Mathematics IV
(Probability and Statistics)

MATH 403

Lecture 9

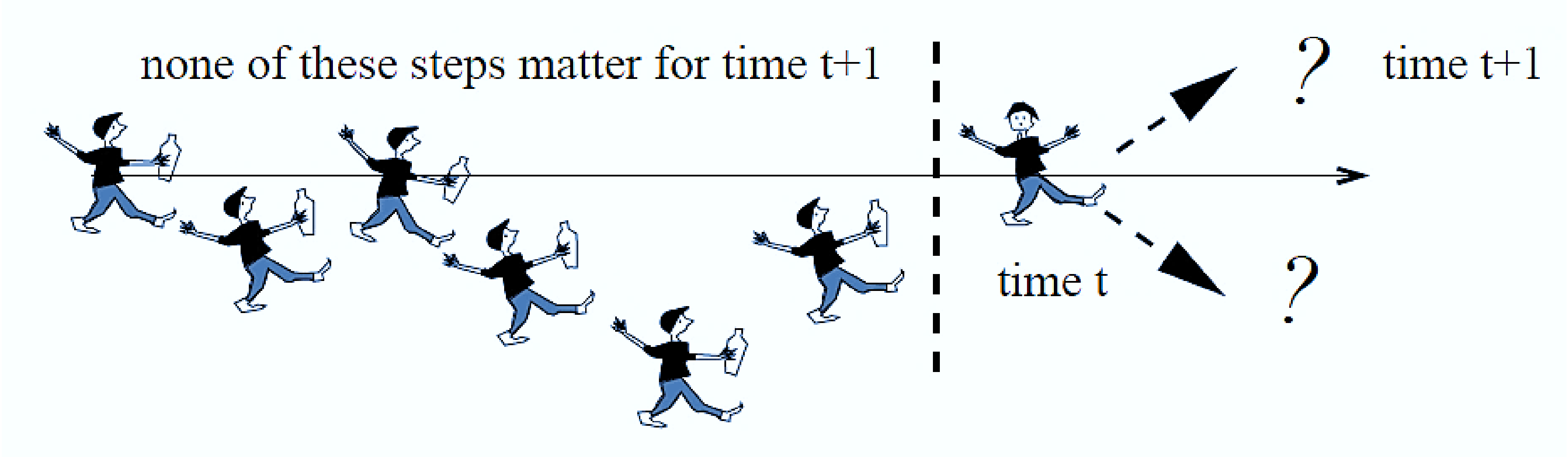
Dr. Phoebe Edward Nashed

Lecture 9 - Outline

- Marcov Chains
- Diagonalization

Introduction to Marcov Chains

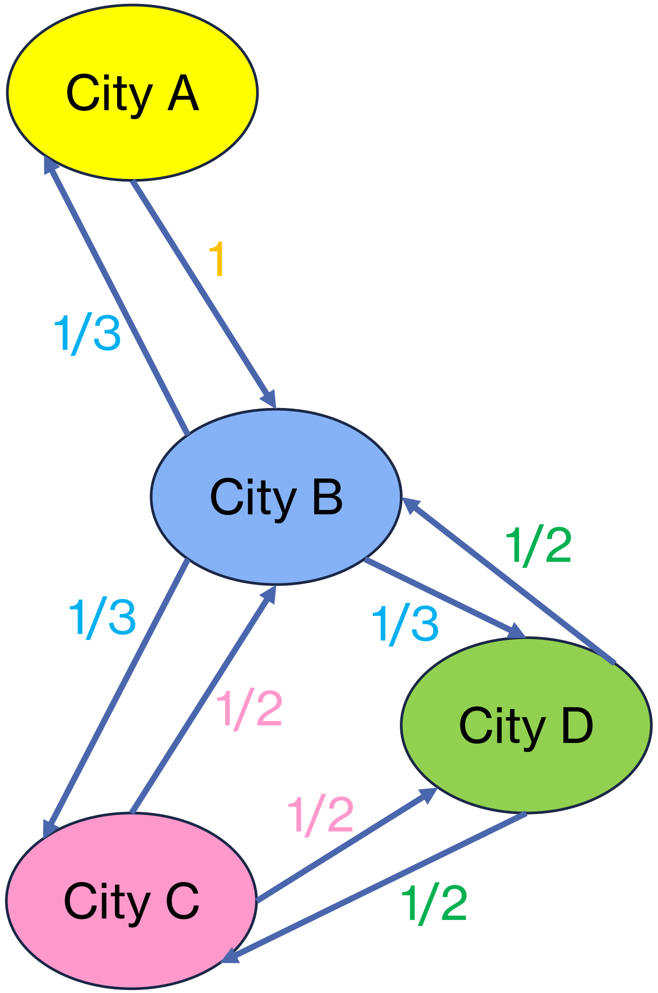
- Markov chain is a sequence of events where the probabilities of the future only depend on the present. It only matters where you are, not where you have been.



