

理論談話会#?

Capturing correlation with a mixed recursive logit model for activity-travel scheduling

Maëlle Zimmermann, Oskar Blom Västberg, Emma Frejinger, Anders Karlström,
Transportation Research Part C: Emerging Technologies, Vol.93, August 2018, pp. 273-291

M1 飯塚卓哉

目次

1. Introduction
2. Literature review
3. Methodology
4. Application
5. In-sample fit and predictions
6. Conclusion

論文概要

- 時空間NW上の経路選択としてactivity-travel schedulingを記述するモデルを改良
- activity, location, timing, transportation mode の選択を表現
- mixed recursive logitモデルを構築することでIIA特性を緩和
- 現実的な時間内で推定可能

1. Introduction

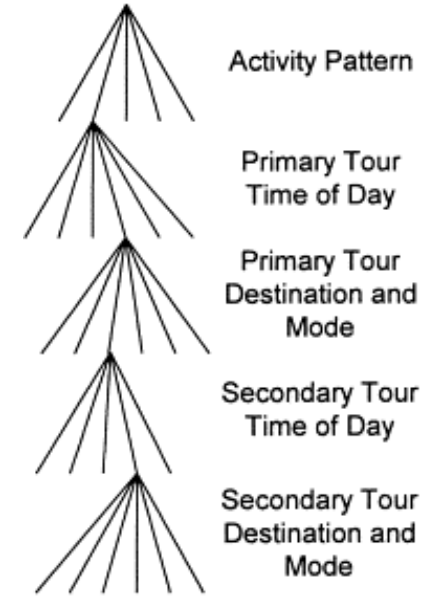
- 「**移動は連続的な time interval 上での活動の派生需要**」
= アクティビティベースのモデルとは、一日の時間（資源制約）を各活動と派生的な移動に割り振るという考え方
- **選択肢集合が膨大**という課題
 - 従来の解き方：各次元の選択をMNLやNLに分解 → 統一的でない
 - 統一的な表現：**multi-state network** = 活動，交通手段などの状態のNW
 - 選択肢集合を必要としないモデル：**RLモデル**

→この組み合わせで書ける
- RLモデルの課題
 - IIA特性
 - 膨大な推定時間

2. Literature review

ヒエラルキー構造を持ったアクティビティモデル

- アクティビティへの参加, タイミング, 場所, 交通手段などの選択をネスト構造にし, NLモデルで推定
- ネストの組み方で様々なバリエーションがある
- 選択枝数が膨大になってしまうのを防ぐために単純化が必要
- アクティビティ選択は複雑な相関構造を持っているので, ネストの構造を実際の行動に即して定義することは困難
- 複雑な多層のネスト構造は, モデル推定が困難



Bowman and Ben-Akiva(2001)

2. Literature review

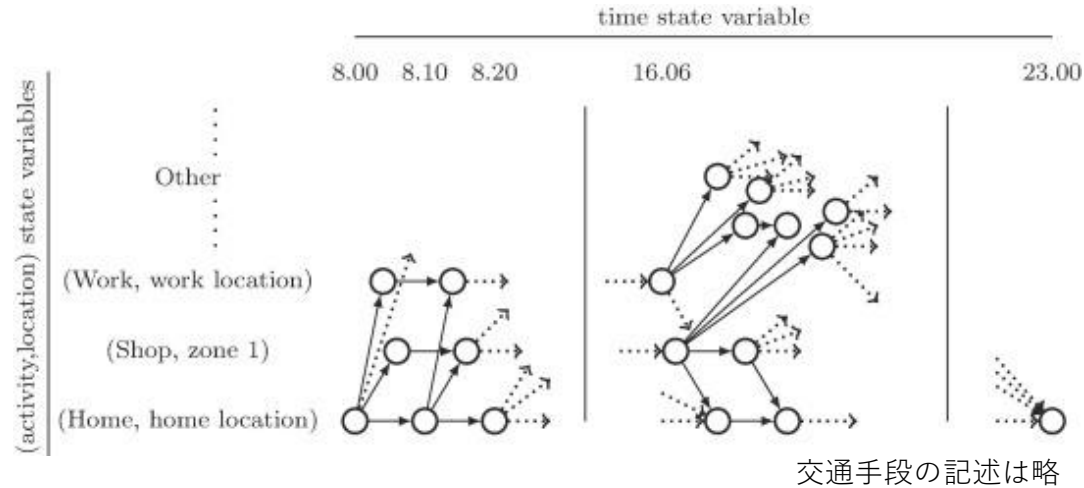
RLによるアクティビティモデル

Activity network

$$G = (\mathcal{A}, \mathcal{V})$$

リンク：交通手段，
次のノードに対応

ノード：時刻，活動，
場所に対応



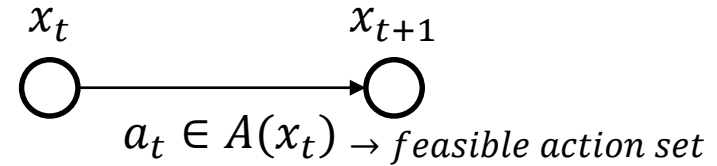
- 1タイムステップ=1分
- アクティビティ = 基本10分継続，水平リンク = さらに10分延長
- 移動時間は10分の倍数でなくても良い（1分の倍数）
- 個人ごとの時空間制約を考慮
- Home → … → Home の移動を扱う

2. Literature review

RLによるアクティビティモデル

Modeling framework

time step $t = 1, \dots, T$



Activity path $\mathbf{a} = (a_0, \dots, a_{T-1})$

- 個人は直近の効用 $u_n(a_t|x_t) = v_n(a_t|x_t) + \mu\epsilon(a_t)$ と将来の期待効用の和を最大化

Bellman方程式の形で再帰的に記述可能

$$V_n(x_t) = E\left(\max_{a_t \in A(x_t)} \{v_n(a_t|x_t) + \mu\epsilon(a_t) + V_n(x_{t+1})\}\right), \quad \text{将来の期待効用} = \text{価値関数}$$

- 誤差項 $\epsilon(a_t)$ に平均値0のi.i.dガンベル分布を仮定すると

$$V_n(x_t) = \mu \log \left(\sum_{a_t \in A(x_t)} e^{\frac{1}{\mu}(v_n(a_t|x_t) + V_n(x_{t+1}))} \right).$$

ガンベル分布における「最大値の期待値」 = ログサム

2. Literature review

RLによるアクティビティモデル

Modeling framework

- 状態 x_t におけるリンク a_t の選択確率はMNL

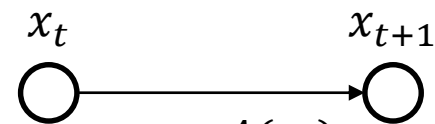
$$P_n(a_t|x_t) = \frac{e^{v_n(a_t|x_t)+V_n(x_{t+1})}}{\sum_{k_t \in A(x_t)} e^{v_n(k_t|x_t)+V_n(x_{t+1})}}$$

→ 前ページの式を用いて変形して, $P_n(a_t|x_t) = e^{v_n(a_t|x_t)+V_n(x_{t+1})-V_n(x_t)}$.

