Multivariate Extreme Value models

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

Outline

1. Introduction
2. Multivariate Extreme Value distribution
3. MEV model
4. Examples of MEV models
5. Cross nested logit model
6. Network MEV model

Logit

Assumptions

- Random utility:

$$U_{in} = V_{in} + \varepsilon_{in}$$

- ε_{in} is i.i.d. EV (Extreme Value) distributed
- ε_{in} is the maximum of many r.v. capturing unobservable attributes, measurement and specification errors.

i.i.d.

- independent and
- identically
- distributed.

Relax the independence assumption

Multivariate distribution

$$\begin{pmatrix} U_{1n} \\ \vdots \\ U_{Jn} \end{pmatrix} = \begin{pmatrix} V_{1n} \\ \vdots \\ V_{Jn} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1n} \\ \vdots \\ \varepsilon_{Jn} \end{pmatrix}$$

that is

$$U_n = V_n + \varepsilon_n$$

and ε_n is a vector of random variables.

Outline

1 Introduction

2 Multivariate Extreme Value distribution

3 MEV model

4 Examples of MEV models

5 Cross nested logit model

6 Network MEV model

Multivariate Extreme Value distribution

Definition

$$\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{Jn})$$

follows a multivariate extreme value distribution if it has the CDF

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{Jn}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{Jn}})},$$

where $G : \mathbb{R}_+^{J_n} \to \mathbb{R}_+$ is a positive function with positive arguments.

Valid CDF must verify three properties

- $F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, -\infty, \ldots, \varepsilon_{J_n n}) = 0.$
- $F_{\varepsilon_n}(+\infty, \ldots, +\infty) = 1.$
- For any set of $\widehat{J}_n \leq J_n$ distinct indices $i_1, \ldots, i_{\widehat{J}_n}$,

$$\frac{\partial^{\widehat{J}_n} F_{\varepsilon_n}}{\partial \varepsilon_{i_1 n} \cdots \partial \varepsilon_{i_{\widehat{J}_n} n}}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) \geq 0.$$

The limit property

Valid CDF

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, -\infty, \ldots, \varepsilon_{J_n n}) = 0.$$

MEV

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})}.$$

Valid G function

$$G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty.$$

The zero property

Valid CDF

$$F_{\varepsilon_n}(+\infty, \ldots, +\infty) = 1.$$

MEV

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})}.$$

Valid G function

$$G(0, \ldots, 0) = 0.$$

The strong alternating sign property

Valid CDF

$$\frac{\partial^{\widehat{J}_n} F_{\varepsilon_n}}{\partial \varepsilon_{i_1 n} \cdots \partial \varepsilon_{i_{\widehat{J}_n} n}}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) \geq 0.$$

MEV

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})}.$$

Valid G function (notation: $G_i = \partial G / \partial y_i$)

- The right-hand side changes sign each time it is differentiated.
- To obtain ≥ 0, G must also change sign each time it is differentiated.
- For any set of $\widehat{J}_n$ distinct indices $i_1, \ldots, i_{\widehat{J}_n}$,

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geq 0.$$

Homogeneity

We need another property: homogeneity

A function G is homogeneous of degree μ, or μ-homogeneous, if

$$G(\alpha y) = \alpha^{\mu} G(y), \; \forall \alpha > 0 \text{ and } y \in \mathbb{R}_{+}^{J_n}.$$

It will imply two results

- the marginals are univariate extreme value distributions,
- the choice model has a closed form.

Marginal distribution

ith marginal distribution

$$F_{\varepsilon_n}(+\infty, \ldots, +\infty, \varepsilon_{in}, +\infty, \ldots, +\infty) = e^{-G(0,\ldots,0,e^{-\varepsilon_{in}},0,\ldots,0)}.$$

If G is μ-homogeneous, we have

$$G(0, \ldots, 0, e^{-\varepsilon_{in}}, 0, \ldots, 0) = e^{-\mu\varepsilon_{in}} G(0, \ldots, 0, 1, 0, \ldots, 0),$$

or equivalently,

$$G(0, \ldots, 0, e^{-\varepsilon_{in}}, 0, \ldots, 0) = e^{-\mu\varepsilon_{in} + \log G(0,\ldots,0,1,0,\ldots,0)},$$

Define $\log G(0, \ldots, 0, 1, 0, \ldots, 0) = \mu\eta$, so that

$$F_{\varepsilon_n}(+\infty, \ldots, +\infty, \varepsilon_{in}, +\infty, \ldots, +\infty) = \exp\left(-e^{-\mu(\varepsilon_{in}-\eta)}\right).$$

Multivariate Extreme Value distribution

CDF

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})},$$

ith marginal: univariate extreme value distribution

$$F_{\varepsilon_n}(+\infty, \ldots, +\infty, \varepsilon_{in}, +\infty, \ldots, +\infty) = \exp\left(-e^{-\mu(\varepsilon_{in}-\eta)}\right).$$

Multivariate Extreme Value distribution

Three conditions on G

- The limit property

$$G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty.$$

- The strong alternating sign property

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geq 0.$$

- Homogeneity (which implies the zero property)

$$G(\alpha y) = \alpha^{\mu} G(y), \ \forall \alpha > 0 \text{ and } y \in \mathbb{R}_+^{J_n}.$$

