

A nested recursive logit model for route choice analysis

Tien Mai, Mogens Fosgerau, Emma Frejinger
Transportation Research Part B, Vol. 75, pp.100-112, 2015

2015/06/19(金)

理論談話会2015#6

B4 三木真理子

目次

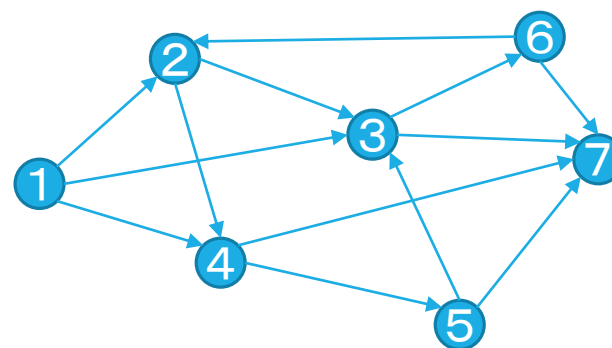
1. NRLモデルの強み
2. NRLモデルの定式化
3. 期待価値関数の計算方法と最尤推定法
4. ケーススタディ
5. アクティビティネットワークでの経路選択問題への適用

経路選択問題の目標と課題

ネットワーク上の経路選択問題の分析には、
一般に離散選択モデルが用いられる。

経路選択モデルを作るうえでの目標

- ①すべてのパラメータを矛盾なく一貫して推定する
- ②将来予測にも有効なものを作る



経路選択モデルの難点

- ①経路の選択肢集合が分析者にはわからない上に、
あるODペアについてすべての可能な経路を列挙するのは不可能
- ②経路の効用は互いに独立ではなく、関連していると思われる

目標を2つともクリアし、難点を2つとも克服したモデルはNRLモデルのみ！

既往研究(選択肢集合抽出型)

現存する多くの経路選択モデル

: 経路の選択肢集合に基づいた経路選択モデル

(例. Guevara and Ben-Akiva(2013a,b) GEVモデルとmixedロジットモデル)



- 難点①経路列挙の不可能性のため、推定や適用の前に選択肢集合を抽出する必要がある
- 抽出された選択肢集合に応じて、選択肢の効用を調整する必要がある
- パラメータを矛盾なく推定することはできるが、予測にどう用いたら良いかはまだわからない

既往研究(選択肢集合抽出型)

難点②経路がリンクの重なりによる相関をもつことと向き合った経路選択モデル

: link-nested logitモデル (Vovsha and Bekhor, 1998)

誤差項の相関を考慮したmixed logitモデル(Frejinger and Bierlaire, 2007)



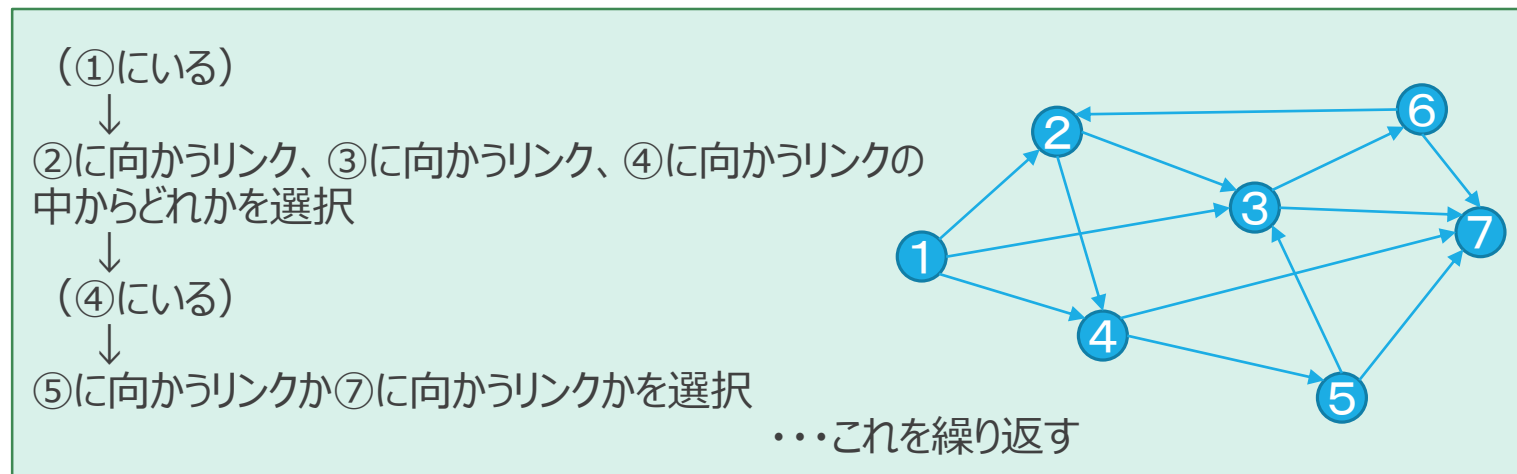
- ・現状、抽出された選択肢集合に基づいて、効用を修正することなしに推定している
⇒選択肢集合からいくつか経路を抜いたり足したりするだけで推定値が変わる
⇒実際の選択肢集合がわからない現実世界で用いるには不向き？
- ・ただし先ほど示したように、効用を調整することはできる