Random Walk

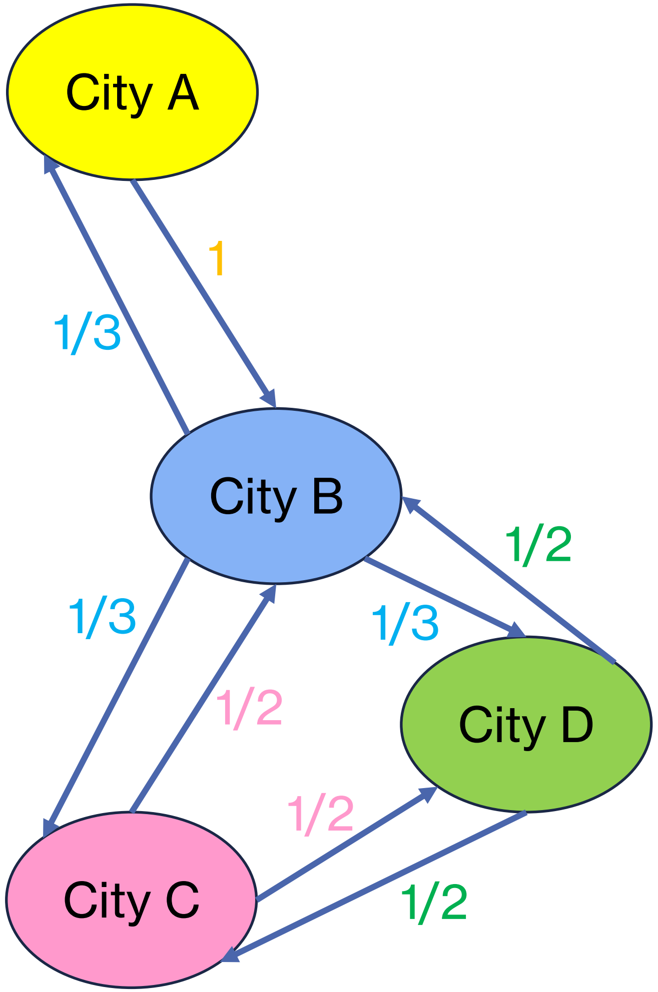
Illustrative Example

- Imagine the shown cities with the interconnection train lines between them.
- Imagine that you are standing for example now at any city of these 4 cities, and that your decision where you will go next depends only on where you are now, not on where you have been previously.



Illustrative Example

- The shown connected graph shows the probabilities of going from any city (state) to the other, given that you know that you are in this city (state) now.
- The sum of all probabilities from a given city to all others (outgoing arrows from any state) is equal to 1.

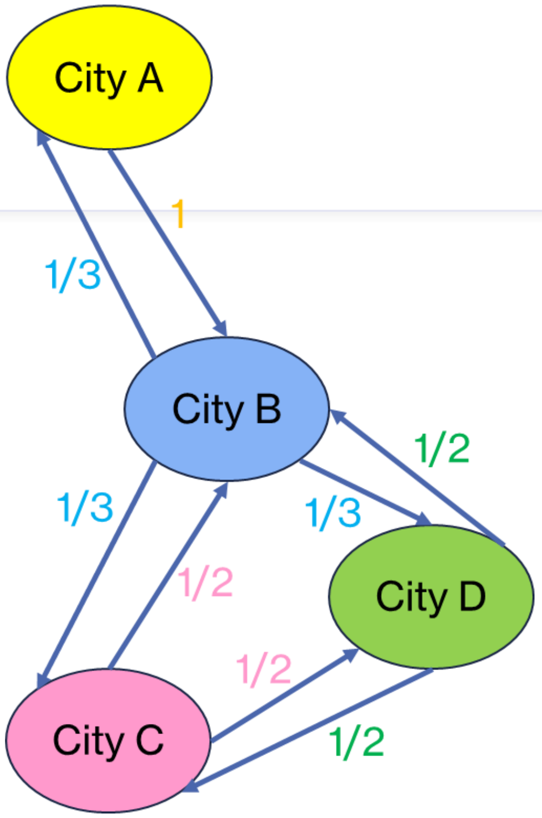


Marcov Chains - Conditions

❖ In a Markov chain, the future depends only upon the present: NOT upon the past. More specifically, a future distribution depends only on the most recent previous distribution. In any Markov process there are two necessary conditions:

1. The total population remains fixed.
2. The probability of transition from any given state should be non-negative.

The Transition Matrix



- If it is known how a population is distributed in any given time interval (initial distribution), the final distribution can be related to the initial one using the tools of linear algebra.

- A matrix ***T***, called a **transition matrix**, describes the probabilistic motion of a population between various states. The individual elements of the matrix reflect the probability that a population moves to a certain state.

$$T = \begin{bmatrix} 0 & 1/3 & 0 & 0 \\ 1 & 0 & 1/2 & 1/2 \\ 0 & 1/3 & 0 & 1/2 \\ 0 & 1/3 & 1/2 & 0 \end{bmatrix}$$

The Transition Matrix

- The two conditions stated above require that in the transition matrix **each column sums to 1**. The transition matrix $\boldsymbol{T}$ is comprised of elements, denoted t_{ij}, relating the motion of a population from state j to state i.

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

- For example, the 3×3 matrix above represents transition probabilities between 3 different states. A given element, t_{23} for example, describes the likelihood that a portion of the population moves from state 3 to state 2.

The State S

- The transition matrix informs us about the probabilities of chaning the state from a given one to any other one. However, we need a means to describe the distribution of the population in any given point in time; we call this the state (of the population).
- The proportions of the population in its various k states are given by a column vector S; known as the **population distribution vector:**

$$S = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{bmatrix}$$

The State S

- The proportions of the population in its various k states are given by a column vector S:

$$S = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{bmatrix}$$

- An element p_i of such a vector, provides the probability that a member of population exists in the i^{th} state. Similar to the requirements on the transition matrix, the sum of the entries in $\boldsymbol{S}$ must add to 1 and be non-negative.

Obtaining the n^{th} State S_n

- Application of a transition matrix to a population vector provides the population distribution at a later time. If the transition matrix remains valid over n time intervals, the population distribution at time n is given by $\boldsymbol{T^{n}S}$. This is simple to demonstrate.

$$\boldsymbol{S_1 = TS_0}$$

$$\boldsymbol{S_2 = TS_1 = T^{2}S_0}$$

$$\vdots$$

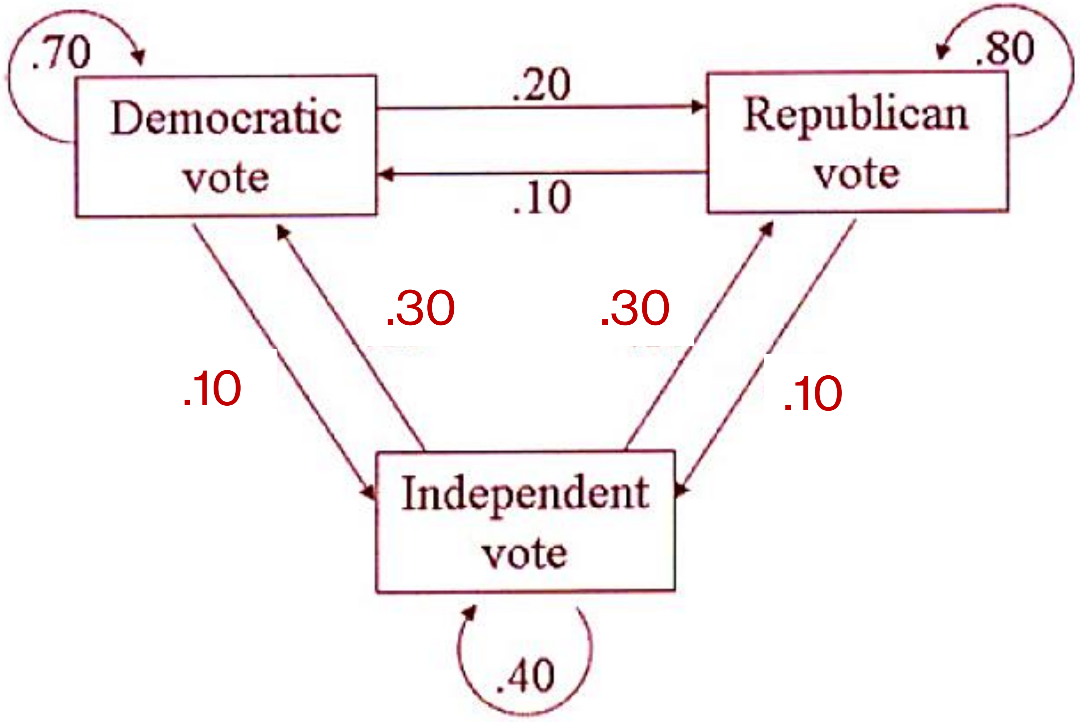
$$\boldsymbol{S_n = TS_{n-1} = T^{n}S_0}$$

