Developments in Statistics and Methodology
A. Ferligoj and A. Kramberger (Editors)
Metodološki zvezki, 9, Ljubljana: FDV 1993

Applications of Interactive Markov Models in the Dynamics of Social Systems

Vesna Omladič[1]

Abstract

In a Markov model a large population of individuals is distributed over a finite number of states. If the transition matrix is not constant and depends on the actual distribution of the individuals, the model is called interactive. In the paper equilibrium points of this kind of models are studied as well as stability of these points. Interactive Markov models have a wide choice of applications in the studies of dynamics of social systems, e.g. consumers' buying behaviour, families' behaviour in choosing their physician, voting behaviour, geographical, social, and occupational mobility. In particular, this kind of models have recently been gaining more attention in the studies of social mobility in fast changing societies. A different approach to the study of interactive voting behaviour of small groups is also presented.

Keywords: Stochastic processes; Differential growth rate; Social mobility; Interaction within small groups.

1 Introduction

Markov stochastic processes have wide application in the studies of dynamics of social systems. In these models a population of individuals is considered classified into k mutually exclusive and exhaustive *categories*, sometimes also called *classes* or *states*. The probabilities p_{ij} that an individual moves from state i to state j between two consecutive points of time are sometimes called *transition probabilities* and form a matrix, called *transition matrix*. The simplest possible model of the kind is the homogeneous stationary Markov model, where it is assumed that all the individuals are moving independently of each other and independently of their previous "career paths", and that all these probabilities are the same for all individuals and constant over time. As any individual must either stay within the same state or move to another, the transition matrix is *stochastic*, meaning that for all i it holds that $\sum_{j=1}^{k} p_{ij} = 1$. Such Markov models are sometimes called *probabilistic* or *stochastic* as opposed to the *deterministic model* in which it is assumed that p_{ij} are

[1]Faculty of Social Sciences, University of Ljubljana, Kardeljeva pl. 5, 61 109 Ljubljana, Slovenia

not probabilities, but, relative frequencies, i.e. proportions of individuals from state i moving in a period of time to state j.

In this paper we will consider a more general model derived from the Markov model. Namely, the transition probabilities (or rates) will be assumed to depend on the size of the state towards which the individuals are moving. This approach takes into account the fact that individuals are following not only their own choices, but, are also influenced by what others are doing. This kind of models have been widely studied and for the sake of simplicity, linear dependence has usually been assumed. The main applications of this model are in the studies of social mobility, where the population is stratified according to some socio-economic or cultural attribute. The usual Markov model assumptions described in the previous paragraph may suffice for the studies of mobility phenomena in stable countries with long-lasting slow-changing economic political and social system. On the other hand, for the fast developing countries of the third world as well as lately for the fast changing societies of the eastern Europe, such models may not be appropriate.

From the wide range of applications of these models let us mention only a few. In the study of voting behaviour, the categories may be chosen to be political parties, time points elections, and "individuals in state i" the voters supporting party i. When studying geographical mobility, the categories may be towns or regions, and the individuals may be families living there over some specified time period. Buying behaviour of a certain population may be studied by classifying them according to the brands of a certain product such as soap or cheese. The time period here may be two consecutive purchases of the product. One of the most frequent applications is in the social or occupational mobility. In these cases the population is stratified in terms of some socio-economic or cultural attribute such as income, education, religion, etc. A closely related and also very frequent application arises in manpower planning where employees of a certain institution are classified according to their position, age, length of service, etc.

In section 2 we present a special type of the so-called interactive Markov models. It is based on a similar model due to Conlisk (1976,1978) which he refers to as *AB*-model. Some apparently new results on the model are given in this section, while section 3 is devoted to illustrative examples of applications. Kulkarni and Kumar (1986,1989) recently developed two models based on similar ideas, but assuming ordinality of the attribute and upwards-only mobility. They also assume different growth rates for populations in different categories. A slightly more general model is presented in our section 4 together with some possibly new results on these models. Section 5 is devoted to an overall different approach to the study of interactive influence. This approach is illustrated on a study of voting behaviour of a population grouped in a large number of small closely connected sub populations.

