ARU discrete choice and generalised entropy models

Mogens Fosgerau

DTU,KTH

June 2015

Outline

- ARUM background, duality
- Models from the dual perspective, GEM
- Discrete choice with GEM
- Model constructions
- Estimation
- Perspectives

Based on joint work with André de Palma: "Demand systems for market shares" (wp available on MPRA)

Outline

ARUM

- Choice set $C = \{1, \cdots, J\}$
- Additive random utility model (ARUM): a utility field defined by $v + \varepsilon$, where
- $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_J)$ is absolutely continuous with finite mean
- v is an additive shift vector, can be specified in accordance with economic theory to depend on observables
- Decision maker maximises random utility
- Choice probability given by

$$P(j|v) = \text{Prob}(v_j + \varepsilon_j = \max_k \{v_k + \varepsilon_k | k \in C\}) \text{ for every } j \in C$$

The surplus

The surplus function is the expected maximum utility

$$G(v) = E \max_j \{v_j + \varepsilon_j\}$$

- (This definition is more straightforward than McFadden's original definition: $\ln G(e^v) = E \max_j \{v_j + \varepsilon_j\}$)
- G is also called a a choice probability generating function
- Choice probabilities: $P(v) = \nabla G(v)$ (Roy's identity)
- Homogeneity:

$$G(v + c) = G(v) + c, \forall c \in \mathbb{R}$$

- G is convex, finite, continuous and closed.
- G is twice continuously differentiable.

Surplus functions

All ARUM surplus functions satisfy a certain set of conditions:

- Homogeneity
- All mixed partial derivatives exist and alternate in sign
- Boundary condition, integrability condition

These conditions are also sufficient for a function to be an ARUM surplus

It is possible to specify ARUM through the surplus function. Examples follow.

Multinomial logit

The basic logit model

$$\begin{aligned} G(v) &= \ln\left(\sum_k e^{v_k}\right) \\ P(j|v) &= \frac{e^{v_j}}{\sum_k e^{v_k}} \end{aligned}$$

Has the IIA property:

$$\frac{P(i|v)}{P(j|v)} = \frac{e^{v_i}}{e^{v_j}}$$

does not depend on other alternatives. We often want something more realistic

Paired combinatorial logit

Any pair of alternatives forms a nest. Large μ_{ik} makes nest dominated by alternative with highest v

$$
\begin{aligned}
G(v) &= \ln\left(\sum_k \sum_{i>k} \alpha_{ik}\left(e^{\mu_{ik} v_i} + e^{\mu_{ik} v_k}\right)^{\frac{1}{\mu_{ik}}}\right), \alpha_{ik} \geq 0, \mu_{ik} \geq 1 \\
P(j|v) &= \frac{\sum_{k \neq j} \alpha_{jk} e^{\mu_{jk} v_j}\left(e^{\mu_{jk} v_j} + e^{\mu_{jk} v_k}\right)^{\frac{1}{\mu_{jk}}-1}}{\sum_k \sum_{i>k} \alpha_{ik}\left(e^{\mu_{ik} v_i} + e^{\mu_{ik} v_k}\right)^{\frac{1}{\mu_{ik}}}}
\end{aligned}
$$

Cross-nested ("sieved")

$$G(v) = \frac{1}{a} \ln \left[\sum_{n=1}^{N} w_n \left(\sum_k \exp\left(\frac{v_k + x_{nk}}{a} \right) \right)^{a^2} \right], 0 < a < 1$$

$$P(j|v) = \frac{\sum_{n=1}^{N} w_n \left(\sum_k \exp\left(\frac{v_k + x_{nk}}{a} \right) \right)^{a^2 - 1} \exp\left(\frac{v_j + x_{nj}}{a} \right)}{\sum_{n=1}^{N} w_n \left(\sum_k \exp\left(\frac{v_k + x_{nk}}{a} \right) \right)^{a^2}}$$

Any RUM may be approximated arbitrarily well with this kind of model (Dagsvik, 1995; Fosgerau et al., 2013).