RL(recursive logit)モデル

経路の選択肢集合を抽出することなく、経路の選択確率を求める

⇒逐次的離散選択モデル

: 選択者は、次に進むリンクをマルコフ性をもつ確率的な過程で選んでいき、結果的に経路の選択が行われる (link-based approach)



これは、すべての可能な経路に対して、MNLで経路選択を行うことと同等！

RLからNRLへ

ということでRLモデルは、

2つの目標はクリアし、難点①を克服したが、難点②はまだ不十分

※リンクサイズLSという指標で、リンクの共有による経路の相関は考慮することはできるが、リンクサイズ以外の要素(リンク旅行時間など)が変化した場合、IIA特性は残っている



スケールパラメータを各リンクに固有なものとして与えることで、IIA特性を緩和
⇒NRL(nested recursive logit)モデル

※ネスティットロジットモデルと同様のアイデアなので、nested RLモデル

(復習) Nested Logit model(NL)

–2種のスケールパラメータを使った選択確率

$$U_{dm} = V_d + V_m + V_{dm} + \varepsilon_d + \varepsilon_{dm}$$

U_{dm} : 選択肢(d,m)の真の効用

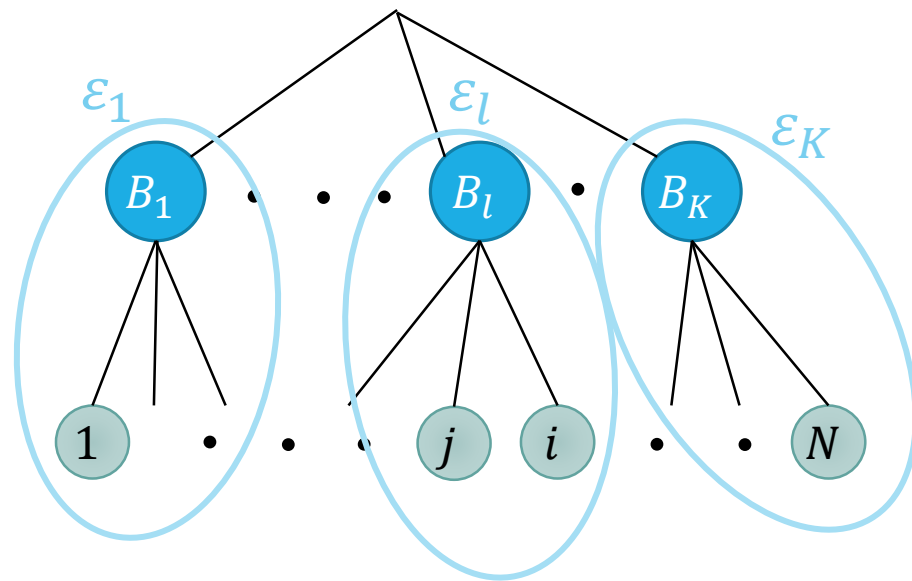
V_d : 目的地選択肢dに特有な効用の確定項

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dに特有な効用の確率項

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IIAガンベル分布に従う



$V'_d = \frac{1}{\mu} \ln \sum_{m'} \exp\{\mu(V_m + V_{dm})\}$, $\varepsilon'_d = \max_{m'}(V_m + V_{dm} + \varepsilon_{dm}) - V'_d$ として、

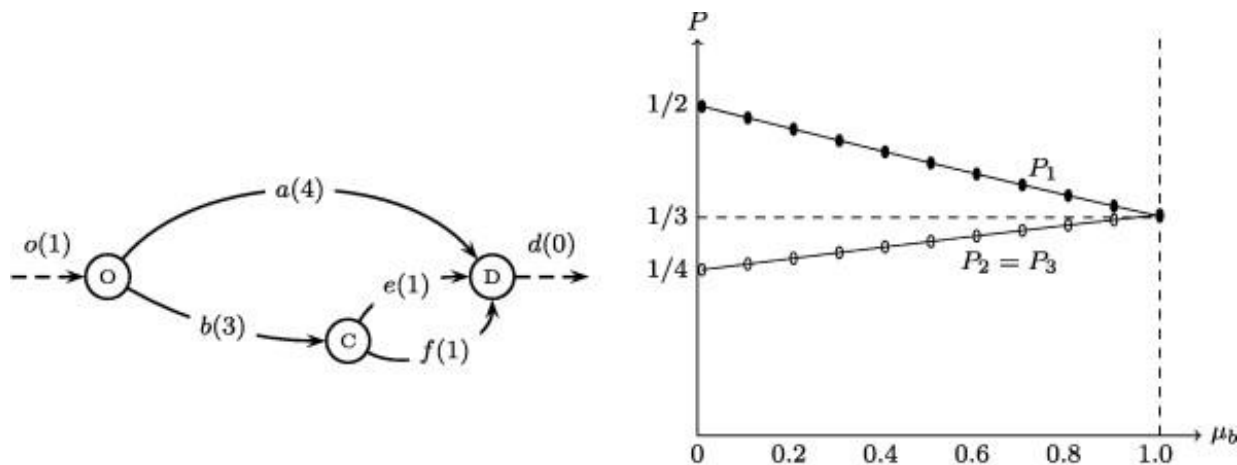
$$P(d, m) = \frac{\exp\{\mu(V_m + V_{dm})\}}{\sum_{m'} \exp\{\mu(V_{m'} + V_{dm'})\}} \cdot \frac{\exp\{\mu_d(V_d + V'_d)\}}{\sum_{d'} \exp\{\mu_d(V_{d'} + V'_{d'})\}}$$

