補足資料 2：入れ子型ロジットモデル、ランダム係数ロジットモデルの推定に関する詳細

最終更新： 2021 年 7 月 14 日

経済セミナー連載「実証ビジネス・エコノミクス」
第 3 回「プライシングの真髄は代替性にあり：
消費者需要モデルの推定［基礎編 2]」
(2021 年 8・9 月号) の付録

1 はじめに

補足資料 2 では本誌で取り上げた入れ子型ロジットモデルとランダム係数ロジットモデルの詳細について説明する。なお、ノーテーションはすべて本誌の用法に従う。

2 赤バス・青バス問題に関する補論

ここでは本誌でも取り上げた赤バス・青バス問題について、少し技術的な観点から説明する。やや数理的な説明になるものの、この点を踏まえると、入れ子型ロジットモデルやランダム係数ロジットモデルにおいて赤バス・青バス問題がなぜ緩和されるか (すなわちなぜ代替性の問題が緩和されるか) を理解しやすくなるであろう。

まず、通勤のための選択肢の集合として、自動車 (car)、赤バス ($redbus$) という 2 つを考える。さらに仮定として、平均効用 $\delta_{car} = \delta_{redbus}$ とする。その結果、ロジットモデルにおいて、両選択肢の選択確率 (マーケットシェア) は 50% となる。

ここで、新たな 3 つ目の選択肢として「青バス ($bluebus$) 」を導入しよう。ここで青バスは、赤バスの一部を青ペンキで塗り替えたものと考えればよい。したがって、青バスと赤バスの平均効用も

$\delta_{bluebus} = \delta_{redbus}$ となる。この状況で、人々の通勤選択はどのように変化するであろうか。青バスが赤バスで塗り替えられたものであることを踏まえるのであれば、「妥当な予測」は

$$s_{bluebus} = s_{redbus} = 0.25, \quad s_{car} = 0.5$$

になるであろう。しかしながら、この状況を多項ロジットモデルで分析すると、

$$s_{bluebus} = s_{redbus} = s_{car} = \frac{1}{3}$$

という予測が得られてしまうのである。

この原因となっているのが、個人・選択肢特有のショック ϵ_{ijt} である。青バスが導入されたあとにおける選択問題は以下となる。

$$\max \left\{ \begin{array}{c} \delta_{bluebus} + \epsilon_{i,bluebus} \\ \delta_{redbus} + \epsilon_{i,redbus} \\ \delta_{car} + \epsilon_{i,car} \end{array} \right\}.$$

いま、平均効用は 3 つの選択肢で共通である。しかしながら、ロジットモデルにおいては個人・選択肢特有のショック ϵ_{ijt} が独立なショックであるとしているため、このショック項によって「赤バス」と「青バス」が「差別化された選択肢」とされてしまっているのである。言い換えると、もし青バスが赤バスの一部を塗り替えたものであるならば、$\epsilon_{i,redbus}$ と $\epsilon_{i,bluebus}$ は強く相関しているべきであろう。

本誌で導入した入れ子型ロジットモデルとランダム係数ロジットモデルはこの問題を解決するべく、個人・選択肢特有のショックに関して相関を許すような構造を導入している。入れ子型ロジットモデルにおいては当該ショックが $\zeta_{igt} + (1-\sigma)\epsilon_{ijt}$ という形で入っており、これは同じグループ内における製品への選好ショックは ζ_{igt} を通じて相関が強くなるのである。一方、ランダム係数ロジットモデルでは当該ショックは $\mu_{ijt} + \epsilon_{ijt}$ である。ここで μ_{ijt} は消費者の製品属性への選好の異質性に起因する項である。仮に燃費を強く気にする消費者がいると、その消費者は燃費性能が高い自動車を強く好むようになり、これは μ_{ijt} に反映される。結果、燃費性能が高い自動車車種同士の選好ショック $\mu_{ijt} + \epsilon_{ijt}$ は相関が強くなるのである。

