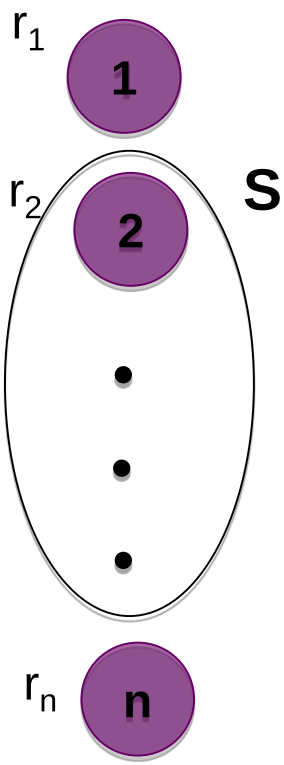
A Markov Chain Substitution Scheme for Approximation of Choice Models

Jose Blanchet
Columbia University

Joint work with
Guillermo Gallego and Vineet Goyal

Assortment Planning Problem



How to Estimate?

Choice Model

$$\pi_j(S) = P(\text{customer selects item j when offer set is } S)$$

How to Optimize?

Assortment Problem

Find S to $\max \sum_{j \in S} r_j \cdot \pi_j(S)$

Choice Model

- Ordered preference list (permutation) of items
- Customer selects the most preferable item available
- Most general choice model: **distribution over all permutations**
- Tradeoffs
  - Complex choice model: hard to optimize
  - Simple choice model: not rich enough

Common Choice Models

- Multinomial Logit model: MNL (Plackett-Luce, MacFadden (1974))
  - utility parameters: w_i for item *i,* attribute w_i+X_i & X_i*'s i.i.d.* Gumbel.
  - choice probabilities $\pi_j(S) = \frac{u_j}{\sum_{i \in S_+} u_i}$ (S_+ = S U {0})

Model Selection: which is the right model?

- True choice model is latent
- We only observe choice or sales data
- Error in model selection can lead to highly suboptimal decisions

Related Work

- Smith and Agrawal (2000), Netessine and Rudi (2003): two-step dynamic substitution.

- Other dynamic substitution & data inference: Saure and Zeevi (2009), Rusmevichientong and Topaloglu (2009).

- Farias et al. (2010)
  - Estimate distribution over permutation with sparsest support consistent data
  - Can efficiently provide estimator under some conditions

- Vulcano, van Ryzin, Ratliff (2012)
  - EM algorithm to estimate a semi-parametric family of choice models

This Talk

- **Markov chain based** “Universal” choice “model” (really a computational tool).

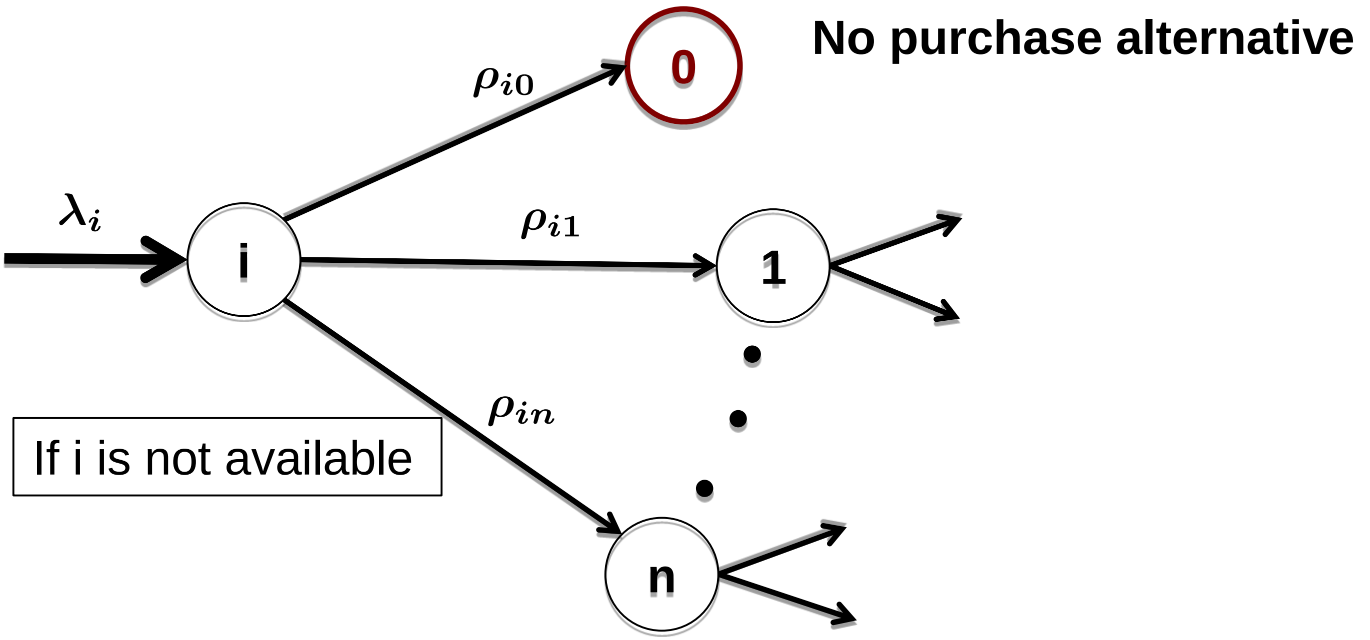
- Can be estimated efficiently
  - $O(n^2)$ parameters

