2014年9月27日@東大

行動モデル夏の学校2014

講義1:行動モデルの基礎理論と応用事例

行動モデルの応用事例＋α

― 首都圏鉄道需要予測を例に ―

東京大学大学院工学系研究科

特任助教　柳沼秀樹

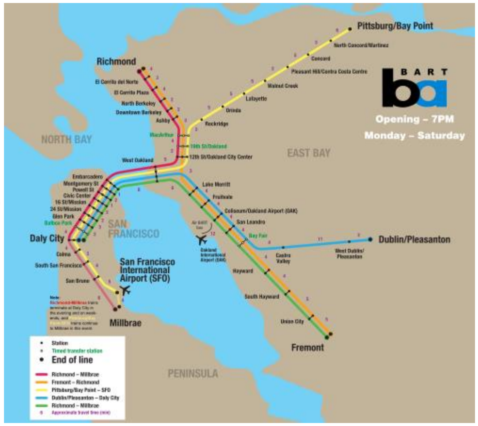
yaginuma@civil.t.u-tokyo.ac.jp

1. 行動モデルの応用範囲

Bay Area Rapid Transit(BART)の需要予測

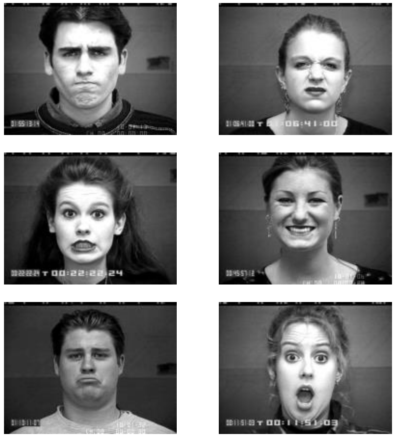
(Domencich and McFadden 1975)

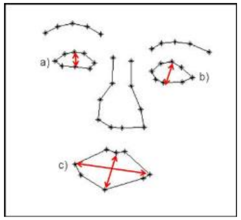
離散選択モデルの応用は
「鉄道需要予測」から始まった！



多様かつ広範囲な適用事例

- 交通需要予測(目的地，交通機関，経路)
- 生活・活動パターン
- 就学・就職・職業選択
- 消費活動(嗜好品，住宅，電話プラン)
- 立地・居住地
- 政党支持，政策評価
- パターン認識…



