Econometrics and Statistics 000 (2017) 1–16

Contents lists available at ScienceDirect

Econometrics and Statistics

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

Semiparametric estimation under shape constraints

Ximing Wu[a,*], Robin Sickles[b]

[a] Department of Agricultural Economics, Texas A&M University, College Station, TX 77845, United States
[b] Department of Economics, Rice University, Houston, TX 77005, United States

ARTICLE INFO

Article history:
Received 26 May 2016
Revised 9 June 2017
Accepted 10 June 2017
Available online xxx

Keywords:
Monotonicity
Concavity
Shape constraints
Semiparametric estimation
Penalized splines
Lorenz curve
Production functions

ABSTRACT

Substantial structure and restrictions, such as monotonicity and curvature constraints, necessary to give economic interpretation to empirical findings are often furnished by economic theories. Although such restrictions may be imposed in certain parametric empirical settings in a relatively straightforward fashion, incorporating such restrictions in semiparametric models is often problematic. A solution to this problem is provided via penalized splines, where monotonicity and curvature constraints are maintained through integral transformations of spline basis expansions. Large sample properties, implementation and inferential procedures are presented. Extension to multiple regressions under the framework of additive models is also discussed. A series of Monte Carlo simulations illustrate the finite sample properties of the estimator. The proposed method is employed to estimate a Lorenz curve of income and a production function with multiple inputs.

1. Introduction

Economic theories can provide useful guidance on the modeling of real world data. Utility functions associated with rational preferences are monotone; furthermore, convex preference implies quasi-concave utility functions. Demand functions of normal goods are downward sloping (Matzkin, 1991; Lewbel, 2010; Blundell et al., 2012). According to duality theory, profit functions are concave in output price and cost functions are monotonically increasing and concave in input price. Convex function estimation is also used extensively in derivative asset pricing models (Broadie et al., 2000; Aï t-Sahalia and Duarte, 2003; Yachew and Härdle, 2006). Researchers, when trying to model economic relationships, often face at least two challenges. One is fidelity to economic theory. Another is flexibility in functional forms (Guilkey et al., 1983; Diewert and Wales, 1987). In addition, these two goals are often at odds: conformity to theories often dictates relatively rigid functional forms, while flexible parameterizations sometimes lead to implausible predictions.

One fruitful approach to tackle this dilemma is to use nonparametric or semiparametric methods subject to the restrictions suggested by economic theory. This is a well-developed literature and has had a number of contributors. Matzkin (1994); Yachew (2003) provide general reviews of this literature. For relatively recent developments, see Hall and Huang (2001); Groeneboom et al. (2001); Horowitz and Mammen (2004); Carroll et al. (2011); Shively et al. (2011); Blundell et al. (2012); Pya and Wood (2015), among others. We follow in this line of research and present a flexible semiparametric estimator with shape constraints. We focus on functional relationships with two shape constraints: monotonicity and concavity (convexity) as this is a class of functions that are frequently modeled in applied economic studies. Functional relationships with either one of these two constraints are special cases of our estimator.

* Corresponding author.
E-mail addresses: xwu@tamu.edu (X. Wu), rsickles@rice.edu (R. Sickles).

http://dx.doi.org/10.1016/j.ecosta.2017.06.001

We base our work on (Ramsay, 1988) monotone smooth estimator and utilize integral transformations defined by some differential equations to impose shape restrictions. A key advantage of this transformation approach is that it transforms a constrained problem into an unconstrained one. We subsequently model the unconstrained problem using penalized spline methods, resulting in a nonlinear semiparametric estimator. We show that careful choice of the transformation and of the model-based penalty can simplify estimation considerably.

We propose an iterative algorithm to calculate the proposed estimator. We establish the consistency of the estimator and present approximate methods for inference and for selecting the smoothing parameter. We then extend our estimator to an additive model. We illustrate the finite sample performance and usefulness of our methods with Monte Carlo simulations and two empirical applications.

The remainder of the paper is organized as follows. Section 2 briefly reviews the relevant literature and then presents our transformation-based model to accommodate shape restrictions. Section 3 proposes a Gauss–Jordan algorithm to solve the estimator and Section 4 discusses methods of model specification. Section 5 extends the model to multiple regressions.Sections 6 and 7 report Monte Carlo simulations and two empirical examples. The last section concludes. The large sample properties of the proposed estimator and technical proofs are relegated to Appendix.

