交通・行動観測に基づく 支配的・従属的活動の特定

鷹野　由利香[1]・原　祐輔[2]

[1]非会員　東北大学　大学院情報科学研究科（〒 980-8579 仙台市青葉区荒巻字青葉 6-6-06-408）
E-mail: takano.yurika.t2@dc.tohoku.ac.jp
[2]正会員　東北大学准教授　大学院情報科学研究科（〒 980-8579 仙台市青葉区荒巻字青葉 6-6-06-408）
E-mail: hara@tohoku.ac.jp（Corresponding author）

本研究ではアクティビティチェインにおける活動間の共起関係に着目し，その生成過程をモデル化した．活動の共起関係を生み出すメカニズムとして，連続する活動の生成確率に影響を与える支配的活動と影響を与えない従属的活動の非対称的な 2 つの潜在属性を定義し，観測データから潜在属性やパラメータを推定する手法を提案した．本手法の検証のために，シミュレーションデータを用いて提案手法の検証を行い，観測データのみから潜在属性やパラメータを推定できることを実証的に確認した．これにより，特定の活動が除去・追加された場合のアクティビティチェインの変化が確率的に推論できることを示した．

Key Words: *observed data*，*travel behavior analysis*，*latent structure analysis*，*stochastic block model*

1. はじめに

情報通信技術の発展や社会情勢の変化を受けて，人々の行動様式は著しく変化している．例えば，E-commerce の普及は多種多様な商品に対するアクセス性を高めることで，実店舗とオンラインショッピングの役割分担が進み，人々の買い物習慣を大きく変化させている[1)2)]．リモートワークの普及は，勤務地と自宅との物理的距離から人々を解放し，通勤時間の削減や柔軟な働き方を実現することで，かつての通勤・通学を伴う義務的な活動の性質を根本的に変化させている．

このような新たなライフスタイルは，自宅やオフィスから業務や消費行動を可能にし，利便性・効率性を向上させた一方で，外出機会の減少をもたらし，物理的な店舗への来客数や滞在時間の減少に繋がることも懸念されている[3)]．その理由として，一日に行う都市活動の多くは，それら自身が必ずしも外出を行う主目的ではないと考えられるためである．例えば，オフィスでの勤務という主目的で出勤することで，オフィス周辺の飲食店で昼食を取るという別の活動が派生することは一般的であるが，その飲食店がいかに気に入っていたとしても，勤務日以外にその飲食店を訪れることは稀だろう．このように勤務目的の活動，食事目的の活動の 2 つが存在しても，上記の例では互いに対等な派生活動を生み出す関係にはない．これは勤務目的のみに当てはまるわけではなく，映画館に行くことが主目的でその帰りに飲食店で夕食をする，買い物が主目的で途中でカフェに立ち寄る等の主目的（外出目的）の活動とそこから派生する活動の組み合わせによって，人々の一日の活動が成り立っていると考えられる．

リモートワークや E-commerce が物理的空間に与える外部性，中心市街地活性化のための様々な施策が実行動に与える影響を評価することは今日的に重要な課題である．特定の活動が除去・追加されることで，その活動の前後の活動が生成される確率や一日の行動パターンはどのように変化するのだろうか．この問いを考えるために，本研究は活動間の共起関係とアクティビティチェインの生成過程に着目する．一日に行われる複数の活動は互いに独立ではなく，ある活動が別の活動を派生させる関係性にあることから，以上のような社会の変化に伴い特定の活動の生成確率が増減することは，派生的に生成された活動の生成確率やアクティビティチェインにも影響を与える可能性が大きいと推測される．

そこで本研究では，観測されたトリップ・アクティビティデータによる行動履歴のみを用いて，それぞれの活動が外出の主目的であるか，主目的の活動から派生して生成された活動であるかを判別するための潜在的な相互依存性を表す属性を特定する手法を構築する．本手法は既存の Activity-based approach (ABA)[4)] と対立するものではなく，時空間的な連続性や順序を踏まえたアクティビティベースモデルの発展に寄与するものである．社会状況の変化や政策への反応感度をもつアクティビティベースモデルは，E-commerce やリモート

ワークが与える外部性の評価をできるだけでなく，中心市街地活性化施策がミクロなレベルでの回遊性の向上や都市内滞在時間に与える影響を評価するための基盤となる．

本研究の貢献は以下の 3 点である．

- 活動の順序組として表現されるアクティビティチェインに対して，連続する活動の共起関係に着目した生成過程のモデル化を行った．各活動には観測できない潜在属性として，外出の主目的となりうる支配的活動と支配的活動から派生的に生成されやすい従属的活動の 2 種類が存在すると仮定し，観測可能な活動属性からこれらの潜在属性を特定する手法を構築した．(**3. (1), (2), (3)**)
- 本研究の仮定の成立下において，アクティビティチェインに含まれる一部の活動が除去，または外的に追加された際に，アクティビティチェインの変化を推論する活用法を提案した．(**3. (4)**)
- 提案手法に基づき，あらかじめ設定した真のパラメータから生成したシミュレーションデータに対して，観測データのみから推定したパラメータが真のパラメータに十分近いことを検証した．この結果から，実データに適用してもパラメータや潜在属性を推定できる可能性を示した．(**4.**)

