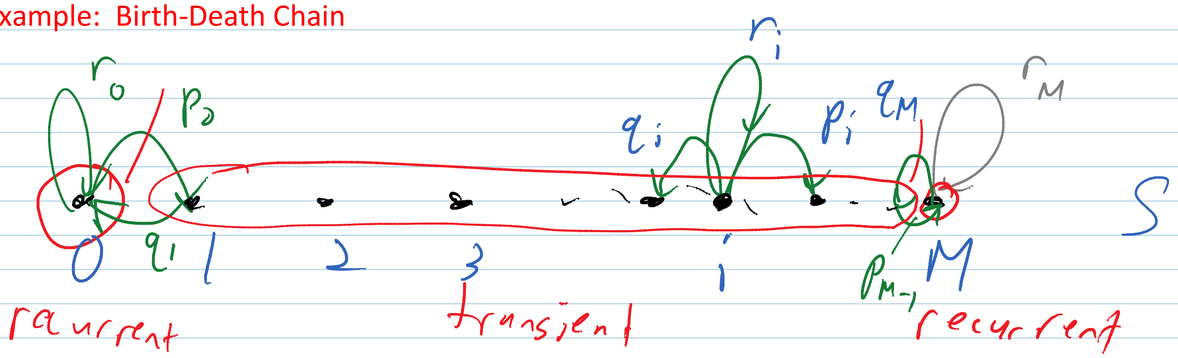
Birth-death chains

Monday, October 19, 2015

2:22 PM

Example: Birth-Death Chain



State space $S = \{0, \dots, M\}$. From any state $1 \leq i \leq M-1$ we allow the following transitions:

- $i \to i+1$ with probability p_i (birth)
- $i \to i-1$ with probability q_i (death)
- $i \to i$ with probability $r_i = 1 - p_i - q_i$

At the boundary states $i = 0, M$, we take the same rules except we forbid leaving the state space, so $p_M = 0$, $q_0 = 0$.

Applications of birth-death chains:

- population models
- random walk in heterogenous environment
- queueing model
- # objects bound to some surface (molecular motors bound to a microtubule)

Let's proceed with the general abstract formalism, and we'll take absorbing boundary conditions which means taking the choice: $q_0 = 0, p_M = 0$.

This choice means that, if the rest of the dynamics has no broken links ($p_i > 0, q_i > 0$) for all $i = 1,2, \dots, M-1$) the communication classes would be:

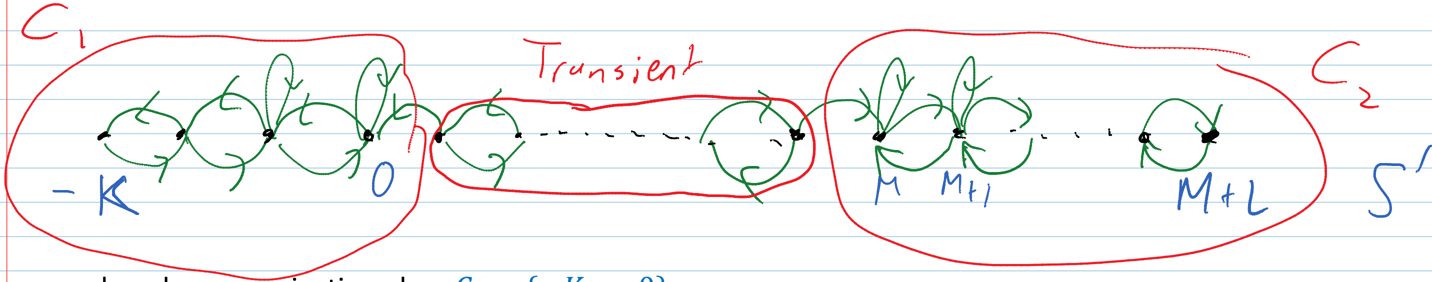
- absorbing state at 0
- absorbing state at M
- class $\{1,2, \dots, M-1\}$ (every state within here can be reached from any other state by simply proceeding stepwise from one to the other)

The class of interior states is transient because there are one-way links from interior states to the absorbing states.

Therefore, if we start at an interior state $i \in \{1,2, \dots, M-1\}$, we can ask the questions that are standard for long-term behavior starting from a transient state:

- Into which recurrent communication class does the Markov chain eventually go? That is, will the Markov chain reach 0 or M, and with what probabilities? (Absorption probabilities)
- Expected cost/reward (i.e., expected time) until absorption at a recurrent state happens.