2 Interactive Markov models

Here is a model, similar to what was called an *AB*-model in Conlisk (1976). We assume that A and B are real $k \times k$ matrices having zeros on the main diagonal. We also suppose that the entries of A are non negative and that its row sums are not greater than one. Let us denote by $q_i(n)$ the proportion of the population under consideration that is in state i at a point in time n. Here, we tacitly assume that the proportions $q_i(n)$ are all non negative and that their sum equals one, so that the k-vector $(q_i(n))_i$ forms a "probability" distribution. Furthermore, let the interactive transition proportions $p_{ij}(n)$ between time points n and $n+1$ be given by

$$p_{ij}(n) = \begin{cases} a_{ij} + b_{ij}q_j(n) & \text{if } i \neq j \\ 1 - \sum_{l\neq i}[a_{il} + b_{il}q_l(n)] & \text{if } i = j \end{cases}. \tag{1}$$

The sequence of proportional distributions then satisfy the following recurrence relations

$$q_j(n+1) = \sum_{i=1}^{k} q_i(n)p_{ij}(n). \tag{2}$$

In order to simplify our approach we shall rewrite these two equations using matrix notation. From now on let us view all the vectors as columns and let us denote by ^{T} the transpose of vectors and matrices. Thus, x^T will stand for the row vector corresponding to the column vector x. Moreover, let us introduce the notation $\mathrm{Diag}\, x$ for the diagonal matrix having components of the vector $x = (x_i)$ on the main diagonal and zeros everywhere else. We will be using throughout the notation $A \geq 0$, respectively $x \geq 0$, meaning that all the components of matrix A, respectively vector x, are non negative. Then, the interactive transition matrix $P(n) = (p_{ij}(n))$, defined by (1) can be written as

$$P(n) = A + B\mathrm{Diag}[q(n)] + I - \mathrm{Diag}[Ae + Bq(n)], \tag{3}$$

where we have denoted by $q(n)$ the vector corresponding to the n-th distribution of proportions $(q_i(n))$ and by e the vector with all components equal to 1. Let us introduce the matrix of "intrinsic" transition proportions $P_0 = A + I - \mathrm{Diag}[Ae]$, in order to simplify this expression even further

$$P(n) = P_0 + B\mathrm{Diag}[q(n)] - \mathrm{Diag}[Bq(n)]. \tag{4}$$

Observe that for any vector x it holds that $\mathrm{Diag}(x)e = x$ to see that $P_0e = Ae + e - \mathrm{Diag}(Ae)e = e$ and therefore, $P(n)e = e + B\mathrm{Diag}[q(n)]e - \mathrm{Diag}[Bq(n)]e = e$. This shows that P_0 is stochastic and that $P(n)$ is stochastic as soon as its entries are non negative. It is easy to get sufficient conditions on matrices A and B to ensure that the matrices $P(n)$ are all stochastic, independently of the choice of distributions $q(n)$. Let X^+ denote the positive part of a matrix X, i.e. the matrix, having the same positive entries as X and zeros everywhere else. We can define the negative part X^- of X similarly, or simply by putting $X^- = X^+ - X$. Using this notation, it is easy to see that $P(n) \geq P_0 - B^- - \mathrm{Diag}(B^+e)$. Hence, the condition

$$P_0 - B^- \geq \mathrm{Diag}(B^+e) \tag{5}$$

forces $P(n)$ to be stochastic automatically for all n. In other words, for those indices i and j for which $b_{ij} < 0$ we must have $a_{ij} + b_{ij} \geq 0$, and, denoting $\beta_i = \sum\{b_{ij} | \text{all}\, j \text{ such that } b_{ij} > 0\}$, it must hold that $\beta_i + \sum_j a_{ij} \leq 1$ for all i.

Using definition (4) for the matrix $P(n)$, we can also rewrite the recurrence relation (2) into

$$q(n+1)^T = q(n)^T[P_0 + B\mathrm{Diag}(q(n)) - \mathrm{Diag}(Bq(n))]. \tag{6}$$

Observe that for stochastic matrices $P(n)$, this relation sends non negative columns $q(n)$ with sum one into non negative columns $q(n+1)$ of sum one, i.e. it sends distributions of proportions into distributions of proportions.

