Journal of Multivariate Analysis 136 (2015) 75–85

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

On testing common indices for two multi-index models: A link-free approach

Xuejing Liu[a], Zhou Yu[b], Xuerong Meggie Wen[a,*], Robert Paige[a]

[a] Department of Mathematics and Statistics, Missouri University of Science and Technology, MO 65409, USA

[b] School of Finance and Statistics, East China Normal University, Shanghai, China

ARTICLE INFO

Article history:
Received 28 May 2014
Available online 21 January 2015

AMS subject classifications:
62H86
62H15
62H25
62G08

Keywords:
Multi-index models
Sufficient dimension reduction
Sliced inverse regression
Sliced average variance estimation
Directional regression

ABSTRACT

We propose a link-free procedure for testing whether two multi-index models share identical indices via the sufficient dimension reduction approach. Test statistics are developed based upon three different sufficient dimension reduction methods: (i) sliced inverse regression, (ii) sliced average variance estimation and (iii) directional regression. The asymptotic null distributions of our test statistics are derived. Monte Carlo studies are performed to investigate the efficacy of our proposed methods. A real-world application is also considered.

1. Introduction

For a regression problem with univariate response Y and p-dimensional predictors $\mathbf{X} = (X_1, \ldots, X_p)^T$, we consider the following generalized multi-index model

$$Y = g(\boldsymbol{\beta}_1^T\mathbf{X}, \ldots, \boldsymbol{\beta}_d^T\mathbf{X}; \epsilon), \tag{1.1}$$

where $g(\cdot)$ is an unknown link function, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_d)$ is a $p \times d$ matrix, $d \le p$, and the random error ϵ is independent with $\mathbf{X}$. Model (1.1) is a very general semiparametric model which includes the multi-index model [12,27] and the single-index model [11,29] with $Y = g(\boldsymbol{\beta}^T\mathbf{X}) + \epsilon$ as special cases. One is usually concerned with estimation of indices $\boldsymbol{\beta}$, the total number of indices d and the link function $g(\cdot)$ [9]. We, however, focus on testing if two multi-index models share identical indices (subspaces). Specifically, consider two d-dimensional multi-index models for two populations (groups):

$$Y = g_1(\boldsymbol{\beta}_1^T\mathbf{X}, \ldots, \boldsymbol{\beta}_d^T\mathbf{X}; \epsilon_1), \quad \textit{for group } 1;$$

$$Y = g_2(\boldsymbol{\xi}_1^T\mathbf{X}, \ldots, \boldsymbol{\xi}_d^T\mathbf{X}; \epsilon_2), \quad \textit{for group } 2. \tag{1.2}$$

Since the identifiable parameters here are the subspaces spanned by the columns of $\boldsymbol{\beta}$ and $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_d)$, rather than $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ themselves, we develop a test of null hypothesis

$$\mathrm{span}(\boldsymbol{\beta}) = \mathrm{span}(\boldsymbol{\xi}), \tag{1.3}$$

* Corresponding author.
E-mail address: wenx@mst.edu (X.M. Wen).

http://dx.doi.org/10.1016/j.jmva.2015.01.009

where both $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ are $p \times d$ matrices. This hypothesis is similar in nature to the null hypothesis of common principal component subspaces for Common PCA considered in [20].

A hypothesis test of this type might be of special interest in many applications involving two datasets, where the same variables are being measured on objects from two different groups, and for which it is of interest to determine how similar the two groups are with respect to the span of the indices of predictor vectors regardless of the unknown link functions.

Consider the AIS dataset discussed by Weisberg [24], which contains information on the lean body mass L and other physical and hematological measurements ($\mathbf{X}$), from 102 male and 100 female elite Australian athletes who trained at the Australian Institute of Sport. We investigate how the relationship between the body fat and various predictors varies with gender. Suppose that subject matter knowledge and prior modeling experience suggest that a d-dimensional multi-index model of the form (1.1) applies to both female and male groups, naturally, we would like to know if the equivalent set of indices of the hematological measurements serve for both genders. Informal comparisons such as those based upon graphical methods can be carried out. However, such comparisons might become unwieldy when d is greater than 2, and the resulting conclusions could be overly subjective. Hence, a formal test seems necessary here. This is the motivation for our development of a test statistic for the null hypothesis in (1.3).

We propose a link-free test for testing hypothesis of (1.3) via a sufficient dimension reduction approach [15,3]. For a regression problem, the scope of sufficient dimension reduction is to seek a minimal set of indices of $\mathbf{X}$, say $\boldsymbol{\beta}^T\mathbf{X} = (\boldsymbol{\beta}_1^T\mathbf{X}, \ldots, \boldsymbol{\beta}_d^T\mathbf{X})$, for which the distribution of $Y|\boldsymbol{\beta}^T\mathbf{X}$ is the same as the original regression $Y|\mathbf{X}$, without assuming a parametric model. Numerous approaches are available in the literature including sliced inverse regression (SIR; [15]), sliced average variance estimation (SAVE; [5]), minimum average variance estimation (MAVE; [28]), directional regression (DR; [18]), likelihood acquired directions (LAD; [4]), and dimension reduction via central solution space [16].