Models from surplus functions

- So...
- We are used to constructing models via the surplus function
- Many explicit models of the nested logit family are known
- We also know things about duality and ARUM

Duality for logit

$$G(v) = \ln\left(\sum_k e^{v_k}\right), \nabla G(v) = \frac{e^v}{\sum_k e^{v_k}} = q$$

- The convex conjugate of G is (minus) the Shannon entropy

$$\begin{aligned} G^*(q) &= \sup_v \{q \cdot v - G(v)\} = \sup_v \left\{q \cdot v - \ln\left(\sum_k e^{v_k}\right)\right\} \\ &= q \cdot \ln q \end{aligned}$$

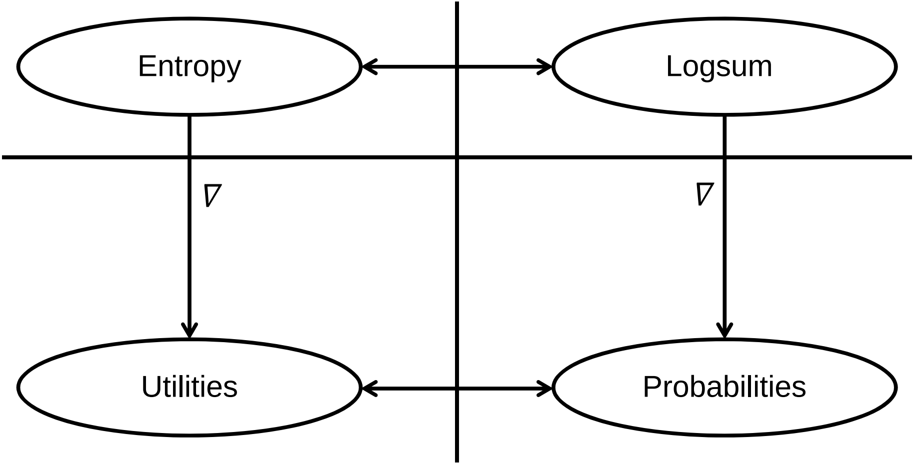
- with gradient

$$\nabla G^*(q) = \ln q + 1 = v + c, c \in \mathbb{R}$$

- and

$$G(v) + G^*(q) = q \cdot v$$

Duality and the logit



Duality for general ARUM

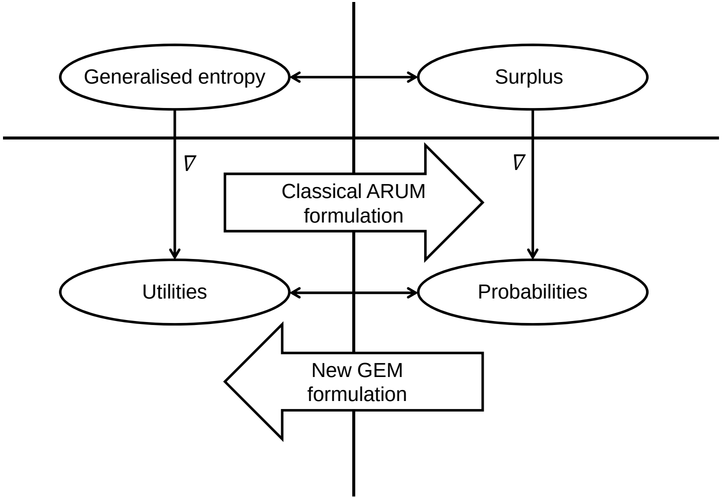
- Let
$$G^*(q) = \sup_v \{q \cdot v - G(v)\}$$
be the convex conjugate of a surplus $G(v)$
- G is a general surplus and therefore $-G^*$ is a **generalised entropy**!
- Let $q = \nabla G(v)$. Then
$$G(v) + G^*(q) = q \cdot v$$
- and
$$q = \nabla G \Leftrightarrow v \in \nabla G^*$$

Outline

This presentation: GEM

- Will present some new stuff!
- Formulate new models based on generalised entropy: GEM
- Models specified in terms on generalised entropy, this is analogous to models specified in terms of surplus
- Comprises dual of all ARUM
- New explicit functional forms
- Allow for endogeneity
- Possible to estimate by regression in simple cases!

