

# 第4章付録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：2025年12月3日

上武康亮・遠山祐太・若森直樹・渡辺安虎

「実証ビジネス・エコノミクス」

第4章「プライシングの真髄は代替性にあり」  
の付録

© 上武康亮・遠山祐太・若森直樹・渡辺安虎

## 1 はじめに

補足資料2では第4章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書籍の用法に従う。

## 2 赤バス・青バス問題に関する補論

ここでは書籍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  $\epsilon_{ijt}$  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。

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

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  $\epsilon_{ijt}$  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。

書籍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  $\zeta_{igt}$  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  $\mu_{ijt}$  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  $\mu_{ijt}$  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

書籍と同様に、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  $g = 0, 1, \dots, G$  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。

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

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してことで、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。

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

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

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

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く。今、マーケットシェアは以下の形で与えられる。

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

$$s_{0t} = \frac{1}{\sum_{g=0}^G D_{gt}^{(1-\sigma)}}.$$

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。

$$s_{jt/g} = \frac{\exp\left(\frac{\delta_{jt}}{1-\sigma}\right)}{D_{gt}}$$

$$\iff \ln(s_{jt/g}) = \frac{\delta_{jt}}{1-\sigma} - \ln(D_{gt}).$$

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ターの識別に関する補足

書籍でも述べたように、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ターを識別するために追加の操作変数が必要となる。言い換えれば、「ランダム係数に関するパラメターを、どのようにデータから識別し推定するのか」という点を明確にする必要がある。個票データが利用できる場合には、標準偏差パラメタや消費者属性に関するパラメタの識別方法について、本書第

2.10.3 項で説明したとおりである。しかし、利用可能なデータが市場レベルの集計データに限られる場合には、個人属性と個人の選択行動との結びつきを直接観察できない。その結果、消費者の選好の異質性を表すパラメタの識別・推定は、個票データがある場合と比べて難しくなることが知られている。

本付録では、ランダム係数に関するパラメタを推定する際の、実践的なアプローチをいくつか紹介する。議論をわかりやすくするため、以下では財の特徴がひとつ（スカラー  $x_{jt}$ ）のみである場合を考え、そのときの効用関数を次のように定式化する。

$$U_{ijt} = \beta_0 + (\beta_1 + \pi_1 y_{it} + \sigma_1 \nu_{it}) x_{jt} + \xi_{jt} + \epsilon_{ijt}.$$

ここで、 $y_{it}$  は観察される消費者属性であり、既知の分布  $y_{it} \sim F_t(\cdot)$  に従うとする。また、 $\nu_{it}$  は観察されない選好の異質性を表し、 $\nu_{it} \sim N(0, 1)$  と仮定する。その他の変数については書籍本文と同じである。このモデルでは、線形パラメタは  $\theta_1 = (\beta_0, \beta_1)$ 、非線形パラメタは  $\theta_2 = (\pi_1, \sigma_1)$  と整理できる。

この効用関数のもとで導出される需要関数について、Salanié and Wolak (2019) のアイデアに基づき、 $\pi_1 = \sigma_1 = 0$  の周りで二次のテイラー展開を行った次の近似式を考える。<sup>\*1</sup>

$$\log \frac{s_{jt}}{s_{0t}} \approx \beta_0 + \beta_1 x_{jt} + \sigma_1^2 a_{jt} + \pi_1 m_t^y x_{jt} + \pi_1^2 v_t^y a_{jt} + \xi_{jt} \quad (1)$$

ここで、

$$a_{jt} = \left( \frac{x_{jt}}{2} - \sum_{k \in J_t} s_{kt} x_{kt} \right) x_{jt},$$

$m_t^y$  は市場  $t$  における消費者属性  $y_{it}$  の平均、 $v_t^y$  はその分散を表す。この近似式を用いることで、非線形パラメタ  $\theta_2 = (\pi_1, \sigma_1)$  に関する識別を、線形回帰モデルに近い枠組みで議論することが可能となる。