3 入れ子型ロジットモデルの導出の詳細

ここでは、入れ子型ロジットモデルにおける導出の詳細について説明する。より厳密な導出方法については、Berry (1994) および Cardell (1997) を参照されたい。

3.1　選択確率

本誌と同様に、グループが全体で $G+1$ あるとし、グループのインデックスを $g=0,1,\cdots,G$ とする。なお、アウトサイドグッズ ($j=0$) は、グループ 0 における唯一の製品とする。このとき、消費者 i の効用を以下のように書く。

$$
\begin{aligned}
u_{ijt} &= \beta_0 + \sum_{k=1}^{K} \beta^k x_{jt}^k - \alpha p_j + \xi_{jt} + \zeta_{igt} + (1-\sigma)\epsilon_{ijt}, \\
&= \delta_{jt} + \zeta_{igt} + (1-\sigma)\epsilon_{ijt}.
\end{aligned}
$$

では、このモデルに基づいて消費者の購買確率を導出しよう。以下ではマーケットのインデックス t を落として議論する。まず、製品をグループ g から購入することを条件付けたとき、ある製品 $j \in \mathcal{G}_g$ を購入する確率は以下となる。

$$
\Pr(choose\ j \mid group\ g) = \frac{\exp(\frac{\delta_j}{1-\sigma})}{\sum_{k\in\mathcal{G}_g} \exp(\frac{\delta_k}{1-\sigma})}.
$$

ここで、グループ g を選んだ場合の**包括的価値** (inclusive value) を以下のように定義する。

$$
I_g = \log\left[\sum_{k\in\mathcal{G}_g} \exp\left(\frac{\delta_k}{1-\sigma}\right)\right].
$$

なお、この包括的価値は、グループ g の選択肢から得られる期待効用として解釈することができる。

この包括的価値を利用することで、グループ g を選択する確率を以下のように求めることができる。

$$
\Pr(group\ g) = \frac{\exp((1-\sigma)I_g)}{\sum_{g=0}^{G} \exp((1-\sigma)I_g)} = \frac{D_g^{(1-\sigma)}}{\sum_{g=0}^{G} D_g^{(1-\sigma)}}.
$$

ここで、 $D_g = \sum_{k\in\mathcal{G}_g} \exp(\frac{\delta_k}{1-\sigma})$ となる。

これらをまとめることで、製品 j を購入する確率は

$$
\begin{aligned}
\Pr(choose\ j) &= Pr(group\ g) \times Pr(choose\ j \mid group\ g) \\
&= \frac{D_g^{(1-\sigma)}}{\sum_{g=0}^{G} D_g^{(1-\sigma)}} \times \frac{\exp\left(\frac{\delta_j}{1-\sigma}\right)}{D_g}
\end{aligned}
$$

となる。なお、アウトサイドグッズ ($j=0$) を選択する確率は、

$$\Pr(outside\ good)=\frac{1}{\sum_k D_k^{(1-\sigma)}}$$

として与えられる。

3.2 推定式

続いて、推定式の導出について見ていこう。今、マーケットシェアは以下の形で与えられる。

$$\begin{aligned} s_{jt} &= \frac{\exp\left(\frac{\delta_{jt}}{1-\sigma}\right)}{D_{gt}^{\sigma}\left[\sum_{g=0}^{G} D_{gt}^{(1-\sigma)}\right]}, \\ s_{0t} &= \frac{1}{\sum_{g=0}^{G} D_{gt}^{(1-\sigma)}}. \end{aligned}$$

まず、上の 2 つの式からロジットモデルの場合と同様の変形を考えよう。

$$\ln(s_{jt})-\ln(s_{0t})=\frac{\delta_{jt}}{1-\sigma}-\sigma\ln(D_{gt})$$

さらに変形を進めるべく、次にインサイドシェア $s_{jt/g}$ に関する次の式に着目しよう。

$$\begin{aligned} s_{jt/g} &= \frac{\exp\left(\frac{\delta_{jt}}{1-\sigma}\right)}{D_{gt}} \\ \Longleftrightarrow \ln(s_{jt/g}) &= \frac{\delta_{jt}}{1-\sigma}-\ln(D_{gt}). \end{aligned}$$

