DOI: 10.7763/IPEDR. 2012. V58. 4

Online versus Classroom – A Mathematical Model for the Future of Education

Yixun Shi

Department of Mathematics, Computer Science and Statistics, Bloomsburg University of Pennsylvania, USA

Abstract. Today, online education is growing rapidly as a new way of education. What are the driving forces behind this trend? What will happen in the future of education? This paper observes a number of reasons that attract people to online education and looks into the trend for the future education. A Markov chain model with changing transition matrices is used to describe the trend. Numerical procedures for establishing the model and estimating changes and long term trend are also presented.

Keywords*:* Online Education, Markov Chain Model, Long Term Trend, Numerical Procedure

1. Introduction

With the rapid development of multimedia technologies, devices, services and educational software, online learning has been rapidly growing as a new way of education. More and more educators and learners have accepted or are getting interested in the concept and practice of online education. Online education is forming a new trend of education globally.

What are the driving forces behind the rapid development of online education? We can list quite a number of them.

- **Needs from the job market.** In the recent decades, numbers of college graduates are continuously increasing, and increasing rapidly. Take China as an example. In the year of 2003, 2.12 million students graduated from colleges in China. In the year of 2009, this number became 6.11 million. By the month of June of 2009, only about 45% of those 6.11 million college graduates had signed a work contract. Two years later, by the month of May of 2011, according to a report of a Chinese educational research organization, only about 64% of the 2009 college graduates were employed. The job market does not need that many college graduates! Instead, the job market needs people with appropriate skills that many college majors do not teach.
- **Flexibility and broad range of online classes.** With the rapid development of technology and education software, online classes and training programs are available in almost all practical fields, with the flexibility of time schedule and pace of learning. Therefore, many college students and/or graduates, with or without a job, are continuing their education by taking various online courses, in order to make themselves better fit today's job market.
- **Needs of continuing education.** In this era of information, existing knowledge is replaced by new knowledge every day. Therefore, continuing education becomes an important part of the professional life in many careers. Take United States as an example. Medical doctors, accountants, teachers, and even car repairers, are required to take new training programs periodically. Almost all such training programs are offered and taken online.
- **Social needs.** Elderly people tend to take online classes in health education; housewives tend to take online classes in home management education; retired people now have all the time to pursuit another academic degree through online degree programs; and all the people get online to see the world, to learn different cultures, and to make new friends. A new trend of elementary and secondary education is home schooling. Take United States as an example, due to the weak moral education and the poor academic achievements of many public schools, more and more families are practicing home schooling for their children at elementary and/or secondary level. A complete education program is available online, and many home schooling families are using the online program for the education of their children for the elementary and secondary grades.

- **Changes in qualities of online education and classical college education.** The companies offering online education services, technologies, and educational software systems are after profits. Therefore, they tend to do their best to improve the quality of their products so they can attract and keep customers. This is exactly what we see in the online education market. The devices and software are more and more user friendly, and education system is more and more effective, the quality of education is getting better and better. On the other hand, the quality of classical college education is in general going down. Take China as an example. Most of Chinese universities are continuously expending their enrollment of freshman every year. Most universities in China have more than one campuses, with one being the "old one" and the others "new campuses". The "old campus" is the original university campus, which is usually conveniently located in the city, and is where the main administration and faculty departments are. The "new campuses" are usually huge and located in suburban, where undergraduate dorms and classrooms are placed. Professors take shuttle buses to the new campuses, normally about one hour each way, to teach classes. After the class, they will be aboard the shuttle bus again and get on their way back to the city campus. Students either communicate with professors via email, or don't communicate at all. Take United States as another example. Due to the budget concerns, most public colleges today are trying to enlarge class size, to increase faculty teaching loads, to reduce the number of permanent professors, and to increase the number of temporary instructors.
- **Cost of online education versus cost of classical college education.** The cost of the classical four year college education is shooting up in the recent years. Today, the average total cost per year in a public university in the Unites States is around \$30000, and an average student will graduate college with around \$22,000 to \$27,000 in debt. On the other hand, the cost of online courses and training programs are affordable and stay stable.
- **Acceptance of academic society and positive responses from industry.** When time goes by, due to above reasons and many other facts, online degree, online course achievements, and online training program certificates are becoming better and better received by the academic society and employers from various industries. Many universities, including a number of top universities, are offering "combined programs" in which a portion of the program is completed online. These have become another strong motivation to drive people to the online education.

