Estimation of the Static Discrete Choice Model

Kodai Yasuda

Graduate School of Economics, Hitotsubashi University

Created: December, 9, 2023
Last Modified: May 1, 2024

Table of Contents

Introduction

- Using the structural approach, we often face the problem called **curse of dimensionality**.
- The curse of dimensionality is the problem that the estimation takes too much time as the number of parameters increases.
- Hence we need to transform the structure of functions, and use some methods to get the solution efficiently.

Table of Contents

Review: Model Setting

- Now we consider the following random coefficient model.

Random Coefficient Model

$$U_{ij} = \begin{cases} -\alpha p_j + \mathbf{X}_j' \boldsymbol{\beta}_i + \xi_j + \epsilon_{ij} & j = 1, \cdots, J \\ \boldsymbol{\epsilon}_{i0} & j = 0 \end{cases}$$

where

- p_j: price of product j
- $\mathbf{X}_j$: vector of product j's characteristics
- ξ_j: unobserved demand shifter of product j
- ϵ_{ij}: idiosyncratic error of consumer i's choice of product j. ϵ_{ij} is assumed to be i.i.d. extreme value type I.

- For simplicity, we can transform the utility function in the case of $j = 1, \cdots, J$ as

$$U_{ij} = \delta_j + \mu_{ij} + \epsilon_{ij}$$

Note that δ_j depends on $\alpha, \bar{\beta}$ and μ_{ij} depends on $\{\sigma_k\}_{k=1}^K$ (see appendix).

Review: Model Setting (cont.)

- Then, the choice probability of consumer i is given by

$$Pr(d_i = j \mid \{\nu_{ik}\}_{k=1,\cdots,K}) = \frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_{l=1}^{J} \exp(\delta_l + \mu_{il})}. \tag{1}$$

- In total, the market share of product j is given by

$$s_j = \underbrace{\int \cdots \int}_{\nu} \prod_{l \neq j} \left[\frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_{l=1}^{J} \exp(\delta_l + \mu_{il})} \right] dF_{\nu_1}(\nu_{i1}) \cdots dF_{\nu_K}(\nu_{iK}).$$

- However, this integral is intractable, so we use some approximation methods to estimate the parameters.

Berry Inversion

- In simpler models, we can get the average utility from the market share.
- For example, in the multinomial logit model, we consider the following utility function.

$$U_{ij} = \begin{cases} \delta_j + \epsilon_{ij} & j = 1, \cdots, J \\ \epsilon_{i0} & j = 0. \end{cases}$$

where δ_j is the deterministic term.

- Then, since the market share is identical to the choice probability, we get the market share s_j as

$$s_j = \begin{cases} \frac{\exp(\delta_j)}{1+\sum_{l=1}^{J} \exp(\delta_l)} & j = 1, \cdots, J \\ \frac{1}{1+\sum_{l=1}^{J} \exp(\delta_l)} & j = 0. \end{cases}.$$

Berry Inversion (cont.)

- Hence we can derive the average utility δ_j from the market share s_j as

$$\ln\left(\frac{s_j}{s_0}\right) = \delta_j.$$

- This inversion is called **Berry Inversion**.
- In the case $\delta_j = -\alpha p_j + \mathbf{X}_j'\boldsymbol{\beta} + \xi_j$, ξ_j is treated as the error term.
- Berry inversion enables us to solve the above equation for ξ_j.

$\rightarrow$ We can estimate the parameters $\alpha, \boldsymbol{\beta}$ by using GMM.

Simulated Market Share of RCL Model

- However, in the case of the RCL model, it is difficult to derive the average utility from the market share analytically;

$$s_j \stackrel{\text{def}}{=} S_j(\boldsymbol{\delta} \mid \boldsymbol{\nu}) \quad \forall j.$$

  where $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_J)$, $\boldsymbol{\nu} = (\boldsymbol{\nu}_1, \cdots, \boldsymbol{\nu}_N)$, and $\boldsymbol{\nu}_i = (\nu_{i1}, \cdots, \nu_{iK})$.
- Then, we use the simulated market share instead of the actual market share.
  - First, we draw $\{\nu_{ik}\}_{k=1}^K$ from the distribution $\{F_{\nu_k}\}_{k=1}^K$ for N times.
  - Then, substituting them into the equation (1) yields the choice probability of each consumer.
  - Finally, we calculate the market share by taking the average of the choice probability over consumers;

$$\hat{S}_j(\boldsymbol{\delta} \mid \boldsymbol{\nu}) = \frac{1}{N} \sum_{i=1}^{N} \frac{\exp(\delta_j + \hat{\mu}_{ij}(\boldsymbol{\nu}_i))}{1 + \sum_{l=1}^{J} \exp(\delta_l + \hat{\mu}_{il}(\boldsymbol{\nu}_i))}$$

- Note that we have $\hat{S}_j(\boldsymbol{\delta} \mid \boldsymbol{\nu}) \to S_j(\boldsymbol{\delta} \mid \boldsymbol{\nu})$ by LLN.

