Adaptive Rank Penalized Estimators in Multivariate Regression

Marten Wegkamp
Florida State University

We introduce a new criterion, the Rank Selection Criterion (RSC), for selecting the optimal reduced rank estimator of the coefficient matrix in multivariate response regression models. The corresponding RSC estimator minimizes the Frobenius norm of the fit plus a regularization term proportional to the number of parameters in the reduced rank model.

The rank of the RSC estimator provides a consistent estimator of the rank of the coefficient matrix; in general the rank of our estimator is a consistent estimate of the effective rank, which we define as the number of singular values of the target matrix that are appropriately large. The consistency results are valid not only in the classic asymptotic regime, when n, the number of responses, and p, the number of predictors, stay bounded, and m, the number of observations, grows, but also when either, or both, n and p grow, possibly much faster than m.

We establish minimax optimal bounds on the mean squared errors of our estimators. Our finite sample performance bounds for the RSC estimator show that it achieves the optimal balance between the approximation error and the penalty term.

Furthermore, our procedure has very low computational complexity, linear in the number of candidate models, making it particularly appealing for large scale problems. We contrast our estimator with the nuclear norm penalized least squares estimator (NNP), which has an inherently higher computational complexity than RSC. We show that NNP has estimation properties similar to those of RSC, albeit under stronger conditions. However, it is not as parsimonious as RSC. We offer a simple correction of the NNP estimator which leads to consistent rank estimation.

We verify and illustrate our theoretical findings via an extensive simulation study.

This is joint work with Florentina Bunea and Yiyuan She.