スケールパラメータをリンク固有にする意義

例として、いくつかのネットワークで経路選択確率を出してみる

※簡単のためにリンク長のみを説明変数としているため、経路の重なり具合は説明変数としては考慮されない

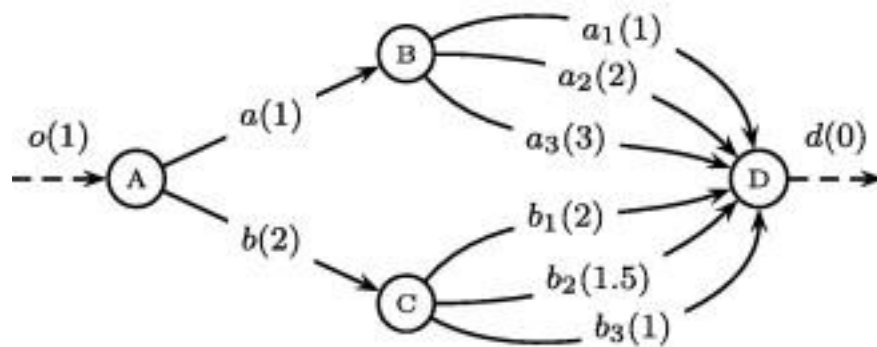
リンクbのスケールパラメータを変化させ、それ以外を1に固定した結果が以下



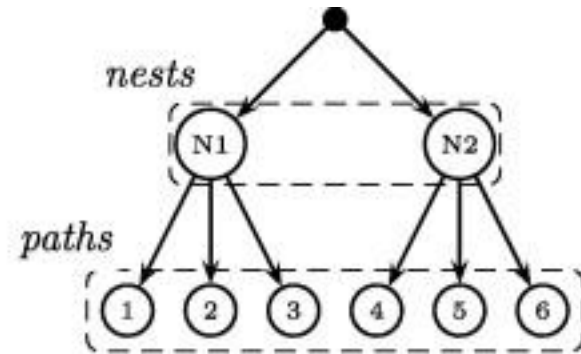
RLモデル(すべてのリンクのスケールパラメータが1)では、どの経路も1/3の選択確率

