Semiparametric Single Index Models

Yasuyuki Matsumura

Graduate School of Economics, Kyoto University

November 6, 2024

Introduction

- A semiparametric single index model is given by

$$Y = g(X^T \beta_0) + u,$$

where

$$\begin{aligned}
&Y \in \mathbb{R} : \text{a dependent variable,} \\
&X \in \mathbb{R}^q : \text{a } q \times 1 \text{ explanatory vector,} \\
&\beta_0 \in \mathbb{R}^q : \text{a } q \times 1 \text{ vector of unknown parameters,} \\
&u \in \mathbb{R} : \text{an error term which satisfies } \mathbb{E}(u \mid X) = 0, \\
&g(\cdot) : \text{an unknown distribution function.}
\end{aligned}$$

Introduction

- Even though x is a $q \times 1$ vector, $x^T \beta_0$ is a scalar of a single linear combination, which is called a single index.
- By the form of the single index model, we obtain

$$\mathbb{E}(Y \mid X) = g(X^T \beta_0),$$

  which means that the conditional expectation of Y only depends on the vector X through a single index $X^T \beta_0$.
- The model is SEMIPARAMETRIC when $\beta \in \mathbb{R}^q$ is estimated with the parametric methods and $g(\cdot)$ with the nonparametric methods.
- Some of the PARAMETRIC single index models are really familiar with us.

Examples of Parametric Single Index Model

- If $g(\cdot)$ is the identity function, then the model turns out to be a linear regression model:

$$Y = g(X^T \beta_0) + u = X^T \beta_0 + u.$$

- If $g(\cdot)$ is the CDF of Normal$(0, 1)$, then the model turns out to be a probit model.
  - See the textbook for further discussions on a probit model.
- If $g(\cdot)$ is the CDF of logistic distribution, then the model turns out to be a logistic regression model.

Identification Conditions

Proposition 8.1 (Identification of a Single Index Model)

For the semiparametric single index model $Y = g(x^T\beta_0) + u$, identification of β_0 and $g(\cdot)$ requires that

- (i) x should not contain a constant/intercept, and must contain at least one continuous variable. Moreover, $\|\beta_0\|=1$.
- (ii) $g(\cdot)$ is differentiable and is not a constant function on the support of $x^T\beta_0$.
- (iii) For the discrete components of x, varying the values of the discrete variables will not divide the support of $x^T\beta_0$ into disjoint subsets.

Identification Condition (i)

- Note that the location and the scale of β_0 are not identified.
- The vector x cannot include an intercept because the function $g(\cdot)$ (which is to be estimated in nonparametric manners) includes any location and level shift.
  - That is, β_0 cannot contain a location parameter.
- Some normalization criterion (scale restrictions) for β_0 are needed.
  - One approach is to set $\|\beta_0\|$.
  - The second approach is to set one component of β_0 to equal one. This approach requires that the variable corresponding to the component set to equal one is continuously distributed and has a non-zero coefficient.
  - Then, x must be dimension 2 or larger. If x is one-dimensional, then $\beta_0 \in \mathbb{R}^1$ is simply normalized to 1, and the model is the one-dimensional nonparametric regression $E(Y \mid x) = g(x)$ with no semiparametric component.

Identification Conditions (ii) and (iii)

- The function $g(\cdot)$ cannot be a constant function and must be differentiable on the support of $x^T\beta_0$.
- x must contain at least one continuously distributed variable and this continuous variable must have non-zero coefficient.
  - If not, $x^T\beta_0$ only takes a discrete set of values and it would be impossible to identify a continuous function $g(\cdot)$ on this discrete support.

Ichimura (1993)'s Method: Semiparametric Least Squares

- Suppose that the functional form of $g(\cdot)$ were known.
- Then we could estimate β_0 by minimizing the least-squares criterion:

$$\sum_{i=1} \left[Y_i - g(X_i^T \beta)\right]^2$$

with respect to β.

- We could think about replacing $g(\cdot)$ with a nonparametric estimator $\hat{g}(\cdot)$.
- However, since $g(z)$ is the conditional mean of Y_i given $X_i^T \beta_0 = z$, $g(\cdot)$ depends on unknown β_0, so we cannot estimate $g(\cdot)$ here.

Ichimura (1993)'s Method: Semiparametric Least Squares

- Nevertheless, for a fixed value of β, we can estimate

$$G(X_i^T\beta) := \mathbb{E}(Y_i \mid X_i^T\beta) = \mathbb{E}(g(X_i^T\beta) \mid X_i^T\beta).$$

- In general $G(X_i^T\beta) \neq g(X_i^T\beta)$.
- When $\beta = \beta_0$ [1], it holds that $G(X_i^T\beta_0) = g(X_i^T\beta_0)$.

[1] Recall that β_0 is the true value of β.

Ichimura (1993)'s Method: Semiparametric Least Squares

- First, we estimate $G(X_i^T\beta)$ with the leave-one-out NW estimator:

$$\hat{G}_{-i}(X_i^T\beta) := \hat{\mathbb{E}}_{-i}(Y_i \mid X_i^T\beta)$$
$$= \frac{\sum_{j\neq i} Y_j K\left(\frac{X_j^T\beta - X_i^T\beta}{h}\right)}{\sum_{j\neq i} K\left(\frac{X_j^T\beta - X_i^T\beta}{h}\right)}.$$

Ichimura (1993)'s Method: Semiparametric Least Squares

- Second, using the leave-one-out NW estimator $\hat{G}_{-i}(X_i^T\beta)$, we estimate β with

$$\hat{\beta} := \arg\min_{\beta} \sum_{i=1}^{n} \left[Y_i - \hat{G}_{-i}(X_i^T\beta)\right]^2 w(X_i)\mathbf{1}(X_i \in A_n) := S_n(\beta).$$

- $w(X_i)$ is a nonnegative weight function.
- $\mathbf{1}(X_i \in A_n)$ is a trimming function to trim out small values of $\frac{1}{nh}\sum_{j\neq i} K\left(\frac{X_j^T\beta - X_i^T\beta}{h}\right)$, so that do not suffer the random dominator problem.

Asymptotic Distribution of Ichimura (1993)'s Estimator

References (1)

- Li, Q. and J. S. Racine, (2007). *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.
- 末石直也 (2024) 『データ駆動型回帰分析：計量経済学と機械学習の融合』日本評論社.
- 西山慶彦，人見光太郎 (2023) 『ノン・セミパラメトリック統計解析 (理論統計学教程：数理統計の枠組み)』共立出版.

References (2)

Useful references also include some lecture notes of the following topic courses:

- ECON 718 NonParametric Econometrics (Bruce Hansen, Spring 2009, University of Wisconsin-Madison),
- セミノンパラメトリック計量分析（末石直也，2014 年度後期，京都大学）.