2022.5.23
理論談話会 #9
10:00-11:00 @セミナーB

A decomposition method for estimating recursive logit based route choice models

Mai, Tien, Fabian Bastin, and Emma Frejinger.
EURO Journal on Transportation and Logistics 7.3 (2018): 253-275.

M1 増橋 佳菜

0. Abstract

既往研究

経路選択モデルの課題

課題 0 経路の選択肢集合が不明

課題 1 **経路同士の相関**

解決策 0

再帰的な定式化（マルコフ連鎖による記述）

→ 選択肢集合を列挙しない

Recursive Logit Model (Fosgerau; 2013)

課題 1 **経路同士の相関**を考慮できない（LS 以外に**IIA特性**が残る）**※不十分 !!!**

課題 2 多数の大規模な連立方程式を解く必要性 → **計算時間コストが高い**

→ **大規模NW・複雑なモデルに適用困難**

本論文の提案

課題 1 → 解決策 1 **複雑なモデル Mixed RLモデル の採用** → IIA特性が緩和

課題 2 → 解決策 2 **Decomposition method (DeC method)を提案** → 超高速化

（RLモデルで実証）

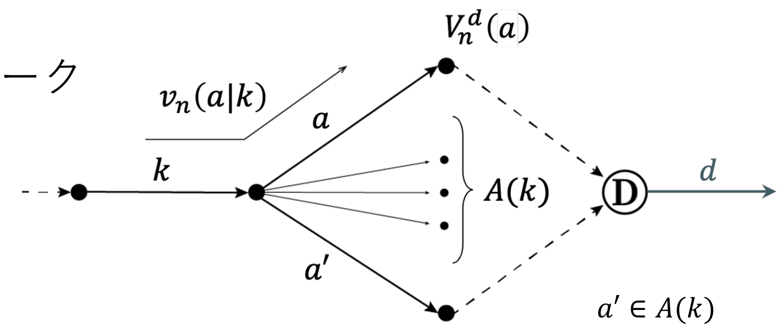
理論的に可能

+α Mixed RL モデルへの適用

実際に可能

2. Recursive Logit Model

- ネットワークの定義
  - $G=(A,\nu)$: 有向リンクで結ばれたネットワーク
  - A: リンク集合、 ν: ノードの集合



- 効用関数
  - 即時効用 $u_n(a|k)=v_n(a|k)+\mu\varepsilon_n(a)$
  - 期待効用 $V_n^d(a)$

- 価値関数 $V_n^d(k)$ **- Bellman方程式**の形
  - $V_n^d(k;\beta)=E\left[\max_{a\in A(k)}\left(v_n(a|k;\beta)+V_n^d(a;\beta)+\mu\varepsilon_n(a;\beta)\right)\right] \quad \forall k\in A$

※ 以下、個人nの表記を省略

- 選択確率$P^d(a|k)$

$$P^d(a|k)=\underbrace{\delta(a|k)}_{\text{接続行列}}\frac{\exp\left(\frac{1}{\mu}\left(v(a|k)+V^d(a)\right)\right)}{\sum_{a'\in A(k)}\exp\left(\frac{1}{\mu}\left(v(a'|k)+V^d(a')\right)\right)}$$

※ 導出過程は参考資料 [6]pp.16参照

全ての可能な経路に対して各ノードでMNL型の選択を行う

スタートアップゼミ2022 #2 参照

2. Recursive Logit Model

IIA特性 (Independence from Irrelevant Alternatives)

2つの選択肢を考えるとき「その2つの選択肢の確定項のみの影響を受け、
選択肢集合に含まれる他の選択肢の影響を受けない」性質。
無関係な選択肢からの選択確率の独立性を保証している 。

- 経路$\sigma = \{k_i\}_{i=0}^{I} = \{\overset{発生}{k_0}, k_1, \dots, \underset{=d}{\overset{吸収}{k_I}}\}$について
- 経路σの選択確率=経路σが観測されたときの尤度$P(\sigma)$

$$P(\sigma) = \prod_{i=0}^{I-1} P(k_{i+1}|k_i)$$

$$= \prod_{i=0}^{I-1} \frac{\exp(\frac{1}{\mu}(v(k_{i+1}|k_i) + V^d(k_{i+1})))}{\boxed{\sum_{a\in A(k)} \exp(\frac{1}{\mu}(v(a|k_i) + V^d(a)))} = e^{\frac{1}{\mu}V^d(k)}}$$

$$= \prod_{i=0}^{I-1} \exp\left((\frac{1}{\mu}\left(v(k_{i+1}|k_i) + V^d(k_{i+1}) - V^d(k_i)\right)\right)$$

$$= \exp\left(-\frac{1}{\mu}V^d(k_0)\right)\prod_{i=0}^{I-1} \exp\left(\frac{1}{\mu}v(k_{i+1}|k_i)\right) \quad (\because V^d(k_I) = 0)$$

$$= \frac{\exp\left(\frac{1}{\mu}v(\sigma)\right)}{\exp\left(\frac{1}{\mu}V^d(k_0)\right)} = \frac{\overbrace{\exp\left(\frac{1}{\mu}v(\sigma)\right)}^{経路\sigma の確定項}}{\underbrace{\sum_{\sigma\prime\in\Omega} \exp\left(\frac{1}{\mu}v(\sigma\prime)\right)}_{可能な経路の確定項和}}$$

（Ω：全経路の集合）

RLモデルは
“無限の選択肢を持つMNLモデル”

→ IIA特性を持つ
経路σ_1と経路σ_2の確率の比は
$v(\sigma_1) - v(\sigma_2)$のみに依存

課題 1 **経路同士の相関**を考慮できない

スタートアップゼミ2022 #2 参照

番外編 LS (リンクサイズ属性) とIIA特性の緩和

- **リンクの重複 → $\boxed{\text{課題 1}}$ 経路同士の相関**

- 既往研究
  - 共通性係数 (c-ligitモデル) (Cascetta et al.;1996)
  - PS属性 (PSLモデル) (Ben-Akiva and Bierlaire; 1999)