Outline

1. Introduction
2. Multivariate Extreme Value distribution
3. MEV model
4. Examples of MEV models
5. Cross nested logit model
6. Network MEV model

MEV: the choice model

CDF of the error terms

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})},$$

Choice model: $P_n(i) =$

$$\int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n}, \varepsilon_{2n}, \ldots, \varepsilon_{J_n n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon.$$

$$\frac{\partial F_{\varepsilon_{1n}, \varepsilon_{2n}, \ldots, \varepsilon_{J_n n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots)$$

$$= e^{-\varepsilon} G_i(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots)$$
$$\exp\left(-G(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots)\right)$$

MEV: the choice model

G is μ-homogeneous

so that $G_i = \partial G / \partial y_i$ is $(\mu - 1)$-homogeneous.

$$
\begin{aligned}
& e^{-\varepsilon} G_i(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots) \\
& \exp\left(-G(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots)\right) \\
= {} & e^{-\varepsilon} e^{-(\mu-1)\varepsilon} e^{-(\mu-1)V_{in}} G_i(\ldots, e^{V_{(i-1)n}}, e^{V_i n}, e^{V_{(i+1)n}}, \ldots) \\
& \quad \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(\ldots, e^{V_{(i-1)n}}, e^{V_{in}}, e^{V_{(i+1)n}}, \ldots)\right) \ .
\end{aligned}
$$

MEV: choice model

We now denote

$$e^V = \left(\ldots, e^{V_{(i-1)n}}, e^{V_{in}}, e^{V_{(i+1)n}}, \ldots\right),$$

and simplify the terms to obtain

$$\frac{\partial F_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots)$$

$$= e^{-\mu\varepsilon} e^{-\mu V_{in}} e^{V_{in}} G_i(e^V) \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(e^V)\right).$$

Therefore

$$P_n(i) = e^{-\mu V_{in}} e^{V_{in}} G_i(e^V) \int_{\varepsilon=-\infty}^{+\infty} e^{-\mu\varepsilon} \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(e^V)\right) d\varepsilon.$$

MEV: choice model

Choice probability

$$P_n(i) = e^{-\mu V_{in}} e^{V_{in}} G_i(e^V) \int_{\varepsilon=-\infty}^{+\infty} e^{-\mu\varepsilon} \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(e^V)\right) d\varepsilon.$$

Define $t = -\exp(-\mu\varepsilon)$, so that $dt = \mu \exp(-\mu\varepsilon) d\varepsilon$:

$$P_n(i) = e^{-\mu V_{in}} e^{V_{in}} G_i(e^V) \frac{1}{\mu} \int_{t=-\infty}^{0} \exp\left(t e^{-\mu V_{in}} G(e^V)\right) dt,$$

which simplifies to

$$P_n(i) = \frac{e^{V_{in}} G_i(e^V)}{\mu G(e^V)}.$$

MEV: choice model

Choice probability

$$P_n(i) = \frac{e^{V_{in}} G_i(e^V)}{\mu G(e^V)}.$$

From Euler's theorem:

$$P_n(i) = \frac{e^{V_{in}} G_i\left(e^V\right)}{\sum_j e^{V_{jn}} G_j\left(e^V\right)}.$$

Logit-like form:

$$P_n(i) = \frac{e^{V_{in} + \log G_i\left(e^V\right)}}{\sum_j e^{V_{jn} + \log G_j\left(e^V\right)}}.$$

MEV: choice model

The multivariate extreme value model:

$$P_n(i) = \frac{e^{V_{in} + \log G_i\left(e^V\right)}}{\sum_j e^{V_{jn} + \log G_j\left(e^V\right)}}.$$

where $G_i = \partial G / \partial y_i$, and G verifies

- (i) the limit property: $G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty$.
- (ii) the strong alternating sign property: for any set of $\widehat{J}_n$ distinct indices $i_1, \ldots, i_{\widehat{J}_n}$,

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geq 0.$$

- (iii) the homogeneity property:

$$G(\alpha y) = \alpha^{\mu} G(y), \ \forall \alpha > 0 \text{ and } y \in \mathbb{R}_+^{J_n}.$$

MEV: choice model

Probability generating function

A function G, which is μ homogeneous, that verifies the MEV properties is called a μ-MEV function.

Expected maximum utility

$$E[\max_{j \in \mathcal{C}_n} U_{jn}] = \frac{1}{\mu}(\log G(e^{V_{1n}}, \ldots, e^{V_{J_n n}}) + \gamma),$$

where γ is Euler's constant

Euler's constant

$$\gamma = -\int_0^{+\infty} e^{-x} \ln x \, dx \approx 0.5772.$$

MEV vs GEV

McFadden (1978)

MEV is called "General Extreme Value model" (GEV)

Jenkinson (1955)

a Generalized Extreme Value distribution (Jenkinson, 1955) is a univariate distribution with CDF

$$F_X(x) = \begin{cases} e^{-(1+\xi((x-\mu)/\sigma))^{-1/\xi})} & -\infty < x \leq \mu - \sigma/\xi \quad \text{for } \xi < 0 \\ & \mu - \sigma/\xi \leq x < \infty \quad \text{for } \xi > 0 \\ e^{-e^{-(x-\mu)/\sigma}} & -\infty < x < \infty \quad \text{for } \xi = 0 \end{cases}$$

