スタートアップゼミ#6

アクティビティモデル概論

M1 飯塚卓哉

目次

目標

アクティビティモデルの代表的なモデルに触れる・考え方を知る

はじめに

Activity based approach

「すべての移動は活動の派生需要である」

Activity based approach

：一日の時間（資源制約）を種々の活動とその派生的な移動に割り振る

＝ scheduling model

⇕

trip based approach

：四段階推定法のようなODベースの分析

Activity based approach 分類

1. ランダム効用最大化理論に基づく計量経済学モデル

   ：活動，場所，タイミング，活動時間，交通手段などの要素それぞれの意思決定を，個人の効用に誤差項を仮定し，確率的に決定

今日の内容

2. ルールベースドコンピュータプロセスモデル

   ：決められたルールの下で活動列を作成していくモデル

   ex） **ALBA-TROSS**, Arentze and Timmermans(2004)

   TASHA, Miller and Roorda

何が難しいか？

選択肢集合が膨大

- 活動
- 場所
- タイミング
- 活動時間
- 交通手段
- （物理的な）経路

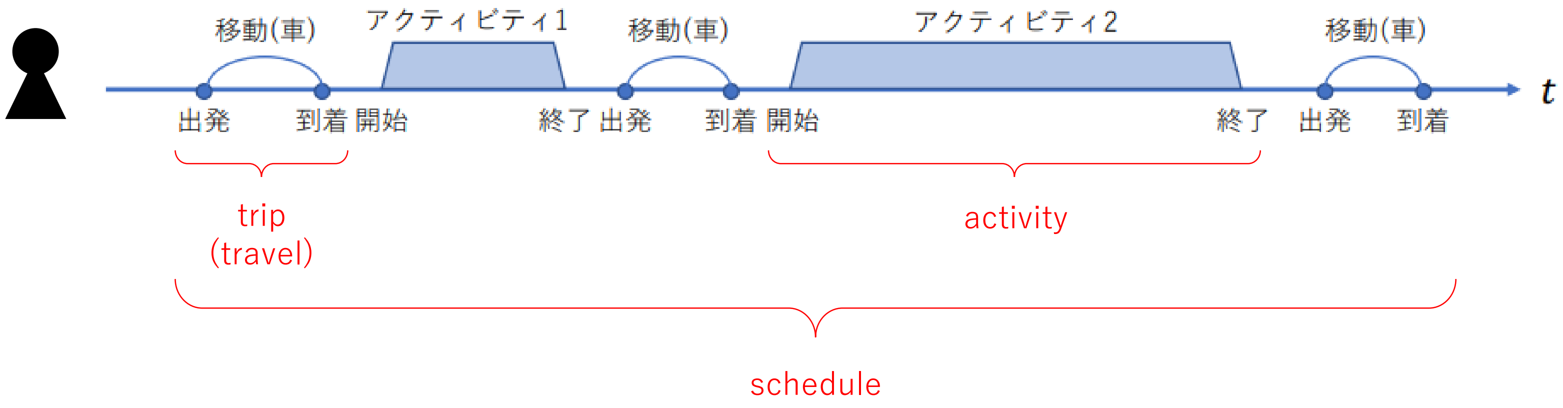
schedule = これらの組み合わせ

⇒ 選択肢集合が膨大という問題にいかに対処するかで,
様々なモデルが提案されてきた

何が難しいか？

意思決定のメカニズム

- pre-trip型：一日の活動が自宅を出発する前にすべて決定されることを仮定
- 逐次選択型：スケジュール内の１つ１つの活動について，活動内容（開始時刻，場所等）を順次決定

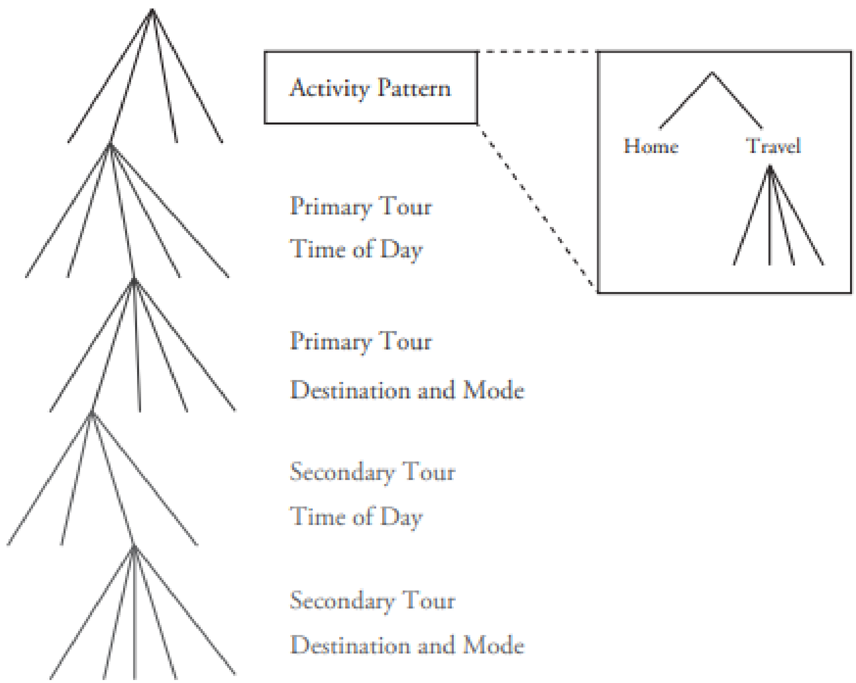


Bowman and Ben-Akiva(2001)