The rest of this article is organized as follows. In Section 2, we give a brief review of sufficient dimension reduction methods. Specifically, we focus on those methods based upon a spectral decomposition [25]. In Section 3, we present our link-free test statistic for null hypothesis (1.3). The asymptotic distribution of our test statistic is also discussed. We illustrate the performance of our method with Monte Carlo studies in Section 4. We then apply our method to the AIS dataset. Brief conclusions and a discussion on the future research directions are given in Section 5. For ease of exposition, we defer some technical details to Appendix.

2. A brief review on sufficient dimension reduction: the spectral decomposition approach

In this section, we give a brief review on how to use sufficient dimension reduction to make inference about span($\boldsymbol{\beta}$) in model (1.1). In particular, we consider three commonly used sufficient dimension reduction methods: SIR, SAVE and DR.

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

$$\mathbf{M}_z\boldsymbol{\eta}_i = \lambda_i\boldsymbol{\eta}_i, \quad i = 1, \ldots, p,$$

where $\mathbf{M}_z$ is the $\mathbf{Z}$ scale method-specific candidate matrix. Assuming the *linearity condition* [15] holds, which is a mild condition imposed on the marginal distribution of the predictors alone, the eigenvectors $(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_d)$ corresponding to the non-zero eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$ form a basis of the $\mathbf{Z}$ scale central subspace $\mathcal{S}_{Y|\mathbf{Z}}$. Then by the invariance property $\mathcal{S}_{Y|\mathbf{X}} = \boldsymbol{\Sigma}^{-1/2}\mathcal{S}_{Y|\mathbf{Z}}$, 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 for $\mathcal{S}_{Y|\mathbf{X}}$. The linearity condition, which basically requires that $\mathrm{E}(\mathbf{X}|\boldsymbol{\beta}^T\mathbf{X})$ is a linear function of $\boldsymbol{\beta}^T\mathbf{X}$, is a common assumption in dimension reduction methods and holds for elliptically contoured predictors [8]. Additionally, Hall and Li [10] showed that as the number of predictors p increases, the linearity condition holds to a reasonable approximation in many problems.

For the three sufficient dimension reduction methods that target $\mathcal{S}_{Y|\mathbf{Z}}$, the corresponding candidate matrices are summarized below:

$$\begin{aligned}
\text{Sliced Inverse Regression:} \quad & \mathbf{M}_z = \mathrm{Var}\{\mathrm{E}(\mathbf{Z}|Y)\}; \\
\text{Sliced Average Variance Estimation:} \quad & \mathbf{M}_z = \mathrm{E}\{I_p - \mathrm{Var}(\mathbf{Z}|Y)\}^2; \\
\text{Directional Regression:} \quad & \mathbf{M}_z = 2\mathrm{E}\{\mathrm{E}^2(\mathbf{Z}\mathbf{Z}^T)\} + 2\mathrm{E}^2\{\mathrm{E}(\mathbf{Z}|Y)\mathrm{E}(\mathbf{Z}^T|Y)\} \\
& \quad + 2\mathrm{E}\{\mathrm{E}(\mathbf{Z}^T|Y)E(\mathbf{Z}|Y)\}\mathrm{E}\{\mathrm{E}(\mathbf{Z}|Y)\mathrm{E}(\mathbf{Z}^T|Y)\} - 2I_p.
\end{aligned}$$

Although in the literature, people tend to work with standardized predictors, for our purpose, it is easier to describe the candidate matrices in terms of the original predictor $\mathbf{X}$. Since we will make use of the eigenprojection corresponding to the non-zero eigenvalues, the $\boldsymbol{\beta}_i = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\eta}_i$ provided by the above approach are orthonormal under the inner-product of $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T\boldsymbol{\Sigma}\mathbf{b}$, but not the regular dot product, which induces unnecessary difficulty to the development of our test statistic. In this paper, we work directly with the original predictor $\mathbf{X}$, and, as such, use the following symmetric candidate matrices $\mathbf{M}$:

$$\begin{aligned}
\text{SIR:} \quad & \mathbf{M} = \boldsymbol{\Sigma}^{-1}\mathrm{Var}\{\mathrm{E}(\mathbf{X}|Y)\}\boldsymbol{\Sigma}^{-1}; \\
\text{SAVE:} \quad & \mathbf{M} = \boldsymbol{\Sigma}^{-1}\mathrm{E}\{\boldsymbol{\Sigma} - \mathrm{Var}(\mathbf{X}|Y)\}^2\boldsymbol{\Sigma}^{-1}; \\
\text{DR:} \quad & \mathbf{M} = \boldsymbol{\Sigma}^{-1}\mathrm{E}\{2\boldsymbol{\Sigma} - \mathrm{E}\big((\widetilde{\mathbf{X}} - \mathbf{X})(\widetilde{\mathbf{X}} - \mathbf{X})^T|Y, \widetilde{Y}\big)\}^2\boldsymbol{\Sigma}^{-1},
\end{aligned}$$