Markov Chains

We've studied the "gambler's ruin" problem of ~~estimating~~ computing the probability that a gambler who wins 1\$ on each flip of heads and loses 1\$ on each flip of tails will reach N dollars before reaching 0 dollars if they start with K.

Recall: The answer is $\frac{K}{N}$.

Definition. A discrete Markov chain consists of a finite state space $1, \ldots, n$ and an $n \times n$ matrix P so that

$$P_{ij} = \text{probability of transition from state } i \text{ to state } j$$

Note: $\sum_j P_{ij} = 1$ for all i.

Example. The gambler's ruin chain. State space $\{0, \dots, n\}$ indicates # of dollars held by gambler.

$$P = \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ 1/2 & 0 & 1/2 & & \vdots \\ \vdots & 1/2 & 0 & 1/2 & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & & & 1 \end{bmatrix}$$

Example. A student ~~on a~~ ~~n~~ ~~carousel with~~ sitting at a circular table with 5 ~~8~~ chairs has equal probability of moving cw or ccw.

Give state space and P.

Observe: ~~the~~ Given a probability distribution $\vec{x} \in \mathbb{R}^n$ on state space,

$$P\vec{x}, P^2\vec{x}, \ldots, P^k\vec{x}, \ldots$$

are the distributions after k steps.

Question: Does this converge?

Definition. A markov chain is <u>irreducible</u> if for every i,j we have an n_{ij} so

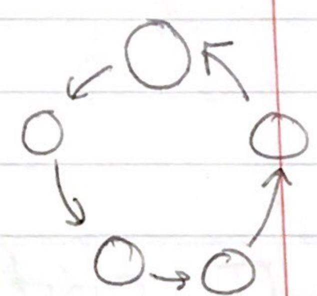
$$(P^{n_{ij}})_{ij} > 0$$

(That is every state can be reached from every other state.)

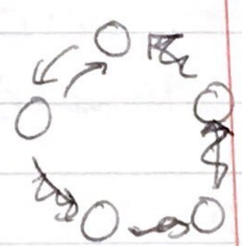
Definition. A state i is <u>aperiodic</u> if

$$\gcd\{n > 0 \mid P_{ii}^{n} > 0\} = 1.$$

Example. The 5 chair round table is aperiodic



$P_{ii}^{5} > 0$



$P_{ii}^{2} > 0$

Since $\gcd\{5, 2\} = 1$, the chain is aperiodic.

What about the 4 chair table?

Theorem. An irreducible aperiodic Markov chain has a unique limiting distribution π.

1) $\lim_{k \to \infty} P^k x = \pi$, for <u>any</u> p.d. x

2) $P\pi = \pi$, (all other eigenvectors have smaller evals)

Proof. Perron-Frobenius theorem. (Not easy!)

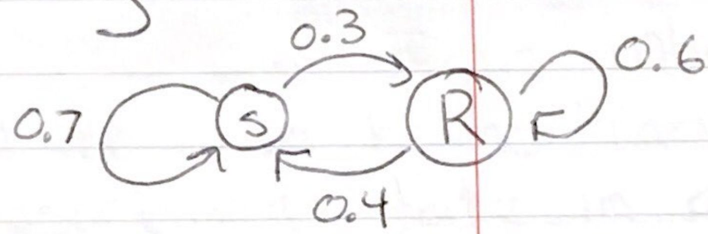
~~[illegible]~~

Application. The limiting distribution of the round table is uniform (for odd # seats)

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \cancel{0}\,\frac{1}{2} & 0 & 0 & \underset{1/2}{\cancel{0}} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

So the limiting dist. is uniform!

Example. Weather data tends to be correlated in time - ~~sun~~ if today is sunny, tomorrow is more likely to be sunny than rainy.



This diagram shows transition probabibities between states.

Find the limiting distribution of weather in this model.

(Steps. Write matrix. Find Perron eigenvector.)

We can't analyze the gambler's ruin because it's not irreducible: states 1 and n only transition to themselves!

Definition. An absorbing state is one that only transitions to itself.

An absorbing Markov chain is a Markov chain where each state can transition to an absorbing state in a finite number of steps.

(A state which is not absorbing is called a transient state.)

We can write the transition matrix for an absorbing chain ~~as~~ with t transient states, r abs. states

$$P = \begin{bmatrix} Q & R \\ 0 & I_r \end{bmatrix}$$

where Q is $t \times t$ and I_r is the $r \times r$ identity matrix.

