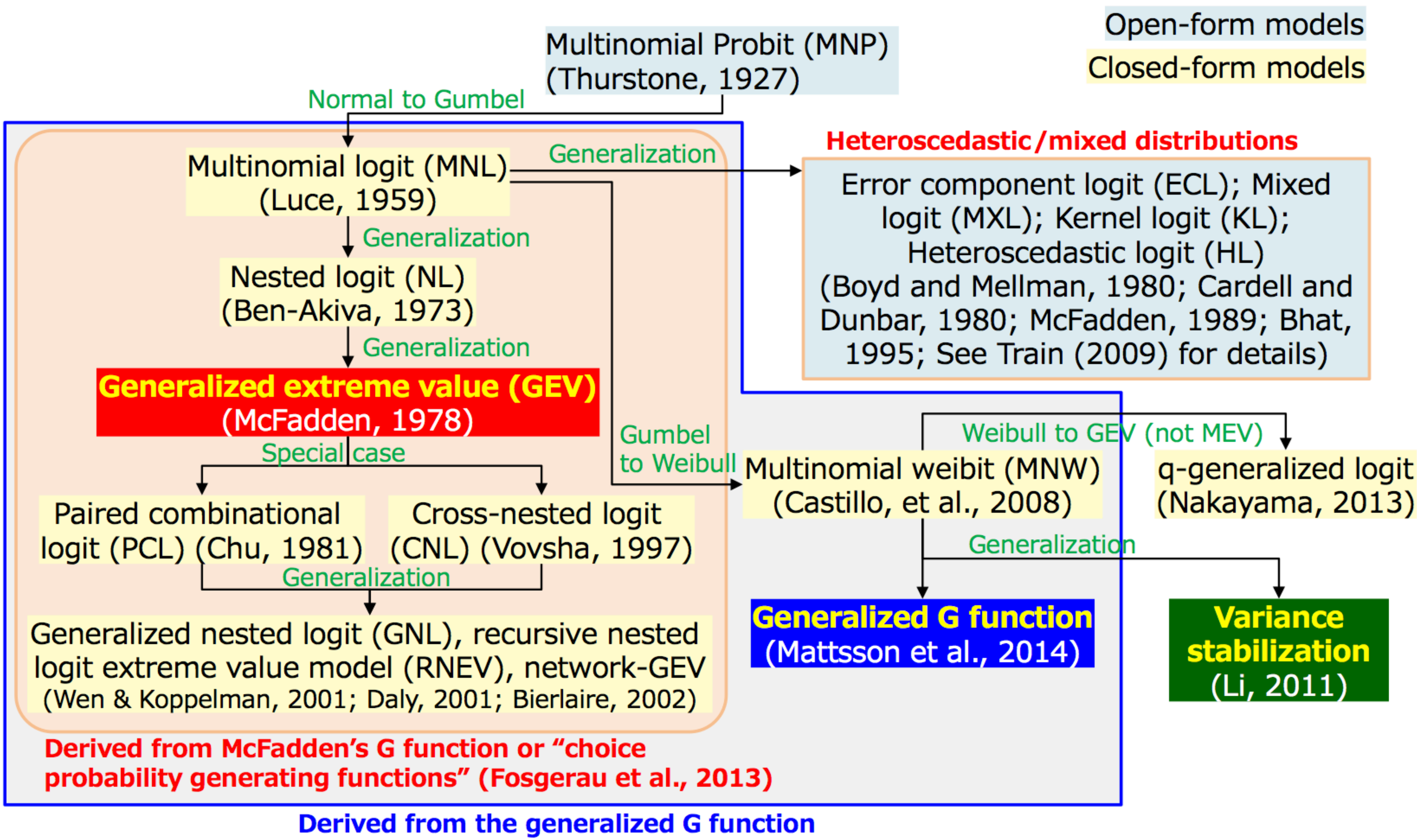
Open-form
discrete choice models

Genealogy of DCMs (again)



Difference of GEV and Non-GEV

GEV model (Closed-form)

Multinomial Logit (MNL)

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

- Luce(1959), McFadden(1974)
- Not consider correlation of choice alternatives' (IIA)
- Easy and fast estimation
- High operability
  (easy evaluation for new additional choice alternative ⇒ benefit of IIA)

Non-GEV model (Open-form)

Multinomial Probit (MNP)

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

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

- Thurstone(1927)
- Consider correlation of choice alternates' based on Variance-Covariance matrix
- Hard and slow estimation
  (need calculation of multi-dimensional interrelation depend on N of alternatives')

Non-GEV model has high power of expression, however parameter estimation cost is high.

