Strategy 4: There's a notion weaker than regularity — Ergodic.

Defn: A Markov chain is ergodic if $\forall i,j \in [n]$ $\exists m$, $(M^m)_{ij} > 0$.

▷ A regular Markov chain is ergodic.

- Converse fails!

- Eg. $M := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow M^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \Rightarrow M^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = M$

$\Rightarrow \ldots$

▷ M is not regular, but is ergodic.

- Qn: Could we turn ergodic into regular?

Lemma: Let $M' := \frac{1}{2}\cdot I + \frac{1}{2}\cdot M$. If M is ergodic then M' is regular.

Pf: • $(M'^{m+1})_{ij} = (M' \cdot M'^{m})_{ij} = \sum_{k=1}^{n} M'_{ik} \cdot (M'^{m})_{kj}$

> 0, if $M'_{ii} \cdot (M'^{m})_{ij} > 0$.

$\Rightarrow$ If, for $i,j \in [n]$: $(M'^{m_{ij}})$ has (i,j)-th entry > 0,
then: for $m := \max_{i,j} m_{ij}$ & $\forall i,j \in [n]$,
$(M'^{m})_{ij} > 0$ simultaneously!

$\Rightarrow$ M' is regular. $\square$

- So, if Internet graph G is <u>connected</u>, then M is ergodic $\implies$ M' is regular.
[G disconnected $\Rightarrow$ M is <u>not</u> ergodic.]

▷ For a <u>connected Internet G</u> the web-surfer either: <u>stays</u> at the current page (with $P=\frac{1}{2}$)
or: <u>moves</u> to a <u>linked</u> page (").

- This <u>explains</u> $\underline{M'} = \frac{1}{2}I + \frac{1}{2}M$ & gives a page-ranking!

- Let's see a <u>harder</u> eg. of Markov Chain:

Cell Genetics (paper in 1958)

- Cell of a certain organism contains N particles of two types — A & B, (genes)
- The cell reproduces by first doubling the particles & then forming a child cell with random N particles.

Qn: What's the gene pool after a long time?

- Define $\underline{p_{jk}} := P(\underbrace{\text{parent has } j}_{E_j} \text{ A-types} \wedge \text{child } \underbrace{\text{has } k}_{E_k} \text{ A-types})$.

- $\underline{\text{States}}$: $E_0, E_1, \dots, E_N$.

▷ $p_{jk} = \binom{2j}{k} \cdot \binom{2N-2j}{N-k} \Big/ \binom{2N}{N}$. ← Hypergeometric distribution

- Transition Matrix $\underline{M} := (p_{jk})_{j,k \in [0..N]}$
is $(N+1) \times (N+1)$. [is a stochastic matrix.]

▷ This makes reproduction a homog. Markov process.

- <u>Defn:</u> $M^n = (p_{jk}^{(n)})_{j,k \in [0 \ldots N]}$, for $n \geq 1$.

▷ The process is <u>not</u> ergodic (<u>not</u> regular).

[Eg. $p_{01}^{(n)} = 0$, $\forall n \geq 1$.]

↳ Stationary distribution may <u>not</u> exist!

▷ A parent in E_j is expected to give birth to a child in the <u>same</u> state.

<u>Pf</u>: • $E[\text{child's state}] = \sum_k p_{jk} \cdot k = \sum_k \frac{\binom{2j}{k} \cdot \binom{2N-2j}{N-k}}{\binom{2N}{N}} \cdot k$

$$= \frac{2j}{\binom{2N}{N}} \cdot \sum_k \binom{2j-1}{k-1} \cdot \binom{2N-2j}{N-k} = \frac{2j \cdot \binom{2N-1}{N-1}}{\binom{2N}{N}} = j. \qquad \square$$

Defn: Such a Markov chain is called <u>Martingale</u>.

▷ $\sum_k p_{jk} \cdot k = j \implies \forall n \geq 1, \ \sum_k p_{jk}^{(n)} \cdot k = j$.

[Exp. state in the <u>descendants</u> is j !]

Pf: $\cdot \sum_k p_{jk}^{(n)} \cdot k =$ j-th-entry in $M^n \cdot \begin{pmatrix} 0 \\ 1 \\ \vdots \\ N \end{pmatrix} = M^{n-1} \cdot \begin{pmatrix} 0 \\ 1 \\ \vdots \\ N \end{pmatrix} = \cdots = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ N \end{pmatrix}$

$\Rightarrow$ j-th-entry in $M^n \cdot \begin{pmatrix} 0 \\ 1 \\ \vdots \\ N \end{pmatrix}$ is j. $\square$

▷ $p_{00} = p_{NN} = 1$, Pf: by defn. $\square$

while other $p_{jk} < 1$.

- E_0 & E_N are absorbing states.

- $\forall k \neq 0, N$: E_k could move to another state; it's not absorbing.

▷ One can decompose the chain M into:

(i) disjoint stationary distributions, each over a subset of $E_1, \dots, E_{N-1}$. ← [closed sets]

(ii) transient states k s.t. $\forall j$, $p_{jk}^{(n)} \to 0$ $(n \to \infty)$.

Qn: What's the prob. of reaching one of these?

– Let's assume that $E_1, \ldots, E_{N-1}$ are transient.

i.e. $p_{jk}^{(n)} \to 0$, $k \in [N-1]$.

▷ Then, $\sum_{k>0} p_{jk}^{(n)} \cdot k = j \Rightarrow p_{jN}^{(n)} \cdot N \to j$ (as $n \to \infty$)

$\Rightarrow p_{jN}^{(n)} \to j/N$, $\forall j \in [N-1]$.

$\Rightarrow p_{j0}^{(n)} \to \left(1 - \frac{j}{N}\right)$.

$\Rightarrow$ ▷ $\begin{cases} j > N/2 \Rightarrow \text{Process tends to } E_N. \\ j < N/2 \Rightarrow \quad \text{"} \quad \text{"} \quad \text{"} \quad E_0. \end{cases}$

$\Rightarrow$ ▷ The dominating gene ultimately wins!