Where is the trend of online education leading us to? Would online education eventually replace our traditional classroom education? Motivated by these questions, this paper is considering a Markov chain dynamical model to picture the future of our education.

In current days, even though online education is growing so rapidly, most of children in the world will still go through the regular classrooms of elementary school, secondary school, and perhaps college, too. Only some children, such as home schooling children in Unites States, choose to go online for their elementary and secondary education. However, with the progress towards perfectness of online education products and full acceptance of society, time will come to a point when every family can make a choice between classroom and online for their children's education programs, just like that today's American parents have choices among public schools, private schools, and home schooling for their children.

Having that said, let us start illustrating our Markov chain dynamical model. We consider the population of all elementary, secondary, and college year students. Suppose a survey is made at the beginning of each year to find out the percentage of students that are enrolled in online education programs. Each survey shows a status vector $S = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where x_1 stands for the percent of the population studying online and x_2 for the percent of the population studying in regular schools. Clearly $0 \le x_1, x_2 \le 1$ and $x_1 + x_2 = 1$. To be more precise, let us use S_1 to denote the initial status vector, that is, the status vector at the beginning of the first year, S_2 to denote the status vector at the beginning of the second year, and so on so forth.

Suppose the successive surveys show that in each year, a certain percent, denoted by $p_{11,}$ of those doing online continue to do it and rest of them, denoted by $p_{21,}$ transfer to regular schools. Meanwhile, a certain percent, denoted by $p_{12,}$ of those attending regular schools switch to online while the rest of the group, denoted by p_{22}, continue to stay in regular schools.

With those notations, we have that

$$S_2 = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} S_1 \tag{1}$$

Similarly we have

$$S_3 = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} S_2 \tag{2}$$

In general, if we assume that the same changes will occur in the population for every year, then we have

$$S_{k+1} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} S_k \text{ , for } k = 1, 2, 3,\ldots \tag{3}$$

The sequence of status vectors { $S_1, S_2, S_3 \ldots$ } forms a Markov chain dynamical system, with the transition matrix

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{4}$$

It is easy to see that this is a regular Markov chain with a stationary status vector

$$S = \lim_{k \to \infty} S_k$$

In other words, if the same changes keep going on every year, then eventually the status vector of the student population will approach the limit vector S, and the status will stay that way forever. Moreover, the stationary status S can be easily obtained by solving the equation

$$PS = S \tag{5}$$

In our real life, however, situations are not that simple. The transition matrix P surely changes from year to year. Therefore, the above simple model would not work well. In this paper, we consider a Markov chain model with changing transition matrices, and try to mathematically model the transition matrices and numerically estimate the stationary status for the trend of education.

2. Models for Changing Transition Matrices

Let $P(k) = \begin{bmatrix} p_{11}(k) & p_{12}(k) \\ p_{21}(k) & p_{22}(k) \end{bmatrix}$ be the transition matrix for the k-th year. That is, $p_{11}(k)$ stands for the percentage of those doing online schooling in the k-the year that continue to do it, $p_{21}(k)$ for the percentage that transfers to regular schools. Similarly, $p_{12}(k)$ stands for the percentage of those attending regular schools in the k-the year that switch to online schooling, while $p_{22}(k)$ for the percentage that continue to stay in regular schools. Values of $p_{11}(k)$, $p_{21}(k)$, $p_{12}(k)$ and $p_{22}(k)$ change from year to time year, and hence they are functions of k. However, it is always true that

$$0 \le p_{11}(k), p_{21}(k), p_{12}(k), \ p_{22}(k) \le 1 \tag{6}$$

and

$$p_{11}(k) + p_{21}(k) = 1, \quad \text{and} \quad p_{12}(k) + p_{22}(k) = 1 \tag{7}$$

With the driving forces described in the previous section, we may expect a growth of the value of $p_{12}(k)$ in the beginning years, reflecting the fact that more and more students are taking online schoolings. The value of $p_{11}(k)$ may stay quite stable and high because those who have started online schooling may very much likely to stay with it as long as it fits their needs. When time goes by, however, the value of $p_{12}(k)$ may start to decease, reflecting the existence of those families who prefer and can afford regular schools. In long run, the transition matrix $P(k)$ may eventually get closer and closer to the identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, meaning that those take online schools for a long time may very likely to stay with it through their school years, and those strongly prefer regular schools will stay with regular schools. The status vector S_k reached in those years will then be approaching the stationary status for the trend of education.