Proposition. The probability of being in absorbing state j after K steps, starting in transient state i is

$$\left[\left(\sum_{i=0}^{K} Q^{k}\right) R\right]_{ij}$$

Proof. Let's examine

$$P = \begin{bmatrix} Q & R \\ 0 & I_r \end{bmatrix}$$

$$P^2 = \begin{bmatrix} Q^2 & QR + R \\ 0 & I_r \end{bmatrix}$$

$$P^3 = \begin{bmatrix} Q^3 & Q^2R + QR + R \\ 0 & I_r \end{bmatrix}$$

$$P^K = \begin{bmatrix} Q^K & \left(\sum_{i=0}^{K} Q^i\right) R \\ 0 & I_r \end{bmatrix}$$

$\square$

~~Corollary 4~~

Proposition. If we start at transient state i, the probability of ending in absorbing state j (eventually) is

$$\lim_{K\to\infty}\left(\sum_{i=0}^{K} Q^i\right)R = \left(\sum_{i=0}^{\infty} Q^i\right)R$$

Proof. We need only prove that

$\sum_{i=0}^{\infty} Q^i$ converges.

Theorem. The geometric series generated by an $n\times n$ matrix A converges $\Longleftrightarrow$ each eigenvalue λ_i of A ~~is~~ has $|\lambda_i|<1$.

In this case $I-A$ is invertible and

$$(I-A)^{-1} = \sum_{K=0}^{\infty} A^K$$

Since the markov chain is absorbing and there are a finite # of states, we may assume there is some power K so that

$$\left(\sum_{i=0}^{K} Q^{i}\right) R$$

has a positive entry in each row. Now

$$P^{K} = \begin{bmatrix} Q^{K} & \left(\sum_{K=0}^{K} Q^{i}\right) R \\ 0 & I_r \end{bmatrix}$$

is still a transition matrix, so its row sums are 1. Thus the row sums of Q^K are all < 1.

Now Q is a ~~positive~~ nonnegative matrix so Q^K is nonnegative as well.

We need

Definition.

~~Gershgorin Circle Theorem.~~

If A is a square matrix let $R_i = \sum_{j \neq i} |A_{ij}|$ be the sum of absolute values of entries away from diagonal in ith row.

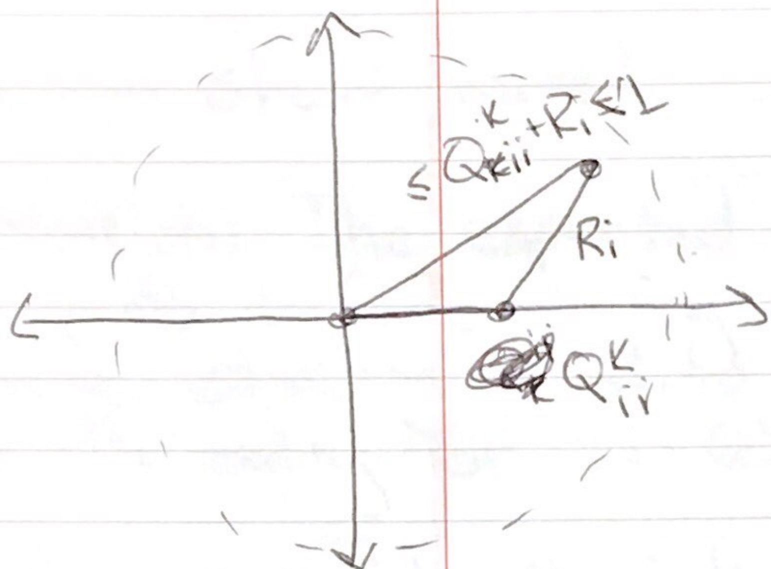
~~The~~ The Gershgorin disks of A are the disks in $\mathbb{C}$ with center A_{ii} and radius R_i.

Gershgorin Circle Theorem.
Every eigenvalue of A lies in a Gershgorin disk.

For Q^k, we know that $Q^k_{ij} \geq 0$ and

$$\sum_j Q^k_{ij} < 1$$

so $0 \leq Q^k_{ii} + R_i \leq 1$. Thus the Gershgorin disk lies inside the disk of radius 1.



Thus all eigenvalues of Q^k (and hence all eigenvalues of Q) have $|\lambda_k| < 1$. $\square$

We have shown

The probability of ending in absorbing state j, starting in transient state i is

$$(I-Q)^{-1} R_{ij}$$

which is pretty amazing!

We can show more!

Proposition. The expected number of steps before being absorbed when starting in state i is the ith entry of $(I-Q)^{-1}\vec{1}$.

Proof. Let $X^{(k)}$ be the indicator for "the chain is in transient state j after k steps".

We saw above that the transition probability between i and j after k steps is $(Q^k)_{ij}$. So

$$E(X^{(k)}) = (Q^k)_{ij}.$$

Now the number of times the chain visits j (starting in i) is given by

$$X^{(0)} + X^{(1)} + \dots + X^{(k)} + \dots$$

We know

$$E\left(\sum_{k=0}^{\infty} X^{(k)}\right) = \sum_{k=0}^{\infty} E(X^{(k)})$$

$$= \sum_{k=0}^{\infty} (Q^k)_{ij} = \left(\sum_{k=0}^{\infty} Q^k\right)_{ij}$$

$$= (I - Q)^{-1}{}_{ij}$$

We now think about the # of steps before absorption.

~~Ever~~ The # of steps before absorption (starting in i) is the sum of # steps in state j (over all j). By linearity of expect.