Journal of Econometrics 174 (2013) 127–143

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

Optimal convergence rates, Bahadur representation, and asymptotic normality of partitioning estimators[☆]

Matias D. Cattaneo [*], Max H. Farrell

Department of Economics, University of Michigan, United States

ARTICLE INFO

Article history:
Received 31 January 2011
Received in revised form
10 December 2012
Accepted 12 February 2013
Available online 24 February 2013

JEL classification:
C14
C21

Keywords:
Nonparametric estimation
Partitioning
Subclassification
Convergence rates
Bahadur representation
Asymptotic normality

ABSTRACT

This paper studies the asymptotic properties of partitioning estimators of the conditional expectation function and its derivatives. Mean-square and uniform convergence rates are established and shown to be optimal under simple and intuitive conditions. The uniform rate explicitly accounts for the effect of moment assumptions, which is useful in semiparametric inference. A general asymptotic integrated mean-square error approximation is obtained and used to derive an optimal plug-in tuning parameter selector. A uniform Bahadur representation is developed for linear functionals of the estimator. Using this representation, asymptotic normality is established, along with consistency of a standard-error estimator. The finite-sample performance of the partitioning estimator is examined and compared to other nonparametric techniques in an extensive simulation study.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Nonparametric estimation of an unknown conditional expectation function and its derivatives is an important problem in econometrics (see, e.g., Ichimura and Todd, 2007 and references therein). In many applications the object of interest is a conditional expectation, its derivative, or functional thereof, while in other cases their nonparametric estimators are employed as a first step in a semiparametric procedure. The implementation of nonparametric estimators requires suitable large sample properties, including sufficiently rapid rates of convergence and known asymptotic distributions. Series- and kernel-based methods are examples whose properties are now well understood.

This paper studies the large sample properties of an estimator of the regression function and its derivatives known as *partitioning*. This estimation strategy is alternatively referred to as *blocking*, *subclassification*, or *stratification*. The estimator is constructed by partitioning the support of the conditioning variables into disjoint cells, which become smaller with the sample size, and within each the unknown regression function (and its derivatives) is approximated by linear least-squares using a fixed-order polynomial basis (other bases are possible). Consistent estimation is achieved as the cells become small enough to remove the error of the parametric approximation. For a recent textbook discussion of this estimation strategy, see Györfi et al. (2002, Chapter 4). After the necessary notation and assumptions are introduced, we provide a detailed comparison between partitioning estimators and other nonparametric estimators in Section 2.2.

The partitioning estimator, although simple and intuitive, has not received a thorough treatment in the econometrics or statistics literature. The available results typically concern mean-square rates for special cases (see, e.g., Kohler et al., 2006 and references therein). The main goal of this paper is to provide a general asymptotic treatment of partitioning estimators. Our analysis yields the following new insights. First, employing simple and intuitive sufficient conditions, in most cases weaker than those in the existing literature, mean-square and uniform convergence rates of the partitioning estimator are established and shown to be optimal. More generally, the uniform convergence rate explicitly

☆ We thank Yves Atchade, Xiaohong Chen, Richard Crump, Guido Imbens, Lutz Kilian, Jim Powell, seminar participants at Berkeley (2008) and Michigan (2009), and conference participants at the 2010 Advances in Econometrics Conference for comments. We are especially grateful to Victor Chernozhukov and Michael Jansson for detailed discussions and suggestions. We also thank the co-Editor, Peter Robinson, an Associate Editor, and four reviewers for detailed comments that greatly improved the paper. An earlier version of this paper was circulated under the title "Optimal Rates of Convergence and Asymptotic Normality of Block Regression Estimators". The authors gratefully acknowledge financial support from the National Science Foundation (SES 0921505 and SES 1122994).

* Correspondence to: Department of Economics, University of Michigan, 238 Lorch Hall, 611 Tappan Street, Ann Arbor, MI 48109-1220, United States. Tel.: +1 734 763 1306; fax: +1 734 764 2769.

E-mail address: cattaneo@umich.edu (M.D. Cattaneo).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.02.002

highlights a natural trade-off between moment assumptions and rate restrictions. Second, we characterize the leading terms of a conditional integrated mean-square error expansion and provide an optimal plug-in selector for the tuning parameter. Third, we derive a uniform Bahadur-type representation of linear functionals of the partitioning estimator, which is used to establish asymptotic normality under simple and intuitive conditions, with a suitable standard-error estimator. We cover both regular and irregular estimands. The applicability of the new results is illustrated with three examples: (i) derivative of the regression function at a point, (ii) partial and full means, and (iii) weighted average derivatives. Our results are also useful in other contexts in econometrics, as discussed in Section 1.1.

The paper proceeds as follows. In the remainder of this section we give the main motivations for our work, discussing in particular the importance of our results for both empirical and theoretical econometrics. Section 2 describes the partitioning estimator formally and also provides a comparison to other nonparametric estimators. Rates of convergence and a general integrated mean-square error expansion for the partitioning estimator are given in Section 3, while a Bahadur-type representation for linear functionals of the estimator and asymptotic normality with valid standard-error estimators are developed in Section 4. The results of a Monte Carlo study are summarized in Section 5. Finally, Section 6 concludes. Proofs are gathered in the Appendix. A supplement is available upon request containing detailed technical proofs and greatly expanded simulation results.