Structured Covariance MNP (1)

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

Prof. Morichi

Prof. Yai

Prof. Iwakura

- Proposed new probit type railway route choice model considering overlapping problem. (1993, 1998)
- This model applied to practical demand forecasting in real Tokyo network, and it used for decision making of railway policy toward 2015. (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

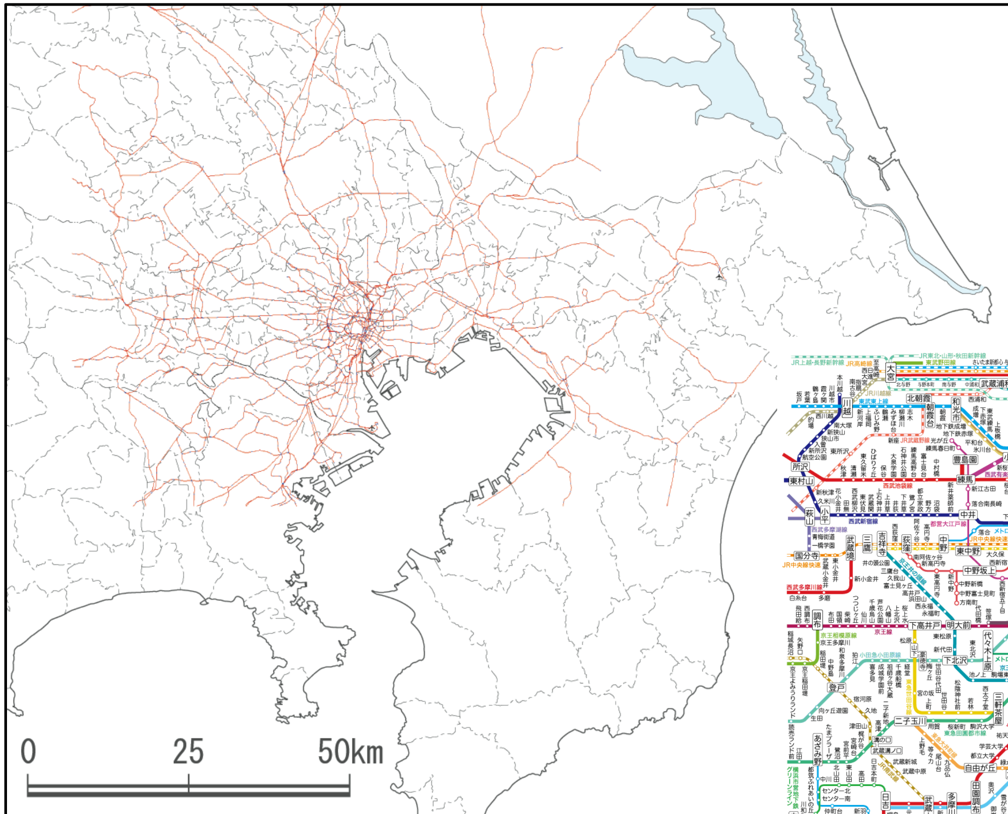
The applications of the multinomial probit model have not been adequately successful in spite of its advantages in flexibility of the model form. Certainly, the complexity of the computational process has deterred its use, compared to the wide applications of the multinomial logit models. Early advances 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.

