Metropolis-Hasting for two-sided logit

Anh Le

April 9, 2016

1 Introduction

We introduce a model that can estimate the utility of both sides, building from their utility function up.

2 Utility

2.1 Officials' utility

Following Logan (1998), we consider the utility function of the two actors, the official and the firm.[1] For official j, the utility of having firm i investing in his country is:

$$U_j(i) = \boldsymbol{\beta}_j' X_i + \epsilon_{1ij} \tag{1}$$

where

β_j is a vector of official j's preference for relevant characteristics of firms
x_i is a vector of firm i's measured values on those characteristics
ϵ_{1ij} is the unobserved component that influences official j's utility

On the other hand, the utility of not having firm i investing is:

$$U_j(\neg i) = b_j + \epsilon_{0ij} \tag{2}$$

where

b_j is the baseline utility of official j without any firm investing
ϵ_{0ij} is the component that influences official j's utility

[1]For ease of exposition, in this section I will refer to country j and official j interchangeably.

For each firm i, official j will make an offer to invest if $U_j(i) > U_j(\neg i)$. Some relevant firm characteristics (i.e. X_i) that the official may consider are: technological intensity, jobs, and capital. The corresponding β's represent the official's preference for these characteristics.

Following the discrete choice literature, we model $\epsilon_{1ij}, \epsilon_{0ij}$ as having the Gumbel distribution. Then, the probability of official j making an offer to firm i takes the familiar binomial logit form:

$$
\begin{aligned}
Pr(o_{ij} = 1) &= Pr(U_j(i) > U_j(\neg i)) && (3)\\
&= Pr(\epsilon_{0ij} - \epsilon_{1ij} < \boldsymbol{\beta}_j' X_i - b_j) && (4)\\
&= \frac{\exp(\boldsymbol{\beta}_j' X_i)}{1 + \exp(\boldsymbol{\beta}_j' X_i)} && (5)
\end{aligned}
$$

where Equation (5) is due to the fact that the difference between two Gumbel-distributed random variables has a logistic distribution. We make the constant term b_j disappear into $\boldsymbol{\beta}_j$ by adding an intercept column to the matrix of firm characteristics X_i.

The opportunity set of firm i is the set of all countries that have made firm i an offer. If we know the preferences of all countries, we can calculate the probability that firm i gets an opportunity set O_i as follows:

$$
\begin{aligned}
p(O_i|\boldsymbol{\beta}) &= \prod_{j \in O_i} p(o_{ij} = 1|\boldsymbol{\beta}) \prod_{j \notin O_i} p(o_{ij} = 0|\boldsymbol{\beta}) && (6)\\
&= \prod_{j \in O_i} \frac{\exp(\boldsymbol{\beta_j'} X_i)}{1 + \exp(\boldsymbol{\beta_j'} X_i)} \prod_{j \notin O_i} \frac{\exp(\boldsymbol{\beta_j'} X_i)}{1 + \exp(\boldsymbol{\beta_j'} X_i)} && (7)
\end{aligned}
$$

In our observed data, since we only observe the final matching of firms and countries, this opportunity set is unobserved. This is the gist of the statistical challenge. TO overcome this issue, we will use Markov chain Monte Carlo to sample from and approximate $p(O_i|\boldsymbol{\beta})$.

2.2 Firms' utility

On the other side, for firm i, the utility of investing in country j is:

$$
V_i(j) = \alpha' W_j + v_{ij} \qquad (8)
$$

where

$$
\begin{aligned}
\alpha &\text{ is a vector of firms' preference for relevant characteristics of countries}\\
W_j &\text{ is a vector of country } j \text{ measured values on those characteristics}\\
v_{ij} &\text{ is the unobserved component that influences firm } i\text{'s utility}
\end{aligned}
$$

Firm i evaluates all the countries that make an offer and chooses the one that brings the highest utility. In our model, the relevant country characteristics are: labor quality, level of development, and market size. Since all firms are considered having homogeneous preferences, α does not have a subscript i. The model can be easily extended so that there is heterogeneous preference among firms.

If v_{ij} is modeled as having a Gumbel distribution, then the probability that firm i will accept the offer of official j out of all the offers in its opportunity set O_i is

$$p(A_i = a_i | O_i, \alpha_i) = \frac{\exp(\alpha' W_{a_i})}{\sum\limits_{j: j \in O_i} \exp(\alpha' W_j)} \tag{9}$$

3 Estimate the actors' preference

While Logan (1996, 1998) successfully reformulated standard discrete choice models to a two-sided setting, the estimation of the two-sided model remains challenging. The key difficulty lies in the fact that we do not know about the full sets of offers that firms receive from countries. Therefore, the likelihood function is incomplete, missing the data on firms' opportunity sets. With an incomplete likelihood function, we cannot use Maximum Likelihood Estimation to estimate firms' and countries' preferences.

To deal with this problem, Logan (1996) used the Expectation-Maximization (EM) algorithm, which iterates between two steps. First, given the current best guess of firms' and countries' preferences, pick values for the unobserved opportunity sets so that we maximize the likelihood. Second, given the current best guess of the unobserved opportunity sets, taken from step 1, pick values for firms' and countries' preferences so that we maximize the likelihood. By iterating between these two steps, the algorithm constantly searches for parameters values that maximize the likelihood.

However, an important downside of the EM algorithm is its lack of standard error. Therefore, while the algorithm is capable of producing the best estimate for the parameters of interest, it is difficult to know how good our best guess really is.

