3) Enough to check that $\vec{r} \cdot P = \vec{r}$

for $r_k = \mathbb{E}[T_{j,k} \mid X_0 = j]$

because $\tilde{\pi} = \dfrac{\vec{r}}{\sum_{k \in S} r_k}$

$\vec{r}$ will thereby be shown to be "invariant measure":

$$r_k \geq 0$$

$$\vec{r} \cdot P = \vec{r}$$

(but no normalization)

$$r_k = \mathbb{E}\left[\sum_{n=0}^{\tau_j(1)-1} I\{X_n = k\} \mid X_0 = j \right]$$

indicator function $= 1$ if arg true, $= 0$ if arg false

$$= \sum_{n=0}^{\infty} \text{Prob}(X_n = k, \tau_j(1) > n \mid X_0 = j)$$

$$\left(\text{used: } \mathbb{E}\, I\{X_n = k\} = 1 \cdot \text{Prob}(X_n = k) + 0 \cdot \text{Prob}(X_n \neq k) = \text{Prob}(X_n = k) \right)$$

$$= \mathbb{E}\left[\sum_{n=0}^{\infty} I\{X_n = k, \tau_j(1) > n\} \mid X_0 = j \right]$$

Side note: Simplest version of the indicator function trick:

Consider N coin tosses, $p = \text{prob}(H)$

$1-p = \text{prob}(T)$

$X = \#$ heads

$$\text{Prob}(X = j) = \binom{N}{j} p^j (1-p)^{N-j}$$

$$\mathbb{E} X = \sum_{j=0}^{N} j \,\text{Prob}(X = j) = \text{a pain}$$

Another way to calculate:

$$X = \sum_{n=1}^{N} I\{Y_n = 1\}$$

where $Y_n = 1$ if coin toss $n = H$

$= 0$ if coin toss $n = T$

$$\mathbb{E} X = \mathbb{E} \sum_{n=1}^{N} I\{Y_n = 1\} = \sum_{n=1}^{N} \mathbb{E}\, I\{Y_n = 1\}$$

$$= \sum_{n=1}^{N} \text{Prob}(Y_n = 1)$$

$$= \sum_{n=1}^{N} p = pN$$

Consider

$$\sum_{k\in S} r_k P_{k\ell} = \sum_{k\in S}\sum_{n=0}^{\infty} \text{Prob}(X_n = k, T_j(1) \geq n \mid X_0 = j) \times \text{Prob}(X_{n+1} = \ell \mid X_n = k)$$

Want to use ~~Prob(A|B) Prob(B|C) = Prob(A and B|C)~~

$$\text{Prob}(A \mid B \text{ and } C)\, \text{Prob}(B \mid C) = \text{Prob}(A \text{ and } B \mid C)$$

Try to get the expression in this form with

$A = \{X_{n+1} = \ell\}$

$B = \{X_n = k \text{ and } T_j(1) > n\}$

$C = \{X_0 = j\}$

This will work out if I can show that

(C) $\text{Prob}(X_{n+1} = \ell \mid X_n = k)$

$= \text{Prob}(X_{n+1} = \ell \mid X_n = k \text{ and } T_j(1) > n \text{ and } X_0 = j)$

This is actually true by the Markov property when $k \neq j$.

(Use $\text{Prob}(X_{n+1} = \ell \mid X_\ell = i_\ell \text{ for } \ell \in T_0)$

$= \text{Prob}(X_{n+1} = \ell \mid X_{\ell_*} = i_{\ell_*})$ with $\ell_* = \max_{\leq n} T_0$

If $k=j$ then $X_n = k$ and $T_j(1) > n$ can't both be true,

so $Prob(X_n = k, T_j(1) > n \mid X_0 = j) = 0$ for $j = k$

so I can make the other factor whatever

So I can make the change (c)