- **Universal approximation** for all random utility models
  - Exact if the underlying model is MNL
  - Good approximation bounds for general random utility model

- **Efficient assortment optimization**

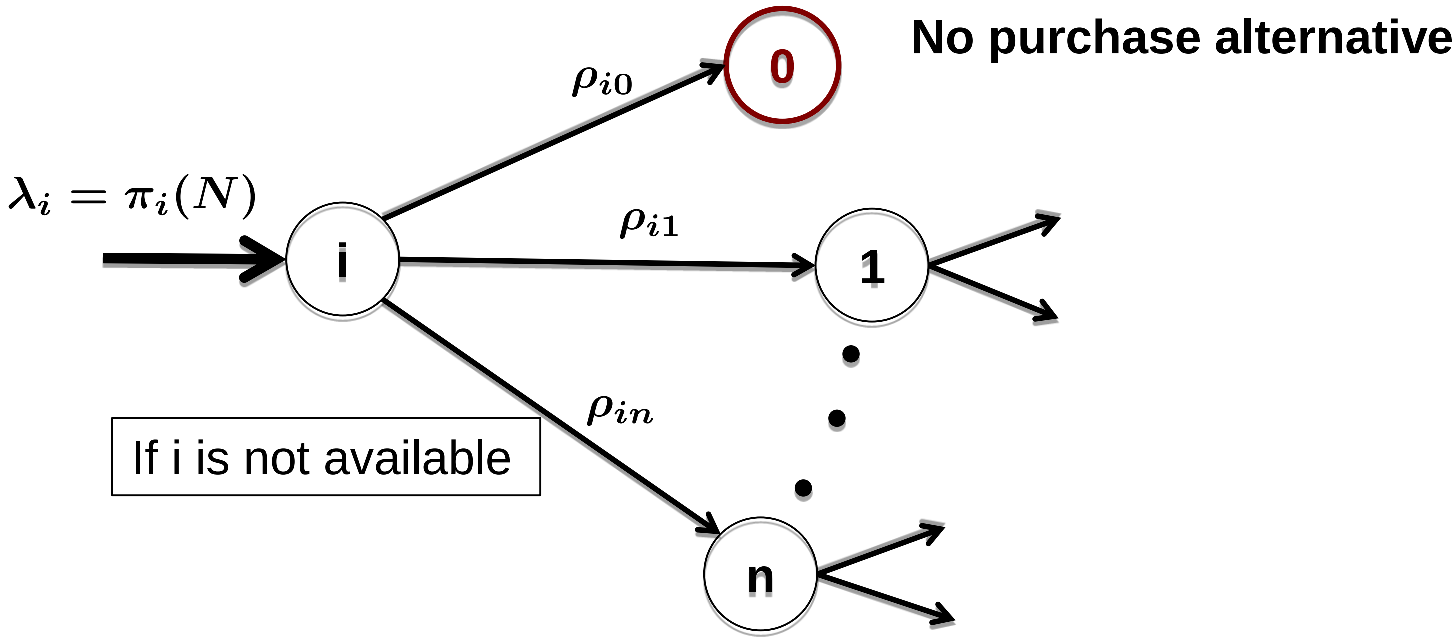
Markov Chain Based Model

- New primitive for substitution behavior



- **No transitions out of state 0**
- **Markovian model**
  - After transition to state j, customer behaves like first choice being j
- **Specified by O(n^2) transition probability parameters**

