

Linear Models Based on Noisy Data and the Frisch Scheme

L. Ning, T. Georgiou, A. Tannenbaum, and S. Boyd

SIAM Review, 57(2):167-197, June 2015.

- [lin_mod_frisch.pdf](#)

We address the problem of identifying linear relations among variables based on noisy measurements. This is a central question in the search for structure in large data sets. Often a key assumption is that measurement errors in each variable are independent. This basic formulation has its roots in the work of Charles Spearman in 1904 and of Ragnar Frisch in the 1930s. Various topics such as errors-in-variables, factor analysis, and instrumental variables all refer to alternative viewpoints on this problem and on ways to account for the anticipated way that noise enters the data. In the present paper we begin by describing certain fundamental contributions by the founders of the field and provide alternative modern proofs to certain key results. We then go on to consider a modern viewpoint and novel numerical techniques to the problem. The central theme is expressed by the Frisch–Kalman dictum, which calls for identifying a noise contribution that allows a maximal number of simultaneous linear relations among the noise-free variables – a rank minimization problem. In the years since Frisch’s original formulation, there have been several insights, including trace minimization as a convenient heuristic to replace rank minimization. We discuss convex relaxations and theoretical bounds on the rank that, when met, provide guarantees for global optimality. A complementary point of view to this minimum-rank dictum is presented in which models are sought leading to a uniformly optimal quadratic estimation error for the error-free variables. Points of contact between these formalisms are discussed, and alternative regularization schemes are presented.