So
$$\sum_{k \in S} r_k P_{kl} = \sum_{n=0}^{\infty} \sum_{k \in S} Prob(X_n = k, T_j(1) > n \mid X_0 = j) \cdot Prob\left(X_{n+1} = l \mid X_n = k, T_j(1) > n, X_0 = j\right)$$

$$= \sum_{n=0}^{\infty} \sum_{k \in S} Prob(X_{n+1} = l, X_n = k, T_j(1) > n \mid X_0 = j)$$

$$= \sum_{n=0}^{\infty} Prob(X_{n+1} = l, X_n \in S, T_j(1) > n \mid X_0 = j)$$

$$= \sum_{n=0}^{\infty} Prob(X_{n+1} = l, T_j(1) > n \mid X_0 = j)$$

Recall $r_l = \sum_{n=0}^{\infty} Prob(X_n = l, T_j(1) > n \mid X_0 = j)$

Let $n' = n+1$

$$\sum_{k \in S} r_k P_{kl} = \sum_{n'=1}^{\infty} Prob(X_{n'} = l, T_j(1) > n'-1 \mid X_0 = j)$$

A) Consider first the case $\ell \neq j$

$$T_j(1) > n'-1 \iff T_j(1) \geq n$$

$$\mathrm{Prob}(X_n = \ell,\ T_j(1) > n \mid X_0 = j)$$

$$= \mathrm{Prob}(X_n = \ell,\ T_j(1) \geq n \mid X_0 = j)$$

Because they differ by the event

$\{X_n = \ell,\ T_j(1) = n\}$ which is contradictory.

Moreover, the missing $n' = 0$ term:

$$\mathrm{Prob}(X_0 = \ell,\ T_j(1) > 0-1 \mid X_0 = j) = \emptyset.$$

Therefore $v_\ell = \sum_{k \in S} v_k P_{k\ell}$ for $\ell \neq j$

because the summations are equivalent

B) Consider case $\ell = j$:

$$V_j = \mathbb{E}[\,T_{j,j} \mid X_0 = j\,] = 1$$

The expression

$$= \sum_{n=0}^{\infty} \mathrm{Prob}(X_{n+1} = j,\ T_j(1) > n \mid X_0 = j)$$

$$= \sum_{n=0}^{\infty} \mathrm{Prob}(T_j(1) = n+1 \mid X_0 = j)$$

$$= \mathrm{Prob}(1 \leq T_j(1) < \infty \mid X_0 = j) = 1$$

So $v_\ell = \sum_{k \in S} v_k P_{k\ell}$ for $\ell = j$ $\square$

This completes the argument that

$\vec{r} \cdot P = \vec{r}$ and therefore shows

that my candidate $\vec{\tilde{\pi}}$ is a

stationary distribution.

Existence ✓

Uniqueness: Resnick, Sec. 2.12

- show that for any other stationary distribution $\vec{\mu}$, that $\vec{\pi} \le \vec{\mu} \Longrightarrow \vec{\pi} = \vec{\mu}$

References for Finite State Markov Chains:

Ch. 1 Lawler: concise

Ch. 2+3 Karlin + Taylor: elaborate

Ch. 2: Resnick, Adventures in Stochastic Processes

What about reducible MC's (finite state)

- proof doesn't work
- but can modify it to prove <u>existence</u> for all finite state MC's

of stationary distribution

- ~~used~~ but uniqueness may fail

Running Markov chains backwards

If specify $\phi_j = Prob(X_0 = j)$ and

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

how calculate $\{X_n\}$ for $n < 0$?

(want the forward evolution for $n < 0$ to look like the forward evolution for $n > 0$)

To go backwards, want to calculate

$$\hat{P}^{(n)}_{ij} = Prob(X_n = j \mid X_{n+1} = i)$$

$$= \frac{Prob(X_n = j \text{ and } X_{n+1} = i)}{Prob(X_{n+1} = i)}$$

$$= \frac{\text{Prob}(X_{n+1} = i \mid X_n = j)\ \text{Prob}(X_n = j)}{\text{Prob}(X_{n+1} = i)}$$

Bayes' rule: $\text{Prob}(A \mid B) = \dfrac{\text{Prob}(B \mid A)\ \text{Prob}(A)}{\text{Prob}(B)}$

$$\hat{\tilde{P}}_{ij}^{(n)} = \frac{P_{ji}\ p_j^{(n)}}{p_i^{(n+1)}} \quad \text{where } p_j^{(n)} = \text{Prob}(X_n = j)$$

Get contamination by "undoing" forgetting of initial data.