・与えられた経路j, jのリンク集合Γ_j, 選択肢集合C, リンクaの長さL_aに対して

$$PS_{jn} = \sum_{a \in \Gamma_j} \frac{L_a}{L_j} \frac{1}{\sum_{i \in C_n} \delta_{ai}}$$

← リンクaを通る経路数

・経路jがどのリンクとも重複していないと1に、重複度大で0に近づく

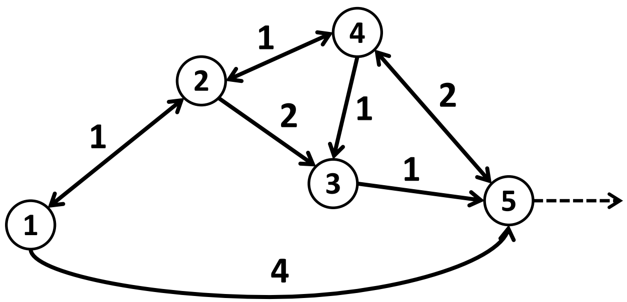
・経路の効用の補正は$\beta_{PS} \ln PS_{jn}$で表し、重複が全くない経路で$\beta_{PS} \ln PS_{jn} = 1$

問題 1：全経路選択肢の列挙が必要
問題 2：リンク加法的ではない

→ RLモデルには適用不可能

番外編 LS (リンクサイズ属性) とIIA特性の緩和

● **Fosgerauの提案：Link Size (LS)の導入**



ズレ
- 普通のMNL→ 全て$\frac{1}{4}$の選択確率
- 左右に半分くらい分かれて$\{1,5\}$に$\frac{1}{2}$くらい

→ ズレを補正するためにリンクの共有を考える

● リンクサイズ属性 **(LS)**

- 重複量の代理として期待リンクフロー$F(a)$を使用する

  cf. リンクaを通る経路数 @PS属性 (PSLモデル)

- 定義 $\boxed{\boldsymbol{LS}^{od} = \boldsymbol{F}^{od}(\tilde{\beta})}$

$\boldsymbol{LS}^{od}$：あるODに対する$(A\times 1)$のベクトル（k番目の要素はリンクkのLS属性）
$\tilde{\beta}$：効用の計算に使用した説明変数のパラメータβの推定値は$\hat{\beta}$とは区別する
$\boldsymbol{F}^{od}(\tilde{\beta})$はスタートアップゼミ2022 #2 資料のp.39の$\left(\boldsymbol{I} - P^d(\tilde{\beta})^T\right)\boldsymbol{F}^{od}(\tilde{\beta}) = \boldsymbol{G}^o$から算出
$\boldsymbol{G}^o$：オリジナルリンクが1、他が0の値を持つ発生フローベクトル
$\boldsymbol{G}^o\boldsymbol{1} = 1$より、$\boldsymbol{F}^{od}$は正規化されたとも

課題 1 **経路同士の相関**を考慮できない（LS 以外に**IIA特性**が残る）

2. Recursive Logit Model

Bellman方程式を解く

$$V_n^d(k) = E\left[\max_{a \in A(k)}\left(v_n(a|k) + V_n^d(a) + \mu\varepsilon_n(a)\right)\right] \quad \forall k \in A$$

※ 以下、個人nの添字を省略し、 $V_n^d(k) = V^d(k)$ とする

- 対数をとる

$$e^{\frac{1}{\mu}V^d(k)} = \begin{cases} \sum_{a \in A} \delta(a|k)\, e^{\frac{1}{\mu}\left(v(a|k)+V^d(a)\right)} & \forall k \in A \\ 1 & k = d \end{cases}$$

- $\boldsymbol{M^d}$ $(|\tilde{A}|\times|\tilde{A}|)$なる接続行列を次のように定義

$$M_{ka} = \delta(a|k)e^{\frac{1}{\mu}v(a|k)} = \begin{cases} e^{\frac{1}{\mu}v(a|k)} & if\ a \in A(k) \\ 0 & k = d \end{cases}$$

($\because d$はダミーリンクなので後続なし)

- $\boldsymbol{b^d}$ $(|\tilde{A}|\times 1)$ と $\boldsymbol{z^d}(|\tilde{A}|\times 1)$ を定義

$${b_k}^d = \begin{cases} 0 & k \neq d \\ 1 & k = d \end{cases} \qquad {z_k}^d = e^{\frac{1}{\mu}V(k)}$$

- 期待効用 z は連立一次方程式 $\boxed{\boldsymbol{z}^d = \boldsymbol{M}^d\ \boldsymbol{z}^d + \boldsymbol{b}^d}$ の解

目的地dごとに1つの線形システム

課題 2 多数の大規模な連立方程式を解く必要性 → **計算時間コストが高い**

→ **大規模NW・複雑なモデルに適用困難**

スタートアップゼミ2022 #2 参照

3. Decomposition method - DeC method

- D : 観測された目的地の集合($d \in D$)
- $(|\tilde{A}|+1)\times(|\tilde{A}|+1)$の行列 $\boldsymbol{M}^0$ と　$(|\tilde{A}|+1)\times(|\tilde{A}|+1)$の行列 $\boldsymbol{U}^d$ と定義

$$M^0{}_{ka} = \begin{cases} M_{ka} & if\ k, a \in A(k) \\ 0 & othrewise \end{cases} \qquad U^d{}_{ka} = \begin{cases} \delta(d|k)e^{\frac{1}{\mu}v(d|k)} & if\ a = d \\ 0 & othrewise \end{cases}$$