2. Model and estimator

Several approaches have been used to impose restrictions in statistical and econometric models. A simple approach is the transformation of variables. For instance, the logarithmic transformation is commonly used to assure positiveness of predicted outcomes and the Box–Cox transformation can offer an even more flexible alternative. In the estimation of production functions, the Cobb–Douglas, constant elasticity of substitution (CES), translog, and generalized Leontief specifications are commonly employed. These functional forms are often chosen because they satisfy certain theoretical properties and also due to their simplicity, as they are either linear in parameters after a simple log transformation or are linear to begin with. Simple parametric forms, however, can sometimes entail nontrivial restrictions. For example, a logarithm transformation of the dependent variable implies multiplicative errors rather than the usual additive ones.

To avoid rigid functional forms, semiparametric and nonparametric methods have been used to accommodate shape restrictions. An ear ly example is (HD, 1955) isotonic estimator, which essentially produces a monotone step function. Mukerjee (1988) and Mammen (1991) developed kernel-based isotonic regression techniques which consist of a kernel smoothing step and an isotonization step to maintain monotonicity. Wu et al. (2015) studied penalized isotonic regression. Instead of isotonization, Hall and Huang (2001) suggested a penalized kernel method to obtain monotonicity. Their method is employed by Henderson et al. (2012); Blundell et al. (2012); Ma and Racine (2013) for various applications or further generalizations. Another popular family of smoothers, the spline-based methods, has been used by Ramsay (1998); Kelly and Rice (1990); Mammen and Thomas-Agnan (1999); Meyer (2008, 2012); Wang et al. (2013), who explored monotone estimators based on shape preserving spline basis functions. Pya and Wood (2015) propose a family of shape constrained additive models. Eilers (2005) and Bollaerts et al. (2006) considered the approach of asymmetric penalties wherein a penalty is activated if a constraint is violated. The technique of rearrangement or data sharpening Braun and Hall (2001) and Chernozhukov et al. (2009) can also be used. Shively et al. (2009) consider a Bayesian approach for nonparametric monotone function estimation of Gaussian regressions, which is generalized to log-concave likelihood functions by Shively et al. (2011). See also Groeneboom et al. (2001) for a theoretical analysis of convex function estimation using least squares and maximum likelihood methods.

Our proposed estimator is inspired by the smooth monotone estimator of Ramsay (1988). Suppose $y = f(x)$ is a smooth monotone function of x. For simplicity, we assume that $x \in [0, 1]$. Ramsay (1988) proposed to model a strictly monotone function via the following integral transformation:

$$f(x) = \int_0^x \exp(r(s))ds, \tag{1}$$

where r is a square integrable function on [0, 1]. Since $f'(x) = \exp(r(x)) > 0$ for all x, the monotone restriction is satisfied. Unlike some penalty-based monotone estimators that impose observation-specific monotonicity, (1) is globally monotone thanks to the positive exponential functional embedded in the integral transformation. We note that this transformation approach was employed by Nelson and Siegel (1987) for parametric modeling of monotonic yield curves.

Since $f''(x) = f'(x)r'(x)$ and $f'(x) > 0$, $f(x)$ is concave if $r'(x) \le 0$ for all x. Our strategy is to use the integration transformation (1) as well to further impose the condition that $r'(x) \le 0$. In particular, we consider the following parameterization

$$f(x) = \int_0^x \exp\left(-\int_0^s g(t)dt\right)ds. \tag{2}$$

It follows that $f'(x) = \exp(-\int_0^x g(t)dt) > 0$ and $f''(x) = -f'(x)g(x)$, implying that $f''(\cdot) \le 0$ if $g(\cdot) \ge 0$. Thus under (2), the monotonicity and concavity constraints are reduced to a simple non-negativity constraint that $g(x) \ge 0$ for all x. Natural candidates of g include $g(x) = x^2$ and $g(x) = \exp(x)$; other choices are certainly possible. Below we will show that $g(x) = x^2$ is particularly appealing for the proposed method on theoretical and practical grounds.