Imposing Economic Constraints in Nonparametric Estimation

Daniel J. Henderson
Associate Professor
Department of Economics
State University of New York at Binghamton
http://bingweb.binghamton.edu/~djhender

December, 2009
1 Outline

1. “Imposing Economic Constraints in Nonparametric Regression: Survey, Implementation and Extension,” with Christopher Parmeter

2. “Regression and Inference Under Smoothness Conditions,” with Subal Kumbhakar, Christopher Parmeter and Kai Sun

3. “Smooth Coefficient Estimation of a Seemingly Unrelated Regression with Cross-Equation Restrictions,” with Subal Kumbhakar, Qi Li and Christopher Parmeter

2  Imposing economic constraints in nonparametric regression

2.1  Introduction

1. Concavity, homogeneity, monotonicity, etc. are all common economic assumptions or consequences of economic assumptions

2. Wide variety of constrained nonparametric estimation strategies have been proposed

3. Existing estimators either deal with a single constraint, lack smoothness or lack the ability to handle categorical data
2.2 Nonparametric regression

\[ y_i = m(x_i) + u_i \]

\[ \hat{m}(x) = \frac{\sum_{i=1}^{n} K \left( \frac{x_i - x}{h} \right) y_i}{\sum_{i=1}^{n} K \left( \frac{x_i - x}{h} \right)} = \frac{1}{n} \sum_{i=1}^{n} w_i(x) y_i \]

\[ w_i(x) = \frac{n K \left( \frac{x_i - x}{h} \right)}{\sum_{i=1}^{n} K \left( \frac{x_i - x}{h} \right)} \]
2.3 Hall and Huang (2001)

Weights can be manipulated so that the estimator satisfies monotonicity

\[ \hat{m}(x \mid p) = \sum_{i=1}^{n} p_i w_i(x) y_i \]

Choose \( \{p_1, p_2, \ldots, p_n\} \) to minimize a distance metric subject to the constraint that \( \partial \hat{m}(x \mid p) / \partial x \geq 0 \)

We also impose the regularity conditions \( p_i \geq 0 \) and \( \sum_{i=1}^{n} p_i = 1 \)

\[ D_\rho(p) = \frac{1}{\rho(1-\rho)} \left[ n - \sum_{i=1}^{n} (np_i)^\rho \right] \]
2.4 Example: concavity of a production function

Concavity of the production function implies diminishing marginal productivity of each input

Concavity implies that the Hessian matrix

$$H = \begin{bmatrix}
\frac{\partial \hat{m}(x)^2}{\partial^2 x_1} & \frac{\partial \hat{m}(x)^2}{\partial x_2 \partial x_1} & \cdots & \frac{\partial \hat{m}(x)^2}{\partial x_1 \partial x_k} \\
\frac{\partial \hat{m}(x)^2}{\partial x_1 \partial x_2} & \frac{\partial \hat{m}(x)^2}{\partial^2 x_2} & \cdots & \frac{\partial \hat{m}(x)^2}{\partial x_2 \partial x_k} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial \hat{m}(x)^2}{\partial x_1 \partial x_k} & \frac{\partial \hat{m}(x)^2}{\partial x_2 \partial x_k} & \cdots & \frac{\partial \hat{m}(x)^2}{\partial^2 x_k}
\end{bmatrix}$$

must be negative semi-definite over the domain of $x$
2.4.1 Problem in practice

1. Does the concavity assumption hold in practice?

2. Construct a data generating process which is concave with an additive random error

3. \( y = \sqrt{x} + u \) gives a production function which should be concave

4. However, simulations show that the nonparametric estimator can have positive second derivatives
1. This is not a special case

2. For a variety of data generating processes, concavity does not always hold

3. The occurrence of concavity cannot be insured with increases in the bandwidth (without losing shape)

4. The occurrence of concavity increases with the sample size
2.4.2 Simulations

Consider the two following data generating processes which should produce a concave production function

1. \( y = \sqrt{x} + u \)

2. \( y = \ln(x) + u \)

Where \( x \sim U \left[0, 1\right] \) and \( u \sim N \left(0, 0.1\right) \)

We drew \( n = 100 \) observations \( B = 100 \) times and take the 95th percentile of \( D(p) \) to check for departures from concavity
2.4.3 Empirical application

Age-earnings profile
   Heckman and Polachek (1974)
   Pagan and Ullah (1999)

Data
   1971 Canadian Census Public Use Tapes
   205 males with 13 years of education
3 Imposing monotonicity nonparametrically in first price auctions

