02/23/04 Limit Distributions for Finite State, Homogeneous, Discrete Time, <u>Irreducible</u> Markov Chains

In the long run what fraction of the time is the system (MC) in a particular state.

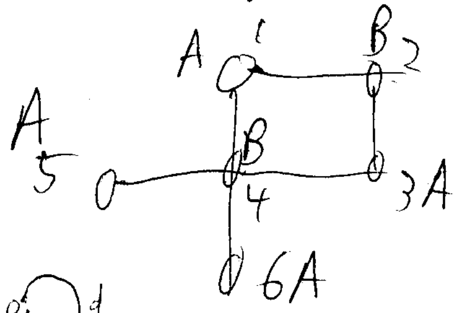
To prepare, need:

<u>Defn</u> The period $d(i)$ of a state $i \in S$ is given by

$$d(i) = \gcd\{n \geq 0, (P^n)_{ii} > 0\}$$

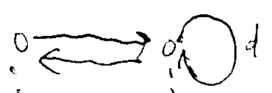
If $d(i) = 1$ then i is said to be aperiodic.

Examples: Random walk on graph (reflecting BC)

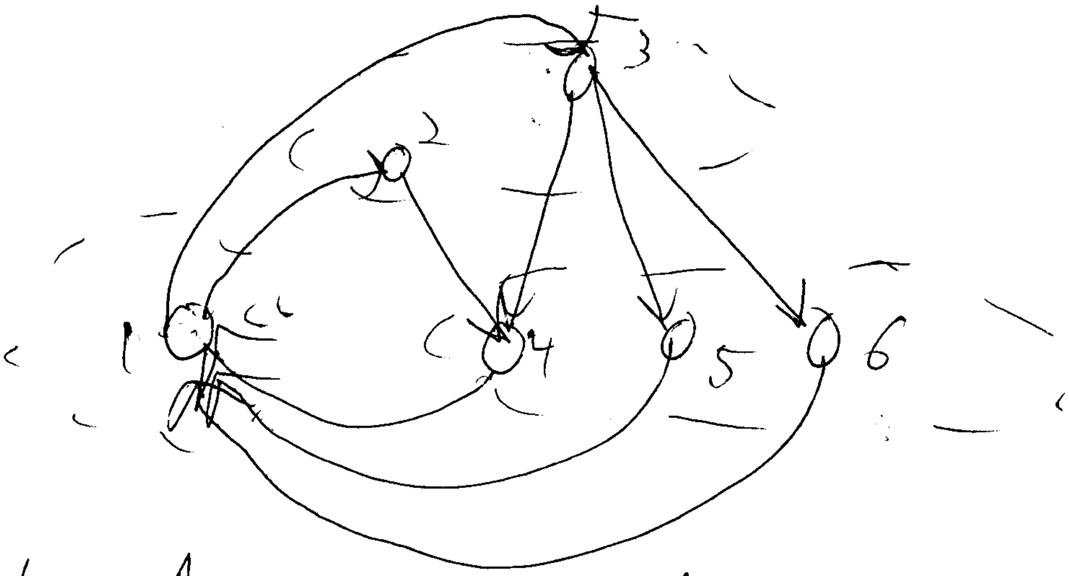


$d(1) = 2$

In fact the period of any state in a given communication class is the same as the period of any other state in the class.



So $d(j) = 2$ for all $j \in S$ in this example.



General idea: A period d Markov chain can be decomposed into d subsets, the MC dynamics proceed from one subset to the next.

We'll focus on <u>aperiodic</u> irreducible MC's for the most part.

<u>Theorem</u>: If $\{X_n\}_{n=0}^{\infty}$ is an aperiodic, irreducible FSDT homogeneous MC w/ ~~initi~~ transition probability matrix P, then it has a unique stationary distribution $\vec{\pi}$ and

$$\lim_{n\to\infty} \text{Prob}(X_n = j) = \pi_j .$$

That is, $\vec{\pi}$ is the <u>limit distribution</u> for the MC.

Idea of Proof:

Proof A): Finite dimensional linear algebra (Lawler Ch. 1)

Thm says $\lim_{n\to\infty} \vec{\phi} \cdot P^n = \vec{\pi}$

$\uparrow$ init. dist.

Why would this be true?

If limit exists, it has to be stat. dist. because $\left(\lim_{n\to\infty} \vec{\phi} \cdot P^n\right) \cdot P = \lim_{n\to\infty} \vec{\phi} \cdot P^{n+1}$

$$= \lim_{n\to\infty} \vec{\phi} \cdot P^n$$

(check that limit exists, normalizable)

Let's write P in Jordan form:

$P = Q J Q^{-1}$ where J has Jordan block form

If diagonalizable $J = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$

eigenvalues $\lambda_1, \dots, \lambda_n$.

