

Markov Chains and Transition Matrices: Applications to Economic Growth and Convergence

Michael Zabek

An important question in growth economics is whether the incomes of the world's poorest nations are either converging towards or moving away from the incomes of the world's richest nations. Economists have tried since the development of growth modeling to answer this fundamental question. This paper will introduce an important technique in Linear Algebra – the use of transition matrices and Markov Chains – to address this longstanding question.

Transition Matrices and Markov Chains:

Consider a case where we wish to track the number of entities in three different categories. Now, imagine that they are constructed as a vector, \bar{p} of quantities in categories one through three: p_1 , p_2 and p_3 :

$$\bar{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \quad (1.1)$$

Now, let's say that the number of entities in each different category in time period n is determined by a linear process on the entities in each process in time period $n-1$. Then, \bar{p}_n can be given as the following:

$$\bar{p}_n = \begin{bmatrix} t_{11}p_{n-1,1} + t_{12}p_{n-1,2} + t_{13}p_{n-1,3} \\ t_{21}p_{n-1,1} + t_{22}p_{n-1,2} + t_{23}p_{n-1,3} \\ t_{31}p_{n-1,1} + t_{32}p_{n-1,2} + t_{33}p_{n-1,3} \end{bmatrix} \quad (1.2)$$

Here, t_{ij} are constants describing the movement of entities between categories between time periods $n-1$ and n . More specifically, t_{ij} shows the proportion of entities in $p_{n-1,j}$ that move to $p_{n,i}$ in the period specified. In fact, we can further simplify this by writing the above in terms of matrix multiplication:

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} p_{n-1,1} \\ p_{n-1,2} \\ p_{n-1,3} \end{bmatrix} = T \bar{p}_{n-1} = \bar{p}_n \quad (1.3)$$

Where:

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \quad (1.4)$$

As I said above, by employing such matrices we are typically trying to model the movement of different entities. Hence, the properties listed below should follow for any matrix T:

1. All entries t_{ij} should be non-negative. Having a certain number of entities in a category $p_{n-1,j}$ should not negatively influence the number of entities in any future category, $p_{n,i}$.
2. The sum of entries in each column in t should equal one. In other words, each entity in $p_{n-1,j}$ for each j should end up in some $p_{n,i}$.

The above two properties imply that the overall number of things in contained in vectors, \bar{p} between two time periods will remain constant. Each entry must either stay in the same category or move in each iteration of the matrix. Entities in one category cannot cause the count in another category to decline. If the two attributes listed above hold then we call T a transition matrix. Formally, we say that:

Definition 1: A Transition Matrix

An $n \times n$ matrix T is a transition matrix for an n-category vector if:

- a. All Entries in T are non-negative
- b. The sum of all entries in each column of T is 1.

Now, consider the case where the transition matrix between any two consecutive periods, k and k-1 is the same as that for another pair, m and m-1 for some interval, say up to some period n. Now, say that we are interested in \bar{p}_k for some k such that $1 \leq k \leq n$. Then, since the above holds, and we may transition from time period m-1 to m for any m as we did above. Then we can say the following by induction:

$$T^{k-1} \bar{p}_1 = \bar{p}_k \quad (1.5)$$

The above equation is for a Markov Chain composed of transition matrix T. As we move on we will see that Markov chains are quite useful in a number of situations – especially our challenge of tracking nations’ movements between different levels of income per-capita.

Regular Markov Chains and Steady States:

Another special property of Markov chains concerns only so-called regular Markov chains. A Regular chain is defined below:

Definition 2: A Regular Transition Matrix and Markov Chain

A transition matrix, T , is a regular transition matrix if for some k , if T^k has no zero entries. Similarly, a Markov Chain composed of a regular transition matrix is called a regular Markov chain.

For any entry, t_{ij} in a regular transition matrix brought to the k th power, T^k , we know that $0 < t_{ij} \leq 1$. Thus, it is easy to see, then that if we multiply T out to any power above k , it will similarly have all positive entries. This is an important property in considering the behavior of a Markov chain in the very long term.

If we consider the behavior of transition matrices in the long term, the following is an interesting and useful theorem that is beyond the scope of this paper to prove. However, its result is useful and I will use it in the remainder of this paper:

Theorem 1: Steady States of Regular Transition Matrices

For a regular transition matrix, T , there exists some unique column vector \bar{s} with strictly positive entries that sum to one such that:

- a. As m becomes large, all of the columns of T^m approach \bar{s} .
- b. $T \bar{s} = \bar{s}$ for the above unique column vector \bar{s} .

