ABSTRACT

Structural equation models (SEMs) are widely used in sciences, ranging from economics to psychology, to uncover causal relationships underlying a complex system under consideration and estimate structural parameters of interest. We study estimation in a class of generalized SEMs where the object of interest is defined as the solution to a linear operator equation. We formulate the linear operator equation as a min-max game, where both players are parameterized by neural networks (NNs), and learn the parameters of these neural networks using the stochastic gradient descent. We consider both 2-layer and multi-layer NNs with ReLU activation functions and prove global convergence in an overparametrized regime, where the number of neurons is diverging. The results are established using techniques from online learning and local linearization of NNs. For the first time we provide a tractable estimation procedure for SEMs based on NNs with provable convergence and without the need for double samples.