2. 既往研究の整理と本研究の位置付け

アクティビティベースの行動分析は都市計画分野において重要度が高まっており，特にそのモデル開発において多くの研究が行われている．Bowman and Ben-akiva [5] では，ツアー単位で活動選択を行う離散選択モデルを提案した．このモデルでは，トリップ間の相互依存関係を考慮する点や，すべてのツアーを主要な活動を含むツアーと副次的な活動を含むツアーに大別し，これらの選択の順序を考慮するという点において，人間の意思決定に即したモデルを構築している．一方，このような離散選択ベースのアクティビティベースモデルは，活動時間の詳細を扱えないという制約があることから，近年では離散連続モデルが適用される研究も増えている．福田・力石 [6] は，離散連続モデルの最新動向を包括的にレビューし，明示的な資源制約条件を考慮した「構造型」，統計的な現象記述を主体とした「誘導型」，そしてミクロ経済理論との調和を図った「折衷型」にモデルを分類できることを示した．構造型の離散連続モデルである Bhat[7] の MDCEV (Multiple discrete-continuous extreme value model) は，複数の選択肢の同時選択とともに活動に割り当てる時間を連続量で記述できるモデルとして，アクティビティベースモデルの発展に大きく貢献した．ただし，このような構造型モデルは，集計的な資源配分の性質を持つため，個別の活動選択の意思決定をモデル化することが難しいという制約がある．このような背景から，Habib [8] の提案した離散連続モデルは，一日の時間軸に基づいた逐次的な選択を考慮しており，スケジューリングの過程を詳細に記述できる点で注目されている．

このように人々の活動そのものに注目したモデル開発やモデルを用いた実証的な分析が行われてきたが，これらの一連の研究において，アクティビティチェイン内の個々の活動間の依存関係についての議論は十分でなく，また特定の活動が一日の活動パターンの形成にどのような影響を与えるのかを明らかにする手法も確立していない点は，昨今のアクティビティベースの行動分析の課題の一つであると考える．そこで本研究では，アクティビティチェインを構成する活動が観測された場合，周辺の活動の生成確率に影響を与える支配的活動と，支配的活動と共起して生成される従属的な活動を特定する手法を構築する．

3. 問題設定と定式化

(1) 問題設定と仮説

本研究では，1 日単位のアクティビティチェインの生成過程をモデル化する．アクティビティチェインは N 個の活動の順序組で表現され，i 番目の活動の活動種類 (e.g. 食事・買い物・仕事) を表す活動種類を $x_i \in M$ とする．ここで M は活動種類の集合を指す．本研究では簡単のため，i 番目の活動の観測可能な属性を活動種類のみとするが，活動時間，活動場所の立地，前後の活動との移動時間等への拡張も可能である．

ここで，このような i 番目，$i+1$ 番目の活動の共起関係を生み出すメカニズムとして，支配的活動と従属的活動という潜在属性を定義する．支配的活動とは，その活動を行わなかった場合に周辺の従属的活動の生成確率を減少させる可能性のある活動であり，従属的活動とは，その活動を行わなかったとしても他の活動の生成確率を変化させない活動である．つまり，支配的活動は従属的活動の生成確率に影響を与える一方で，従属的活動は支配的活動の生成確率に影響を与えない非対称な関係性を規定する．以上の仮説より，N 個の活動で表されるアクティビティチェインの i 番目の活動は $z_i \in K$ という潜在属性を持つと仮定できる．ここで，K は潜在属性の集合であり，i 番目の活動と j 番目の活動の共起関係はチェイン型ネットワークの接続関係として $y_{ij} \in \{0, 1\}$ として表す．

これらの仮定の下で，観測可能な $(\boldsymbol{x}, \boldsymbol{y})$ として表現されるアクティビティチェインは以下の 2 つの仮説に従って生成されるとした．

- i 番目の潜在属性は，前後の活動の潜在属性との共起関係に従って生成される．
- 各潜在属性はそれぞれ固有の活動生成分布を持ち，それに従い活動種類が生成される．

(2) 活動生成のフレームワーク

以上の仮定に従い活動の生成過程を示す．

まず，各活動の潜在属性の生成過程について記載する．各活動 i に対して，潜在属性 z_i はパラメータ $\boldsymbol{\theta}$ の多項分布によって生成される．活動 i が潜在属性 k を持つ確率は θ_k と表され，各活動における潜在属性が割り当てられる確率は独立であるとすると，潜在属性 $\boldsymbol{z}$ の同時分布は，

$$P(\boldsymbol{z}|\boldsymbol{\theta}) = \prod_{k\in K} \theta_k^{n_k} \tag{1}$$