Duality and the logit



Generalised entropy

Define a **generalised entropy** by

$$\Omega(q) = \begin{cases} -q \cdot \ln S(q), & q \in \Delta \\ -\infty, & q \notin \Delta \end{cases}$$

where S is a *flexible generator*

$S(\cdot) : [0, \infty]^J \longrightarrow [0, \infty]^J$ is labelled a **flexible generator** when it satisfies the following conditions

1. $S(\cdot)$ is continuous, and homogenous of degree 1
2. $\Omega(q)$ is concave
3. $\sum_{j=1}^{J} q_j \frac{\partial \ln S^{(j)}(q)}{\partial q_k} = 1$
4. $S(\cdot)$ is invertible

$S(q) = q$ implies $\Omega(q) = -q \cdot \ln q$, the Shannon entropy

Demand theorem

- Denote $H(\cdot) \equiv S^{-1}(\cdot)$
- Think about demand for a representative consumer with utility $u = y + q \cdot v + \Omega(q)$ and budget constraint $1 \cdot q = 1$

Theorem

Utility maximization leads to the demand system with interior solution

$$q(v) = \left(\frac{H^{(1)}(e^v)}{\sum_{j=1}^{J} H^{(j)}(e^v)}, \ldots, \frac{H^{(J)}(e^v)}{\sum_{j=1}^{J} H^{(j)}(e^v)} \right)$$

- This generalises the known case connecting the Shannon entropy to logit demand
- $S(q) = q \implies H(e^v) = e^v \implies$ logit!

Inverse demand theorem

Think again about a representative consumer with utility $u = y + q \cdot v + \Omega(q)$ and budget constraint $1 \cdot q = 1$

Theorem

Demand q corresponds to v if and only if $v = \ln S(q) + c$ *for some* $c \in \mathbb{R}$ *(depends on q)*

- Utility can be computed up to a constant directly from demand, given a flexible generator $S(\cdot)$
- This result is useful for estimation

Outline

Discrete choice and generalised entropy

- We now return to ARUM and see how this fits in
- Will show that all ARUM can be represented by GEM

Constructing flexible generator from ARUM

- Begin from an ARUM $v + \varepsilon$ with surplus G
- Define $H\left(e^{v}\right) = \nabla_{v}\left(e^{G(v)}\right)$

Lemma

$H(\cdot)$ is invertible

- $\nabla G(v) = \frac{H(e^{v})}{1 \cdot H(e^{v})}$
- The exponential function is there due to homogeneity to ensure the information about levels carries over

Constructing flexible generator from ARUM

- Define $S(q) = H^{-1}(q)$
- Let $G^*(q) = \sup_v \{q \cdot v - G(v)\}$ be the convex conjugate of $G(\cdot)$
- Let ε^* be the residual for the chosen alternative

Theorem

S is a flexible generator, $-G^*$ is a generalised entropy,

$$G^*(q) = \begin{cases} q \cdot \ln S(q), & q \in \Delta \\ +\infty, & q \notin \Delta \end{cases}$$

$G(v) = \sup_q \{q \cdot v - G^*(q)\}$ *and* $E(\varepsilon^* | v) = -G^*(q)$ *when* $q = \nabla G(v)$.

- This motivates the definition of GEM

Outline

Construction of flexible generators

- To be useful, we must be able to create flexible generators for use in applications
- This is analogous to defining models from the nested logit family
- We want to be able to construct flexible generators S to get new models with desirable properties

Invertibility of flexible generators

(Ruzhansky and Sugimoto, 2014)

- The main issue in constructing flexible generators is to ensure that S is invertible
- Local invertibility is easy to guarantee, global invertibility is a bit more tricky
- The following lemma helps here.

Lemma

Let $J \geq 3$ and let $S(\cdot) : (0,\infty)^J \to (0,\infty)^J$ be continuously differentiable, linearly homogenous with a Jacobian determinant that never vanishes and with $\inf_{q\in\Delta} \|S(q)\| > 0$. Then $S(\cdot)$ is invertible.