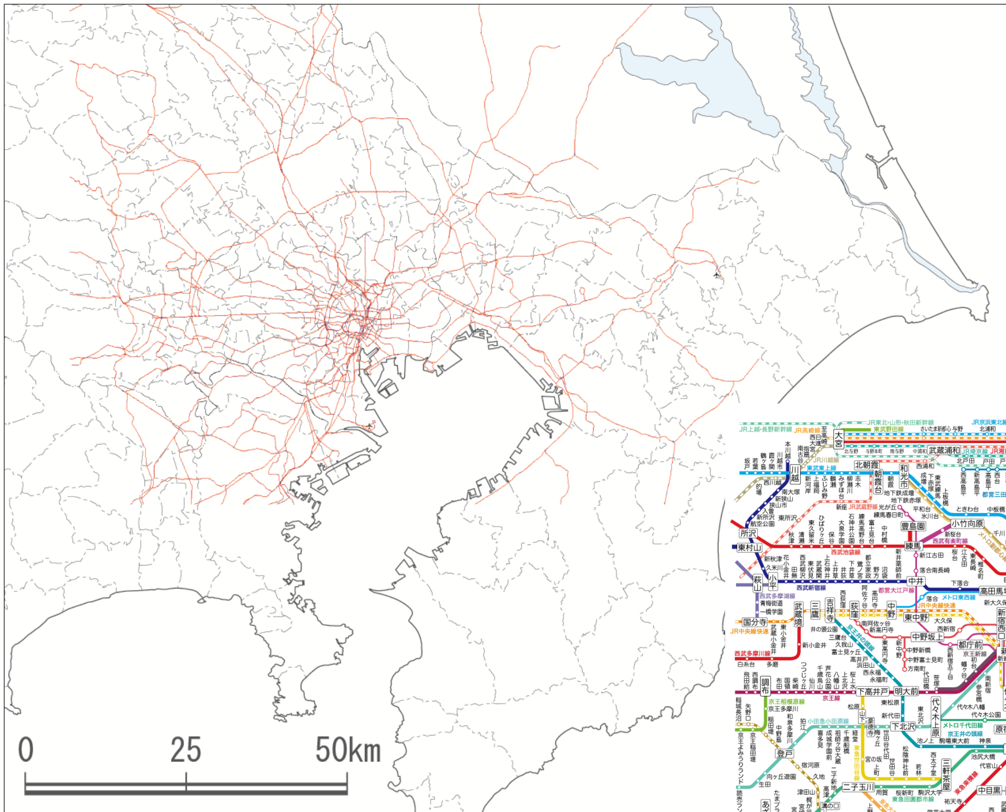


$$P(i)=\frac{\exp(\mu V_i)}{\sum_{j\in C}\exp(\mu V_j)}$$

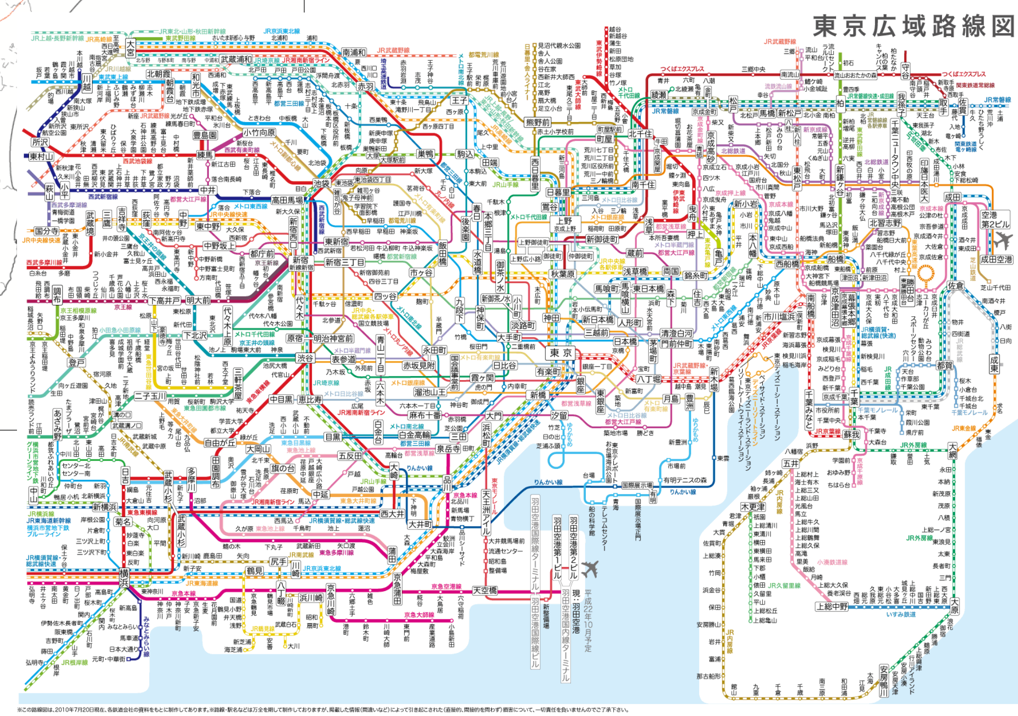
離散的で効用関数さえ特定できれば多くの人間行動に応用可能！！

2. 首都圏鉄道の概要

世界で最も広域･高密度･高頻度な鉄道網

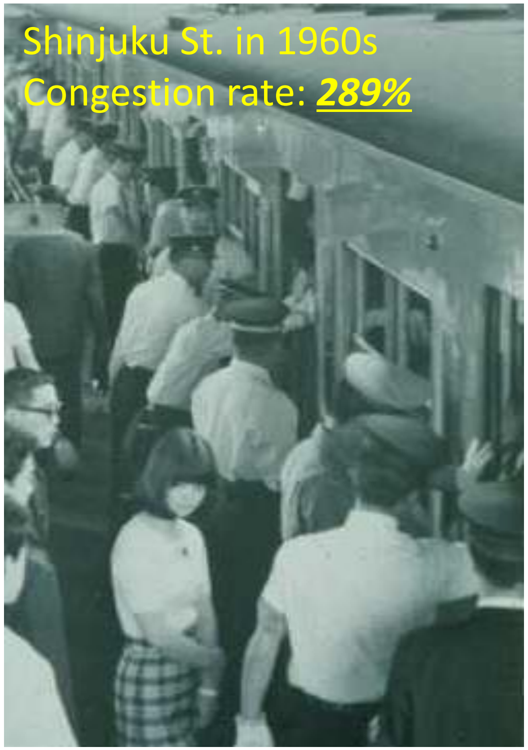


首都圏:東京駅から半径70km
路線数:約130路線
駅数:約1800駅
利用者数:4000万人/日
複数の運営主体が存在



2. 首都圏鉄道の混雑問題

1960年代から現在まで続く解決すべき問題



【参考:運輸政策研究機構(H5)】

- ロンドン : 149%
- パリ : 152%
- ニューヨーク : 71%

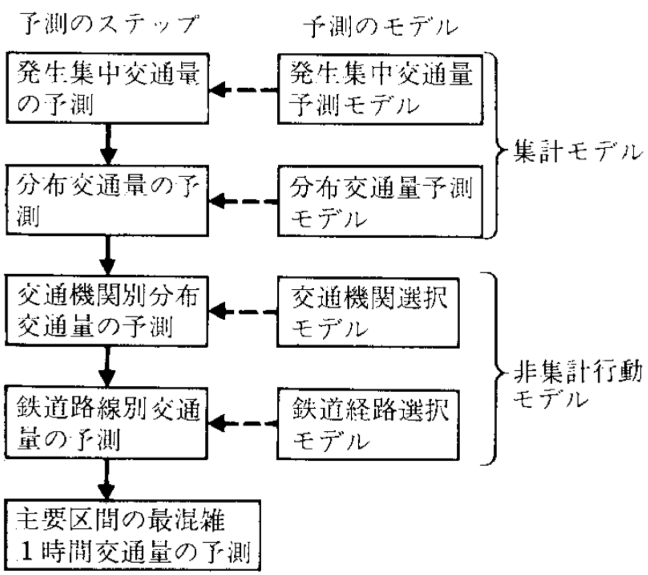
2. 首都圏鉄道の需要予測

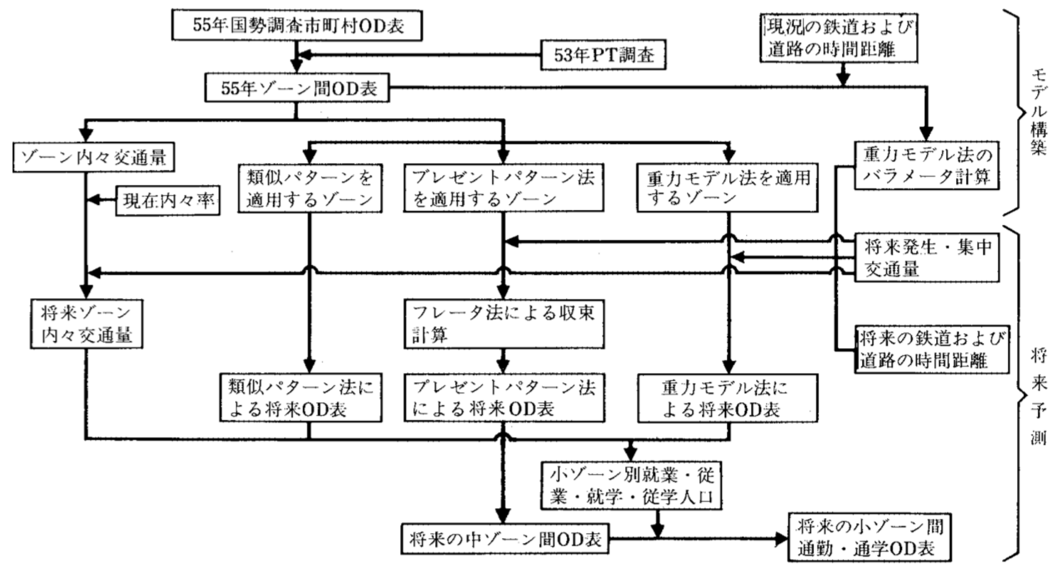
➢ <u>都市交通審議会15号答申(1972)</u>

集計モデル(グラビティモデル)を用いた交通量予測

➢ <u>運輸政策審議会7号答申(1985)</u>

機関分担・経路配分に非集計ロジットを適用





➢ <u>運輸政策審議会18号答申(2000)</u>

経路配分に構造化プロビットを適用（当時の最新手法を即適用！）

3. MNP概要

<u>多項ロジットモデル(MNL)</u>

$$P(i) = \frac{\exp(\mu V_i)}{\sum_{j \in C} \exp(\mu V_j)}$$

- Luce(1959), McFadden(1974)
- Closed-form
- 選択肢間の相関を無視(IIA)

通常の最適化アルゴリズムでパラメータ推定が可能であり、ソースコードの開発が容易

⇒爆発的に各分野に普及

<u>多項プロビットモデル(MNP)</u>

$$P(i) = \int_{\varepsilon_1=-\infty}^{\varepsilon_i+V_i-\varepsilon_1} \cdots \int_{\varepsilon_i=-\infty}^{\infty} \cdots \int_{\varepsilon_J=-\infty}^{\varepsilon_i+V_i-\varepsilon_J} \phi(\varepsilon) d\varepsilon_J \cdots d\varepsilon_1$$