まず、消費者属性の係数  $\pi_1$  の識別について考えよう。式 (1) から明らかなように、もし消費者属性の分布  $F_t(\cdot)$  が市場間で共通である場合、変数  $m_t^y x_{jt}$  と  $x_{jt}$  の間に共線性が生じることになる。その結果、係数  $\beta_1$  と  $\pi_1$  を別々に識別することができなくなる。言い換えれば、社会経済属性の分布が市場間で異なっていることが、係数  $\pi_1$  を識別・推定するうえで重要な役割を果たすことがわかる。

つづいて、 $\sigma_1$  の識別について考えよう。このパラメタは変数  $a_{jt}$  の係数として現れる。変数  $a_{jt}$  は

$$a_{jt} = \left( \frac{x_{jt}}{2} - \sum_{k \in J_t} s_{kt} x_{kt} \right) x_{jt}$$

と定義されており、これは製品  $j$  の品質  $x_{jt}$  が、他の製品の品質（の荷重平均）からどの程度離れている

---

<sup>\*1</sup> Salanié and Wolak (2019) は、この近似式を用いて推定を行う FRAC (Fast, Robust, Approximately Correct) 推定法を提案している。

るか、すなわち製品の差別化の程度を捉える指標となっている。

この点を踏まえると、もし製品集合  $J_t$  が市場間で大きく変わらず、どの市場でも似通った製品が販売されている場合には、変数  $a_{jt}$  と  $x_{jt}$  の間に共線性が生じてしまう。その結果、パラメタ  $\sigma_1$  の識別は困難となる。

この点をさらに直観的に説明しよう。標準偏差パラメタが大きいとは、特定の特徴を好む消費者が、その特徴を共有する別の製品も強く好むことを意味する。たとえば、エコカーを好む消費者が「プリウス」だけでなく「アクア」など類似したモデルを選択しやすいといった状況がその典型である。

このような選好の連動性を識別するには、市場間で製品集合が異なる状況があることが重要である。仮に、ある市場で特定の製品が販売されていないとすれば、その製品を本来選ぶはずであった消費者が、どの製品に代替するかを観察できる。このスイッチングのパターンは、どの製品同士に強い代替関係があるかを示すため、標準偏差パラメタの識別に直接結びつく。具体的には、ある市場で燃費性能の高いモデルが販売されておらず、その市場において燃費性能が類似した別のモデルの販売が相対的に増加しているとすれば、消費者が燃費性能という特性を基準に代替行動を行っていることがわかる。このような代替の強さは、選好の異質性の大きさ、すなわち標準偏差パラメタの大きさに反映される。

市場間での製品集合の違いは、より技術的には除外制約 (exclusion restriction) として理解できる。特定のモデルが市場に存在しないという事実は、他の製品の直接効用には影響を与えない。しかし、選択可能な製品集合が異なることによって、消費者の選択確率には間接的な影響が生じる。このようなモデルの不在に伴う代替行動が、どの製品間で効用差がどの程度縮まるかを反映し、それが標準偏差パラメタの識別に利用されるのである。<sup>\*2</sup>

これらの考え方は、BLP (1995) の操作変数や、Gandhi and Houde (2019) が提案した差別化操作変数にも明確に反映されている。これらの操作変数は、市場においてどの程度「似通った製品」が利用可能であるかという構造的な情報を利用し、それによって選好の異質性に関するパラメタを識別するという発想に基づいている。

具体的には、Gandhi and Houde (2019) の差別化操作変数  $Z_{jt}$  は以下のように与えられる。

$$Z_{jt} = \begin{cases} \sum_{j' \neq j} \left( d_{jt,j'}^k \right)^2, \\ \sum_{j' \neq j} d_{jt,j'}^k \cdot d_{jt,j'}^{k'}. \end{cases}$$