NRLモデルではスケールパラメータの違いによって選択確率に変化が起こっている

スケールパラメータをリンク固有にする意義



ネットワーク例



経路選択枝のネスト

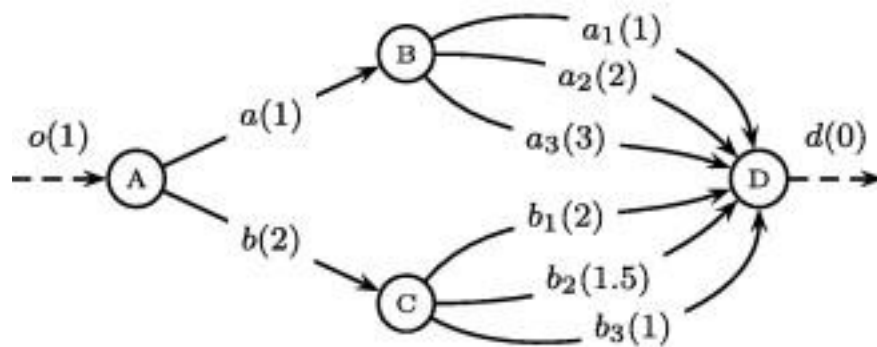
上のネットワークで、

リンクaのスケールパラメータを0.5、リンクbのスケールパラメータを0.8、

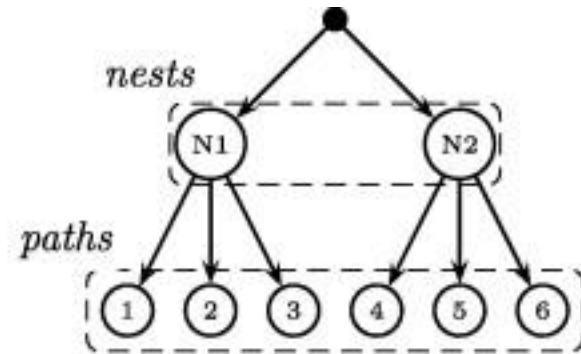
それ以外のリンクのスケールパラメータを1とする。

ここから、リンクa1,a2,b1,b2をそれぞれ取り除いた場合、他のリンクの選択確率がどのように変化するかを見る

スケールパラメータをリンク固有にする意義



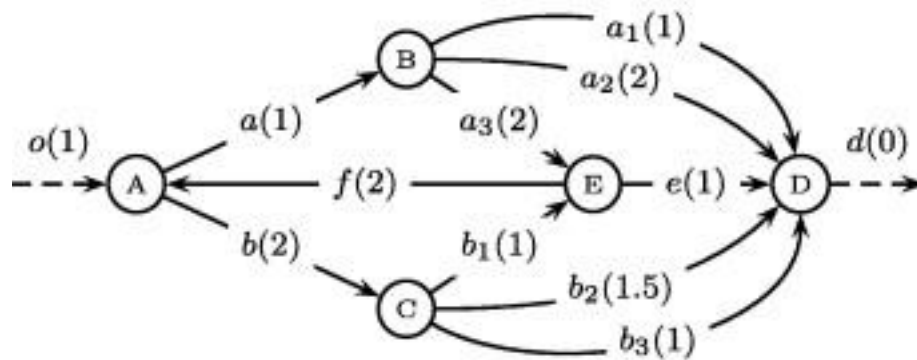
ネットワーク例



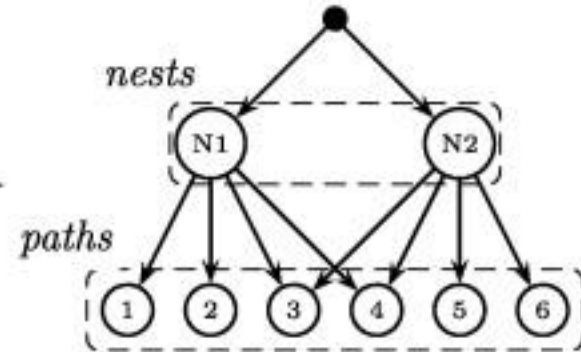
経路選択枝のネスト

Paths	Original	Probabilities with link removed			
		a1	a2	b1	b2
1	0.54	-	0.65(+20%)	0.55(+1%)	0.56(+4%)
2	0.15	0.38(+151%)	-	0.16(+1%)	0.16(+4%)
3	0.04	0.05(+93%)	0.05(+20%)	0.05(+1%)	0.05(+4%)
4	0.02	0.05(+93%)	0.03(+15%)	-	0.03(+19%)
5	0.06	0.12(+93%)	0.07(+15%)	0.17(+6%)	-
6	0.17	0.33(+93%)	0.20(+15%)	0.18(+6%)	0.21(+19%)

スケールパラメータをリンク固有にする意義



ネットワーク例



経路選択枝のネスト

Paths	Original	Probabilities with link removed		
		a1	b3	f
1:[o,a,a1,d]	0.54	-	0.60(+12%)	0.54(+0.7%)
2:[o,a,a2,d]	0.15	0.38(+150%)	0.17(+12%)	0.15(+0.7%)
3:[o,a,a3,e,d]	0.05	0.11(+148%)	0.05(+11%)	0.04(-1.3%)
4:[o,b,b1,e,d]	0.03	0.05(+86%)	0.05(+90%)	0.02(-6.7%)
5:[o,b,b2,d]	0.06	0.12(+93%)	0.12(+91%)	0.06(+1.4%)
6:[o,b,b3,d]	0.17	0.33(+93%)	-	0.17(+1.4%)

結論:IIA特性は緩和され、代替のされ方はネットワーク構造に依存する

経路選択モデルの比較

経路選択モデルを比較する4つの視点

- ①すべてのパラメータを矛盾なく推定できるか
- ②将来予測にも有効かどうか
- ③選択肢集合を限定する必要がないかどうか
- ④経路の重なりなどによる効用の相関を記述できているかどうか

選択肢集合抽出型（従来の大多数）⇒①のみクリア。④は努力しだい。

RLモデル⇒①、②、③はクリア。④はLSによる部分的な解決にとどまる。

NRLモデル⇒①、②、③、④のすべてをクリア！

ということで、NRLを定式化していく

ネットワークの定義

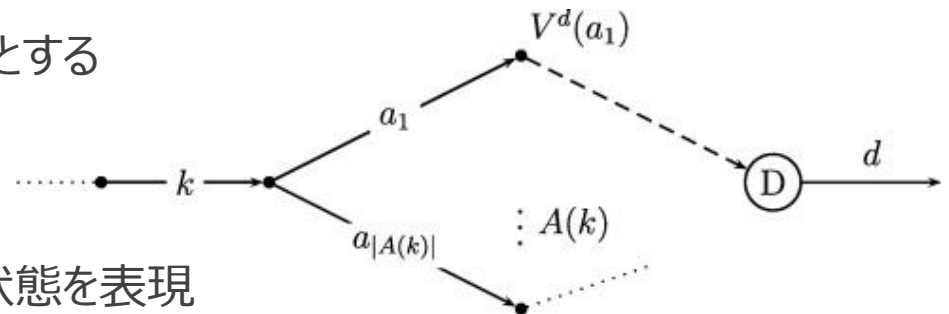
$G = (A, \nu)$: 方向を持つリンクで結ばれたネットワーク

A : リンクの集合

ν : ノードの集合

※サイクルのあるネットワークでもOK

それぞれのリンク $k \in A$ について、
 k の終点から伸びるリンクの集合を $A(k)$ とする



それぞれの目的地 D について、
 D からダミーリンク d を伸ばすことで吸収状態を表現

この結果、全リンクの集合 \tilde{A} は以下ようになる

$$\tilde{A} = A \cup \{d\}$$

リンクの効用関数の定義

あるリンク k の終点ノードにおいて、次のリンクを選ぶときの原理
⇒効用最大化

k から a に移るとき、およびリンク a の通過に関する効用(IU)
+
リンク a を選んだ場合の、リンク a 通過後から目的地まで経路の期待効用(EU)

あるリンク k の終点ノードで、 $a \in A(k)$ を選択するとき、

$$\text{IU: } u^n(a|k; \beta) = v^n(a|k; \beta) + \mu_k \epsilon(a) \quad (1)$$

$$\text{EU: } V^d(k; \beta) = \mathbb{E} \left[\max_{a \in A(k)} \left(v^n(a|k; \beta) + V^d(a; \beta) + \mu_k \epsilon(a) \right) \right] \quad (2)$$

$v^n(a|k; \beta)$: IUの確定項 (ダミーリンク d 以外では負の値)

