MAT 222 Linear Algebra
Week 12
Lecture Notes 1

Murat Karaçayır

Akdeniz University
Department of Mathematics

20th May 2025

Markov Processes

- Suppose that citizens in a country are classified according to their votes in the last election: Some proportion are democrat, some proportion are republican, and the rest are libertarian.
- Assume that the probabilities that each voter will vote in a given way in the next election are known. For example, a democrat will vote for Democrat Party with a probability 0.7, for Republican Party with probability 0.2 and for Libertarian Party with probability 0.1.
- Suppose at the initial time, the distribution of the votes are as follows: 55 percent D, 40 percent R, 5 percent L. this can be represented by the vector

$$\mathbf{x}_0 = \begin{pmatrix} 0.55 \\ 0.40 \\ 0.05 \end{pmatrix}$$

- This is know as the **state** of the system at time 0.
- At time 1, the state will be

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{pmatrix} 0.7 & 0.1 & 0.3 \\ 0.2 & 0.8 & 0.3 \\ 0.1 & 0.1 & 0.4 \end{pmatrix} \begin{pmatrix} 0.55 \\ 0.40 \\ 0.05 \end{pmatrix} = \begin{pmatrix} 0.440 \\ 0.445 \\ 0.115 \end{pmatrix}$$

- Here, P is known as a **transition matrix** or **stochastic matrix**
- Thus, in the next election 44 percent will vote for D, 44.5 percent will vote for R and 11.5 percent will vote for L.
- The state at time k (k-th election) can be calculated by $\mathbf{x}_k = P\mathbf{x}_{k-1}$.
- Such processes are called **Markov process** or **Markov chain**.

Markov Processes

- Markov processes can also be though of as acting over individual elements.
- For example, suppose that a citizen has voted for Democrats in the last election. We can represent this fact by the state matrix

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

- Then, the probability that it will again vote for Democrats in the next election can be read off from $\mathbf{x}_1$ as follows:

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{pmatrix} 0.7 \\ 0.2 \\ 0.1 \end{pmatrix}$$

- This shows that at time 1 (the next election), he will again vote for Democrats with probability 0.7, for Republicans with probability 0.2 and for Libertarians with probability 0.1.
- Similarly, the probability distribution three elections later is

$$\mathbf{x}_3 = P^3\mathbf{x}_0 = \begin{pmatrix} 0.450 \\ 0.411 \\ 0.139 \end{pmatrix}$$

- Thus, he/she will again vote for Democrats with probability 0.45, and so on.

Steady-State of a Markov Chain

- An interesting question about a Markov Process is this: What will be the long term behavior of the system? Given a certain initial state, what is the probability distribution of states in the long run?
- We had come across a similar problem in the study of population models with predator-prey interactions.
- Note that once the system has reached a state $\mathbf{x}$ such that $P\mathbf{x} = \mathbf{x}$, it will forever stay in that state.
- This means that 1 is an eigenvalue of P and $\mathbf{x}$ is the eigenvector corresponding to 1, with the property that its entries add up to 1. (It is known that 1 is always an eigenvalue of a stochastic matrix)
- For example, for the voters' problem eigenvectors of P corresponding to eigenvalue 1 are of the form $t(9, 15, 4)$. Scaling it we obtain $\mathbf{x} = (\frac{9}{28}, \frac{15}{28}, \frac{4}{28})$ as the steady-state.
- Thus, in the long run, approximately 32 percent will vote for Democrats, 54 percent for Republicans and 14 percent for Libertarians. Note that these proportions do not depend on the initial state.

Steady-State of a Markov Chain

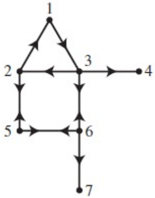
- An important question is this: Does every Markov process converge to a steady state?
- Not every one of them. For example the Markov process with the stochastic matrix

$$P = \begin{pmatrix} 1 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 1 \end{pmatrix}$$

  does not converge to its steady state.
- On the other hand, if the stochastic matrix P is **regular**, then the process converges to the steady state. A matrix P is said to be regular if all the entries in P^k are positive for some value of k.

Random Walks on Directed Graphs

- Many processes can be modeled by directed graphs.
- As an example, consider the directed graph below.



- Some edges have arrows on them and some do not.
- If vertex i has an arrow directed away from it to m different vertices, then the probability of moving to one of these states from state i is $1/m$.
- The transition matrix of this directed graph can easily be constructed.
- This type of Markov processes are called a **random walk**.
- Note that there is no exit from state 7 in this graph. It is called a **dangling node**.
- This random walk will not converge to its steady state. (Why not?)

World Wide Web and Google's PageRank Index

- The Internet (World Wide Web) can also be modeled as a directed graph consisting of n vertices. (There are n web pages in total)
- According to this model, each web page is repesented by a vertex. If web page i has a hyperlink to web page j, this is represented by an arrowed edge from web page i towards web page j.
- If a web page does not have any hyperlinks to other web pages, it is a dangling node according to this model.
- Google's founders Sergey Brin and Lawrence Page devised an algorithm in 1998 to rank web pages with respect to their importance based on the following rough evaluation: If a web page has many hyperlinks directed to it, this web page is important; otherwise it is not.
- According to this approach, the steady state of the Markov process corresponding to World Wide Web will tell the ordering of the available web pages with respect to their importance.

World Wide Web and Google's PageRank Index

- In order to deal with the problem of dangling nodes, they made this arrangement: If a surfer has moved to a dangling website (it does not have any hyperlink to another web site), the he/she will pick any page in the Web with equal probability and move to that page. This replaces the corresponding column with a column consisting of all $1/n$ in the transition matrix.
- In order to deal with cycles in which two pages only hyperlink to each other, a second adjusment is made as follows: If the surfer is at web page j, the event that it chooses a hyperlinked web page is assigned a probability p. The surfer then chooses a random web page (including the hyperlinked ones) with probability $1-p$.
- These adjusments updates the transition matrix as follows: If it was P^* before the second arrangement, it is

  $$G = p\,P^* + (1-p)K,$$

  finally, where K is the $n \times n$ matrix whose every entry is $1/n$.
- The matrix G is called the **Google matrix** and its every entry is nonzero. Therefore it will converge to its steady state.

Google's PageRank: Example

- Consider a miniature version of Web corresponding to the 7-vertex directed graph shown before.
- The stochastic matrix of this graph is

$$P* = \begin{pmatrix} 0 & 1/2 & 0 & 1/7 & 0 & 0 & 1/7 \\ 0 & 0 & 1/3 & 1/7 & 1/2 & 0 & 1/7 \\ 1 & 0 & 0 & 1/7 & 0 & 1/3 & 1/7 \\ 0 & 0 & 1/3 & 1/7 & 0 & 0 & 1/7 \\ 0 & 1/2 & 0 & 1/7 & 0 & 1/3 & 1/7 \\ 0 & 0 & 1/3 & 1/7 & 1/2 & 0 & 1/7 \\ 0 & 0 & 0 & 1/7 & 0 & 1/3 & 1/7 \end{pmatrix}$$

- After applying the adjusments with $p = 0.85$ (This is the current value used by Google), we obtain the regular Google matrix (for this example) G, whose steady state is

$$\mathbf{x} = \begin{pmatrix} .116293 \\ .168567 \\ .191263 \\ .098844 \\ .164054 \\ .168567 \\ .092413 \end{pmatrix}$$

- Thus, the web pages are ranked from most important to least important as follows: Page 3, Page 2 $=$ Page 6, Page 5, Page 1, Page 4, Page 7.