3.1 Introduction

1. Nonparametric methods relax functional form assumptions of an unknown model

2. Typical nonparametric estimation does not employ or impose restrictions often given by economic theory

3. Without monotonicity one cannot adduce the equilibrium bidding strategy associated with the corresponding value distribution
3.2 Theoretical background

Your bid \((b)\) comes from maximizing

\[
\max_b \pi (b) = (v - b) F' (\beta (b)^{-1})^{n-1}
\]

The first-order condition is

\[
(n - 1) (v - b) F' (\beta (b)^{-1})^{n-1} f (\beta (b)^{-1}) \frac{\partial \beta (b)^{-1}}{\partial b} - F (\beta (b)^{-1})^{n-1} = 0
\]

Monotonicity of the bid function allows one to use \(\frac{\partial \beta (b)^{-1}}{\partial b} = 1 / \beta' (v)\), and by noting that \(\beta (v) = b\)

\[
\beta (v) = v - \frac{\int F (u)^{n-1} du}{F (v)^{n-1}}
\]
3.3 Guerre, Perrigne and Vuong (2000)

Manipulation of the differential equation leads to

\[ v_i = \xi(b_i, G, n) \equiv b_i + \frac{1}{n-1} \frac{G(b_i)}{g(b_i)} \]

The equation now expresses the individual private value \( v_i \) as a function of the individuals equilibrium bid, its distribution, its density and the number of bidders

Interest lies in both the values and the density of the values
3.3.1 Four step approach of GPV

1. Estimate \( g(b) \) using kernel methods

\[
\hat{g}(b) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{b_i - b}{h}\right)
\]

2. Estimate \( G(b) \) using the empirical CDF

\[
\hat{G}(b) = \frac{1}{n} \sum_{i=1}^{n} 1(b_i \leq b)
\]

3. Construct \( \hat{v}_i = \hat{\xi}(b_i, G, n) \) using the above estimates

4. Estimate the density and distribution of values \( f(\hat{v}_i) \) and \( F(\hat{v}_i) \)
3.3.2 Problem in practice

1. Does the monotonicity assumption hold?

2. Draw values from a density to construct fake bids and see if monotonicity is present

3. The log-normal distribution provides a theoretically monotonic equilibrium bidding strategy

4. However, simulations show that it can indeed become non-monotonic
1. This is not a special case

2. For a variety of (monotonic) distributions, monotonicity does not always hold in practice

3. The occurrence of monotonicity *does* increase with the bandwidth

4. The occurrence of monotonicity increases with the sample size
3.4 Monotone estimation of the bid function

1. Estimate $g(b)$ using kernel methods

$$
\hat{g}(b \mid p) = \frac{1}{h} \sum_{i=1}^{n} p_i K \left( \frac{b_i - b}{h} \right)
$$

2. Estimate $G(b)$ as

$$
\hat{G}(b \mid p) = \frac{1}{n} \int_{-\infty}^{b} \hat{g}(u \mid p) \, du
$$

3. Construct $\hat{\nu}_i = \hat{\xi}(b_i, G, n)$ using the above estimates

4. Estimate the density and distribution of values $f(\hat{\nu}_i)$ and $F(\hat{\nu}_i)$
Selecting weights

We choose \( \{p_1, p_2, \ldots, p_n\} \) to minimize a distance metric subject to the constraint that \( \frac{\partial \hat{\xi}(b, G, n)}{\partial b} \geq 0 \)

We also impose the regularity conditions \( p_i \geq 0 \) and \( \sum_{i=1}^{n} p_i = 1 \)

Our distance metric is the power divergence measure

\[
D_{\rho}(p) = \frac{1}{\rho(1-\rho)} \left[ n - \sum_{i=1}^{n} (np_i)^{\rho} \right]
\]
3.5 Simulations

100 Draws from a Log-Normal and 100 draws from a Gamma, each repeated 100 times. We do this for $\rho = 0, 0.5$ and $1.0$ to check the sensitivity of our measure to departures from monotonicity.
3.6 Experimental data

Dyer, Kagel and Levin (1989) data

They created the values as draws from the uniform density

They do experiments with 3 and 6 bidders
3.7 Extensions

Bandwidth selection

Reserve prices

Heterogenous auctions
4 Conclusions

1. Relatively simple and fast way to impose multiple economic constraints with nonparametric estimation

2. These methods can be extended to other areas of econometrics (e.g., parametric regression, symmetric density estimation)

3. Empirical results show that the methods work well in relatively small samples