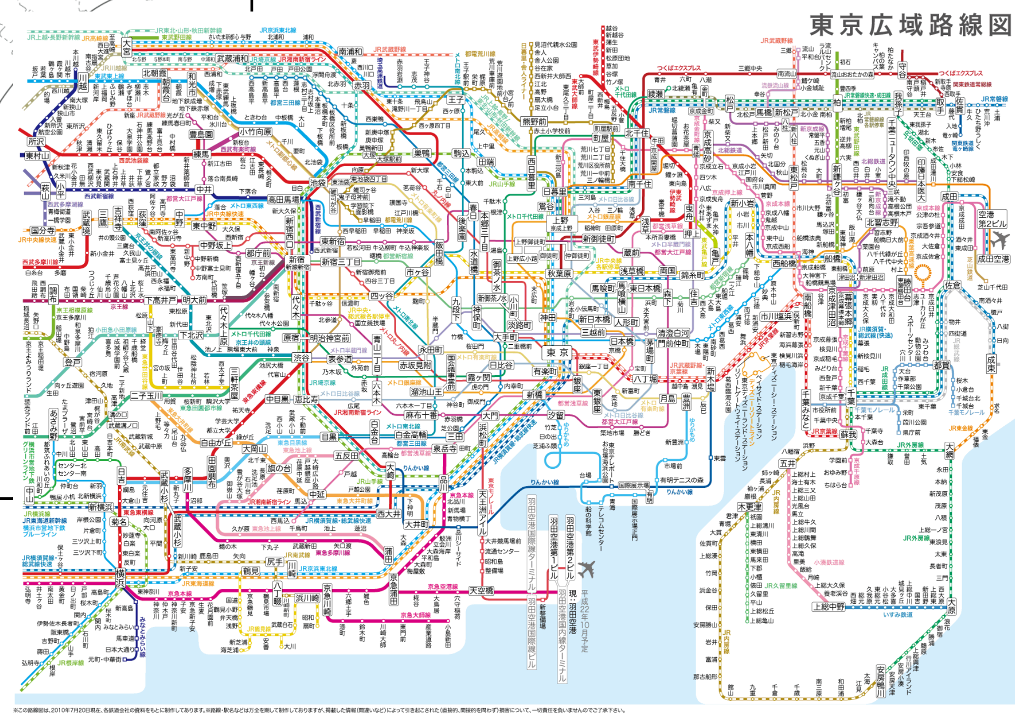
Recently, there have been advances in multinomial probit estimation (McFadden, 1989; Pakes and Pollard, 1989; Bunch, 1991; Bolduc and Ben-Akiva, 1991; Bolduc, 1992; Geweke *et al.*, 1994). The method of simulated moments proposed by McFadden seems to encourage multinomial probit applications because 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

Structured Covariance MNP (2)

Tokyo Metropolitan has highly dense railway network！
⇒ route overlapping problem

Railway line：about 130
Station：about 1800
Passengers：40miliion/day
Cong. rate: max over 200%

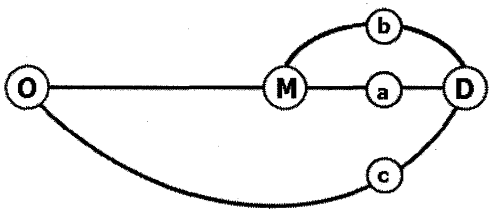




Structured Covariance MNP (3)

In the overlap network that has correlation between routes, Logit model is susceptible to error by IIA property.

Overlap ＝correlation



Probit is better ?

- Difficult to setting covariance matrix for each OD pair
  ⇒ structured covariance by divide into two error
- Difficult to parameter estimation. (multi-dimensional Integral)
  ⇒ reduce computational time using simulation methods

$$U_i = V_i + \varepsilon_i$$

Error of depend on route length

$$\varepsilon_i = \varepsilon_i^{Length} + \varepsilon_i^{Route}$$

Error of route specific

Structured Covariance MNP (4)

Variance-Covariance structure in Error term

Error of depend on route length

Error of route specific

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

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

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

Simplify use cov. ratio

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

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

Error of depend on route length

Variance of route utility increases in proportion to the route length.

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

Covariance between routes increases in proportion to the length of route overlap.

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

Error of route specific

・independent of each route (cov=0)

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

Estimate only cov. ratio !

L_r :length of route r

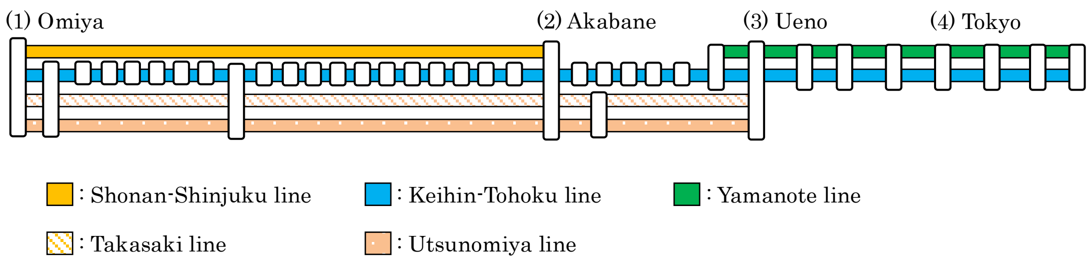
L_{rq} :overlap length between route r and q