Table of Contents

Introduction to the BLP method

- To estimate the parameters, some methods have been proposed.
- In this section, we introduce the most popular method: **the BLP method**.
- The general procedure of the BLP method is as follows:
  - By using the contraction mapping, get the average utility $\hat{\boldsymbol{\delta}}^*$.
  - From the average utility $\hat{\boldsymbol{\delta}}^*$, we get the vector of the error term $\boldsymbol{\xi} = \{\xi_j\}_{j=1}^J$.
  - Get the optimal $\hat{\boldsymbol{\theta}}$ which minimize the GMM objective function.

Logic of the BLP method

- Given $\boldsymbol{\sigma} = \{\sigma_k\}_{k=1}^{K}$, calculate the simulated market share $\hat{S}_j(\boldsymbol{\delta} \mid \boldsymbol{\nu})$.
- Get the averaged utility $\hat{\boldsymbol{\delta}}^* = \{\delta_j\}_{j=1}^{J}$ such that the simulated market share is close to the actual market share.
  - Use the following contraction mapping:

$$\hat{\delta}_j^{(t+1)} = \hat{\delta}_j^{(t)} + \ln s_j - \ln \hat{S}_j(\boldsymbol{\delta}^{(t)} \mid \boldsymbol{\nu}).$$

    Note that
    - If $s_j \geq \hat{S}_j(\boldsymbol{\delta}^{(t)} \mid \boldsymbol{\nu})$, we have $\hat{\delta}_j^{(t+1)} \geq \hat{\delta}_j^{(t)}$.
    - If $s_j < \hat{S}_j(\boldsymbol{\delta}^{(t)} \mid \boldsymbol{\nu})$, we have $\hat{\delta}_j^{(t+1)} < \hat{\delta}_j^{(t)}$.
  - Stop the iteration when the difference between $\hat{S}_j(\boldsymbol{\delta}^{(t)} \mid \boldsymbol{\nu})$ and s_j is sufficiently small, say 10^{-6}.
- Derive $\boldsymbol{\xi} = \{\xi_j\}_{j=1}^{J}$ by solving $\xi_j = \alpha p_j - \mathbf{X}_j' \boldsymbol{\beta}_i + \hat{\delta}_j^*$. $\boldsymbol{\xi}(\alpha, \bar{\boldsymbol{\beta}})$ depends on $\boldsymbol{\sigma}$.

Logic of the BLP method (cont.)

- Minimize the following GMM objective function:

$$Q(\boldsymbol{\theta}) = \mathbf{m}(\boldsymbol{\xi}(\boldsymbol{\theta}))'\mathbf{W}\mathbf{m}(\boldsymbol{\xi}(\boldsymbol{\theta}))$$

  where
  - $\boldsymbol{\theta} = \{\underbrace{\alpha, \bar{\boldsymbol{\beta}}}_{\tilde{\boldsymbol{\theta}}}, \boldsymbol{\sigma}\}$
  - $\mathbf{m}(\boldsymbol{\xi}) = \frac{1}{TJ}\sum_{t=1}^{T}\sum_{j=1}^{J}\xi_{tj}(\boldsymbol{\theta}) \cdot z_{tj}^{d}$: the sample analog of the moment condition
  - $\mathbf{W}$: weighting matrix
  - $(z_{tj}^{1}, \cdots, z_{tj}^{D})_{t=1,\cdots,T,j=1,\cdots,J}$: instrumental variables.
- Get the estimated parameters $\hat{\boldsymbol{\theta}}$ such that