Estimating Markov Chain Model



$$\rho_{ij} = \frac{\pi_j(N \setminus \{i\}) - \pi_j(N)}{\pi_i(N)}$$

Fraction of customers who select j given the first choice is i

Estimating Markov Chain Model

- **Suppose distribution over permutation σ given by p(σ)**

$$\lambda_i = P(\sigma(1) = i)$$

$$\rho_{i,j} = P(\sigma(2) = j | \sigma(1) = i)$$

Estimating Markov Chain Model

$$\rho_{ij} = \frac{\pi_j(N \setminus \{i\}) - \pi_j(N)}{\pi_i(N)}$$

Fraction of customers who select j given the first choice is i

- **Data required to estimate the model**

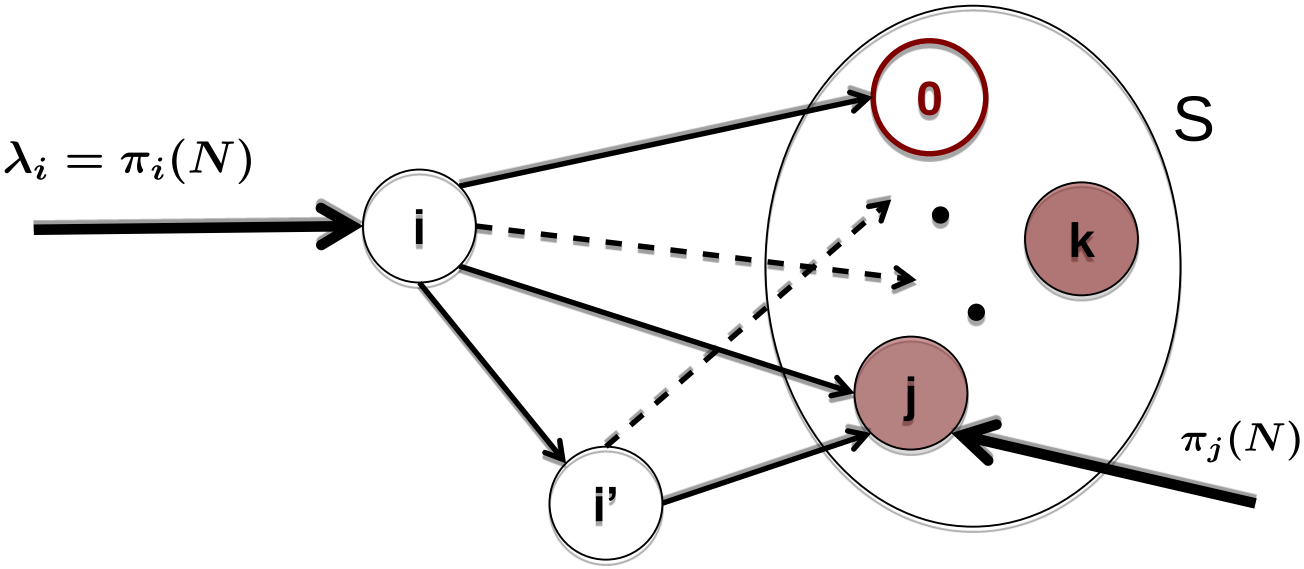
$$\pi_j(N),\ \pi_j(N \setminus \{i\}),\ \forall i, j$$

- **Choice probability data for n offer sets**

- **Given S we can estimate**

$$\rho_{ij} = \frac{\pi_j(S \setminus \{i\}) - \pi_j(S)}{\pi_i(S)},\ \text{if } i, j \in S$$

Computing Choice Probability Estimates



- **Define Markov chain for offer set S: M(S)**
  - All states in S (including 0) are absorbing states
- **Estimate of choice probability of item j in S**

$$\hat{\pi}_j(S) = P(M(S) \text{ absorbs in state } j \text{ given initial arrival probabilities } \lambda_i)$$

- **Can be computed efficiently for any j, S**
  - No closed form expression

Approximation Bounds: MNL Model

- **Suppose underlying model is MNL with parameters u_i for all i**

$$\rho_{ij} = \frac{u_j}{\sum_{\ell \in N_+} u_\ell - u_i} = \pi_j(N \setminus \{i\})$$

Theorem 1 *If the underlying model is MNL with parameters u_i for all $i = 0, \ldots, n$. Then for any offer set $S \subseteq [n]$ and $j \in S$,*

$$\hat{\pi}_j(S) = \frac{u_j}{\sum_{i \in S_+} u_i} = \pi_j(S)$$

Approximation Bounds: Other Models

- **McFadden and Train (2000)**
  - Every random utility based model can be approximated arbitrarily closely by a mixture of MNL
- **Suffices to prove approximation bounds for mixture models**
- **Consider a mixture of MNL model with K segments**
  - Probability of segment k is θ_k
  - Parameters for segment k: $u_0^k, \ldots, u_n^k$
  - Assume wlog. $u_0^k + \ldots + u_n^k = 1$
- **Choice Probability:**

$$\pi_j(S) = \sum_{k=1}^{K} \theta_k \frac{u_j^k}{\sum_{i \in S_+} u_i^k}$$

Approximation Bounds: MMNL Model

MMNL model (with K segments)

$$\rho_{ij} = \sum_{k=1}^{K} P(\text{ segment } k \mid \text{first choice is } i) \cdot \frac{u_j^k}{1 - u_i^k}$$