$\xi = 0$ Type 1 EV distribution
$\xi > 0$ Type 2 EV distribution
$\xi < 0$ Type 3 EV distribution

Distribution of the utility functions

$$U_n = (U_{1n}, \ldots, U_{J_n n}) = (V_{1n} + \varepsilon_{1n}, \ldots, V_{Jn} + \varepsilon_{Jn})$$

CDF

$$F_{U_n}(\xi_1, \ldots, \xi_{J_n}) = \Pr(U_n \leq \xi_n) = e^{-G\left(e^{V_{1n}-\xi_1}, \ldots, e^{V_{J_n n}-\xi_{J_n}}\right)}.$$

Marginal distributions: extreme value

- Mean: $V_{jn} + \frac{\log G(0,\ldots,1,\ldots,0)+\gamma}{\mu}$
- Variance: $\pi^2/6\mu^2$, for each j

Variance-covariance matrix

$$\begin{aligned}\mathrm{Cov}(\varepsilon_{in},\varepsilon_{jn}) &= \mathrm{E}[\varepsilon_{in}\varepsilon_{jn}] - \mathrm{E}[\varepsilon_{in}]\,\mathrm{E}[\varepsilon_{jn}] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi_i\xi_j \frac{\partial^2 F_{\varepsilon_n}(\xi_i,\xi_j)}{\partial\xi_i\partial\xi_j} d\xi_i d\xi_j - \gamma^2,\end{aligned}$$

where $\mathrm{E}[\varepsilon_{in}] = \gamma$, $F_{\varepsilon_n}(\xi_i,\xi_j) = F_{\varepsilon_n}(\ldots,+\infty,\xi_i,+\infty,\ldots,+\infty,\xi_j,+\infty,\ldots)$ is the bivariate marginal cumulative distribution, and

$$\frac{\partial^2 F_{\varepsilon_{in},\varepsilon_{jn}}(\xi_i,\xi_j)}{\partial\xi_i\partial\xi_j} = F_{\varepsilon_{in},\varepsilon_{jn}}(\xi_i,\xi_j) e^{-\xi_i} e^{-\xi_j} (G_i^{ij} G_j^{ij} - G_{ij}^{ij})$$

where

$$G_i^{ij} = \frac{\partial G(\ldots,0,e^{-\xi_i},0,\ldots,0,e^{-\xi_j},0,\ldots)}{\partial y_i}$$

and

$$G_{ij}^{ij} = \frac{\partial^2 G(\ldots,0,e^{-\xi_i},0,\ldots,0,e^{-\xi_j},0,\ldots)}{\partial y_i \partial y_j}.$$

Outline

1. Introduction
2. Multivariate Extreme Value distribution
3. MEV model
4. Examples of MEV models
5. Cross nested logit model
6. Network MEV model

MEV models

Example: $G(y) = \sum_{i=1}^{J} y_i^{\mu}$, $\mu > 0$

1. $$G(\alpha y) = \sum_{i=1}^{J} (\alpha y_i)^{\mu} = \alpha^{\mu} \sum_{i=1}^{J} y_i^{\mu} = \alpha^{\mu} G(y)$$
2. $$\lim_{y_i \rightarrow +\infty} G(y) = +\infty, \; i = 1, \ldots, J$$
3. $$\frac{\partial G}{\partial y_i} = \mu y_i^{\mu-1} \text{ and } \frac{\partial^2 G}{\partial y_i \partial y_j} = 0$$

G complies with the theory

MEV models

Example: $G(y) = \sum_{i=1}^{J} y_i^{\mu}$, $\mu > 0$

$$
\begin{aligned}
F(\varepsilon_1, \ldots, \varepsilon_J) &= e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})} \\
&= e^{-\sum_{i=1}^{J} e^{-\mu \varepsilon_i}} \\
&= \prod_{i=1}^{J} e^{-e^{-\mu \varepsilon_i}}
\end{aligned}
$$

Product of i.i.d EV

Logit Model

MEV models

Example: $G(e^{V_1},\ldots,e^{V_J}) = \sum_{i=1}^{J} e^{\mu V_i}$, $\mu > 0$

$$P(i) = \frac{e^{V_i + \ln G_i(e^{V_1},\ldots,e^{V_J})}}{\sum_{j\in C} e^{V_j + \ln G_j(e^{V_1},\ldots,e^{V_J})}} \text{ with } G_i(x) = \mu x_i^{\mu-1}$$

$$\begin{aligned} e^{V_i + \ln G_i(e^{V_1},\ldots,e^{V_J})} &= e^{V_i + \ln \mu + (\mu-1)\ln e^{V_i}} \\ &= e^{\ln \mu + \mu V_i} \end{aligned}$$

$$P(i) = \frac{e^{\ln \mu + \mu V_i}}{\sum_{j\in C} e^{\ln \mu + \mu V_j}} = \frac{e^{\mu V_i}}{\sum_{j\in C} e^{\mu V_j}}$$

MEV models

Example: $G(e^{V_1},\ldots,e^{V_J}) = \sum_{i=1}^{J} e^{\mu V_i}$, $\mu > 0$

$$\begin{aligned} \mathrm{E}[\max_{j\in\mathcal{C}_n} U_{jn}] &= \frac{1}{\mu}\left(\ln G(e^{V_1},\ldots,e^{V_J}) + \gamma\right) \\ &= \frac{1}{\mu}\ln\sum_{i=1}^{J} e^{\mu V_i} + \frac{\gamma}{\mu} \end{aligned}$$

MEV models

Example: Nested logit

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

with $\mu > 0$, $\mu_m > 0$.