Averaging

This lemma shows that a flexible generator may be constructed by averaging. Only one of the elements has to be invertible

Lemma

Let $T_1(\cdot), ..., T_K(\cdot)$ satisfy Conditions 1-3, where the Jacobian of each $\ln T_k(\cdot)$ *is symmetric and positive semidefinite and positive definite for at least one k. Let $T_k^{(j)}(q) \geq q_j$ for each k and j. Let $\alpha_1, ..., \alpha_K$ be positive numbers that sum to 1. Then*

$$S(\cdot) = \prod_{k=1}^{K} (T_k(\cdot))^{\alpha_k}$$

is a flexible generator.

A mapping created by averaging the identity $T_1(q) = q$ with some T_2 that satisfies the conditions of the lemma is always invertible and hence it is a flexible generator

Permutation

A new flexible generator may be constructed from an old by permuting the alternatives.

Lemma

Let S be a flexible generator, let $m \in \mathbb{R}^J$ and let $A = \{a_{ij}\} \in \mathbb{R}^J \times \mathbb{R}^J$ be invertible with $a_{ij} \geq 0$ and $\sum_i a_{ij} = 1$. Then

$$T : q \to \exp\left(A^\top \left[\ln\left(S\left(Aq\right)\right)\right] + m\right)$$

is a flexible generator.

Expansion of an alternative

A flexible generator may be plugged into another to expand an alternative

Lemma

Let T_1, T_2 be flexible generators with $T_1 : \mathbb{R}^{J_1} \rightarrow \mathbb{R}^{J_1}$ and $T_2 : \mathbb{R}^{J_2} \rightarrow \mathbb{R}^{J_2}$. Then $S : \mathbb{R}^{J_1+J_2-1} \rightarrow \mathbb{R}^{J_1+J_2-1}$ defined for $q^1 \in \mathbb{R}^{J_1}$ and $q^2 \in \mathbb{R}^{J_2-1}$ by

$$S^{(j)}\left(q^1, q^2\right) = \begin{cases} T_1^{(j)}\left(\frac{q^1}{1 \cdot q^1}\right) T_2^{(j)}\left(1 \cdot q^1, q^2\right), & j \leq J_1 \\ T_2^{(j)}\left(1 \cdot q^1, q^2\right), & J_1 < j \leq J_1 + J_2 - 1 \end{cases} \tag{1}$$

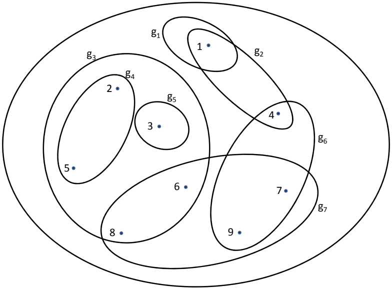
is a flexible generator with inverse given by $H\left(e^{v^1}, e^{v^2}\right) = \left(sT_1^{-1}\left(e^{v^1}\right), q^2\right)$, *where* $\left(s, q^2\right) = T_2^{-1}\left(1, e^{v^2}\right)$.

Concrete examples

- The preceeding results show how new flexible generators may be created from old with a range of general operations
- The next step is to provide some concrete examples of flexible generators
- We already know one example, namely $S(q) = q$

Nesting operation

- A nest g is a set of goods for which a term $\left(\sum_{i\in g} q_i\right)^{\mu_g}$ enters generalised entropy (and then the utility of a representative consumer)
- $\mu_g \in]0,1]$ is a nesting parameter
- The closer μ_g is to 1, the more the goods in nest g act in utility as one single good and they become closer to being perfect substitutes



General nesting construction

Let nests $g \in \mathcal{G}$ be subsets of the set of alternatives $\{1, ..., J\}$.
Let $S(\cdot) = \left(S^{(1)}(\cdot), ..., S^{(J)}(\cdot)\right)$ be given by

$$S^{(j)}(q) = \prod_{\{g \in \mathcal{G} | j \in g\}} \left(\sum_{i \in g} q_i\right)^{\mu_g},$$

where $\sum_{\{g \in \mathcal{G} | j \in g\}} \mu_g = 1$ for all j and $\mu_g > 0$ for all $g \in \mathcal{G}$.