The main problem usually considered with the kind of interactive models is to find critical points of equation (6), sometimes also called stationary or equilibrium points. By definition, a column q of proportions (i.e. a non negative column with sum one) is said to be a *critical point* of this model if $q(n) = q$ forces $q(n+1) = q$ by relation (6). In other words, q has to be a solution of the system of quadratic equations

$$q^T = q^T[P_0 + B\mathrm{Diag}q - \mathrm{Diag}(Bq)]\,. \tag{7}$$

This system has always a solution of the desired kind by the Brouwer's fixed point theorem, however, the solution is not necessarily unique. Another interesting question is, whether a critical point q is *stable*, this means that there is a neighbourhood of q such that as soon as a population reaches a distribution $q(n)$ at a certain point of time n within this neighbourhood, then, the sequence of distributions converges to critical point q.

The following theorem is based on a simple observation:

Theorem 1 *Let* 1 *be the only eigenvalue of* P_0 *of modulus* 1 *and let its algebraic multiplicity be* 1. *If* B *is symmetric and satisfies condition (5), then (7) has a unique non negative solution* q *with sum* 1 *which is the only critical point of model (6). Furthermore, this point is stable and for any starting distribution of proportions* $q(0)$ *the sequence* $q(n)$ *converges to* q.

This theorem is an extension of a result of Conlisk (1978). He proves a similar result under more restrictive assumption that P_0 is *primitive*, i.e. that a power of P_0 has all the components strictly positive. Note that any primitive stochastic matrix P_0 satisfies the conditions of the theorem by Perron-Frobenius theorem (Schaeffer 1974, Theorem I.6.5). However, the condition is weaker since there are many examples of non primitive matrices which still satisfy these conditions. To get a simple example, choose p, $0 < p < 1$, and write

$$P_0 = \begin{bmatrix} p & 0 & \cdots & 0 & 1-p \\ 0 & p & \cdots & 0 & 1-p \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p & 1-p \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

It is clear that P_0 and all of its powers are upper triangular so that no one can have all the components strictly positive. However a simple computation reveals that P_0 satisfies the conditions of our theorem.

Let us show the validity of theorem 1. For any two vectors x and y it holds that $x^T \text{Diag} y = y^T \text{Diag} x$. Therefore, $q^T \text{Diag}(Bq) = (Bq)^T \text{Diag} q = q^T B^T \text{Diag} q$ which implies that equation (6) can be rewritten into

$$q(n+1)^T = q(n)^T[P_0 + (B - B^T)\text{Diag}(q(n))], \tag{8}$$

and, similarly, equation (7) into

$$q^T = q^T[P_0 + (B - B^T)\text{Diag}(q)]. \tag{9}$$

By the assumption of the theorem, matrix B is symmetric which means that $B^T = B$. Hence, $q(n+1)^T = q(n)^T P_0$ and a simple inductive argument yields $q(n)^T = q(0)^T P_0^n$. Denote by $\lambda_1, \lambda_2, \ldots, \lambda_m$ the distinct eigenvalues of P_0, where we may suppose by the assumptions of the theorem that $\lambda_1 = 1$ and that $|\lambda_i| < 1$ for $i > 2$. Using standard linear algebra techniques such as Jordan form, say, P_0 can be written in the form $P_0 = E_1 + E_2 + \cdots + E_m$, where E_i's are the so-called spectral parts of P_0, i.e. they are all commuting with P_0, the product of any two of them is zero, and each E_i has λ_i as the only non-zero eigenvalue. This implies by a simple inductive argument again that $P_0^n = E_1^n + E_2^n + \cdots + E_m^n$ holds for all indices n. Now, as all λ_j for $j > 2$ have absolute value strictly smaller than one, the corresponding E_j^n must converge to zero with increasing n. On the other hand, the algebraic multiplicity of eigenvalue $\lambda_1 = 1$ of matrix P_0 is one which forces E_1 to be a projection (i.e. an idempotent) and, consequently, we must have that $E_1^n = E_1$ for all n. It thus follows that P_0^n converges with increasing n to E_1 and therefore, $q(n)^T$ converges to $q(0)^T E_1$. Using again some standard linear algebra, the form of projection E_1 is xy^T, where x and y^T are the respective right and left eigenvectors of P_0 corresponding to eigenvalue 1 normalized so that $y^T x = 1$. Recalling that P_0 is a stochastic matrix, we can take $x = e$ for the right eigenvector and normalize y^T accordingly to get $q(0)^T E_1 = [q(0)^T e]y^T = y^T$. We have thus seen that for any starting distribution $q(0)^T$, the sequence of distributions $q(n)^T$ converges to the limit distribution y^T. This is therefore the only critical point which is also stable.