で表される．ここで，n_k は，アクティビティチェインの集合 L の中で潜在属性 k を持つ活動の数であり，次のように表す．

$$n_k = \sum_{i\in N} \delta(z_i = k) \tag{2}$$

$\delta(\cdot)$ はクロネッカーのデルタである．

次に各活動の活動種類の生成過程についてである．活動種類の生成確率は潜在属性によって決まり，各活動 i の潜在属性が z_i である時の活動種類 $x_i \in M$ はパラメータ $\boldsymbol{\pi}_{z_i}$ の多項分布によって生成される．潜在属性 z_i が活動種類 m を持つ確率は π_{m,z_i} と表され，各活動に活動種類が割り当てられる確率は潜在属性 z_i が決まった元で独立であるとすると，活動種類 $\boldsymbol{x}$ の同時分布は以下のように表される．

$$P(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{\pi}) = \prod_{m\in M}\prod_{k\in K} \pi_{m,k}^{n_{m,k}} \tag{3}$$

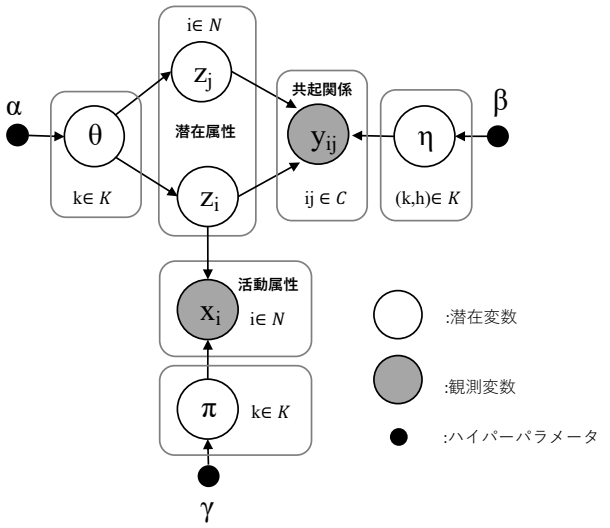
$n_{m,k}$ は潜在属性 k を持つ活動種類 m の数であり，次のように表す．

$$n_{m,k} = \sum_{i\in N} \delta(x_i = m)\delta(z_i = k) \tag{4}$$

最後に活動間の接続関係の生成過程について記載する．活動種類の生成確率は前後に隣接する活動の潜在属性によって決まる．つまり，隣接する活動との組み合わせ集合を C とすると，活動の組 $ij \in C$ に対して，接続関係 y_{ij} は潜在属性 z_i と z_j によって決まり，$\eta(z_i, z_j)$ の 2 項分布によって生成される．よって，各活動における隣接した活動との接続関係を表す確率は互いに独立であるとすると，$\boldsymbol{y}$ の同時分布は，以下のように表せる．

$$P(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{\eta}) = \prod_{k,h\in K} \eta(k,h)^{n_{k,h}} \tag{5}$$

$n_{k,h}$ は L の中で潜在属性 k と h が接続する数であり，



図–1 拡張 Stochastic Block Model のグラフィカルモデル

次のように表す．

$$n_{k,h} = \sum_{i\in N}\sum_{j\in N\setminus\{i\}} \delta(y_{i,j} = 1)\delta(z_i = k)\delta(z_j = h) \tag{6}$$

なお，バックスラッシュ「\」は，集合から特定の要素を取り除く記号である．

このように潜在属性，活動種類，接続関係の生成過程を定義することで，あるアクティビティチェインが与えられた時，パラメータ $\boldsymbol{\theta}$，$\boldsymbol{\pi}$，$\boldsymbol{\eta}$ の下でのアクティビティチェイン生成の尤度は以下のように表される．

$$P(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}|\boldsymbol{\theta},\boldsymbol{\pi},\boldsymbol{\eta}) = P(\boldsymbol{z}|\boldsymbol{\theta})P(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{\pi})P(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{\eta}) \tag{7}$$

また，上記の多項分布のパラメータの $\boldsymbol{\theta}$ と $\boldsymbol{\pi}$ は，それぞれハイパーパラメータの $\boldsymbol{\alpha}$ と $\boldsymbol{\gamma}$ のディリクレ分布から生成され，二項分布のパラメータである $\boldsymbol{\eta}$ はハイパーパラメータ $\boldsymbol{\beta}$ のベータ分布から生成される．この生成過程における変数間の依存関係をグラフィカルに表したのが**図-1** である．**図-1** からも明らかなように，潜在属性 $\boldsymbol{z}$ が生成された下では，活動種類 $\boldsymbol{x}$ と接続関係 $\boldsymbol{y}$ は互いに条件付きで独立となる．

本研究で構築した Stochastic Block Model は稲垣ら[9]で提案された拡張 Stochastic Block Model に類似する．拡張 Stochastic Block Model では，ノードの特徴値が関係性ネットワークの生成に与える影響を明示的に考慮している点が，Nowicki and Snijders [10] による Stochastic Block Model と異なる．また，稲垣らは，中心市街地の店舗集積全体を対象にネットワークを考慮し，店舗間の共起関係を捉える Stochastic Block Model を構築してるが，本研究はチェーン型の無向グラフを対象としてネットワークを考慮した．