Theorem

Then $S(\cdot)$ satisfies C1-C3 (everything except invertibility). Moreover, the Jacobian of $\ln S(\cdot)$ is symmetric and positive semidefinite, and for each j, $S^{(j)}(q) \geq q_j$. If the Jacobian of $\ln S(\cdot)$ is positive definite, then $S(\cdot)$ has an inverse and $S(\cdot)$ is a flexible generator.

Example of general nesting construction

Consider $J \geq 3$ with all possible nests with 1 or 2 alternatives as elements, e.g. for $J = 3$:

$$\mathcal{G} = \{\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}.$$

Each alternative is in J nests and we let $\mu_g = 1/J$.
The $S(\cdot)$ is given by

$$S^{(j)}(q) = q_j^{\frac{1}{J}} \prod_{i \neq j} (q_i + q_j)^{\frac{1}{J}}.$$

This is a flexible generator

Nested logit is a special case

Partition the set of alternatives $\{1, ..., J\}$ into nests $g \in \mathcal{G}$ and denote by g_j the nest that contains alternative j. Let

$$S^{(j)}(q) = q_j^{\mu_{g_j}} \left(\sum_{i \in g_j} q_i \right)^{1-\mu_{g_j}}, j \in g_j,$$

where $\mu_g \in]0,1]$ are parameters. Then $S(\cdot)$ is a flexible generator
Demand is

$$q_j = \frac{\tilde{q}_j}{\sum_{g \in \mathcal{G}} \sum_{k \in g} \tilde{q}_k} = \frac{e^{\frac{v_j}{\mu_{g_j}}}}{\sum_{i \in g_j} e^{\frac{v_i}{\mu_{g_j}}}} \frac{e^{\mu_{g_j} \ln\left(\sum_{i \in g_j} e^{\frac{v_i}{\mu_{g_j}}}\right)}}{\sum_{g \in \mathcal{G}} e^{\mu_g \ln\left(\sum_{i \in g} e^{\frac{v_i}{\mu_g}}\right)}},$$

which is a nested logit model (McFadden, 1978)

Cross-nesting construction example

Let $\mu_0, \mu_1, \mu_2 > 0$, $\mu_0 + \mu_1 + \mu_2 = 1$. Let $\sigma_c(j)$ be the set of products that are grouped together with product j on criteria $c = 1, 2$. Denote $I_c(j) = \sum_{i \in \sigma_c(j)} q_i$ and define $S(\cdot)$ by

$$S^{(0)}(q) = q_0$$
$$S^{(j)}(q) = q_j^{\mu_0} I_1(j)^{\mu_1} I_2(j)^{\mu_2}, j > 0.$$

Then $S(\cdot)$ is a flexible generator

E.g. cars grouped by make, body type, fuel type

Invertible nesting construction

Theorem

Let $S(\cdot)$ be given by nesting construction, where the number of nests is equal to the number of alternatives. Define $W = diag\left(\mu_{g_1}, .., \mu_{g_J}\right)$ and the incidence matrix $M_{J \times J} = \left\{1_{\{j \in g\}}\right\}$, where rows correspond to alternatives and columns correspond to nests, and suppose that M is invertible. Then $S(\cdot)$ has an inverse, that is

$$\ln S(q) = v \Leftrightarrow q = \left(M^{\top}\right)^{-1} \exp\left(W^{-1} M^{-1} v\right),$$

and $S(\cdot)$ is a flexible generator.

Invertible nesting example

- $J \geq 3,\ M = \left\{1_{\{i \neq j\}}\right\}$
- Each alternative is in $J-1$ nests and we associate weights $\mu_g = 1/\left(J-1\right)$ with each nest
- Then

$$q_i = \frac{\sum_{j=1}^{J} e^{-(J-1)v_j} - \left(J-1\right) e^{-(J-1)v_i}}{\sum_{j=1}^{J} e^{-(J-1)v_j}}$$

- Zero demands may arise

Theorem

This is not ARUM consistent. Generalised entropy (as defined) comprises non-ARU models

Outline

Estimation from market shares

- Now consider how GEM may be estimated using observations of market shares
- This can be done just by regression!
- I will show some examples
- First talk about how to allow correction for endogeneity using ideas from Berry (1994)

