IS604_Final_Writeup

Charley Ferrari, Paul Garaud, Dieudonne Ouedraogo

December 12, 2015

Introduction

Markov Chains have a long history of being used as theoretical models for studying human mobility. Pan and Nagurney (Pan, Nagurney 1994) provide a useful survey of past attempts at using Markov Chains, and advance the thinking by allowing for the possibility of non-homogenous markov chains. In their model and review of the literature, geographical locations are seen as states of the markov chain. Previous attempts at modeling population using markov chains involved a single transition matrix that could be calculated. Pan and Nagurney, by introducing costs, introduce the possibility of a nonhomogenous markov chain that still tends towards equilibrium.

Pan and Nagurney's models can give insight into the dynamics of population mobility, but are limited by their deterministic nature. Costs can influence the time it takes to reach equilibrium and represent imperfect information or time lags associated with moving decisions in the real world, but once equilibrium is reached population is assumed to remain static. The authors point out that equilibrium does not necessarily mean there is no more migration, but simply that the probabilities of moving between locations have equalized, and net migration remains at zero.

In this paper we provide an alternative way of looking at a population model using markov chain monte carlo. By introducing randomness into the model, we hope to gain some insight into how exactly our population will converge to equilibrium, and what sort of variation will occur once equilibrium is acheived. In this way, we hope to model both migration induced by comparative utilities across regions, and movement due to random factors.

The choice of markov chain monte carlo led to some necessary departures from Pan and Nagurney's model. Each iteration of Pan and Nagurney's markov chain represents an equilibrium between comparative utilities and the cost of moving, and this equilibrium is reached by considering a flow of migrants. The iterations in our model do not represent equilibrium, but are rather defined steps towards an equilibrium population distribution. Each step represents the possibility of one person moving. Because of this, cost is not dependent on a flow of people, but is rather assumed to be a static, dependent only on the source and destination.

Pan and Nagurney look at network equilibrium models as one-step models. They arrive at equilibrium, but do not allow one to view chain migration nor the dynamics of migration. By using markov chain monte carlo, introducing randomness allows us another way to view these features of a migration model.

Literature Review

Our model is based directly off of the Pan and Nagurney model, repurposed for use with markov chain monte carlo.

The authors define utility functions, u_i for each location i, which are dependent on the populations of each region. Costs are a function of the flow between two locations, so each bilateral pair i and j has its own cost c_{ij}.

Starting from an initial population p_0, the flow between locations i and j can be discovered by solving the below equation:

$$u_i(p_i^1) + c_{ij}(f_{ij}^1) = u_j(p_j^1)$$

For each i and j.

The two region example shown in Pan and Nagurney's paper assumed utility functions of $u_1(p) = -p_1 + 8$ and $u_2(p) = -p_2 + 14$. Costs were simply based on flow, so $c_{12}(f_{12}) = 2f_{12}$ and $c_{21}(f_{21}) = 2f_{21}$.

The solution to this toy model is explained and diagrammed below:

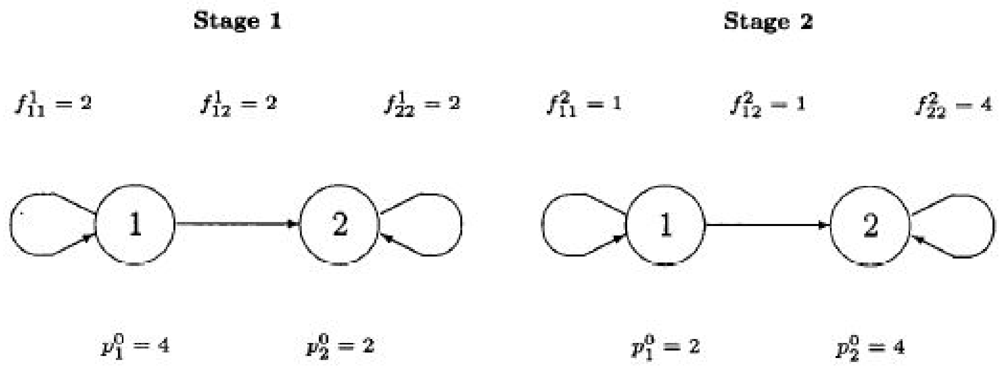

Figure 1. A transition map of the migration process.

See **Figure 1** where a transition map is used to describe the migration process, with two associated transition matrices A_1 and A_2 given as follows:

$$A_1 = \begin{bmatrix} 2/4 & 2/4 \\ 0 & 2/2 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 4/4 \end{bmatrix}.$$

In this case, the transition matrix is the same (and the derivation shows in what way they will differ.)

Other attempts at explaining migration using markov chains include A. Constant's and K. Zimmerman's paper "The Dynamics of Repeat Migration: A Markov Chain Analysis." (Constant, Zimmerman 2012.) Using a Markov Chain approach to model repeat migration, this paper aims to describe the dynamics of repeated exit and entry between two countries. The paper estimates the transition probabilities using the formula $P(E_{t+1} = i|E_t = j) = \frac{e^{\beta} i j^{x} m t}{\sum e^{\beta} i k^{x} m t}$ using empirical data from Germany. These irepresent the coefficients on the personal characteristics variables for transition state i (note that only two transition state coefficient vectors are calculated as the other two are complements and calculated by taking 1 - P) and are estimated using logistic regression.

The paper then converts micro level longitudinal migration data at the individual level to person-years, which are treated as separate observations, to estimate these conditional probabilities. The end result are empirically-derived formulae that describe the personal characteristics that influence the decision to (repeat) migrate.

This approach is an interesting twist on the traditional Markov Chain analysis by detailing a particular case and integrating real world information about migrants. This could be useful for future considerations of our model as a guide to link real world data to our currently theoretical utility functions.