J.L bowman, M.E Ben-Akiva, Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A: Policy and Practice, Volume 35, Issue 1, January 2001, Pages 1-28

概要

- 離散選択モデルを用いたActivity-based model
- pre-trip型のモデル
  ：一日の活動が自宅を出発する前にすべて決定されることを仮定
- 活動パタンと，パタンを構成する主活動の開始時刻，場所，交通手段をネスト構造にして，Nested Logit (NL) モデルで一日の活動パタンの生成を定式化
  ＝ Hierarchical structure



定式化

- 一日の活動パタン（スケジュール）が選択される確率 $p(schedule)$

$$\boxed{p(schedule) = p(pattern)p(tours|pattern)}$$

特定の活動パタンを選択する確率

特定の活動パタンの下で，特定のツアーの集合を選択する確率

注意すること

- $p(pattern)$ と， $p(tours|pattern)$ は独立ではない
- 活動パタンの魅力（効用）はパタンに関連しているツアーによって得られる最大効用の期待値に依存している

パタンを特徴づける要素

A）一日中家にいるという選択肢も含む一日の活動の主要な活動

B）回数・目的・一連の活動停止を含むその日の代表活動ツアーのタイプ

C）二次ツアーの回数と目的

例

(a) 実際の旅程

7:30 AM	ゾーンAの家からゾーンBの仕事場まで一人で運転していった
Noon	ランチを食べに歩いて行き，仕事場に戻った
4:40 PM	仕事場を出て，ゾーンCの銀行に行った
5:00 PM	銀行を出て家に帰った
7:00 PM	家族と一緒にゾーンCのショッピングモールまで運転していった
10:00 PM	家に帰った

(b) モデル上の表現

活動パタン

主要な活動	work
主要ツアーのタイプ	home-work-other-work-other-home
二次ツアーの回数と目的	1 tour, purpose ’other’

主要ツアー

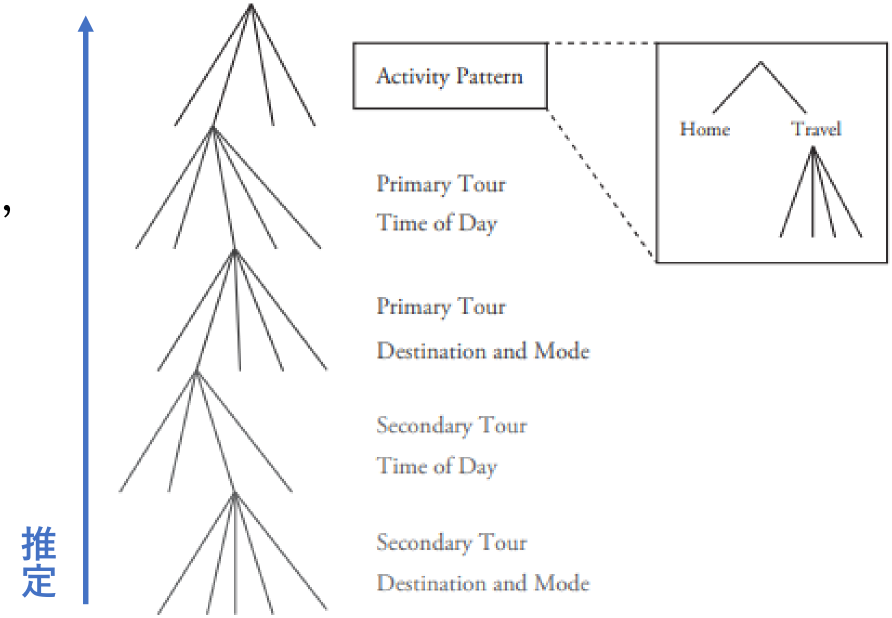
主要ストップ	場所：ゾーンB 手段：一人で運転 時間：AMピーク PMピーク
workベースのサブツアー	場所：ゾーンB 手段：徒歩 時間：midday midday
workストップの後	場所：ゾーンC 時間：PMピーク

二次ツアー

主要ストップ	場所：ゾーンC 手段：乗客を乗せて運転 時間：夕方 夕方

定式化

- 最下層から順に推定を行う
- 上層：一日中家 or 外出の二項選択
  下層：アクティビティパタンの組み合わせ，二次ツアーの回数と目的により分類した活動パタンのデータセットにより構成された選択肢
- 下層は上層によって条件づけられる