- So:

$$\boldsymbol{S_n = T^n\, S_0}$$

Example 1

- A population of voters are distributed between the Democratic (**D**), Republican (**R**), and Independent (**I**) parties. Each election, the voting population $\boldsymbol{S} = [\mathbf{D}, \mathbf{R}, \mathbf{I}]$ obeys the redistribution shown in the next diagram.



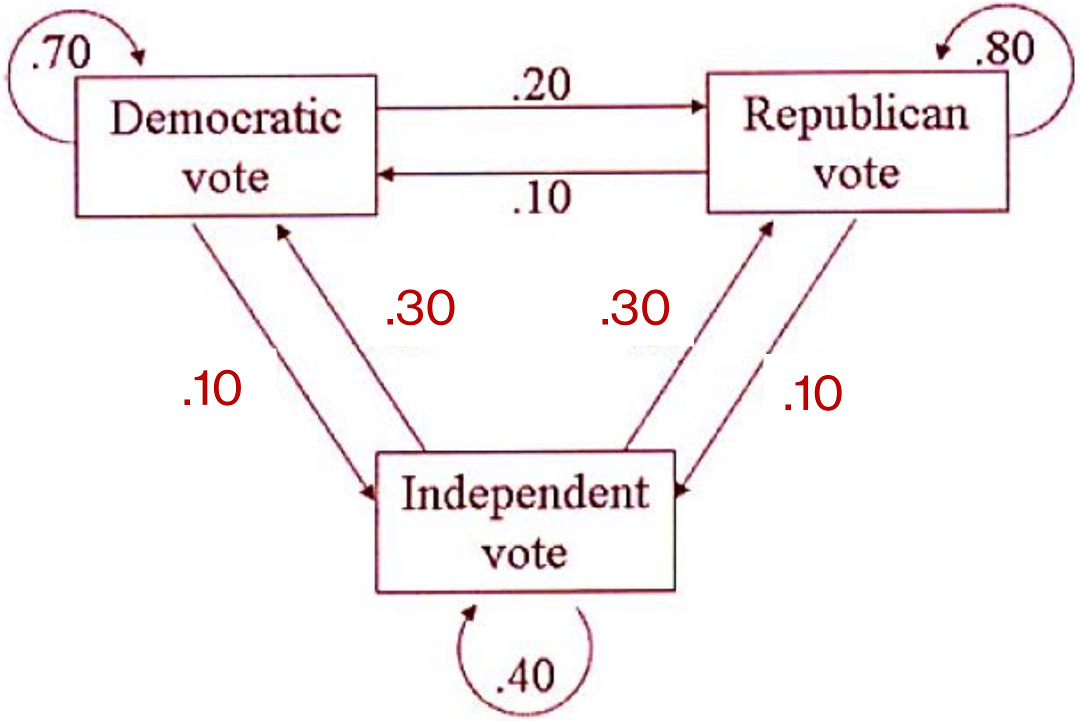
Voter shift between two elections

Example 1

- For example, in an upcoming election, of those that voted Republican in the **previous** election, 80% will remain Republican, 10% will vote Democrat, and the remainin 10% will vote Independent.

