Markov Systems with Fuzzy States for Describing Students' Educational Progress in Greek Universities

Symeonaki, Maria
Panteion University of Social and Political Sciences, Department of Social Policy
136 Syggrou Av.
Athens (17671),Greece
E-mail: msymeonaki@me.com

Kalamatianou, Aglaia
Panteion University of Social and Political Sciences, Department of Sociology
136 Syggrou Av.
Athens (17671),Greece
E-mail: AglaiaKalamatianou@gmail.com

ABSTRACT

In this paper the theory of non homogeneous Markov systems (NHMS) with fuzzy states is used for describing students' educational progress in Greek Universities. More specifically, a model for projecting students' transitions among "progress levels" related to academic years is provided. The progress levels, assumed fuzzy, are defined and the relevant membership functions are given in terms of the degree of the difficultly of the course units. An application of the proposed model is given using data drawn from a Greek public University.
keywords *Markov systems, Markov processes, fuzzy states, students' educational progress*

Introduction

The idea of a non homogeneous Markov system (NHMS) with fuzzy states introduced firstly in Symeonaki and Stamou (2004a) is an attempt to deal with a number of real applications in manpower planning and in population dynamics in general, when one is faced with the fact of fuzzy states, which represent states of the system that cannot be precisely defined. The classification of the state space according to traditional methods (member, non-member) introduces uncertainty that is better dealt perceiving states as having imprecise boundaries that facilitate gradual transition from membership to nonmembership and vice versa. The rate of convergence, the asymptotically attainable structures and the sensitivity of such systems are examined in Symeonaki and Stamou (2006), while some aspects of input control are given in Symeonaki and Stamou (2004b).

In this paper the concept of a NHMS with fuzzy states is used in order to describe and model the transitions of Greek university students among progress levels that are related to academic years and represent students' educational progress.

Modeling the educational progress of students and predicting enrollments and degrees awarded by Universities is fundamental to higher education planning. In recent years, various approaches of modeling such as processes have been suggested (Markov, regression or simulation). The Markovian approach appears to be one of the most widely applied. There, it is assumed that each student occupies a given state at time t and makes a transition from state to state at time $t+1$ (the first and the last state may represent enrollments and graduations correspondingly, while other states represent educational progress). In Harden and Tcheng (1971) transition matrices are computed from available historical data, and enrollment projections are calculated by repetitive multiplication of a given enrollment distribution by the transition matrix. Gani (1963) was the first to produce formulae for estimating student enrollments and degrees awarded in Australian Universities. In doing so, he constructed a simplified model for student progress through a university course from which the formulae were derived. Since then there has been numerous methods to forecast university enrollments and degrees awarded in the literature (see for example Song and Chissom (1994), Sah and Degtriarev (2005), among others).

The present work is an attempt for modeling transitions of the Greek university students between progress levels, assuming that the state space is fuzzy. This is a very realistic assumption because in the Greek university system academic years do not really represent students' educational progress. Students can move to a higher academic year without having successfully passed all course units that normally correspond to a previous academic year. This gives room to students first enrolled in the same academic year to belong to different progress levels at the end of the year. Consider, for example, two freshmen (at time t); the first one is successful in all his/her exams and the second one is not successful at all. These two students cannot actually belong to the same progress level at time $t+1$ (end of the first academic year). This process leads students to have a certain distribution of duration of studies (explained more in the next section) that can be used to estimate degrees awarded. Thus, by estimating students' progress level we can project, at an early stage, the number of degrees awarded.

In the above context, the remainder of the paper is organized as follows: In the next Section, the basic elements of the model describing the student system are defined and the expected number of students at each progress level, at time t, is given. The fuzzy states of the student system are provided and the relevant membership functions are estimated in terms of the degree of difficulty of the course units. The last Section provides a numerical example illustrating the methodology of estimating the membership functions denoting the participation to the different level of progress, with data drawn from the archives of a Greek public University, the Panteion University of Social and Political Sciences.