- アクティビティパス $\mathbf{a} = \{a_t\}_{t=0}^{T-1}$ の選択確率は

$$\begin{aligned} P_n(\mathbf{a}|x_0) &= \prod_{t=0}^{T-1} e^{v_n(a_t|x_t)+V_n(x_{t+1})-V_n(x_t)} \\ &= \frac{e^{v_n(\mathbf{a}|x_0)}}{e^{V_n(x_0)}}, \quad \leftarrow \text{MNLに帰着} \end{aligned}$$

time step $t = 1, \dots, T$



Activity path $\mathbf{a} = (a_0, \dots, a_{T-1})$

2. Literature review

RLによるアクティビティモデル

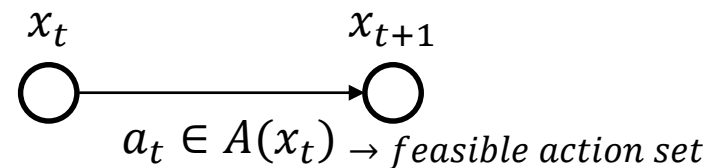
利点

- マルコフ配分をすればいいのでシミュレーションが容易
- 価値関数がログサム形で表されるので政策評価も容易
- pre-trip型のモデルと異なり、予期しないイベントへのreschedulingがシミュレーション可能
- 選択肢集合の限定が不要

欠点

- MNLモデルと等価であるため、IIA特性がある
→ この緩和：**mixed RL モデル**

time step $t = 1, \dots, T$



Activity path $\mathbf{a} = (a_0, \dots, a_{T-1})$

3. Methodology

Challenges

Activity network

→ ノード数, リンク数, パラメータ数が多くなりがち

→ 個人ごとにNWの制約が異なる

→ 計算量が課題

∴ RLの推定 = NFXP法: 現在のパラメータ試行値に対して価値関数とその微分を, 状態・個人ごとに計算する必要がある

例) 40,000リンク, 説明変数15個の自転車ネットワークにおけるNRLモデル推定

→ 2週間 Zimmermann et al. (2017)

もし, 1000個の場所, 8つの活動, 4つの交通手段 → 最大32,000,000リンク

時間を60時点に離散化したとしたら → 1,920,000,000リンク

3. Methodology

Challenges

- 経路選択モデルにおける相関 = physicalな経路の重なり
- アクティビティパス選択モデルにおける相関 = さまざまな相関
 - 同じ交通手段を使うことの相関
 - (異なる時間に) 同じ活動内容を行うことの相関
 - 交通手段と活動内容の相関

etc...

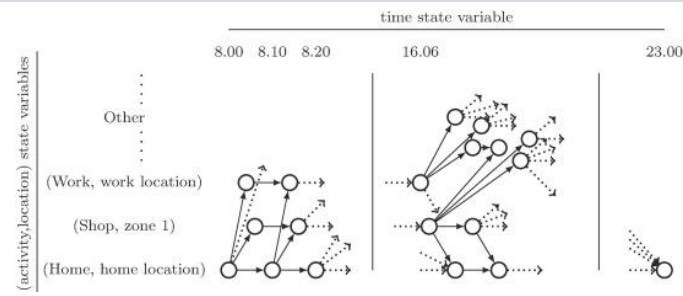
→ 単純なネスト構造では記述不可能

3. Methodology

Mixed Recursive Logit (MRL) model

- リンクの選択 = 活動, 場所, 交通手段の選択
 $a_t = (p, l, m)$

交通手段を例に相関の記述を説明 (個人nの添え字を略)



- 直近の効用 $u(a_t|x_t) = v(a_t|x_t) + \zeta_m(a_t)$
 $= v'z_m + \epsilon(a_t)$

$\epsilon(a_t)$: i.i.dガンベル分布に従う誤差項
 z_m : 交通手段 $m' \in \{1, \dots, M\}$ について $m = m'$ のとき $z_{m'} = 1$ となるダミー変数
 v : 平均0, 分散共分散行列 Σ の正規分布に従う M 次元のランダムベクトル

他の次元の選択肢 $k \in \{1, \dots, K\}$ が加わった場合

- $\zeta_m(a_t) = v'z_m + \eta'y_k + \epsilon(a_t)$ $M + K$ 次元ベクトル $[v; \eta]$ は分散共分散行列 Σ に従う
- 場所の選択など, 選択肢数が多い時 → 隣接点を集約してランダムパラメータ数を制限する

- パラメータベクトル β_m が正規分布 $N(\bar{\beta}_m, \Sigma)$ に従うという解釈も可能
 → この時の $v_n =$ 個人ごとの平均値からの偏差. つまり, $\beta_{m,n} = \bar{\beta}_m + v_n$

効用確定項

3. Methodology

MRLモデルの特性

$$\Sigma = \begin{pmatrix} \sigma_1^2 & & \cdots & & \sigma_{1,M} \\ & \ddots & & & \\ \vdots & & \sigma_m^2 & & \vdots \\ & & & \ddots & \\ \sigma_{M,1} & & \cdots & & \sigma_M^2 \end{pmatrix}$$

- σ_m^2 : 同じ交通手段を使うことによる相関
- $\sigma_{m,m'}$: 異なる交通手段間の相関
- $\sigma_{m,p}$: 異なる次元間の相関は非対角項により表現

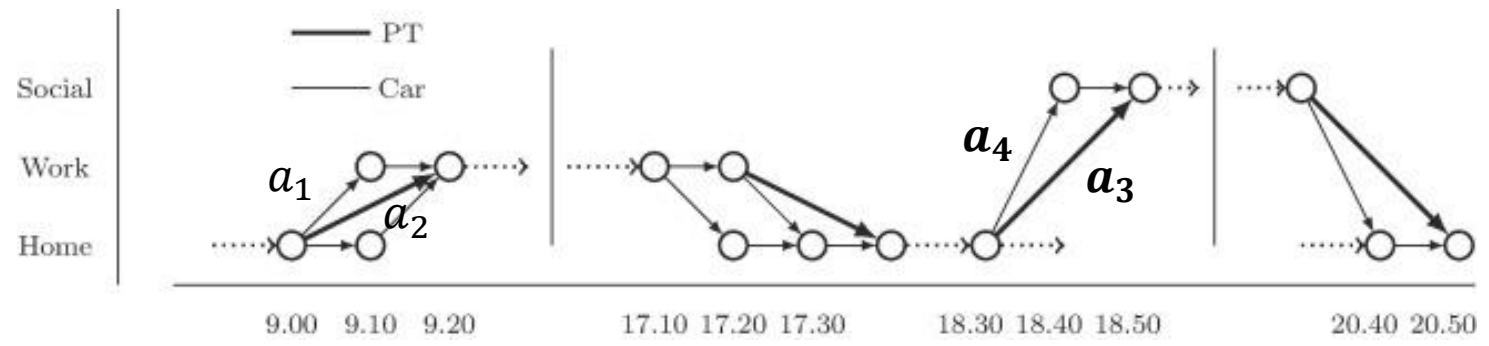
- 経路の効用 $u(\mathbf{a}|x_0) = \sum_{t=0}^{T-1} \{v(a_t|x_t) + \zeta(a_t)\}$

→ $\sum_{t=0}^{T-1} \zeta(a_t)$ の項は、一日の異なる時間帯での類似した行動に関する共分散を生成

→ アクティビティパス間の相関も記述可能

3. Methodology

■ 実例



- car, public transport (PT), social activity の3つの誤差項を定義
 = 3つの選択肢固有定数項 ($ASC_{car}, ASC_{PT}, c_{social}$) を考え, これが平均値 ($\bar{\beta}_{car}, \bar{\beta}_{PT}, \bar{\beta}_{social}$),
 分散共分散行列 $\Sigma = \begin{pmatrix} \sigma_{car}^2 & 0 & 0 \\ 0 & \sigma_{PT}^2 & 0 \\ 0 & 0 & \sigma_{social}^2 \end{pmatrix}$ を持つことと同値
- $x = 9:00$ に家にいる状態, $a_1 = car$ で仕事に行く活動, $a_2 = PT$ で仕事に行く活動とすると
 効用確定項は