- The transition matrix describing this tendency of voter shift is given as

$$T = \begin{array}{c|ccc} \text{Party} & D & R & I \\ \hline D & .70 & .10 & .30 \\ R & .20 & .80 & .30 \\ I & .10 & .10 & .40 \end{array}$$



Example 1

- In the 2004 presidential election, the voters were distributed according to the distribution vector

$$\boldsymbol{S_0} = \begin{bmatrix} 0.48 \\ 0.51 \\ 0.01 \end{bmatrix}$$

$$T = \begin{array}{c|ccc} \text{Party} & D & R & I \\ D & .70 & .10 & .30 \\ R & .20 & .80 & .30 \\ I & .10 & .10 & .40 \end{array}$$

- If the transition matrix $\boldsymbol{T}$ dictates the changes between two primary elections, we can expect the outcome of the 2008 election as follows:

$$\boldsymbol{S_1} = \boldsymbol{TS_0} = \begin{bmatrix} 0.70 & 0.10 & 0.30 \\ 0.20 & 0.80 & 0.30 \\ 0.10 & 0.10 & 0.40 \end{bmatrix} \begin{bmatrix} 0.48 \\ 0.51 \\ 0.01 \end{bmatrix} = \begin{bmatrix} 0.390 \\ 0.507 \\ 0.103 \end{bmatrix}$$

Example 1

$$\boldsymbol{S_1} = \boldsymbol{TS_0} = \begin{bmatrix} 0.70 & 0.10 & 0.30 \\ 0.20 & 0.80 & 0.30 \\ 0.10 & 0.10 & 0.40 \end{bmatrix} \begin{bmatrix} 0.48 \\ 0.51 \\ 0.01 \end{bmatrix} = \begin{bmatrix} 0.390 \\ 0.507 \\ 0.103 \end{bmatrix}$$

$$T = \begin{array}{cc} \text{Party} & \begin{array}{ccc} D & R & I \end{array} \\ \begin{array}{c} D \\ R \\ I \end{array} & \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \end{array}$$

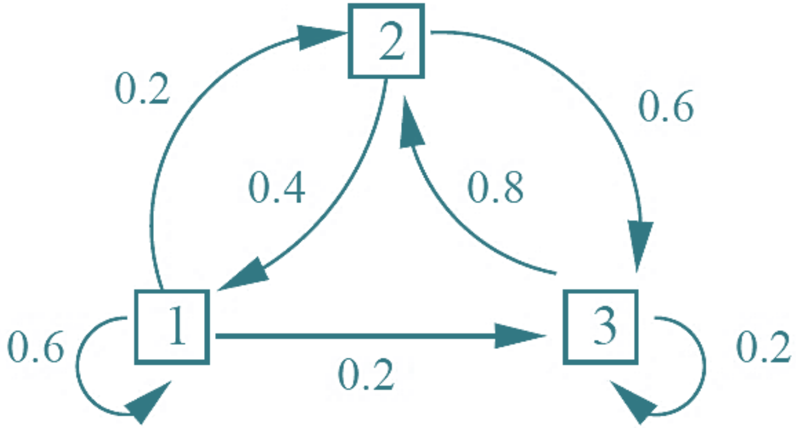
- More explicitly, 70% of the original Democrats (48%) remain Democrat, 10% of the 51% Republican population will vote Democrat, and 30% of the 1% Libertarian population also vote Democrat:
  - $0.70(0.48) + 0.10(0.51) + 0.30(0.01) = 0.390$
- The same goes for the shift of votes to the Republican and Libertarian parties.

Example 2

- A Markov chain has three states, **1**, **2** and **3**. The probabilities of going from one state to another are shown in the following diagram.

(i) Find the transition matrix $\boldsymbol{T}$.

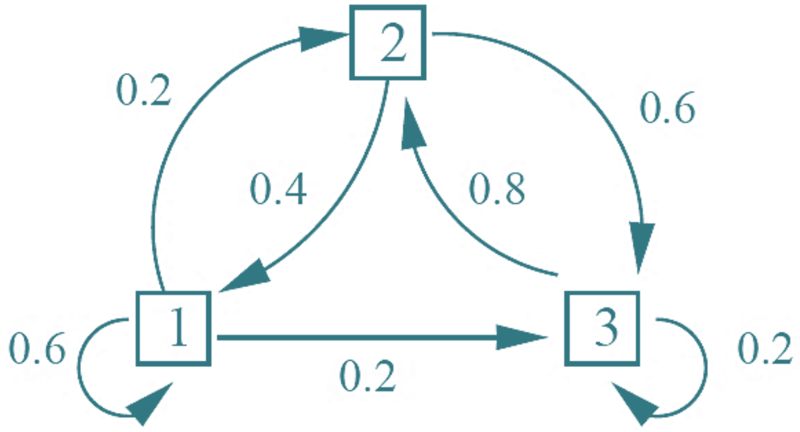
(ii)Find $\boldsymbol{S_1}$ for the initial-state $\boldsymbol{S_0} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}$.