これら 2 つの式を変形して D_{gt} を消すことで、ロジットモデルと同様の線形の式が得られる。

$$\begin{aligned} \ln(s_{jt})-\ln(s_{0t}) &= \delta_{jt}+\sigma\ln(s_{jt/g}), \\ &= \beta_0+\sum_{k=1}^{K}\beta^k x_{jt}^k-\alpha p_j+\sigma\ln(s_{jt/g})+\xi_{jt}. \end{aligned}$$

4 ランダム係数パラメーターの識別に関する補足

本誌で述べたように、ランダム係数ロジットモデルを推定する際には、ランダム係数の標準偏差パラメーターを推定するための追加的な操作変数が必要となる。言い換えると、「ランダム係数のパラメーターをどのようにデータから識別・推定するか？」を考える必要がある。

この識別の寄与するデータの変動の 1 つが、マーケット間の製品集合の差である。たとえば、マー

ケット A とマーケット B という 2 つの市場を考えよう。今、マーケット B ではある燃費性能が良いモデル α「のみ」が販売されていないものの、それ以外の条件は同じとしよう[*1]。このとき、マーケット B ではモデル α が販売されていないゆえに、そのモデルを買ったであろう人たちは他のモデルを購入することになる。この状況において、もし「モデル α と似たような燃費性能が良いモデル」の販売数がそうでないものよりも増えていたとすると、消費者は燃費性能に関する代替性を重視していることを示唆するであろう。これはランダム係数ロジットモデルにおいては、「燃費性能に関する選好の異質性が大きい (つまり標準偏差パラメターが大きい) 」という形で捉えられるのである。

このようなマーケット間での製品集合の差を捉えている操作変数が、BLP 操作変数や Gandhi and Houde (2020) の差別化操作変数となっている。これらの操作変数は、「マーケットにおいて、似通った製品がどの程度利用可能か？」という情報に基づいており、その情報がランダム係数パラメターの識別・推定に寄与するのである。より詳細な議論については Gandhi and Houde (2020) を参照されたい。

なお、このようなマーケット間の製品集合の差が本当に外生的であるかについては議論の必要性がある。価格付けのような比較的短期間の意思決定を考える際には問題ないかもしれないが、新製品の投入や参入・退出などの中長期的な状況を考えると、製品集合自体も内生的なものとなるであろう。これら参入・退出については、連載第 5 回以降で扱っていく予定である。

5 BLP アルゴリズムの詳細

5.1 復習：セットアップ

- 詳細は本誌記事を参照されたい。
- 市場 $t = 1, \cdots, T$ における消費者 i は、選択肢 $j \in \{0, 1, \cdots, J_t\}$ に直面している。
- 間接効用は以下のように与えられる。

$$u_{ijt} = -\alpha_i p_{jt} + \beta_i' x_{jt} + \xi_{jt} + \epsilon_{ijt}.$$

- 係数 (α_i, β_i) は以下のように与えられる。

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i,\ D_i \sim F(D),\ v_i \sim G(v).$$

- 線形パラメター (α, β) の数を K_1、非線形パラメターの個数を K_2 とする。

[*1] 以下、少し技術的な説明を補足しておこう。このような製品集合の差は、除外制約 (exclusion restriction) としても考えることができる。すなわち、「モデル α が存在しない」ということは他の製品から得られる効用に直接影響はしないものの、最終的な購入の意思決定においてはモデル α が存在しないゆえに他の製品の購入確率に影響するのである。

- 市場シェアは以下のように与えられる。

$$s_{jt} = \int\int \frac{\exp(\delta_{jt} + \mu_{ijt}(v_i, D_i))}{1 + \sum_{j=1}^{J_t} \exp(\delta_{jt} + \mu_{ijt}(v_i, D_i))} dF(D_{it})dG(v_i).$$