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}).$$

Recap

- In the RCL model, we cannot derive the average utility from the market share analytically.
- Hence, we derive it numerically by using the BLP method.
- The BLP method estimates the parameter which minimizes the objective GMM function by using the contraction mapping.
- The optimization of the GMM function consists of two steps in the typical setting of the BLP method:
  1. By using $\boldsymbol{\xi}(\alpha, \bar{\boldsymbol{\beta}})$, get the parameter $\tilde{\boldsymbol{\theta}} = (\alpha, \bar{\boldsymbol{\beta}})$ which minimizes the GMM function (denotes $\hat{\tilde{\boldsymbol{\theta}}}$).
  2. Use $\tilde{\boldsymbol{\theta}}$ for the contraction mapping
     $$\boldsymbol{\xi}^{(t+1)} = \boldsymbol{\xi}^{(t)} + \ln \mathbf{s} - \ln \mathbf{S}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\sigma}^{(t)}),$$
     then we get the fixed point $\boldsymbol{\xi}^*$. Substituting it into the GMM function, we get the optimal $\hat{\boldsymbol{\sigma}}$ as
     $$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma}} Q(\tilde{\boldsymbol{\theta}}, \boldsymbol{\sigma}).$$
     Repeat two steps until both GMM functions are sufficiently small.
- However, there are some advanced algorithms to solve the optimization problem more efficiently. For example, see Dubé, Fox, and Su 2012.

Table of Contents

Endogeneity Problem

- In the discrete choice model, we must consider the endogeneity of price.
  - The price p_j may be correlated with the unobserved demand shock ξ_j. i.e., $Cov(p_j, \xi_j) \neq 0$.
- There are some ways to solve this problem.
  - Control variables
    - We can control the fixed effects of products by introducing some control variables.
    - For example, we can add time-invariant variables such as the brand dummy.
    - However, we cannot control the effects from the time-varying variables.
  - Instrumental variables

Instrumental Variables

- Instrumental variables are the most common way to solve the endogeneity problem.
- We introduce some variables z_j which are correlated with the price but uncorrelated with the unobserved demand shock. i.e.,
  - $E[\xi_j \mid z_j] = 0$ and $E[p_j \mid z_j] \neq 0$.
- The following instruments are proposed as the popular instruments.
  - BLP instruments
  - Supply side instruments
  - Hausman-Nevo instruments
  - Arellano-Bond instruments

BLP Instruments

- BLP instruments are the instruments proposed by Berry, Levinsohn, and Pakes 1995.
- BLP proposed the following instruments:
  - x_j^k: the k-th characteristic of product j
  - $\sum_{j' \neq j, j' \in \mathcal{J}_f} x_{j'}^k$: the sum of the k-th characteristic of products in the **same firm** as product j (without product j itself)
  - $\sum_{j' \neq j, j' \notin \mathcal{J}_f} x_{j'}^k$: the sum of the k-th characteristic of products in the **other firms** (without product j itself)
- As more advanced instruments, Gandhi and Houde 2019 proposed the **differentiated Instrumental Variables**.

Extension to Supply Side

- Firms obviously take the key role in the price setting.
- Hence, the marginal cost of product is strongly associated with its price.
- However, there are some problems in the estimation of the supply side.
  - It is difficult to get the data of the marginal cost.
  - The cost shifters are often common to many products in a market.
    - 上武 et al. 2021 pointed out that the cost shifters in the car market are often common across most of firms; e.g., the cost shifters of steel.

References

Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium". In: *Econometrica (1986-1998)* 63.4 (July 1995), p. 841.

Dubé, Jean-Pierre, Jeremy T Fox, and Che-Lin Su (2012). "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation". In: *Econometrica* 80.5, pp. 2231–2267.

Gandhi, Amit and Jean-François Houde (2019). "Measuring Substitution Patterns in Differentiated-Products Industries". In: *IDEAS Working Paper Series from RePEc.*

上武, 康亮, 祐太 遠山, 直樹 若森, and 安虎 渡辺 (2021). "需要を制する者はプライシングを制す". In: vol. 2. 実証ビジネス・エコノミクス. 日本評論社.

Transformation of the utility function of the RCL model

$$
\begin{aligned}
U_{ij} &= -\alpha p_j + \mathbf{X}_j' \boldsymbol{\beta}_i + \xi_j + \epsilon_{ij} \\
&= -\alpha p_j + \sum_{k=1}^{K} x_{jk}(\bar{\beta}_k + \sigma_k \nu_{ik}) + \xi_j + \epsilon_{ij} \\
&= \underbrace{-\alpha p_j + \sum_{k=1}^{K} x_{jk}\bar{\beta}_k + \xi_j}_{\delta_j} + \underbrace{\sum_{k=1}^{K} x_{jk}\sigma_k \nu_{ik}}_{\mu_{ij}} + \epsilon_{ij} \\
&= \delta_j + \mu_{ij} + \epsilon_{ij}.
\end{aligned}
$$

where

- ν_{ik}: unobserved idiosyncratic demand shock for consumer i, characteristic k