Based on above assumptions, many mathematical functions may be used to establish models for $p_{11}(k)$ and $p_{12}(k)$. Here we describe two such models.

- The upside-down normal distribution model $M_{11}(k) = 1-\frac{\alpha_1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}(\frac{k-\mu_1}{\sigma_1})^2}$ for $p_{11}(k)$. Three parameters α_1, μ_1, and σ_1 are involved in the formula of this model. The model is formed by subtracting from 1 a portion of the density function of the normal distribution $N(\mu_1, \sigma_1)$. The curve of $M_{11}(k)$, for example, when α_1= 10, μ_1=20, and σ_1= 10, is shown in Fig. 2.1 below. Note that the value of α_1 determines how low the value of $M_{11}(k)$ can get, μ_1 determines when it will get to that low value, and σ_1 determines how long it will take for the value of $M_{11}(k)$ to approach 1.
- The normal distribution model $M_{12}(k) = \frac{\alpha_2}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}(\frac{k-\mu_2}{\sigma_2})^2}$ for $p_{12}(k)$. This model is formed by using a portion of the density function of the normal distribution $N(\mu_2, \sigma_2)$. Again, three parameters α_2, μ_2, and σ_2 are involved. In Fig. 2.2 below, we display the curve of $M_{12}(k)$ when α_2= 5, μ_2= 18, and σ_2= 8.

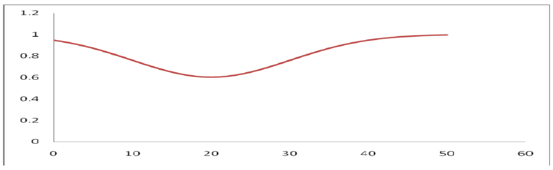

Figure 2.1. Curve of $M_{11}(k)$

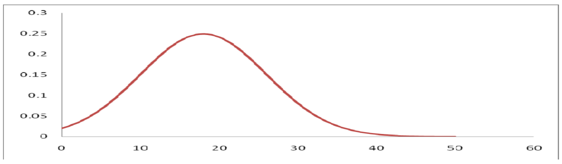

Figure 2.2. Curve of $M_{12}(k)$

Since models $M_{11}(k)$ and $M_{12}(k)$ each involves three parameters, in order to numerically estimate these parameters, surveys need to be conducted for at least four years. For example, suppose a survey is made at the beginning of the first year and then at the beginning of n following years (with $n \geq 3$) and the values of $p_{11}(k)$ and $p_{12}(k)$, for $k = 1, 2, \ldots n$, are found from these surveys, then by numerically solving the nonlinear system

$$p_{11}(k) = 1 - \frac{\alpha_1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}(\frac{k-\mu_1}{\sigma_1})^2} \quad \text{for } k = 1, 2, \ldots n \tag{8}$$

We can determine the parameters α_1, μ_1, σ_1 in $M_{11}(k)$, and thus establish a model for $p_{11}(k)$. Similarly, by solving the nonlinear system

$$p_{12}(k) = \frac{\alpha_2}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}(\frac{k-\mu_2}{\sigma_2})^2} \quad \text{for } k = 1, 2, \ldots n \tag{9}$$

We can establish the model $M_{12}(k)$ for $p_{12}(k)$. Many numerical methods are available for solving nonlinear systems (8) and (9). See, for example, [1], [2] and [3].

3. Numerical Estimates of Transition Matrices and Stationary Status

In this section, let us use a simulated example to describe the numerical procedure for estimating the transition matrix and stationary status for the trend of future education. Suppose a survey is made at the beginning of each year for five years. Suppose the first survey shows that the initial status $S_1 = \begin{bmatrix} 0.01 \\ 0.99 \end{bmatrix}$, that is, at that time 1% of the population are taking online schools while 99% are not. Suppose the succeeding surveys show that the transition matrix for these four months are:

$$P(1) = \begin{bmatrix} 0.967 & 0.233 \\ 0.033 & 0.767 \end{bmatrix}, \quad P(2) = \begin{bmatrix} 0.945 & 0.290 \\ 0.055 & 0.710 \end{bmatrix}, \quad P(3) = \begin{bmatrix} 0.919 & 0.350 \\ 0.081 & 0.650 \end{bmatrix}, \quad P(4) = \begin{bmatrix} 0.894 & 0.399 \\ 0.106 & 0.601 \end{bmatrix}.$$