(3) Stochastic block model のパラメータ推定

ここでは，観測可能なアクティビティ$(\boldsymbol{x},\boldsymbol{y})$の集合が観測されたもとでの，潜在属性$\boldsymbol{z}$とパラメータ$\boldsymbol{\theta}$,$\boldsymbol{\pi}$,$\boldsymbol{\eta}$の事後分布を導出する．事後分布の導出には，Collapsed Gibbs Sampling を用いることで，対象の確率モデルから一部の確率変数を周辺化し，効率的なサンプリングが可能となる．

a) 潜在属性 z の事後分布の導出

まず，潜在属性$\boldsymbol{z}$の事後分布はベイズの定理を用いることで，以下のような結合分布で表せる．

$$\begin{aligned}&P(z_i=k|\boldsymbol{y},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&\propto P(z_i=k,\boldsymbol{y},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\end{aligned}\tag{8}$$

さらにベイズの定理により以下のように展開できる．

$$\begin{aligned}&P(z_i=k,\boldsymbol{y},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&=P(\boldsymbol{y}\backslash\{y_i\},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x}\backslash\{x_i\}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&\quad\cdot P(z_i=k|\boldsymbol{y}\backslash\{y_i\},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x}\backslash\{x_i\},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&\quad\cdot P(x_i=m|z_i=k,\boldsymbol{y}\backslash\{y_i\},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x}\backslash\{x_i\},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&\quad\cdot P(y_i|x_i=m,z_i=k,\boldsymbol{y}\backslash\{y_i\},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x}\backslash\{x_i\},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\end{aligned}\tag{9}$$

ここでy_iは，活動iの前後に隣接する活動y_{ij}の集合を指し，mは活動種類で，観測データから得られる情報である．

次に，**図-1** に従い定数因子を除外し，変数の依存関係に基づいて整理すると，以下のように表せる．

$$\begin{aligned}&P(z_i=k|y,z\backslash\{z_i\},x,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&\propto P(z_i=k|\boldsymbol{\alpha})\cdot P(x_i=m|z_i=k,\boldsymbol{\gamma})\\&\quad\cdot P(y_i|z_i=k,\boldsymbol{z}\backslash\{z_i\},\boldsymbol{\beta})\\&=P(z_i=k|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\alpha})\cdot P(x_i=m|z_i=k,\boldsymbol{\pi})P(\boldsymbol{\pi}|\boldsymbol{\gamma})\\&\quad\cdot P(y_i|z_i=k,\boldsymbol{z}\backslash\{z_i\},\boldsymbol{\eta})P(\boldsymbol{\eta}|\boldsymbol{\beta})\end{aligned}\tag{10}$$

次に,$\boldsymbol{\theta}$，$\boldsymbol{\pi}$，$\boldsymbol{\eta}$を周辺化し，これらの変数を消去すると以下のように表せる．

$$\begin{aligned}&P(z_i=k|\boldsymbol{y},\boldsymbol{z}\backslash\{z_i\},\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})\\&\propto\int\int\int P(z_i=k|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\alpha})\cdot P(x_i=m|z_i=k,\boldsymbol{\pi})\\&\quad\cdot P(\boldsymbol{\pi}|\boldsymbol{\gamma})P(y_i|z_i=k,\boldsymbol{z}\backslash\{z_i\},\boldsymbol{\eta})P(\boldsymbol{\eta}|\boldsymbol{\beta})\,d\boldsymbol{\theta}\,d\boldsymbol{\pi}\,d\boldsymbol{\eta}\\&=\int P(z_i=k|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\alpha})d\boldsymbol{\theta}\\&\quad\cdot\int P(x_i=m|z_i=k,\boldsymbol{\pi})P(\boldsymbol{\pi}|\boldsymbol{\gamma})d\boldsymbol{\pi}\\&\quad\cdot\int P(y_i|z_i=k,\boldsymbol{z}\backslash\{z_i\},\boldsymbol{\eta})P(\boldsymbol{\eta}|\boldsymbol{\beta})d\boldsymbol{\eta}\\&=\mathbb{E}_{(\boldsymbol{\theta}|\boldsymbol{\alpha})}\theta_k\cdot\mathbb{E}_{(\boldsymbol{\pi}|\boldsymbol{\gamma})}\pi_{k,m}\cdot\mathbb{E}_{(\boldsymbol{\eta}|\boldsymbol{\beta})}\eta(z_i,z_j)\end{aligned}\tag{11}$$

よって，潜在属性z_iの事後分布はパラメータ$\boldsymbol{\theta}$，$\boldsymbol{\pi}$，$\boldsymbol{\eta}$の期待値に比例することがわかる．

b) パラメータ θ，π，η の期待値の導出

次に，上記の導出に従い，パラメータ$\boldsymbol{\theta}$，$\boldsymbol{\pi}$，$\boldsymbol{\eta}$の期待値の導出を行う．

まず，$\boldsymbol{\theta}$の期待値について，以下のように式展開ができる．

$$\begin{aligned}&\mathbb{E}_{(\boldsymbol{\theta}|\boldsymbol{\alpha})}\theta_k=\int P(z_i=k|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\alpha})d\boldsymbol{\theta}\\&=\prod_{k=1}^{K}\theta_k^{\sum_{i=1}^{n}\delta(z_i=k)}\prod_{k=1}^{K}\theta_k^{\alpha_k-1}\prod_{k=1}^{K}\theta_k^{n_k+\alpha_k-1}\end{aligned}\tag{12}$$