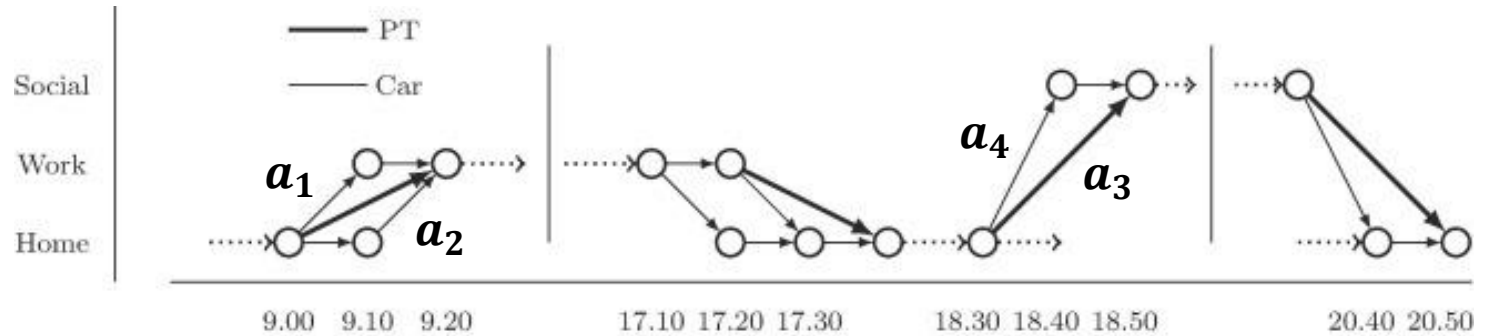
$$v_n(a_1|x) = \beta^T X_n(a_1|x) + ASC_{car,n} = \beta^T X_n(a_1|x) + \bar{\beta}_{car} + v_{car,n},$$

$$v_n(a_2|x) = \beta^T X_n(a_2|x) + ASC_{PT,n} = \beta^T X_n(a_2|x) + \bar{\beta}_{PT} + v_{PT,n},$$

$X_n(a|x)$: 説明変数ベクトル

3. Methodology

■ 実例



- 4つのアクティビティパスの効用

$$u_n(\mathbf{a}_1|x_0) = \beta^T X_n(\mathbf{a}_1|x_0) + 2\bar{\beta}_{car} + 2\nu_{car,n} + \epsilon_n,$$

$$u_n(\mathbf{a}_2|x_0) = \beta^T X_n(\mathbf{a}_2|x_0) + 2\bar{\beta}_{car} + 2\nu_{car,n} + \epsilon_n,$$

$$u_n(\mathbf{a}_3|x_0) = \beta^T X_n(\mathbf{a}_3|x_0) + 4\bar{\beta}_{PT} + 4\nu_{PT,n} + \bar{\beta}_{social} + \nu_{social,n} + \epsilon_n,$$

$$u_n(\mathbf{a}_4|x_0) = \beta^T X_n(\mathbf{a}_4|x_0) + 4\bar{\beta}_{car} + 4\nu_{car,n} + \bar{\beta}_{social} + \nu_{social,n} + \epsilon_n,$$

- 4つのアクティビティパスの分散共分散行列 M

$$M = \begin{pmatrix} 4\sigma_{car}^2 & 4\sigma_{car}^2 & 0 & 8\sigma_{car}^2 \\ 4\sigma_{car}^2 & 4\sigma_{car}^2 & 0 & 8\sigma_{car}^2 \\ 8\sigma_{car}^2 & 0 & 16\sigma_{PT}^2 + \sigma_{social}^2 & \sigma_{social}^2 \\ 8\sigma_{car}^2 & 0 & \sigma_{social}^2 & 16\sigma_{car}^2 + \sigma_{social}^2 \end{pmatrix}.$$

3. Methodology

MRLの最尤推定法

NFXP法：価値関数の計算とパラメータの更新を繰り返す → 今回は使わない
代替法 → **選択肢サンプリング**

- θ を混合分布 $f(\beta|\theta)$ のパラメータとする = 今回は θ は β の正規分布の平均値と標準偏差
- サンプル \tilde{C}_n の下でのアクティビティパス \mathbf{a} の選択確率は

$$P_n(\mathbf{a}|\tilde{C}_n, \theta) = \int \left(\frac{q(\tilde{C}_n|\beta)}{q(\tilde{C}_n|\theta)} \right) \frac{e^{u(\mathbf{a}|\beta) + \log(q_n(\tilde{C}_n|\mathbf{a}))}}{\sum_{j \in \tilde{C}_n} e^{u(j|\beta) + \log(q_n(\tilde{C}_n|j))}} f(\beta|\theta) d\beta.$$
$$= W_n$$

- アクティビティパスの観測数 = N : $\{\mathbf{a}\}_{n=1, \dots, N}$ のとき対数尤度関数は

$$\mathcal{L} = \sum_{n=1}^N \log \int W_n \frac{e^{u(\mathbf{a}_n|\beta) + \log(q_n(\tilde{C}_n|\mathbf{a}_n))}}{\sum_{\mathbf{a} \in \tilde{C}_n} e^{u(\mathbf{a}|\beta) + \log(q_n(\tilde{C}_n|\mathbf{a}))}} f(\beta|\theta) d\beta.$$

$$W_n = \frac{\sum_{\mathbf{a} \in \tilde{C}_n} P_n(\mathbf{a}|\beta, C_n) q(\tilde{C}_n|\mathbf{a})}{\sum_{\mathbf{a} \in \tilde{C}_n} P_n(\mathbf{a}|\theta, C_n) q(\tilde{C}_n|\mathbf{a})}. \quad \text{が full choice set } C_n \text{ に依存しているため, 計算不可能}$$

→ Guevara and Ben-Akiva(2013)に基づき, $W_n = 1$ と近似

<https://www.sciencedirect.com/science/article/pii/S0191261513001471>

3. Methodology

MRLの最尤推定法

対数尤度関数

$$\mathcal{L} = \sum_{n=1}^N \log \int \frac{e^{u(a_n|\beta) + \log(q_n(\tilde{C}_n|a_n))}}{\sum_{a \in \tilde{C}_n} e^{u(a|\beta) + \log(q_n(\tilde{C}_n|a))}} f(\beta|\theta) d\beta. \quad \rightarrow \text{真値はモンテカルロシミュレーションによる求解が必要}$$

積分値を離散点 β_r の周りで平均することで近似可能

離散点 β_r の値は分布 $f(\beta|\theta)$ からランダムに選ぶか、積分範囲上で等間隔に選ぶ

擬似対数尤度関数

$$\mathcal{L} \mathcal{L} = \sum_{n=1}^N \log \left\{ \frac{1}{R} \sum_{r=1}^R \frac{e^{u(a_n|\beta_r) + \log(q_n(\tilde{C}_n|a_n))}}{\sum_{a \in \tilde{C}_n} e^{u(a|\beta_r) + \log(q_n(\tilde{C}_n|a))}} \right\}.$$

この手法によるパラメータ推定の妥当性については Guevara and Ben-Akiva(2013)を参照

- 今回は β_r はHalton数列から準ランダムに生成
- サンプル \tilde{C}_n は初期パラメータに対するRLモデルを用いたシミュレーションによって生成しても、任意の経路生成アルゴリズムによって生成してもいい

4. Application

Data – Stockholm 2004

- 1日の6:00-23:00のアクティビティダイアリーを記録
- 仕事をしていて、6:00より後に出勤し、20:00より前に帰宅する人に限定
- 1日のすべてのトリップを車で行うか、車を一度も使わない人に限定
- 3,150個のアクティビティスケジュールを取得