A somewhat deeper argument shows that practically the same is still valid for an "almost symmetric" matrix B. The following theorem presents the details:

Theorem 2 *If P_0 is as in theorem 1, B satisfies condition (5), and $B - B^T$ is small enough, then there exists a unique stable critical point of model (6). Furthermore, for any initial distribution $q(0)$ the sequence of distributions $q(n)$ converges to this critical point.*

The exact formulation and the proof of this theorem go beyond the scope of this work and will be published elsewhere.

3 Some applications of interactive models

In this section we will give a number of illustrative examples from the studies of dynamics of social systems for which the interactive Markov model, as presented in the previous section, may be applied. These models should, of course, not be taken for granted. Whether and how much does this model fit the particular situation is a matter of serious empirical and statistical testing. Since the kind of verifying would go beyond the framework of this paper we will not go into it.

First, let us consider a population of consumers and give a possible model for studying their buying behaviour with respect to a certain fixed product. Assume that there are k brands of this product on the market and let the time period be the time between two successive purchases of the product under consideration. Let us assume that there is an "intrinsic" probability a_{ij} that a consumer buying brand i would buy brand j on his next purchase. This probability is based on the satisfaction of the consumer with the i-th brand, and on the advertisement, price, availability, and design of the j-th and other brands, and on other facts which are independent of the actual distribution of the consumers over the brands. However, the other consumers may have some influence on his choice as well. Namely, another consumer, already using brand j, may be recommending that product, if he is satisfied with it, or, he may by trying to turn other consumers away from it, if he is not. The influence of the consumers of the j-th brand on the next choice of those buying the i-th brand will grow with their size: the bigger the group, the bigger their influence. Let us denote by q_j the proportion of consumers buying the j-th brand on a certain purchase. We are assuming for simplicity that the transition probability depends linearly on q_j. Thus, a term $b_{ij}q_j$ has to be added to the "intrinsic" probability a_{ij} for the transition from the i-th to the j-th brand. Here, the coefficient b_{ij} is positive if the consumers of the j-th brand have an attracting influence on the consumers of the i-th brand on the average, and it is negative in the opposite case.

Here is a simple numerical example to illustrate this application. Let the consumers be faced with three brands of a product. The first two brands are comparable and for them the probability that a consumer will stay with the brand is twice as high as the probability that it will change to the other. The third brand has appeared recently on the market and due to its low price, attractive design, and noticeable advertisements, a consumer of the first two brands decides to try it with probability $\frac{1}{2}$. However, those who have already tried it are disappointed and will buy on their next purchase one of the old brands giving both of them an equal chance. The "intrinsic" model is described by the transition matrix

$$P_0 = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

which has the left normalized eigenvector equal to $q^T = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ which means that in a long run of selling under the same conditions the three brands will split the market in thirds. Let us now adjust the model for the interaction of unsatisfied buyers of the third brand. Assume that the coefficients b_{13} and b_{23} are both equal to $-\frac{1}{2}$ which is the strongest possible negative influence which is still ensuring stochasticity of the matrices $P(n)$ by (5). We thus have

$$B = \begin{bmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B - B^T = \begin{bmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.$$

A simple computation shows that in this case equation (9) has a unique solution, which can also be computed. We get $q^T = (\frac{1}{2}(\sqrt{3}-1), \frac{1}{2}(\sqrt{3}-1), 2-\sqrt{3}) = (0.366, 0.366, 0.268)$ which means that the third company obtains in a long run a somewhat smaller market share (by 6.5 %) in this interactive model as compared to the one, where no interactions between consumers were assumed.