$\epsilon(a)$: IUの確率項 (iidガンベル分布に従う)

β : 未知パラメータ

μ_k : リンク k に固有の確率項のスケールパラメータ (正の値)

$P^d(a|k)$ を求める

$$P^d(a|k) = \delta(a|k) \frac{\exp\left(\frac{1}{\mu_k} (v(a|k) + V^d(a))\right)}{\sum_{a' \in A(k)} \exp\left(\frac{1}{\mu_k} (v(a'|k) + V^d(a'))\right)}$$

ここで、

$$V^d(k; \beta) = \mathbb{E} \left[\max_{a \in A(k)} (v^n(a|k; \beta) + V^d(a; \beta) + \mu_k \epsilon(a)) \right]$$

より、

$$\begin{aligned} \frac{1}{\mu_k} V^d(k) &= \mathbb{E} \left[\max_{a \in A(k)} \left(\frac{1}{\mu_k} (v(a|k) + V^d(a)) + \epsilon(a) \right) \right] \\ &= \ln \left(\sum_{a \in A(k)} \exp \left(\frac{1}{\mu_k} (v(a|k) + V^d(a)) \right) \right) \end{aligned}$$

よって、

$$\sum_{a \in A(k)} \exp \left(\frac{1}{\mu_k} (v(a|k) + V^d(a)) \right) = e^{\frac{1}{\mu_k} V^d(k)} \quad (3)$$

左の結果から、

$$P^d(a|k) = \delta(a|k) e^{\frac{1}{\mu_k} (v(a|k) + V^d(a) - V^d(k))} \quad (4)$$

$$\delta(a|k) = \begin{cases} 1 & \text{if } a \in A(k) \\ 0 & \text{otherwise} \end{cases}$$

この式は、 $\forall k, a \in \tilde{A}$ において成り立つ。

経路の選択確率

$[k_0, k_1, \dots, k_l]$ というリンクで成り立つ経路 σ が選択される確率は、

$P^d(a|k) = \delta(a|k) e^{\frac{1}{\mu_k}(v(a|k) + V^d(a) - V^d(k))}$ より、

$$P(\sigma) = \prod_{i=0}^{l-1} e^{\frac{1}{\mu_{k_i}}(v(k_{i+1}|k_i) + V^d(k_{i+1}) - V^d(k_i))} \quad (5)$$

ちなみに、2つの経路 $\sigma_1 = [k_0, k_1, \dots, k_{l_1}]$ と $\sigma_2 = [h_0, h_1, \dots, h_{l_2}]$ の選択確率の比は

$$\frac{P(\sigma_1)}{P(\sigma_2)} = \frac{\prod_{i=0}^{l_1-1} e^{\frac{1}{\mu_{k_i}}(v(k_{i+1}|k_i) + V^d(k_{i+1}) - V^d(k_i))}}{\prod_{i=0}^{l_2-1} e^{\frac{1}{\mu_{h_i}}(v(h_{i+1}|h_i) + V^d(h_{i+1}) - V^d(h_i))}}$$

複数の $V^d(k)$ がある \rightarrow 確率の比にネットワーク全体が影響する \Rightarrow IIA特性はもはや保たれていない！

NRLモデルの難しさ

NRLモデルを実装するうえで、困難なことが2つある

①期待値関数 $V^d(k)$ の計算

非線形の行列計算。再帰的に求める必要があるため計算負荷が高い。

→効率的な求解方法の提案

②最尤推定法

尤度関数の定義と、

そのヘッセ行列を効率的に求められることを示す。

↙
今回は触れません

①期待値関数 $V^d(k)$ の計算

$$(3) \sum_{a \in A(k)} \exp\left(\frac{1}{\mu_k} (v(a|k) + V^d(a))\right) = e^{\frac{1}{\mu_k} V^d(k)} \quad \text{について、仮定より、}$$
$$V^d(d) = 0$$

よって、(3)は以下のようにかける

$$e^{\frac{1}{\mu_k} V^d(k)} = \begin{cases} \sum_{a \in A} \exp\left(\frac{1}{\mu_k} (v(a|k) + V^d(a))\right) & \forall k \in A \\ 1 & k = d \end{cases} \quad (6)$$

ここで $|\tilde{A}| \times |\tilde{A}|$ 行列 M^d と、 $|\tilde{A}| \times 1$ ベクトル z^d を以下のように定義する

$$M_{ka}^d = \delta(a|k) e^{\frac{v(a|k)}{\mu_k}}, \quad z_k^d = e^{\frac{V(k)}{\mu_k}}, \quad k, a \in \tilde{A} \quad (7)$$

この z^d を求めることで、 $V^d(k)$ を計算することを目指す

①期待値関数 $V^d(k)$ の計算

— z^d の計算

$$e^{\frac{1}{\mu_k} V^d(k)} = \begin{cases} \sum_{a \in A} \exp\left(\frac{1}{\mu_k} (v(a|k) + V^d(a))\right) & \forall k \in A \\ 1 & k = d \end{cases} \quad (6)$$