$$\phi(\varepsilon) = \frac{1}{\left(\sqrt{2\pi}\right)^{J-1} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\varepsilon \Sigma^{-1} \varepsilon'\right)$$

- Thurstone(1927)
- Open-form
- 選択肢間の相関を表現

選択肢数−1の多重積分が必要であり、パラメータ推定は非常に煩雑

⇒計算が困難であるため
近年まで伸び悩む…

4. 構造化プロビットモデル(1)

Multinomial Probit with Structured Covariance for Route Choice Behavior,
Transportation Research Part B, Vol.31, No.3, pp195-207, 1997.

森地茂 先生

屋井鉄雄 先生

岩倉成志 先生

首都圏鉄道での経路重複問題への適用を前提にプロビットモデルの誤差項を構造化．基本的アイディアは1993年の和文論文から見られる．

※当時，世界中で類似研究が盛んに行われており，
最もアツイ研究テーマ
Ex: C-Logit (1997), CNL (1997), MXL (2000)

Pergamon

Transpn Res.-B, Vol. 31, No. 3, pp. 195–207, 1997

0191-2615/97 $17.00 + 0.00

PII: S0191-2615(96)00025-2

MULTINOMIAL PROBIT WITH STRUCTURED COVARIANCE FOR ROUTE CHOICE BEHAVIOR

TETSUO YAI
Department of Civil Engineering, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152, Japan

SEIJI IWAKURA
Japan Transport Economics Research Center, 3-18-19 Toranomon, Minato-ku, Tokyo, Japan

and

SHIGERU MORICHI
Department of Civil Engineering, University of Tokyo, Hongo,Bunkyo-ku, Tokyo 112, Japan