- 全てのdについて $\boxed{\boldsymbol{M}^d = \boldsymbol{M}^0 + \boldsymbol{U}^d}$

$$\underbrace{\begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,|\mathcal{A}|+1} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,|\mathcal{A}|+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}}_{\boldsymbol{M}^d} = \underbrace{\begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & 0 \\ m_{2,1} & m_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}}_{\boldsymbol{M}^0} + \underbrace{\begin{pmatrix} 0 & 0 & \cdots & m_{1,|\mathcal{A}|+1} \\ 0 & 0 & \cdots & m_{2,|\mathcal{A}|+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}}_{\boldsymbol{U}^d}$$

- $V^d(d) = 0$ から、$z^d_{|A|+1} = e^{\frac{1}{\mu}V^d(d)} = e^0 = 1$ and

$$\boldsymbol{U}^d\boldsymbol{z}^d = \begin{pmatrix} 0 & 0 & \cdots & m_{1,|\mathcal{A}|+1} \\ 0 & 0 & \cdots & m_{2,|\mathcal{A}|+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} z^d_1 \\ z^d_2 \\ \vdots \\ z^d_{|\mathcal{A}|+1} \end{pmatrix} = \begin{pmatrix} m_{1,|\mathcal{A}|+1} \\ m_{2,|\mathcal{A}|+1} \\ \vdots \\ 0 \end{pmatrix}$$

以上から

行列 $\boldsymbol{U}^d$ の最終列

$$\boldsymbol{z}^d = \boldsymbol{M}^d\,\boldsymbol{z}^d + \boldsymbol{b}^d = (\boldsymbol{M}^0 + \boldsymbol{U}^d)\boldsymbol{z}^d + \boldsymbol{b}^d = \boldsymbol{M}^0\boldsymbol{z}^d + \boldsymbol{U}^d\,\boldsymbol{z}^d + \boldsymbol{b}^d = \boldsymbol{M}^0\boldsymbol{z}^d + \boldsymbol{t}^d + \boldsymbol{b}^d$$

3. Decomposition method - DeC method

仮定1　$k \in A, a \in A(k), a \neq D$ のもとで、即時効用$v(a|k)$は独立かつ目的地$d \in D$に対して独立

→ **個人または目的地に固有の属性は存在しない**

$$\boldsymbol{M^0}\text{は}\forall \boldsymbol{d} \in \boldsymbol{D}\text{において独立} \leftarrow M^0{}_{ka} = \begin{cases} M_{ka} = \begin{cases} e^{\frac{1}{\mu}v(a|k)} & if\ a \in A(k) \\ 0 & othrewise \end{cases} \\ 0 \end{cases}$$

- $(|\tilde{A}| + 1) \times |D|$の行列$\boldsymbol{Z}$を定義　と　$(|\tilde{A}| + 1) \times |D|$の行列$\boldsymbol{B}$を定義

$$\boldsymbol{Z} = [z^{d_1}, z^{d_2}, \ldots, z^{d_{|D|}}] \ \ for\ \forall d \in D \qquad \boldsymbol{B} = \boldsymbol{t}^d + \boldsymbol{b}^d \ \ for\ \forall d \in D$$

- これにて以下の新しい連立一次方程式を得る

$$\boldsymbol{Z} = \boldsymbol{M}^0\boldsymbol{Z} + \boldsymbol{B} \Leftrightarrow \boxed{(\boldsymbol{I} - \boldsymbol{M}^0)\boldsymbol{Z} = \boldsymbol{B}} \cdots (5)$$

- $\boldsymbol{I} - \boldsymbol{M}^0$が可逆(逆行列を持つ)場合、一意の解$\boldsymbol{Z}$を持つ
- $\boldsymbol{I} - \boldsymbol{M}^0$が可逆となる条件はFosgerau(2013) から拡張可能

- LL関数の勾配とHessian行列も、 $\boldsymbol{Z}$のJacobianを導出することで得られる

ベクトルZはすべての目的地に対する全価値関数が含まれている

解決策 2　**一次方程式(5)を<u>1度だけ</u>解けばよい→ 計算時間コスト削減**

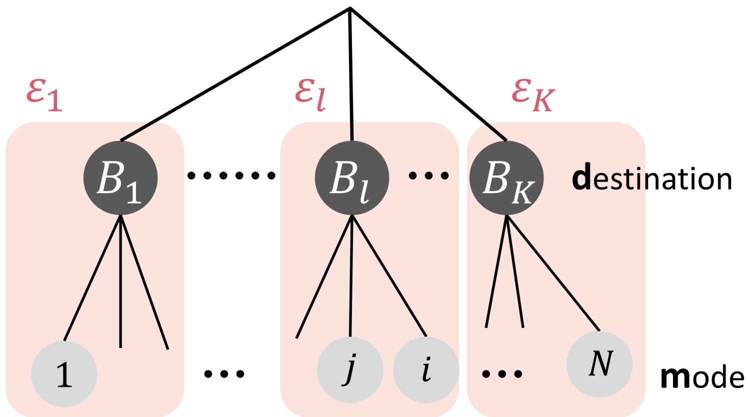
→ 大規模NW・<u>複雑なモデル</u>への適用可能性

復習編　Nested Logit Model (NL)

＝2 種類のスケールパラメータを使った選択確率の記述

例）交通手段選択

$d = 1,2, \dots, l, \dots, K$
$m = 1,2, \dots, j, i, \dots, N$



$$U_{dm} = V_d + V_m + V_{dm} + \varepsilon_d + \varepsilon_{dm}$$
U_{dm}: 選択肢(d,m)の真の効用