---

<sup>\*2</sup> ただし、市場間の製品集合の違いが常に外生的であるとは限らない点には注意を要する。短期的には、価格設定などと比較して製品集合が固定的であるとみなせる状況も多いが、中長期的には新製品の投入や市場からの退出など企業の戦略的行動が関与するため、製品集合自体が内生的に決まる可能性がある。このような内生性の問題をどのように扱うかについては、より発展的な議論が必要であり、これは参入・退出を扱う章で取り上げることとする。

ここで、

$$d_{jt,j'}^k = x_{jt}^k - x_{j't}^k$$

は製品属性  $k$  に関する、製品  $j$  と  $j'$  の属性差を表す。この操作変数は、先述した変数  $a_{jt}$  と構造的に類似している点にも注意されたい。

以上が、集計データのみを用いた場合におけるランダム係数パラメタの識別に関する議論である。実務的な応用においては、マイクロデータに基づく情報を追加的なモーメント（いわゆるマイクロモーメント）として利用することで、ランダム係数パラメタの推定精度を高めようとするアプローチもしばしば用いられる。このアイデアは Petrin (2002) によって導入され、その後 Conlon and Gortmaker (2025) が包括的な整理と議論を行っている。以下では、代表的なマイクロモーメントの例をいくつか取り上げる。<sup>\*3</sup>

1つ目の例は、Petrin (2002) で用いられた「特定の製品を購入した消費者グループの社会経済属性」である。Petrin (2002) はミニバンの需要分析を主題としており、「ミニバンを購入した消費者の所得や家計サイズ」といった情報をマイクロモーメントとして利用することで、これらの属性がミニバンに対する選好にどのような影響を及ぼすのかを推定している。

もう1つの例は、BLP (2004) が使用した「第一希望・第二希望の製品の間にみられる属性の相関」である。彼らは、BLP (1995) で用いた集計データに加えて、消費者サーベイを用いて「セカンドチョイス」（消費者が“2番目に欲しかった製品”）に関する情報を収集した。実際に購入された製品とセカンドチョイスの製品の属性の相関を調べることで、消費者が製品を代替する際にどの属性次元を重視しているのかを推測することが可能となる。

最後に、マイクロモーメントを利用しない追加的な推定アプローチを紹介する。1つ目は、GMM 推定を行う際に、供給側の価格付けに関する最適化条件を追加的なモーメントとして利用する方法である。オリジナルの BLP (1995) では、この供給側モーメントを用いて推定を行っている。2つ目は、最適操作変数 (Optimal Instrumental Variables) を構築する方法である。Chamberlain (1987) の条件付きモーメントの考え方に基づき、最適なモーメント条件を構築し、それを用いて推定するというアイデアである。ただし、BLP モデルは非線形であるため、最適モーメントを解析的に求めることは困難である。

---

<sup>\*3</sup>マイクロモーメントが用いられる典型的な状況は、個票データが存在するものの、そのデータから得られる情報が消費者の選択行動について部分的な内容に限られている場合である。たとえば、「ミニバンを購入した世帯の平均所得」のように、個々の消費者が具体的にどの製品を購入したかまでは観察できないケースがこれに該当する。これに対し、個人レベルでの購買行動（どの消費者がどの製品を購入したか）を詳細に観察できる場合には、第2章で扱ったように、マイクロデータに基づく尤度関数を構築した推定をメインに行うことができる。詳細な方法論については Goolsbee and Petrin (2004) や Grieco et al. (2025) を参照されたい。なお、第2章との重要な相違点として、価格の内生性をどのように考慮するかという問題が挙げられる。

り、実際には近似的な計算手法を用いる必要がある。詳細については、Reynaert and Verboven (2014) を参照されたい。<sup>\*4</sup>

