

Abstract: This paper presents an effective search for explanatory variables that correlate with response variable even being masked by other explanatory variables, which correspond to the explanatory variables that show no correlation but appear to be significant by including certain subset of explanatory variables in multiple regression model. Instead of all possible subset regression, the proposed search only requires pairwise Pearson's correlation coefficients between all explanatory variables and response variable, thereby enabling rapid computation. Multiple testing is utilized, followed by constructing shortest paths to response variable. A simple procedure for quantifying the potential correlations that are hidden is also presented. The proposed method is useful for explanatory analysis and works even under multicollinearity and when the number of explanatory variables is much larger than the sample size. Simulation studies demonstrate that the proposed method performs well for a high-dimensional model with severe multicollinearity, in which the elastic net fails. The practical usefulness is illustrated by a real environmental data. (This is a joint work with Dr. Masao Ueki of Kurume University.)