Statistica Sinica: Preprint
doi:10.5705/ss.202016.0167

Bayesian model selection procedures based on nonlocal alternative prior densities are extended to ultrahigh dimensional settings and compared to other variable selection procedures using precision-recall curves. Variable selection procedures included in these comparisons include methods based on g-priors, reciprocal lasso, adaptive lasso, scad, and minimax concave penalty criteria. The use of precision-recall curves eliminates the sensitivity of our conclusions to the choice of tuning parameters. We find that Bayesian selection procedures based on nonlocal priors are competitive to all other procedures in a range of simulation scenarios, and we subsequently explain this favorable performance through a theoretical examination of their consistency properties. When certain regularity conditions apply, we demonstrate that the nonlocal procedures are consistent for linear models even when the number of covariates p increases sub-exponentially with the sample size n. Methods based on Zellner's g-prior are also found to be competitive with penalized likelihood methods in identifying the true model, but the posterior distribution on the model space induced by this method is much more dispersed than the posterior distribution induced on the model space by the nonlocal prior methods.

Statistica Sinica