Here the stationary distributions of the closed/recurrent classes are trivial, but one could contemplate a variation of the birth-death chain where say, the Markov chain had extended state space $S' = \{-K, -K+1, \dots, -1, 0, 1, 2, \dots, M, M+1, \dots, M+L\}$ with birth-death dynamics as above, and $p_i > 0$ for $i \in S \setminus \{0, M+L\}$ and $q_i > 0$ for $i \in S \setminus \{-K, M\}$, but $p_0 = p_{M+L} = 0$ and $q_M = q_{-K} = 0$.



- closed communication class $C_1 = \{-K, \dots, 0\}$
- closed communication class $C_2 = \{M, \dots M+L\}$

  Then one could combine the absorption probabilities that we are about to calculate (i.e., the probability the MC eventually ends up in C_1 or C_2) together with the stationary distributions of the MC restricted to C_1, C_2 respectively to compute long-term properties of the MC by using law of total probability.

  - For example, let's suppose the stationary distributions restricted to C_1, C_2 are $\pi^{(L)}, \pi^{(R)}$, respectively, and the absorption probabilities to hit state 0 (or M) first starting from state i is U_{i0} (or U_{iM})

    Then C_1, C_2 will under the above assumptions serve as state spaces for irreducible MC. with suitable assumptions (i.e. $r_j > 0$ for some $j \in \{-K, \dots, -1\}$ and for some $j \in \{M, \dots, M+L\}$), they will also be aperiodic. Then we can apply limit distribution and LLN of MC ideas.

    - If $k \in C_1$, $\lim_{n\to\infty} P(X_n = k | X_i = 0) = U_{i0}\pi_k^{(L)}$
      - because $= \lim_{n\to\infty} P(X_n \in C_1 | X_i = 0) P(X_n = k | X_n \in C_1, X_i = 0)$
    - If we have some cost/reward function $f(k)$ for $k \in S'$, then

Corrected from lecture

      $\lim_{n\to\infty} \frac{1}{n} \sum_{n'=1}^{n} f(X_{n'}) =$
      - $\sum_{k\in C_1} \pi_k^{(L)} f(k)$ with probability U_{i0}
      - $\sum_{k\in C_2} \pi_k^{(R)} f(k)$ with probability U_{iM}
      - Sketch of proof: Let A_i denote the event that $X_n \in C_i$ eventually. That is, $A_i = \{\omega \in \Omega : \exists n \geq 0 \ s.t. X_n(\omega) \in C_i\}$.
        $P(A_1) = U_{i0}$ and $P(A_2) = U_{iM}$, and $P(A_1) + P(A_2) = 1$ (since Markov chains have zero probability to stay forever in a finite set of transient states). Therefore, in terms of indicator functions, $1_{A_1} + 1_{A_2} = 1$(almost surely), so

$$\lim_{n\to\infty} \frac{1}{n} \sum_{n'=1}^{n} f(X_{n'}) = \lim_{n\to\infty} \frac{1}{n} (1_{A_1} + 1_{A_2}) \sum_{n'=1}^{n} f(X_{n'}) =$$

$$= \lim_{n\to\infty} \frac{1}{n} 1_{A_1} \sum_{n'=1}^{n} f(X_{n'}) + \lim_{n\to\infty} \frac{1}{n} 1_{A_2} \sum_{n'=1}^{n} f(X_{n'})$$

$$= 1_{A_1} \sum_{k\in C_1} \pi_k^{(L)} f(k) + 1_{A_2} \sum_{k\in C_2} \pi_k^{(R)} f(k)$$

We will just show how to do the calculation of the absorption probabilities; the calculation of expected cost/reward until absorption (i.e., expected time) can be found in the texts using similar ideas.

Let's write the probability transition matrix:

$$P = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ \\ \\ M \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & \cdots & & & M \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & & & & \\ q_1 & r_1 & & & & 0 & \\ & q_2 & r_2 & p_2 & & & \\ & & q_3 & r_3 & p_3 & & \\ & 0 & & & & & 0 \\ & & & & & & 0 \\ & & & & q_{M-1} & r_{M-1} & p_{M-1} \\ & & & & & & \end{pmatrix} \end{array}$$

The structure of the Markov chain is simple enough that we can actually get an analytical solution for the absorption probabilities, and for deriving analytical results, it's usually easier to work with the recursion form of the equation for absorption probabilities:

$$U_{ij} = P_{ij} + \sum_{k \in T} P_{ik} U_{kj}$$

If we write out these equations based on the probability transition matrix, we have (note we will only compute U_{i0} because $U_{i0} + U_{iM} = 1$; there are only two possible recurrent states to be absorbed in.)