1.1. *Motivation and preliminary discussion*

Studying the large-sample properties of partitioning estimators may be interesting and important for a variety of reasons, some theoretical and others methodological. The partitioning estimator has specific features and asymptotic optimality properties that make it a useful addition to the econometrics toolkit: a complement, not a substitute, to the arsenal of nonparametric procedures commonly employed in econometrics. This estimator is attractive because it is very tractable and enjoys useful asymptotic representations leading to intuitive results, as well as other features that may be useful in econometric applications.

In particular, the partitioning estimator is potentially discontinuous in finite samples (just like nearest-neighbor estimators). This specific characteristic may be an advantage from a practical point of view, and could also lead to an estimator with desirable theoretical properties. The "binning" underlying the partitioning estimator arises naturally in many economic problems, where units (people, firms, etc.) in the same bin share similar economic behavior, and therefore partitioning-based inference procedures have been proposed to retain this natural interpretability (see applications below). From a theoretical perspective, we are interested in understanding the asymptotic properties of partitioning given its potential discontinuity in finite samples, and how they compare with results for other nonparametric procedures. We briefly discuss three implications of this discontinuity, which make the partitioning estimator theoretically and practically interesting in our view.

1. *Shape restrictions: convergence rates*. Nonparametric estimation typically assumes the estimand is smooth and most estimators are constructed imposing some of the underlying smoothness assumed. The partitioning estimator does not impose smoothness and therefore allows us to understand what effects imposing this shape restriction may have on asymptotic properties, which arguably is of theoretical interest. For instance, we establish optimal uniform convergence rates for partitioning, showing (by example) that imposing smoothness is not necessary for this result. This finding is not ex-ante obvious in our view, especially given other known results (see Section 2.2).
2. *Shape restrictions: bias–variance trade-off*. From a more practical perspective, removing the smoothness restriction may be interpreted as "freeing up restrictions". This means that the estimator will have a different bias–variance behavior in finite samples. To fix ideas, consider the linear partitioning and linear regression spline estimators of a univariate regression function. For each sample size, both are (piecewise linear) least squares fits, and differ only in that the spline is required to be continuous (see Section 2.2). That is, the linear spline is a restricted least squares problem compared to the partitioning estimate. From linear model results, it follows that the spline has larger bias than the partitioning estimate, but smaller variance.[1] Neither can be strictly superior or inferior, based on the usual bias–variance trade-off, and in fact the partitioning estimator may have better properties from a theoretical point of view.
3. *Diagnostics*. The potential discontinuity of the partitioning estimator in finite samples makes it a useful complement to existing smooth estimates already available in the literature. Specifically, the partitioning estimate may be used as a diagnostic check on the underlying smoothness assumptions imposed by other procedures, particularly if such assumptions are in question for a certain region of the support. Furthermore, the discontinuous partitioning estimate can be used to characterize the overall variability of the data relative to a smoothed-out estimate (see the regression discontinuity application below for an example).

Further motivation for our work stems from the role of partitioning estimators in empirical economics. Perhaps originating with the regressogram of Tukey (1947), partitioning-based procedures have been suggested in many contexts where "binning" has a natural interpretation, despite their formal properties being unknown in most cases. We close this section by briefly discussing four examples where partitioning estimation arises in econometrics: as an exploratory device, a nonparametric estimator, and two semiparametric cases.

Application (*Regression Discontinuity*)**.** Partitioning estimators are used heuristically in the regression discontinuity (RD) design for two purposes: (i) to plot a smoothed-out cloud of points along with global polynomial fits of the underlying regression function for control and treatment units, and (ii) to investigate whether the data suggests the presence of other possible discontinuities in the underlying conditional expectation of potential outcomes, as a form of falsification test. Imbens and Lemieux (2008) review the RD literature, and explicitly advocate partitioning (calling it a "histogram-type estimate") to assess the plausibility of the RD design. Our general result in Theorem 3 is employed in Calonico et al. (2012) to derive an optimal choice of partition length in this context, thereby providing a systematic way of plotting RD data. □

Application (*Portfolio Sorting*)**.** In understanding anomalous asset returns, a common approach is "portfolio sorts", in which assets are partitioned into homogeneous groups according to characteristics that may drive anomalies. A number of informal and formal analyses are then performed on the sorted assets, including tests of monotonicity and comparison of extremes. See, e.g., Fama and French (2008). Our results may be used to develop formal nonparametric inference for this type of application. □

[1] This claim assumes that the estimators are misspecified in finite samples, as is the case with nonparametric estimators in general. This remains true when comparing to kernel-based estimators.

Download English Version:

https://daneshyari.com/en/article/5096297

Download Persian Version:

https://daneshyari.com/article/5096297

Daneshyari.com