(Received 8 September 1994; in revised form 19 June 1996)

Abstract—We propose another version of the multinomial probit model with a structured covariance matrix to represent any overlapped relation between route alternatives. The fundamental ideas of the model were presented in Yai *et al.* (1993) and Yai and Iwakura (1994). The assumptions introduced in the model may be more realistic for route choice behaviors on a dense network than the strict assumption of the independent alternative property of the multinomial logit model. As the nested logit model assumes an identical dispersion parameter between two modeling levels for all trip makers, the model has difficulty in expressing individual choice-tree structures. To improve the applicability of the multinomial probit model to route choice behaviors, we introduce a function which represents an overlapped relation between pairs of alternatives and propose a multinomial probit model in which the structured covariance matrix uses the function in order to consider the individual choice-tree structures in the matrix and the estimatability of the new alternative's covariances. After examining the applicability of the multinomial probit model using empirical route choice data in a Tokyo metropolitan region, we also propose a method for evaluating consumer benefits on complicated networks based on the multinomial probit model. © 1997 Elsevier Science Ltd

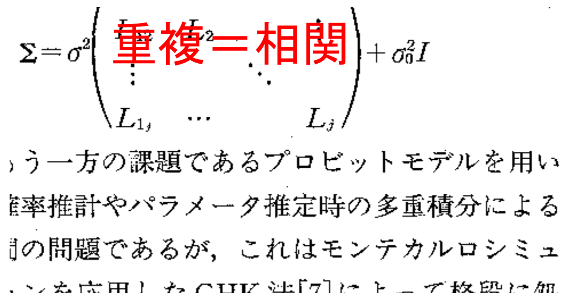
1. INTRODUCTION

The applications of the multinomial probit model have not been adequately successful in spite of its advantages in flexibility of the model form. Certainly, the complexity of the computational process has deterred its use, compared to the wide applications of the multinomial logit models. Early advances in the estimation method of the multinomial probit model were achieved before the early 80s, by Daganzo (1977), Lerman and Manski (1981), Daganzo and Sheffi (1982) and Sheffi *et al.* (1982). Their work discussed alternative methods for estimating the covariance matrix simultaneously with utility function parameters. No accurate method was found during these earlier advances and thus the multinomial probit model was not widely applied (Horowitz *et al.*, 1982; Horowitz, 1991). In the 1980s, most discrete choice models were calibrated by the multinomial logit model or expansion forms of the multinomial logit such as the nested logit model. Although most results were satisfactory in representing travel behaviors of modal choices, several behaviors which do not satisfy the assumptions of the multinomial logit model exist. Most probably, the cause of such behaviors is the interdependency of choice alternatives.

Recently, there have been advances in multinomial probit estimation (McFadden, 1989; Pakes and Pollard, 1989; Bunch, 1991; Bolduc and Ben-Akiva, 1991; Bolduc, 1992; Geweke *et al.*, 1994). The method of simulated moments proposed by McFadden seems to encourage multinomial probit applications because of its computational efficiency in seeking model parameters. Bolduc focused on the estimation of the multinomial probit model with a large choice set using auto-regressive errors with distance related functions among alternatives for simplifying its covariance matrix. Bunch simplified the multinomial probit model's covariance matrix with his transformation method which lessens the estimation problem. Geweke *et al.* compared several

4. 構造化プロビットモデル(2)

首都圏鉄道では経路重複が存在するため，IIAを仮定したロジットでは便益計測時に誤差が発生
⇒プロビットが有効



課題

- OD毎に経路選択集合が異なるため，共分散が設定できない
  ⇒経路長に依存する誤差と経路固有の誤差に分解
- 計算コストが非常に高い（選択肢数−1の多重積分）
  ⇒シミュレーション法(GHK)による高速化

$$U_i = V_i + e_i$$

経路長に依存する誤差

$$e_i = e_i^{Length} + e_i^{Route}$$

経路固有の誤差

考え方はMXLと同じ！

$$U_{in} = V_{in} + [\eta_{in} + \varepsilon_{in}]$$

4. 構造化プロビットモデル(3)

誤差項とその共分散

経路長に依存する誤差

経路固有の誤差

$$\varepsilon_r = \varepsilon_r^{\ 1} + \varepsilon_r^{\ 0}$$

$$\Sigma = \Sigma^1 + \Sigma^0$$

経路長に依存する誤差

・経路長が長いほど効用の分散は大

$$Var(\varepsilon_r^{\ 1}) = L_r \sigma^2$$

・重複区間が長いほど2経路間の共分散は大

⇒共分散は重複距離の分散に等しい

$$Cov(\varepsilon_r^{\ 1}, \varepsilon_q^{\ 1}) = L_{rq}\sigma^2$$

経路固有の誤差

・経路毎に独立に発生(共分散＝0)

$$Cov(\varepsilon_r^{\ 0}, \varepsilon_q^{\ 0}) = \sigma_0^{\ 2}, \quad q = r$$
$$= 0, \quad q \neq r$$

$$\Sigma = \sigma^2 \begin{pmatrix} L_1 & L_{12} & \cdots & L_{1R} \\ L_{12} & L_2 & \cdots & L_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ L_{1R} & L_{2R} & \cdots & L_R \end{pmatrix} + \sigma_0^2 I$$

