Full Law Identification In Graphical Models Of Missing Data: Completeness Results

Razieh Nabi, John Hopkins University, Computer Science Department

Abstract

Missing data has the potential to affect analyses conducted in all fields of scientific study, including healthcare, economics, and the social sciences. Several approaches to unbiased inference in the presence of non-ignorable missingness rely on the specification of the target distribution and its missingness process as a probability distribution that factorizes with respect to a directed acyclic graph. In this paper, we address the longstanding question of the characterization of models that are identifiable within this class of missing data distributions. We provide the first completeness result in this field of study – necessary and sufficient graphical conditions under which, the full data distribution can be recovered from the observed data distribution. We then simultaneously address issues that may arise due to the presence of both missing data and unmeasured confounding, by extending these graphical conditions and proofs of completeness, to settings where some variables are not just missing, but completely unobserved.