

ABSTRACT

THESIS: Dimension Selection Criteria For Predictor Envelopes in Univariate Regression

STUDENT: Sam Bales

DEGREE: Master of Science

COLLEGE: Science and Humanities

DATE: May 2020

PAGES: 34

Ordinary Least Squares is a ubiquitous multivariate statistical technique. Predictor envelopes have the potential to result in multifold gains in efficiency over OLS. Envelopes achieve these gains by partitioning the variability of the predictions into material and immaterial portions to the response. This requires the estimation of a subspace the predictor covariance matrix, and specifying the correct subspace dimension is critical for performance. Selecting too large of an envelope prevents the modeler from taking full advantage of the efficiency gains, and estimating too small of an envelope can result in biased estimates. Thus, three dimension selection criteria—Akaike’s Information Criteria (AIC), Bayesian Information Criteria (BIC), and Likelihood Ratio Testing (LRT)— listed in order of most conservative (largest subspace) to least conservative (smallest subspace) were investigated via a simulation study and a data application.

The simulation study was designed to answer two main questions: when do predictor envelopes offer advantages over OLS, and in those situations, what is the best dimension selection criteria to use. Simulation results indicate that envelopes offer advantages when predictor variance is low, response variation is high, in small sample sizes, and when the predictors are not normally distributed. Moreover, when envelopes offer advantages over OLS, LRT is the best dimension selection criterion. Intuition about when envelopes will offer advantages can be based on a preliminary analysis of the predictors and the correlation structure of the model. Predictor envelopes were applied to Fama-Macbeth regression for

estimating the risk premia for financial securities, resulting in an average decrease in the standard error of estimated coefficients by 50X.