Homogeneity

$$G(\alpha y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} (\alpha y_i)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} = \alpha^{\mu} \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

Limit property

$$\lim_{y_i \to +\infty} G(y) = +\infty, i = 1, \ldots, J$$

MEV models

Example: Nested logit

$$G(y)=\sum_{m=1}^{M}\left(\sum_{i=1}^{J_m} y_i^{\mu_m}\right)^{\frac{\mu}{\mu_m}}$$

with $\mu > 0$, $\mu_m > 0$.

Strong alternating sign property

$$\frac{\partial G}{\partial y_i}=\frac{\mu}{\mu_m}\mu_m y_i^{\mu_m-1}\left(\sum_{i=1}^{J_m} y_i^{\mu_m}\right)^{\frac{\mu}{\mu_m}-1}\geq 0$$

If $\mu \leq \mu_m$, then

$$\frac{\partial^2 G}{\partial y_i y_j}=\mu\mu_m y_i^{\mu_m-1} y_j^{\mu_m-1}(\frac{\mu}{\mu_m}-1)\left(\sum_{i=1}^{J_m} y_i^{\mu_m}\right)^{\frac{\mu}{\mu_m}-2}\leq 0$$

MEV models

So far, we have seen that

- the logit model is a MEV model,
- the nested logit model is also a MEV model:

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

- If $\frac{\mu}{\mu_m} \leq 1$, then G complies with the theory
- Are there other such models?

Outline

1. Introduction
2. Multivariate Extreme Value distribution
3. MEV model
4. Examples of MEV models
5. Cross nested logit model
6. Network MEV model

Cross-Nested logit model

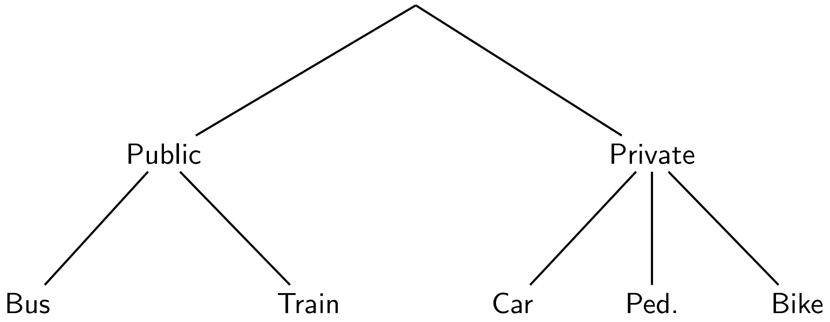
Probability generating function

$$G(y_1,\ldots,y_J) = \sum_{m=1}^{M} \left(\sum_{j} ({\alpha_{jm}}^{1/\mu} y_j)^{\mu_m} \right)^{\frac{\mu}{\mu_m}},$$

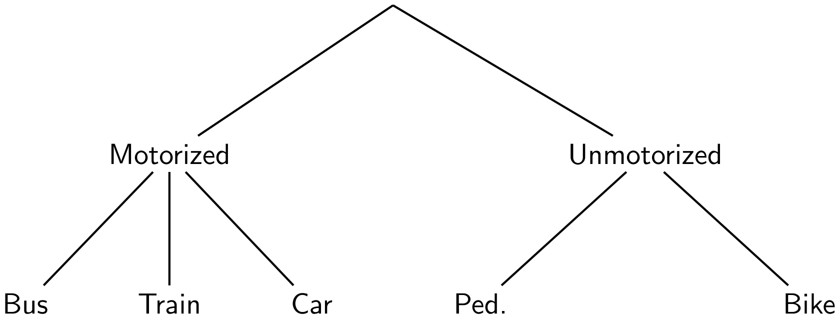
with $\frac{\mu}{\mu_m} \leq 1$, $\alpha_{jm} \geq 0$, and $\forall j, \exists m$ s.t. $\alpha_{jm} > 0$