Example 2 - Solution

(i) Find the transition matrix $\boldsymbol{T}$.

(ii)Find $\boldsymbol{S_1}$ for the initial-state $\boldsymbol{S_0} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}$.

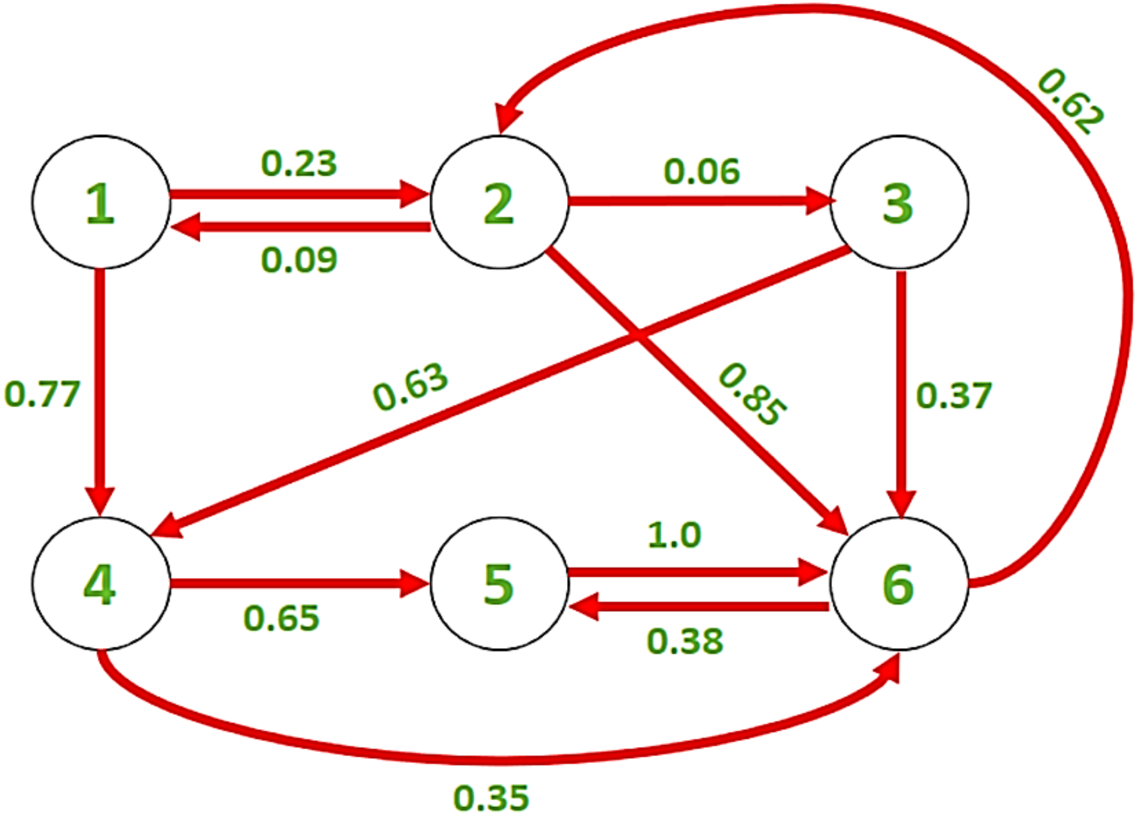


- ❖ $\boldsymbol{T} = \begin{bmatrix} 0.6 & 0.4 & 0 \\ 0.2 & 0 & 0.8 \\ 0.2 & 0.6 & 0.2 \end{bmatrix}$

- ❖ $\therefore \boldsymbol{S_1} = \boldsymbol{TS_0} = \begin{bmatrix} 0.6 & 0.4 & 0 \\ 0.2 & 0 & 0.8 \\ 0.2 & 0.6 & 0.2 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.1 \\ 0.4 \end{bmatrix}$

