Computing a QRE Branch
Turocy (2005) & Turocy (2010)

Environment and Notation

- Set of players $N = \{1, \ldots, n\}$
- S_i - set of strategies available to each player $i \in N$
- J_i - the number of strategies in S_i
- $s_{ij} \in S_i$ with $1 \leq j \leq J_i$
- The set of strategy profiles $S = \times_{i=1}^{n} S_i$ is the Cartesian product of S_i
- $J = \sum_{i=1}^{n} J_i$
- $u_i : S \rightarrow \mathbb{R}$ is the payoff/utility function for player i
- $\boldsymbol{\pi_{ij}}$ - the probability assigned to a strategy s_{ij} of player i
- Δ_i - the set of probability distributions (mixed strategy) over S_i
- $\Delta = \times_{i=1}^{n} \Delta_i$, $\pi \in \Delta$ – mixed strategy profile

Environment and Notation

- (s_{ij}, π_{-i}) – mixed strategy profile where player i plays s_{ij} with probability 1, while all other players play according to the mixed strategy π
- $\bar{u}_{ij}(\pi) \equiv u_i(s_{ij}, \pi_{-i})$ – expected payoff/utility to player i from playing his jth strategy, holding fixed his opponents' mixed strategies π_{-i}

- In the quantal response framework players observe a noisy evaluation of the strategy values
  - $\hat{u}_{ij}(\pi) = \bar{u}_{ij}(\pi) + \epsilon_{ij}$
  - where the ϵ_{ij} are random variables drawn according to some joint distribution. When ϵ_{ij} are drawn from an extreme value distribution with parameter λ, a logistic quantal response equilibrium profile (logit equilibrium) is given by

$$\pi_{ij} = \frac{e^{\lambda \bar{u}_{ij}(\pi)}}{\sum_{k=1}^{J_i} e^{\lambda \bar{u}_{ik}(\pi)}}$$

Computational Problem

- Trace out a path of mixed strategy profiles π for different λ
- The logit equilibria are expressed as the zeros of a system of equations

$$H(\pi, \lambda) = 0$$

- $\sum_{i=1}^{n} J_i$ equations comprising the system $H(\pi, \lambda)$ are indexed $H_{ij}(\pi, \lambda)$ corresponding to the jth strategy of player i

$$H_{ij}(\pi, \lambda) = e^{\lambda \bar{u}_{ij}(\pi)} - \pi_{ij} \sum_{k=1}^{J_i} e^{\lambda \bar{u}_{ik}(\pi)} = 0$$

- n equations comprising the system $H(\pi, \lambda)$ are indexed $H_i(\pi, \lambda)$

$$H_i(\pi, \lambda) = 1 - \sum_{k=1}^{J_i} \pi_{ik} = 0$$

Computational Problem Simplified

- For each agent i we can choose some reference strategy $s_{ik} \in S_i$, then

$$\frac{\boldsymbol{\pi_{ij}}}{\boldsymbol{\pi_{ik}}} = \frac{e^{\lambda\,\bar{u}_{ij}(\pi)}}{e^{\lambda\,\bar{u}_{ik}(\pi)}}$$

- We can also convert to logs:

$$\boldsymbol{ln(\pi_{ij}) - ln(\pi_{ik}) = \lambda[u_{ij}(\pi) - u_{ik}(\pi)]}$$

- Then the system $H(\pi, \lambda) = 0$ is defined by $\sum_{i=1}^{n}(J_i - 1)$ equations of the form

$$\boldsymbol{H_{ijk}(\pi, \lambda) = ln(\pi_{ij}) - ln(\pi_{ik}) - \lambda[u_{ij}(\pi) - u_{ik}(\pi)] = 0}$$

- And n equations of the form

$$\boldsymbol{H_i(\pi, \lambda) = 1 - \sum_{k=1}^{J_i} \pi_{ik} = 0}$$

Computational Problem Simplified

- Let $p_{ij} = \ln(\pi_{ij})$
- Then the system $H(p, \lambda) = 0$ is defined by $\sum_{i=1}^{n}(J_i - 1)$ equations of the form

$$\boldsymbol{H_{ijk}(p, \lambda) = p_{ij} - p_{ik} - \lambda[u_{ij}(e^p) - u_{ik}(e^p)] = 0}$$

- And n equations of the form

$$\boldsymbol{H_i(p, \lambda) = 1 - \sum_{k=1}^{J_i} e^{p_{ik}} = 0}$$

These n equations can also be imposed as a constraint for minimization algorithm

The objective function is:

$$\boldsymbol{F(p, \lambda) = H(p, \lambda)^{\intercal} H(p, \lambda)}$$

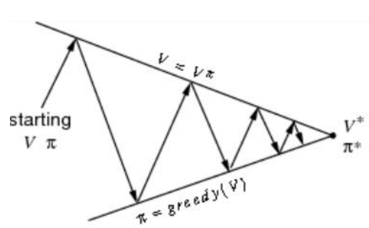
A quantal response equilibrium is at 0 (minimum) of $F(p, \lambda)$

Computational Steps

1. Let $\{\lambda_t\}$ be a sequence such that $\lambda_t > \lambda_{t+1}$
2. Initialize
   - $\lambda_0 = 0$
   - $\pi_{ij} = \frac{1}{J_i}$
   - $p_{ij} = \ln(\pi_{ij})$
3. Find minimum of $F(\lambda_t, \pi)$ using π_{t-1} as initial guess

A different approach to QRE: Policy Iteration

- Policy iteration consists of two simultaneous, interacting processes, one making the value function consistent with the current policy (policy evaluation), and the other making the policy "greedy" with respect to the current value function (policy improvement)
  - In policy iteration these two alternate, each completing before the other begins.



For our problem policy is π_i, and value associated with that policy is $u_i(\pi)$

For a given λ

1. Initialize $\pi_{ij} = \frac{1}{J_i}$
2. Evaluate $u_i(\pi)$
3. Improve
   - $\pi'_{ij} = \frac{e^{\lambda\,\bar{u}_{ij}(\pi)}}{\sum_{k=1}^{J_i} e^{\lambda\,\bar{u}_{ik}(\pi)}}$
   - If $\pi' = \pi$ then stop; else go to 2.

- In practice the improvement step is incremental
  - π'=(1-α)π+$\alpha\pi'$