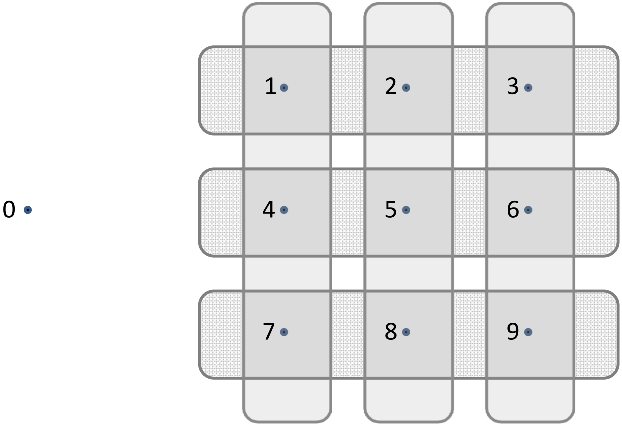
Estimation Berry (1994) style

- J products and an outside good
- q_j depends only on mean utility levels $v = (v_1, ..., v_J)$, where $v_j = z_j \cdot \beta + \xi_j$.
- z_j : vector of variables
- β : vector of parameters to be estimated
- ξ_j: unobserved demand characteristic of product j, mean independent of z and independent across markets. Soaks up endogeneity, perhaps related to price and unobserved quality
- $v_0 = 0$

$$\ln S^{(j)}(q) - \ln S^{(0)}(q) = z_j \cdot \beta + \xi_j$$

- Given a specific form for $S(\cdot)$, this may be estimated using IV

Example: cross-nested model



Cross-nested model

- Insert $S^{(j)}(q) = \begin{cases} q_0, & j = 0 \\ q_j^{\mu_0} l_1(j)^{\mu_1} l_2(j)^{\mu_2}, & j > 0 \end{cases}$ into
- $\ln S^{(j)}(q) - \ln S^{(0)}(q) = z_j \cdot \beta + \xi_j$ and get
- regression with alternative specific disturbances $\tilde{\xi}_j$:

$$\ln q_j = z_j \cdot \tilde{\beta} - \tilde{\mu}_1 \ln \left(\sum_{i \in \sigma_1(j)} q_i \right) - \tilde{\mu}_2 \ln \left(\sum_{i \in \sigma_2(j)} q_i \right) + \delta \ln q_0 + \tilde{\xi}_j$$

- Endogeneity: q_j appears on the RHS
- Instruments:

$$1, z_j, \sum_{i \in \sigma_1(j)} z_i, \sum_{i \in \sigma_2(j)} z_i, \sum_i z_i$$

and squares of these

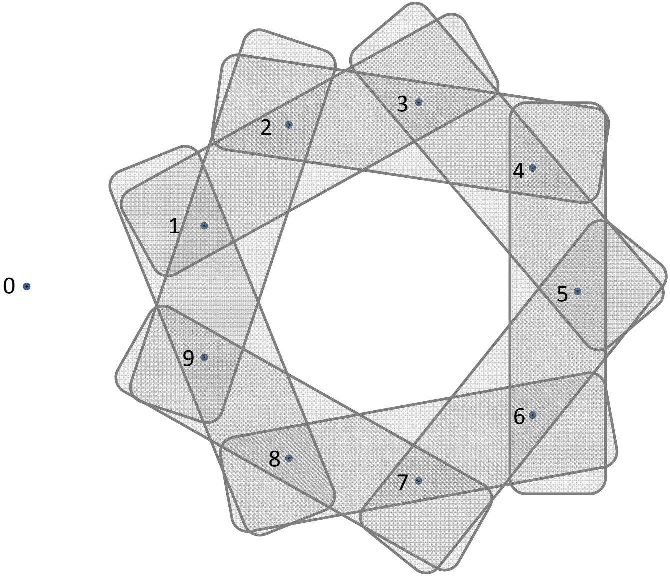
Simulation example - cross-nested model

- $z_j \sim N(0,1)$, $\tilde{\xi}_j \sim 0.5 \cdot N(0,1)$
- 1000 datasets with 100 observations in each
- F-stats for excluded instruments in the first-stage: >100 for $\ln\left(\sum_{i\in\sigma_1(j)} q_i\right)$ and $\ln\left(\sum_{i\in\sigma_2(j)} q_i\right)$, >30 for $\ln q_0$