Example 3

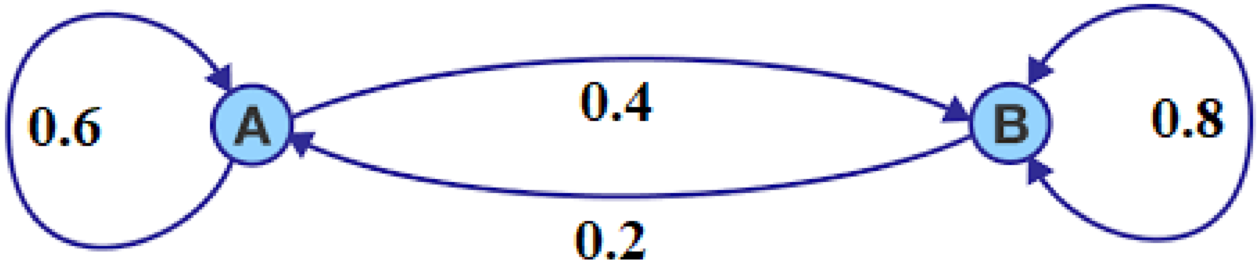
- A Markov chain has six states. The probabilities of going from one state to another are shown in the following diagram. **Write** the transition matrix $\boldsymbol{T}$.
- Solution:



- $$\boldsymbol{T} = \begin{bmatrix} 0 & 0.09 & 0 & 0 & 0 & 0 \\ 0.23 & 0 & 0 & 0 & 0 & 0.62 \\ 0 & 0.06 & 0 & 0 & 0 & 0 \\ 0.77 & 0 & 0.63 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.65 & 0 & 0.38 \\ 0 & 0.85 & 0.37 & 0.35 & 1 & 0 \end{bmatrix}$$

Example 4

- A Markov chain has two states, **A** and **B**. The probabilities of going from one state to another are shown in the following diagram.



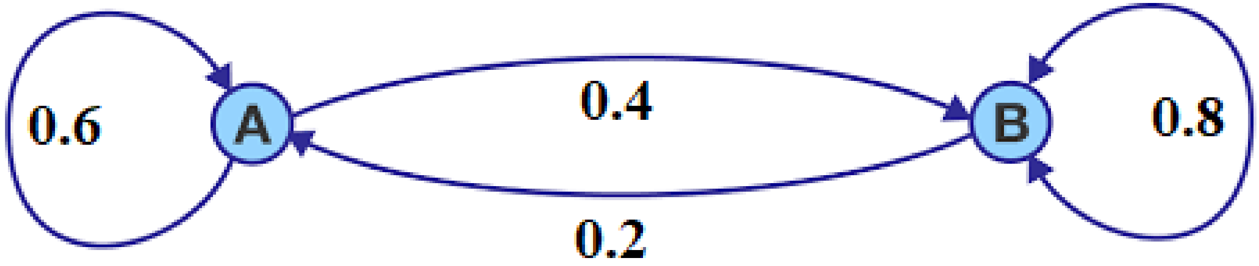
(i) Find the transition matrix $\boldsymbol{T}$.

(ii) Compute $\boldsymbol{S_{20}}$ for the initial-state $\boldsymbol{S_0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Example 4 - Solution

(i) Find the transition matrix $\boldsymbol{T}$.

(ii)Compute $\boldsymbol{S_{20}}$ for the initial-state $\boldsymbol{S}$



- $\boldsymbol{T} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix}$

- $\boldsymbol{S_{20}} = \boldsymbol{T^{20}}\ \boldsymbol{S_0}$

- We will use the **diagonalization** process to find $\boldsymbol{T^{20}}$.

Diagonalization – Step 1

- **Step 1:** Find the eigenvalues of T.

$$|T - \lambda \boldsymbol{I}| = \begin{vmatrix} (0.6-\lambda) & 0.2 \\ 0.4 & (0.8-\lambda) \end{vmatrix}$$

$$= (0.6-\lambda)(0.8-\lambda) - 0.08 = \lambda^2 - 1.4\lambda + 0.4 = 0.$$

- We see that the eigenvalues are $\lambda = 1$ and $\lambda = 0.4$.

