

SOMS Friday Seminar
April 4, 2008
SMC 305
2:30-3:25

Adaptive Ridge Selector (ARiS)

Artin Armagan
Statistics Ph.D. candidate

ABSTRACT

We introduce a new shrinkage variable selection operator which we term Adaptive Ridge Selector (ARiS). This approach is inspired by the Relevance Vector Machine (RVM) of Tipping (2001), which uses a Bayesian hierarchical linear model to do sparse estimation. RVM was originally introduced to obtain sparse solutions in the case of kernel regression where one has many highly correlated bases (features).

Extending the RVM algorithm, we include a proper prior distribution for the precisions of the regression coefficients along with a hyper- parameter to be chosen. Based upon this model, we derive the full set of conditional posterior distributions for parameters as would typically be done when applying Gibbs sampling. However, instead of simulating samples from the posterior distribution in order to estimate posterior means of quantities, we apply the Lindley-Smith mechanism (Lindley 1972). This approach sequentially maximizes the conditional distributions, in order to find the joint maximum of the posterior distribution given the value of the hyper-parameter. An empirical Bayes method is proposed for choosing this hyper-parameter leading to ARiS-eB. Having moved from a Bayes argument, we also look at the problem from a penalized least squares estimation angle.

From the conventional viewpoint, the proposed method eliminates the need for combinatorial search techniques over a discrete model space, converting the model selection problem into the maximization of the marginal likelihood over a one dimensional continuous space.

Close similarities exist between this estimator obtained and the lasso-type shrinkage estimators. The lasso (Tibshirani 1996) and its variants, as will be thoroughly discussed, use 1-norm for regularization leading to sparse solutions. The estimator proposed here is contrasted with various other shrinkage estimators along with simulation studies and real data examples.

Inference is also possible using a very straight forward Gibbs sampling procedure after the active variables are determined in the model. The model is also extended to handle departures from normality in the likelihood.

Refreshments provided after seminar