The NHMS with fuzzy states for describing students' progress in Greek Universities

To facilitate understanding of the problem considered in this paper we give first some details concerning studies in Greek universities. The majority of the university curricula are of the type of four-academic years or eight semesters (exceptions correspond to medical, veterinary, engineering and agricultural sciences). Graduation is possible at the end of the prescribed period of time if a certain number of course units, is successfully completed by the students. This number of course units and its distribution to the academic years is decided by the corresponding department. However, according to the legislation covering university studies a student can enroll to a higher academic year without having successfully passed all course units that are normally taught in the previous academic year. Students are allowed to take exams for an unlimited number of times till successfully complete the certain number of course units that is necessary for graduation. It is apparent that there exists a prescribed minimum time for graduation but there is no corresponding upper limit. The above process leads students registered in the same academic year to potentially have different levels of progress, while considering graduation time students can graduate at the minimum prescribed time, can have a long period for graduation or can remain as perpetual students. The distribution of such as duration of studies is fully examined in Kalamatianou and McClean (2003). Given this, the question now arises,

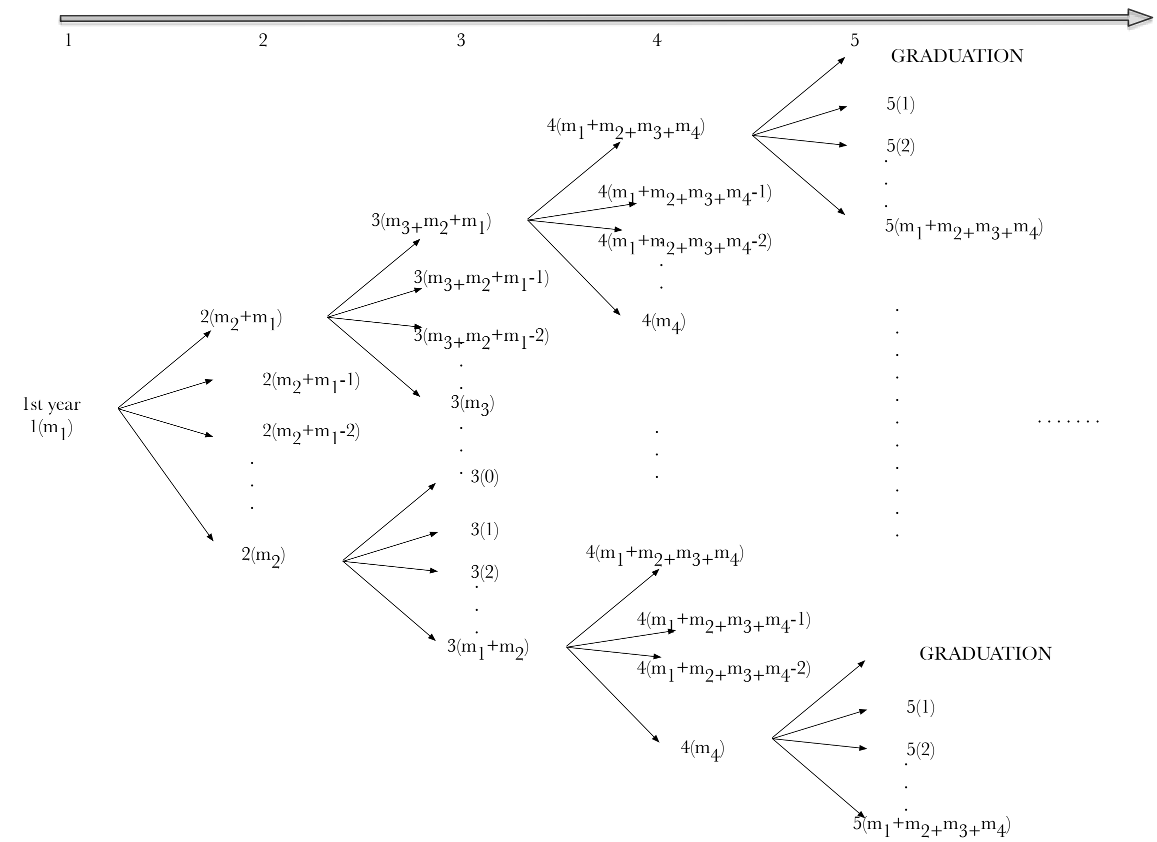

Figure 1: Representing the transitions.

as to whether modeling educational progress helps us predict, in early time, the duration of studies and finally the degrees awarded. In this paper we consider the case of four-academic years curricula, but the model could easily be adjusted in the case of five or more years of studies.