分散比で表現

$$\Sigma = \sigma_0^2 \begin{pmatrix} \eta L_1 + 1 & \eta L_{12} & \cdots & \eta L_{1R} \\ \eta L_{12} & \eta L_2 + 1 & \cdots & \eta L_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \eta L_{1R} & \eta L_{2R} & \cdots & \eta L_R + 1 \end{pmatrix}$$

$$\eta = \frac{\sigma^2}{\sigma_0^2}$$

分散比のみを推定するだけで良い！

L_r :経路rの経路長

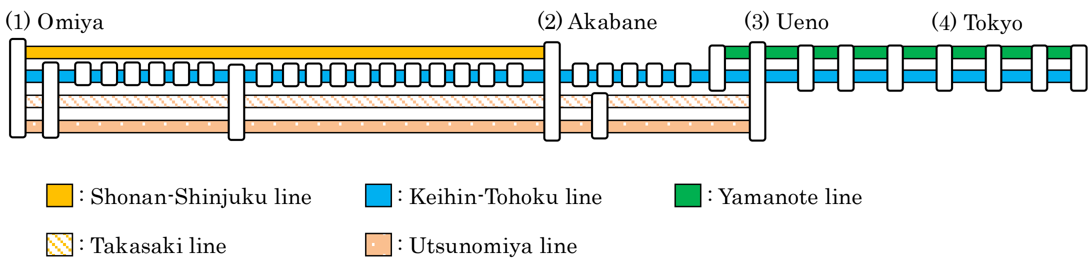
L_{rq} :経路rとqの重複経路長

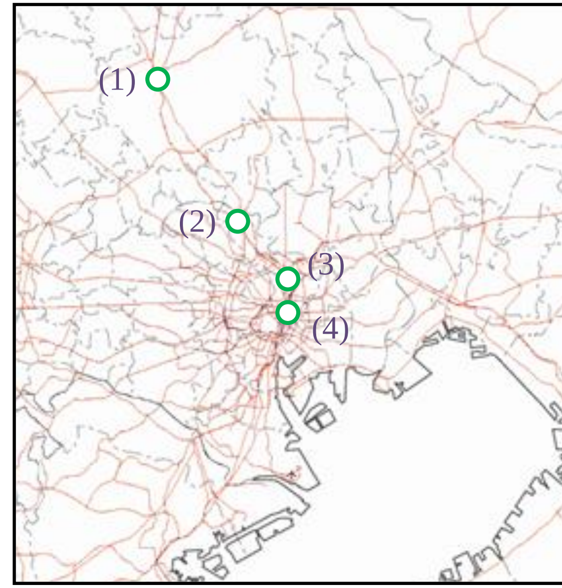
σ^2 :単位距離あたりの分散

5. 構造化プロビットモデル(4)

運輸政策18号答申に構造化プロビットを導入(岩倉2010)

例：大宮ー神田間での再現





パラメータ推定結果

説明変数	通勤目的 推定値	通勤目的 t値
乗車時間(分)	−0.0943	−8.09
アクセス・イグレス時間(分)	−0.127	−11.7
乗換時間(分)	−0.112	−10.7
運賃(円)	−0.002	−3.98
混雑指標	−0.00869	−3.34
分散パラメータ	0.436	2.71
尤度比	0.39	
サンプル数	1218	

現況再現結果

発着駅	各経路の選択確率	実績(H7センサス)		MNL 独立性非考慮		MNPSC 独立性考慮	
大宮駅	i 東北→山手	33%	48%	28%	52%	27%	47%
→	j 東北→京浜東北	15%		24%		20%	
神田駅	k 京浜東北	53%		47%		52%	
	選択確率計	100%		100%		100%	

経路重複の緩和により高い予測精度を実現

(全駅間断面交通量は実績の±10%)

6. まとめ

- 首都圏の鉄道需要予測では，高い表現力を有する構造化プロビットが用いられている．また，当該モデルは目的選択や出発時刻選択にも適用可能．
- MXLよりも優れてる？（清水1999）

