Review of: "On Optimal Linear Prediction"

David C. Schneider[1]

1 Memorial University of Newfoundland, St. John's, Canada

Potential competing interests: No potential competing interests to declare.

On Optimal Linear Prediction by Inge Svein Helland.

https://doi.org/10.32388/I2OXKX   Preprint 1. 22 October 2024

Review by David Schneider,

Department of Ocean Science, Memorial University, St. John's, Canada

The stated purpose of the article is

"to prove that, under certain assumptions in a linear prediction setting, optimal methods based upon model reduction and even an optimal predictor can be provided."

This is not clear. I take it to mean that under certain assumptions, an optimal predictor can be provided in a linear model setting, based on the reduction of multiple explanatory variables by partial least squares (PLS) regression.

The use of partial least squares for the reduction of multiple variables was unfamiliar to me, so I looked it up. PLS regression was developed by Herman Wold in the 1970s by extending multiple linear regression to a latent variable approach. More specifically, the covariance structures in two spaces (i.e., the X and Y spaces) are modeled so that variables X and variables Y are projected to a new space. This is called projection to latent (hidden) structure, a concept familiar to psychologists and social scientists.

In trying to find a clear statement of the main result, I looked at the conclusion section at the end of the article. It contained opinions, with no statement of the main conclusion of the article. The main result is stated in Theorem 12, where the premise is in English, with a conclusion in symbolic form. The conclusion needs to be stated in English.

The method of partial least squares (PLS) regression, as developed by a chemometrician, is stated as the emerging and recommended method by some statisticians. The statement needs to be supported by references that recommend usage.

PLS regression is further stated as having grown popular among very many applied researchers. The statement is unsubstantiated. The use of PLS in psychology and the social sciences, where latent variable analysis is familiar, needs to be established by citation. Where latent variable analysis is less familiar, as in many of the natural sciences, citations to the use of PLS are needed, beyond its proposed use by a chemometrician.

In section 2, Quantum Foundations, the article states *“the fundamental notion in my approach towards quantum foundations is that of a theoretical variable connected in a given context to an observer or to a communicating group of observers”*

This strikes me as inconsistent with Heisenberg’s statement of the physical content of quantum mechanics (Heisenberg, W. (1927) *Zeitschrift für Physik*, 43, 172-198. English translation as: The physical content of quantum kinematics and mechanics. Pp 62-84 in

J.A. Wheeler, W.H. Zurek, Eds. *Quantum Theory and Measurement,* Princeton University Press, 1983.). The use of “quantum foundation” needs to be clarified relative to the usage of “quantum measurement” as established nearly a century ago by Heisenberg.

Communicating groups of observers, as in the statement above, are not to be found in Heisenberg. So the notion that communicating groups are an analog of measurement at quantum scales needs to be established with reference to Heisenberg 1927. If this has been done elsewhere, it needs to be summarized briefly, with an appropriate citation.

The notion of communicating groups is the basis for section 3.

Section 3. An addition to statistical inference theory; two statisticians

“For the purpose of this article, let B be a very experienced statistician, and let him be a Bayesian with a very open mind. Let A be some statistician who is inspired by some ideals; without much loss of generality, we assume that this can be modeled by B.”

In the quantum interpretation of Heisenberg, there are no communicating groups. There are physical variables for which quantum uncertainty applies only in the case of conjugate variables.

It is a mystery as to how the inequality of the statisticians—very experienced vs idealistic-- is an analog to quantum measurement. Is it an analogy by object and related object? Or is it another among the many types of analogy? The basis for the analogy needs to be stated explicitly.

“As a basis for their joint thinking, let there be a concrete statistical problem with a data set X.”

In the quantum interpretation of Heisenberg, there is no joint thinking. Turning to data, how is data set X obtained? By statistician A? By statistician B? By a third party such as a researcher? This matters. Statistician A, statistician B, and the researcher may well develop differing measurement protocols. In science, the report of a finding requires a clear statement of the measurement protocol that can be applied by another researcher. In the absence of such a statement, the study is irreproducible It is not science. Given this asymmetry, does the analogy of communicating groups with quantum measurement hold? For quantum measurement, there will be a written protocol only on the part of the observer, not the object.

“But instead of starting with only a concrete prior distribution, he assumes some symmetry in the space $\Omega\phi$ “

The article relies heavily on what is now called Bayes' theorem. There is no theorem in either of the only two publications by Thomas Bayes, both published posthumously. In Bayes (1763), there is only a rule for calculating an uncertainty interval. The theorem attributed to Bayes was first stated by Laplace (1774) as follows:

Principle

"If an event can be produced by a number n of different causes, the probabilities of these causes given the event are to each other as the probabilities of the event given the causes, and the probability of the existence of each of these is equal to the probability of the event given that cause, divided by the sum of all the probabilities of the event given each of these causes."

(Translation from French by S.M. Stigler (1986 Statistical Science 1: 364-378)

Laplace's principle is equivalent to what is now called Bayes' theorem, with all causes being *a priori* equally likely (Stigler 1986 *Statistical Science* 1: 359-378). The first proof of Laplace's principle occurs in Keynes' 1921 *Treatise on Probability* (Schneider 2021 *Cambridge Journal of Economics* 45: 951–966).

Given that what is now called Bayes' theorem is due to Laplace and was only much later proven by Keynes, would it not be more appropriate for an open-minded Priorist to update their belief on the origin of the theorem? Priorist with a very open mind would be more appropriate than Bayesian with a very open mind.