Now, let $i = 1, 2, 3, 4$ indicate the academic year and m_i the number of course units being taught during that year. The proposed model assumes that students are stratified into different crisp (nonfuzzy) states according to their progress, which is measured in terms of the students' academic performance. These crisp states are called *progress states.* Let $i(l)$ denote the progress state of a student of the $i - th$ academic year that would have to pass l course units, by the end of that year. Consider for example the first year of studies (progress state $1\,(m_1)$); at the end of this year a student can either continue his/her studies or voluntarily leave the system. If the student carries on, he/she can do so, having passed all his/her exams (progress state $2(m_2)$), having passed none of the course units (progress state $2(m_1 + m_2)$), having missed one (progress state $2(m_2 + 1)$), etc. Accordingly, a student can move to the $(i + 1) - th$ academic year, having to pass $\left(\sum_{l=1}^{i+1} m_l - j\right)$ for $j = 0, 1, 2, ..., \sum_{k=1}^{i} m_k$ and $i = 1, 2, 3$, course units. Figure 1 represents all possible transitions of the students. It is apparent that the state space of the student system is $S = \{(1(m_1), 2(m_2), 2(m_2 + 1), , ..., 2(m_2 + m_1), 3(m_3), ..., 3(m_3 + m_2 + m_1), 4(m_4), ..., 4(m_4 + m_3 + m_2 + m_1), ...\}$. Obviously, the state space of the NHMS consists of a rather large number of states and the transitions between them are not actually associated with these exact states. For this reason it is more appropriate to perceive the states as having imprecise boundaries and to assume that the system consists of four fuzzy states, each denoting a level of progress related to the corresponding academic year. We define by $F = \{F_1, F_2, F_3, F_4\}$ the fuzzy state space (i.e. the set of all fuzzy states), where F_1 represents the "1st LEVEL OF PROGRESS", F_2 the "2nd LEVEL OF PROGRESS", F_3 the "3rd LEVEL OF PROGRESS" and F_4 the "4th LEVEL OF PROGRESS". F_r, $r = 1, 2, 3, 4$, is assumed to be a fuzzy set on S, i.e. the state space of the system is formed by fuzzy subsets of a primary space of the crisp states of S. Let also $\mu_{F_r}(\cdot) : S \rightarrow [0, 1]$ denote the membership function of a fuzzy set F_r, for $r = 1, 2, 3, 4$. It is assumed that $F = \{F_1, F_2, ..., F_N\}$ defines a fuzzy partition on $S = \{1, 2, ..., k\}$ such that $\sum_{r=1}^{N} \mu_{F_r}(\cdot) = 1$, i.e. F is a Ruspini partition on S (Ruspini (1969)). This is not restrictive, since in numerous real applications it is rather appropriate to use the condition $\sum_{r=1}^{N} \mu_{F_r}(\cdot) = 1$. The basic parameters of the proposed model are now given:

1. $N_{F_r}(F, t)$ is the number of students in fuzzy state F_r at time t, $r = 1, 2, 3, 4$, $t = 0, 1, 2, ...$, calculated as the fuzzy cardinality of F_r,

2. $p_{F_r F_s}(F, t)$ is the transition probability from fuzzy state F_r to fuzzy state F_s at time t, $r = 1, 2, 3, 4$, $s = 1, 2, 3, 4$,

3. $p_{oF_s}(F, t)$ is the probability that a new student entering the system at time t goes to fuzzy state F_s, $s = 1, 2, 3, 4$,

4. $p_{F_r N+1}(F, t)$ is the probability that a student being at fuzzy state F_r at time t leaves the system $r = 1, 2, 3, 4$,

5. $q_{F_r F_s}(F, t)$ is the total transition probability from fuzzy state F_r to fuzzy state F_s at time t, i.e. $q_{F_r F_s}(F, t) = p_{F_r F_s}(F, t) + p_{F_r N+1}(F, t)\, p_{oF_s}(F, t)$, $r = 1, 2, 3, 4$, $s = 1, 2, 3, 4$,

6. $T(t)$ is the total number of students serving the system at time t, and

7. $\Delta T(t) = T(t) - T(t-1)$.