全ての旅行は主要ツアーと，0回以上の二次ツアーで構成
→ 二次ツアーは主要ツアーの結果に依存する条件付確率

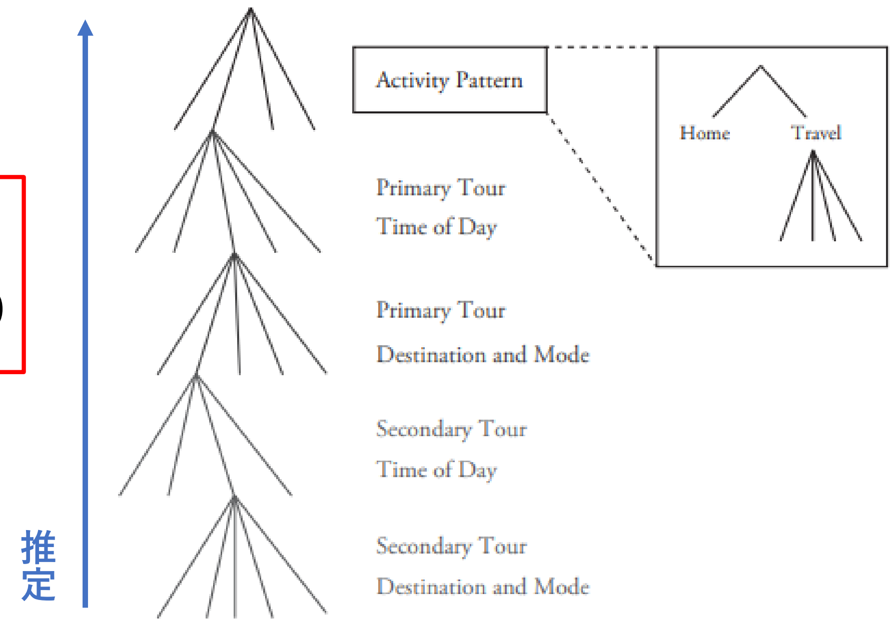
ツアーの選択確率

$$\boxed{\begin{aligned} p(tours|pattern) = p(primary\ tours|pattern) \times \\ p(secondary\ tours|primary\ tours) \end{aligned}}$$

定式化

二次ツアーは互いに独立だとすると，
二次ツアーの条件付確率は

$$p(secondary\ tours|primary\ tours) = \prod_{t=1}^{T} p(secondary\ tours_t|primary\ tours)$$



それぞれの確率を代入すると

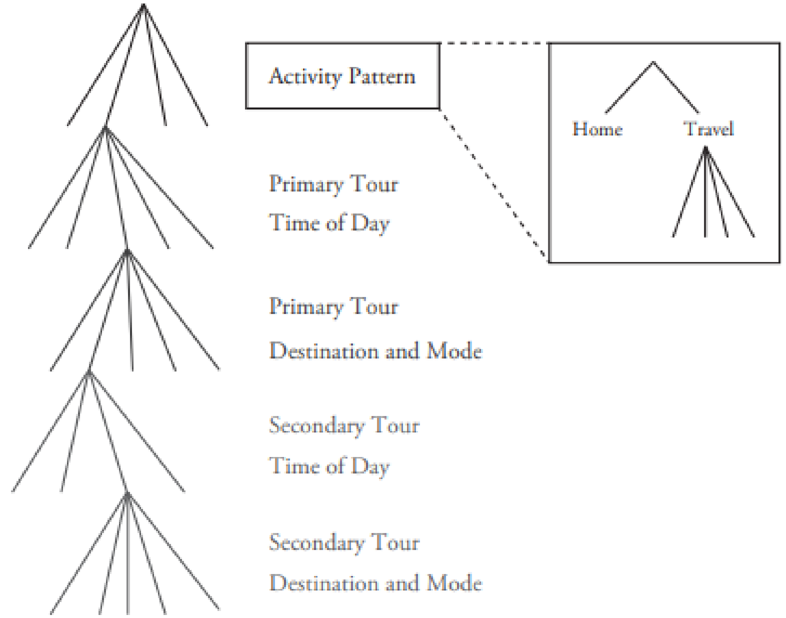
$$p(schedule) = p(pattern)p(primary\ tours|pattern) \times \prod_{t=1}^{T} p(secondary\ tours_t|primary\ tours)$$

→ 最下層から順に推定可能！

優れた点と課題

優れた点

- 主要ツアーと二次ツアーというある程度人間の意思決定メカニズムに即したモデルを単純なヒエラルキー構造で表せている



課題

- ツアーにの Time of day 選択
→ 一日を複数の時間帯に分割したものを選択肢とする
= 選択肢集合の膨大さ

- ツアーの場所と交通手段選択 ＝ 選択肢集合の膨大さ

- 活動選択の相関は一方向のヒエラルキー構造では表現しきれない

補足

Nested Logit (NL) Model

- 科研本や過去のゼミ資料を見ながら勉強してみてください
- スタートアップゼミ第3回で，渡したコードの中にNLモデルの推定コードもあるので見てみてください
- Bowman and Ben-Akiva型のアクティビティモデルの推定を去年の行動モデル夏の学校で東工大チームが行っているのでぜひ見てみてください

