Lecture 17: Markhov Chains

Example:

(a) If today is sunny there is a 70% chance that tomorrow will be sunny.

(b) If today is cloudy the chances are 80% that tomorrow will be cloudy.

Suppose today (Saturday) is sunny what is the probability that next Saturday's weather is also sunny.

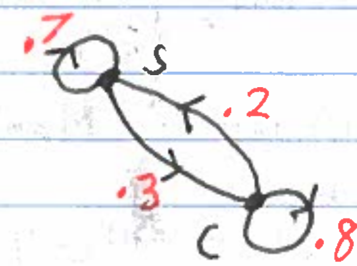
Idea:

Let

$s^{(K)}$ = probability that day K is sunny

$c^{(K)}$ = probability that day K is cloudy

Properties of Probability



1. $s^{(K)} + c^{(K)} = 1$
2. $s^{(K+1)} = .7s^{(K)} + .2c^{(K)}$

   $c^{(K+1)} = .3s^{(K)} + .8c^{(K)}$

Let $\vec{v}^{(K)} = \begin{bmatrix} s^{(K)} \\ c^{(K)} \end{bmatrix}$ it follows that

$$\vec{v}^{(K+1)} = \begin{bmatrix} s^{(K+1)} \\ c^{(K+1)} \end{bmatrix} = \begin{bmatrix} .7 & .2 \\ .3 & .8 \end{bmatrix} \begin{bmatrix} s^{(K)} \\ c^{(K)} \end{bmatrix}$$

transition matrix

state vector

$$\begin{array}{c c} & \begin{array}{cc} S & C \end{array} \\ \begin{array}{c} S \\ C \end{array} & \begin{bmatrix} .7 & .2 \\ .3 & .8 \end{bmatrix} \end{array}$$

Sat: $\vec{v}^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Sun: $\vec{v}^{(2)} = \begin{bmatrix} .7 & .2 \\ .3 & .8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .7 \\ .3 \end{bmatrix}$

Mon: $\vec{v}^{(3)} = \begin{bmatrix} .7 & .2 \\ .3 & .8 \end{bmatrix} \begin{bmatrix} .7 \\ .8 \end{bmatrix} = \begin{bmatrix} .55 \\ .45 \end{bmatrix}$

Tue: $\vec{v}^{(4)} = \begin{bmatrix} .475 \\ .525 \end{bmatrix}$

Wed: $\vec{v}^{(5)} = \begin{bmatrix} .438 \\ .563 \end{bmatrix}$

Th: $\vec{v}^{(6)} = \begin{bmatrix} .419 \\ .581 \end{bmatrix}$

Fr: $\vec{v}^{(7)} = \begin{bmatrix} .410 \\ .591 \end{bmatrix}$

Sat: $\vec{v}^{(8)} = \begin{bmatrix} .405 \\ .595 \end{bmatrix}$

The iterates converge fairly rapidly to $\vec{v}^{*} = \begin{bmatrix} .4 \\ .6 \end{bmatrix}$ which is called the stationary distribution.

* Why does this happen??

Definition - A vector $\vec{v} \in \mathbb{R}^n$ is called a probability vector if each entry v_i satisfies $0 \le v_i \le 1$ and $v_1 + \dots + v_n = 1$.
(v_i represents probability of being in state i).

Definition - A Markov chain is represented by a linear system

$$\vec{v}^{(k+1)} = T\vec{v}^{(k)}$$

satisfying

1. $\vec{v}^{(1)}$ is a probability vector
2. $0 \le T_{ij} \le 1$,
3. $T_{1j} + T_{2j} + \dots + T_{nj} = 1$.

(T_{ij} represents probability of transitioning from state j to state i)

We call T a transition matrix.

Theorem - If T is a transition matrix then there exists a stationary state $\vec{v}^*$ such that $T\vec{v}^* = \vec{v}^*$.

proof:
Let $M = T^T$. Therefore, M and T have the same eigenvalues and the sum of each row of M is
$\therefore$ Consequently,

$$M \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} M_{11} + M_{12} + \cdots + M_{1n} \\ \vdots \\ M_{n1} + M_{n2} + \cdots M_{nn} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

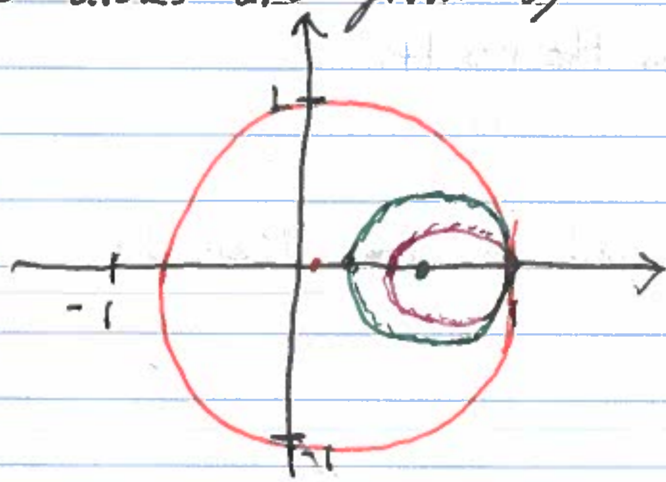
■

Theorem - The eigenvalues of T satisfy $|\lambda_j| < 1$.

proof:
Let $M = T^T$. Each Gershgorin disk satisfies

$$D_i = \{ z : |z - M_{ii}| < 1 - M_{ii} \}$$

Since the rows of M add up to one. Consequently, the disks are given by



■

<u>Theorem</u> — If only one eigenvalue of T satisfies $|\lambda^*| = \lambda^* = 1$ and all other eigenvalues are distinct then

$$\lim_{k\to\infty} \vec{U}^{(k+1)} = \lim_{k\to\infty} T\vec{U}^{(k)} = \lim_{k\to\infty} T^k \vec{U}^{(1)} = \vec{U}^*.$$