⇒計算時間が4倍速く計算時間の変動は1/10かつ推定パラメータの変動係数もMXLより小さい

- 無論，Open−formなので計算コストは高いが，シミュレーション法により高速化が可能．

⇒近年ではGHKやMCMC等のシミュレーションに加えて，MACML等の解析近似手法もある

- 構造化による重複は考慮できるが，今後は選択肢集合（経路集合）の設定方法が課題．

⇒Hyperpath（一般化費用が等しい経路群）を用いた経路列挙の可能性を模索中

[付録] 首都圏鉄道施策の変遷

首都圏における答申の概要

答申名	発表年次	目標年次	背景	施策
都市交通審議会答申第 1 号	1956	1975	急激な発展と都市化に伴う交通需要の増大	新規路線整備，相互直通運転
都市交通審議会答申第 4 号	1960	—	モータリゼーションの進展に伴う路面交通問題の顕在化	路面電車の撤去
都市交通審議会答申第 6 号	1962	1975	さらなる交通需要の増大と混雑問題の深刻化	1 号答申（1956）の路線見直し，5 路線の新規開発
都市交通審議会答申第 9 号	1966	1985	横浜周辺地域における人口増加とさらなる混雑への懸念	最混雑区間における混雑時 1 時間の混雑率が 200 %を越える路線について，新規整備もしくは複々線化を検討
都市交通審議会答申第 10 号	1968	1975	急激な都心部の拡大と副都心部の発展による需要の増大に対する緊急対策	地下鉄計画の見直しと追加
都市交通審議会答申第 15 号	1972	1985	通勤・通学の長距離化	大規模延伸による郊外との接続，新規路線整備（副都心線），空港アクセス鉄道整備，輸送力拡大，高速化による時間短縮
運輸政策審議会答申第 7 号	1985	2000	人口増加と都市構造の変化による混雑と通勤・通学の長時間化	新規国鉄路線（埼京線，京葉線，常磐新線）整備，羽田空港アクセス路線整備，貨物船整備
運輸政策新議会答申第 18 号	2000	2015	副都心への業務地集積，混雑の二極化，ピーク時の速度低下，高齢化社会への移行	供給施策による混雑緩和，相互直通による速達性の向上，バリアフリー化・シームレス化の推進

[付録] MNPの推定方法

開発当初から多重積分の計算がネックとなり、各種の数値計算アルゴリズムが提案

	選択肢数	手法	
1970			
	3～4	• 級数近似法 • Clarkの近似法etc…	解析的近似
1990			
	3～10	• シミュレーション法 • ベイズ推定	乱数近似
	10以上		
2010		• MACML	

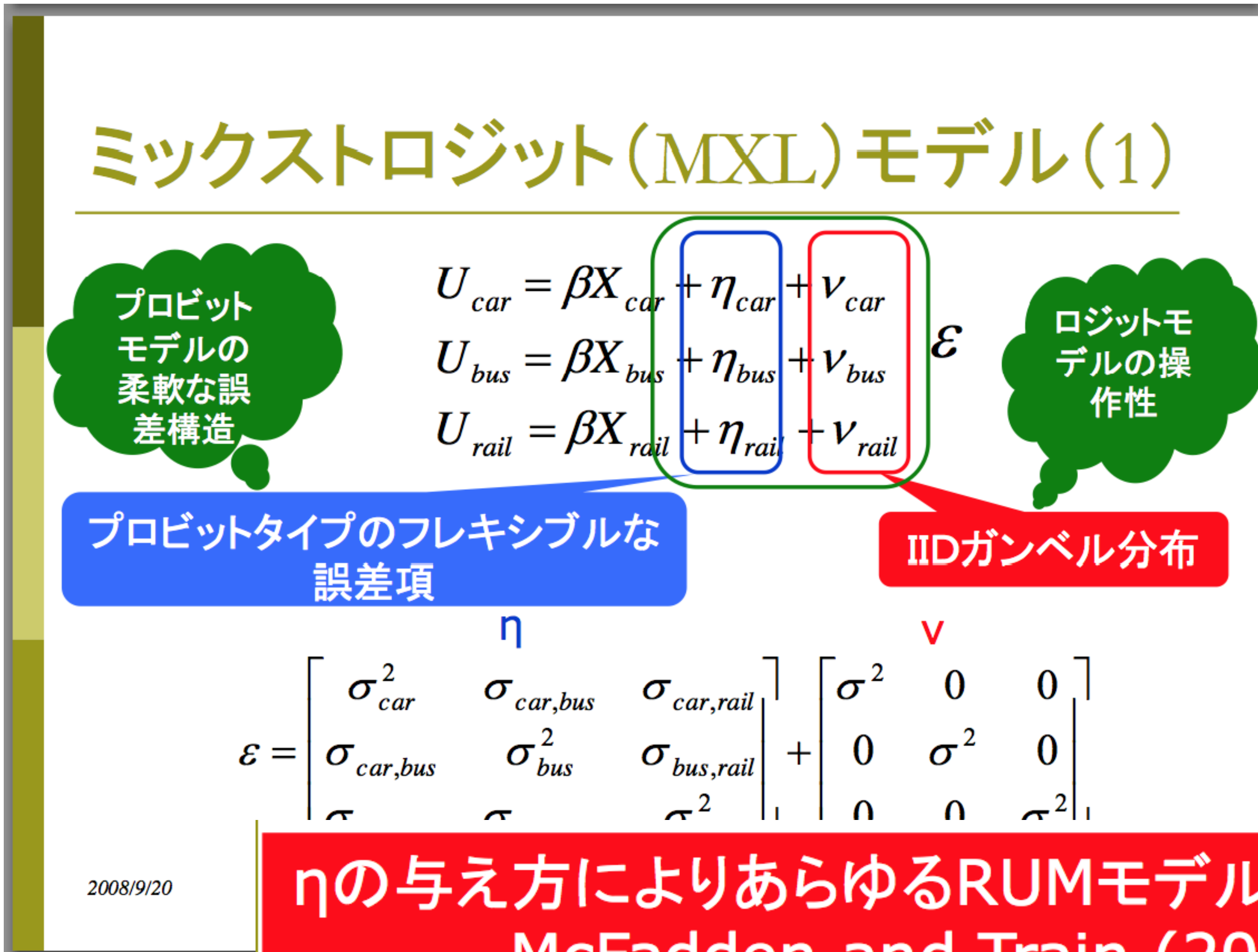
Bhat, C.R.: The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.923-939, 2011.

