *factor analysis* for this purpose. Although perhaps no other multivariate statistical technique ever experienced more revisions and variations, Thurstone was never satisfied that it accomplished his goal. When he retired, he left "the box problem" as a challenge to future scientists: if we have a series of boxes, and we make many measurements of them, (e.g., their width, height, surface area, volume, etc.) can we discover that, underlying all of them, are the dimensions *length, width* and *height*?

In 1938, Young and Householder in a classic article showed that the measurement required to find the dimensionality underlying a set of points in space was a matrix of their interpoint distances. The principle axes of the scalar products of the dissimilarities matrix were the dimensions sought in Thurstone's box problem.

Warren Torgerson modified this approach slightly by expressing all the vectors of the scalar products matrix from the center of the set of points by a procedure he called "double centering". The principle axes of this "double-centered scalar products" matrix were the dimensions of the space in which the points were projected with origin at the center of the set of points. Torgerson referred to this procedure as "Multidimensional Scaling."

Galileo researchers implemented the Young Householder Torgerson procedure for obtaining the double-centered scalar products matrix from a matrix of dissimilarities among the points in the Galileo algorithm. Kim Blaine Serota and Richard A. Holmes implemented the algorithm in FORTRAN, using a method discovered by Karl Jacobi in 1849 and implemented in FORTRAN by Johannes Van de Geer in 1971, to calculate the eigenvalues and eigenvectors of this space. Although new to sociologists and communication researchers, the mathematical procedure of projecting distances onto principle axes had been a core procedure in physics for establishing inertial reference frames for well over a hundred years by this time.

While one of the central algorithms for multidimensional scaling became a part of the Galileo software, Galileo and multidimensional scaling parted ways at this point.

At the root of the divorce[3] were two core issues: first, Galileo theorists placed their primary trust in measurements, which they believed to mean exclusively comparison to some standard. MDS adherents, on the other hand, thought that the measurement of cognitive structures and processes was fundamentally impossible, and so whatever numbers resulted from measurements were considered untrustworthy.

[3] It should be noted that, while Galileo researchers were very active in reading the MDS literature, attending MDS conferences, and consulting with MDS experts, the corollary was not true and MDS researchers had little or no interest in or knowledge of Galileo research.

Second, paired comparison data, particularly when well measured, seldom fit into low dimensional Euclidean spaces, but rather seemed to lie in higher dimensional Riemann spaces. Galileo researchers, committed to the data above all else, took this as a finding and began modifying the Galileo software to work in a general Riemannian space (Woelfel & Barnett, 1982). Psychometricians, on the other hand, began devising mathematical algorithms that would modify observed values in such a way that the resulting modified dissimilarities would fit into low-dimensional Euclidean spaces.

The basic goal of MDS researchers was to make low (preferably two) dimensional maps. These so-called "non-metric" MDS algorithms, probably because they provided the hope that correct solutions could be found even though the original data were wrong, displayed a brief period of considerable popularity, particularly in market and advertising research, but ultimately proved less than useful and, several years later, some of the most prominent advocates of non-metric procedures conceded that the original Young Householder Torgerson techniques were generally more reliable[4].

Galileo researchers, on the other hand, moved toward increasing precision of measure, increasing study of the ways of tracking motions of points through high dimensional Riemann space, and understanding the dynamics of such motions. In fact, on the flyleaf of their 1980 Galileo book, Woelfel & Fink say, "Most of the work done by communication researchers deals with how one operates on the multidimensional coordinate system yielded by the metric scaling algorithm rather than with how the coordinate system is generated."

This work included abstract theoretical work, such as study of how simple messages combine, how the self is defined across situations, modeling cultural processes as inertial processes in the reference frame, measuring cognitive aspects of the diffusion of innovations, measuring the inertial masses of concepts, the relationship of the network of concepts in the space to the networks of neurons in brains and culture, relationships between the meanings of texts and concept locations in the Galileo space, and many more abstract areas.

On the applied side, Galileo researchers use the system to model product market share, track election campaigns, and design effective advertising and marketing strategies in many substantive areas across many disciplinary boundaries.

All through this period, Galileo researchers referred to the system in diverse ways, most often using the term "multidimensional scaling," but also sometimes using "metric

---

[4] At the time the Galileo system was being developed, arguments in sociology and communication revolved around whether mathematical methods were appropriate for the study of human cognitive processes. A large part of *The Measurement of Communication Processes: Galileo Theory and Method* (Woelfel & Fink, 1980) was devoted to arguing for the utility of mathematics in studying human cognitive processes. It wasn't until many years of experience with psychometricians that I realized there were those who were enamored of mathematics to the exclusion of measurements. That same way of thinking is now prominent in network analysis.

multidimensional scaling" to differentiate Galileo from the then more well known non-metric models. In this they were adopting the psychometric usage, which, unfortunately, conflicts with established mathematical conventions. The "metric" algorithm, so called to distinguish it from the "non-metric" systems which treat measurements as if they are ordered values rather than magnitudes, is, in fact, not metric in the mathematical sense.

In mathematics, a metric space is one in which the triangle inequalities[5] are met, and is therefore Euclidean. This seldom turns out to be the case when the Young Householder Torgerson method is employed, since, empirically, the triangle inequalities rule is frequently violated in ways that make good sense and the resulting space is usually non-Euclidean, not metric.

Referring to the Galileo procedures as metric MDS, therefore, may make sense according to psychometric terminology, but it is clearly wrong in the much larger mathematical community.

Even among psychometricians themselves, however, use of the term MDS to refer to Galileo is a cause of considerable confusion.

There is more to the meaning of words than their dictionary definition. In the 1980's, a joint conference of Galileo researchers and prominent psychometricians, including Robert Pruzek and James Ramsay (at that time president of the Psychometric Society) was held at the University at Albany. At first, the two groups had considerable difficulty understanding each other, until Ramsay explained that, to the psychometricians, "MDS" referred to a specific set of people who did specific things, but what the Galileo researchers were doing was not related to that work.

He suggested we don't use the term "multidimensional scaling", but instead refer to ourselves by the name people associate with us -- *Galileo*. That still seems prudent as identifying Galileo work with a different research thread from a different discipline only seems to cause confusion. Better to say that Galileo, while it shares some ancestral roots with multidimensional scaling, goes far beyond MDS into the analysis of cognitive and cultural processes.

---

[5] In any triangle drawn on a flat, Euclidean plane, any two sides must be at least as long as the third. If this is not true, the plane cannot be flat.

References

Haller, A., & Woelfel, J. (1972). Significant others and their expectations: concepts and instruments to measure interpersonal influence on status aspirations. *Rural Sociology, 37*(4), 591-622.

Wisan, G. (1972). *The construction of a continuous manifold for social objects: A prolegomenous study*. University of Illinois at Urbana-Champaign, Urbana, IL.

Woelfel, J. (1972). Significant others and their role relationships to students in a high school population. *Rural Sociology, 37*(1), 86-97.

Woelfel, J., & Barnett, G. (1982). "Multidimensional scaling in Riemann space," *Quality and Quantity, 16*, 461-491.

Woelfel, J., & Fink, E. (1980). *The measurement of communication processes: Galileo theory and method.* New York: Academic Press.

Woelfel, J., & Haller, A. (1971). Significant others: The self-reflexive act and the attitude formation process. *American Sociological Review 36*(1), 74-87.

Woelfel, J., Woelfel, J. C., & Woelfel, M. L. (1980). Standardized vs. unstandardized data matrices: which type is best for factor analysis? *Quality & Quantity, 14*, 1-13.

Young, G., & Householder, A. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika, 3*, 19-22.