各トリップの持つ情報

- ①出発時間, ②到着時間, ③交通手段, ④到着地で行った活動, ⑤到着地の場所, ⑥継続時間

個人属性

Age (years)	[%]	Income (SEK)	[%]	Household	[%]	Working hours	[%]
12-18	1.9	0-15,000	1.8	Single	28.2	< 6	5.0
19-24	1.6	15,001-25,000	15.1	Couple	30.1	6-8	21.2
25-39	31.3	25,001-40,000	24.3	Single w. child	2.6	8-10	66.0
40-64	63.8	40,001-55,000	44.1	Couple w. child	39.2	> 10	7.8
> 64	1.3	> 55,000	14.7				

Gender	[%]	Owns car	[%]	Owns PT card	[%]
Female	58.4	Yes	20.9	Yes	42
Male	42.6	No	79.1	No	58

加えて勤務時間は固定かどうかも訊いた



- 個人の時空間制約の設定
- 効用関数の説明変数

4. Application

記述のルール

状態 x_t の要素

時刻 $t \in [5 \text{ am}, 11 \text{ pm}]$: 1分ごとのタイムステップに離散化

場所 $l \in L$: スtockホルムを1240ゾーンに分割し, そのうちの一つ

活動 $p \in P$: social, recreational, shop small, shop medium, shop large, home, work, escort children

Errand indicator $e \in \{0,1,2,3\}$: 必須の用事をいくつ終了させたか

車の有無 $\delta_{car} \in \{0,1\}$: 0なら車は使えない. 1で, かつ車で外出した場合帰るまで車

活動 a_t の要素

活動 $p \in P$: 変えても変えなくてもいい

場所 $l \in L$: 変えても変えなくてもいい

交通手段 $m \in M$: 車, 公共交通, 徒歩, 自転車のいずれか

- 活動の継続時間は10分単位, 移動にかかる時間は1分単位
- 時空間制約は, 不可能な選択肢の効用を $-\infty$ にすることで表現する

4. Application

効用関数

個人 n の状態 $x_t = (t, l, p, e, \delta_{car})$ における行動 $a_t = (p', l', m)$ の効用確定項 $v_n(a_t|x_t)$
= 目的地 l' , 交通手段 m , 活動 p' の効用の和と定義

交通手段 m の効用 $v_{n,m}(l, l', t)$

料金

待ち時間

$$v_{n,car}(l, l', t) = ASC_{car} + \theta_{t,car} T_{car}(l', l, t) + \theta_c C_{car}(l', l, t)$$

$$v_{n,PT}(l, l', t) = ASC_{PT} + \theta_{t,PT} T_{PT}(l', l, t) + \theta_{wait,PT} T_{wait,PT}(l', l, t) + \theta_c C_{PT}(l', l, t)$$

$$v_{n,bike}(l, l', t) = ASC_{bike} + \theta_{t,bike} T_{bike}(l', l, t)$$

$$v_{n,walk}(l, l', t) = ASC_{walk} + \theta_{t,walk} T_{walk}(l', l, t) + \theta_{samezone} \delta_{samezone}$$

旅行時間

ゾーン内移動ダミー

他に個人属性を導入し、いくつかのパターンで推定を行った

4. Application

交通手段選択 推定結果

Parameter	MNL		MNL bike dummy		Mixed		Mixed bike dummy		Mixed bike dummy covariance		
	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	
<i>Constants for choosing a specific mode of transport on a trip</i>											
Car ASC	$\bar{\beta}$	-2.498	-22.897	-2.532	-24.588	-2.756	-19.745	-2.746	-23.832	-2.743	-23.564
	σ	0	-	0	-	1.228	13.523	0.869	15.305	0.873	15.165
PT ASC	$\bar{\beta}$	-3.691	-36.802	-3.502	-37.620	-4.231	-25.391	-3.568	-36.895	-3.637	-36.683
	σ	0	-	0	-	0.969	9.096	0.006	0.394	0.002	0.244
Walk ASC	$\bar{\beta}$	-1.708	-15.651	-1.648	-16.185	-2.142	-15.772	-2.009	-16.818	-2.013	-17.155
	σ	0	-	0	-	1.147	15.414	-0.898	-14.026	-0.901	-14.539
Bike ASC	$\bar{\beta}$	-3.267	-14.968	-7.176	-21.820	-4.337	-19.126	-7.231	-21.348	-7.220	-21.348
$\sigma_{\text{Walk-PT}}$		0	-	0	-	0	-	0	-	0.257	5.781
<i>Additional constants for choosing a specific mode conditional on bike being the first trip on the tour</i>											
Bike ASC δ_{bike}		0	-	5.120	17.521	0	-	4.364	13.924	4.362	13.949
PT ASC δ_{bike}		0	-	-2.031	-6.692	0	-	-2.593	-7.937	-2.581	-7.846
Walk ASC δ_{bike}		0	-	-0.975	-3.608	0	-	-1.500	-6.018	-1.635	-6.582
<i>Additional constants for choosing a specific mode conditional on socio-demographics or same zone trips</i>											
Walk ASC same zone		-0.598	-4.517	-0.640	-5.162	-0.572	-4.865	-0.540	-4.854	-0.558	-4.949
Bike ASC Female		0.143	0.574	0.455	2.922	0.326	1.263	0.415	2.414	0.431	2.513
Car ASC Female		-0.290	-6.173	-0.255	-5.955	-0.424	-5.473	-0.359	-6.235	-0.355	-6.134
<i>Parameters for travel time, cost (car and PT) and wait time (PT)</i>											
Cost		-0.017	-6.207	-0.018	-6.954	-0.002	-0.262	-0.019	-5.553	-0.017	-5.417
Car time		-0.080	-18.688	-0.079	-19.730	-0.111	-14.281	-0.082	-17.711	-0.085	-17.226
PT time		-0.040	-5.485	-0.044	-6.613	-0.081	-7.906	-0.059	-7.386	-0.060	-7.495
PT wait time		0.008	0.853	0.010	1.178	0.048	3.836	0.023	2.375	0.025	2.504
Walk time		-0.050	-23.171	-0.049	-24.691	-0.055	-23.755	-0.051	-25.588	-0.052	-25.702
Bike time		-0.050	-9.336	-0.035	-7.957	-0.057	-8.416	-0.041	-8.230	-0.042	-8.190
Bike time age ≥ 24		-0.016	-1.499	-0.008	-2.072	-0.017	-1.562	-0.008	-1.556	-0.009	-1.835
Bike time Female		-0.024	-3.067	-0.026	-3.867	-0.030	-3.186	-0.025	-3.532	-0.026	-3.634
Log-likelihood		-23,671		-22,091		-22,564		-21,865		-21,855	

4. Application

効用関数

活動 p の効用

- 時刻 $t' = t + T_m(l, l', t)$ に活動 p を開始
- この時の効用は活動を始める時刻に依存する定数 $c_p(t')$ と、活動を始める時刻と継続時間に依存する効用 $v_{n,p}(t', \Delta_{tp})$ によって決まる
- いくつかの活動にタイムステップ T_k によって異なるパラメータ $\theta_{p,T_k}, c_{p,T_k}$ を考える
例) c_{work,T_k} for $T_k \in \{6 AM, 7 AM, 8 AM, 9 AM, 10 AM\}$
- それ以外の活動はパラメータ数を制限するため、時間非依存の定数 c_p とする
- 時刻 t に活動を開始する限界効用 $v(p, t)$ はパラメータ θ_{p,T_k} の最近点 T_j, T_{j+1} ($t \in (T_j, T_{j+1})$) の間での線形補間により与えられる
- この時、継続時間 Δ_{tp} の活動の効用 $v_{n,p}(t, \Delta_{tp}) = \int_t^{t+\Delta_{tp}} v(p, \tau) d\tau$
- 定数 c_p に個人属性（年齢、子どもがいるダミー）を導入し、いくつかのパターンで推定を行った