V_d：目的地選択肢dに特有な効用の確定項

V_m：手段選択肢mに特有な効用の確定項

V_{dm}：選択肢(d,m)に特有な効用の確定項

ε_d：目的地選択肢dに特有な効用の確率項
→ スケールパラメータμ_dの$i.i.d.$ガンベル分布に従う

ε_{dm} :選択肢(d,m)に特有な効用の確率項
→ スケールパラメータμの$i.i.d.$ガンベル分布に従う

$$V'_d = \frac{1}{\mu} \ln \sum_{m'} \exp\{\mu(V_{m'} + V_{dm'})\}$$
$$\varepsilon'_d = \max_{m'}(V_m + V_{dm} - \varepsilon_{dm}) - V'_d$$
とおいて、選択確率は

$$P(d,m) = \frac{\exp\{\mu(V_m + V_{dm})\}}{\sum_{m'} \exp\{\mu(V_{m'} + V_{dm'})\}} \cdot \frac{\exp\{\mu_d(V_d + V'_d)\}}{\sum_d \exp\{\mu_d(V_d + V'_d)\}}$$

スタートアップゼミ2022 #2 参照

番外編　NRLモデル (Mai et al. 2015 a)

RLモデル

解決 0 経路の選択肢集合が不明

課題 1 **経路同士の相関**　← **不十分 !!!** ※ **LS** 以外の要素に対して**IIA特性**が残る

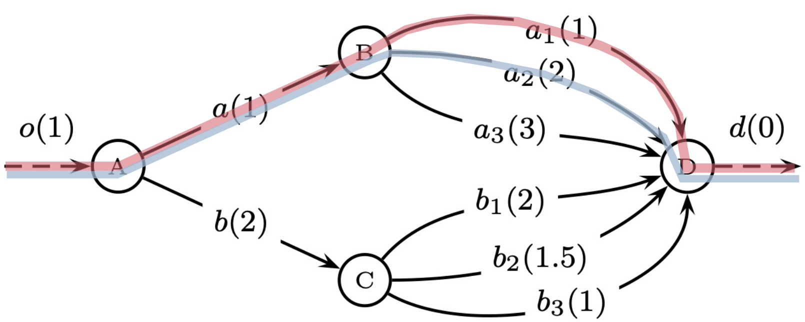
Nested RLモデル

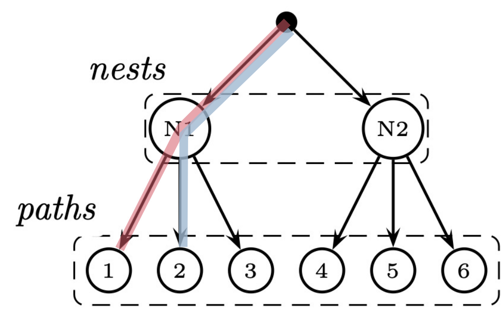
スケールパラメータを各リンクに固有に与えることでIIA特性を緩和（＝選択確率が変化）

-Mai et al. 2015 の検証-

スケールパラメータの設定

$\mu_a = 0.5 \quad \mu_b = 0.8 \quad \text{others}\ \mu = 1$





		Probabilities woth link removed			
Paths	Original	a_1	a_2	b_1	b_2
1：[o,a,a_1,d]	0.54	-	0.65	0.55	0.56
2：[o,a,a_2,d]	0.15	0.38	-	0.16	0.16
3：[o,a,a_3,e,d]	0.04	0.05	0.05	0.05	0.05
4：[o,b,b_1,e,d]	0.02	0.05	0.03	-	0.03
5：[o,b,b_2,d]	0.06	0.12	0.07	0.17	-
6：[o,b,b_3,d]	0.17	0.33	0.2	0.18	0.21

- **IIA特性は緩和**
- **代替のされ方はNW構造に依存**

4-2. 番外編　Mixed logit の定式化と理解

Mixed logit とは

- 誤差項に**正規分布**を仮定したプロビット系列のモデルと
  誤差項に**ガンベル分布**を仮定したモデルを組み合わせたもの
- MNLの選択確率のパラメータを確率分布によって扱うことで
  母集団内でのばらつきを表現

→ **誤差相関**や異分散性を表現可能

定式化

ωで評価されたロジット型確率式

$$L_{ni}(\omega) = \frac{e^{V_{ni}(\omega)}}{\sum_{j=1}^{J} e^{V_{nj}(\omega)}}$$

→選択確率

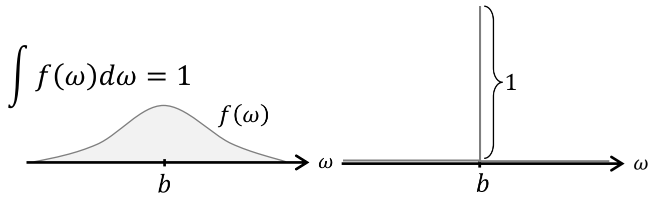
$$P_{ni} = \int L_{ni}(\omega) f(\omega) d\omega$$

確率密度関数(＝確率分布)

$V_{ni}(\omega)$:効用の確定部分(ωの関数)
ω:ランダム係数

確率密度関数 $f(\omega)$

- $f(\omega)$は離散、連続どちらもあり得る
  （ただし多くの場合で連続型を用いる）
  ※ 後述： $f(\omega)$はパラメータθの関数 $f(\omega|\theta)$
- ωの平均をbとすると
  $f(\omega) = 1\ for \omega = b, 0\ for\ \omega \neq b$
  となるとき、通常のロジットモデル



4-2.番外編 Mixed logit の定式化と理解

パラメータが個人nによって異なると考えるとき

個人ごとに異なるω_n（真の値）を実際に観測することはできない
→ 確率密度関数$f(\omega)$によって表現

→ 個人nがJ個の選択肢から選択肢jを選択したときの効用

$$U_{nj} = V_{nj} + \varepsilon_{nj} = \omega'_n x_{nj} + \varepsilon_{nj}$$

ω_nは個人によって異なる→ $f(\omega)$に従ってω'_nとなる、とおく

cf. ω_nが定数ω'_nであるとすれば確率式はロジットモデルと同様に
＝**パラメータが個人nで共通**

$$L_{ni}(\omega) = \frac{e^{\omega'_n x_{nj}}}{\sum_{j=1}^{J} e^{\omega'_n x_{nj}}}$$ とかける。

