Journal of Multivariate Analysis 109 (2012) 61–72

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On model-free conditional coordinate tests for regressions

Zhou Yu *, Lixing Zhu, Xuerong Meggie Wen

School of Finance and Statistics, East China Normal University, China
Department of Mathematics, Hong Kong Baptist University, Hong Kong
Department of Mathematics and Statistics, Missouri University of Science and Technology, United States

ARTICLE INFO

Article history:
Received 13 April 2011
Available online 26 February 2012

AMS subject classifications:
62H15
62H25
62H86

Keywords:
Conditional coordinate test
Sufficient dimension reduction
Sliced inverse regression

ABSTRACT

Existing model-free tests of the conditional coordinate hypothesis in sufficient dimension reduction (Cook (1998) [3]) focused mainly on the first-order estimation methods such as the sliced inverse regression estimation (Li (1991) [14]). Such testing procedures based on quadratic inference functions are difficult to be extended to second-order sufficient dimension reduction methods such as the sliced average variance estimation (Cook and Weisberg (1991) [9]). In this article, we develop two new model-free tests of the conditional predictor hypothesis. Moreover, our proposed test statistics can be adapted to commonly used sufficient dimension reduction methods of eigendecomposition type. We derive the asymptotic null distributions of the two test statistics and conduct simulation studies to examine the performances of the tests.

1. Introduction

For parametric regressions, hypothesis testing for predictor contributions to the response is a well developed research area. For instance, in the linear models, t test is often applied to check the contribution of every predictor. However, for semiparametric models, this topic has not yet received enough attention because how to construct a test therein is a challenge. To attack this problem, Cook [4] investigated this issue in a dimension reduction framework.

For a typical regression problem with a univariate random response Y and a vector of random predictors $\mathbf{X} = (X_1, \ldots, X_p)^T \in \mathbb{R}^p$, the goal is to understand how the conditional distribution $Y|\mathbf{X}$ depends on the value of $\mathbf{X}$. The spirit of sufficient dimension reduction [14,3] is to reduce the dimension of $\mathbf{X}$ without loss of information on the regression and without requiring a pre-specified parametric model. Assuming the following semiparametric regression model: $Y = g(\boldsymbol{\beta}_1^T\mathbf{X}, \boldsymbol{\beta}_2^T\mathbf{X}, \ldots, \boldsymbol{\beta}_d^T\mathbf{X}, \epsilon)$, where $g(\cdot)$ is an unknown function and ϵ is an unknown random error independent of $\mathbf{X}$, we can see that the conditional distribution of $Y|(\boldsymbol{\beta}_1^T\mathbf{X}, \ldots, \boldsymbol{\beta}_d^T\mathbf{X})$ is the same as that of $Y|\mathbf{X}$ for all values of $\mathbf{X}$. Hence, these $\boldsymbol{\beta}$'s provide a parsimonious characterization of the conditional distribution of $Y|\mathbf{X}$. We call them the *effective (sufficient) directions* [14,3]. When d is small which is often the case in real applications, the original regression problem (data) can be effectively reduced by projecting $\mathbf{X}$ along these effective directions.

More formally, we search for subspaces $\mathcal{S} \subseteq \mathbb{R}^p$ such that $Y \perp\!\!\!\perp \mathbf{X} | P_{\mathcal{S}}\mathbf{X}$ where $\perp\!\!\!\perp$ indicates independence, and $P_{(\cdot)}$ stands for a projection operator with respect to the standard inner product. The intersection of all such $\mathcal{S}$ is defined as the *central subspace*, denoted as $\mathcal{S}_{Y|X}$ [3], which almost always exists in practice under mild conditions [25]. We assume the existence of the central subspaces throughout this article. Sufficient dimension reduction is concerned with making inferences for the central subspace. $d = \dim(\mathcal{S}_{Y|X})$ is called the structural dimension of the regression. Unlike other nonparametric approaches, sufficient dimension reduction can often avoid the curse of dimensionality. Many sufficient dimension reduction methods enjoy $\sqrt{n}$ convergence rates since they exploit the global features of the dependence of Y on $\mathbf{X}$.

* Corresponding author at: East China Normal University, School of Finance and Statistics, Shanghai, China.
E-mail address: yz19830224@gmail.com (Z. Yu).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2012.02.004

Sufficient dimension reduction has been a promising field during the past decades. It has attracted considerable interests, and many methods have been developed. Among them, sliced inverse regression (SIR; [14]), sliced average variance estimation (SAVE; [9]), minimum average variance estimation [22], inverse regression estimation [7] and directional regression (DR; [17]) are the most widely investigated methods in the literature. All these aforementioned methods except Xia et al. [22] mainly focus on the estimation of the central subspace.

Other than estimating the central subspace, it is also of interest to evaluate the predictor effects in a model free setting. Cook [4] considered two types of hypotheses to test the significance of subsets of predictors under the framework of sufficient dimension reduction. The first type is the *Marginal Coordinate Hypothesis*: $\mathcal{P}_{\mathcal{H}}\mathcal{S}_{Y|X} = \mathcal{O}_p$ versus $\mathcal{P}_{\mathcal{H}}\mathcal{S}_{Y|X} \neq \mathcal{O}_p$. The second type is called the *Conditional Coordinate Hypothesis*:

$$\mathcal{P}_{\mathcal{H}}\mathcal{S}_{Y|X} = \mathcal{O}_p \quad \text{versus} \quad \mathcal{P}_{\mathcal{H}}\mathcal{S}_{Y|X} \neq \mathcal{O}_p \quad \text{given } d; \tag{1.1}$$

where $\mathcal{H}$ is an r-dimensional ($r \leq p-d$) user-selected subspace of the predictor space and $\mathcal{O}_p$ indicates the origin in $\mathbb{R}^p$. For example, suppose that $\mathbf{X}^T = (\mathbf{X}_1^T, \mathbf{X}_2^T)$, where $\mathbf{X}_1 \in \mathbb{R}^r$ and $\mathbf{X}_2 \in \mathbb{R}^{p-r}$, and we would like to test if $\mathbf{X}_1$ makes any contributions to the regression $Y|\mathbf{X}$, we then consider these two types of hypotheses tests with $\mathcal{H} = \text{Span}((I_r, 0)^T)$. Although in general, $\mathcal{H}$ need not correspond to a subset of predictors (coordinates).