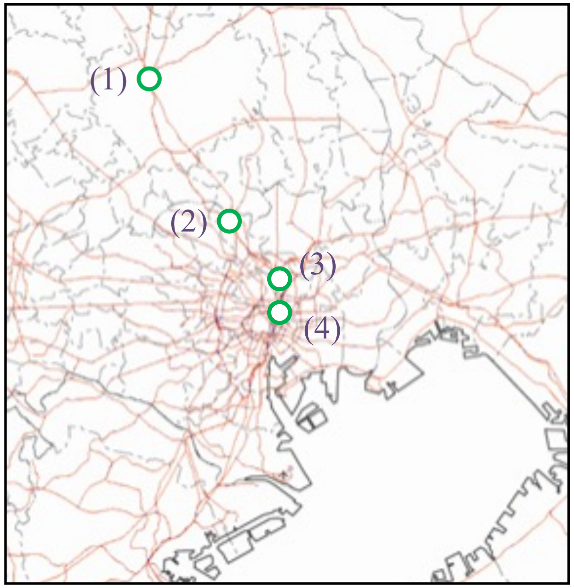
σ^2 :variance of unit length

Structured Covariance MNP (5)

Apply to the SCMNL for The 18th master plan for urban railway network in TMA (2000)

Ex：Oomiya to Kanda station





Estimation results

	parameter	t-value
in-vehicle time	−0.0943	−8.09
access/egress time	−0.127	−11.7
transfar time	−0.112	−10.7
cost	−0.002	−3.98
congestion index	−0.00869	−3.34
η	0.436	2.71
Adj−ρ 2	0.39	
# of sample	1218	

Prediction results

	Obs		MNL		SCMNP	
Utsunomiya + Yamanote	33%	48%	28%	52%	27%	47%
Utsunomiya + Keihin-Tohoku	15%		24%		20%	
Keihin-Tohoku	53%		47%		52%	

To achieve a high prediction accuracy by the relaxation of route overlap (Obs ±10% in all route)

Mixed Logit Model

Mixed Loigt (Train 2000)

High flexible structure using two error term.

Utility function

$$U_i = V_i + \boxed{\eta_i} + \boxed{\nu_i}$$

ν dist.: assume any G function

・IID Gamble (Logit Kernel) ⇒ MNL
・any G function (GEV Kernel) ⇒ NL, PCL, CNL, GNL…

η dist.: basically assume "*Normal dist.*"

In the case of normal distribution takes a non-realistic value, it can assume a variety of probability distribution (triangular distribution, cutting normal distribution, lognormal distribution, Rayleigh distribution, etc.).

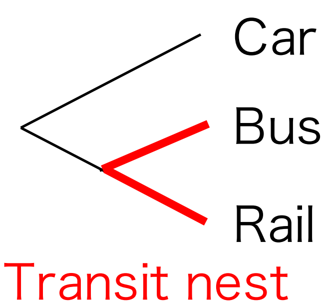
- Error Component: approximate to any GEV model
- Random Coefficient: Consider the heterogeneity

Error Component: NL (1)

Approximation of Nested Logit (NL)

Describe the nest (covariance) using structured $\boldsymbol{\eta}$.

Ex: model choice



Normal ⇒ nest

$$U_{car} = \beta \mathbf{X}_{car} + \qquad + \nu_{car}$$
$$U_{bus} = \beta \mathbf{X}_{bus} + \sigma_{transit}\eta_{transit} + \nu_{bus}$$
$$U_{rail} = \beta \mathbf{X}_{rail} + \sigma_{transit}\eta_{transit} + \nu_{rail}$$

IID Gamble ⇒ Logit

Choice prob. (open-form)

$$P_{rail} = \int_{\eta_{transit}} \frac{e^{V_{rail}+\sigma_{transit}\eta_{transit}}}{e^{V_{car}} + e^{V_{bus}+\sigma_{transit}\eta_{transit}} + e^{V_{rail}+\sigma_{transit}\eta_{transit}}} f\left(\eta_{transit}\right) d\eta_{transit}$$

$$\eta_{transit} \approx N(0,1)$$

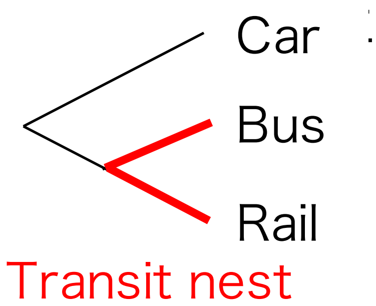
Choice prob. (Simulated)

$$P_{rail} = \frac{1}{N}\sum_{N} \frac{e^{V_{rail}+\sigma_{transit}\eta_{tranist}^{N}}}{e^{V_{car}} + e^{V_{bus}+\sigma_{transit}\eta_{tranist}^{N}} + e^{V_{rail}+\sigma_{transit}\eta_{tranist}^{N}}}$$