## 5 BLP アルゴリズムの詳細

### 5.1 復習：セットアップ

- 詳細は書籍を参照されたい。
- 市場  $t = 1, \dots, T$  における消費者  $i$  は、選択肢  $j \in \{0, 1, \dots, J_t\}$  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  $(\alpha_i, \beta_i)$  は以下のように与えられる。

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

- 線形パラメター  $(\alpha, \beta)$  の数を  $K_1$ 、非線形パラメターの個数を  $K_2$  とする。
- 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

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

- 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。

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

---

<sup>\*4</sup> Conlon and Gortmaker (2020) では、モンテカルロ実験を通じて、供給側のモーメント条件を用いる点及び最適操作変数の重要性について議論している。

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。詳細は書籍を参照されたい。

## 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。 $L \geq K$ .
  - $\theta = (\theta_1, \theta_2)$  、  $\theta_1$  は線形パラメター、  $\theta_2$  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。
- $Z = (Z'_{11}, \dots, Z'_{JT})$  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  $\xi_{jt}$  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、 $\xi(\theta) = (\xi_1(\theta), \dots, \xi_N)'$  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  $\theta$  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く。
- まず、非線形パラメター  $\theta_2$  を所与とすると、Berry インバージョンから平均効用  $\delta_{jt}$  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  $\delta_{jt}$  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  $\xi_{jt}$  について線形の式となっている。
- したがって、 $\theta_2$  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  $\theta_1$  が得られるのである。

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

- 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  $\theta_2$  のみの関数となる。
- $\theta_2$  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 について

- 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。

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

- ここで

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

- この導出は陰関数定理を用いている。詳しくは Nevo (2000, Appendix)<sup>\*5</sup> 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、

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

## 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。

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

- ここで、

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

<sup>\*5</sup> 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.pdf](https://web.archive.org/web/20171116012208/http://faculty.wcas.northwestern.edu/~ane686/supplements/Ras_guide.pdf))

そして

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

- 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  $\sqrt{\text{diag}(\Sigma)/N}$  となる。なお、 $\text{diag}(\Sigma)$  は行列  $\Sigma$  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、 $\hat{\xi} = (\hat{\xi}_1, \dots, \hat{\xi}_N)'$  は  $(N \times 1)$  ベクトル。
- 同様にして、 $G$  の標本対応は

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

## 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 Press.
- Berry, S. (1994) "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2): 242–262.
- Berry, S., Levinsohn, J., and Pakes, A. (1995) "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4): 841–890.
- Berry, S., Levinsohn, J., and Pakes, A. (2004) "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112(1): 68–105.
- 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.
- Chamberlain, G. (1987) "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics*, 34(3): 305–334.
- Conlon, C. and Gortmaker, J. (2020) "Best Practices for Differentiated Products Demand Estimation with PyBLP," *RAND Journal of Economics*, 51(4): 1108–1161.
- Conlon, C. and Gortmaker, J. (2025) "Incorporating Micro Data into Differentiated Products Demand Estimation with PyBLP," *Journal of Econometrics*, 105926.
- Gandhi, A. and Houde, J.-F. (2019) "Measuring Substitution Patterns in Differentiated-Products Industries," NBER Working Paper, No. 26375.
- Goolsbee, A. and Petrin, A. (2004) "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 72(2): 351–381.
- Grieco, P., Murry, C., Pinkse, J., and Sagl, S. (2025) "Conformant and Efficient Estimation of Discrete Choice Demand Models," *Review of Economic Studies*, forthcoming.
- Nevo, A. (2000) "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9(4): 513–548.
- Petrin, A. (2002) "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4): 705–729.
- Reynaert, M. and Verboven, F. (2014) "Improving the Performance of Random Coefficients Demand

Models: The Role of Optimal Instruments,” *Journal of Econometrics*, 179(1): 83–98.

Salanié, B. and Wolak, F. (2019) “Fast, ‘Robust’, and Approximately Correct: Estimating Mixed Demand Systems,” NBER Working Paper, No. 25726.