Here is another possible application of interactive Markov models. Assume the population of all the families of a city choosing their family physician. There are k doctors available in the city and the families are divided into k classes with respect to their choice among these doctors. The time period of one observation might be a year. There is again an intrinsic transition probability a_{ij} for the family signed up with the i-th physician during a period, to become dissatisfied with him and select j-th one for the next period of time. Again, b_{ij} will denote the corresponding coefficient of interaction. Thus, $b_{ij}q_j$ denotes the "aping" part of this transition probability.

To illustrate this example, let us assume a small town, where there are only three doctors of approximately the same knowledge and capability. A family, signed up with a physician in a period, will stay with him for the next period as well with probability $\frac{1}{2}$ and will give an equal chance to the other two otherwise. Besides that, an overcrowdedness in the waiting room of any of the three doctors turns away new patients with intensity $-\frac{1}{4}$ towards any of the other two. Therefore, we have

$$P_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & 0 & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & 0 \end{bmatrix}.$$

It is easy to see that P_0 has the left Perron eigen vector equal to $q^T = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ and this must also be the limiting distribution of this model without interactions. However, when we are taking the interaction into consideration as well, we observe that $B^T = B$ and by the discussion following theorem 1 in the previous section, we observe, that in this case, matrix B, although it does have an influence on the transition matrix $P(n)$, it has no influence on the sequence $q(n)$. Therefore, by theorem 1 we have a unique equilibrium point again and by discussion this point equals again the point $q^T = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$.

The phenomenon used in the proof of theorem 1 and illustrated by this example may need some more comment. We have seen that a symmetric matrix of "aping trends" has no influence on the actual sequence of distributions of the population. We could describe the reason why this happened very simply. Namely, if j-th class of population is attracting the i-th one with intensity b_{ij}, while the i-th one is attracting the j-th one with the same intensity $b_{ji} = b_{ij}$, then, this attraction is causing more actual individuals of the population to be moving from the i-th to the j-th class as well as the other way around. However, in the average, the "aping" effects are in this case cancelling each other and the next size of either i-th or j-th class is the same as if there were no effects of the kind at all.

In the third example consider a population of people living in a small town commuting every day to the near-by bigger city to work. They have k modes of travel to choose among. Here, the transition probabilities are given, similarly as in the previous two examples, by $a_{ij} + b_{ij}q_j(n)$. For a concrete numeric example assume $k = 3$ and let the three possible choices for commuting be train, bus, and car. The intrinsic transition probabilities are given by

$$P_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}.$$

We have assumed here that non-car commuters rarely change to car. However, a car commuter finds it hard to decide for some other way of travel, but if he does, he prefers the train. The stationary distribution of this non-interactive model is $q^T = (\frac{11}{24}, \frac{7}{24}, \frac{1}{4}) = (0.458, 0.292, 0.250)$. Now, let us introduce some interaction into this model. Assume that an increasing amount of car commuters causes overcrowdedness of the roads which is discouraging train commuters from changing to car travel with intensity $-\frac{1}{6}$. On the other hand, when the proportion of train travellers is increasing, the railway company may offer a better and more attractive train service for a more affordable price. This is causing car commuters to be attracted to the train with intensity $\frac{1}{6}$. The matrix of the "aping trends" is therefore given by

$$B = \begin{bmatrix} 0 & 0 & -\frac{1}{6} \\ 0 & 0 & 0 \\ \frac{1}{6} & 0 & 0 \end{bmatrix},$$

and because it is skew-symmetric, we necessarily have $B - B^T = 2B$. A simple computation shows that the critical point is again unique and is given by $q^T = (\frac{1}{2}, \frac{3}{10}, \frac{1}{5}) = (0.5, 0.3, 0.2)$. A comparison with the non-interactive stationary point shows that taking into account the above described interaction in the long run causes 5 % of population that belongs to car commuters to change their mode of travel to work. Most of them are taking the train, but some of them decide for the bus.