For $2 \leq i \leq M-2$

$$U_{i0} = \underbrace{0}_{P_{i0}} + \underbrace{p_i U_{i+1,0}}_{k=i+1} + \underbrace{q_i U_{i-1,0}}_{k=i-1} + \underbrace{r_i U_{i,0}}_{k=i}$$

$i=1$: $$U_{10} = \underbrace{q_1}_{P_{i0}} + \underbrace{p_1 U_{2,0}}_{k=2} + \underbrace{r_1 U_{1,0}}_{k=1}$$

$$i = M-1: \quad U_{M-1,0} = 0 + q_{M-1} U_{M-2,0} + r_{M-1} U_{M-1,0}$$

$$P_{i,0}$$

Notice that in general the equations for U_{ij} are only coupled in i, not in j.

How do we solve these equations analytically? These equations can be thought of as second order linear difference equations in analogy to second order linear differential equations. The solution procedure we will adopt is analogous to how we solve second order differential equations of the form:

$$y'' + a(x)\, y' = 0$$

Define $u = y'$

$$u' + a(x)\, u = 0$$

We will find that our set of linear difference equations has an analogous structure and will be solved by successive integration of two first order linear difference equations. For more on solving linear difference equations, see Lawler Ch. 0 or Bender & Orszag, Advanced Mathematical Methods for Scientists and Engineers Ch. 3.

To expose the structure of our second order difference equations, we will make 2 simplifications:

1) First, define $U_{00} = 1$, $U_{M0} = 0$. Then all equations can be written in the common form:

$$U_{i0} = q_i U_{i-1,0} + r_i U_{i,0} + p_i U_{i+1,0}$$

$$\text{for } 1 \le i \le M-1$$

2) Substitute the relation: $r_i = 1 - p_i - q_i$

$$U_{i0} = q_i U_{i-1,0} + (1 - p_i - q_i) U_{i,0} + p_i U_{i+1,0}$$

$$= U_{i,0} + q_i (U_{i-1,0} - U_{i,0}) + p_i (U_{i+1,0} - U_{i,0})$$

$$0 = q_i (U_{i-1,0} - U_{i,0}) + p_i (U_{i+1,0} - U_{i,0})$$

This is analogous to a second order differential equation of the form $y'' + a(x)y' = 0$

Following the linear DE analogy, we introduce the difference variable:

$$V_i = U_{i,0} - U_{i-1,0} \quad \text{for} \quad 1 \le i \le M$$

$$0 = q_i(-V_i) + p_i V_{i+1}$$

This can be solved directly by recursion (analogous to just integrating a first order linear differential equation even with variable coefficients):

$$V_{i+1} = \frac{q_i}{p_i} V_i$$

"integrating factor"

$$V_i = \underbrace{\prod_{k=1}^{i-1} \left(\frac{q_k}{p_k}\right)}_{\gamma_i} V_0$$

What is V_0? Need to apply boundary conditions. We can extract a condition on the difference variable from the boundary conditions on U by imposing a summation constraint on V (just like an integral constraint on a derivative in differential equations).

$$\sum_{i=1}^{M} V_i = \sum_{i=1}^{M} (U_{i,0} - U_{i-1,0})$$

telescoping sum

$$= U_{M,0} - U_{0,0} = 0 - 1$$

$$\sum_{i=1}^{M} V_i = \sum_{i=1}^{M} (U_{i,0} - U_{i-1,0})$$ telescoping sum

$$= U_{M,0} - U_{0,0} = 0 - 1$$

$$\sum_{i=1}^{M} V_i = -1$$

$$\sum_{i=1}^{M} \gamma_i V_0 = -1$$

$$V_0 = -\frac{1}{\sum_{i=1}^{M} \gamma_i}$$

$$V_i = \gamma_i V_0 = \frac{-\gamma_i}{\sum_{k=1}^{M} \gamma_k}$$

Now we need to sum up the differences V_i to get the U_{i0}:

$$U_{i0} - \underbrace{U_{00}}_{1} = \sum_{l=1}^{i} (U_{l0} - U_{l-1,0})$$

$$= \sum_{l=1}^{i} V_l$$

$$= \sum_{l=1}^{i} \gamma_l V_0$$

$$= \frac{-\sum_{l=1}^{i} \gamma_l}{\sum^{M} \cdots}$$

$$\frac{{}_{\ell=1} \qquad {}_{\ell}}{\sum_{k=1}^{M} \gamma_k}$$

$$U_{i0} = 1 - \frac{\sum_{\ell=1}^{i} \gamma_\ell}{\sum_{k=1}^{M} \gamma_k} = \frac{\sum_{\ell=i+1}^{M} \gamma_\ell}{\sum_{k=1}^{M} \gamma_k}$$

where $$\gamma_k = \prod_{m=1}^{k-1} \left(\frac{q_m}{p_m}\right)$$

denotes a sort of ``ease'' with which the Markov chain can move from state k to state 1.

Recall U_{i0} is the probability that the Markov chain will, starting from state i, eventually reach state 0 (and therefore class C_1 if we think of the extended Markov chain on state space S')