Note that $p_{F_r F_s}(F, t)$, $p_{oF_s}(F, t)$, $p_{F_r N+1}(F, t)$ and $q_{F_r F_s}(F, t)$ are estimated in Symeonaki and Stamou (2004) using the concept of the probability of a fuzzy event introduced in Zadeh (1968). What we require to find is the total number of students at fuzzy state F_i, at time t, i.e. we need to estimate the population vector: $\mathbf{N}(F, t) = [N_{F_1}(t), N_{F_2}(t), N_{F_3}(t), N_{F_4}(t)]$. This is given by:

$$\mathbf{N}(F, t+1) = \mathbf{N}(F, t)\, \mathbf{Q}(F, t) + \Delta T(t)\, \mathbf{p}_o(F, t). \quad (1)$$

Now, assume that md_i denotes the mean deference of the grades achieved by the students in the $i-$th course unit minus the average grade in the rest of the units, i.e.

$$md_i = \frac{\sum_{s=1}^{n} \left(y_{si} - \frac{\sum_{k=1}^{l, k \neq i} y_{sk}}{l} \right)}{n} \quad (2)$$

where l denotes the number of units that a student passes in a corresponding academic year, n is the number of students that pass the exams of course unit i in that year and y_{si} is the grade of the $s-th$ student, $s = 1, 2, ..., n$, in the $i-th$ unit. Note that the mean difference, denoted here by md_i was firstly introduced for estimating the difficulty of an examination in Kelly (1976), and is now one of the commonly used examinee linear models for estimating relative difficulty of examinations in different subjects (Coe *et al* (2008)).

We now provide the way of estimating the membership functions, $\mu_{F_r}(\cdot)$, of the fuzzy sets F_r.

1. $\mu_{F1}(1(m_1)) = 1$, i.e. a student that is at the first year of studies, has to follow m_1 course units and is regarded as a student of a "1st LEVEL OF PROGRESS" with a membership degree equal to one. In other words $\mu_{F1}(1(m_1)) = 1$, corresponds to the degree that a student being at state $1(m_1)$ is compatible with the concept of a 1st LEVEL OF PROGRESS student represented by the fuzzy state F_1. Moreover, $\mu_{F2}(1(m_1)) = 0$, $\mu_{F3}(1(m_1)) = 0$ and $\mu_{F4}(1(m_1)) = 0$, which means that the student is considered to be at the 2nd, 3rd or 4th LEVEL OF PROGRESS with a membership degree equal to zero.

2. $\mu_{F2}(2(m_2)) = 1$, i.e. a student that is at the second year of studies and has successfully passed all course units of the first academic year, belongs to the "2nd LEVEL OF PROGRESS" with a membership degree equal to one. Moreover, $\mu_{F1}(2(m_2)) = 0$, $\mu_{F3}(2(m_2)) = 0$ and $\mu_{F4}(2(m_2)) = 0$, meaning that he/she is considered to be at the 1st, 3rd and 4th LEVEL OF PROGRESS with a membership degree equal to zero.

3. $\mu_{F1}(2(m_2+1)) = \frac{1+|md_i|}{m_1}$, i.e. a student that is at the second year of studies and has successfully passed all course units of the first academic year except one (which has a degree of difficulty equal to md_i) is regarded as a 2nd LEVEL OF PROGRESS student with a membership degree equal to $\frac{1+|md_i|}{m_1}$. Moreover, $\mu_{F2}(2(m_2+1)) = 1 - \mu_{F1}(2(m_2+1))$, $\mu_{F3}(2(m_2+1)) = 0$ and $\mu_{F4}(2(m_2+1)) = 0$. For example, consider a student that moves from the 1st year to the 2nd who has successfully passed all course units of the first academic year except one (out of 8 $(m_1 = 8)$). Let also that the md of that course unit is equal to -0.77. Therefore, the student will be considered to be at fuzzy state F_1 ("1st LEVEL OF PROGRESS") with a membership function equal to $\mu_{F1}(2(m_2+1)) = \frac{1+0.77}{8} = 0.2125$. Simultaneously, the student will be at state "2nd LEVEL OF PROGRESS" with a membership function equal to 0.7875, etc.