より、

$$z_k^d = \begin{cases} \sum_{a \in A} M_{ka}^d (z_a^d)^{\mu_a / \mu_k} & \forall k \in A \\ 1 & k = d \end{cases} \quad (7)$$

$$M_{ka}^d = \delta(a|k) e^{\frac{v(a|k)}{\mu_k}}$$

$$z_a^d = e^{\frac{V(a)}{\mu_a}}$$

この z^d を、繰り返し計算によって求める

①期待値関数 $V^d(k)$ の計算

— z^d の計算

$|\tilde{A}| \times |\tilde{A}|$ 行列 $X(z)$ を以下のように定義する

$$X(z)_{ka} = z_a^{\mu_a/\mu_k} \quad \forall k, a \in \tilde{A} \quad (8)$$

このとき、(7) $z_k^d = \begin{cases} \sum_{a \in A} M_{ka}^d (z_a^d)^{\mu_a/\mu_k} & \forall k \in A \\ 1 & k = d \end{cases}$ は、

$$z = [M \circ X(z)]e + b \quad (9)$$

b : $|\tilde{A}| \times 1$ ベクトル. リンク d は1、それ以外は0

e : $|\tilde{A}| \times 1$ ベクトル. 全要素が1

\circ : 要素同士を掛け合わせる操作

のようにかける

①期待価値関数 $V^d(k)$ の計算

— z^d の計算

厳密解 z^d は(9) $z = [M \circ X(z)]e + b$ を満たす。これを繰り返し計算で求める。
初期値を z^0 として、

$$z^{i+1} \leftarrow [M \circ X(z^i)]e + b \quad (10)$$

以上の操作で次の z^i を求める。ある閾値 γ を決めて、

$$\|z^{i+1} - z^i\|^2 < \gamma$$

を満たすまで計算を繰り返す。

(9)式が解を持つ場合、有限解の操作で必ず収束し、収束の速さは z^0 に依存する

↑
RLモデルでの解($\mu_k = \mu \forall k \in \tilde{A}$ と仮定した場合の解)を入れると速い

↳ $V^d(k)$ が線形の関数で計算できるため、計算負荷が低く求められる

①期待価値関数 $V^d(k)$ の計算

— z^d の計算

実際の $V^d(k)$ は、パラメータ β の値に依存する

⇒最尤推定法で β を計算するたびに、 $V^d(k)$ も繰り返し解く必要がある



閾値 γ を、最初は大きめの値で設定し、計算を繰り返すたびにより小さい値に更新していくことで、計算負荷を減らす

②最尤推定法

Rust(1987)のNested fixed point algorithm(NFXP)を使って、
逐次的離散選択モデルを推定する

仕組み：尤度関数の最大化を行うことでパラメータを推定する。

尤度関数の最大化をするアルゴリズムの中で、先ほどのやり方で $V^d(k)$ を求める

推定するパラメータ： β とスケールパラメータ μ_k

→すべてのスケールパラメータをそれぞれ推定することはできないので、

μ_k を β の関数 $\mu_k(\beta)$ と仮定する

観測された経路集合($n=1,2,\dots,N$)について、尤度関数LLを定義すると、

$$\begin{aligned} LL(\beta) &= \sum_{n=1}^N \ln P(\sigma_n, \beta) \\ &= \sum_{n=1}^N \sum_{t=0}^{l_n} \frac{1}{\mu_{k_t}} (v^n(k_{t+1}|k_t) + V^n(k_{t+1}) - V^n(k_t)) \end{aligned}$$

実ネットワークと使用サンプル

スウェーデンのボルレンゲという街の実ネットワークで経路選択を試みる
(リンク数7459、ノード数3077、
交通量の増加による旅行時間の変化はないものと仮定)

<推定に用いるサンプル>

1832トリップ
(観測された経路のうち、もっともリンク数の少ないもので5リンク)

到着地ノード：466個

ODペア：1420種類

37,000回以上のリンク選択が行われている

これらを、RLモデルとNRLモデル双方で推定。
双方について、リンクサイズ(LS)を考慮したものと考慮しないものの2種類のモデルがある
⇒計4種類のモデルで比較

リンク効用関数の定義

$$\begin{aligned}v^{RL}(a|k; \beta) &= v^{NRL}(a|k; \beta) \\ &= \beta_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k)\end{aligned}$$

$$\begin{aligned}v^{RL-LS}(a|k; \beta) &= v^{NRL-LS}(a|k; \beta) \\ &= \beta_{TT}TT(a) + \beta_{LT}LT(a|k) + \beta_{LC}LC(a) + \beta_{UT}UT(a|k) + \beta_{LS}LS(a)\end{aligned}$$

$TT(a)$:リンク a の旅行時間

$LT(a|k)$:左折ダミー(リンク a と k のなす角が 40° 以上 177° 以下)

$LC(a)$:リンク固有定数

$UT(a|k)$:Uターンダミー(リンク a と k のなす角が 177° 以上)

$LS(a)$:リンクサイズ(対象経路のODに絞った場合のリンクの期待フロー)

β は未知パラメータ。初期値のみこちらが与える。

IUの定義

<スケールパラメータ μ_k について>

すべてを未知パラメータとして推定するのは不可能
 $\Rightarrow \mu_k$ をリンクの説明変数の関数として考える

$\mu_k > 0$ より、 $\mu_k = e^{\lambda_k}$ とおけて、

$$\lambda_k = \omega_{TT}TT(k) + \omega_{LS}LS(k) + \omega_{OL}OL(k)$$

$TT(k)$:リンク k の旅行時間

$LS(k)$:リンクサイズ(対象経路のODに絞った場合のリンクの期待フロー)

