Dimension reduction for non-elliptically distributed predictors

Yuexiao Dong

The Pennsylvania State University
yud101@stat.psu.edu

Abstract

Sufficient dimension reduction methods often require stringent conditions on the joint distribution of the predictor, or, when such conditions are not satisfied, rely on marginal transformation or reweighting to fulfill them approximately. For example, a typical dimension reduction method would require the predictor to have elliptical or even multivariate normal distribution. We reformulate the commonly used dimension reduction methods, via the notion of "central solution space", so as to circumvent the requirement such strong assumptions, while at the same time preserve the desirable properties of the classical methods, such as $\sqrt{n}$-consistency and asymptotic normality.

Most methods in the sufficient dimension reduction literature are based on inverse conditional moments. Those methods can be further divided into two categories: first-order methods that depend on moments or inverse conditional moments such as $E(XY^k)$ and $E(X|Y)$; second-order methods that depend on moments or inverse conditional moments such as $E(XY^k)$, $E(X|Y)$, $E(Y^kXX^T)$, and $E(XX^T|Y)$, where k is an integer. First-order methods include Sliced Inverse Regression (SIR; Li, 1991), Ordinary Least Squares (OLS; Li and Duan, 1989), Parametric Inverse Regression (PIR; Bura and Cook, 2001), Canonical Correlation (Fung et al., 2002), and Kernel Inverse Regression (KIR; Zhu and Fang, 1996; Ferre and Yao, 2005). Second-order methods include Principal Hessian Directions (PHD; Li, 1992 and Cook, 1998b), Sliced Average Variance Estimator (SAVE; Cook and Weisberg, 1991), SIRII (Li, 1991), Contour Regression (Li et al., 2005), and Directional Regression (DR; Li and Wang, 2007). The efficacy of aforementioned methods relies on the elliptical distribution assumption of the predictor X. The main focus of this presentation is to generalize all these methods to work for non-elliptically distributed predictors. The new methods will be compared with existing methods by simulation, and the results show a significant improvement over traditional dimension reduction methods when the elliptical assumption is violated.