Based on these surveys, we have that $p_{11}(1) = 0.967$, $p_{11}(2) = 0.945$, $p_{11}(3) = 0.919$, and $p_{11}(4) = 0.894$. By solving the nonlinear system (8) with $n = 4$, we get α_1= 1, μ_1=5, and σ_1= 3. Therefore the model for $p_{11}(k)$ in this case is $M_{11}(k) = 1 - \frac{1}{\sqrt{2\pi}\cdot 3}e^{-\frac{1}{2}(\frac{k-5}{3})^2}$. Similarly, the surveys show that $p_{12}(1) = 0.233$, $p_{12}(2) = 0.290$, $p_{12}(3) = 0.350$, $p_{12}(4) = 0.399$. Solving the nonlinear system (9) with these values we see that α_2= 6, μ_2=6, and σ_2= 5. Thus the model for $p_{12}(k)$ will be $M_{12}(k) = \frac{6}{\sqrt{2\pi}\cdot 5} e^{-\frac{1}{2}(\frac{k-6}{5})^2}$. With these two models we may estimate the yearly transition matrix which represents the changes in each year, using the formula

$$P(k) \approx \begin{bmatrix} M_{11}(k) & M_{12}(k) \\ 1-M_{11}(k) & 1-M_{12}(k) \end{bmatrix} \tag{10}$$

The following Table 3.1 shows the estimates of entries in the transition matrices in the 5-th, 10-th, 15-th, and 20-th years. Note that $P(k)$ approached $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ when k gets bigger and bigger. Using the Markov chain formula $S_{k+1} = P(k)S_k$, we can see the trend for the future. Table 3.2 below shows the estimates of S_k for k from 1 to 30. For convenience, in the table the status vector S_k is displayed as a row vector instead of a column vector. From the table we can see that the stationary status will be about $\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix}$. That is, when time goes by, the percentage of the population taking online schools will approach 95%. It is also interesting to mention that, like in the regular Markov chain with constant transition matrix, the stationary status in this model does not depend on the initial status vector. For example, if $S_0 = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$ instead of $\begin{bmatrix} 0.01 \\ 0.99 \end{bmatrix}$, the stationary status will still be $\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix}$; and this remains true for any value of initial status vector.

Table 3.1: Estimates of entries in the transition matrices

k	$p_{11}(k)$	$p_{21}(k)$	$p_{12}(k)$	$p_{22}(k)$	k	$p_{11}(k)$	$p_{21}(k)$	$p_{12}(k)$	$p_{22}(k)$
5	0.867	0.133	0.469	0.531	15	0.999	0.001	0.095	0.905
10	0.967	0.033	0.348	0.652	20	1.000	0	0.009	0.991

Table 3.2: Estimates of Sk for k from 1 to 30

k	S_k	k	S_k	k	S_k	k	S_k	k	S_k
1	[0.01, 0.99]	5	[0.70, 0.30]	9	[0.80, 0.20]	13	[0.89, 0.11]	17	[0.94, 0.06]
2	[0.24, 0.76]	6	[0.74, 0.26]	10	[0.82, 0.18]	14	[0.91, 0.09]	18	[0.94, 0.06]
3	[0.45, 0.55]	7	[0.77, 0.23]	11	[0.85, 0.15]	15	[0.92, 0.08]	19	[0.94, 0.06]
4	[0.60, 0.40]	8	[0.78, 0.22]	12	[0.87, 0.13]	16	[0.93, 0.07]	20 to 30	[0.95, 0.05]

4. References

[1] R. Fletcher, *Practical Methods of Optimization*, 2nd Ed., John Wiley & Sons, New York, 1987.

[2] G. R. Noubary and Y. Shi, "Difference equation models for estimating athletic records", *The Journal of Computational and Applied Mathematics*, 91, pp 107 -- 122, 1998.

[3] Y. Shi, "A globalization procedure for solving nonlinear systems of equations", *Numerical Algorithms*, 12, pp. 272-- 286, 1996.