Diagonalization – Step 2

- **Step 2:** Find the corresponding eigenvectors of T.

$$\boxed{(\boldsymbol{T}-\lambda\boldsymbol{I})X=\boldsymbol{O}} \Longrightarrow \begin{bmatrix}(0.6-\lambda) & 0.2\\ 0.4 & (0.8-\lambda)\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\xRightarrow{\lambda=1} \begin{bmatrix}-0.4 & 0.2\\ 0.4 & -0.2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\left[\begin{array}{cc|c}1 & -0.5 & 0\\ 0.4 & -0.2 & 0\end{array}\right] \xRightarrow{-0.4R_1+R_2} \left[\begin{array}{cc|c}1 & -0.5 & 0\\ 0 & 0 & 0\end{array}\right]$$

- Let $\boldsymbol{y}=\boldsymbol{t}$, we find $x=0.5t$.

- This yields that the eigenvectors corresponding to $\lambda=1$ are

$$X_1=\begin{bmatrix}0.5t\\ t\end{bmatrix}=t\begin{bmatrix}0.5\\ 1\end{bmatrix}$$

Diagonalization – Step 2

- **Step 2:** Find the corresponding eigenvectors of T.

$$\boxed{(\boldsymbol{T}-\boldsymbol{\lambda I})\boldsymbol{X}=\boldsymbol{O}} \Longrightarrow \begin{bmatrix}(0.6-\lambda) & 0.2\\ 0.4 & (0.8-\lambda)\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\xRightarrow{\boldsymbol{\lambda=0.4}} \begin{bmatrix}0.2 & 0.2\\ 0.4 & 0.4\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\left[\begin{array}{cc|c}1 & 1 & 0\\ 0.4 & 0.4 & 0\end{array}\right] \xRightarrow{-0.4R_1+R_2} \left[\begin{array}{cc|c}1 & 1 & 0\\ 0 & 0 & 0\end{array}\right]$$

- Let $\boldsymbol{y}=\boldsymbol{t}$, we find $x=-t$.
- This yields that the eigenvectors corresponding to $\lambda=0.4$ are

$$X_2=\begin{bmatrix}-t\\ t\end{bmatrix}=t\begin{bmatrix}-1\\ 1\end{bmatrix}$$

Diagonalization – Steps 3 & 4

$$X_1 = \begin{bmatrix} 0.5t \\ t \end{bmatrix} = t\begin{bmatrix} 0.5 \\ 1 \end{bmatrix}, \qquad X_2 = \begin{bmatrix} -t \\ t \end{bmatrix} = t\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

- **Step 3**: Form an eigenvector matrix ***P*** whose columns are these 2 eigenvectors.

$$\boldsymbol{P} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

- **Step 4:** Form the eigenvalue matrix ***D*** (diagonal matrix) whose main diagonal elements are the eigenvalues. ($\lambda_1 = 1$ and $\lambda_2 = 0.4$)

$$\boldsymbol{D} = \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}$$

- Note the **order** of the eigen values should be corresponding to that of the eigenvectors.

Diagonalization – Step 5

- **Step 5**: Finish the diagonalization.
- Finally, we can diagonalize the matrix $\boldsymbol{T}$ as $\boldsymbol{T = PDP^{-1}}$. Hence,

$$\boldsymbol{T^{20} = PD^{20}P^{-1}}$$

$$\therefore \boldsymbol{T^{20}} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (1)^{20} & 0 \\ 0 & (0.4)^{20} \end{bmatrix} \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$$

$$\boldsymbol{T^{20}} \cong \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix}$$

$$\therefore \boldsymbol{S_{20} = T^{20}\, S_0} = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}$$

Diagonalization – Remarks

- Note that:

- $\begin{bmatrix} a & c \\ d & b \end{bmatrix}^{-1} = \frac{1}{a\,b-c\,d}\begin{bmatrix} b & -c \\ -d & a \end{bmatrix}$

- In addition, the sum of the eigenvalues equals the sum of the diagonal entries, this means that

$$\lambda_1 + \lambda_2 = a + b.$$

- Remarkably, it can be shown that *any* transition matrix obeying conditions 1 and 2 must have $\lambda = 1$ as an eigenvalue.

Marcov Chains – Final Comment

- Finally, we can conclude that the application of linear algebra and matrix methods to Markov chains provides an efficient means of monitoring the progress of a dynamical system over discrete time intervals.

Thank You ☺