Generalised entropy models

Mogens Fosgerau   André de Palma

December 2015

Outline

ARUM

- Decision maker makes discrete choice to maximise random utility $v + \varepsilon = (v_1, \cdots, v_J) + (\varepsilon_1, \cdots, \varepsilon_J)$
- The surplus function is the expected maximum utility

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

- It is possible and convenient to specify ARUM through the surplus function

Multinomial logit and duality

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

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

$$G^*(q) = \sup_v \{q \cdot v - G(v)\} = q \cdot \ln q$$

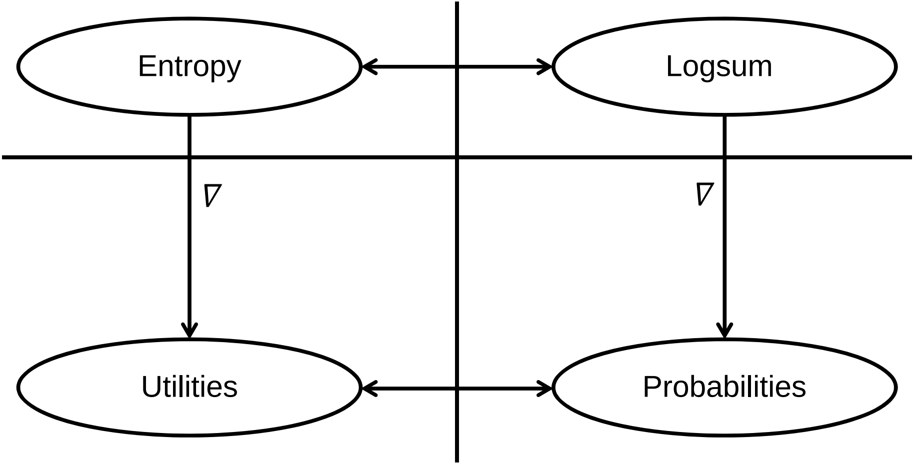
with gradient

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

and

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

Duality and the logit



Duality for general ARUM

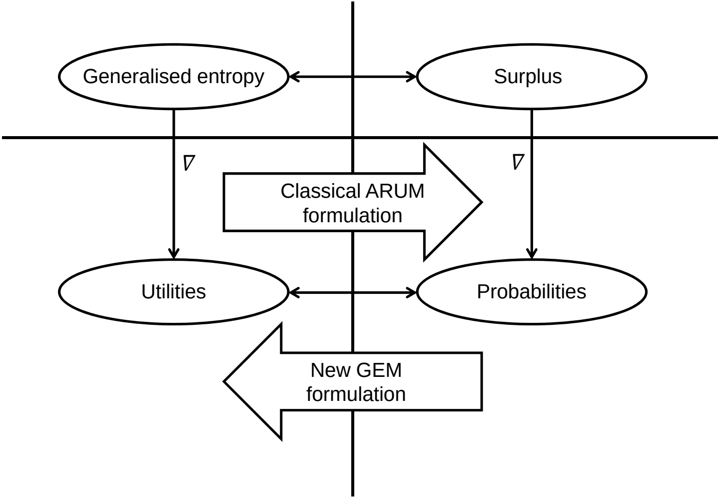
- Let
$$G^*(q) = \sup_v \{q \cdot v - G(v)\}$$
be the convex conjugate of a surplus $G(v)$
- G is a general surplus and therefore $-G^*$ is a **generalised entropy**!
- Let $q = \nabla G(v)$. Then
$$G(v) + G^*(q) = q \cdot v$$
- and
$$q = \nabla G \Leftrightarrow v \in \nabla G^*$$

Outline

Generalised entropy models

- Models specified in terms on generalised entropy, (analogous to models specified in terms of surplus)
- Comprise dual of all ARUM, and more
- New explicit functional forms
- Allow for endogeneity
- Possible to estimate by regression from market shares

Duality and ARUM



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*

$S(\cdot) : [0, \infty]^J \to [0, \infty]^J$ 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