  （ホームページからコードもダウンロードできます）

離散連続モデル by Habib

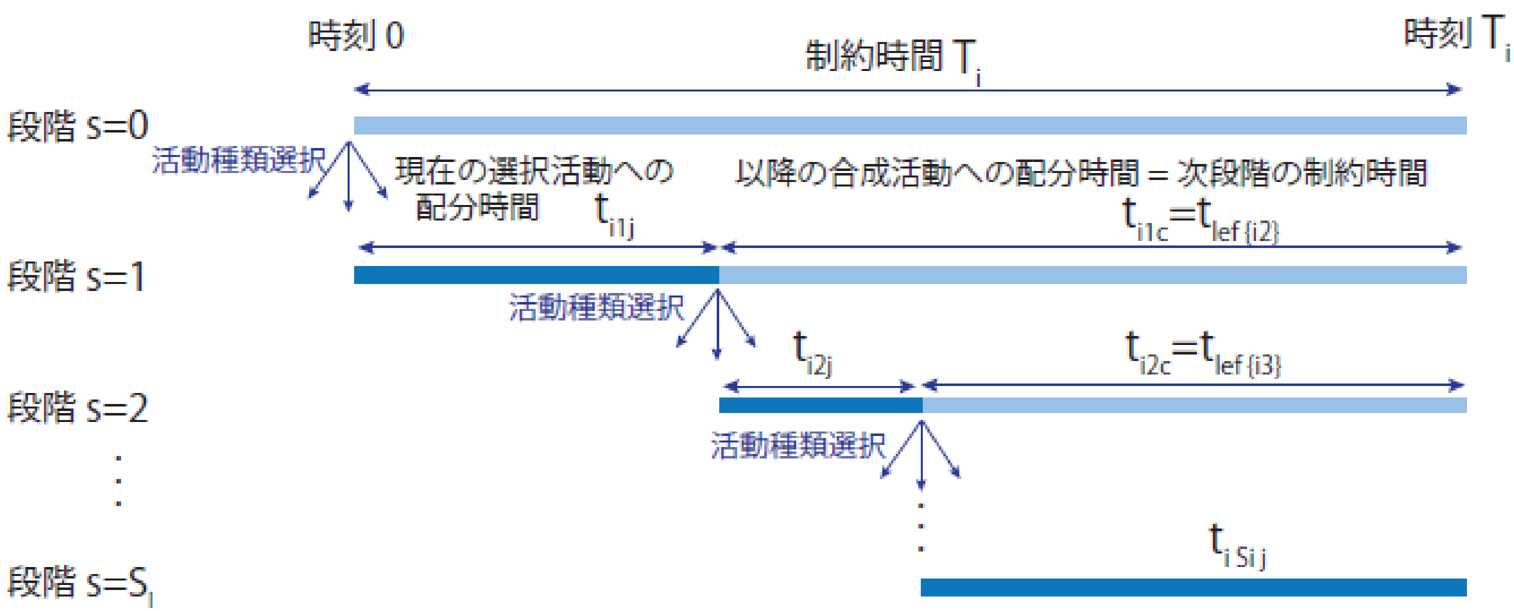
Khandker M. Nurul Habib, A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling, Transportation, January 2011, Volume 38, Issue 1, pp 123–151

概要

考え方

- 逐次的な活動決定＝一日の時間を各活動に順番に割り振っている
- 「どのような活動をするか」＝離散選択
- 「その活動にどれだけの時間を使うか」＝連続選択
- 「どのような活動をするか」と「その活動にどれだけの時間を使うか」は，相互に関係している

→ 逐次的離散連続モデル



定式化

離散選択の効用関数

活動選択肢 j を選択した場合

確定項　誤差項

$$U_j = V_j + \varepsilon_j$$
$$= \beta_j x_j + \varepsilon_j$$

連続選択の効用関数

消費所間 t_k を選択した場合

$$U(t_k) = \sum_{k=1}^{2} \frac{1}{\alpha_k} \exp(\psi_k z_k + \varepsilon'_k)(t_k^{\alpha_k} - 1)$$

$k = 1$：選択された活動
$k = 2$：利用可能な時間予算下の残り時間（合成活動）
α_k：飽和パラメータ(< 1)
ε'_k：効用の誤差項　～ガンベル分布
z_k：説明変数ベクトル
ψ_k：パラメータ

- $\psi_k z_k + \varepsilon'_k$：消費時間0におけるベースライン効用
- α_k：消費時間の増加に伴って限界効用が逓減していくことを表す

定式化

連続選択の効用関数

消費所間 t_k を選択した場合　$U(t_k) = \sum_{k=1}^{2} \frac{1}{\alpha_k} \exp(\psi_k z_k + \varepsilon_k')(t_k^{\alpha_k} - 1)$

→ 残り時間で行える活動を要約した概念 $k = 2$（Hicksian合成財と同義）とのトレードオフの中で活動 $k = 1$ に割り振る時間を効用最大化で決定

制約条件　$t_j + t_c = T$

t_j：現在の選択活動に配分された時間
t_c：残り時間
T：全体の時間

時間予算制約をラグランジュ関数を用いて連続選択の効用関数と統合：

$$l = \sum_{k=1}^{2} \frac{1}{\alpha_k} \exp(\psi_k z_k + \varepsilon_k') \, (t_k^{\alpha_k} - 1) - \lambda \left[\sum_{k=1}^{2} t_k - T \right]$$

Kuhn-Tucker条件から：

$$\exp(\psi_k z_k + \varepsilon_k') \, t_k^{\alpha_k - 1} - \lambda = 0, \quad \text{if } t_k > 0, \; k = 1, 2$$

$$\exp(\psi_k z_k + \varepsilon_k') \, t_k^{\alpha_k - 1} - \lambda < 0, \quad \text{if } t_k = 0, \; k = 1, 2$$