ω_nが$f(\omega)$に従って変動するためωに対して積分する必要があり、

選択確率は

$$P_{ni} = \int \left(\frac{e^{\omega'_n x_{ni}}}{\sum_{j=1}^{J} e^{\omega'_n x_{nj}}} \right) f(\omega) d\omega$$

Mixed logitにはIIA特性がない

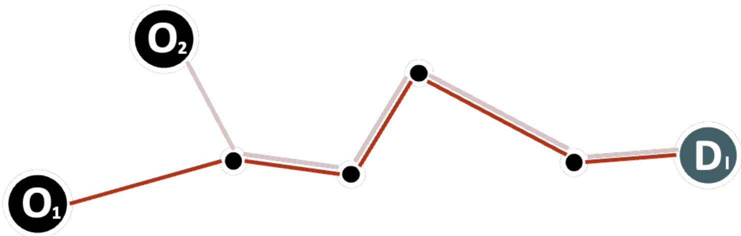
$\frac{P_{ni}}{P_{nj}}$は選択肢i, j以外を含む全データに依存

解決策 1 **理論的には**、Mixed logit型のRLを用いれば、**IIA特性を緩和可能**

4-2. 番外編　Mixed logit と Error Component ; 誤差相関

異なる選択肢間の**誤差項の相関**を考える場合

課題 1　**2つの経路は無関係に独立???**



$$U_{nj} = \alpha' x_{nj} + \mu'_n z_{nj} + \varepsilon_{nj}$$

x_{nj} ,z_{nj} : jに関する観測値ベクトル

α : 固定された係数ベクトル

μ_n : 平均0の変動する係数ベクトル

ε_{nj} : $i.i.d.$の誤差項

Nested Logit の考え方

ネスト構造を表現するダミー変数

$$\mu'_n z_{nj} = \sum_{k=1}^{K} \mu_{nk} d_{jk}$$

$$d_{jk} = \begin{cases} 1 & \text{選択肢 } j \text{がネスト} k \text{に含まれる} \\ 0 & \text{選択肢 } j \text{がネスト} k \text{に含まれない} \end{cases}$$

$$\mu_{nk} \sim N(0, \sigma_k)$$

効用関数は、

$$U_{nj} = \alpha' x_{nj} + \sum_{k=1}^{K} \mu_{nk} d_{jk} + \varepsilon_{nj}$$

4-2. 番外編　Mixed logit と Error Component ; 誤差相関

異なる選択肢間の**誤差項の相関**を考える場合

$$U_{nj} = \alpha' x_{nj} + \boxed{\mu'_n z_{nj}} + \boxed{\varepsilon_{nj}} \xrightarrow[\text{NL Analogy}]{} U_{nj} = \alpha' x_{nj} + \boxed{\sum_{k=1}^{K} \mu_{nk} d_{jk}} + \boxed{\varepsilon_{nj}}$$

$\mu_{nk} \sim N(0, \sigma_k)$

z_{nj}も誤差項なので、
ε_{nj}とまとめて効用の確率項η_{nj}とおく

$$\eta_{nj} = \boxed{\mu'_n z_{nj}} + \boxed{\varepsilon_{nj}}$$

選択肢iと選択肢jの誤差相関は

$$Cov(\eta_{ni}, \eta_{nj}) = E(\mu'_n z_{ni} + \varepsilon_{ni})(\mu'_n z_{nj} + \varepsilon_{nj}) \longrightarrow$$

ネストkに含まれる選択肢の共分散は

$$Cov(\eta_{ni}, \eta_{nj}) = E(\mu_{nk} + \varepsilon_{ni})(\mu_{nk} + \varepsilon_{nj}) = \sigma_k$$

$$Var(\eta_{ni}) = E(\mu_{nk} + \varepsilon_{ni})^2 = \sigma_k + \frac{\pi^2}{6}$$

ネストk内の相関係数

$$\frac{\sigma_k}{\sigma_k + \frac{\pi^2}{6}}$$ ：ネストkに依存

各ネストの誤差項の分散が等しい場合は
$\sigma_k = \sigma \quad for\ k = 1,2, \ldots, K$

$$\frac{\sigma}{\sigma + \frac{\pi^2}{6}} = Const.$$

4. Mixed Recursive Logit Model

混合型　　　再帰的　　　ロジット

- 経路 $\sigma = \{k_i\}_{i=0}^{J} = \{k_0, k_1, \dots, k_J\}$の選択確率

$$P(\sigma;\beta) = e^{-V^d(k_0;\beta)} \prod_{i=1}^{J} e^{v(k_{i+1}|k_i;\beta)}$$

- Mixed (MN)L の枠組みでは確定項 $v(k_{i+1}|k_i;\beta)$自体がランダムな要素を含む

$$\boxed{P(\sigma;\theta) = E_\omega[P(\sigma;\beta)] = \int P(\sigma;\theta,\omega) f(\omega) d\omega} \quad \dots (6) \qquad (\beta = \beta(\theta,\omega))$$

例： βがK次元の正規分布のランダムベクトルの場合

$$\omega = (\omega_1, \dots, \omega_K)\ with\ \ \omega_i \sim N(0,1)$$

θでβの平均 (means)・標準偏差 (SD：Standard Deviation) を決定

$f(\cdot)$：確率密度
θ　：LL関数(対数尤度関数)最大化によって推定されるパラメータ

- 経路選択結果の観測集合 $n = 1,2, \dots, N$に対して、 $LL(\theta) = \dfrac{1}{N}\sum_{n=1}^{N} \ln P(\sigma_n;\theta)$

個体nあたり1つの多次元積分＝解析解は得られない
→ 求積法 or シミュレーションによって**数値的に近似**する

課題 2　**複雑なモデル**において計算コストが膨大過ぎる

4. Mixed Recursive Logit Model

再掲：$P(\sigma;\theta)=E_{\omega}[P(\sigma;\beta)]=\int P(\sigma;\theta,\omega)f(\omega)d\omega$ … (6)