- ここで

$$\delta_{jt} = -\alpha p_{jt} + \beta' x_{jt} + \xi_{jt}.$$
$$\mu_{ijt}(v_i, D_i) = [\Pi D_i + \Sigma v_i]' \begin{bmatrix} p_{jt} \\ x_{jt} \end{bmatrix}.$$

- 今後の表記のために、積分の中の確率を $s_{ijt} = \frac{\exp(\delta_{jt}+\mu_{ijt})}{1+\sum_{j=1}^{J_t}\exp(\delta_{jt}+\mu_{ijt})}$ とする。
- また、この積分自体には解析的な解がないため、シミュレーションによるモンテカルロ積分を行う。

$$s_{jt} = \frac{1}{R}\sum_{r=1}^{R} \frac{\exp\left(\delta_{jt} + \mu_{ijt}(v^r, D^r)\right)}{1 + \sum_{j=1}^{J_t} \exp\left(\delta_{jt} + \mu_{ijt}(v^r, D^r)\right)}.$$

ここで R はシミュレーションにおいて生成した乱数の個数である。

- Berry インバージョンを行うことで、データにおけるマーケットシェアのベクトルから、平均効用のベクトルを導出することができる。この操作を

$$\delta_{jt} = S_{jt}^{-1}\left(\{s_{jt}\}_{j=1}^{J_t}; \theta_2\right)$$

とする。具体的には、縮小写像アルゴリズムを用いて計算する。詳細は本誌を参照されたい。

5.2 GMM 目的関数

- 表記の定義
  - K_1：線形パラメターの個数。
  - K_2：非線形パラメターの個数 (ランダム係数パラメター)。
  - $K = K_1 + K_2$.
  - N：サンプルサイズ。これは、すべての市場におけるすべての製品数を足し合わせたもの。$N = \sum_{t=1}^{T} J_t$.
  - L：外生変数を含む、操作変数の個数。$L \geq K$.
  - $\theta = (\theta_1, \theta_2)$ 、θ_1 は線形パラメター、 θ_2 は非線形パラメター。

- 非線形 GMM における目的関数は以下となる。

$$\begin{aligned} J(\theta) &= \left(\frac{1}{N}\xi(\theta)'Z\right) W \left(\frac{1}{N}Z'\xi(\theta)\right) \\ &= \left(\frac{1}{N}\sum_{j,t}\xi_{jt}(\theta)z_{jt}\right)' W \left(\frac{1}{N}\sum_{j,t}\xi_{jt}(\theta)z_{jt}\right) \end{aligned}$$

  - z_{jt} は操作変数をまとめたベクトル。 $(L \times 1)$ ベクトル。
  - $Z = (Z'_{11}, \cdots, Z'_{JT})$ は操作変数をまとめた $(N \times L)$ 行列。
  - W は $(L \times L)$ の荷重行列 (weighting matrix).
  - ここで誤差項 ξ_{jt} はパラメターの関数として計算することができ、

$$\xi_{jt}(\theta) = S_{jt}^{-1}\left(\{s_{jt}\}_{j=1}^{J_t}; \theta_2\right) - (-\alpha p_{jt} + \beta' x_{jt})$$

  となる。また、 $\xi(\theta) = (\xi_1(\theta), \cdots, \xi_N)'$ を $(N \times 1)$ ベクトルとしてまとめる。