定式化

合成活動の時間配分は0でないとすると

$$\lambda = \exp(\psi_c z_c + \varepsilon_c')t_c^{\alpha_c - 1}$$

これを活動 j でのKKT条件式に代入すると

$$\exp(\psi_j z_j + \varepsilon_j')t_j^{\alpha_j - 1} = \exp(\psi_c z_c + \varepsilon_c')t_c^{\alpha_c - 1}$$

$$(\psi_j z_j + \varepsilon_j') + (\alpha_j - 1)\ln(t_j) = (\psi_c z_c + \varepsilon_c') + (\alpha_c - 1)\ln(t_c)$$

$$[\psi_j z_j + (\alpha_j - 1)\ln(t_j)] + \varepsilon_j' = [\psi_c z_c + (\alpha_c - 1)\ln(t_c)] + \varepsilon_c'$$

$$V_j' + \varepsilon_j' = V_c' + \varepsilon_c'$$

両辺の対数を取る

同様に，$t_j = 0$の場合 $V_j' + \varepsilon_j' < V_c' + \varepsilon_c'$

以上より，

$$\varepsilon_j' - \varepsilon_c' = V_c' - V_j' \quad \text{for} \quad t_j > 0$$

$$\varepsilon_j' - \varepsilon_c' < V_c' - V_j' \quad \text{for} \quad t_j = 0$$

連続選択の効用確定項

ただし，$V_k' = \psi_k z_k + (\alpha_k - 1)\ln(t_k), \quad k = j, c$

定式化

離散選択の選択確率

ランダム効用最大化理論では，活動 j の選択確率は

$$\begin{aligned} Pr(U_j > \max_{n=1,2,\ldots,A,n\neq j} U_n) &= Pr(V_j > [\max_{n=1,2,\ldots,A,n\neq j} U_n] - \varepsilon_j)) \\ &= Pr(V_j > (V_n + \varepsilon_n) - \varepsilon_j) \\ &= Pr(V_n < V_j + (\varepsilon_j - \varepsilon_n)) \end{aligned}$$

n：2番目に効用の高い活動

誤差項εに平均値0，スケールパラメータ1のi.i.d. ガンベル分布を仮定すると，ガンベル分布の誤差項の差はロジスティック分布に従うので，

$$Pr((\varepsilon_n - \varepsilon_j) < (V_j - V_n)) = \frac{\exp(V_j)}{\exp(V_j) + \sum_{n\neq j} \exp(V_n)}$$

$$\text{therefore} \quad Pr(\varepsilon_n < (V_j - V_n + \varepsilon_j)) = \frac{\exp(\gamma_j x_j)}{\exp(\gamma_j x_j) + \sum_{n\neq j} \exp(\gamma_n x_n)}$$

定式化

連続選択の選択確率

連続選択は選択した活動 j と，残りの合成活動 c に時間を配分することである.

連続選択の誤差項の分布を平均値0，スケールパラメータσのi.i.d. ガンベル分布とすると，その差はロジスティック分布に従うので,

確率密度関数（PDF） $Pr((\varepsilon_j' - \varepsilon_c') = (V_c' - V_j')) = \frac{1}{\sigma} \exp(\frac{-(V_c' - V_j')}{\sigma}) [1 + \exp \frac{-(V_c' - V_j')}{\sigma}]^{-2}$

累積分布関数（CDF） $Pr((\varepsilon_j' - \varepsilon_c') < (V_c' - V_j')+) = [1 + \exp \frac{-(V_c' - V_j')}{\sigma}]^{-1}$

$$\text{therefore} \quad Pr(\varepsilon_j' < (V_c' - V_j + \varepsilon_c')) = [1 + \exp \frac{-(V_c' - V_j')}{\sigma}]^{-1}$$

効用の確定項は以下のように定義されることが分かっている（Bhat(2008)）

$$V_j' = \psi_j z_j + (\alpha_j - 1)\ln(t_j)$$

$$V_c' = (\alpha_c - 1)\ln(t_c)$$

定式化

変数変換の公式を用いると確率密度関数（PDF）は以下のように変形される

$$
\begin{aligned}
Pr(t = t_j) &= (\frac{\delta((V_c' - V_j'))}{\delta t_j})\frac{1}{\sigma}\exp(\frac{-(V_c' - V_j')}{\sigma})[1 + \exp\frac{-(V_c' - V_j')}{\sigma}]^{-2} \\
&= (\frac{1-\alpha_j}{t_j} + \frac{1-\alpha_c}{t_c})\frac{1}{\sigma}\exp(\frac{-(V_c' - V_j')}{\sigma})[1 + \exp\frac{-(V_c' - V_j')}{\sigma}]^{-2}
\end{aligned}
$$

誤差項にi.i.d. ガンベル分布を仮定することで，離散選択，連続選択
それぞれの選択確率を定式化できた

定式化

活動選択と時間配分の確率統合

活動選択と時間配分の同時確率を求めたい → その際，相関を考慮したい

but：スケールパラメータσを，推定される係数パラメータと区別することはできない

→ スケールパラメータを$\sigma = 1$にする（相関が考慮できない）

誤差項は両方とも特定の周辺分布を持っているので，これを標準正規分布の逆関数Φ^{-1}を用いて標準正規分布に変換することができる

$$\varepsilon_j^* = J_1(\varepsilon_j) = \Phi^{-1}\left[(\varepsilon_n - \varepsilon_j) < (V_j - V_n)\right]$$
$$\varepsilon_k^* = J_2(\varepsilon_j') = \Phi^{-1}\left[(\varepsilon_j' - \varepsilon_c') < (V_j' - V_c')\right]$$