なお，$n_k=\sum_{i=1}^{n}\delta(z_i=k)$とした．ここで，(12) は$n_k+\alpha_k$をパラメータとするディリクレ分布であるので，以下のように展開できる．

$$\begin{aligned}&\prod_{k=1}^{K}\theta_k^{n_k+\alpha_k-1}=\frac{\prod_{k=1}^{K}\theta_k^{n_k+\alpha_k-1}}{\int\prod_{k'=1}^{K}\theta_k^{n_{k'}+\alpha_{k'}-1}d\boldsymbol{\theta}}\\&=\frac{\Gamma(\sum_{k=1}^{K}n_k+\alpha_k)}{\prod_{k'=1}^{K}\Gamma(n_{k'}+\alpha_{k'})}=\frac{n_k+\alpha_k}{\sum_{k'=1}^{K}n_{k'}+\alpha_{k'}}\end{aligned}\tag{13}$$

最後の展開では，$\boldsymbol{\alpha}=(\alpha_1,\alpha_2,...,\alpha_K)$をパラメータとするディリクレ分布において，

$$\mathbb{E}[\theta_k]=\frac{\alpha_k}{\alpha_0}\qquad where\quad\alpha_0=\sum_{k=1}^{K}\alpha_k\tag{14}$$

が成り立つことを用いた．

次に$\boldsymbol{\pi}$の期待値の導出においても以下のように式が展開できる．

$$\begin{aligned}&\mathbb{E}_{(\boldsymbol{\pi}|\boldsymbol{\gamma})}\pi_{k,m}=\int P(x_i=m|z_i=k,\boldsymbol{\pi})P(\boldsymbol{\pi}|\gamma)d\boldsymbol{\pi}\\&=\prod_{x_i=1}^{M}\pi_{k,m}^{\sum_{i=1}^{n}\delta(z_i=k,w_i=m)}\prod_{x_i=1}^{M}\pi_{k,x_i}^{\gamma_k-1}=\prod_{x_i=1}^{M}\pi_{k,m}^{n_{k,m}+\gamma_k-1}\end{aligned}\tag{15}$$

と表される． なお$n_{k,m}=\delta(z_i=k,w_i=m)$とした．$\boldsymbol{\theta}$と同様に$n_{k,m}+\gamma_k$をパラメータとするディリクレ分布に従うことを用いて計算することで，以下のように$\boldsymbol{\pi}$の期待値を求められる．

$$\mathbb{E}_{(\boldsymbol{\pi}|\boldsymbol{\gamma})}\pi_{k,m}=\frac{n_{k,m}+\gamma_k}{\sum_{k'=1}^{K}n_{mk'}+\gamma_{k'}}\tag{16}$$

最後に，$\boldsymbol{\eta}$の期待値も以下のように展開できる．

$$\begin{aligned}&\mathbb{E}_{(\boldsymbol{\eta}|\boldsymbol{\beta})}\eta(z_i,z_j)=\int P(y_i|z_i=k,z\backslash\{z_i\},\boldsymbol{\eta})P(\boldsymbol{\eta}|\boldsymbol{\beta})d\boldsymbol{\eta}\\&=\prod_{h\in K}\eta(k,h)^{\sum_{j=1}^{n}\delta(y_{ij}=1)\delta(z_j=h)}\prod_{h\in K}\eta(k,h)^{\boldsymbol{\beta}-1}\\&=\prod_{h\in K}\eta(k,h)^{n_{k,h}+\boldsymbol{\beta}-1}\end{aligned}\tag{17}$$

なお$n_{k,h}=\delta(y_{ij}=1)\delta(z_j=h)$とした．$\boldsymbol{\theta}$，$\boldsymbol{\pi}$の場合と同様に式 (17) は，$n_{k,h}+\boldsymbol{\beta}$をパラメータとするディリクレ分布に従うため，以下のように$\boldsymbol{\eta}$の期待値を求められる．

$$\mathbb{E}_{(\boldsymbol{\eta}|\boldsymbol{\beta})}\eta(z_i,z_j)=\prod_{h'\in K}\frac{n_{k,h}+\beta_{k,h}}{\sum_{k,h'\in C}n_{k,h'}+\beta_{k,h'}}\tag{18}$$

以上より, 式 (11) に式 (13) 式 (16) 式 (18) を代入することで，潜在属性 $\boldsymbol{z}$ の事後分布が得られる．

$$P(z_i = k|\boldsymbol{y}, \boldsymbol{z}\backslash\{z_i\}, \boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \propto \frac{n_k + \alpha_k}{\sum_{k'=1}^{K} n_{k'} + \alpha_{k'}} \cdot \frac{n_{k,m} + \gamma_m}{\sum_{k'=1}^{K} n_{k'} + \gamma_{k'}} \cdot \prod_{h' \in K} \frac{n_{k,h} + \beta_{k,h}}{\sum_{k,h' \in C} n_{k,h'} + \beta_{k,h'}} \tag{19}$$