Comparing equation in b. above with that describing an eigenvalue of T , $T \bar{v} = \lambda \bar{v}$ where λ is a scalar, we see that it describes a situation where \bar{s} is an eigenvector of T with eigenvalue 1. It would be interesting to test if this is true for all Markov Chains. The following will prove that this is the case:

Proof 1:

Consider some transition matrix (not necessarily a regular one) T .

Now, we set the characteristic polynomial of the matrix T equal to zero:

$$\det(T - \lambda I) = 0 \quad (2.1)$$

If 1 is an eigenvalue then it one will be a solution to this equation, so:

$$\det(T - I) = 0 \quad (2.2)$$

Now, since T is a transition matrix, the sum of entries in each column equals one. If we subtract I from T then we will be subtracting one from the sum of each column. Thus the sum of each column of $T - I$ equals zero since:

$$\sum_{j=1}^n t_{ij} - 1 = 1 - 1 = 0 \quad (2.3)$$

Thus, we know that if we add all of the t_{ij} in the jth column of the matrix together, then we will get zero for each t_{ij} . So, each row can be turned into a row of zeros by adding every other row to it. If we perform this operation for each row then we will get the zero matrix, whose determinant is, of course, zero. Then, by the properties of determinants we know that all of the row operations that we performed will not change the value of the determinant. Thus we know that equation 2.2. holds for any transition matrix, T.

End of Proof.

The World Income Distribution:

The Wealth of Nations:

Having developed the transition matrix and the Markov Chain, I return to the question that originally motivated this development – how can we model movements in the wealth of nations?

Jones (1997) investigates the relative movements of countries between income brackets. He uses a Markov Chain technique based in the preceding discussion to show how nations moved between different income categories in the period from 1960 to 1988.

For this paper I will update Jones' work by using Penn World Table data (Heston, Summers and Aten (2006)) from 1980 to 2000 to calculate a 3×3 transition matrix and an initial condition column vector. I will first present the data as classifications of countries. From there I will present a second specification designed to proxy for the income levels of individual citizens.

Following Jones' convention, I divided countries into income brackets based on the largest and most influential economy throughout the period – the United States. I define \tilde{y}_i as the

ratio of nation i 's Real GDP per worker to that of the United States in each period. I obtain the following initial vectors:

$$\bar{v}_{1980} = \begin{bmatrix} v_{high} \\ v_{middle} \\ v_{low} \end{bmatrix}_{1980} = \begin{bmatrix} 0.282051282 \\ 0.384615385 \\ 0.333333333 \end{bmatrix}; \bar{v}_{2000} = \begin{bmatrix} v_{high} \\ v_{middle} \\ v_{low} \end{bmatrix}_{2000} = \begin{bmatrix} 0.301282 \\ 0.339744 \\ 0.358974 \end{bmatrix} \quad (1.1)$$

Where v_{high} contains the proportion of countries for which $\tilde{y}_i > 0.4$, v_{middle} represents the proportion of countries where $0.1 > \tilde{y}_i \geq 0.4$ and v_{low} represents countries where $\tilde{y}_i \leq 0.1$. Comparing individual countries' movements, we obtain the following transition matrix, $T_{countries}$:

$$T_{countries} = \begin{bmatrix} 0.909091 & 0.1166667 & 0 \\ 0.090909 & 0.7166667 & 0.115385 \\ 0 & 0.1666667 & 0.884615 \end{bmatrix} \quad (1.2)$$