参考文献

1) 八十島義之助：東京の通勤鉄道線網計画に関する研究，土木学会論文集，Vol.371, pp.31-43, 1986.
2) 森地茂監修，東京圏鉄道整備研究会編：首都圏の鉄道のあゆみと未来[解説]，運輸政策研究機構，2000.
3) 屋井鉄雄, 岩倉成志，伊藤誠: 鉄道ネットワークの需要と余剰の推計法について，土木計画学研究・論文集，No.11, pp.81-88, 1993.
4) Yai, T., Iwakura, S. and Morichi, S.: Multinomial Probit with Structured Covariance for Route Choice Behavior, Transportation Research Part B, Vol.31, No.3, pp195-207, 1997.
5) 岩倉成志: 東京圏の都市圏鉄道計画における構造化プロビットモデル, オペレーションズリサーチ, Vol. 55, pp. 159-163, 2010
6) 清水哲夫, 屋井鉄雄: Mixed Logit Modelとプロビットモデルの推定特性に関する比較分析, 土木計画学研究・論文集, No. 16, pp. 587-590, 1999.
7) 屋井鉄雄, 中川隆広, 石塚順一, シミュレーション法による構造化プロビットモデルの推定特性, 土木学会論文集, No. 604/IV-41, pp. 11-21, 1998.
8) 屋井鉄雄,中川隆広: 構造化プロビットモデルの発展性，土木計画学研究・論文集，No.13, pp.563-570, 1996.
9) 屋井鉄雄, 清水哲夫, 坂井康一,小林亜紀子：非IIA型選択モデルの選択肢集合とパラメータ特性, 土木学会論文集, No. 702/IV-55, pp. 3-13, 2002.

[オマケ] Mixed Logit のすごさ

夏の学校にて（倉内先生発表資料）



ηの与え方によりあらゆるRUMモデルが近似可能
McFadden and Train (2000)

[オマケ] 夏の学校のすごさ


ホーム > bin studies -都市と移動- > 初学者用サンプルプログラム

BinN studies シリーズ

ネットワーク行動学 -都市と移動- 初学者用サンプルプログラム 公開ページ

ネットワーク行動学 -都市と移動- は，都市計画・交通計画をはじめとした幅広い分野の研
ト・学習ポータルサイトです． 初学者用サンプルプログラム公開ページでは，プローブデータ
グ，行動モデル，ネットワークモデルとその適用例について，データとプログラムコード，そ
す．

サンプルデータ/サンプルコード

プログラム言語R，Javaの使い方，ならびに各モデルのサンプルコードとサンプルデータ，および仕様書です．
コードとデータはzipファイルをダウンロードしてご利用ください．

R

Rの導入		仕様書	2014年09月01日更新
Rのインストール方法について解説しています．			
Multinomial Logit(MNL)モデル(交通手段選択)	zipファイル	仕様書	2014年09月01日更新
選択肢相関を考慮しない最も基本的なモデルです．			
Nested Logit(NL)モデル(交通手段選択)	zipファイル	仕様書	2014年09月01日更新
選択肢間の相関を扱うことのできるモデルです．			
Mixed Logit(MXL)モデル(交通手段選択)	zipファイル	仕様書	2014年09月01日更新
確率項の誤差項をわけることで，選択肢間の相関をより多く考慮することのできるモデルです．			
Cross Nested Logit(CNL)モデル(経路選択)	zipファイル	仕様書	2014年09月01日更新
NLの特徴を緩和し，選択肢間の相関をより多く考慮することができるモデルです．			

Mixed Logit(MXL)モデル(交通手段選択)

確率項の誤差項をわけることで，選択肢間の相関をより多く考慮することのできるモデルです．

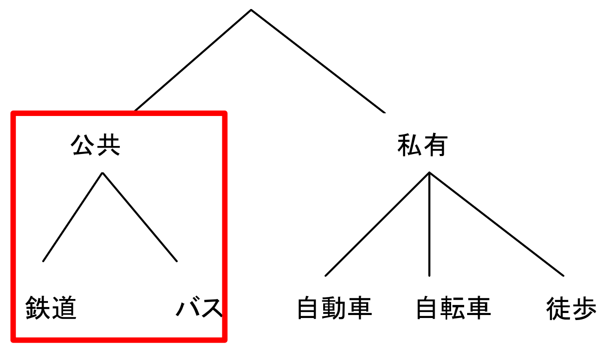
（恐れずに）
使ってみよう！

URL:http://bin.t.u-tokyo.ac.jp/kaken/program.html

[オマケ] MXLでのNLとCNL

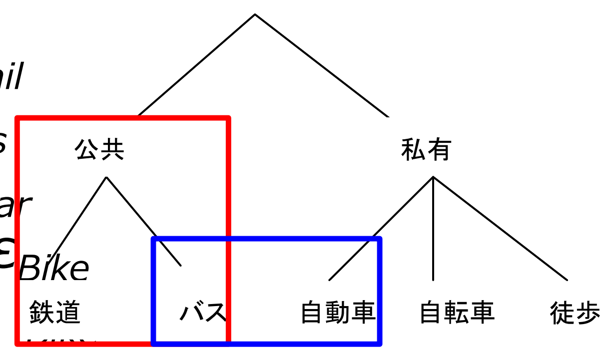
NL型

$$
\begin{aligned}
U_{Rail} &= V_{Rail} + \sigma_{public}\,\eta_{public} + \varepsilon_{Rail}\\
U_{Bux} &= V_{Bus} + \sigma_{public}\,\eta_{public} + \varepsilon_{Bus}\\
U_{Car} &= V_{Car} + \varepsilon_{Car}\\
U_{Bike} &= V_{Bike} + \varepsilon_{Bike}\\
U_{Walk} &= V_{Walk} + \varepsilon_{Bike}
\end{aligned}
$$



