Detailed Balance and Branching Processes

Monday, October 20, 2008
12:04 PM

Looking for a detailed balance solution for a stationary distribution is a useful technique.

Stationary distribution: S1) $\pi_j \geq 0$ for $j \in S$

S2) $\sum_{i \in S} \pi_i P_{ij} = \pi_j$ for $j \in S$

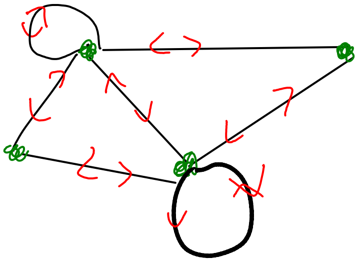
S3) $\sum_{j \in S} \pi_j = 1$

Detailed balance: D1) $\pi_j \geq 0$ for $j \in S$

D2) $\pi_i P_{ij} = \pi_j P_{ji}$ for $i, j \in S$

D3) $\sum_{j \in S} \pi_j = 1$

Summing D2 over $i \in S$ gives S2 so any detailed balance solution gives rise to a stationary distribution.

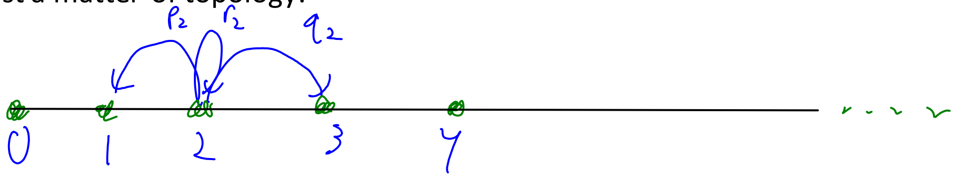


Note that a stationary distribution must satisfy, at each node (state), that the total probability flux out of a node is equal to the total probability flux into a node at each time step.

In a detailed balance solution, the probability flux is balanced in both directions along each connection between nodes (states).

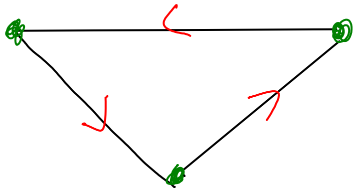
This becomes interesting in modeling physical systems with time-reversibility, which generally speaking refers to systems with atomic or molecular resolution. One has to be careful because initial conditions can break time-reversal symmetry when you look at statistics. However, once the system reaches thermal equilibrium (to a decent approximation), the statistical dynamics should again look time-reversible. Being in thermal equilibrium is like being in a stationary distribution. If a system has a time-reversibility property and is in a stationary distribution (thermal equilibrium), then it must satisfy detailed balance.

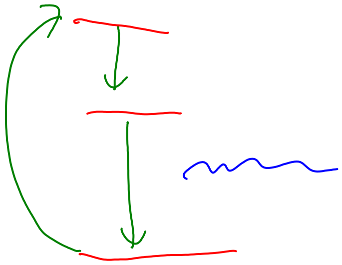
But last time we looked a birth-death chain that had nothing to do with physics, yet the detailed balance solution also worked there. Why? Just a matter of topology.



In retrospect, we see here that violating detailed balance would imply a non-decaying net probability flux to infinity, which is incompatible with having a positive recurrent system (which is what the existence of a stationary distribution would imply).

Not every Markov chain needs to have detailed balance. For example a 3-state laser cycle:





By the way, notice that from the first homework, when we time-reverse a Markov chain in a stationary distribution, the time-reversed Markov chain has the probability transition matrix

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

$$= \frac{P_{ji}\,\pi_i}{\pi_j} \qquad \text{(Bayes' law)}$$

$= P_{ij}$ if the system is in detailed balance.

Branching processes

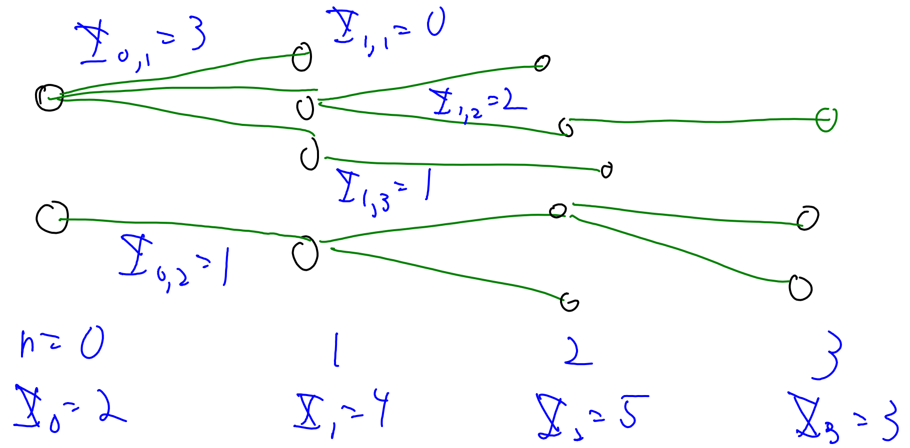
Lawler Sec. 2.4
Resnick Sec. 1.4

These are a special class of countable-state, discrete-time Markov chains, but it is useful to introduce some special tools to analyze them.

State space $S = \mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$

X_n is the number of "agents" in the system at epoch n.

At each epoch, each agent gives rise to a certain number of offspring according to some probability distribution, and each agent gives rise to offspring according to the same probability distribution and independently of one another. Moreover the offspring generated at each epoch by each agent are again independent and identically distributed random variables.