4 Update opportunity set

Target distribution for a firm i

$$p(O_i | A_i, \alpha, \boldsymbol{\beta}) = \frac{p(O_i, A_i, \alpha, \boldsymbol{\beta})}{p(A_i, \alpha, \boldsymbol{\beta})} \tag{10}$$

We propose a new O_i^* by randomly sample a new offer, j^* for each firm. If the new offer is not already in the current opportunity set, we add the offer to the set. If it already is in the opportunity set, we remove it from the set.

We then calculate the Metropolis-Hasting acceptance ratio:

$$MH_O = \frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{p(O_i^*,A_i,\alpha,\boldsymbol{\beta})}{p(A_i,\alpha,\boldsymbol{\beta})} \times \frac{p(A_i,\alpha,\boldsymbol{\beta})}{p(O_i,A_i,\alpha,\boldsymbol{\beta})} \tag{11}$$

$$= \frac{p(O_i^*,A_i,\alpha,\boldsymbol{\beta})}{p(O_i,A_i,\alpha,\boldsymbol{\beta})} \tag{12}$$

$$= \frac{p(A_i|O_i^*,\alpha)p(O_i^*|\boldsymbol{\beta})}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})} \tag{13}$$

$$\tag{14}$$

where the factorization of the likelihood in Equation (13) is due to the fact that the acceptance of firm i only depends on what is offered to it and what is its preference, $p(A_i|O_i^*,\alpha)$; what is offered to i depends on the preferences of all countries, $p(O_i^*\ \boldsymbol{\beta})$.

If we plug in Equation (9) and Equation (7)

$$\frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j) + \exp(\alpha' W_{j^*})} \times \exp(\boldsymbol{\beta}'_{j^*} X_i) \tag{15}$$

where j^* is the index of the newly sampled job. This is the case when the newly proposed job is not already offered, so it's added to the opportunity set.

When the newly proposed job is already offered, so it's removed from the opportunity set, we have

$$\frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j) - \exp(\alpha' W_{j^*})} \times -\exp(\boldsymbol{\beta}'_{j^*} X_i) \tag{16}$$

5 Update firms' parameters, α

Target distribution:

$$p(\alpha|A,O,\boldsymbol{\beta}) = \frac{p(O,A,\alpha,\boldsymbol{\beta})}{p(A,O,\boldsymbol{\beta})} \tag{17}$$

We propose a new α^* using a symmetric proposal distribution that sample α^* in a box whose boundary is $\alpha^* \pm \epsilon_\alpha$

Metropolis-Hasting acceptance ratio:

$$MH_\alpha = \frac{p(\alpha^*|A,O,\boldsymbol{\beta})}{p(\alpha|A,O,\boldsymbol{\beta})} = \frac{p(A_i|O_i,\alpha^*)p(O_i|\boldsymbol{\beta})}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})} \tag{18}$$

$$= \frac{p(A_i|O_i,\alpha^*)}{p(A_i|O_i,\alpha)} \tag{19}$$

where Equation (19) is due to the flat prior (so $\frac{p(\alpha^*)}{p(\alpha)} = 1$) and the symmetric proposal distribution (so $\frac{p(\alpha^*|\alpha)}{p(\alpha|\alpha^*)} = 1$)

If we plug in Equation (9),

$$MH_\alpha = \prod_i \left[\frac{\exp(\alpha^{*\prime} W_{a_i})}{\exp(\alpha' W_{a_i})} \times \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)} \right] \tag{20}$$

$$= \prod_i \left[\exp(\epsilon'_\alpha W_{a_i}) \times \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)} \right] \tag{21}$$

Finally, we log transform the MH acceptance ratio for numerical stability.

$$\log MH_\alpha = \sum_i \left[\epsilon'_\alpha W_{a_i} + \log\left(\sum_{j:j\in O_i} \exp(\alpha' W_j) \right) - \log\left(\sum_{j:j\in O_i} \exp(\alpha^{*\prime} W_j) \right) \right] \tag{22}$$

6 Update countries' parameters, β

Target distribution:

$$p(\boldsymbol{\beta}|A, O, \alpha) = \frac{p(O, A, \alpha, \boldsymbol{\beta})}{p(A, O, \alpha)} \tag{23}$$

We propose a new $\boldsymbol{\beta}^*$ using a symmetric proposal distribution that sample $\boldsymbol{\beta}^*$ in a box with side length ϵ_β

Metropolis-Hasting acceptance ratio:

$$MH_\beta = \frac{p(\beta^*|A, O, \alpha)}{p(\beta|A, O, \alpha)} = \frac{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta}^*)}{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta})} \tag{24}$$

$$= \frac{p(O_i|\boldsymbol{\beta}^*)}{p(O_i|\boldsymbol{\beta})} \tag{25}$$

where Equation (24) is due to the flat prior on β and the symmetric proposal distribution.

We plug in Equation (7),

$$MH_\beta = \prod_i \left[\prod_{j\in O_j} \frac{\exp(\beta_j^{*\prime} X_i)}{\exp(\beta'_j X_i)} \times \prod_j \frac{1 + \exp(\beta_j^{*\prime} X_i)}{1 + \exp(\beta'_j X_i)} \right] \tag{26}$$

$$\log MH_\beta = \sum_i \left[\sum_{j\in O_i} \beta_j^{*\prime} X_i - \beta'_j X_i + \sum_j \log(1 + \exp(\beta_j^{*\prime} X_i)) - \log(1 + \exp(\beta'_j X_i)) \right] \tag{27}$$

In the MCMC implementation, since $\boldsymbol{\beta}$ is high dimensional, we conduct multiple block Metropolis Hastings, updating several β's at one time.

References

Logan, J. A. (1996). Opportunity and Choice in Socially Structured Labor Markets. *American Journal of Sociology 102*(1), 114.

Logan, J. A. (1998). Estimating Two-Sided Logit Models. *Sociological methodology 28*(1), 139–173.