Generalization of the nested-logit model

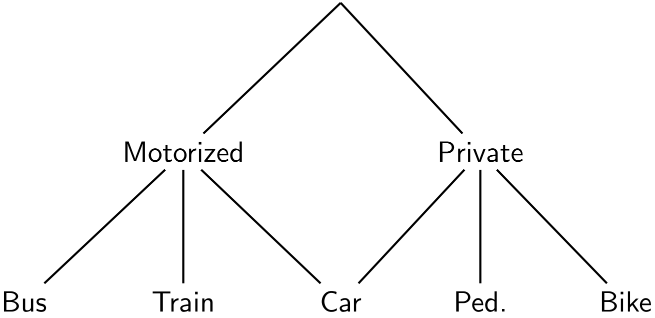
Nested Logit Model



Nested Logit Model



Cross Nested Logit Model



Cross-Nested Logit Model

Choice model

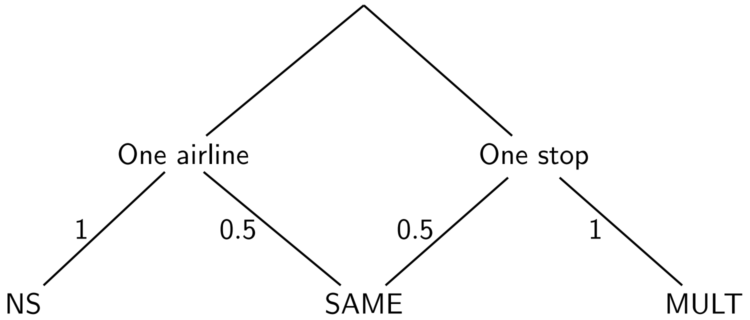
$$P(i|\mathcal{C}) = \sum_{m=1}^{M} \frac{\left(\sum_{j\in\mathcal{C}} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}\right)^{\frac{\mu}{\mu_m}}}{\sum_{n=1}^{M}\left(\sum_{j\in\mathcal{C}} \alpha_{jn}^{\mu_n/\mu} e^{\mu_n V_j}\right)^{\frac{\mu}{\mu_n}}} \frac{\alpha_{im}^{\mu_m/\mu} e^{\mu_m V_i}}{\sum_{j\in\mathcal{C}} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}}.$$

which can nicely be interpreted as

$$P(i|\mathcal{C}) = \sum_{m} P(m|\mathcal{C}) P(i|m).$$

Airline itinerary choice example

Cross-nested logit



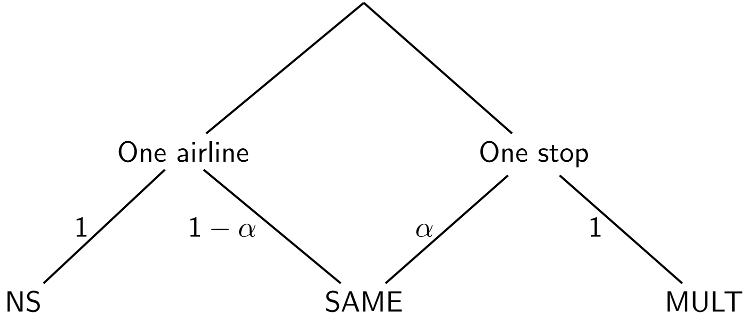
Airline itinerary choice example

Parameter number	Description	Coeff. estimate	Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-0.674	0.185	-3.64	0.00
2	One stop–multiple airlines	-1.10	0.175	-6.29	0.00
3	Round trip fare ($100)	-1.55	0.170	-9.10	0.00
4	Elapsed time (0–2 hours)	-0.783	0.210	-3.72	0.00
5	Elapsed time (2–8 hours)	-0.177	0.0627	-2.82	0.00
6	Elapsed time ($>$ 8 hours)	-0.832	0.274	-3.03	0.00
7	Leg room (inches), if male (non stop)	0.0904	0.0305	2.97	0.00
8	Leg room (inches), if female (non stop)	0.174	0.0302	5.77	0.00
9	Leg room (inches), if male (one stop)	0.0998	0.0227	4.40	0.00
10	Leg room (inches), if female (one stop)	0.0640	0.0200	3.20	0.00
11	Being early (hours)	-0.128	0.0175	-7.30	0.00
12	Being late (hours)	-0.0747	0.0154	-4.86	0.00
13	More than two air trips per year (one stop–same airline)	-0.241	0.120	-2.01	0.04
14	More than two air trips per year (one stop–multiple airlines)	-0.0964	0.132	-0.73	0.47
15	Round trip fare / income ($100/$1000)	-17.9	7.68	-2.34	0.02
16	$\mu_{\text{One airline}}$	1.11	0.122	0.86[1]	0.39
17	$\mu_{\text{One stop}}$	2.38	0.392	3.51[1]	0.00

[1] t-test against 1

Airline itinerary choice example

Cross-nested logit: estimate α



Airline itinerary choice example

Invalid estimation results

- μ parameter for "One airline" $= 0.785$.
- Should be greater or equal to 1.0.
- We reject the model.
- We constrain the μ parameter to 1.0.

Airline itinerary choice example

Parameter number	Description	Coeff. estimate	Asympt. std. error	t-stat	p-value
1	One stop, same airline dummy	-0.703	0.165	-4.27	0.00
2	One stop, multiple airlines	-0.975	0.172	-5.67	0.00
3	Travel time (hours) (0–2 hours)	-0.806	0.214	-3.76	0.00
4	Travel time (hours) (2–8 hours)	-0.182	0.0593	-3.07	0.00
5	Travel time (hours) ($\geq$ 8 hours)	-0.866	0.271	-3.20	0.00
6	Round trip fare ($100) / Income ($1000)	-18.8	7.53	-2.50	0.00
7	Round trip fare ($100)	-1.54	0.150	-10.26	0.00
8	More than two air trips per year (one stop, same airline)	-0.244	0.123	-1.99	0.05
9	More than two air trips per year (one stop, multiple airlines)	-0.109	0.131	-0.83	0.41
10	Leg room (inches), if female (non-stop)	0.179	0.0296	6.06	0.00
11	Leg room (inches), if male (non-stop)	0.0918	0.0309	2.97	0.00
12	Leg room (inches), if female (one-stop)	0.0607	0.0187	3.24	0.00
13	Leg room (inches), if male (one-stop)	0.0952	0.0211	4.52	0.00
14	Being early (hours)	-0.127	0.0157	-8.10	0.00
15	Being late (hours)	-0.0711	0.0141	-5.03	0.00
16	μ One stop	2.19	0.320	3.72[1]	0.00
17	α One stop / One stop, same airline	0.798	0.0889	8.98	0.00

