Variable Selection in Causal Inference Using Penalization: Paper Summary

This summary introduces a method designed for variable selection in a doubly robust average treatment effect estimator, focusing on simple additive linear propensity score and outcome models, particularly in the context of non-mediated treatment effects. Paper can be found here. The doubly robust nature of the estimator ensures consistency even if either the propensity score or outcome regression model is misspecified. The authors propose a strategy for variable selection, emphasizing the penalization of a reparametrized joint likelihood of outcome and treatment models to address the challenge of missing important confounders, especially in small to moderate sample sizes.

The method introduces the concept of non-ignorable confounders, weak predictors of outcome or treatment with a better chance of being selected. The Maximum Penalized Likelihood Estimator (MLP) is employed for variable selection, prioritizing a trade-off between bias and variance to construct a parsimonious model. Traditional variable selection models, such as backward elimination and permutation methods, may overlook weakly associated covariates, leading to increased model misspecification. As an illustrative example, the authors discuss scenarios where a predictor weakly predicts both the outcome and the treatment, highlighting that the proposed method effectively penalizes such variables and their corresponding parameters, in contrast to a scenario where a variable barely predicts the outcome but strongly predicts the treatment assignment.

The methodology involves estimating propensity score parameters using shrinkage, creating a random variable (Observed Treatment - Propensity Score), and fitting a regression model for the outcome based on the defined random variable and selected confounders for the propensity score model.

For the production of doubly robust estimators, the authors employ LASSO and Smoothly Clipped Absolute Deviation penalty (SCAD), comparing their method with Inverse Probability Weighting (IPW) and Regression model separately, as well as an Oracle model with all true predictors. Simulation results for sample sizes of 300 and 500, with predictors exceeding N, indicate that doubly robust estimators using LASSO and SCAD achieve Mean Squared Error (MSE) and variance similar to the Oracle, with some bias but not significantly different from the Oracle method. Applied data analysis of economic data reinforces these findings.