Very many other applications in the studies of dynamics of social systems could have been proposed here: various socio-cultural and occupational mobilities, regional migration, political and other affiliations, etc. In some of these the above described model may turn out to be useful. However, there are many other ways of taking into account interactive effects. In the next section we will describe an improved model of the kind, while in section 5 a completely different approach will be presented.

4 Models with differential growth rates

In this section we will study a model proposed by Kulkarni, Kumar (1986,1989). In those papers, an interactive model is introduced, similar to the one presented in section 2, in connection with some applications to social mobility. The main difference between the two models is the possibility that different states may have different growth rates. The kind of models may have some advantages in applications over the ones presented in section 2 especially in studies of dynamics of social systems in rapidly changing societies. Just for the flavour of possible applications assume a study of buying behaviour of a new product, which has been introduced lately through a number of brands. The consumers are being recruited from the population of people who have not been using this product before. But, different brands may have substantially different success in attracting new customers to the product, so that the population of buyers of a certain brand may grow at substantially different rate as the group of buyers of some other brand. Let us improve the above presented model in such a way that it will include the interactive model of Kulkarni and Kumar (1989). As we shall see the improved model will differ only slightly from the already presented one and similar methods may apply.

Let again $p_{ij}(n)$ denote the interactive transition proportions and let they be defined by (1). But, now, let ν_i denote the growth rate of the i-th class and let us improve the recurrence relation (2) into $\sum_i \nu_i q_i(n) p_{ij}(n)$. However, the so obtained k-vector is not a distribution of proportions any more. If $P(n)$ satisfies condition (5) it is still positive, but, may not sum to 1 anymore. Actually, the sum of its components equals $\sum_i \nu_i q_i(n)$. Therefore, in order to get a distribution, we have to normalize it by dividing it by this sum. In this way we get the improved relation (2)

$$q_j(n+1) = \frac{\sum_{i=1}^{k} \nu_i q_i(n) p_{ij}(n)}{\sum_{i=1}^{k} \nu_i q_i(n)}. \tag{10}$$

To rewrite this main equation of the model in matrix notation we have to introduce, in addition to what has already been introduced in section 2, the matrix of growth rates $N = \mathrm{Diag}\nu_i$. Thus, $P(n)$ can be given by (3) and (4), while $q(n)$ has to follow the recurrence relation

$$q(n+1)^T = [q(n)^T N e]^{-1} q(n)^T N P(n), \tag{11}$$

or, if we expand, using (3):

$$q(n+1)^T = [q(n)^T N e]^{-1} q(n)^T N \Big[I + A + B\mathrm{Diag}(q(n)) - \mathrm{Diag}[Ae + Bq(n)]\Big]. \tag{12}$$

We are again interested in similar questions as in section 2.

Let us first give some sufficient conditions for existence and stability of critical points of the model, given by (12). The following two theorems are based on similar observations as theorems 1 and 2. Namely, we can again rearrange equation (12) to obtain

$$q(n+1)^T = [q(n)^T N e]^{-1} q(n)^T \Big[N P_0 + (NB - B^T N)\mathrm{Diag}(q(n))\Big]. \tag{13}$$

Using this relation we can prove the following two theorems:

Theorem 3 *Let the Perron-Frobenius eigenvalue of the matrix NP_0 be the only one of that absolute value and let it have algebraic multiplicity* 1 *as well as a strictly positive right eigenvector. If NB is symmetric and B satisfies condition (5), then (12) has a unique critical point q. Furthermore, this point is stable and for any starting distribution of proportions $q(0)$ the sequence $q(n)$ converges to q.*

Theorem 4 *If P_0 is as in theorem 3, B satisfies condition (5), and $NB - B^T N$ is small enough, then there exists a unique stable critical point of model (12). Furthermore, for any initial distribution $q(0)$ the sequence of distributions $q(n)$ converges to this critical point.*

5 Interaction within small groups

In this section we will consider a substantially different approach to interaction which may be useful in situations where members of the population are not all in similar contacts with each other, but the interaction is much more intensive within some smaller groups of population than it is between these groups. In situations where this is true for only a part of the population, we may use these methods for this part and some other for the rest of population, of course.