[1] t-test against 1

Airline itinerary choice example

Cross Nested logit

- Number of parameters: 17
- Final log likelihood: -1611.670

Nested logit

- Number of parameters: 16
- Final log likelihood: -1613.858
- Special case of the cross nested: $\alpha = 1$

Testing

- t-test: $\alpha = 1$ is rejected (test=2.27).
- Likelihood ratio: $-2(-1613.858 - (-1611.670)) = 4.32$
- Nested is rejected: $\chi^2_{1,0.05} = 3.84$.

Correlation matrix of the cross nested logit model

Bivariate marginal cumulative distribution

$$F_{\varepsilon_i,\varepsilon_j}(\xi_i,\xi_j) = \exp\left(-\sum_{m=1}^{M}\left((\alpha_{im}^{\frac{1}{\mu}}e^{-\xi_i})^{\mu_m} + (\alpha_{jm}^{\frac{1}{\mu}}e^{-\xi_j})^{\mu_m}\right)^{\frac{1}{\mu_m}}\right).$$

Correlation matrix

$$\Sigma_{\text{CNL}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.695 \\ 0 & 0.695 & 1 \end{pmatrix}$$

Notes

- In this case, block diagonal structure, as the nested logit model.
- But it does not mean it is a nested logit model.
- Contrarily to probit models, MEV models are not characterized by the structure of their correlation matrix.

Outline

1. Introduction
2. Multivariate Extreme Value distribution
3. MEV model
4. Examples of MEV models
5. Cross nested logit model
6. Network MEV model

Inheritance theorem

Context

- Choice set $\mathcal{C}$ with J alternatives.
- M subsets of alternatives $\mathcal{C}_m$, $m = 1, \ldots, M$.
- J_m is the number of alternatives in subset m.
- Let $G^m : \mathbb{R}_+^{J_m} \longrightarrow \mathbb{R}, m = 1, \ldots, M$ be a μ_m-MEV function on $\mathcal{C}_m$, for each m.

Theorem

$$G : \mathbb{R}_+^J \longrightarrow \mathbb{R} : y \rightsquigarrow G(y) = \sum_{m=1}^{M} \left(\alpha_m G^m([y]_m)\right)^{\frac{\mu}{\mu_m}}$$

is a μ-MEV function if $\alpha_m > 0$, $\mu > 0$ and $\mu_m \geq \mu$, $m = 1, \ldots, m$, where $[y]_m$ denotes a vector of dimension J_m with entries y_i, where the indices i correspond to the elements in $\mathcal{C}_m$.

MEV models

Features

- Provide a great deal of flexibility
- Require significant imagination
- Require heavy proofs

Network MEV

Daly & Bierlaire (2006)

- Extension of the tree representation for nested logit
- Investigate new MEV models
- Provide the proof once for all

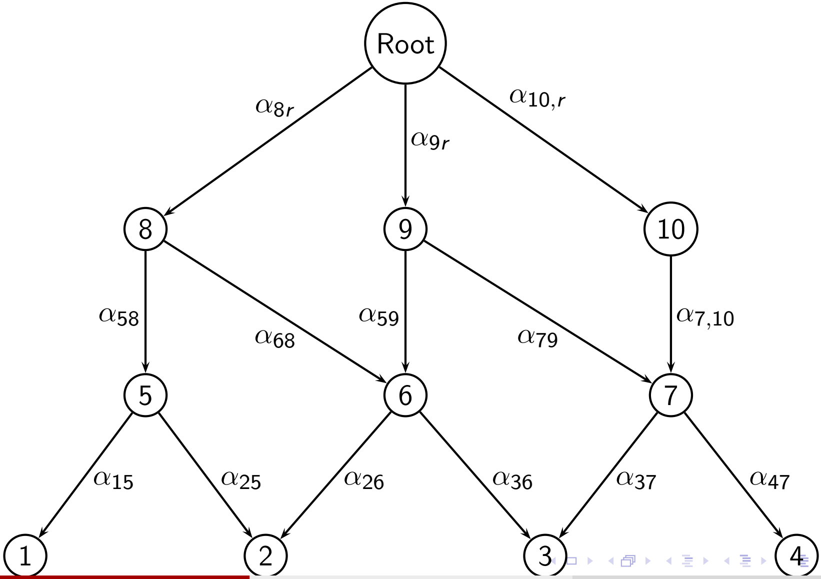
Network MEV

Network

- Consider a network with nodes $i, j, k, \ldots$, and links connecting nodes.
- No circuit.
- One node without predecessor: root.
- J nodes without successor: alternatives.
- All other nodes are called: nests.
- Each nest m is associated with a nest parameter μ_m.
- The parameter associated with the root is μ. It cannot be identified and is normalized to 1.
- Each arc linking node m to node p is associated with a parameter α_{mp}, which captures the level of membership, in a similar way as the α parameters of the cross nested logit model.