定式化

活動 j と配分時間 t_j の同時選択確率は，変換された標準正規分布 $J_1(\varepsilon_j)$ と $J_2(\varepsilon_j')$ と，その相関係数 ρ_{jt} によって定まる2変量正規分布 $BVN[J_1(\varepsilon_j),\ J_2(\varepsilon_j'),\ \rho_{jt}]$ に従う．

その確率は

$$
\begin{aligned}
&Pr(\text{time} = t_j \cap \text{Activity type} = j)\\
&= Pr(t = t_j \cap \varepsilon \le J_1(\varepsilon_j))\\
&= (\frac{1-\alpha_j}{t_j} + \frac{1-\alpha_c}{t_c})\frac{1}{\sigma}\exp(\frac{-(V_c' - V_j')}{\sigma})[1+\exp\frac{-(V_c' - V_j')}{\sigma}]^{-2}\Phi(\frac{J_1(\varepsilon_j) - \rho_{jt}J_2(\varepsilon_j')}{\sqrt{1-\rho_{jt}^2}})
\end{aligned}
$$

この時，尤度関数は

$$
\begin{aligned}
L_i = \prod_{j=1}^{n}((\frac{1-\alpha_{ji}}{t_{ji}} + \frac{1-\alpha_{ci}}{t_{ci}})\frac{1}{\sigma}\exp(\frac{-(V_{ci}' - V_{ji}')}{\sigma})[1+\exp\frac{-(V_{ci}' - V_{ji}')}{\sigma}]^{-2}\\
\times\Phi(\frac{J_1(\varepsilon_{ji}) - \rho_{jt}J_2(\varepsilon_{ji}')}{\sqrt{1-\rho_{jt}^2}}))^{D_{ji}}
\end{aligned}
$$

D_{jt}：選択した活動で1をとるダミー変数

適用例

同時選択確率と尤度関数が定式化できたので，活動データがあれば推定可能

活動

1. 基盤的な必要活動：睡眠，選択，ごはん，...
2. 仕事，学校
3. 世帯の仕事
4. 送迎や届け物
5. 買い物
6. サービス：病院，銀行，自動車整備，...
7. 余暇活動
8. 社会的活動：バー，長電話，スポーツ，...
9. そのほか

1日の活動を以上の9カテゴリの活動に分類
→24時間のタイムフレーム内での配分として推定

Table 1 Estimated parameters of weekend activity scheduling model

Variables	Activity type	Parameter	t-Statistics
Activity type choice model component			
Constant			
	Household obligations	1.3574	1.485
Star time in hours from mid-night			
	Drop-off/pick-up	−0.3584	−8.663
	Shopping	−0.1556	−2.921
	Services	−0.3429	−5.808
	Recreation/entertainment	0.0123	0.807
Number of activities already performed from beginning of the day			
	Basic needs	−0.0687	−4.894
	Drop off/pick up	0.2110	3.915
	Shopping	0.1746	3.008
	Services	0.2178	2.836
	Social	0.0721	2.466
Total travel time (minutes)			
	Basic needs	−0.1755	−9.496
	Work/school	−0.0221	−1.529
	Household obligations	−0.1478	−8.324
	Recreation/entertainment	−0.1571	−7.886
	Other	−0.0682	−3.095
Household size: number of people in the household			
	Work/school	−0.1775	−1.993
	Recreation/entertainment	−0.1281	−2.531
	Social	−0.2404	−2.406
	Other	−0.2591	−2.516
Logarithm of age in years			
	Work/school	−0.5769	−6.688
	Household obligations	−0.5350	−2.21
	Shopping	−0.4803	−4.333
	Recreation/entertainment	−0.2723	−3.539
	Social	−1.0267	−5.233
	Other	−0.3649	−2.709
Logarithm of yearly income in Canadian dollars (2002–2003)			
	Household obligations	−0.0503	−3.152
	Services	−0.1296	−5.294
	Social	0.0332	0.868
	Other	−0.0646	−1.916
Employment status: non-full time job			
	Shopping	−0.2498	−0.893
	Recreation/Entertainment	0.2199	1.164
	Social	0.9300	2.555
Number of children in household			
	Household obligations	0.0850	1.553

Table 1 continued

Variables	Activity type	Parameter	t-Statistics
	Drop off/pick up	0.0929	1.273
	Shopping	−0.4030	−3.225
Time expenditure model component			
Variance (Sigma)		0.4952	26.354
Constant			
	Basic needs	0.8195	1.273
	Social	1.8545	−3.225
	Other	3.9394	26.354
Star time in hours from mid night			
	Work/school	−0.0248	−0.923
	Household obligations	0.1071	6.749
	Drop off/pick up	0.0729	2.31
	Shopping	0.0682	2.955
	Services	0.0602	2.483
	Recreation/entertainment	0.0375	0.954
Number of activities already performed from beginning of the day			
	Basic needs	0.0471	3.456
	Social	−0.0343	−1.184
Total travel time (minutes)			
	Household obligations	−0.0170	−1.558
	Drop off/pick up	0.0275	1.926
	Shopping	0.0382	2.832
	Services	0.0249	1.321
	Social	0.0295	2.962
	Other	0.0340	2.567
Household size: number of people in the household			
	Work/school	0.0557	0.727
	Household obligations	0.2049	4.758
	Recreation/entertainment	0.0868	2.471
Logarithm of age in years			
	Work/school	0.4589	3.005
	Drop off/pick up	−0.3269	−2.597
	Recreation/entertainment	0.4949	3.882
	Social	−0.2097	−0.753
	Other	−0.6673	−1.329
Logarithm of yearly income in Canadian dollars (2002–2003)			
	Work/school	0.0466	1.674
	Household obligations	0.0207	1.64
Number of automobile in household			
	Household obligations	−0.1242	−1.224
Logarithm of duration (years) of living in the city			
	Household obligations	−0.1306	−2.655
	Recreation/entertainment	−0.1013	−1.875