A numerical example

We now provide a numerical example with data drawn from the archives of the Panteion University of Social and Political Sciences. We consider three students of the Department of Sociology and we observe them at the stage where they have finished the first two semesters and move to the second academic year of studies. The first student (A) failed to pass the exams only in one course unit (C_1), the second student (B) passed six course units and the third student (C) failed in all course units except one. The reason we examine these students is that they are typical examples that correspond to the three categories of duration of studies detected in Kalamatianou and McClean (2003), mentioned earlier. Student A belongs to the category of those that graduate just after the completion of the minimum required time. Student B belongs to the category of the students whose graduation process takes a long time and student C belongs to the category of the perpetual students. For these three students the membership functions indicating the participation in the 1st , 2nd, 3rd and 4th LEVEL OF PROGRESS are estimated. The procedure illustrated in Table 1 concerns student A and course unit C_1, which was selected by 103 students. Column 2 provides the corresponding grades of these students at C_1. The third column gives for each student the average grades in all other units except C_1. The difference is given at the fourth column. Then, the md of unit C_1 is equal to 0.0899 (Equation (2)). Therefore, the specific student is considered to be at the different fuzzy states with a membership function equal to: $\mu_{F2}(m_2+2) = \frac{1+0.0899}{12} = 0.909$, ("2nd LEVEL OF PROGRESS"), $\mu_{F1}(m_2+2) = 0.091$, ("1st LEVEL OF PROGRESS"), $\mu_{F3}(m_2+2) = 0$, ("3rd LEVEL OF PROGRESS"), $\mu_{F4}(m_2+2) = 0$, ("4th LEVEL OF PROGRESS"). Table 2 provides the outcomes for students A, B and C. The values of the membership functions abide by the former

Table 1: Estimating the degree of difficulty of course unit C_1

Student	Grade	Average grades in other units	Difference
1	7	6.49	0.51
2	7	6.32	0.68
⋮	⋮	⋮	⋮
103	6	5.60	0.40
mean difference			0.0899

Table 2: Membership functions and duration of studies

Student	Membership function μ_{F1}	Membership function μ_{F2}	Duration of studies (in months)
A	0.091	0.909	42
B	0.480	0.520	102
C	0.990	0.01	-

categorization of students and highlights the possibility of predicting the duration of studies from the students' progress at the first year, i.e. the membership values m_{F1}.

REFERENCES

R. Coe, J. Searle, P. Barmby, K. Jones and S. Higgins. Relative difficulty of examinations in different subjects, CEM Centre, Durham University, 2008.

A. Gani. Formulae for projecting Enrollments and Degrees awarded in Universities, *Journal of the Royal Statistical Society*, 126(3), 400-409, 1963.

W. Harden and M. Tcheng. Projection enrollment distribution with enrollment ceilings by Markov processes, *Socio-Economic Planning Sciences*, 5(5), 467-473, 1971.

A. Kalamatianou and S. McClean. The perpetual student: Modeling duration of undergraduate studies based on lifetime-type educational data, *Lifetime Data Analysis*, 9(4), 311-330, 2003 A. Kelly. A study of comparability of external examinations in different subjects, *Research in Education*, 16, 37-63, 1976.

E. Ruspini. A new approach to clustering, *Information and Control*, 15, 1969, 22-32.

M. Sah and K. Degtiarev. Forecasting enrollment model based on First-Order Fuzzy time series, *World Academy of Science, Engineering and Technology I*, 132-135, 2005

Q. Song and B. Chissom. Forecasting enrollments with fuzzy time series - Part II, *Fuzzy Sets and Systems*, 62, 1-8, 1994.

M. Symeonaki and G. Stamou. Theory of Markov systems with fuzzy states, *Fuzzy Sets and Systems*, 143, 427-445, 2004.

M. Symeonaki and G. Stamou. Input control in non Homogeneous Markov systems with fuzzy states, *International Mathematical Journal*, 5(6), 567-575, 2004.

M. Symeonaki and G. Stamou. Rate of convergence, Assymptotically attainable structures and Sensitivity in non homogeneous Markov systems with fuzzy states, *Fuzzy Sets and Systems*, 157, 143-159, 2006.

L. Zadeh. Probability measures of fuzzy evants, *J. Math. Anal. Appl.*, 23, 421-427, 1968.