Statistical Science
2007, Vol. 22, No. 2, 264–265
DOI: 10.1214/088342306000000367
Main article DOI: 10.1214/088342306000000420

William H. Kruskal and the Development of Coordinate-Free Methods

Morris L. Eaton

Soon after joining the University of Chicago's Statistics Department, in the fall of 1966, I became aware of Bill Kruskal's lecture notes on topics he wryly referred to as "a coordinate-free approach to linear this and that." Graduate students raved about his course and the three-inch set of lecture notes that was clearly valued property. Bill's interest in applying vector space methods had been spurred by Jimmie Savage in the mid-1950s. Indeed to quote Bill, "Once Jimmie said a few magic words, it all became plain, but it needed writing up" (see Zabell, 1994, page 293). My guess is that the "writing up" had been going on for almost a decade prior to my arrival in Chicago.

By 1966 the notes were quite polished and consisted of nine or so chapters of linear statistical model theory that were used in a two-quarter course at Chicago. The care with which the notes were prepared, very characteristic of Bill, certainly suggested an intent to publish a book based on them. He had joined the Chicago faculty in 1950 and had finished his Ph.D. in 1955. It thus seems reasonable to conjecture that serious work on the notes was begun in the late 1950s (an early research paper in the area is Kruskal, 1961). In discussions with him in 1967, I sensed a flagging interest in the coordinate-free area, perhaps because of his many other interests—historical topics, measures of association, university administration and governmental statistics. In retrospect, his description of himself as an "overperfectionist" may provide some insight into the lack of a book based on his notes. The statistics community is certainly poorer because of this.

My interest in the coordinate-free approach to linear statistical problems was motivated by at least two things: first, a predilection for elegant mathematics applied to statistics and second, the hint that such an approach could beneficially be brought to bear on multivariate analysis. It was with some trepidation that I approached Bill in early 1967 with a request to teach "his course" in the 1967–1968 academic year. He was thrilled that a young colleague had taken an interest in the coordinate-free methods and most likely, although unspoken, was pleased to get a break from the course. After teaching the course for two years running, I had a fairly firm grip on the material.

Before describing the influence of the coordinate-free approach on other research areas, let me briefly summarize the notes and Bill's approach. The initial chapter was an extended review of finite-dimensional inner product spaces. Paul Halmos's marvelous treatment of finite-dimensional vector theory is the obvious origin of this chapter, with full attribution of course. [A preliminary edition of the Halmos text, *Finite-Dimensional Vector Spaces*, was first copyrighted in 1942 by Princeton University Press. The text was published by Van Nostrand in 1958 and is currently in the Undergraduate Texts in Mathematics series published by Springer-Verlag (Halmos, 1974). Halmos was also at the University of Chicago during part of the 1950s and, with Jimmie Savage, had published the famous Halmos–Savage theorem in 1949 (Halmos and Savage, 1949).] After the vector space review followed the introduction of random vectors, mean vectors, covariance operators and the normal distribution. Linear model material including the Gauss–Markov theorem, hypothesis testing, confidence intervals and analysis of variance (ANOVA) examples rounded out most of the remaining chapters.

An early description of the Gauss–Markov theorem in a coordinate-free setting occurs in Kruskal (1961). This paper contains, in rather condensed form, a variety of the topics covered in the 1966 version of the lecture notes. Even in hindsight, it is difficult for me to assess the enormous influence Bill's notes had on both my mathematical skills and my research development. However, the direct effect of Kruskal (1968), a marvelous paper, is relatively easy to describe. In coordinate-free language, here is a statement of the main result of that paper:

> The Gauss–Markov and least squares estimators are the same if and only if the linear manifold of the mean vector is an invariant subspace of the covariance.

Morris L. Eaton is Professor Emeritus, School of Statistics, University of Minnesota, Minneapolis, Minnesota 55455, USA (e-mail: eaton@stat.umn.edu).