$P^n = Q J^n Q^{-1}$

For large n, Jordan blocks with:

$|\lambda| > 1$: amplified

$|\lambda| = 1$: persists

$|\lambda| < 1$: decay

If P^n is to project all vectors $\vec{\phi}$ onto the vector $\vec{\pi}$ (which is an eigenvector with eigenvalue 1), then need... <u>all</u> other eigenvalues satisfy $|\lambda| < 1$.

What do we need to guarantee this?

Irreducibility (otherwise multiple eigenvalues $= 1$)

Aperiodicity avoids eigenvalues λ with $|\lambda| = 1$ but $\lambda \neq 1$.

So given aperiodicity + irreducibility,
how prove this?

Perron-Frobenius Theorem (Karlin + Taylor App. 2
Lawler Ex 1.15)

If matrix A has all entries $A_{ij} > 0$
then there exists an eigenvalue λ
of maximal amplitude such that:
i) λ is real and positive
ii) λ is a simple eigenvalue
iii) The unique eigenvector associated
to λ has all entries ≥ 0.

Use this for the proof of the limit thm:
Irreducible + aperiodic $\Rightarrow$ There is an $N < \infty$
such that for $n \geq N$, $(P^n)_{ij} > 0$ for all $i, j \in S$.

P may have some 0 entries (so can't use
PFT on P)
but P^n has all positive entries for $n \geq N$.

Apply PFT to P^n

P^n has a simple positive real eigenvalue λ which has maximal amplitude.

I also know that $\vec{\pi}$ is an eigenvector of P^n w/ eigenvalue 1.

Is $\lambda = 1$ or is $\lambda > 1$?

Observe that $\sum_{j \in S} (P^n)_{ij} = 1$.

Since the eigenvector $\vec{v}$ corresponding to λ has all nonnegative entries,

$$\underbrace{\left(P^n \vec{v}\right)_i}_{\lambda^n v_i} = \sum_{j \in S} (P^n)_{ij} v_j \le \underbrace{\|v\|_\infty}_{\text{max component of } v} \sum_{j \in S} (P^n)_{ij}$$

$$\lambda^n v_i \le \|v\|_\infty$$

This implies $\lambda^n \le 1 \Rightarrow \lambda \le 1 \Rightarrow \lambda = 1$.

So therefore, $\vec{\pi}$ is the unique eigenvector of eigenvalue $\lambda = 1$, it has all nonnegative entries, and all other eigenvalues μ satisfy $|\mu| < 1$. This shows $\vec{\pi}$ is the limit distribution

Comment: General principle is that amount of time you have to wait for transient effects from initial data to die out and for the MC to reach its limit distribution is n such that $|\lambda_2|^n \ll 1$ where λ_2 is next biggest eigenvalue (after 1).

If λ_2 is near $1 \Rightarrow$ metastable states (C. Schütte, Berlin)

Another remark: One can show that the approach to stat dist. can be described in terms of information theory/entropy.

Define the relative entropy of a prob. dist. $\vec{p}$ with respect to stat dist $\vec{\pi}$ is

$$I(\vec{p}, \vec{\pi}) = \sum_{j \in S} p_j \ln(p_j / \pi_j)$$

This is a Lyapunov function for MC, meaning

$$I(\vec{p}^{(n+1)}, \vec{\pi}) \leq I(\vec{p}^{(n)}, \vec{\pi}) \text{ and}$$

$$\lim_{n \to \infty} I(\vec{p}^{(n)}, \vec{\pi}) = 0$$

Haken, Synergetics

Proof B): Coupling: Trendy and generalizes to ∞-state MC

$\boxed{\text{Resnick 2.13}}$

We know that we have stat. dist. $\vec{\pi}$

$\vec{\pi} \cdot P = \vec{\pi}$.

If we have a MC $\{Y_n\}_{n=0}^{\infty}$ which starts w/ dist. $\vec{\pi}$: $\text{Prob}(Y_0 = j) = \pi_j$ and evolves w/ P, then obviously

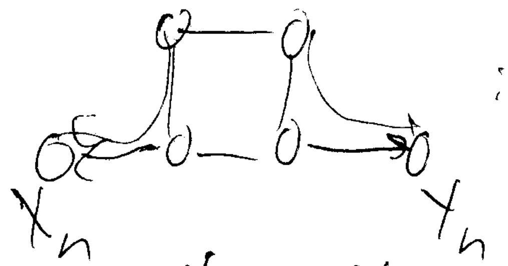
$$\lim_{n \to \infty} \text{Prob}(Y_n = j) = \pi_j.$$

Consider a MC $\{X_n\}_{n=0}^{\infty}$ with same trans matrix P but arbitrary initial dist $\phi_j = \text{Prob}(X_0 = j)$.