$OL(k)$: $|A(k)|$ リnk k の終点ノードから伸びるリンク数

よって、IUは以下のように定義される

$$u^{RL}(a|k; \beta) = v^{RL}(a|k; \beta) + \mu \epsilon(a)$$

$$u^{RL-LS}(a|k; \beta) = v^{RL-LS}(a|k; \beta) + \mu \epsilon(a)$$

$$u^{NRL}(a|k; \beta, \omega) = v^{RL}(a|k; \beta) + e^{\lambda_k} \epsilon(a)$$

$$u^{NRL-LS}(a|k; \beta, \omega) = v^{RL-LS}(a|k; \beta) + e^{\lambda_k} \epsilon(a)$$

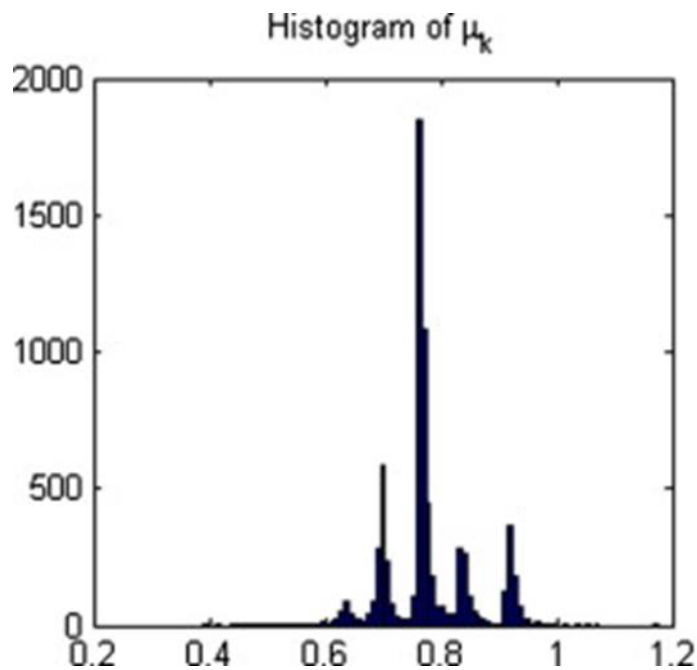
推定結果

Parameters	RL	NRL	RL-LS	NRL-LS
β_{TT}	-2.494	-1.854	-3.060	-2.139
Rob. Std. Err.	0.098	0.132	0.103	0.145
Rob. t-test(0)	-25.45	-14.05	-27.709	-14.75
β_{LT}	-0.933	-0.679	-1.057	-0.748
Rob. Std. Err.	0.03	0.043	0.029	0.047
Rob. t-test(0)	-31.10	-15.79	-36.448	-15.91
β_{LC}	-0.411	-0.258	-0.353	-0.224
Rob. Std. Err.	0.013	0.016	0.011	0.015
Rob. t-test(0)	-31.62	-16.13	-32.091	-14.93
β_{UT}	-4.459	-3.340	-4.531	-3.301
Rob. Std. Err.	0.114	0.2	0.126	0.207
Rob. t-test(0)	-39.11	-16.7	-35.960	-15.95
β_{LS}	-	-	-0.227	-0.155
Rob. Std. Err.	-	-	0.013	0.013
Rob. t-test(0)	-	-	-17.462	-11.92
ω_{TT}	-	0.515	-	0.341
Rob. Std. Err.	-	0.255	-	0.288
Rob. t-test(0)	-	2.02	-	1.18
ω_{LS}	-	-0.674	-	-0.581
Rob. Std. Err.	-	0.093	-	0.09
Rob. t-test(0)	-	-7.25	-	-6.46
ω_{OL}	-	-0.086	-	-0.092
Rob. Std. Err.	-	0.015	-	0.016
Rob. t-test(0)	-	-5.73	-	-5.75
$LL(\beta)$	-6303.9	-6187.9	-6045.6	-5952.0

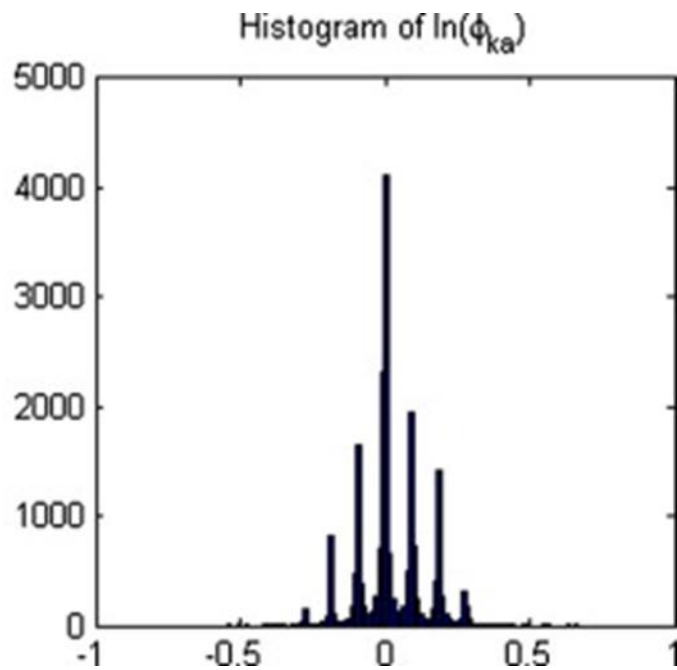
• LSを確定項に入れたほうが、尤度が高い
 ※リンクの共有具合によって、経路の効用を修正するために導入された変数だが、たとえ誤差項の相関を認めていたとしても、確定項にLSを加えたほうが説明力が上がる