5.3 GMM 推定

- GMM 推定では、目的関数 $J(\theta)$ を最小化するようなパラメター θ を求める。
- $J(\theta)$ は非線形の関数であるが、$\hat{\theta}_1 = (\alpha, \beta)$ は解析的に求めることが可能である。この点について見ていこう。
- まず、非線形パラメター θ_2 を所与とすると、Berry インバージョンから平均効用 δ_{jt} を $S_{jt}^{-1}\left(\{s_{jt}\}_{j=1}^{J_t}; \theta_2\right)$ という式で求めることができる。
- 平均効用 δ_{jt} は $\delta_{jt} = -\alpha p_{jt} + \beta' x_{jt} + \xi_{jt}$ という定式化であり、価格 p_{jt} と製品属性 x_{jt}、そして誤差項 ξ_{jt} について線形の式となっている。
- したがって、θ_2 を所与として Berry インバージョンから得られた平均効用 $\delta_{jt}(\theta_2)$ を「被説明変数」とみなし、式 $\delta_{jt} = -\alpha p_{jt} + \beta' x_{jt} + \xi_{jt}$ について外生変数を含んだ操作変数 z_{jt} を用いた線形 GMM を行うことで θ_1 が得られるのである。

$$\hat{\theta}_1 = (X'ZWZ'X)^{-1}X'ZWZ'\delta(\theta_2).$$

- この $\hat{\theta}_1$ を目的関数に代入することで、目的関数は非線形パラメター θ_2 のみの関数となる。
- θ_2 については解析的な解がないため、数値計算を行って求めることとなる。

5.3.1 ［Advanced］GMM 目的関数の Gradient に関して

- サンプルコードでは、非線形パラメターの個数が少ないため利用していないものの、一般に数値計算の際には、パラメターに関する勾配ベクトル (Gradient) がわかると、計算が早くかつ安定する。
- ここでは Gradient の導出を行おう。

$$\underbrace{\frac{\partial J(\theta_2)}{\partial \theta_2}}_{(K_2 \times 1)} = 2 \cdot \left(\underbrace{D\delta(\theta_2)}_{(N \times K_2)} \right)' ZWZ'\xi(\theta).$$

- ここで

$$\begin{aligned}
\underbrace{D\delta(\theta_2)}_{N \times K_2} &= \begin{pmatrix} \frac{\partial \delta_1}{\partial \theta_{21}} & \cdots & \frac{\partial \delta_1}{\partial \theta_{2K_2}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \delta_N}{\partial \theta_{21}} & \cdots & \frac{\partial \delta_N}{\partial \theta_{2K_2}} \end{pmatrix} \\
&= - \underbrace{\begin{pmatrix} \frac{\partial s_1}{\partial \delta_1} & \cdots & \frac{\partial s_1}{\partial \delta_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_N}{\partial \delta_1} & \cdots & \frac{\partial s_N}{\partial \delta_N} \end{pmatrix}^{-1}}_{(N \times N)} \underbrace{\begin{pmatrix} \frac{\partial s_1}{\partial \theta_{21}} & \cdots & \frac{\partial s_1}{\partial \theta_{2K_2}} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_N}{\partial \theta_{21}} & \cdots & \frac{\partial s_N}{\partial \theta_{2K_2}} \end{pmatrix}}_{(N \times K_2)}.
\end{aligned}$$

- ここの導出は陰関数定理を用いている。詳しくは Nevo (2000, Appendix)[*2]を参照されたい。
- ここで、行列の中身の微分について

$$\begin{aligned}
\frac{\partial s_{jt}}{\partial \delta_{jt}} &= \frac{1}{R} \sum_{i=1}^{R} \frac{\partial s_{rjt}}{\partial \delta_{jt}} = \frac{1}{R} \sum_{i=1}^{R} s_{ijt}(1 - s_{ijt}), \\
\frac{\partial s_{jt}}{\partial \delta_{mt}} &= \frac{1}{R} \sum_{i=1}^{R} \frac{\partial s_{rjt}}{\partial \delta_{mt}} = -\frac{1}{R} \sum_{i=1}^{R} s_{ijt} s_{imt}.
\end{aligned}$$

- また、

$$\frac{\partial s_{jt}}{\partial \sigma^k} = \frac{1}{R} \sum_{i=1}^{R} \frac{\partial s_{rjt}}{\partial \sigma^k} = \frac{1}{R} \sum_{i=1}^{R} s_{ijt} \left(x_{jt}^k v_i^k - \sum_{m=1}^{J} x_{mt}^k v_i^k s_{imt} \right).$$

[*2] Appendix to "Practitioner's Guide to Estimation of Random: Coefficients Logit Models of Demand" – Estimation: The Nitty-Gritty (`https://web.archive.org/web/20171116012208/http://faculty.wcas.northwestern.edu/~ane686/supplements/Ras_guide_appendix.pdf`).