Assumptions

Network MEV


Root
α_{8r}
α_{9r}
$\alpha_{10,r}$
8
9
10
α_{58}
α_{68}
α_{59}
α_{79}
$\alpha_{7,10}$
5
6
7
α_{15}
α_{25}
α_{26}
α_{36}
α_{37}
α_{47}
1
2
3
4

Network MEV model

Choice model

- Recursively defined.
- Associate with each node a subset $\mathcal{C}_m$ and a μ_m-MEV function G^m.

Alternative i

- Subset: $\mathcal{C}_i = \{i\}$.
- Normalize $\mu_i = 1$.
- 1-MEV function: $G^i : \mathbb{R} \to \mathbb{R} : G(y) = y$

Network MEV model

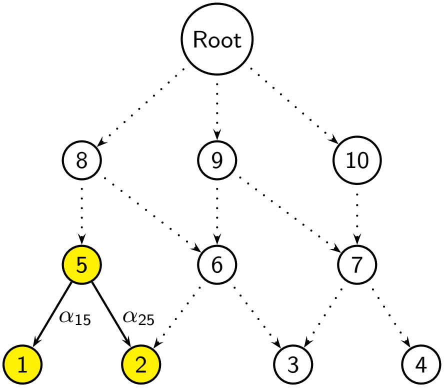
Nest m: list of successors I_m

- Subset: $\mathcal{C}_m = \bigcup_{p \in I_m} \mathcal{C}_p$.
- μ_m-MEV function:

$$G^m : \mathbb{R}^{|\mathcal{C}_m|} \to \mathbb{R} : G^m(y) = \sum_{p \in I_m} \left(\alpha_{pm} G^p(y)\right)^{\frac{\mu_m}{\mu_p}} .$$

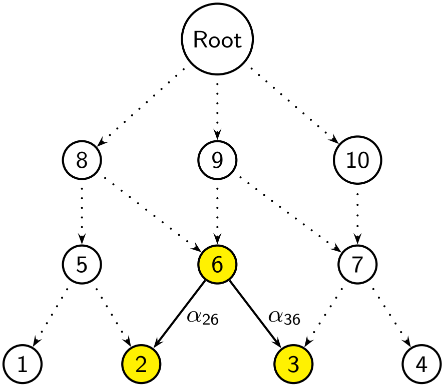
- Validity: inheritance theorem.

Illustrative example



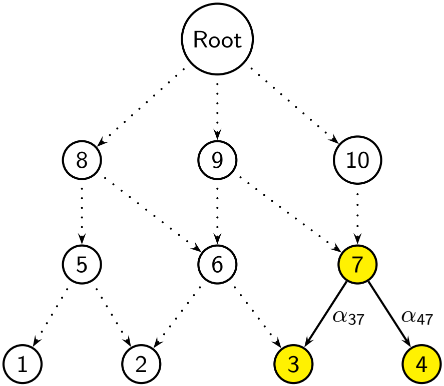
$$G^5(y_1, y_2) = (\alpha_{15} y_1)^{\mu_5} + (\alpha_{25} y_2)^{\mu_5}.$$

Illustrative example



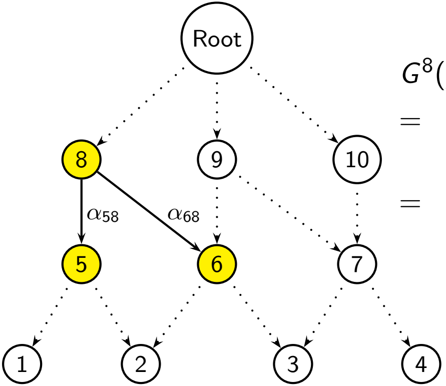
$$G^6(y_2, y_3) = (\alpha_{26} y_2)^{\mu_6} + (\alpha_{36} y_3)^{\mu_6},$$

Illustrative example



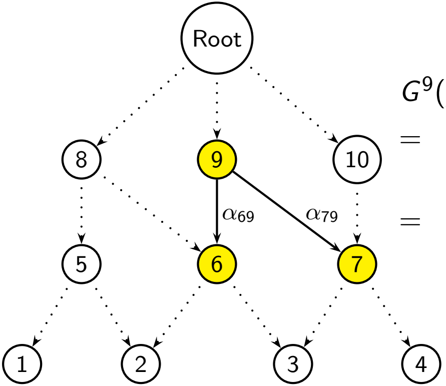
$$G^7(y_3, y_4) = (\alpha_{37} y_3)^{\mu_7} + (\alpha_{47} y_4)^{\mu_7}.$$