このように，$(\boldsymbol{x},\boldsymbol{y})$ の組で表されたアクティビティチェインの集合 L が観測されたもとで，潜在属性 $\boldsymbol{z}$ の事後分布を求めることができる．同様に，$\boldsymbol{\theta}$，$\boldsymbol{\pi}$，$\boldsymbol{\eta}$ の事後分布についても求めることができる．

(4) 本モデルの活用方策

観測可能なアクティビティチェイン $(\boldsymbol{x}, \boldsymbol{y})$ が与えられた時，本手法を用いて各活動の潜在属性 z_i を確率的に推定することができるが，本研究の活用法はこれだけではない．推定したパラメータ $(\boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\eta})$，および潜在属性 $\boldsymbol{z}$ により，i 番目の活動を除去した際に前後の活動の生成に与える影響や，その結果として生成されるアクティビティチェインに与える影響を確率的に推論できる．

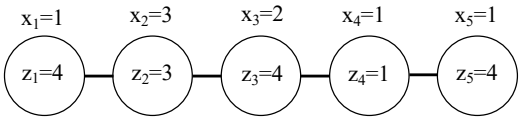
このようなモデル方策を説明するために，**図-2** のような潜在属性 $k \in \{1,2,3,4\}$, 活動種類 $x \in \{1,2,3\}$ から構成される仮想的なアクティビティチェインを考える．このアクティビティチェインは 5 つの活動から構成されており，$1 \to 3 \to 2 \to 1 \to 1$ の活動種類が順に行われたことが観測されている．さらに，3.(3) 節に従い，パラメータ $(\boldsymbol{\theta},\boldsymbol{\pi},\boldsymbol{\eta})$ と各潜在属性が確率的に推定できる．

ここで，仮に前から 2 番目の活動が行われなくなった場合のアクティビティチェインについて考えよう．前後の活動の生成有無を場合分けしアクティビティパターンを列挙すると**図-3** で示す 7 通りが考えられる．またこれらの生成確率は，パラメータ $\boldsymbol{\theta}$，$\boldsymbol{\pi}$，$\boldsymbol{\eta}$ を元に同様にして計算できるため，2 番目の活動が除かれた場合におけるアクティビティチェインの生成確率は以下のように計算できる．

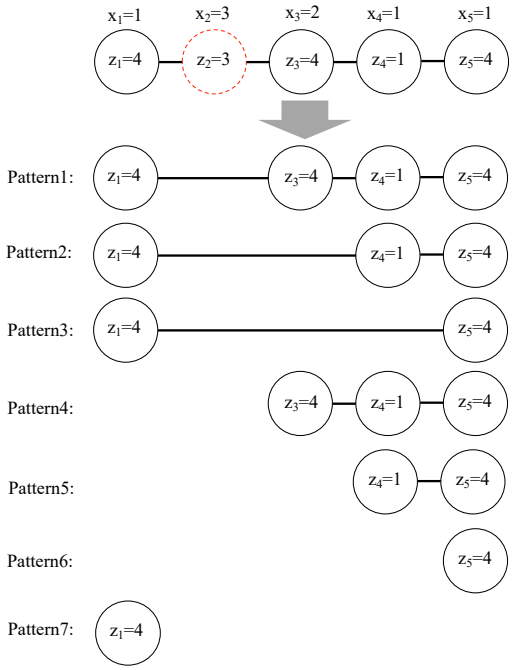
$$P(\text{pattern } k|\text{remove } x_2) = \frac{P(\text{pattern } k|\boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\eta})}{\sum_{k' \in K'} P(\text{pattern } k'|\boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\eta})} \tag{20}$$

ここで K' を $i = 2$ の活動が除かれた場合に生成しうるアクティビティチェインの集合とする．

このように，本モデルを活用することで，特定の活動が実施されなくなった場合の代替アクティビティチェインを推論でき，特定の活動が周辺活動の生成確率に影響を与える支配的活動であるか，独立な従属的活動であるかを捉えることが可能となる．



図–2 4 種類の潜在属性と 3 種類の活動属性から成る仮想的なアクティビティチェイン



図–3 $i = 2$ の活動が除かれた場合に生成しうるアクティビティチェイン

4. データを用いた検証

最後に，3. で構築した Stomatic block model を用いて，活動の接続関係に基づく潜在属性の特定が可能であることをシミュレーションデータを通して実証的に検証する．

(1) シミュレーションデータの作成

シミュレーションデータの作成では，簡単のために活動種類を $\boldsymbol{x} \in \{1,2\}$ の 2 種類，潜在属性を $\boldsymbol{z} \in \{1,2,3\}$ の 3 種類を設定し, さらに活動属性と潜在属性は表-1 のような対応関係を持つと仮定した．これは，潜在属性 $z = 3$ からは全ての活動種類が生成しうるのに対し，$z = 1,2$ からは対応する活動属性のみが必ず生成されることを示す．つまり，全ての活動を生成する可能性のある $z = 3$ が従属的活動であり，$z = 1,2$ が支配的活動である．この関係に従い，$\boldsymbol{\pi}$ のパラメータを設定した．