4. Application

活動選択 推定結果

Parameter	MNL		MNL bike dummy		Mixed		Mixed bike dummy		Mixed bike dummycovariance	
	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test
<i>Utility to arrive at work at specific time, linear between parameters</i>										
Work ASC 6 AM	0.946	2.512	1.045	2.772	1.575	3.850	1.457	3.711	1.491	3.778
Work ASC 7 AM	0.506	2.751	0.507	2.796	0.674	3.372	0.648	3.377	0.677	3.517
Work ASC 8 AM	0	-	0	-	0	-	0	-	0	-
Work ASC 9 AM	-1.310	-8.144	-1.273	-8.099	-1.494	-8.344	-1.422	-8.383	-1.450	-8.464
Work ASC 10 AM	-5.092	-13.627	-5.008	-13.637	-5.573	-12.854	-5.344	-13.364	-5.416	-13.273
<i>Constants for starting activities</i>										
Home ASC	0	-	0	-	0	-	0	-	0	-
Shop ASC	-6.767	-41.591	-6.769	-42.708	-6.658	-37.692	-6.643	-39.807	-6.595	-39.251
Social ASC	-9.046	-46.112	-9.091	-48.419	-8.935	-45.025	-8.946	-47.059	-8.901	-46.557
Recreative ASC	-7.723	-51.624	-7.737	-53.690	-7.645	-47.022	-7.652	-50.457	-7.608	-49.783
Other ASC	-7.191	-48.493	-7.191	-50.119	-7.121	-47.839	-7.107	-49.839	-7.056	-49.248
<i>Additional constants for starting activities dependent on socio-demographics</i>										
Shop ASC children	-0.171	-1.786	-0.197	-2.206	-0.229	-2.199	-0.240	-2.504	-0.243	-2.535
Freetime ASC age ≤30	0.211	1.649	0.235	1.895	0.161	1.190	0.162	1.324	0.170	1.387
Freetime ASC age ≥60	-0.471	-3.308	-0.444	-3.243	-0.438	-2.814	-0.412	-2.861	-0.407	-2.863
Trip ASC Own Car	0.081	1.348	0.085	1.552	-0.198	-2.675	-0.104	-1.679	-0.115	-1.861
<i>Utility per minute of activity participation</i>										
Shop time	-0.021	-14.137	-0.021	-14.131	-0.021	-13.299	-0.021	-13.919	-0.021	-13.932
Social time	-0.000	-0.066	-0.000	-0.133	-0.000	-0.089	-0.000	-0.213	-0.000	-0.198
Recreative time	0	-	0	-	0	-	0	-	0	-
Other time	-0.009	-6.108	-0.009	-6.436	-0.008	-5.919	-0.008	-6.424	-0.008	-6.393
Freetime time children	-0.003	-3.511	-0.003	-3.488	-0.004	-3.989	-0.003	-3.963	-0.003	-3.971
<i>Utility per minute of time spent at home, marginal utility is linear between the time periods specified</i>										
Home time 6 AM	0.042	8.397	0.044	8.971	0.047	8.173	0.047	8.850	0.048	8.812
Home time 7 AM	0.040	11.688	0.039	11.864	0.043	11.607	0.041	11.781	0.042	11.952
Home time 8 AM	0.019	6.006	0.018	5.977	0.020	5.783	0.020	6.092	0.021	6.202
Home time 9 AM	0.016	3.260	0.015	3.080	0.019	3.450	0.017	3.323	0.018	3.409
Home time 1 PM	-0.012	-10.839	-0.012	-11.317	-0.013	-11.721	-0.012	-11.776	-0.012	-11.632
Home time 5 PM	0.003	3.503	0.003	3.326	0.002	2.065	0.002	2.696	0.003	2.878
Home time 7 PM	0.002	2.529	0.002	2.352	0.001	1.427	0.001	1.614	0.002	1.660
Home time 9 PM	0.018	12.858	0.018	13.482	0.019	12.761	0.019	13.298	0.018	13.248

4. Application

効用関数

場所 l の効用

- サイズパラメータ $\theta_{p,LSM}$, その場所で可能なアクティビティの数 $\gamma_{p,s}$ によって決定
- 効用関数

$$v_{n,p}(l) = \theta_{p,LSM} \log(\sum_{s=1}^{S_p} x_{p,l,s} e^{\gamma_{p,s}})$$

- S_p はアクティビティ p のサイズ変数の数, サイズ変数 $x_{p,l,s}$ は例えば場所 l のあるセクター内の従業員数など

場所選択 推定結果

Parameter	MNL		MNL bike dummy		Mixed		Mixed bike dummy		Mixed bike dummy covariance		
	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	
<i>Log-sum parameters for size attributes. Enters utility as $\theta_{LSM} \text{Size}$ in $\theta_{LSM} \text{Size} \cdot \log \sum_s e^{\gamma_s \cdot N_{s,location}}$</i>											
Social LSM size	0.017	1.964	0.017	2.001	0.020	1.528	0.020	1.619	0.020	1.611	
Recreative LSM size	0.057	1.766	0.060	1.871	0.081	2.158	0.078	2.160	0.080	2.178	
Other LSM size	0.318	5.663	0.309	5.828	0.357	6.240	0.343	6.375	0.348	6.449	
Shop LSM size	0.485	33.491	0.484	33.392	0.487	14.299	0.486	17.289	0.487	17.227	
<i>Parameters for size attributes. Enters utility as γ_i in $\theta_{LSM} \text{Size} \cdot \log \sum_s e^{\gamma_s \cdot N_{s,location}}$</i>											
Rec.	Population	-100	-	-100	-	-100	-	-100	-	-100	-
	No employed rec.	5.907	9.819	5.809	9.154	6.154	11.715	6.028	10.881	6.030	10.999
Other	No employed OE	-100	-	-100	-	-100	-	-100	-	-100	-
Shop	Population	-100	-	-100	-	-100	-	-100	-	-100	-
	No employed shop	3.585	13.587	3.611	13.798	3.685	12.754	3.663	13.259	3.682	13.279
Social	Population	-100	-	-100	-	-100	-	-100	-	-100	-

4. Application

相関構造

交通手段選択における相関構造を記述するために
選択肢固有定数項 ($ASC_{car}, ASC_{PT}, ASC_{walk}, ASC_{bike}$) を
平均値ベクトル ($\bar{\beta}_{car}, \bar{\beta}_{PT}, \bar{\beta}_{walk}, \bar{\beta}_{bike}$),

$$\text{分散共分散行列 } \Sigma = \begin{pmatrix} \sigma_{car}^2 & 0 & 0 & 0 \\ 0 & \sigma_{PT}^2 & \sigma_{walk,PT} & 0 \\ 0 & \sigma_{walk,PT} & \sigma_{walk}^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

を持つ分布に従ってランダムに分布させる.

パラメータ $ASC_{car}, ASC_{PT}, ASC_{walk}$ が

それぞれ独立な分布 $N(\bar{\beta}_{car}, \sigma_{car}^2), N(\bar{\beta}_{PT}, \sigma_{PT}^2), N(\bar{\beta}_{walk}, \sigma_{walk}^2)$ に従って分布するとした場合と比較.