推定結果の例

Table 1 continued

Variables	Activity type	Parameter	t-Statistics
	Other	0.2185	1.573
Number of children in household			
	Household obligations	−0.1027	−2.138
Satiation parameter			
Constant			
	Household obligations	−0.1383	−2.36
	Drop off/pick up	−0.0427	−0.42
	Shopping	−0.1093	−1.21
	Recreation/entertainment	−0.4619	−5.51
	Social	−0.2111	−1.83
	Other	−0.2455	−2.09
	Composite activity	1.2654	9.99
Continuous start time in hours from mid night			
	Recreation/entertainment	0.0234	3.74
	Social	0.0173	2.45
Start hour: time of the day			
	Composite activity:		
	Before 6 AM	−0.8773	−8.40
	6:01 AM to 7 AM	−0.2232	−2.13
	7:01 AM to 8 AM	−0.1707	−1.82
	8:01 AM to 9 AM	−0.1693	−1.90
	9:01 AM to 10 AM	−0.2023	−2.32
	10:01 AM to 11 AM	−0.2079	−2.45
	11:01 AM to 12 noon	−0.2214	−2.74
	12:01 noon to 1 PM	−0.1891	−2.46
	1:01 PM to 2 PM	−0.1500	−1.99
	2:01 PM to 3 PM	−0.0514	−0.70
	3:01 PM to 4 PM	−0.0714	−1.03
	4:01 PM to 5 PM	−0.0712	−1.06
	After 5 PM	–	–
Correlation coefficient between activity type choice and time expenditure			
Constant		−0.4423	−3.43
Loglikelihood of full model			−8595.7618
Loglikelihood of constant-only model			9877.646
Adjusted Rho-square value			0.12

優れた点と課題

優れた点

- 逐次的な活動選択を記述可能
- 時間を離散化する必要なく，連続量として扱うことができる

課題

- 場所の選択も組み込む場合，選択肢集合が膨大になる問題
- 活動間の移動についてどう扱うか
- 活動の時間制約は，実際はあとやりたい活動がどれだけ残っているかに依存する

→ 人間の活動選択は必ずしも逐次的ではないのでは？

Recursive Logit (RL) model

Fosgerau, M., Frejinger, E., & Karlstrom, A. (2013). A link based network route choice model with unrestricted choice set. Transportation Research Part B: Methodological, 56, 70-80.