5.4 漸近分散 (Asymptotic Variance)

- 非線形 GMM の漸近分散は以下のように与えられる。 詳細については、末石 (2015) や Hayashi (2000) などを参照されたい。

$$\sqrt{N}(\hat{\theta}-\theta) \xrightarrow{d} N\left(0,(G'WG)^{-1}G'W\Omega WG(G'WG)^{-1}\right).$$

- ここで、

$$\underbrace{\Omega}_{(L\times L)} = E[\xi_{jt}z_{jt}z'_{jt}\xi_{jt}].$$

そして

$$\begin{aligned}
\underbrace{G}_{(L\times K)} &= E\left[z_{jt}\frac{\partial\xi_{jt}(\theta)}{\partial\theta}\right] \\
&= E\left[z_{jt}\frac{\partial(\delta_{jt}(\theta_2)-x'_{jt}\theta_1)}{\partial\theta}\right] \\
&= \left(\underbrace{E\left[z_{jt}(-x'_{jt})\right]}_{(L\times K_1)}, \underbrace{E\left[z_{jt}\frac{\partial\delta_{jt}(\theta_2)}{\partial\theta}\right]}_{(L\times K_2)}\right).
\end{aligned}$$

- この結果から、推定パラメター $\hat{\theta}$ は漸近的に $\hat{\theta}\sim N(\theta,\Sigma/N)$ と近似できる。この結果に基づいてパラメタの漸近標準誤差を計算することができる。具体的には、漸近標準誤差は $\sqrt{diag(\Sigma)/N}$ となる。なお、$diag(\Sigma)$ は行列 Σ の対角要素を示す。

5.5 漸近分散の推定

- 漸近分散の推定は、期待値を対応する標本平均で置き換えることとなる。これを標本対応 (sample analogue) と呼ぶ。
- すなわち、

$$\underbrace{\hat{\Omega}}_{(L\times L)} = \frac{1}{N}\sum_{j,t}\hat{\xi}_{jt}z_{jt}z'_{jt}\hat{\xi}_{jt}.$$

  - ここで、$\hat{\xi}_{jt}$ は推定値に基づく残差であり、$\hat{\xi}_{jt}=\delta_{jt}(\hat{\theta}_2)-x'_{jt}\hat{\theta}_1$ である。
  - また、$\hat{\xi}=(\hat{\xi}_1,\cdots,\hat{\xi}_N)'$ は $(N\times 1)$ ベクトル。

- 同様にして、G の標本対応は

$$
\begin{aligned}
\hat{G} &= \left(\frac{1}{N} \sum_{j,t} z_{jt} \frac{\partial \delta_{jt}(\theta_2)}{\partial \theta},\ \frac{1}{N} \sum_{j,t} z_{jt}(-x'_{jt}) \right) \\
&= \frac{1}{N} Z' \left(\underbrace{-X}_{(N \times K_1)},\ \underbrace{D\delta(\theta_2)}_{(N \times K_2)} \right).
\end{aligned}
$$

参考文献

末石直也 (2015)『計量経済学：ミクロデータ分析へのいざない』日本評論社。

Hayashi, F (2000) *Econometrics*, Princeton University Prress.

Berry, S. (1994) "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2): 242–262.

Cardell, N. S. (1997) "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 13(2): 185–213.

Gandhi, A. and Houde, J.-F. (2020) "Measuring Substitution Patterns in Differentiated-Products Industries," NBER Working Paper, No.26375.