4. Application

状態空間の拡張

→ 連続的な活動・移動の選択の間の一貫性，相互依存を捉える

bike tour dummy $\delta_{bike} \in \{0,1\}$: 自転車で家を出たとき1を取るダミー変数
→ 自転車で家に帰る確率が高いはず

$\left\{ \begin{array}{l} ASC_{bike} | \delta_{bike} : \text{自転車で家を出たときの自転車の選択肢固有定数} \\ ASC_{PT} | \delta_{bike} : \text{自転車で家を出たときの公共交通の選択肢固有定数} \\ ASC_{walk} | \delta_{bike} : \text{自転車で家を出たときの徒歩の選択肢固有定数} \end{array} \right.$

→ 交通機関選択の選択肢固有定数項を $ASC_m + ASC_m | \delta_{bike} \cdot \delta_{bike}$ で置き換える
= 一日のツアー内の交通機関選択の一貫性を捉えることが可能

4. Application

推定結果

1. MNLモデル
 2. MNLモデル（状態空間の拡張あり）
 3. MXLモデル
 4. MXLモデル（状態空間の拡張あり）
 5. MXLモデル（共分散 $\sigma_{walk,PT}$, 状態空間の拡張あり） の5パターンで推定
- Halton数列から500個の数字を取り出し，擬似対数尤度関数を計算
 - 経路の選択肢集合はBlom Västberg et al. (2016)で推定されたパラメータ値を初期値として抽出
 - 各選択肢集合につき，観測された経路選択肢以外に600個の代替選択肢を抽出

4. Application

推定結果の解釈 (交通手段選択)

Parameter	MNL		MNL bike dummy		Mixed		Mixed bike dummy		Mixed bike dummy covariance		
	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	Est.	t-test	
<i>Constants for choosing a specific mode of transport on a trip</i>											
Car ASC	$\bar{\beta}$	-2.498	-22.897	-2.532	-24.588	-2.756	-19.745	-2.746	-23.832	-2.743	-23.564
	σ	0	-	0	-	1.228	13.523	0.869	15.305	0.873	15.165
PT ASC	$\bar{\beta}$	-3.691	-36.802	-3.502	-37.620	-4.231	-25.391	-3.568	-36.895	-3.637	-36.683
	σ	0	-	0	-	0.969	9.096	0.006	0.394	0.002	0.244
Walk ASC	$\bar{\beta}$	-1.708	-15.651	-1.648	-16.185	-2.142	-15.772	-2.009	-16.818	-2.013	-17.155
	σ	0	-	0	-	1.147	15.414	-0.898	-14.026	-0.901	-14.539
Bike ASC	$\bar{\beta}$	-3.267	-14.968	-7.176	-21.820	-4.337	-19.126	-7.231	-21.348	-7.220	-21.348
$\sigma_{Walk-PT}$		0	-	0	-	0	-	0	-	0.257	5.781
<i>Additional constants for choosing a specific mode conditional on bike being the first trip on the tour</i>											
Bike ASC δ_{bike}		0	-	5.120	17.521	0	-	4.364	13.924	4.362	13.949
PT ASC δ_{bike}		0	-	-2.031	-6.692	0	-	-2.593	-7.937	-2.581	-7.846
Walk ASC δ_{bike}		0	-	-0.975	-3.608	0	-	-1.500	-6.018	-1.635	-6.582
<i>Additional constants for choosing a specific mode conditional on socio-demographics or same zone trips</i>											
Walk ASC same zone		-0.598	-4.517	-0.640	-5.162	-0.572	-4.865	-0.540	-4.854	-0.558	-4.949
Bike ASC Female		0.143	0.574	0.455	2.922	0.326	1.263	0.415	2.414	0.431	2.513
Car ASC Female		-0.290	-6.173	-0.255	-5.955	-0.424	-5.473	-0.359	-6.235	-0.355	-6.134
<i>Parameters for travel time, cost (car and PT) and wait time (PT)</i>											
Cost		-0.017	-6.207	-0.018	-6.954	-0.002	-0.262	-0.019	-5.553	-0.017	-5.417
Car time		-0.080	-18.688	-0.079	-19.730	-0.111	-14.281	-0.082	-17.711	-0.085	-17.226
PT time		-0.040	-5.485	-0.044	-6.613	-0.081	-7.906	-0.059	-7.386	-0.060	-7.495
PT wait time		0.008	0.853	0.010	1.178	0.048	3.836	0.023	2.375	0.025	2.504
Walk time		-0.050	-23.171	-0.049	-24.691	-0.055	-23.755	-0.051	-25.588	-0.052	-25.702
Bike time		-0.050	-9.336	-0.035	-7.957	-0.057	-8.416	-0.041	-8.230	-0.042	-8.190
Bike time age ≥ 24		-0.016	-1.499	-0.008	-2.072	-0.017	-1.562	-0.008	-1.556	-0.009	-1.835
Bike time Female		-0.024	-3.067	-0.026	-3.867	-0.030	-3.186	-0.025	-3.532	-0.026	-3.634
Log-likelihood		-23,671		-22,091		-22,564		-21,865		-21,855	

4. Application

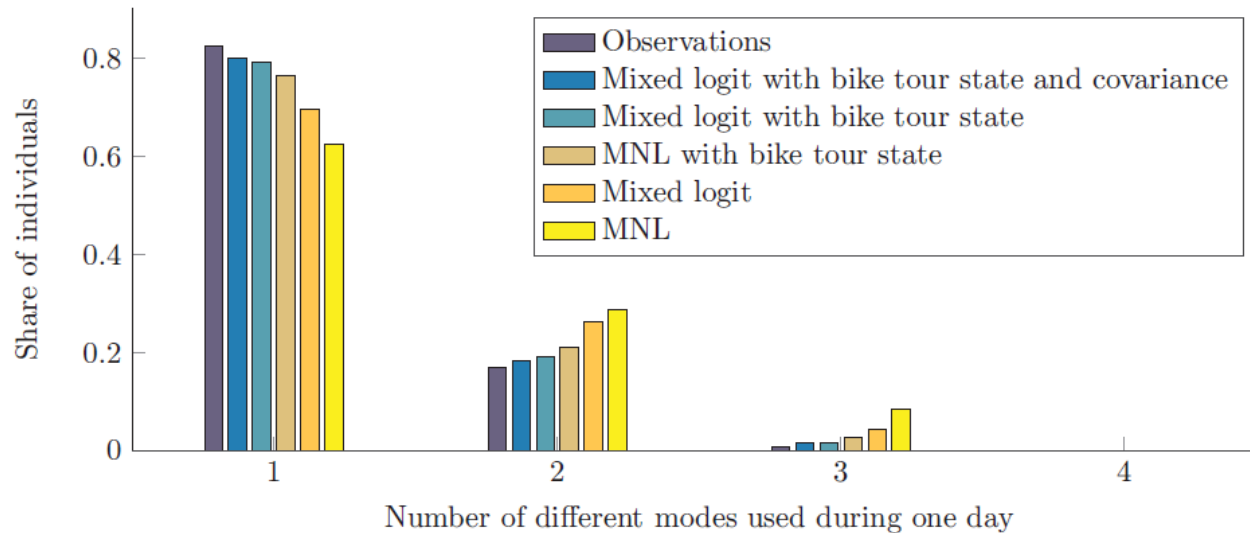
推定結果の解釈（交通手段選択）

- 選択肢固有定数項は有意に0とは異なる → 選好の異質性
- MXLモデル（状態空間の拡張あり）では，PTの標準偏差は0とほぼ同じ
= 係数 $ASC_{PT}|\delta_{bike}$ がPTの選好の誤差を良く吸収している
- 5番目のモデルでは共分散 $\sigma_{walk,PT}$ は負
= walkとPTのどちらかを好めば，もう片方は好まない傾向
= 個人は単一の交通手段を選好する傾向があるので他の2つの組み合わせでも同様の傾向があると予想される
- 状態空間の拡張を行ったときの結果は予想通り
- 女性は車を選択しずらく，自転車を選択する傾向があるということなども分かる
- モデルの適合度はMNLよりMXLの方が高い
- 推定時間はMNLで30分，MXLで3時間