In interpreting this branching process, you can either think of each agent living for only one epoch, or alternatively, that if it survives, it counts as one of its own offspring.

$$X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k}$$ (random sum)

where the $\{Y_{n,k}\}$

are independent, identically distributed random variables drawn from a specified probability distribution

$$p_j = P(Y_{n,k} = j)$$

for $j = 0,1,2,\ldots$

This mathematical definition of the branching process is called the Galton-Watson process, which is the simplest version of a more general modeling framework for branching processes which can include such things as:

- age structure
- keep track of the population and branching events in continuous time
- multiple types of agents

For such extensions, see specialized books on branching processes such as Branching Processes in Biology by Kimmel and Axelrod.

Applications of branching processes (including the more sophisticated extensions)

- genealogy
- populations without hindrances to growth (particularly under asexual reproduction)
- proliferation of viruses such as AIDS which infects T-lymphocytes
- diversity of mutated samples in a PCR experiment
  - Kary Mullis Dancing Naked in the Mind Field
- evolutionary studies of genomes
- photomultiplier tube cascades
- nuclear fission
- earthquake triggering
- queuing models

Mathematical analysis of branching processes

The usual tools are bit problematic to employ because it's awkward to write down the probability transition matrix, but easy to write down the stochastic update rule.

On the other hand, the random sum structure in the stochastic update rule allows us to bring in a very useful probability tool for dealing with sums of random variables.

Probability generating functions (Resnick 1.3, Karlin & Taylor 1.1E)

Given a nonnegative integer-valued random variable X, we define its probability generating function as follows:

$$\mathcal{P}_X(s) = \mathbb{E}\, s^X = \sum_{j=0}^{\infty} p_j s^j$$

$$\text{where} \quad p_j = P(X = j)$$

This is basically a Z-transform of the sequence $\{p_0, p_1, p_2, \ldots, p_n, \ldots\}$

which is like a Laplace transform except one is summing over a sequence rather than integrating over a continuous function.

The basic idea is that the probability generating function is a complete alternative representation of the probability distribution, just as a Fourier/Laplace transform of a function is a complete alternative representation of that function. The original probability distribution can be recovered from the probability generating function through an inversion formula.

$$p_j = \frac{1}{j!}\left(\frac{d}{ds}\right)^j \mathcal{P}_X(s)\Big|_{s=0}$$

Also, the probability generating function converges for at least the interval $[0, 1]$ using the fact that $\sum_{j=0}^{\infty} p_j = 1$

Note that even though the probability generating function can be defined for complex values of s, typically one focuses just on nonnegative real values of s.

Remark: Similar transforms (called moment generating functions, characteristic functions) can be defined for random variables with arbitrary probabilty distributions but we won't need them here.

What are probability generating functions good for?

It can simplify the computation of moments.

$$\langle X^n \rangle = \sum_{j=0}^{\infty} j^n p_j = \left(s \frac{d}{ds}\right)^n \mathcal{P}_X(s)\Big|_{s=1}$$

$$\|$$

$$\langle X^n \rangle = \sum_{j=0}^{\infty} j^n p_j = \left(s\frac{d}{ds}\right)^n \mathcal{P}_X(s)\Big|_{s=1}$$

$$= \left(s\frac{d}{ds}\right)^n \sum_{j=0}^{\infty} p_j s^j \Big|_{s=1}$$

use theorems on power series diff →

$$= \sum_{j=0}^{\infty} p_j \left(s\frac{d}{ds}\right)^n s^j \Big|_{s=1}$$

$$= \sum_{j=0}^{\infty} p_j j^n s^j \Big|_{s=1}$$

$$= \sum_{j=0}^{\infty} p_j j^n$$

Example where this is useful: Binomial distribution

$$p_j = \binom{m}{j} p^j (1-p)^{m-j} \quad \text{for } j = 0, 1, \ldots, m$$

$$\langle X \rangle = \sum_{j=0}^{m} \binom{m}{j} p^j (1-p)^{m-j} j = \ldots$$

$$\mathcal{P}_X(s) = \sum_{j=0}^{m} \binom{m}{j} p^j (1-p)^{m-j} s^j$$

$$= \sum_{j=0}^{m} \binom{m}{j} (ps)^j (1-p)^{m-j}$$

binomial theorem

$$= (ps + 1 - p)^m$$

$$\langle X \rangle = s\frac{d}{ds} \mathcal{P}_X(s)\Big|_{s=1} = mps(ps + 1 - p)^{m-1}\Big|_{s=1}$$

$$=$$

$$\langle X \rangle = s \frac{d}{ds} \mathcal{P}_X(s) \Big|_{s=1} = m p s (ps + 1 - p)^{m-1} \Big|_{s=1}$$

$$= m p$$

$$\langle X^2 \rangle = \left(s \frac{d}{ds} \right)^2 \mathcal{P}_X(s) \Big|_{s=1} = m(m-1) p^2 s^2 (ps + 1 - p)^{m-2} + m p s (ps + 1 - p)^{m-1} \Big|_{s=1}$$

$$= m(m-1) p^2 + m p$$

$$\sigma_X^2 = \langle X^2 \rangle - \langle X \rangle^2 = m(m-1) p^2 + m p - (m p)^2$$

$$= m p - m p^2 = m p (1-p)$$

$$\sigma_X = \sqrt{m p (1-p)}$$