The situation usually assumed in these models will be described in terms of a specific example, namely, the study of voting behaviour, however, these methods may be useful for some other applications, such as consumers' buying behaviour. The assumption is that the voting population (or a part of it) splits into a large number of small groups of individuals having substantial impact on each other, while communications between individuals of different groups are minimal and their influence is negligible as compared to the mutual influence within the group. Examples of this kind of situation are easy to find. Groups might be workers employed in small plants or shops or maybe inhabitants of small villages. The model presented here is a discrete time modification of the continuous time model given in Coleman (1964).

To develop the model let us first assume that there is no interaction within a group of size k. We are interested in their perspective vote with respect to a fixed candidate A. Let us split the campaign time into smaller time units and assume that the "intrinsic" probability that an individual decided not to vote for candidate A at a certain point of time will change his mind during one period of time is α. Similarly assume that an individual decided to vote for candidate A will change his mind between two consecutive points of time is β. The states of the Markov chain are the total number of individuals decided to vote for candidate A. We will compute the transition matrix of this chain as well as its stationary distribution.

In order to compute the transition matrix, let us assume that candidate A has at a certain time point i out of k potential votes in the group and let us see what are his chances on the next point of time, concretely, let us compute the probability that he will then have j potential votes. This can happen in a number of disjoint ways, namely, he can loose $i - r$ votes and gain $j - r$ votes, and, since r cannot exceed neither i nor j and must not be negative, while $j - r$ must not exceed $k - i$, we have that r goes from $\max\{0, i + j - k\}$ to $\min\{i, j\}$. For any particular r the lost votes may be chosen in $\binom{i}{r}$ ways and the gained ones in $\binom{k-i}{j-r}$ ways, while the probability of any of these ways equals $\alpha^{j-r}(1-\alpha)^{k+r-i-j}\beta^{i-r}(1-\beta)^r$. Therefore, the transition probability equals

$$p_{ij} = \sum_{r=\max\{0,i+j-k\}}^{\min\{i,j\}} \binom{k-i}{j-r}\binom{i}{r}\alpha^{j-r}(1-\alpha)^{k+r-i-j}\beta^{i-r}(1-\beta)^r. \tag{15}$$

A straightforward computation shows that this matrix is stochastic. It is a matter of simple observation that it is also primitive and must have a left strictly positive eigenvector which must represent the stationary distribution of this chain, if properly normed. It turns out that this distribution is given by

$$q_i = \binom{k}{i}\frac{\alpha^i\beta^{k-i}}{(\alpha+\beta)^k},$$

which can be checked out by a not so long computation. Note that this is a binomial distribution for parameter $\frac{\alpha}{\alpha+\beta}$.

Let us now generalize this model for the case of mutual interaction. The influence of one person on a chosen person will be measured by a parameter γ. Thus, if a person is not voting for candidate A and if there are i individuals voting for this candidate and $k-i-1$ people besides the considered one not voting for him, then the probability that the observed person will change his or her mind during one period of time is $(\alpha+i\gamma)/(1+(k-1)\gamma)$ and the probability that he or she will stay with the same decision is $(\alpha+(k-i-1)\gamma)/(1+(k-1)\gamma)$. Similarly, if somebody believes to vote for candidate A together with $i-1$ other individuals, while $k-i$ are decided otherwise, the probability that this person will change his or her mind is $(\beta+(k-i)\gamma)/(1+(k-1)\gamma)$ and the probability that the individual will stick with the decision for the next period is $(\alpha+(i-1)\gamma)/(1+(k-1)\gamma)$.

