

Abstract: Instrumental variables (IVs) are a commonly used technique for causal inference from observational data. In practice, weak IVs can render the approach ineffective and cause estimates to be biased. To overcome the problem of weak instruments, literature has suggested using optimal instruments. Estimating optimal instruments from data in most cases is infeasible, and even low order approximations could lead to basis functions that grow exponentially. To this end, we formulate the problem of estimating strong instrumental variable as a machine learning problem. We develop a novel algorithm, called *MLIV algorithm*, which allows learning of optimal instruments and causal inference to be simultaneously performed from sample data. We demonstrate the consistency and asymptotic normality of our estimators, regardless of the machine learning approach used. Simulations and application to real-world data demonstrate that the algorithm is highly effective for causal inference from observational data.