Error Component: NL (2)

Approximation of Nested Logit (NL)

Note that variance-covariance matrix is inconsistent with normal NL



Normal NL

$$\begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \underline{\sigma^2} & \sigma^2_{transit} \\ 0 & \sigma^2_{transit} & \underline{\sigma^2} \end{pmatrix}$$

Diagonal elements (variance of Bus and Rail) is bigger than σ_{transit}

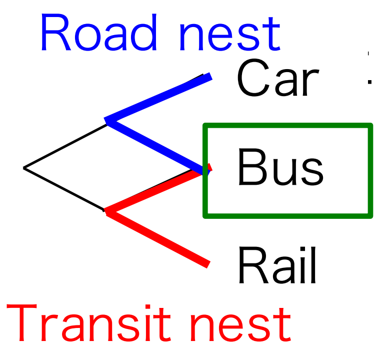
Approximated NL based on MXL

$$\overset{\eta}{\begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma^2_{transit} & \sigma^2_{transit} \\ 0 & \sigma^2_{transit} & \sigma^2_{transit} \end{pmatrix}} + \overset{\nu}{\begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}} = \overset{\varepsilon}{\begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \underline{\sigma^2_{transit} + \sigma^2} & \sigma^2_{transit} \\ 0 & \sigma^2_{transit} & \underline{\sigma^2_{transit} + \sigma^2} \end{pmatrix}}$$

Error Component: CNL

Approximation of Cross Nested Logit (CNL)

Describe the nest (covariance) using structured $\boldsymbol{\eta}$.



$$
\begin{aligned}
U_{car} &= \beta X_{car} & & + \sigma_{road}\eta_{road} + \nu_{car} \\
U_{bus} &= \beta X_{bus} & + \sigma_{transit}\eta_{transit} & + \sigma_{road}\eta_{road} + \nu_{bus} \\
U_{rail} &= \beta X_{rail} & + \sigma_{transit}\eta_{transit} & \qquad\qquad\quad + \nu_{rail}
\end{aligned}
$$

$$\eta_{transit}, \eta_{road} \approx N(0,1)$$

$$
\underbrace{\begin{pmatrix} \sigma^2_{road} & \sigma^2_{road} & 0 \\ \sigma^2_{road} & \sigma^2_{road} + \sigma^2_{transit} & \sigma^2_{transit} \\ 0 & \sigma^2_{transit} & \sigma^2_{transit} \end{pmatrix}}_{\boldsymbol{\eta}}
+
\underbrace{\begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}}_{\boldsymbol{\nu}}
$$

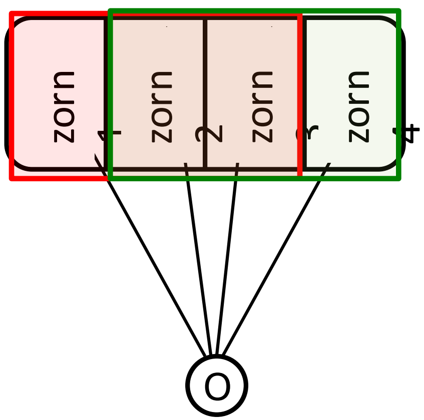
$$\boldsymbol{\varepsilon}$$

$$
= \begin{pmatrix} \sigma^2_{road} + \sigma^2 & \sigma^2_{road} & 0 \\ \sigma^2_{road} & \sigma^2_{road} + \sigma^2_{transit} + \sigma^2 & \sigma^2_{transit} \\ 0 & \sigma^2_{transit} & \sigma^2_{transit} + \sigma^2 \end{pmatrix}
$$

Error Component: SCL

Approximation of Spatial Correlation Logit

Describe the spatial correlation using structured $\boldsymbol{\eta}$.