Needless to say, if we compute $T_{countries} \cdot \bar{v}_{1980}$ then we will get $\bar{v}_{2000} = \begin{bmatrix} 0.301282 \\ 0.339744 \\ 0.358974 \end{bmatrix}$. If we

compute $T_{countries}^t \cdot \bar{v}_{1980}$ for some integer t then we will find the predicted distribution of countries in year $1980 + 20 \cdot t$, $\bar{v}_{1980+20 \cdot t}$. Since $T_{countries}^2$ has all non-zero entries, $T_{countries}$ is a regular transition matrix and we can compute the steady state of this matrix. We will begin by computing the eigenvector for the eigenvalue, v , such that $v=1$ for the above matrix. To find the eigenvector we solve the below:

$$(T_{countries} - I)\bar{s} = 0 \quad (1.3)$$

$$\begin{bmatrix} 0.909091-1 & 0.1166667 & 0 \\ 0.090909 & 0.7166667-1 & 0.115385 \\ 0 & 0.1666667 & 0.884615-1 \end{bmatrix} \bar{s} = 0 \quad (1.4)$$

$$\begin{bmatrix} -0.090909 & 0.1166667 & 0 \\ 0.090909 & -0.2833333 & 0.115385 \\ 0 & 0.1666667 & -0.115385 \end{bmatrix} \bar{s} = 0 \quad (1.5)$$

Then, row-reducing, we find that:

$$\bar{s} = \begin{bmatrix} 0.6085305276 \\ 0.4741783619 \\ 0.6849215152 \end{bmatrix} \quad (1.6)$$

Thus:

$$T_{countries} \cdot \begin{bmatrix} 0.6085305276 \\ 0.4741783619 \\ 0.6849215152 \end{bmatrix} \approx \begin{bmatrix} 0.6085305276 \\ 0.4741783619 \\ 0.6849215152 \end{bmatrix} \quad (1.7)$$

The \approx is inserted because of rounding errors. Listing \bar{s} as a steady state does not make sense in our context, however, because the properties of our transition matrix make it impossible for us to lose countries. The sum of the values in our column vector have to always equal one. So it would be nice if we could find some way to use this eigenvector by scaling it in some way. Luckily, we can say the following about $T_{countries}$:

$$x \cdot T_{countries} \cdot \begin{bmatrix} 0.6085305276 \\ 0.4741783619 \\ 0.6849215152 \end{bmatrix} \approx x \cdot \begin{bmatrix} 0.6085305276 \\ 0.4741783619 \\ 0.6849215152 \end{bmatrix} \quad (1.8)$$

For any scalar x we can find the correct scalar with which to multiply \bar{s} by solving the following equation:

$$x \cdot 0.6085305276 + x \cdot 0.4741783619 + x \cdot 0.6849215152 = 1 \quad (1.9)$$

This reduces to $x = 0.5657291845$. Thus, we multiply the above to scale the eigenvector. The resulting value, by Theorem 1, will be the steady state of our matrix:

$$\bar{s}_{real} = 0.5657291845 \cdot \begin{bmatrix} 0.6085305276 \\ 0.4741783619 \\ 0.6849215152 \end{bmatrix} = \begin{bmatrix} 0.3446328630 \\ 0.2685443689 \\ 0.3878958439 \end{bmatrix} \quad (1.10)$$

Thus, in the long term, we can see that more countries will fall within the lowest category than the highest, if the transition matrix that that we computed for the period from 1980 to 2000 hold. This is a discouraging result for developing nations.

Jones' results, presented in Table 1 and using the same technique as well as dataset are far more positive than ours. Jones uses an earlier period when fewer countries were a.) in existence and b.) large enough to collect good GDP and foreign exchange data. Since I used later dates, my sample now includes a larger number of smaller countries. This may bias my results.

The Wealth of People:

To correct for the above, seek to model the proportion of the world population that lives in countries within each income category. To satisfy the requirements of Markov matrices, I will assume a constant world population. The world population is obviously expanding rapidly, so my assumption is relatively unrealistic. However, the use of Markov matrices will enable me to study trends in the distribution of income through the period, a relatively interesting phenomenon. I also assume zero migration across countries – another unrealistic, but convenient assumption. I begin with the following initial vectors and transition matrix calculated from the Heston, Summers and Aten dataset in terms of Real GDP per capita¹ (for non-economists, this basically means how much the average person in the country lives on per year):

$$\bar{p}_{1980} = \begin{bmatrix} 0.173594 \\ 0.206345 \\ 0.620062 \end{bmatrix}; \bar{p}_{2000} = \begin{bmatrix} 0.175218 \\ 0.45978 \\ 0.365002 \end{bmatrix} \quad (1.11)$$

$$T_{people} = \begin{bmatrix} 0.935844 & 0.061845 & 0 \\ 0.064156 & 0.868254 & 0.434608 \\ 0 & 0.069901 & 0.565392 \end{bmatrix} \quad (1.12)$$

In the period specified it seems that many countries moved out of the lowest category, but that very few actually progressed to the highest category, from examining the transition matrix and changes in the vectors. The transition matrix shows that it was more likely for an individual nation in the highest income category to move into the middle income category than for a middle income country to move into the upper income category.