また，支配的活動及び従属的活動同士の生成確率は互いに独立の関係にあり，支配的活動は従属的活動の生成確率に影響を与える一方で従属的活動は支配的活

表–1 潜在変数と活動属性の対応関係

潜在属性 z	生成しうる活動属性 x
1	1
2	2
3	1,2,3

動の生成確率に影響を与えないという仮定に従い $\boldsymbol{\eta}$ のパラメータを設定した．具体的には，一日単位のアクティビティチェインに対して潜在属性 $z = h$ に接続している活動の中で $z = k$ の占める割合を，$p(z = k|z = h)$ と示すこととすると，$\boldsymbol{\eta}$ のパラメータの設定において，$p(z = 1|z = 3) > p(z = 3|z = 1)$，$p(z = 2|z = 3) > p(z = 3|z = 2)$ が成立するように設定した．一方で，その他の独立の関係にある潜在属性間の接続割合は，$p(z = k|z = h) \approx p(z = h|z = k)$ となるように設定した．

さらに，$\boldsymbol{\theta}$ についても適当な値を設定し，$\boldsymbol{\theta},\boldsymbol{\pi},\boldsymbol{\eta}$ のパラメータに従って，アクティビティチェインの集合 L を生成した．生成したデータは 1 アクティビティチェインあたり 6 つの活動からなる 3,000 人・日のアクティビティチェインであり，データ内の総活動数は 18,000 である．

(2) 推定結果

推定においては，Collapsed Gibbs Sampling における繰り返し計算毎にアクティビティチェインの集合 L 生成の対数尤度を計算した．約 150 回程度でその対数尤度が収束したことを確認し，適切に推定がなされたと判断した．また，潜在属性の数は $|K| = 3$ とした．

a) $\boldsymbol{\theta}$ の推定結果

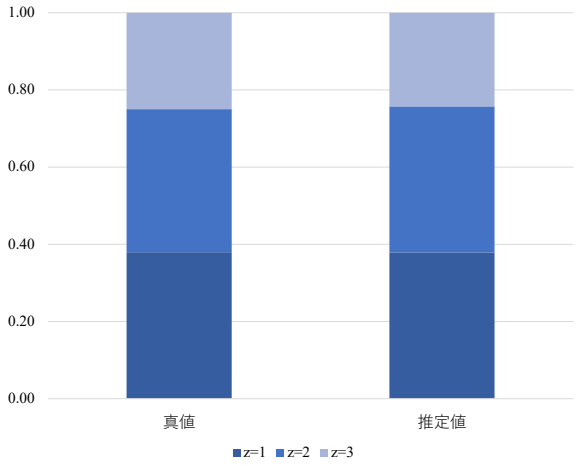
まずパラメータ $\boldsymbol{\theta}$ の推定結果を図-4 に示す．$\boldsymbol{\theta}$ の推定値はシミュレーションデータ作成時に設定した真値と非常に類似しており，上記の推定方法によって適切に推定ができていることが確認できる．

b) $\boldsymbol{\pi}$ の推定結果

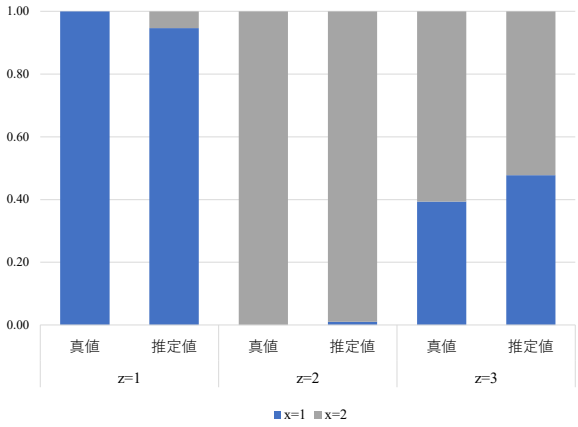
次に，$\boldsymbol{\pi}$ の推定結果を図-5 に示す．図-5 より，z=3 のみが x=1,2 の両方の活動属性を生成しうる確率をもち，一方で，z=1 については，活動属性 x=1 のみ，z=2 については，活動属性 x=2 のみにおいて生成確率が突出して高くなっていることが示せたことから，支配的活動と従属的活動を明確に区別できる結果が得られた．また真値と推定値の差分は小さく，適切に推定されていることが確認できる．

c) $\boldsymbol{\eta}$ の推定結果

最後に，$\boldsymbol{\eta}$ の推定結果を図-6 に示す．図-6 は，x 軸に真値，y 軸に推定値をプロットすることで，これらの対応関係を表す．全ての真値と推定値の組は 45 度線の周