- 数値的に近似すると（求積法 or シミュレーション法）

$$\boxed{\tilde{P}_{R_n}(\sigma_n;\theta)=\sum_{i=1}^{R_n}\nu_{n,i}\,P\big(\sigma_n;\theta,\omega_{n,i}\big)} \quad \ldots (7)$$

$\omega_{n,i}$：数値積分の分点
$\nu_{n,i}$：数値積分の重み

- 特にモンテカルロ法では、積分分点$\omega_{n,i}$はωの分布からランダムに描かれる

$$\tilde{P}_{R_n}(\sigma_n;\theta)=\frac{1}{R_n}\sum_{i=1}^{R_n}P\big(\sigma_n;\theta,\omega_{n,i}\big) \quad \ldots (8)$$

※ モンテカルロ法では抽選数 $R_n\to\infty$で真の最尤推定値に収束する
(Bastin et al. 2006)

- DeC法は(7)の形の任意の近似に適用可能
- ただしDeC法の最大の強みである「**解くべき線形システムを1本にする**」を満たすため以下の条件で制限

条件：積分分点は観測経路nに関わらず共通

$$R_n\equiv R,\qquad \nu_{n,i}\equiv\nu_i,\qquad \omega_{n,i}\equiv\omega_i,\qquad n=1,2,\ldots,N$$

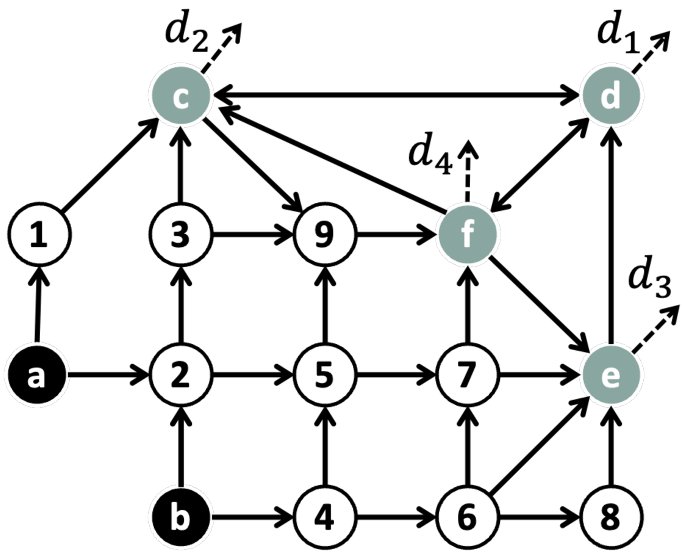
※ 選択確率のシミュレーション・バイアスは $R\to\infty$で$\big|\tilde{P}_{R_n}(\sigma_n;\theta)-\tilde{P}_{R_m}(\sigma_m;\theta)\big|\to 0$ (証明は論文中) → **検証**

力不足

解決策 2 <u>**Mixed RL モデル**</u>では異なる積分分点ごとに線形システム(5)を解く必要 → **計算コスト 膨大**

解決策 2+α + 積分分点が**観測経路n間でみな同じ**であるという条件 → **解くべき一次方程式を1つに**

4-1. Mixed Recursive Logit Model - IRNとCRNの比較実験



2つの起点ノード：$\{a, b\}$
4つの目的地ノード：$\{d, c, e, f\}$
8つの観測ODペア
8つの観測ODペア各々に5つの経路→40経路

$TT(a)$：リンクaの移動時間([0,1])

- シミュレーションにおける乱数生成方法

  IRL（independent random numbers）
  ＝観測OD毎に$\omega_{n,i}$を独立に抽選（既往研究）

  CRL（common random numbers）
  ＝観測OD共通のセットで抽選 ($\omega_{n,i} \equiv \omega_i$)（提案）

- 即時効用　$u(a|k) = \beta_1 TT(a) + \beta_2 + \varepsilon(a)$

  β_1, β_2: 正規分布　ε : $i.i.d.$ガンベル

  $\beta_1 \sim N(\theta_{TT}, \sigma_{TT}^2)$　$\beta_2 \sim N(0, \sigma^2)$

- モンテカルロドローを50、200、1000回で比較
- 100回の独立した推定の平均値を推定値として採用

実験結果

1. 抽選回数が十分に大（1000回）→ 5%の有意水準で最終尤度はIRNとCRNで有意な差はない
   → CRNは標準的なIRNに代わる方法
2. **DeC法は計算時間を10～50倍も短縮（高速化）**

   ※ IRN→CRNはLL計算コストに影響微小（最適化アルゴリズムの反復回数をわずかに削減するのみ）

課題整理

～4章

課題 1 → **解決策 1** 複雑なモデル Mixed RLモデル の採用 → IIA特性が緩和 **理論的に可能**

課題 2 → **解決策 2+α** Decomposition method (DeC method)を提案 → 簡易NWで高速化を確認

実際に可能？
複雑な実際のデータ（複雑なNW・膨大な観測OD）で**計算コスト的に**推定可能？
複雑なモデル（Mixed RL モデル）で**計算コスト的に**推定可能 ??? **推定精度**は??

