DISCUSSION III ON FITTING EQUATIONS TO SENSORY DATA

ARMAND V. CARDELLO

U.S. Army Natick RD&E Center
Science and Technology Directorate
Natick, MA 01760-5020

LARRY L. LESHER

Geo-Centers, Inc.
Natick, MA 01760

"Science presents itself as a two-faced bipartite, endeavor looking at once toward the formal, analytic, schematic features of model-building, and toward the concrete, empirical, experiential observations by which we test the usefulness of a particular representation. Schematics and empirics are both essential to science, and full understanding demands that we know which is which".

S.S. Stevens (1968, pg. 865)

The above statement captures the essence of an important and recurrent theme in science, a theme that the paper by Moskowitz now raises within the context of the modeling of sensory data. When the statement was offered, the field of psychophysics was in the throws of a controversy between measurement theorists and statisticians over the "allowable" statistics to be applied to sensory data. "Theorists" argued that the schematic model, i.e. the mathematical theory of measurement, required that statistical procedures on data be limited to those mathematical operations consistent with the scale properties of the data. "Statisticians", taking the more empirical position, argued that data are simply numbers, not beholden to theory, and that all statistical operations should be permissible. This fundamental dichotomy, between schema and empirics, between theory and practical data-driven application, is again laid bare in the paper by Moskowitz.

At the practical/empirical level, there are several elements of Moskowitz's paper that warrant consideration. First is the simple fact that only one of the two data sets provides convincing evidence of a "paradox". The differences among the equations for predicting the vegetable soup hold-out samples (Table 1) are marginal at best. Second is Moskowitz's approach to model validation using holdout samples derived from cluster analysis. His approach uses x samples out of the total n samples to build the model. Thus, x/n samples are used to create a model to predict the remaining 1 - (x/n) of the samples (holdouts) for purposes of validation. A better approach, using available software and a simple looping routine, is as follows: Let n be the total number of samples available. For i = 1 to n, select the ith sample as your holdout sample. Use the remaining n-1 samples to build your model and predict the ith sample value from this model. Repeat this process until all n models have been built. Calculate correlations as in Table 2 and plot actual and predicted values as in Fig. 1. This jackknife approach has several advantages. It uses (n-1)/n of the samples to build each model. It does not require a cluster analysis as the basis for selecting holdout samples. Each sample is used in (n-1)/n of the models, and each is used as a holdout only once. Another benefit of this approach is that the researcher can readily identify a particular sample that has a unique or hard to predict attribute profile.

A third issue concerns the outliers in the data comparing the actual versus the predicted values of the holdout samples (Fig. 1). The differences among the correlations in Table 2 can well be attributed to these outliers. It would be interesting to know why some of these correlations (image and liking) are so much worse for the quadratic model than for the linear model. Do the holdout samples corresponding to these outliers fall in the same sample space as those that were used for model building? Could this be a function of the specific samples that were chosen for model building? Would this same pattern occur if a different set of samples were chosen for model building? And would this same pattern occur and would the same conclusions be drawn if the model validation procedure described in the previous paragraph was used?

While the above points are important and interesting, there is a still more fundamental issue that Moskowitz's paper raises. That issue is the one mentioned at the outset and concerns the relative merits of theory-driven model building versus empirical or data-driven model building.

In theory-driven model building, the researcher is motivated by the nomothetic pursuit, i.e. the search for universal relationships that can be used to describe/predict all phenomena within a general class; albeit, with some occasional loss of accuracy in predicting perturbations in a small subset of cases. This approach views the goal of science as making sense out of the world, not simply describing it; and theory is seen as the basis of any scientific attempt at understanding. Occasional "paradoxes" and nonconformance of the data to the predictive model are to be expected, especially when working in disciplines where the data are generally noisy, as is the case with sensory data. In empirical model building, on the other hand, the approach is more idiographic. Here, accuracy of prediction for a limited and specific set of data is paramount. This is more often the situation encountered in industry, where the "bottom line" is directly related to the ability to predict success for a specific set of products.

If one looks at the evolution of any scientific discipline, the nature of model building can be seen to shift from data-driven modeling or "curve-fitting" approaches to more theory-driven modeling. In his paper, Moskowitz sites the example of the shift that occurred when sensory scientists began to accept the "power law" as the "true" underlying functional relationship between sensory magnitude and physical intensity. That shift did not occur overnight, but rather, required the accumulation of a large base of empirical data, before a consensus emerged among scientists for this new "schema". In several critical places in his paper, Moskowitz reminds the reader that theory and substantial data tell us that liking is nonlinearly related to sensory and instrumental variables (at least for those perceptual dimensions commonly associated with food stimuli). He is, of course, referring to the proposition first put forth in 1879 and commonly known as "Wundt's law". For Moskowitz, who was the author of much of the contemporary research on the relationship between sensory attribute ratings and liking, the schema of a nonlinear relationship is well supported by the empirical data, and it is on this basis that he argues for the use of the quadratic model for relating sensory and liking data. While he and many other sensory researchers have assimilated the data showing a nonlinear relationship into their schematic view of how hedonics operate, others have not. For the latter researchers, the empirical data on this relationship and, especially, its generality, are not sufficiently convincing.

The issue of the generality of the sensory/liking relationship brings us to another interesting aspect of Moskowitz's paper, i.e. the modeling of "image" attributes. Here, as Moskowitz points out, there is no theory to guide the way. His personal solution to this dilemma is to use the quadratic model. However, unlike his focus on the "theoretical" argument for applying the quadratic model to sensory/liking data, his rationale here is a more practical one, i.e. "a quadratic model will not distort a curve into a plane". This practical argument makes sense, since one could readily conceive of some "image" variables being linearly related to sensory attributes, while others are nonlinearly related. So, what is to be lost by following Moskowitz's suggestion to fit the quadratic model to these data? The answer, of course, is parsimony, i.e. why risk the attribution of a complex model to the data when a simpler one would suffice? Indeed, parsimony is often the reason why linear models are still fit to sensory attribute/liking data, in spite of the considerable data showing a nonlinear relationship. While parsimony in scientific explanation is laudable, it is not always conducive to the advancement of theory in a field. In many cases it runs the risk of being shortsighted, especially when applied to the description/prediction of a subset of cases for which a more general and complex model will explain all cases within the class.

Like any good scientific paper, Moskowitz's "paradoxical" data and the statistical approaches that he describes serve as fodder for further thought and analysis. For statistically-oriented readers, the paper encourages consideration of alternative approaches for testing a model's validity and effectiveness. For practicing sensory scientists, the paper inspires consideration of the often disparate goals involved in modeling sensory data. And for those who consider such philosophical issues important for understanding progress (or lack thereof) in a field, it is a reminder that the relationship between schema and empirics in sensory science is just as challenging today, as it was 30 years ago.

REFERENCE

STEVENS, S.S. 1968. Measurement, statistics, and the schemapiric view. Science *161*, 849–856.