5. In-sample fit and predictions

In-sample fit (交通手段選択に着目して)

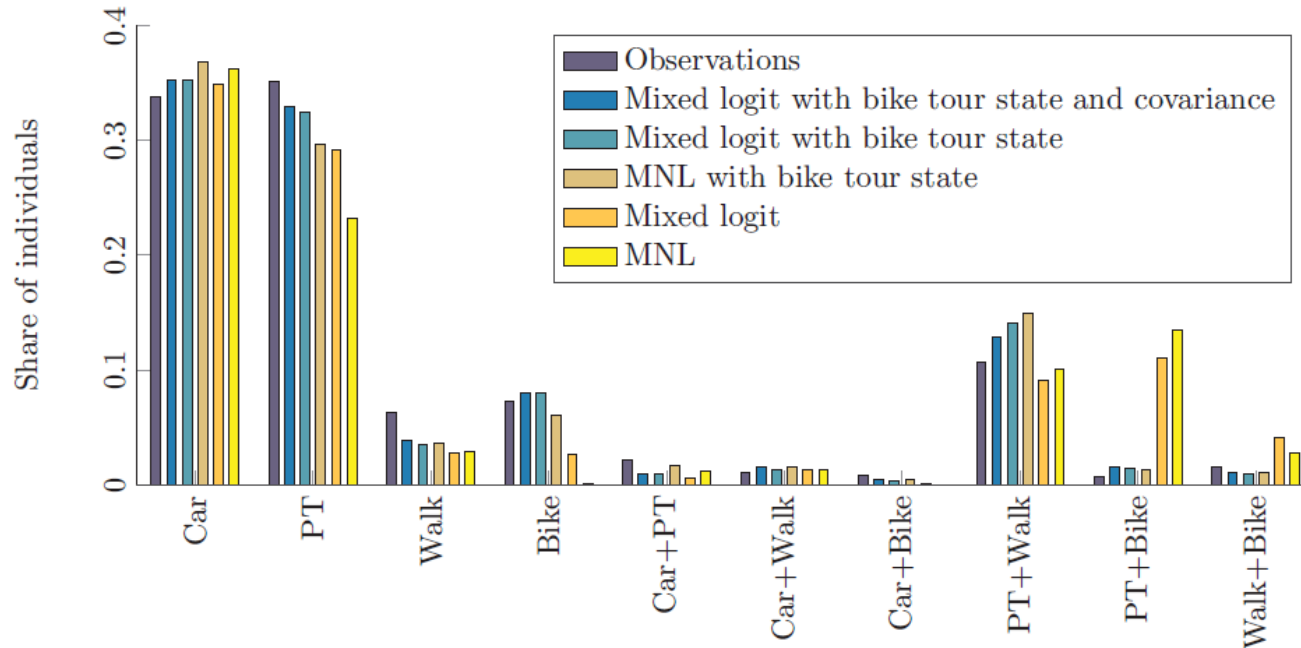
- RLの選択確率式に基づいて、一人につき1000個のスケジュールを生成する
- mixed RLの場合、個人ごとに、推定された平均 $\bar{\beta}$ と標準偏差 σ で決まる分布から一つ値を抽出し、パラメータ値とする



- 標準的なMNLやMXLは実際の観測よりも多種の交通手段を一日に利用してしまう

5. In-sample fit and predictions

In-sample fit (交通手段選択に着目して)



- 特にPT+Bike, Walk+Bikeの選択確率はMNLやMXLで不当に高い
→ 状態空間の拡張がモデルの当てはまりを大きく改善している
- 実際はモデルのoverspecificationにつながる可能性がある→のちに議論

5. In-sample fit and predictions

代替パターン

- 17:00~19:00のPTの課金を考える
- ある7パターンのスケジュールの選択確率の変化をwith/withoutで比較

Change in choice probability of alternatives after price increase.

Alternatives				Change in choice probability (%)		
				MNL	MNL with δ_{bike}	Mixed logit with δ_{bike}
Alternative 1				-100.00	-100.00	-100.00
5:00	8:30	Home	PT			
8:44	17:10	Work	PT			
17:33	23:00	Home				
Alternative 2				+35.05	+27.65	+39.03
5:00	7:40	Home	PT			
7:54	16:20	Work	PT			
16:43	23:00	Home				
Alternative 3				+35.00	+27.51	+36.73
5:00	8:30	Home	PT			
8:44	17:10	Work	Bike			
17:33	23:00	Home				
Alternative 4				+35.00	+27.51	+28.87
5:00	8:30	Home	PT			
8:44	17:10	Work	Walk			
18:18	23:00	Home				
Alternative 5				+34.92	+27.44	+21.46
5:00	7:40	Home	Walk			
8:40	16:20	Work	Walk			
18:15	23:00	Home				
Alternative 6				+34.79	+27.29	+18.83
5:00	7:50	Home	PT			
8:04	16:30	Work	Walk			
17:38	18:28	Home	Walk			
18:50	19:20	Shop Small	Walk			
19:42	23:00	Home				
Alternative 7				+35.04	+27.61	+41.18
5:00	7:00	Home	PT			
7:14	15:40	Work	PT			
16:10	16:30	Other	PT			
16:44	23:00	Home				
Alternative 8				+34.79	+27.30	+34.24
5:00	8:10	Home	Bike			
8:26	16:50	Work	Bike			
17:20	23:00	Home				

← 課金時間にPTを使うスケジュールは選択されなくなる

← 全ての選択肢の選択確率が同じ割合で変化する
= IIA特性の問題点

5. In-sample fit and predictions

代替パターン

- 17:00~19:00のPTの課金を考える
- ある7パターンのスケジュールの選択確率の変化をwith/withoutで比較

Change in choice probability of alternatives after price increase.

Alternatives				Change in choice probability (%)		
				MNL	MNL with δ_{bike}	Mixed logit with δ_{bike}
Alternative 1				-100.00	-100.00	-100.00
5:00	8:30	Home	PT			
8:44	17:10	Work	PT			
17:33	23:00	Home				
Alternative 2				+35.05	+27.65	+39.03
5:00	7:40	Home	PT			
7:54	16:20	Work	PT			
16:43	23:00	Home				
Alternative 3				+35.00	+27.51	+36.73
5:00	8:30	Home	PT			
8:44	17:10	Work	Bike			
17:33	23:00	Home				
Alternative 4				+35.00	+27.51	+28.87
5:00	8:30	Home	PT			
8:44	17:10	Work	Walk			
18:18	23:00	Home				
Alternative 5				+34.92	+27.44	+21.46
5:00	7:40	Home	Walk			
8:40	16:20	Work	Walk			
18:15	23:00	Home				
Alternative 6				+34.79	+27.29	+18.83
5:00	7:50	Home	PT			
8:04	16:30	Work	Walk			
17:38	18:28	Home	Walk			
18:50	19:20	Shop Small	Walk			
19:42	23:00	Home				
Alternative 7				+35.04	+27.61	+41.18
5:00	7:00	Home	PT			
7:14	15:40	Work	PT			
16:10	16:30	Other	PT			
16:44	23:00	Home				
Alternative 8				+34.79	+27.30	+34.24
5:00	8:10	Home	Bike			
8:26	16:50	Work	Bike			
17:20	23:00	Home				

←PTを使用するのは変わらず，帰宅時間をずらす選択肢は選択確率が大きく上がる

←PTを使用しない選択肢はあまり上がらない

5. In-sample fit and predictions

代替パターン

- 17:00~19:00のPTの課金を考える
- ある7パターンのスケジュールの選択確率の変化をwith/withoutで比較

Change in choice probability of alternatives after price increase.