CNL型

$$
\begin{aligned}
U_{Rail} &= V_{Rail} + \sigma_{public}\,\eta_{public} + \varepsilon_{Rail}\\
U_{Bux} &= V_{Bus} + \sigma_{public}\,\eta_{public} + \sigma_{Road}\,\eta_{Road} + \varepsilon_{Bus}\\
U_{Car} &= V_{Car} + \sigma_{Road}\,\eta_{Road} + \varepsilon_{Car}\\
U_{Bike} &= V_{Bike} + \varepsilon_{Bike}\\
U_{Walk} &= V_{Walk} +
\end{aligned}
$$

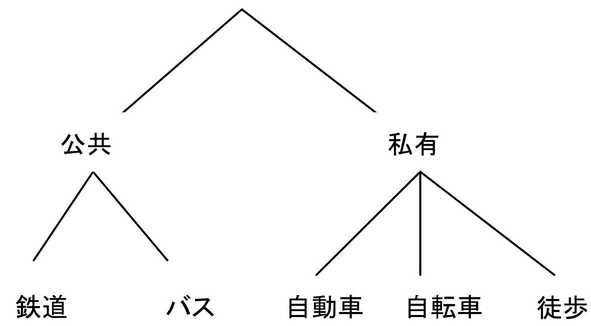


η_{public} は標準正規乱数, $N(0,1)$

[オマケ] 誤差構造の差異

Nested Logit

$$
\begin{bmatrix}
\sigma^2 & \sigma^2_{public} & 0 & 0 & 0 \\
\sigma^2_{public} & \sigma^2 & 0 & 0 & 0 \\
0 & 0 & \sigma^2 & 0 & 0 \\
0 & 0 & 0 & \sigma^2 & 0 \\
0 & 0 & 0 & 0 & \sigma^2
\end{bmatrix}
$$



異なる！！

Mixed Logit

$$
\begin{array}{ccc}
\eta & & \varepsilon \\
\begin{bmatrix}
\sigma^2_{public} & \sigma^2_{public} & 0 & 0 & 0 \\
\sigma^2_{public} & \sigma^2_{public} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
& + &
\begin{bmatrix}
\sigma^2 & 0 & 0 & 0 & 0 \\
0 & \sigma^2 & 0 & 0 & 0 \\
0 & 0 & \sigma^2 & 0 & 0 \\
0 & 0 & 0 & \sigma^2 & 0 \\
0 & 0 & 0 & 0 & \sigma^2
\end{bmatrix}
\end{array}
=
\begin{bmatrix}
\sigma^2_{public} + \sigma^2 & \sigma^2_{public} & 0 & 0 & 0 \\
\sigma^2_{public} & \sigma^2_{public} + \sigma^2 & 0 & 0 & 0 \\
0 & 0 & \sigma^2 & 0 & 0 \\
0 & 0 & 0 & \sigma^2 & 0 \\
0 & 0 & 0 & 0 & \sigma^2
\end{bmatrix}
$$

[オマケ] 計算例（NL型）

パラメータ設定

```
## 目的地までの所要時間
d1 <- x[5]

## 料金
f1 <- x[6]

## ネストパラメータもどき
sig1 <- x[7]
```

シミュレーション設定

```
##R回シミュレーションを行うループ
for (i in 1:R) {

  ## 対数尤度のための変数を宣言
  LL = 0

  ##乱数からパラメータをシミュレーションで発生
  sig <- sig1 * rand1[,i]
```

効用関数の設定

```
る.                                        # 時間                      # 料金
xp(d1*Data$総所要時間train/100 +f1*Data$費用train/100 + sig + b1*matrix(1,nrow =hh,ncol=1))
xp(d1*Data$総所要時間bus/100   +f1*Data$費用bus/100   + sig + b2*matrix(1,nrow =hh,ncol=1))
xp(d1*Data$所要時間car/100                          + b3*matrix(1,nrow =hh,ncol=1))
xp(d1*Data$所要時間bike/100                         + b4*matrix(1,nrow =hh,ncol=1))
xp(d1*Data$所要時間walk/100                                                         )
```

推定結果

```
> ##パラメータ推定値
> print(b)
[1]  0.491643035 -1.608301398 -0.837015629 -0.771393905 -8.707322307  0.085114809
[7]  0.009606772
> ## t値
> print(tval)
[1]  1.7090005 -4.0422896 -3.8190829 -3.6785025 -9.7626046  1.3687167  0.0725635
>
```