構造推定入門: 補足資料

応用社会科学RAブートキャンプ

講師: 遠山祐太 (早稲田大学)

最終更新: 2023-08-30 02:04:15

イントロダクション

補足資料

- 本スライドでは、トピックス・レクチャーの補足資料を掲載する。
- 内容的に、学部上級ー大学院レベルのものも含むため、必ずしも理解する必要はない。
- 上級の内容に興味がある受講生向け。

パラメタの識別

離散選択モデルにおけるパラメタの識別

- 識別: モデルパラメタがデータから一意に特定できること。
- 離散選択モデルにおいて選択確率は**効用の差**のみに依存している。

$$P(d_i = j) = \Pr\left(\{\epsilon_{ij}\}_{j=1}^{J} : V_{ij} + \epsilon_{ij} \geq V_{ik} + \epsilon_{ik} \; \forall k \neq j\right)$$

- 効用の定式化において2種類の**正規化(normalization)**が必要となる。
  - 1: location : 定数を足しても同じ (i.e., $U_{ij} + \alpha$ for all j)
  - 2: scale : U_{ij} について定数倍しても同じ (i.e., αU_{ij} for all j for some $\alpha > 0$)

Location Normalization

- 新しい定式化として $\tilde{U}_{ij} = U_{ij} + \alpha$ を考える。
- すると、

$$\begin{aligned} Prob(\tilde{U}_{ij} > \tilde{U}_{ik}, \forall k) &= Prob(U_{ij} + \alpha > U_{ik} + \alpha, \forall k) \\ &= Prob(U_{ij} > U_{ik}, \forall k) \end{aligned}$$

- 含意: ある選択肢から得られる効用の水準を正規化する必要がある。
- 実証分析では**アウトサイドオプション($j = 0$)**の効用をゼロと置くことが多い。

Scale Normalization

- 新しい定式化として $\tilde{U}_{ij} = \lambda U_{ij}$ を考える。ただし $\lambda > 0$.
- すると

$$\begin{aligned} Prob(\tilde{U}_{ij} > \tilde{U}_{ik}, \forall k) = & \quad Prob(\lambda U_{ij} > \lambda U_{ik}, \forall k) \\ = & \quad Prob(U_{ij} > U_{ik}, \forall k) \end{aligned}$$

- 含意: 選考ショックの分散について基準化する必要がある。
  - ロジットモデルでは $Var(\epsilon_{ij}) = \pi^2/6$ としている。
- 別の例：二項プロビットモデル $U_{i1} = \beta x_i + \epsilon_{i1}, U_{i0} = 0$ with $\epsilon_{i1} \sim N(0, \sigma^2)$.

$$Prob(d_i = 1) = \Phi\left(\frac{-\beta x_i}{\sigma}\right)$$

σ と β は別々に識別されない。

ランダム係数ロジットモデルの詳細

ランダム係数ロジットモデル

- 前のモデルにおいて、消費者間の異質性は選考ショック $\epsilon_{i,j}$ のみに反映。
- 消費者の異質性の入れ方
  - 観察される属性(社会経済属性) を入れる。
  - 製品属性の選好に関して、観察されない異質性を考える -> **ランダム係数**
- 定式化

$$u_{ij} = \beta_i X_j + \epsilon_{ij}$$

  - ランダム係数 $\beta_i \sim f(\beta)$.
  - X_j: 製品属性のベクトル
  - ϵ_{ij} は以前と同様。

選択確率

- ランダム係数 β_i を所与とすると、選択確率は

$$Pr(d_i = j|\beta_i) = \frac{\exp\left(\beta_i X_j\right)}{\sum_{j=1}^{J} \exp\left(\beta_i X_j\right)}$$

- しかし分析者は β_i を観察できない。従って、 β_i について平均をとったものを考える。

$$Pr(d_i = j) = \int \frac{\exp\left(\beta_i X_j\right)}{\sum_{j=1}^{J} \exp\left(\beta_i X_j\right)} f(\beta_i) d\beta_i$$

推定における問題点

- 単純化のために一変数を考える。ランダム係数の分布を以下と考える。$\beta \sim N(\mu, \sigma^2)$.
- 目標：ランダム係数の分布パラメタ (μ, σ^2).
- 個人の選択確率は

$$Pr(d_i = j) = \int \underbrace{\frac{\exp\big((\mu + \sigma\nu_i)X_{ij}\big)}{\sum_{j=1}^{J} \exp\big((\mu + \sigma\nu_i)X_{ij}\big)}}_{Pr(d_i=j|\nu_i,\theta)} dG(\nu_i)$$

ここで $G(\nu_i)$ は標準正規分布 -> $(\mu + \sigma\nu_i) \sim N(\mu, \sigma^2)$

- 問題点：積分があるため、この選択確率を解析的な形で求めることができない。

解決策：シミュレーションによる近似

- 選択確率における積分をモンテカルロ積分する。
- Step 1: 乱数 ν を $G(\nu)$ から生成する。 ν^r とラベル付する。.
- Step 2: 乱数 ν^r に基づいてロジット確率を計算する。

$$Pr(d_i = j|\nu^r) = \frac{\exp\big((\mu + \sigma\nu^r)X_j\big)}{\sum_{j=1}^{J} \exp\big((\mu + \sigma\nu^r)X_j\big)}$$

