Economics Letters 156 (2017) 162–167

Contents lists available at ScienceDirect

Economics Letters

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

A martingale-difference-divergence-based test for specification[☆]

Liangjun Su[*], Xin Zheng

School of Economics, Singapore Management University, Singapore

HIGHLIGHTS

- We propose a novel martingale-difference-divergence-based test for specification.
- The test does not require any nonparametric estimation.
- The test is applicable even if we have many covariates in the regression model.
- The test has superb finite sample performance and dominates its competitors.

ARTICLE INFO

Article history:
Received 20 January 2017
Received in revised form
30 March 2017
Accepted 5 May 2017
Available online 12 May 2017

JEL classification:
C12
C14
C21

Keywords:
Distance covariance
Integrated conditional moment test
Martingale difference divergence
Martingale difference correlation
Specification test

ABSTRACT

In this paper we propose a novel consistent model specification test based on the martingale difference divergence (MDD) of the error term given the covariates. The MDD equals zero if and only if error term is conditionally mean independent of the covariates. Our MDD test does not require any nonparametric estimation under the alternative and it is applicable even if we have many covariates in the regression model. We establish the asymptotic distributions of our test statistic under the null and a sequence of Pitman local alternatives converging to the null at the usual parametric rate. Simulations suggest that our MDD test has superb performance in terms of both size and power and it generally dominates several competitors. In particular, it is the only test that has well controlled size in the presence of many covariates and reasonable power against high frequency alternatives as well.

1. Introduction

In this paper we propose a new test for the correct specification of a parametric conditional mean model based on a variant of the *martingale difference divergence* (MDD hereafter) measure of conditional mean dependence between two random variables. In a sequence of papers, Székely et al. (2007), Székely and Rizzo (2009) and Székely and Rizzo (2014) propose to use distance covariance and distance correlation to measure the dependence between two random vectors which exhibit various nice properties. Such measures have been explored for feature screening in high dimensional regressions; see, e.g., Li et al. (2012). When one of the two random variables is a scalar, Shao and Zhang (2014, SZ hereafter) propose to use MDD to measure the conditional mean dependence of the scalar random variable given a random vector (see the definition of MDD in (2.4) in the next section). Like the relationship between covariance and correlation, the MDD can also be rescaled to ensure that it lies between 0 and 1, yielding the *martingale difference correlation* (MDC) measure of a scalar variable given a random vector. MDD measures the departure of the conditional mean independence between a scalar response variable and a vector of covariates, which is a natural extension of the distance correlation measure proposed by Székely et al. (2007). MDD and MDC have many nice properties. First, both of them are nonnegative and equal zero if and only if the scalar response variable is conditionally mean independent of the covariates. This suggests that we can propose a test for the conditional mean independence hypothesis which is widely used in econometrics and statistics. Second, both measures have a closed-form formula that is only involved with certain expectation and norm calculations so that they can be

☆ Su gratefully acknowledges the Lee Kong Chian Fund for Excellence.
* Correspondence to: School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore.
E-mail address: ljsu@smu.edu.sg (L. Su).

http://dx.doi.org/10.1016/j.econlet.2017.05.002
0165-1765/

easily estimated from the data based on the sample analogue principle. Third, the measures are dimension-free in the sense that the dimension of the conditioning variable is allowed to be large but finite. Indeed, SZ use MDC as a method to conduct high-dimensional variable selection to screen out variables that do not contribute to the conditional mean of the response variable given the covariates.

One drawback of SZ's original MDD and MDC measure is that when they are used for variable screening, both the response variable and covariates need to be observed. Therefore, we propose a variant of MDD that is used to measure the conditional mean independence of a scalar random error term given the covariates. With this variant, we propose a new consistent test for the null hypothesis that a parametric conditional mean model is correctly specified. Under the null hypothesis, the error term from the correctly specified model is conditionally mean independent of the regressors in the model and has mean zero. Since the error term is not observed, we propose to estimate it from the null model and construct a test statistic based on the sample analogue of this new MDD measure. We study the asymptotic distributions of the test statistic under the null and under a sequence of Pitman local alternatives. Our test shares many nice properties that a typical nonsmoothing test might have. First, its limiting distribution under the null is a mixture of central chi-square distributions that is not asymptotically pivotal. So we propose a wild bootstrap method to obtain the bootstrap p-value or critical value. Second, our test has nontrivial asymptotic power against local alternatives converging to the null at the usual parametric rate. More importantly, our test is free of the choice of any smoothing parameter (e.g., the bandwidth in kernel-based tests or the number of sieve approximating terms in sieve-based tests) and it does not suffer from the curse of dimensionality associated with kernel- or sieve-based tests. In principle, our test works for any finite dimensional regression problem where the number of covariates, q, can be large. But for the derivation of our asymptotic distribution theory, we still need restrict q to be fixed. We conduct some Monte Carlo simulations and compare our test with some popular tests in the literature. Our simulation results indicate that our MDD-based test generally outperforms its competitors, especially for the case of high frequency alternatives and for the case of many covariates (e.g., $q = 10, 20$). To the best of our knowledge, this paper is the first to consider consistent model specification test in the presence of many covariates where existing tests tend to fail due to the notorious curse of dimensionality.

The rest of the paper is organized as follows. We introduce the hypotheses and the test statistic in Section 2. We study the asymptotic distributions of the test statistic under the null hypothesis and under a sequence of Pitman local alternatives in Section 3. We compare the MDD test with several popular tests through Monte Carlo simulations in Section 4. Section 5 concludes. The proofs of all results are relegated to the online supplementary Appendix.

Notation. For any matrix or vector A, $\|A\|$ denotes its Euclidean norm. The operators $\xrightarrow{p}$ and $\xrightarrow{d}$ denote convergence in probability and distribution, respectively.

