Rudolf Beran

Multiple penalty regression: fitting and extrapolating a discrete incomplete multi-way layout

Abstract The discrete multi-way layout is an abstract data type associated with regression, experimental designs, digital images or videos, spatial statistics, gene or protein chips, and more. The factors influencing response can be nominal or ordinal. The observed factor level combinations are finitely discrete and often incomplete or irregularly spaced. This paper develops low risk biased estimators of the means at the observed factor level combinations; and extrapolates the estimated means to larger discrete complete layouts. Candidate penalized least squares (PLS) estimators with multiple quadratic penalties express competing conjectures about each of the main effects and interactions in the analysis of variance decomposition of the means. The candidate PLS estimator with smallest estimated quadratic risk attains, asymptotically, the smallest risk over all candidate PLS estimators. In the theoretical analysis, the dimension of the regression space tends to infinity. No assumptions are made about the unknown means or about replication.