~~proof~~:

Since the eigenvalues of T are distinct the eigenvectors $\vec{U}^*, \vec{V}_2, \ldots, \vec{V}_n$ with corresponding eigenvalues $1, \lambda_2, \lambda_3, \ldots, \lambda_n$ are linearly independent and thus form a basis. Therefore,

$$\begin{aligned}
\vec{U}^{(k+1)} &= T\vec{U}^{(k)} \\
&= T^k \vec{U}^{(1)} \\
&= T^k (c_1 \vec{U}^* + c_2 \vec{V}_2 + \ldots + c_n \vec{V}_n) \\
&= c_1 T^k \vec{U}^* + c_2 T^k \vec{V}_2 + \ldots + c_n T^k \vec{V}_n \\
&= c_1 \vec{U}^* + c_2 \lambda_2^k \vec{V}_2 + \ldots + c_n \lambda_n^k \vec{V}_n.
\end{aligned}$$

Taking the limit we have

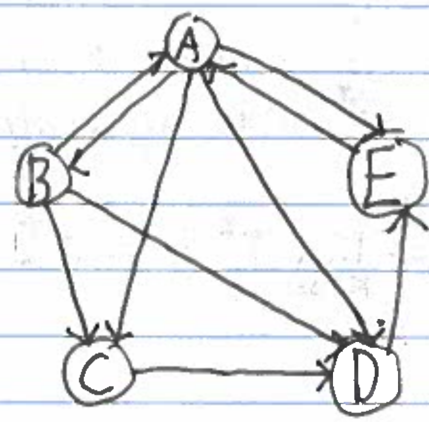
$$\lim_{k\to\infty} \vec{U}^{(k+1)} = c_1 \vec{U}^*$$

Normalizing, we obtain the result.

(This technique is called the Power Method)

Example (Page Rank):

Consider the following internet



- On a website A, B, C, D, E each link is followed with equal probability

$$T = \begin{array}{c|ccccc} & A & B & C & D & E \\ \hline A & 0 & \frac{1}{3} & 0 & 0 & 1 \\ B & \frac{1}{4} & 0 & 0 & 0 & 0 \\ C & \frac{1}{4} & \frac{1}{3} & 0 & 0 & 0 \\ D & \frac{1}{4} & \frac{1}{3} & 1 & 0 & 0 \\ E & \frac{1}{4} & 0 & 0 & 1 & 0 \end{array}$$

An associated quantity is the adjacency matrix which tracks if nodes are connected

$$A = \begin{array}{c|ccccc} & A & B & C & D & E \\ \hline A & 0 & 1 & 0 & 0 & 1 \\ B & 1 & 0 & 0 & 0 & 0 \\ C & 1 & 1 & 1 & 0 & 0 \\ D & 0 & 1 & 0 & 0 & 0 \\ E & 1 & 0 & 0 & 1 & 0 \end{array}$$

Definition- A transition matrix is regular if some power T^K contains no zero entries, then it is regular.

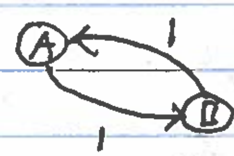
Theorem- If T is a regular transition matrix, then it admits a unique stationary probability $\vec{U}^*$ with eigenvalue $\lambda = 1$.
Moreover, for all initial states

$$\lim_{K\to\infty} \vec{U}^{(K+1)} = \lim_{K\to\infty} T\vec{U}^{(K)} = \lim_{K\to\infty} T^K \vec{U}^{(1)} = \vec{U}^*.$$

Example:

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

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



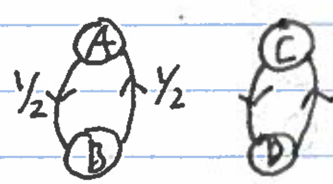
$$T^3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow$$ Not transitive. No stationary probability

$\vdots$

(T is a periodic matrix).

Example:

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



$$T^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \rightarrow$$ Not transitive

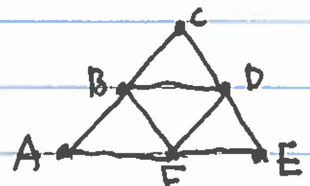
There are three stationary probabilities:

$$\vec{U}_1^* = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix}, \quad \vec{U}_2^* = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \vec{U}_3^* = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$$

*Depending on $\vec{U}^{(1)}$, can limit to any of these states.

Example:

An insect crawls along the edges of the following triangular lattice



Upon arriving at a vertex, there is an equal probability of moving to another vertex. Determine, on average, how often each vertex is visited.

The transition matrix is given by

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} 0 & 1/4 & 0 & 0 & 0 & 1/4 \\ 1/2 & 0 & 1/2 & 1/4 & 0 & 1/4 \\ 0 & 1/4 & 0 & 1/4 & 0 & 0 \\ 0 & 1/4 & 1/2 & 0 & 1/2 & 1/4 \\ 0 & 0 & 0 & 1/4 & 0 & 1/4 \\ 1/2 & 1/4 & 0 & 1/4 & 1/2 & 0 \end{bmatrix} \end{array} = A.$$

Logically, we should be twice as likely to be on vertex B, D, or F than on A, C, or E. If we let p be probability on corner of triangle then

$$3p + 6p = 1$$
$$\Rightarrow p = 1/9$$

The stationary vector is

$$\vec{v}^* = \begin{bmatrix} 1/9 \\ 2/9 \\ 1/9 \\ 2/9 \\ 1/9 \\ 2/9 \end{bmatrix}$$