When computing the transition probabilities we have to take into account the fact that the decisions of individuals are not being taken independently. To illustrate the complication, let us compute the probability in a three-person group that no-one will be willing to vote for candidate A at the end of the period, provided that nobody has been at the beginning. This probability is equal to the one that no one of the three voters are changing their minds. One voter is not changing his or her mind with probability $(1-\alpha+2\gamma)/(1+2\gamma)$. However, once knowing that one of the voters hasn't changed for candidate A, the conditional probability for one of the rest becomes $(1-\alpha+\gamma)/(1+\gamma)$. Finally, after we already know that two of them remained of the same opinion, the third one's probability to stick to the same choice becomes $1-\alpha$. Thus, the wanted probability is the product of the three:

$$\frac{(1-\alpha)(1-\alpha+\gamma)(1-\alpha+2\gamma)}{(1+\gamma)(1+2\gamma)}.$$

We would, of course, like to compute all the transition probabilities and it turns out that they are all quotients of products of similar expressions. Thus, to simplify these expressions we introduce the following notation:

$$\Pi_i^j(\xi) = \begin{cases} 1 & \text{if } j < i \\ \prod_{r=i-1}^{r=j-1}(\xi + r\gamma) & \text{otherwise} \end{cases}.$$

Using it, the probability, computed in the previous paragraph, becomes

$$\frac{\Pi_1^3(1-\alpha)}{\Pi_1^3(1)}.$$

We will omit all the details of the tedious and involved considerations leading to the development of transition probabilities of this model. Here is the final result, using the above notation:

$$p_{ij} = \sum_{r=\max\{0,i+j-k\}}^{\min\{i,j\}} \binom{i}{r}\binom{k-i}{j-r} \frac{\Pi_1^{j-r}(\alpha)\,\Pi_{-r+1}^{k-i-j}(1-\alpha)\,\Pi_{k-i+1}^{k-r}(\beta)\,\Pi_1^{r}(1-\beta)}{\Pi_1^k(1)}.$$

It is not so easy to check that this defines a stochastic matrix, but, it is true. And, it is even harder to find out its positive left eigenvector corresponding to eigenvalue 1, which represents the equilibrium point of this Markov chain. However, such an eigenvector exists and it is even possible to compute it. The stationary distribution is given by probabilities:

$$q_i = \binom{k}{i}\frac{\Pi_1^i(\alpha)\,\Pi_1^{k-i}(\beta)}{\Pi_1^k(\alpha+\beta)}.$$

We can easily recognize in the above equilibrium distribution the well-known Polya's distribution for k draws from an urn with parameters $\frac{\alpha}{\alpha+\beta}$ and $\frac{\gamma}{\alpha+\beta}$. The same distribution was obtained as stationary by Coleman (1964) in his continuous time model treating a similar problem. He also claims that this distribution has been confirmed by some experimental data.

6 Conclusion

Although the theory of Markov models has been often used in the studies of dynamics of social systems, there are many questions left unanswered. One of the area which may be gaining in its popularity and where many problems are yet unsolved are the interactive Markov models. This calls for a more systematic treatment of the subject.

References

[1] Bartholomew, D. J. (1981): *Mathematical Methods in Social Science.* New York: John Wiley.

[2] Bartholomew, D. J., Forbes, A. W. (1979): *Statistical Techniques for Manpower Planning.* New York: John Wiley.

[3] Coleman, J. S. (1964): *Introduction to Mathematical Sociology.* London: Collier-MacMillan.

[4] Conlisk, J. (1976): Interactive Markov chains. *J. of Math. Sociology*, **4**, 157-185.

[5] Conlisk, J. (1978): A stability theorem for an interactive Markov chain, *J. of Math. Sociology*, **6**, 163-168.

[6] Grinold, R. C., Marshall, K. T. (1977): *Manpower Planning Models*. New York: Elsevier North-Holland.

[7] Kulkarni, M. S., Kumar, A. (1986): A mathematical model of progressive literacy. *J. of Math. Sociology*, **12**, 275-298.

[8] Kulkarni, M. S., Kumar, A. (1989): Interactive Markovian models of progressive trends. *J. of Math. Sociology*, **14**, 45-65.

[9] Omladič, V. (1992): Nekateri modeli zaposlovanja in promocije. *Blejsko metodološko srečanje '91, Metodološki zvezki*, **8**, Ljubljana: FDV, 123-135.

[10] Schaefer, H. H. (1974). *Banach Lattices and Positive Operators*. New York: Springer-Verlag.