- Step 3: この手順を R 回繰り返し、その平均を取る。

$$\hat{Pr}(d_i = j|\theta) = \frac{1}{R}\sum_{r=1}^{R} Pr(d_i = j|\nu^r, \theta)$$

Simulated Maximum Likelihood Estimator

- シミュレーションした対数尤度の最適化を行う。

$$SLL(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log \hat{Pr}(d_i = j)$$

- Simulated MLEと呼ばれる。詳細はTrain Chater 10 を参照。
- シミュレーションの方法として、モンテカルロ積分以外にもいくつか方法がある。
  - Importance sampling
  - Halton draw
  - Gaussian quadrature

きのこ・たけのこ事例の尤度関数

- 参考：Train Chapter 6.7 "Panel Data"
- パラメタ θ をもつ個人 i が、選択 y_{ik} を行う確率を $P_{i,k}\left(j = y_{i,k}|\theta\right)$ とする。
- 各個人は 5 回選択を行うが、その選択を通じて選好パラメタ θ は共通とする。
- パラメタ θ をもつ個人 i がデータで観察される一連の選択 $\{y_{i,k}\}_{k=1}^{5}$ を行う確率は

$$L_i(\theta|\{y_{i,k}\}_{k=1}^{5}) = \prod_{k=1}^{5} P_{i,k}\left(j = y_{i,k}|\theta\right)$$

尤度関数続き

- パラメタ自体は観察されないので、分布で積分を取る必要がある。

$$P_i(\Omega) = \int L_i(\theta | \{y_{i,k}\}_{k=1}^{5}) f(\theta | \Omega) d\theta$$

  - ここで $f(\theta|\Omega)$ はパラメタの分布。
  - Ω は分布のパラメタ。 $\Omega = (\theta_{Kino}, \sigma_{Kino}^2, \theta_{Take}, \sigma_{Take}^2, \theta_\beta, \sigma_\beta^2)$
  - モンテカルロ積分などで近似を行う必要がある。

- 尤度関数は

$$L(\Omega) = \prod_{i=1}^{N} P_i(\Omega)$$

いわゆる「構造推定 VS 誘導形」について

いわゆる「誘導系」対「構造推定」の議論

- 「誘導系」アプローチ：ミクロ実証研究における主流
  - データの良いVariationを見つける （自然実験、操作変数法、回帰不連続デザイン）
  - 回帰分析フレームワークで、因果効果(トリートメント効果)を推定
  - 統計学でのCausal inference methodに依拠
  - 自然実験アプローチとも呼ばれる。
- 「誘導系」・「構造推定」アプローチは1990年代半ばくらいから広がっていた。
- それと同時に、これら手法の間の対立も発生。

対立の概略 (個人的見解を多分に含む)

- 「誘導系」から「構造推定」への批判
  - 経済モデルを明示的に定式化している。仮定がとても強い。
  - モデルパラメタの推定を、どのようにやっているか不明瞭
  - シミュレーションによる結果がブラックボックス。
- 「構造推定」から「誘導系」への批判
  - 背後のモデルがないので、推定している因果効果パラメタの解釈が不明瞭
  - 均衡効果を考慮できない (いわゆるSUTVAの仮定)
  - 「自然実験」がある「理想的」なデータの状況の分析ばかりになるのではないか？

個人的見解：これからの時代は両方必要！！

- 「誘導系」と「構造推定」は代替的ではなく、**補完的**
- 「誘導系」でも、単に因果効果を推定するだけではなく、**背後のメカニズム**の説明が必要
    - 経済モデルの重要性
- 「構造推定」では「自然実験」をうまく活用してパラメタを推定する。
    - RCTと構造推定を合わせる研究も増えている。
    - 例１：Todd and Wolpin (2006)
    - 例２：Ito, Ida, and Tanaka (2016)
    - 例３：Kawaguchi, Uetake, and Watanabe (2021)

手法論争に関する関連文献

- "SYMPOSIUM: CON OUT OF ECONOMICS" appeared in Spring 2010 volume of Journal of Economic Perspectives
  - Angrist and Pischke による、いわゆる構造推定アプローチ（特にIOやマクロ）への批判。
  - それに対する反論 (Keane, Sims, Nevo, Whinston)
- "Structural vs. atheoretic approaches to econometrics"
  - Michael Keane (労働経済学者) による、「自然実験アプローチ」への批判
- "Structural vs. Reduced Form:" Language and Models in Empirical Economics"
  - Phil Haile (IO)による論考。
  - そもそも「誘導系 VS 構造推定」という呼び方や分け方が不適切ではないか？という議論。
  - 多少読みづらいかもしれないが、**個人的に強くオススメ**