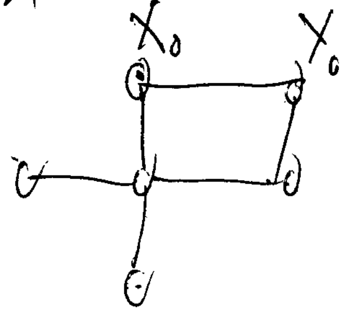
If I can show that $\text{Prob}(X_n = Y_n \text{ for some } n < \infty)$ then the Markov property will show that $\lim_{n \to \infty} \text{Prob}(X_n = j) = \lim_{n \to \infty} \text{Prob}(Y_n = j) = \pi_j$

⋮

How show $\mathrm{Prob}(X_n = Y_n \text{ for some } n < \infty)$?
- aperiodicity + irreducibility



: reducible



: X_n and Y_n can be out-of-phase if periodic.

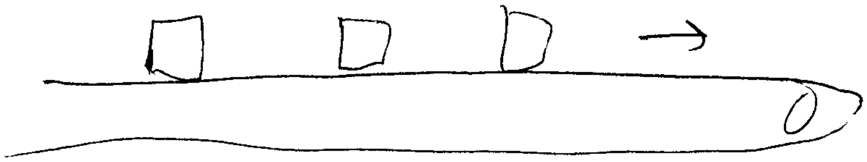
Resnick: View $Z_n = (X_n, Y_n)$

This is a powerful idea:
Goes from existence of a stat. dist. to prove it is the unique, long-time limit of system.

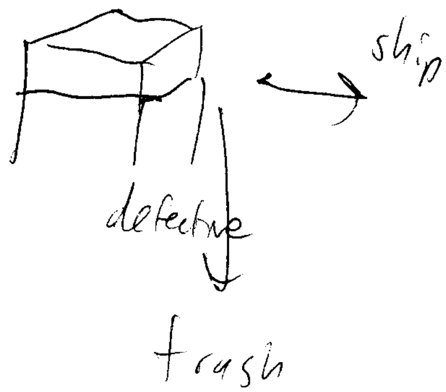
Mattingly: Two-dim. Navier-Stokes eqn

Example: Quality Inspection Protocols

Production line



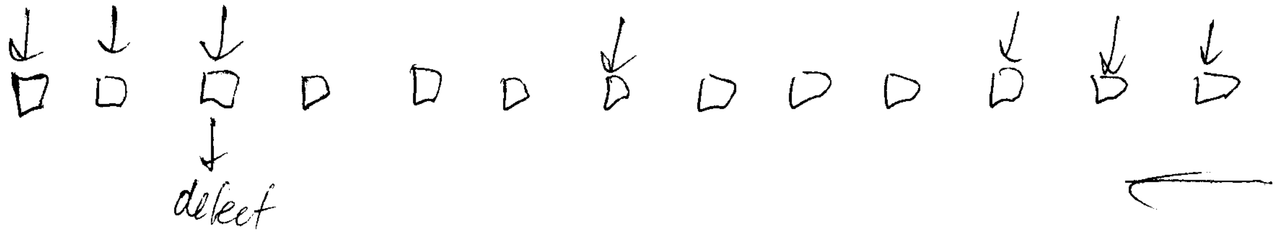
Inspection station



Product ships if not inspected or if inspected but not defective.

Regular adaptive sampling scheme: Start by inspecting every product until you get M good inspections in a row. Then I switch to sampling only 1 in every r products, with regular spacing. Whenever a defect is found, revert to inspecting every product, until M good consecutive products are found again.

$M = 3 \qquad r = 4$



Questions (for any inspection protocol):

1.) How often is an inspected product defective?

→ 2) What fraction of shipped products are defective?

→ 3) What fraction of products are inspected? (effort)

With answers, could try to optimize w.r.t. M and r or compare schemes.

Use Markov chain ideas to make predictions about these questions.

Have to model unpredictability in the status of products.

Simplest version: Each product is defective w/ prob. p, independent of the status of any other product.

Markov Chain Model:

We will choose state space $S = \{0, \dots, M\}$ as describing the state of the machine. We will let X_n be the state of machine after the nth inspection.

$X_n = j$ means after nth inspection, the machine has seen j consecutive good products.

(Don't need to include state of product since it's iid & act as input noise to machine state.)

$$P_{ij} = \mathrm{Prob}(X_{n+1} = j \mid X_n = i)$$

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ \\ M \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & \cdots & & M \end{array} \\ \begin{pmatrix} p & 1-p & 0 & & & 0 \\ p & 0 & 1-p & 0 & & \\ p & 0 & 0 & 1-p & 0 & \\ \vdots & 0 & & & \ddots & 0 \\ p & 0 & & & 0 & 1-p \\ p & 0 & \cdots & & 0 & 1-p \end{pmatrix} \end{array}$$