Illustrative example



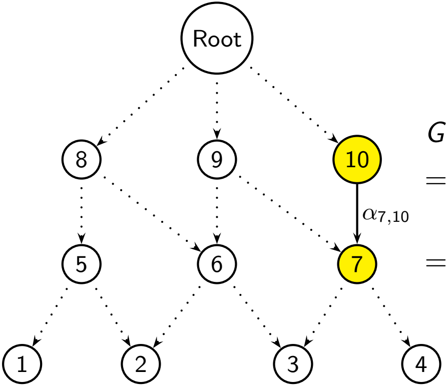
$$G^8(y_1, y_2, y_3)$$
$$= \left(\alpha_{58} G^5(y_1, y_2)\right)^{\frac{\mu_8}{\mu_5}} + \left(\alpha_{68} G^6(y_2, y_3)\right)^{\frac{\mu_8}{\mu_6}}$$
$$= \left(\alpha_{58}((\alpha_{15} y_1)^{\mu_5} + (\alpha_{25} y_2)^{\mu_5})\right)^{\frac{\mu_8}{\mu_5}}$$
$$+ \left(\alpha_{68}((\alpha_{26} y_2)^{\mu_6} + (\alpha_{36} y_3)^{\mu_6})\right)^{\frac{\mu_8}{\mu_6}}.$$

Illustrative example



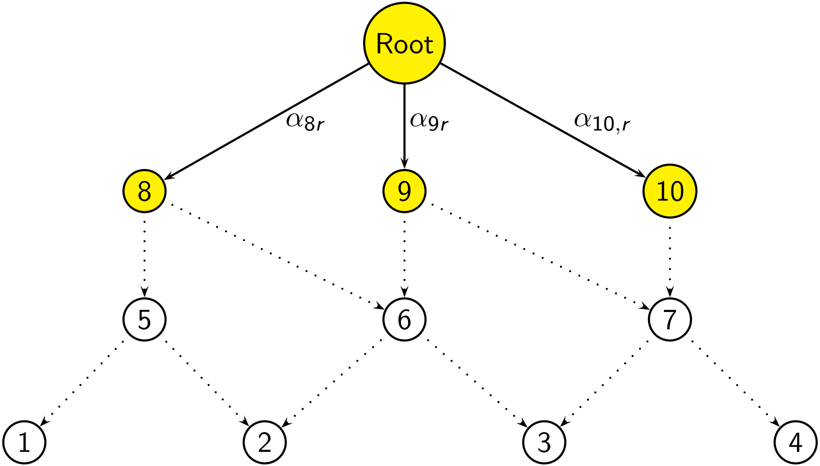
$$G^9(y_2, y_3, y_4)$$
$$= \quad (\alpha_{69} G^6(y_2, y_3))^{\frac{\mu_9}{\mu_6}} + (\alpha_{79} G^7(y_3, y_4))^{\frac{\mu_9}{\mu_7}}$$
$$= \quad (\alpha_{69}((\alpha_{26} y_2)^{\mu_6} + (\alpha_{36} y_3)^{\mu_6}))^{\frac{\mu_9}{\mu_6}}$$
$$+(\alpha_{79}((\alpha_{37} y_3)^{\mu_7} + (\alpha_{47} y_4)^{\mu_7}))^{\frac{\mu_9}{\mu_7}}.$$

Illustrative example



$$
\begin{aligned}
& G^{10}(y_3, y_4) \\
= \quad & \left(\alpha_{7,10} G^7(y_3, y_4)\right)^{\frac{\mu_{10}}{\mu_7}} \\
= \quad & \left(\alpha_{7,10}\left((\alpha_{37} y_3)^{\mu_7} + (\alpha_{47} y_4)^{\mu_7}\right)\right)^{\frac{\mu_{10}}{\mu_7}}.
\end{aligned}
$$

Illustrative example



Illustrative example

Complete model

$$G(y_1,y_2,y_3,y_4) = (\alpha_{8r}G^8(y_1,y_2,y_3))^{\frac{\mu}{\mu_8}} + (\alpha_{9r}G^9(y_2,y_3,y_4))^{\frac{\mu}{\mu_9}} + (\alpha_{10r}G^{10}(y_3,y_4))^{\frac{\mu}{\mu_{10}}},$$

that is

$$G(y_1,y_2,y_3,y_4) = (\alpha_{8r}((\alpha_{58}((\alpha_{15}y_1)^{\mu_5} + (\alpha_{25}y_2)^{\mu_5}))^{\frac{\mu_8}{\mu_5}} + (\alpha_{68}((\alpha_{26}y_2)^{\mu_6} + (\alpha_{36}y_3)^{\mu_6}))^{\frac{\mu_8}{\mu_6}}))^{\frac{\mu}{\mu_8}} + (\alpha_{9r}((\alpha_{69}((\alpha_{26}y_2)^{\mu_6} + (\alpha_{36}y_3)^{\mu_6}))^{\frac{\mu_9}{\mu_6}} + (\alpha_{79}((\alpha_{37}y_3)^{\mu_7} + (\alpha_{47}y_4)^{\mu_7}))^{\frac{\mu_9}{\mu_7}}))^{\frac{\mu}{\mu_9}} + (\alpha_{10r}((\alpha_{7,10}((\alpha_{37}y_3)^{\mu_7} + (\alpha_{47}y_4)^{\mu_7}))^{\frac{\mu_{10}}{\mu_7}}))^{\frac{\mu}{\mu_{10}}}.$$

Network MEV model

Comments

- Normalization of the parameters can be complicated depending on the network topology.
- In practice, tree structures should be kept simple.
- Typical applications: multiple level nested logit or cross-nested logit.