• ω_{LS} と ω_{OL} には説明力があり、負の値
 ⇒フローが大きく、続くリンクが多いほど誤差項の分散が小さい

推定結果



μ_k は $OL(k)$ のため離散的なピークを取る
1を超えるものもわずかながらある



$\Phi_{ka} = \mu_a / \mu_k$ は0に関して対称な分布

将来予測の有効性

検証の仕方

- ① サンプルを2つに分ける
80%は推定に用いる
推定されたモデルによる予測確率を、残り20%で評価する
- ② 残しておいた20%を、同じ大きさの40個のサンプル集合に分ける
- ③ それぞれのサンプル集合に対し、当該サンプル集合以外のサンプルを使ってパラメータ $\hat{\beta}_i$ を推定し、以下の err_i を計算することで予測の誤差を測る

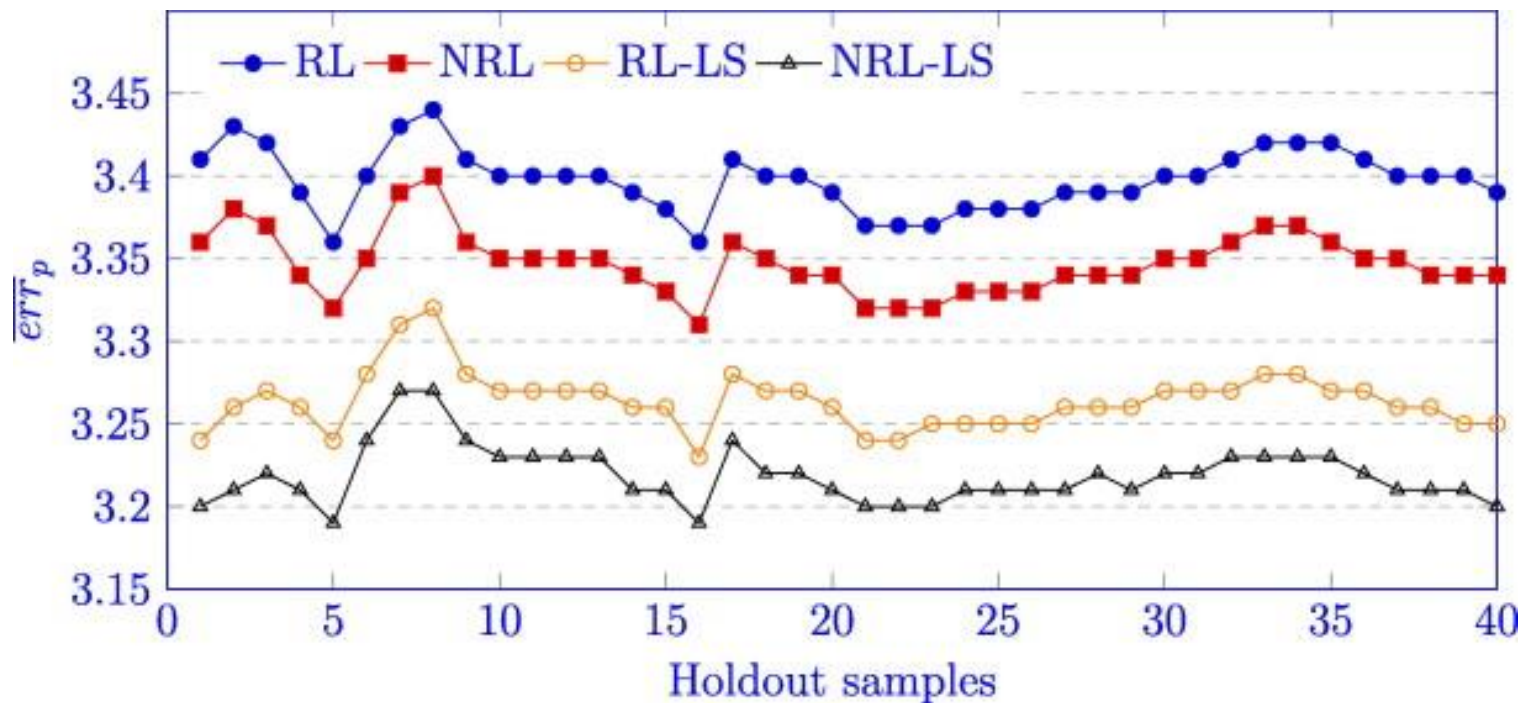
$$err_i = -\frac{1}{|PS_i|} \sum_{\sigma_j \in PS_i} \ln P(\sigma_j, \hat{\beta}_i)$$

尤度関数→負の値
⇒尤度が高いと絶対値が小さくなる
⇒ err_i は小さくなる

err_i はサンプル集合*i*に依存する確率変数なので、平均を用いる

$$\overline{err}_p = \frac{1}{p} \sum_{i=1}^p err_i \quad \forall 1 \leq p \leq 40$$

将来予測の有効性



LSなしよりもLSありのほうが、RLモデルよりもNRLモデルのほうが、予測精度が高い

まとめ

- NRLモデルの利点を確認
(経路列挙を行うことなく、経路の相関を考慮したうえで経路選択モデルを定式化している)
- 期待関数を求めるアルゴリズムを確認
- 実ネットワークへの適用を行い、説明力の向上と将来予測の有効性を確認