Special case: $S(q) = q$, $\Omega(q) = -q \cdot \ln q$, the Shannon entropy

Demand theorem

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

Theorem

Utility maximization leads to the demand system with interior solution

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

- This generalises the known case connecting the Shannon entropy to logit demand
- $S(q) = q \Longrightarrow H(e^v) = S^{-1}(e^v) = e^v \Longrightarrow$ 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

All ARUM can be represented as GEM

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

Lemma

$H(\cdot)$ is invertible

- Demand: $\nabla G(v) = \frac{H(e^{v})}{1 \cdot H(e^{v})}$

All ARUM can be represented as GEM

- 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* $q = \nabla G(v)$.

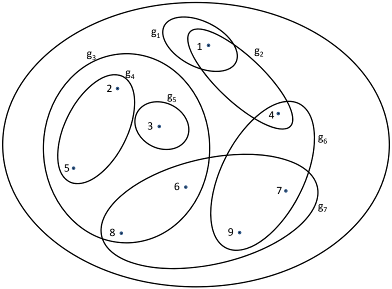
Outline

Construction of flexible generators

- The paper comprises a number of results for constructing new flexible generators from old with a range of general operations
- The simplest flexible generator: $S(q) = q$
- Will show a few more

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

If the Jacobian of $\ln S(\cdot)$ *is positive definite, then* $S(\cdot)$ *has an inverse and* $S(\cdot)$ *is a flexible generator.*

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

$$S^{(0)}(q) = q_0$$

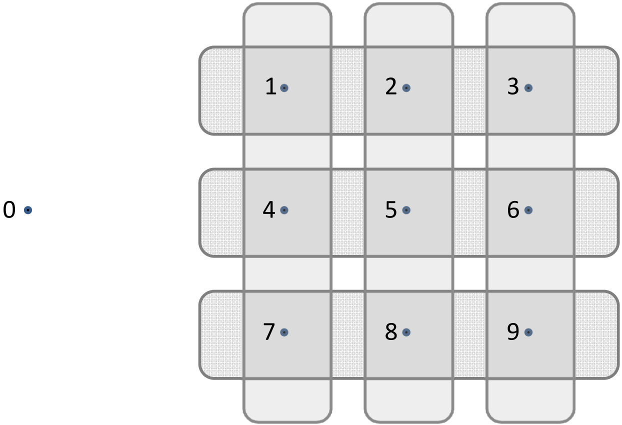
$$S^{(j)}(q) = q_j^{\mu_0} I_1(j)^{\mu_1} I_2(j)^{\mu_2}, j > 0.$$

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

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

Outline

Example: cross-nested model



Cross-nested model

- Rewrite $z_j\beta + \xi_j = \ln S^{(j)}(q) + c$ to 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
- Regression on market shares with 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
- Can accommodate endogeneity of z_j with further instruments

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

Estimation from discrete choices

- 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
- Estimation by maximum likelihood requires computation of q given v
- Need a way to invert S that is feasible within a maximum likelihood routine

NFXP algorithm

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

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

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

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

Speed of convergence to the fixed point

This concerns the special case when the flexible generator is an average of the identity with something else. $d_r(q)$ is Kullback-Leibler distance from r to q

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 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 $\ln S\left(q\right) = v + c$.
- The independent variables were generated as i.i.d. standard normal.
- The likelihood was computed using the fixed point theorem and was maximized numerically (NFXP)

Simulation results

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

Perspectives and questions

- It is desirable to allow **many** dummies to soak up endogeneity: BLP, residential sorting. This is a computational challenge. Perhaps use fixed-point algorithms.
- 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, to an MEV ARUM?
- How does GEM fit into a dynamic discrete choice framework?

Perspectives

- GEM is a new universe of models
- ARUM "inside out"!
- Some attractive features of dual formulation
  - Regression on market shares
  - Handle endogeneity
  - Specify substitution patterns in a structured and transparent way through nesting operation
  - Hope to find more

"Demand systems for market shares", WP available on MPRA