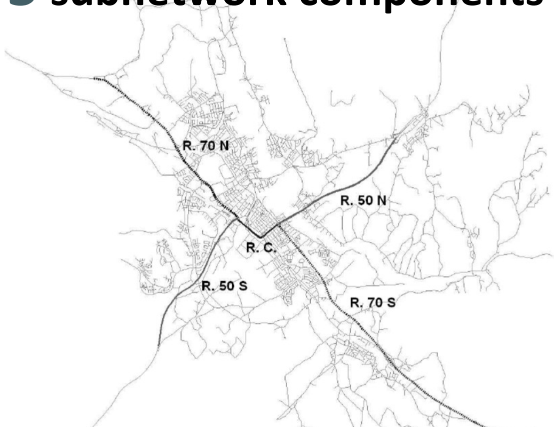
5章＋6章　複雑な実際の Data setにおけるMixed RLモデルを用いた推定

約30倍高速

		Data set 1 **Car**		Data set 2 **Bicycle**	
ネットワーク	リンク数		7459		42,372
	ノード数		3077		16,352
サンプル	経路数		21,452		127,001
	目的地		1832		648
モデルパラメータ数			4		14
Estimation method		**DeC**	**Oridiginal**	**DeC**	**Original**
計算時間 (sec)	価値関数	2	90	10	235
	LL及び勾配計算	10	284	54	1605
	総推定時間	230	6670	4824	144,110

6-1. Mixed RLモデルに対するDeC法の計算結果 -モデルの仕様

5 subnetwork components



Borlänge Road Network and **Subnetwork Definition**
(Frejinger and Bierlaire, 2007)

$TT(a)$: リンクaの移動時間
$LT(a|k)$: 左旋回ダミー(リンクaからkへの旋回角度 α_{LT}が $40° < \alpha_{LT} < 177° \rightarrow 1$)
$UT(a|k)$: Uターンダミー(リンクaからkへの旋回角度 α_{LT}が $\alpha_{LT} > 177° \rightarrow 1$)
$LC(a)$: リンク定数 (本実験ではすべてのリンクで1と設定)
$LS(a)$: リンクサイズ (Fosgerau:2013)

$\varepsilon(a) \sim i.i.d.$ガンベル

$\beta^*_{TT} \sim N(\mu_{TT}, \sigma^2_{TT})$

$\zeta \sim N(0,1)$
$F(a)_q$
size=5のvector
q : subnetwork

$$F(a)_q = \begin{cases} 0 & if\ a \notin q \\ \sqrt{TT(a)} & otherwise \end{cases}$$

$$T(\sigma) = \mathrm{diag}(\sigma_{R.50N}, \sigma_{R.50S}, \sigma_{R.70N}, \sigma_{R.70S}, \sigma_{R.C})$$
$$\sigma = \{\sigma_{R.50N}, \sigma_{R.50S}, \sigma_{R.70N}, \sigma_{R.70S}, \sigma_{R.C}\}$$

- **RL**

$$u^{RL}(a|k;\beta) = \beta_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k) + \varepsilon(a)$$

- **RL with LS属性 ; RL-LS**

$$u^{RL\cdot LS}(a|k;\beta) = \beta_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k) + \beta_{LS}LS(a) + \varepsilon(a)$$

- **Mixed RL with random parameters ; MRL-RP**

$$u^{MRL}_{RP}(a|k;\beta) = \beta^*_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k) + \varepsilon(a)$$

- **Mixed RL with EC ; MRL-EC**

$$u^{MRL}_{EC}(a|k;\beta,\sigma) = \beta_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k) + F(a)^T T(\sigma)\zeta + \varepsilon(a)$$

- **Mixed RL with EC and random parameters ; MRL-REC**

$$u^{MRL}_{REC}(a|k;\beta,\sigma) = \beta^*_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k) + F(a)^T T(\sigma)\zeta + \varepsilon(a)$$

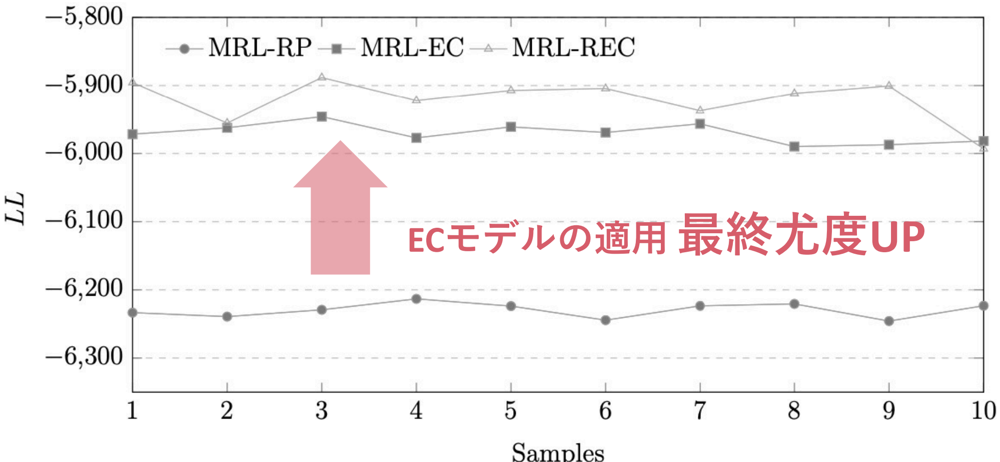
6-2. Mixed RLモデルに対するDeC法の計算結果 -適合度検証

複雑な実際のデータ（複雑なNW・膨大な観測OD）で**計算コスト的に**推定可能？
複雑なモデル（Mixed RL モデル）で**計算コスト的に**推定可能??? **推定精度**は??