$$
\begin{aligned}
U_{zorn1} &= \beta \mathbf{X}_{zorn1} + \sigma\,\eta_1 + \sigma\,\eta_2 + \nu_{zorn1} \\
U_{zorn2} &= \beta \mathbf{X}_{zorn2} + \boxed{\sigma\,\eta_1 + \sigma\,\eta_2 + \sigma\,\eta_3} + \nu_{zorn2} \\
U_{zorn3} &= \beta \mathbf{X}_{zorn3} + \boxed{\sigma\,\eta_2 + \sigma\,\eta_3 + \sigma\,\eta_4} + \nu_{zorn3} \\
U_{zorn4} &= \beta \mathbf{X}_{zorn4} + \sigma\,\eta_3 + \sigma\,\eta_4 + \nu_{zorn4}
\end{aligned}
$$

$$
\begin{array}{ccccc}
\boldsymbol{\eta} & & \boldsymbol{\nu} & & \boldsymbol{\varepsilon} \\
\begin{pmatrix}
2\sigma_0^2 & \sigma_0^2 & & 0 \\
\sigma_0^2 & 3\sigma_0^2 & \sigma_0^2 & 0 \\
0 & \sigma_0^2 & 3\sigma_0^2 & \sigma_0^2 \\
0 & 0 & \sigma_0^2 & 2\sigma_0^2
\end{pmatrix}
& + &
\begin{pmatrix}
\sigma^2 & 0 & 0 & 0 \\
0 & \sigma^2 & 0 & 0 \\
0 & 0 & \sigma^2 & 0 \\
0 & 0 & 0 & \sigma^2
\end{pmatrix}
& = &
\begin{pmatrix}
\sigma_0^2 + \sigma^2 & \sigma_0^2 & 0 & 0 \\
\sigma_0^2 & 2\sigma_0^2 + \sigma^2 & \sigma_0^2 & 0 \\
0 & \sigma_0^2 & 2\sigma_0^2 + \sigma^2 & \sigma_0^2 \\
0 & 0 & \sigma_0^2 & \sigma_0^2 + \sigma^2
\end{pmatrix}
\end{array}
$$

$$\eta_1, \eta_2, \eta_3 \approx N(0,1)$$

Error Component: HL

Approximation of heteroscedastic Logit

Assume the different error variance in each alternatives'

※Identification problem occur in the case of not fixed one of the parameters to zero at least.

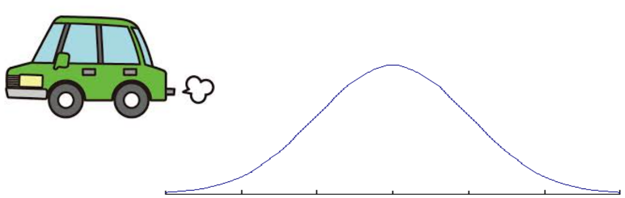
$$U_{car} = \beta X_{car} + \sigma_{car}\eta_{car} + \nu_{car}$$
$$U_{bus} = \beta X_{bus} + \sigma_{bus}\eta_{bus} + \nu_{bus}$$
$$U_{rail} = \beta X_{rail} + \sigma_{rail}\eta_{rail} + \nu_{rail}$$
$$\eta_{car}, \eta_{bus}, \eta_{rail} \approx N(0,1)$$

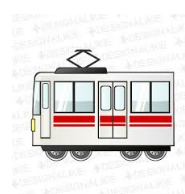
ε

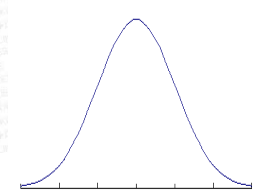
$$\begin{pmatrix} \sigma_{car}^2 + \sigma^2 & 0 & 0 \\ 0 & \boxed{\sigma_{bus}^2} + \sigma^2 & 0 \\ 0 & 0 & \sigma_{rail}^2 + \sigma^2 \end{pmatrix}$$

FIX ! $\sigma_{bus}^2 = 1$

Assume heteroscedastic in error







・Car: Low travel time reliability
⇒Error variance is large

・Rail: High travel time reliability
⇒Error variance is small

※consider only heteroscedasitc（IID assumption is not relaxed)

Random Coefficient (1)