Table: Parameter estimates in simulation with cross-nested model

	$\widetilde{\beta}$	$-\widetilde{\mu}_1$	$-\widetilde{\mu}_2$	δ
True parameters	2	-0.2	-0.8	2
Avg. IV estimates	2.00	-0.20	-0.79	1.99
Std.errs.	0.04	0.05	0.08	0.06
Avg. OLS estimates	1.76	0.10	-0.41	1.59
Std.errs.	0.04	0.04	0.05	0.05

Example: ordered model



Ordered model

- Regression:

$$\ln q_j = z_j \cdot \tilde{\beta} - \tilde{\mu}_1 \ln \left(\sum_{j-2 \leq i \leq j} q_i \right) - \tilde{\mu}_2 \ln \left(\sum_{j-1 \leq k \leq = j+1} q_i \right)$$
$$- \tilde{\mu}_3 \ln \left(\sum_{j \leq k \leq j+2} q_i \right) + \delta \ln q_0 + \tilde{\xi}_j$$

- Instruments:

$$1, z_j, \sum_{j-2 \leq i \leq j} z_i, \sum_{j-1 \leq k \leq = j+1} z_i, \sum_{j \leq k \leq j+2} z_i$$

and squares of these

Ordered model results

- 1000 datasets with 100 observations in each
- $z_j \sim N(0,1)$ and $\tilde{\xi}_j \sim 0.5 \cdot N(0,1)$
- F-stats for the excluded instruments in the first-stage regression very high.

Table: Parameter estimates in simulation with ordered model

	$\widetilde{\beta}$	$-\widetilde{\mu}_1$	$-\widetilde{\mu}_2$	$-\widetilde{\mu}_3$	δ
True parameters	2.50	-0.50	-0.50	-0.50	2.50
Avg. IV estimates	2.49	-0.49	-0.49	-0.49	2.49
Std.errs.	0.06	0.08	0.08	0.08	0.08
Avg. OLS estimates	2.16	-0.10	-0.36	-0.10	1.91
Std.errs.	0.06	0.05	0.06	0.06	0.06

Estimation from discrete choices

- Apply the generalised entropy model to discrete choices
- Demonstrate feasibility of estimation by maximum likelihood

Setup

- An individual chooses good j with probability q_j satisfying $v = \ln S(q) + c$ for some flexible generator S and with $c \in \mathbb{R}$ ensuring that probabilities sum to 1
- If the generalised entropy is the convex conjugate of an ARUM surplus function, then q are the corresponding discrete choice probabilities
- Generalised entropies that are not ARUM consistent may still correspond to nonadditive random utility models, i.e. models where utilities are not just sums but more general functions of v_j and ε_j (Matzkin, 2007) (additivity is just a convenience)
- Alternatively, individuals could be seen as making random choices with probabilities that are the result of utility maximization (Fudenberg et al., 2014)

Maximum likelihood with NFXP algorithm

- Estimation by maximum likelihood
- Must compute the likelihood q given v
- Need a way to invert S that is feasible within a maximum likelihood routine
- The next theorem indicates how the likelihood may computed by using an iterative process to solve a fixed-point problem.
- Use the Kullback and Leibler (1951) distance function to evaluate the distance from the fixed point r to some q

$$d_r(q) = r \cdot \ln\left(\frac{r}{q}\right)$$

- This is a convex function with minimum at r with $d_r(r) = 0$. Hence $d_r(q)$ will be larger the further q is from r.

Convergence to the fixed point

Theorem

Let S be a flexible generator constructed by general nesting and let $r \in \Delta$ satisfy $v = \ln S(r) + c$ for some $c \in \mathbb{R}$. Then the mapping

$$w(q) = \left\{ \frac{q_i e^{v_i} / S^{(i)}(q)}{\sum_j q_j e^{v_j} / S^{(j)}(q)} \right\} \tag{2}$$

has r as unique fixed point and iteration of (2) from any starting point in Δ converges to r.