計算コスト

DeCを用いない場合数ヶ月かかるMixed RLモデルを用いた推定が**3-5日程度で可能に**＝**実際に可能**

※ 抽選回数=500回



10回のMCシミュレーションによる最終尤度の平均値

尤度比検定の結果

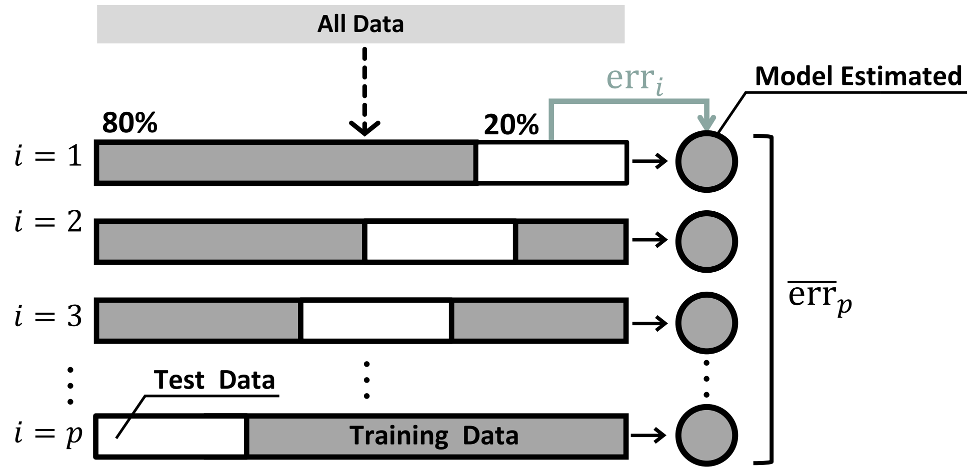
Models	χ^2	p value
RL & MRL-RP	148.42	3.84e−34
RL & MRL-EC	667.22	6.00e−142
MRL-RP & MRL-REC	616.16	6.52e−131
MRL-EC & MRL-REC	97.36	5.78e−23

適合度検証

尤度比検定の結果から、Mixed RLモデルはRLモデルよりも**適合度が高い**

6-3. Mixed RLモデルに対するDeC法の計算結果 -予測性能

● **クロスバリデーション**による**予測性能**比較

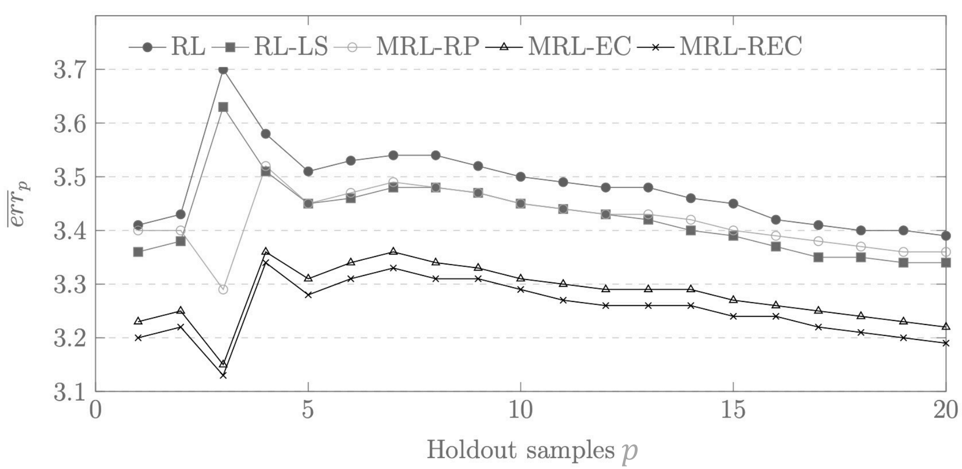


- **検定誤差** ※ $|PS_i|$: サンプルサイズ

$$\mathrm{err}_i = -\frac{1}{|PS_i|} \sum_{\sigma_j \in PS_i} \ln \tilde{P}^R(\sigma_j, \hat{\theta}_i) \quad 1 \le i \le 20$$

- **検定誤差の平均値**

$$\overline{\mathrm{err}}_p = \frac{1}{p} \sum_{i=1}^{p} \mathrm{err}_i \quad \forall\, 1 \le p \le 20$$



● Mixed RLモデルはRLモデルよりも**予測性能**が優れている

● 特に**MRL-(R)ECモデル**は他のモデルに比較して
- ・サンプル内適合度が著しく高い
- ・サンプル外適合度も高い

7. まとめと課題の整理

本論文の成果まとめ

課題 1 → 解決策 1　Mixed RLモデル の採用 → IIA特性が緩和　理論的に可能

課題 2 → 解決策 $2^{+\alpha}$　DeC法を提案→ 超高速化 Mixed RLモデルで実証

実際に可能

MRL-(R)EC モデル
- 有意に適合度が高い
- 標本外適合も良好=**予測性能**

→
- 空間相関のモデル化
- パネルデータでの応用

課題の整理

- **個人や目的地の異質性に対応できない**
  - DeC法適用のための”条件”(p.16参照)を設定しているため
  - 属性が目的地固有の場合 例:LS属性
  - 効用が旅行者の特性に依存する場合 例：年齢、収入etc.

    ＊個人の異質性の考慮は行動原理の重要事項なので致命的では?

    ＊属性が各クラスで個々に独立であるようにグループ化し、各クラスにDeC法を適用
- **多変量極限分布の誤差項を持つモデルには未対応**
  - 価値関数がそもそも線形方程式の形ではなく、DeC法を適用できない

参考資料

RL model と 経路相関

[1]夏学期ゼミ2021 #8 発表資料 http://bin.t.u-tokyo.ac.jp/szemi21/#6-2

[2]理論輪読会2020 #5-2 発表資料 http://bin.t.u-tokyo.ac.jp/rzemi20/#5-2

Nested RL model

[3]夏の理論合宿2018 #1-5 発表資料 http://bin.t.u-tokyo.ac.jp/summercamp2018/

[4]理論談話会2016 #4-1 発表資料 http://bin.t.u-tokyo.ac.jp/rzemi16/

[5] Mai, T., Fosgerau, M., & Frejinger, E. (2015). A nested recursive logit model for route choice analysis. *Transportation Research Part B*: Methodological, 75, 100-112.

[6]理論談話会2015 #6 発表資料 http://bin.t.u-tokyo.ac.jp/rzemi15/file/6_miki.pdf

クロスバリデーション

[7] 【AI・機械学習】ホールドアウト検証とK分割交差検証(K-foldクロスバリデーション)｜モデル性能の評価 https://di-acc2.com/analytics/ai/6498/