Bruce E. Hansen
(University of Wisconsin)

Model Averaging, Sieve Regression, and Minimax Efficiency

This talk will review recent developments in the theory of model averaging for regression estimation. We introduce an averaging nonparametric sieve regression estimator, which can achieve a lower integrated mean-squared error (IMSE) than least-squares estimators. We develop a uniform asymptotic approximation for the IMSE, introduce a cross-validation criterion for selection of the averaging weights, and show that that cross-validation weight selection is asymptotically IMSE equivalent to the infeasible best averaging sieve approximation.

We also examine the asymptotic risk of nested least-squares averaging estimators when the averaging weights are selected to minimize a penalized least-squares criterion. We find conditions under which the asymptotic risk of the averaging estimator is globally smaller than the unrestricted least-squares estimator. For the Mallows averaging estimator under homoskedastic errors the condition takes the simple form that the regressors have been grouped in sets of four or larger. This condition is a direct extension of the classic theory of James-Stein shrinkage. This discovery suggests the practical rule that implementation of averaging estimators be restricted to models in which the regressors have been grouped in this manner. Our simulations show that this new recommendation results in substantial reduction in mean-squared error relative to averaging over all nested sub-models.

Averaging between two nested models is equivalent to shrinkage. We explore the efficiency properties of shrinkage in high dimensional parametric models using a local asymptotic framework as in Hjort and Claeskens (2003). We show that if the shrinkage dimension exceeds two, the asymptotic risk of shrinkage estimators is strictly less than that of the MLE. This reduction holds globally in the parameter space. We also provide a new local asymptotic minimax efficiency bound. We show that shrinkage estimators asymptotically achieve this local asymptotic minimax bound when the shrinkage dimension is high. This theory is a combination and extension of standard asymptotic efficiency theory (Hájek, 1972) and local minimax efficiency theory for Gaussian models (Pinsker, 1980).