Hence, both the marginal coordinate hypothesis and the conditional coordinate hypothesis can be used to test the contributions of selected predictors without requiring a pre-specified model about the original regression $Y|\mathbf{X}$. When d, the structural dimension of the regression, is specified as a modeling device, or inferences on d result in a clear estimate, a conditional coordinate hypothesis test will be the natural choice. Otherwise, a marginal coordinate hypothesis would be tested. We would expect that the conditional coordinate hypothesis will provide us with greater power when a correct d is given prior to testing predictors. On the other hand, when d is misspecified, a conditional coordinate hypothesis test might lead to misleading results, while the marginal coordinate hypothesis test should be considered. Although simulation studies provided in Section 4 suggest that the misspecification of d need not be a worrisome issue in practice.

Based on a nonlinear least squares formulation of the sliced inverse regression estimation, Cook [4] constructed asymptotic tests for the marginal and conditional coordinate hypotheses. Cook and Ni [7] showed how to test marginal (conditional) coordinate hypotheses using various quadratic inference functions, which stimulated the tests of conditional independence hypotheses based on the minimum discrepancy approach [7] and the covariance inverse regression estimation [8].

All the aforementioned tests are based on the first moment of the inverse regression of $\mathbf{X}|Y$ that are called the first-order sufficient dimension reduction methods. Note that these tests for the predictor contributions might be invalid when the response surface is symmetric about the origin since these first-order sufficient dimension reduction methods themselves would fail in such cases. Therefore, it is of great interest to consider coordinate tests using the second-order sufficient dimension reduction methods which involve both the first and second moments of the inverse regression of $\mathbf{X}|Y$. However, the commonly used second-order sufficient dimension reduction methods such as the sliced average variance estimation [9], and the directional regression [17], are very different from those first-order methods which could be derived from quadratic inference functions. Hence, the asymptotic tests developed by Cook and Ni [7] are not directly applicable. Shao et al. [21] provided a marginal coordinate test based on the sliced average variance estimation. But to the best of our knowledge, there are no methods available in the literature for testing of the conditional coordinate hypotheses of (1.1) with second-order dimension reduction methods. To address this issue, we in this article present two new tests of conditional coordinate hypotheses which could be adapted to essentially all existing sufficient dimension reduction methods of the eigendecomposition type, including both the sliced inverse regression estimation and the sliced average variance estimation methods.

The rest of the paper is organized as follows. Section 2 revisits several moment based sufficient dimension reduction methods. In Section 3, we construct two new tests and present their asymptotic null distributions. Sections 4 and 5 are concerned with simulation studies and a real data application. We conclude with a brief discussion in Section 6. For easy of exposition, the proofs of the asymptotic results are deferred to the Appendix A.

2. Sufficient dimension reduction methods revisited

Let $\boldsymbol{\mu} = E(\mathbf{X})$, $\boldsymbol{\Sigma} = \text{Var}(\mathbf{X})$, and $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$ be the standardized predictor. Many moment based sufficient dimension reduction methods can be formulated as the following eigendecomposition problem:

$$\mathcal{M}\boldsymbol{\eta}_i = \lambda_i\boldsymbol{\eta}_i, \quad i = 1, \ldots, p, \tag{2.2}$$

where $\mathcal{M}$ is the $\mathbf{Z}$ scale method-specific candidate matrix. Under certain conditions imposed only on the marginal distribution of the predictor, the eigenvectors $(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_d)$ corresponding to the nonzero eigenvalues $\boldsymbol{\lambda}_1 \geq \cdots \geq \boldsymbol{\lambda}_d$ form a basis of the $\mathbf{Z}$ scale central subspace $\mathcal{S}_{Y|Z}$. Then by the invariance property $\mathcal{S}_{Y|X} = \boldsymbol{\Sigma}^{-1/2}\mathcal{S}_{Y|Z}$ as described by Cook [3], $\boldsymbol{\beta} = (\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\eta}_1, \ldots, \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\eta}_d)$ forms a basis of $\mathcal{S}_{Y|X}$.

As most of commonly used sufficient dimension reduction methods that target $\mathcal{S}_{Y|Z}$ are of candidate matrices satisfying the above eigendecomposition, we only list some as follows:

Sliced Inverse Regression: $\mathcal{M} = \text{Var}\{E(\mathbf{Z}|Y)\}$;

Sliced Average Variance Estimation: $\mathcal{M} = E\{I_p - \text{Var}(\mathbf{Z}|Y)\}^2$;

Directional Regression: $\mathcal{M} = 2E\{E^2(\mathbf{Z}\mathbf{Z}^T|Y)\} + 2E^2\{E(\mathbf{Z}|Y)E(\mathbf{Z}^T|Y)\} + 2E\{E(\mathbf{Z}^T|Y)E(\mathbf{Z}|Y)\}E\{E(\mathbf{Z}|Y)E(\mathbf{Z}^T|Y)\} - 2I_p$.