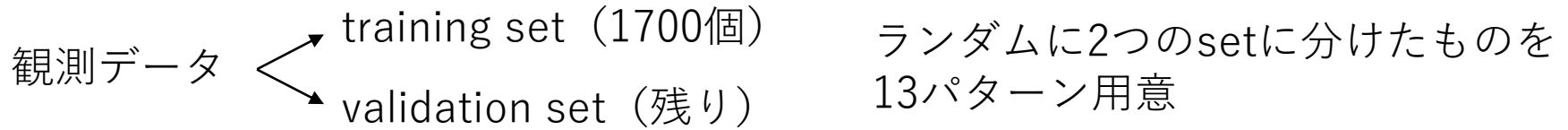
Alternatives				Change in choice probability (%)		
				MNL	MNL with δ_{bike}	Mixed logit with δ_{bike}
Alternative 1				-100.00	-100.00	-100.00
5:00	8:30	Home	PT			
8:44	17:10	Work	PT			
17:33	23:00	Home				
Alternative 2				+35.05	+27.65	+39.03
5:00	7:40	Home	PT			
7:54	16:20	Work	PT			
16:43	23:00	Home				
Alternative 3				+35.00	+27.51	+36.73
5:00	8:30	Home	PT			
8:44	17:10	Work	Bike			
17:33	23:00	Home				
Alternative 4				+35.00	+27.51	+28.87
5:00	8:30	Home	PT			
8:44	17:10	Work	Walk			
18:18	23:00	Home				
Alternative 5				+34.92	+27.44	+21.46
5:00	7:40	Home	Walk			
8:40	16:20	Work	Walk			
18:15	23:00	Home				
Alternative 6				+34.79	+27.29	+18.83
5:00	7:50	Home	PT			
8:04	16:30	Work	Walk			
17:38	18:28	Home	Walk			
18:50	19:20	Shop Small	Walk			
19:42	23:00	Home				
Alternative 7				+35.04	+27.61	+41.18
5:00	7:00	Home	PT			
7:14	15:40	Work	PT			
16:10	16:30	Other	PT			
16:44	23:00	Home				
Alternative 8				+34.79	+27.30	+34.24
5:00	8:10	Home	Bike			
8:26	16:50	Work	Bike			
17:20	23:00	Home				



←PTとwalkの共分散は負
 =walkを使う選択肢は特に選択確率が上がりにくい

5. In-sample fit and predictions

交差検証 (out-of-sample fit)



→ training setで推定したパラメータのvalidation setへの当てはまりを検証

対数尤度損失

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

T_i : validation set i

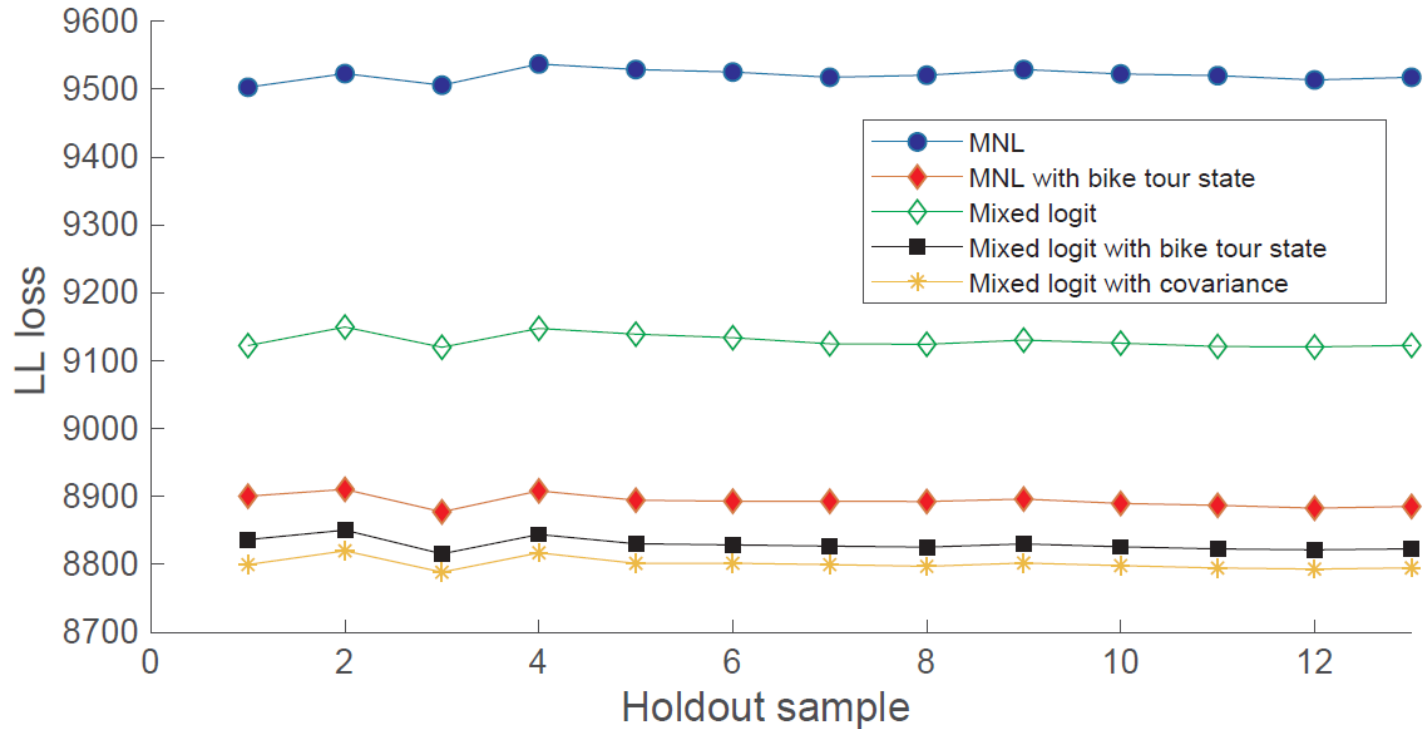
$\hat{\beta}_i$: training set i から推定したパラメータ

13回の平均

$$\overline{err}_p = \frac{1}{p} \sum_{i=1}^p err_i \quad \forall 1 \leq p \leq 13.$$

5. In-sample fit and predictions

交差検証



状態空間の拡張をすることで、サンプル外への当てはまりもよくなる

6. Conclusion

成果

- 相関構造を扱うモデルで、現実的な時間内での推定を可能にした
- 従来のモデルよりもIIA特性を緩和し、モデルの予測精度を向上させた
- 選択肢サンプリングという手法を用いても、サンプル外へのモデルの適合度も従来のモデルより高く、柔軟な選択肢の代替パターンを再現できた

今後の展望

- 現実には、毎日行う必要のない活動が多く存在
→ モデルのタイムスパンを数日間に拡張することが妥当
- 移動時間を外生的に与えるのではなく、リンクフローの関数とするのが妥当
→ Dynamic Traffic Equilibrium (DTA) modelと組み合わせたモデルが検討可能