概要

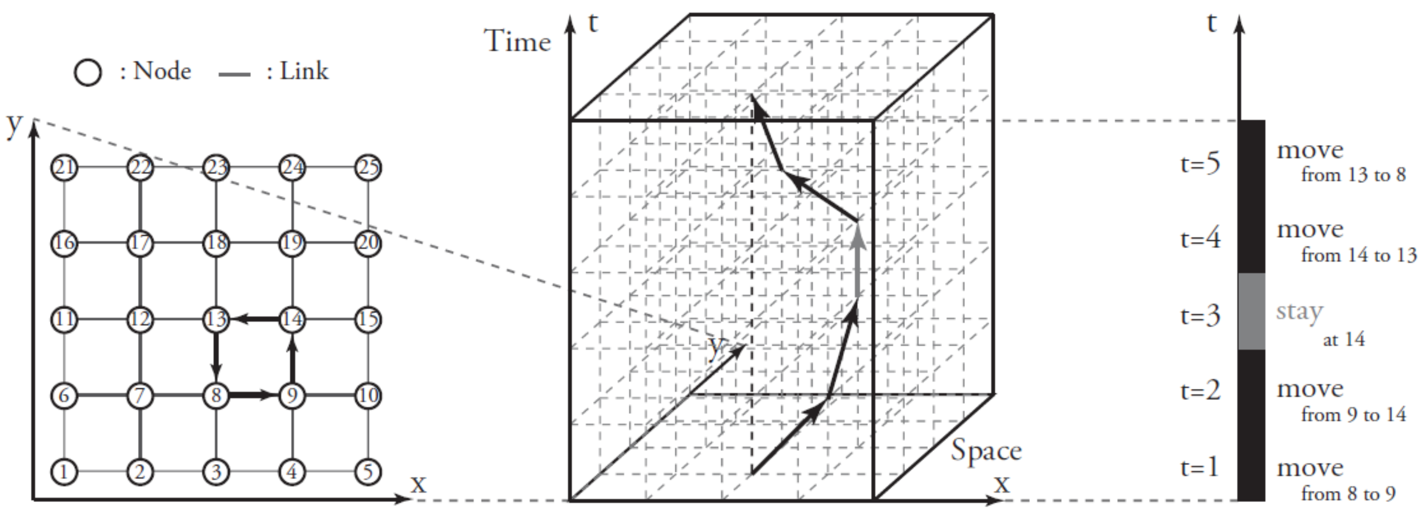
Recursive Logit (RL) model：

<u>経路選択肢を列挙することなく</u>，逐次的なリンクの選択によって経路選択行動を記述するモデル　→ **選択肢集合が膨大**という問題をクリア

考え方

- Time-Space NWにおいて時間軸方向の移動（＝滞在）を明示的に考える

　→ Time-Space NW上の経路選択は活動経路選択とみなせる（サイクリック構造を持たない）



大山・羽藤（2016）移動軌跡情報に基づく時間構造化ネットワーク上の交通配分，第53回土木計画学研究発表会・講演集（CD-ROM）より抜粋.

概要

時空間（プリズム）制約

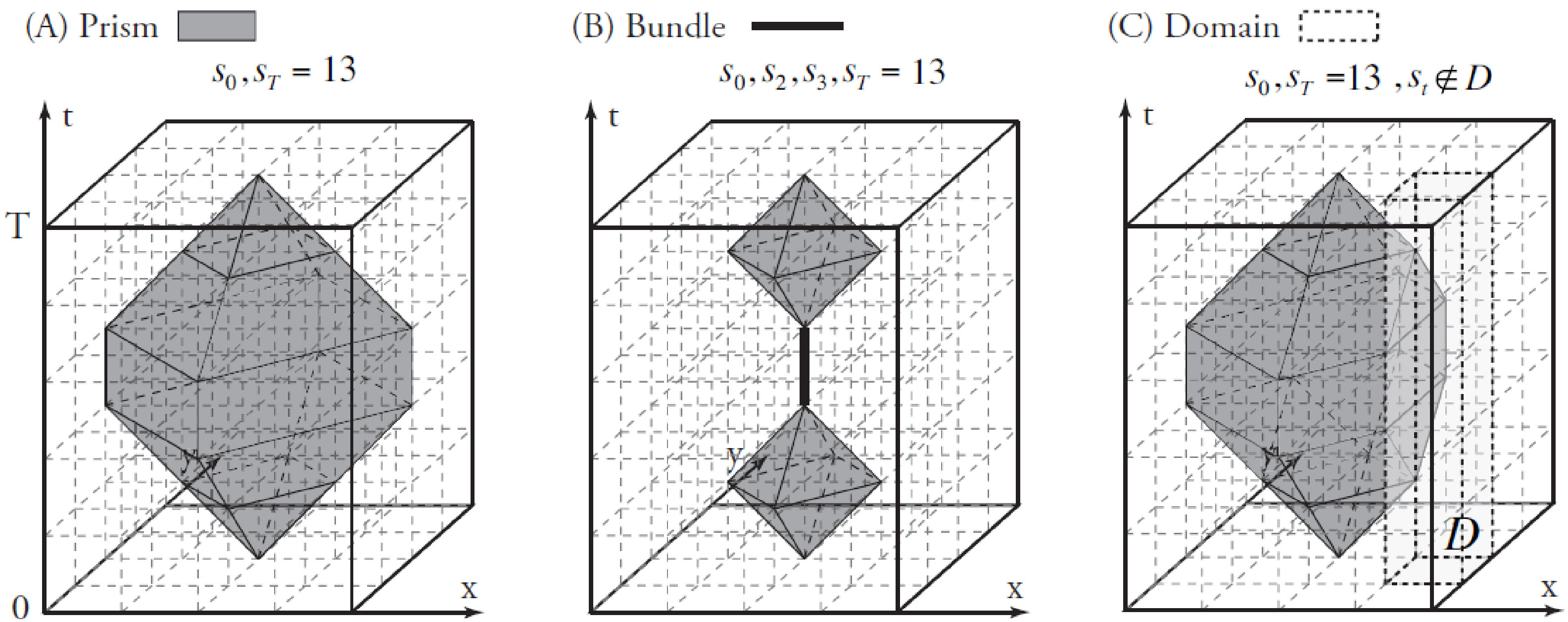


図 2　時空間制約による活動経路集合の限定

遠すぎるところには制約時間内に行けない

ある時間に特定の場所にいなければならない（仕事など）

ある時間に特定の範囲内には入れない（店舗の営業時間など）

大山・羽藤（2016）時空間制約と経路相関を考慮した歩行者の活動配分問題，都市計画学会論文集, Vol. 51, No. 3, pp. 680-687 より抜粋.

概要

さらに考えるべきこと

- 活動経路選択 ＝ 多次元の同時選択

  ex）活動内容，活動場所，活動開始時刻，継続時間，交通手段

  → 多次元の選択の統一的記述が必要

- 活動経路における経路の”相関”とは？

  → 単純な経路のフィジカルな重なりだけではない

- 計算時間の問題

  → RLモデルの推定には計算時間の問題がしばしば発生する

明日の理論談話会で詳しく説明します

まとめ

- アクティビティモデルは色々ある
- 現状のモデルの課題を理解してモデルを改良していく（より人間の行動を当てれるようにする）ことも大切
- 自分の分析・評価したい問題に適したアクティビティモデルを選ぶ（つくる）ことも大切