Theorem 2 *If the underlying model is mixture of MNL with K segments, probabilities θ_k of segment k, and MNL parameters $u_0^k, \ldots u_n^k$. For any offer set $S \subseteq [n]$, let $\alpha = \max_k u^k(\bar{S})$. Then for any $j \in S$,*

$$\pi_j(S)(1-\alpha^2) \leq \hat{\pi}_j(S) \leq (1+\alpha^2/(1-\alpha))$$

For α=0.5, we get a 0.75-approximation for choice probabilities

Approximation Bounds: MMNL Model

- **An example (2 classes of customers completely asymmetric utilities) shows that bounds are sharp.**
- **Numerical experiment with random u_j's & report average over 500 randomly picked offer sets S (of sizes 30% to 60% n).**

		Average worst case relative error in choice probabilities	
Case	**n**	**K=log(n)**	**errMC(%)**
1	10	3	**3.1**
2	20	3	**2.4**
3	30	4	**2.5**
4	40	4	**2.4**
5	60	5	**1.9**
6	80	5	**1.6**
7	100	5	**1.6**
8	150	6	**1.2**
9	200	6	**1.1**
10	500	7	**0.8**
11	1000	7	**0.6**

Approximation Bounds: MNL Model

- **Suppose underlying model is MNL with parameters u_i for all i**

$$\rho_{ij} = \frac{u_j}{\sum_{\ell \in N_+} u_\ell - u_i} = \pi_j(N \setminus \{i\})$$

Theorem 1 *If the underlying model is MNL with parameters u_i for all $i = 0, \ldots, n$. Then for any offer set $S \subseteq [n]$ and $j \in S$,*

$$\hat{\pi}_j(S) = \frac{u_j}{\sum_{i \in S_+} u_i} = \pi_j(S)$$

Approximation Bounds: Other Models

- **McFadden and Train (1996)**
  - Every random utility based model can be approximated arbitrarily closely by a mixture of MNL

- **Suffices to prove approximation bounds for mixture models**

- **Consider a mixture of MNL model with K segments**
  - Probability of segment k is θ_k
  - Parameters for segment k: $u_0^k, \ldots, u_n^k$
  - Assume wlog. $u_0^k + \ldots + u_n^k = 1$

- **Choice Probability:**

$$\pi_j(S) = \sum_{k=1}^{K} \theta_k \frac{u_j^k}{\sum_{i \in S_+} u_i^k}$$

Approximation Bounds: MMNL Model

MMNL model (with K segments)

$$\rho_{ij} = \sum_{k=1}^{K} P(\text{ segment } k \mid \text{first choice is } i) \cdot \frac{u_j^k}{1 - u_i^k}$$

Theorem 2 *If the underlying model is mixture of MNL with K segments, probabilities θ_k of segment k, and MNL parameters $u_0^k, \ldots u_n^k$. For any offer set $S \subseteq [n]$, let $\alpha = \max_k u^k(\bar{S})$. Then for any $j \in S$,*

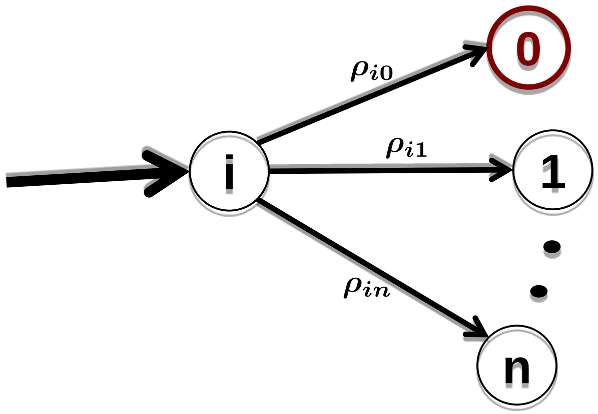
$$\hat{\pi}_j(S) \geq \pi_j(S)(1 - \alpha^2)$$

For α=0.5, we get a 0.75-approximation for choice probabilities

Assortment Optimization

- **Optimization Problem**

$$\max_{S \subseteq [n]} \sum_{j \in S} r_j \cdot \hat{\pi}_j(S)$$



r_i vs $\sum_{j \neq i} \rho_{ij} r_j$

$$g^0 = r$$

$$g_i^t = \max\left(r_i, \sum_{j \neq i} \rho_{ij} g_j^{t-1} \right)$$

Theorem 3 *Suppose $g = \lim_{t \to \infty} g^t$. Let*

$$S = \{ j \in [n] \mid g_j = r_j \}.$$

Then S is an optimal assortment with respect to choice model $\hat{\pi}$.

Conclusions

- **Choice model selection and assortment optimization problem**

- **Present Markov chain based universal choice model**
  - Simultaneous approximation for all random utility models (under mild assumptions)
  - Polynomial time assortment optimization

- **Future directions**
  - Additional constraints (eg. capacity) in assortment optimization

Questions?