<u>But</u> one can get useful backward evolution by getting rid of initial data

→ initialize with stationary distribution
- because then $p_j^{(n)} = \pi_j$ for all n

$$\hat{\tilde{P}}_{ij} = \frac{P_{ji}\ \pi_j}{\pi_i}$$

Physical meaning of this?

Suppose I have a "microreversible system"

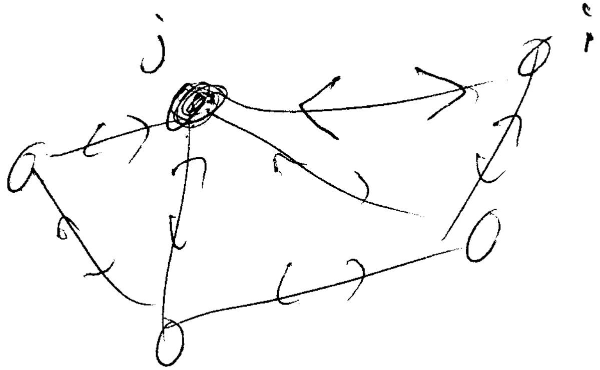
$$\tilde{P}_{ij} = P_{ij}$$

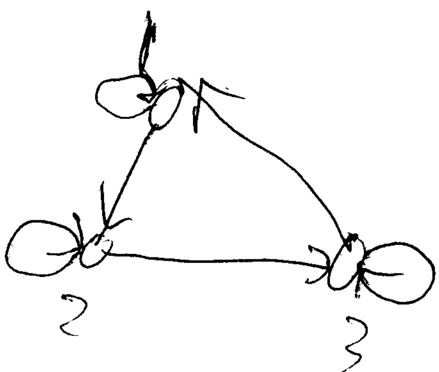
That means

$$\Pi_i P_{ij} = P_{ji} \Pi_j$$ "detailed balance"

↓

applies to systems in thermal equilibrium





: Non-detailed balance

Some comments about computing stationary distribution

A) $\vec{\pi} \cdot P = \vec{\pi}$

$\sum_{j \in S} \pi_j = 1$; solve as eigenvalue problem.

$\pi_j \geq 0$

B) Haken, Synergetics Sec. 4.6-4.8

- easy to solve for $\vec{\pi}$ when detailed balance
- graphical Kirchoff procedure for calculating $\vec{\pi}$ w/o detailed balance

C) Resnick Sec. 2.14:

$$\vec{\pi} = \vec{1} \cdot (I - P + ONE)^{-1}$$

D) Resnick Sec. 2.13:

Calculate $\vec{\pi}$ by using stochastic update rule, manipulate w/ gen fns.

- useful for ∞ states

Brief note on statistical questions:

- How should I choose $\vec{\phi}$ and P? (given data)
- How do I know if my model is "good"?

Guttorp, <u>Stochastic Modeling of Scientific Data</u> Sec. 2.1

Idea: Maximum likelihood method

I) Suppose we try to model the given data as a set of i.i.d. variables with distribution $\vec{\phi}$.

How choose $\vec{\phi}$? (after choosing state space)

What's the likelihood to observe data given my model?

$$L = \prod_{i=1}^{M} \phi_i^{a_i}$$

where a_i = # occurrences of state i in the data.

Choose $\vec{\phi}$ to maximize L.

$$\Rightarrow \phi_i = \frac{a_i}{\sum_{i \in S} a_i}$$

II) Markov chain model: How choose $\vec{\emptyset}$ and P?

$$L = \prod_{\substack{i \in S \\ j \in S}} P_{ij}^{a_{ij}} \emptyset_k$$

where k = initial state

a_{ij} = # transitions $i \to j$ in the data

$1\ 3\ 5\ 4\ 2$: $\emptyset_1 P_{13} P_{35} P_{54} P_{42}$

Optimize L with respect to P
(usually ignore $\vec{\emptyset}$)

$$P_{ij} = \frac{a_{ij}}{\sum_{j \in S} a_{ij}}$$

Was the use of a Markov chain a significant improvement over using iid rv?

- use χ^2 statistic to see if L^{mc}/L^{iid} is large enough to have statistically significant value, given extra degrees of freedom.

How calculate long-time properties of MC?

Need some classification of Markov chains.

Period of a state i:

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

prob to come back at time n.