To examine the matrix over the long term, I will compute the steady state vector as before. The eigenvector corresponding to the eigenvalue of one for this matrix as computed by software is:

$$\bar{s}_1 = \begin{bmatrix} 0.6894142881 \\ 0.7151760614 \\ 0.1150266950 \end{bmatrix} \quad (1.13)$$

¹ Note: Jones uses the same data set, however he uses Real GDP per worker, chain weighted. I use chain weighted Real GDP per capita. This should not be a large distinction, however it is worth noting.

Then, as we did before, imposing our condition that the entries of the steady state vector sum to one, or that $x \cdot 0.6894142881 + x \cdot 0.7151760614 + x \cdot 0.1150266950 = 1$ for $\bar{s}_{real} = x \cdot \bar{s}_1$, gives the following steady state:

$$\bar{s}_{real} = \begin{bmatrix} 0.4536763330 \\ 0.4706291392 \\ 0.0756945281 \end{bmatrix} \quad (1.14)$$

This is much more encouraging. Table 2 reports vector values as well as steady states for the transition matrices and predicted values in specified years.

Conclusion:

As a few simple calculations have demonstrated, Markov chains have real power in a number of applied situations. Obviously this paper has only scratched the surface of the topic of Markov chains. However, it has showed some interesting properties of recent changes in the world income distribution. Without delving too far into the realm of growth theory, the above results show that the 1980's were a period of improvement for most people.

Many of the changes in the incomes of the poor from 1980 to 2000 occurred in relatively large nations at the bottom end of the world income distribution. This is apparent from the differences in my results for my two different transition matrices representing income changes from 1980 to 2000.

Spectacular increases in wealth throughout Asia are masked in the first specification by measuring individual countries and not individual citizens. If we are to use the second specification, my preferred specification, then we can see how positive changes throughout the period were for individuals.

Moving back from our particular application, Markov chains also have important applications to the fields of physics, biology, statistics and to many stochastic processes. They are an immensely powerful and, actually, a quite simple and intuitive concept.

Appendix 1: Tables

Table 1, Jones' Results:

World Income Distributions, Using Markov Transition Method

<i>Interval</i>	<i>1960</i>	<i>1988</i>	<i>Predicted</i>		
			<i>2010</i>	<i>2050</i>	<i>Long-Run</i>
$\tilde{y} \leq .05$	15	17	15	12	8
$.05 < \tilde{y} \leq .10$	19	13	13	11	8
$.10 < \tilde{y} \leq .20$	26	17	14	13	11
$.20 < \tilde{y} \leq .40$	20	22	23	23	24
$.40 < \tilde{y} \leq .80$	17	22	23	26	30
$\tilde{y} > .80$	3	9	12	15	19

Note. Entries in the table reflect the percentage of countries with relative incomes in each interval.

Table 2, Author's results:

Proportions of countries in each category:

	<i>1980</i>	<i>2000</i>	<i>2020</i>	<i>2040</i>	<i>Steady State</i>
<i>High</i>	0.282	0.301	0.314	0.321	0.345
<i>Medium</i>	0.385	0.340	0.312	0.295	0.269
<i>Low</i>	0.333	0.359	0.374	0.383	0.388

Proportions of people in each category:

	<i>1980</i>	<i>2000</i>	<i>2020</i>	<i>2040</i>	<i>Steady State</i>
<i>High</i>	0.174	0.175	0.192	0.215	0.454
<i>Medium</i>	0.206	0.460	0.569	0.610	0.471
<i>Low</i>	0.620	0.365	0.239	0.175	0.076

References:

Anton, Howard and Robert Busby, Contemporary Linear Algebra. John Wiley and Sons, 2003.

Fraleigh, John and Raymond Beauregard, Linear Algebra. Third edition. Addison-Wesley, 1995.

Heston, Alan, Robert Summers and Bettina Aten, *Penn World Table Version 6.2*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

Jones, Charles I. "On the Evolution of the World Income Distribution," *Journal of Economic Perspectives*, Summer 1997, 11, 19-36.