- Intuitively, the numerator of (2) adjusts each q_i in the direction that makes $v = \ln S(q) + c$ true, while the denominator ensures that $1 \cdot w(q) = 1$.

Speed of convergence to the fixed point

Theorem

(continued) If S has the form

$$S^{(j)}(q) = q_j^{\mu_0} \prod_{\{g \in \mathcal{G} | j \in g, g \neq \{j\}\}} \left(\sum_{i \in g} q_i \right)^{\mu_g}$$

for some $\mu_0 > 0$, then $d_r(w(q)) \leq (1 - \mu_0) d_r(q)$.

- The second half of the theorem concerns the special case when the flexible generator is an average of the identity with something else.
- Beginning from q^0 and iterating such that $q^n = w\left(q^{n-1}\right), n \geq 1$ the theorem shows that $d_r(q^n) \leq (1 - \mu_0)^n d_r\left(q^0\right)$, which means that the distance to the fixed point decreases exponentially

Simulation experiment

- Simulated data from the cross-nested structure shown before. No outside option
- Utilities are $v_j = \alpha x_{1j} + \beta x_{1j} x_2$, where x_{1j} represents an alternative specific characteristic, while x_2 represents individual specific variation.
- 100 replications with 1000 individuals in each
- Each individual selects 1 among the 9 alternatives in the model with probabilities q, where $\ln S(q) = v + c$.
- The independent variables were generated as i.i.d. standard normal. The likelihood was computed using the fixed point theorem and was maximized numerically.

Simulation results

The true parameters are well recovered

Table: Maximum likelihood estimates in discrete choice simulation with cross-nested model

	α	β	μ_1	μ_2
True parameters	0.500	0.500	0.200	0.500
Avg. estimates	0.498	0.498	0.208	0.495
Std.dev.	0.050	0.050	0.043	0.055

Outline

Estimation from individual choice data

- It is desirable to allow **many** dummies to soak up endogeneity
- Estimating large models with many dummies is a computational challenge
- There are some ideas to be explored, using fixed-point algorithms to compute dummies
- Nested pseudo likelihood seems feasible (**?**)

Some open questions

- Can more specific flexible entropies be found? Partial membership, distance to centres?
- Is there a result that any GEM/ARUM may be approximated by a specific GEM? (Just as cross-nested logit may approximate any RUM)
- Is it possible to characterise those GEM that are dual to an ARUM?
- How does GEM fit into a dynamic discrete choice framework?

Perspectives

- GEM is a new universe of models
- So far unexplored
- ARUM "inside out"!
- Some attractive features:
  - Regression on market shares
  - Handle endogeneity
  - Specify substitution patterns in a structured and transparent way through nesting operation

References

Aguirregabiria, V. and Magesan, A. (2013) Euler Equations for the Estimation of Dynamic Discrete Choice Structural Models *Advances in Econometrics* **31**, 3–44.

Berry, S. (1994) Estimating discrete-choice models of market equilibrium *The RAND Journal of Economics* **25**(2), 242–262.

Dagsvik, J. (1995) How large is the class of generalized extreme value random utility models? *Journal of Mathematical Psychology* **39**(1), 90–98.

Fosgerau, M., McFadden, D. and Bierlaire, M. (2013) Choice probability generating functions *Journal of Choice Modelling* **8**, 1–18.

Fudenberg, D., Iijima, R. and Strzalecki, T. (2014) Stochastic Choice and Revealed Perturbed Utility *Working Paper* .

Kullback, S. and Leibler, R. A. (1951) On Information and Sufficiency *The Annals of Mathematical Statistics* **22**(1), 79–86.

Matzkin, R. L. (2007) Chapter 73 Nonparametric identification *in* J. J. H. a. E. E. Leamer (ed.), *Handbook of Econometrics* Vol. 6, Part B Elsevier pp. 5307–5368.

McFadden, D. (1978) Modelling the choice of residential location *in* A. Karlquist, F. Snickars and J. W. Weibull (eds), *Spatial Interaction Theory and Planning Models* North Holland Amsterdam pp. 75 –96.

Ruzhansky, M. and Sugimoto, M. (2014) On global inversion of homogeneous maps *Bulletin of Mathematical Sciences* pp. 1–6.