2. The hypotheses and statistic

In this section we state the hypotheses and introduce the test statistic.

2.1. The hypotheses

We consider the following parametric regression model

$$Y_i = g(X_i;\beta) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where Y_i is a scalar dependent variable, X_i is a $q \times 1$ vector of covariates, β is a $d \times 1$ vector of unknown parameters, and ε_i is the unobserved error term. We assume that the functional form of $g(\cdot;\cdot)$ is known up to the finite dimensional parameter β. We are interested in testing the correct specification of $g(\cdot;\cdot)$. That is, we test the null hypothesis

$$\mathbb{H}_0 : P\{E(Y_i|X_i) = g(X_i;\beta_0)\} = 1 \quad \text{for some } \beta_0 \in \mathcal{B} \tag{2.2}$$

versus the alternative hypothesis

$$\mathbb{H}_1 : P\{E(Y_i|X_i) = g(X_i;\beta)\} < 1 \quad \text{for all } \beta \in \mathcal{B}, \tag{2.3}$$

where $\mathcal{B}$ is the parameter space.

2.2. Test statistic

To motivate our test statistic, we follow SZ and consider the MDD of ε given X whose square is defined by

$$\mathrm{MDD}(\varepsilon|X)^2 = \int_{\mathbb{R}^q} \left|\mathbb{E}\left[\varepsilon \exp(\mathbf{i}s'X)\right] - \mathbb{E}(\varepsilon)\,\mathbb{E}\left[\exp(\mathbf{i}s'X)\right]\right|^2 \times W(s)ds, \tag{2.4}$$

where $\mathbf{i} = \sqrt{-1}$, $W(s) = \frac{1}{c_q\|s\|^{(1+q)}}$, $c_q = \frac{\pi^{(1+q)/2}}{\Gamma((1+q)/2)}$, and $\Gamma(\cdot)$ is the complete gamma function: $\Gamma(z) = \int_0^\infty t^{z-1}\exp(-t)\,dt$. Let $(\varepsilon^\dagger, X^\dagger)$ be an independent copy of (ε, X). By Theorem 1 in SZ, we have

$$\mathrm{MDD}(\varepsilon|X)^2 = -\mathbb{E}\left\{[\varepsilon - \mathbb{E}(\varepsilon)]\left[\varepsilon^\dagger - \mathbb{E}(\varepsilon^\dagger)\right]\|X - X^\dagger\|\right\}, \tag{2.5}$$

and $\mathrm{MDD}(\varepsilon|X)^2 = 0$ if and only if $\mathbb{E}(\varepsilon|X) = \mathbb{E}(\varepsilon)$.

In our setup, ε denotes the error term in a regression such that $\mathbb{E}(\varepsilon) = 0$ is always maintained. This motivates us to consider the following variant of $\mathrm{MDD}(\varepsilon|X)^2$

$$\mathrm{MDD}^*(\varepsilon|X)^2 = -\mathbb{E}\left[\varepsilon\varepsilon^\dagger\|X - X^\dagger\|\right] + 2\mathbb{E}\left[\varepsilon\|X - X^\dagger\|\right]\mathbb{E}\left[\varepsilon^\dagger\right]. \tag{2.6}$$

The following proposition establishes the properties of $\mathrm{MDD}^*(\varepsilon|X)^2$ that serve as the basis of our test statistic.

Proposition 2.1. *Let $(\varepsilon^\dagger, X^\dagger)$ be an independent copy of (ε, X), where ε is a scalar random variable and X is a $q \times 1$ random vector. Suppose that $0 < \mathbb{E}[\varepsilon^2] < \infty$ and $0 < \mathbb{E}[\|X\|^2] < \infty$. Then*

(i) *$\mathrm{MDD}^*(\varepsilon|X)^2 \geq 0$;*
(ii) *$\mathrm{MDD}^*(\varepsilon|X)^2 = 0$ if and only if $\mathbb{E}(\varepsilon|X) = 0$ almost surely (a.s.).*

An important implication of Proposition 2.1 is that we can test (2.2) by testing whether $\mathrm{MDD}^*(\varepsilon_i|X_i)^2 = 0$, where $\varepsilon_i = Y_i - g(X_i;\beta_0)$. In practice, ε_i is not observed. We propose to estimate the model (2.1) by the nonlinear least squares (NLS) to obtain the NLS estimator $\hat\beta$ of β. Let $\hat\varepsilon_i = Y_i - g(X_i;\hat\beta)$. We propose to estimate $n\mathrm{MDD}^*(\varepsilon|X)^2$ by the following object

$$T_n = -\frac{1}{n}\sum\sum_{1\le i\ne j\le n}\hat\varepsilon_i\hat\varepsilon_j\kappa_{i,j} + \frac{2}{n}\sum\sum_{1\le i\ne j\le n}\hat\varepsilon_i\kappa_{i,j}\frac{1}{n}\sum_{k=1}^n\hat\varepsilon_k, \tag{2.7}$$

where $\kappa_{i,j} \equiv \|X_i - X_j\|$. In the special case where $g(X_i;\beta)$ is linear in X_i and β, i.e., $g(X_i;\beta) = (1, X_i')\,\beta$, we have $\sum_{i=1}^n\hat\varepsilon_i = 0$ and

$$T_n = -\frac{1}{n}\sum\sum_{1\le i\ne j\le n}\hat\varepsilon_i\hat\varepsilon_j\kappa_{i,j} \equiv T_n^\ell. \tag{2.8}$$

Other than this case, $\sum_{i=1}^n\hat\varepsilon_i$ is generally nonzero and second term in (2.7) is necessary.