図–4 $\boldsymbol{\theta}$ の推定結果



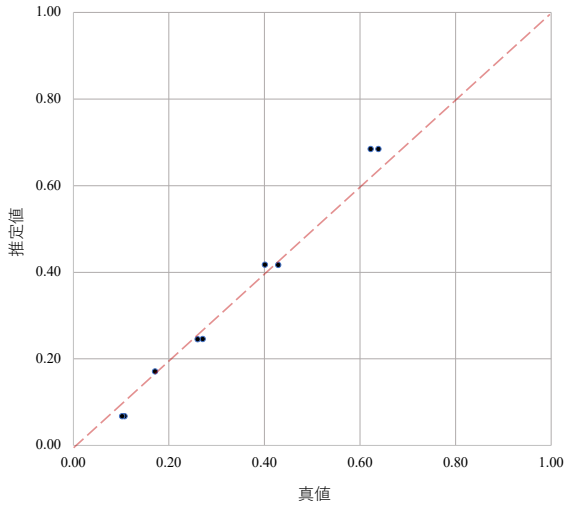
図–5 π の推定結果

辺にプロットされており，推定値は真値に近い値として適切に推定されていることが確認できる．

以上より，観測された活動属性 $\boldsymbol{x}$ のみから，パラメータ $\boldsymbol{\theta},\boldsymbol{\pi},\boldsymbol{\eta}$ を推定できることを確認した．

5. おわりに

本研究では，活動を生み出すメカニズムとして，支配的活動と従属的活動を定義し，これらの共起関係に基づいたアクティビティチェインの生成過程を定式化した．さらに，アクティビティチェインに含まれる各活動をノード，活動間の接続関係をリンクとしたチェーン型の無向ネットワークを仮定することで，観測データから得られる情報のみ (e.g. 活動種類) から属性の持つ性質や属性間の関係，属性と活動種類との関係を踏まえたパラメータの導出と推定が可能であることをシミュレーションデータを用いた検証により実証的に示



図–6 η の推定結果

した．また，今回は簡単のため，観測可能な情報を活動種類のみとしているが，活動時間や移動時間，個人属性といった様々な変数を用いたモデルの拡張が可能である．さらに，本モデルの適用方策は潜在属性の特定にとどまらず，特定の活動が除去・追加されることによる 1 日の行動パターンの変化を確率的に推論できる可能性を示した．このような活動間の相互依存関係を捉える手法の開発は，連鎖的な活動を促進させる都市計画を設計する上で重要な知見となる．

本稿ではシミュレーションデータのみを用いて，手法の確認・検証を行なっているが，今後は実際に観測されたアクティビティチェインデータを用いて，各活動の支配的活動・従属的活動の推論を行い，それぞれの活動が除去・追加された時に 1 日のアクティビティチェインがどのように変化するか検証を行う．

謝辞： 本研究は JST 創発的研究支援事業 JPMJFR225U の助成を受けた研究の一部です．ここに記し，感謝を表します．

REFERENCES

1) 谷口守, 橋本成仁, and 植田拓磨: 個人行動特性に配慮した買物行動のサイバー空間への潜在的な代替性把握, 土木学会論文集 D, Vol.66, No.2, pp.290–299, 2010.
2) Jaller, M. and Pahwa, A.: Evaluating the environmental impacts of online shopping: A behavioral and transportation approach, Transportation Research Part D: Transport and Environment, Vol.80, pp.102223, 2020.
3) 谷口守, 阿部宏史, and 蓮実綾子: サイバーウォークにおける空間抵抗特性とそのタウンウォークとの代替性, 土木計画学研究・論文集, Vol.20, pp.477–483, 2003.
4) Bhat, C. R. and Koppelman, F. S.: Activity-based modeling of travel demand, Handbook of transportation Science, pp. 35–61, Springer, 1999.
5) Bowman, J. L. and Ben-Akiva, M. E.: Activity-based disaggregate travel demand model system with activity schedules, Transportation research part a: policy and practice, Vol.35, No.1, pp.1–28, 2001.
6) 福田大輔 and 力石真: 離散-連続モデルの研究動向に関するレビュー, 土木学会論文集 D3 (土木計画学), Vol.69, No.5, pp.I_497–I_510, 2013.
7) Bhat, C. R.: A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions, Transportation Research Part B: Methodological, Vol.39, No.8, pp.679–707, 2005.
8) Habib, K. M. N.: A random utility maximization (rum) based dynamic activity scheduling model: Application in weekend activity scheduling, Transportation, Vol.38, No.1, pp.123–151, 2011.
9) 稲垣和哉, 原祐輔, and 桑原雅夫: 店舗の空間的・業種的共起関係に基づいた中心市街地の潜在構造分析, 土木学会論文集 D3 (土木計画学), Vol.72, No.5, pp.I_473–I_485, 2016.
10) Nowicki, K. and Snijders, T. A. B.: Estimation and prediction for stochastic blockstructures, Journal of the American statistical association, Vol.96, No.455, pp.1077–1087, 2001.

(Received 2023. 10. 13)
(Accepted 2023. 10. 13)

IDENTIFICATION OF DOMINANT AND DEPENDENT ACTIVITIES BASED ON OBSERVED ACTIVITY-TRAVEL BEHAVIOR

Yurika TAKANO and Yusuke HARA