Taste heterogeneity of decision maker

Parameters defined homogeneously in population. However, decision maker ***n*** has different taste (= heterogeneity)

$$U_{car,n} = \beta T_{car,n} + \varepsilon_{car,n} \rightarrow U_{car,n} = \beta_n T_{car,n} + \varepsilon_{car,n}$$

<u>Segmentation (observable heterogeneity)</u>

・Constant by gender：

male's constant: $\alpha_0 + \alpha_1$

$$U_{car,n} = \left[(\alpha_0) + \alpha_1 * male_n\right] + \beta_1 T_{car,n} + \varepsilon_{car,n}$$

Female's constant; α_0

・parameter by gender：

male's parameter: β_1

$$U_{car,n} = \alpha_0 + \{\beta_1\} * male_n * T_{car,n} + \{\beta_2\} * (1 - male_n) * T_{car,n}$$

Female's parameter: β_2

※1 - $male_n$ = $female_n$

Random Coefficient (2)

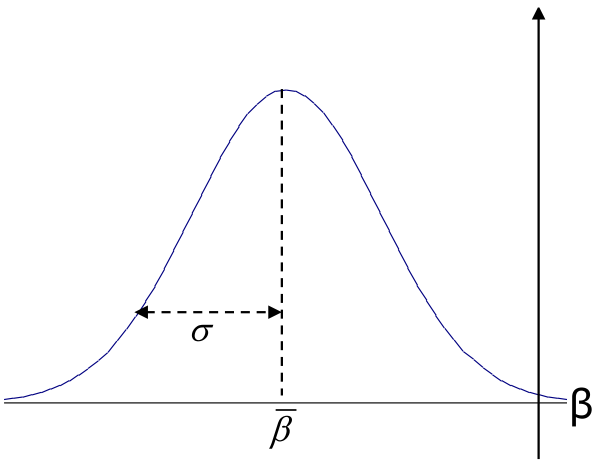
Parameter distribution (unobservable heterogeneity)

Assume the heterogeneity of parameter

⇒In the case of parameter following Normal dist., we estimate the dist.'s hyper-parameter (mean and variance).

$$U_{car,n} = \boxed{\beta_n} T_{car,n} + \nu_{car,n}$$

$$\beta_n \approx N\left(\bar{\beta}, \sigma^2\right)$$



$$U_{car,n} = \bar{\beta} T_{car,n} + \sigma \eta_n T_{car,n}$$

$$U_{bus,n} = \bar{\beta} T_{bus,n} + \sigma \eta_n T_{bus,n}$$

$$U_{rail,n} = \bar{\beta} T_{rail,n} + \sigma \eta_n T_{rail,n}$$

$\eta_n \approx N(0,1)$ $\quad \bar{\beta}, \sigma : unknown\ parameter$

Hyper-parameter can describe using observable variables

$$\bar{\beta}_n = \gamma_0 + \gamma_1 income_n$$

β depend on observable income variable

Summary of open-form models

Strengths

❖ Describe correlation between alternatives' by EC
- MNP: all alternatives' (relax and reduce by structuring)
- MXL: depend on approximated model

❖ Describe heterogeneity by RC
- Segmentation, parameter distribution…

Limitations

❖ High calculation cost in parameter estimation
- Open-form model has high dimensional integration.
- Recently, proposed high speed estimation methods

Ex: Bayesian estimation (MCMC) ⇒ see Train's book

MACML: analytical integration by Bhat et al.(2011)

Reference

- Yai, T., Iwakura, S., & Morichi, S.: Multinomial probit with structured covariance for route choice behavior. *Transportation Research Part B: Methodological*, *31*(3), pp.195-207, 1997.
- Morichi, S., Iwakura, S., Morishige, S., Itoh, M., & Hayasaki, S.: Tokyo metropolitan rail network long-range plan for the 21st century. *Transportation Research Board*, (01-0475), 2001.
- Train, K. E.: *Discrete choice methods with simulation*. Cambridge university press, 2009.
- Revelt, D., & Train, K.:.Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics*, *80*(4), pp.647-657, 1998.
- Ben-Akiva, M., Bolduc, D., & Walker, J.: Specification, identification and estimation of the logit kernel (or continuous mixed logit) model. *Department of Civil Engineering Manuscript, MIT, 2001*.
- Walker, J. L., Ben-Akiva, M., & Bolduc, D.: Identification of parameters in normal error component logit-mixture (NECLM) models. *Journal of Applied Econometrics*, *22*(6), pp.1095-1125, 2007.