Next class is Tuesday, March 18.

A formula for the expected cost/reward (additive functional) will follow from a first step analysis that is close to the first step analysis for the absorption probability. Again, it sets up a recursion.

To implement the first step analysis, we need an analogy to the law of total probability, called the law of total expectation

\[ E(Y) = \sum_k E(Y|B_k) P(B_k) \quad \text{provided that } \{B_k\} \text{ form a partition of sample space.} \]

We will deploy this with a common condition added everywhere:

\[
E(Y|C) = \sum_k E(Y|B_k, C) P(B_k|C)
\]

As before, we consider the first factor by cases.

i) \( k \notin T \); \( T = 1 \) \( \text{ (because } T = \min \{ n \geq 0 : X_n \notin T \} \)
\[ (i) \quad \text{for } t \geq 1 \]

**First factor**

\[
\mathbb{E} \left( \sum_{n=0}^{T-1} f(X_n) \bigg| X_t = k, X_{t-1} = i \right)
\]

\[
= \mathbb{E} \left( f(i) + f(k) + \sum_{n=2}^{T-1} f(X_n) \bigg| X_t = k, X_{t-1} = i \right)
\]

\[
= f(i) + f(k) + \mathbb{E} \left( \sum_{n=2}^{T-1} f(X_n) \bigg| X_t = k \right)
\]

by Markov property.

\[
= f(i) + \mathbb{E} \left( \sum_{n=1}^{T-1} f(X_n) \bigg| X_t = k \right)
\]

\[
= f(i) + \mathbb{E} \left( \sum_{n=1}^{T-1} f(X_n) \bigg| X_t = k \right)
\]

\[
= f(i) + w_k \quad \text{(with time shift, from epoch } 0 \text{ to } 1)
\]

**First factor**

\[
= f(i) + w_k
\]

Putting these two cases together:

\[
w_i = \sum_{ket} f(i) P_{ik} + \sum_{ket} (f(i) + w_k) P_{ik}
\]
Recursion formula for the expected value of the additive functional (cost/reward)

\[ w_i = f(i) + \sum_{k \in \mathcal{T}} p_{ik} w_k \]

for \( i \in \mathcal{T} \)

These results, and the proof generalizes to additive functionals \( f \) that are actually random, provided their randomness is conditionally independent given the realization of the Markov chain. This can be written in matrix form, using the canonical decomposition.

\[
\tilde{w} = \begin{pmatrix} w_i \\ \vdots \\ w_M \end{pmatrix} \quad \tilde{f} = \begin{pmatrix} f(M-M_T) \\ \vdots \\ f(M) \end{pmatrix}
\]

\( M_T \) = \# transient states

\[
\tilde{w} = \tilde{f} + \Omega \tilde{w}
\]

\[
(I-\Omega) \tilde{w} = \tilde{f}
\]

\[
\tilde{w} = (I-\Omega)^{-1} \tilde{f}
\]
independent given the realization of the Markov chain.

\[ \{ f(X_n) \mid X_n \} \text{ independent} \]

Example is the number of defective products shipped per inspection in our quality control model.

The above results carry over with the replacement

\[ f(i) \rightarrow E f(i) \]

Same as our extension of LLN for MC from before.

Some ideas on how these theorems on absorption probability and expected cost/reward can be applied to many situations in analyzing long-time properties of MCs.

Common applications of the cost/reward formula, beyond models for which cost/reward explicitly defined, is:

- Expected number of epochs spent in transient states before hitting a recurrent state can be computed by choosing \( f \equiv 1 \).
- To compute the expected number of visits to a transient state \( k \), starting from another transient state \( i \), choose:

\[
    f(j) = \delta_{jk} = \begin{cases} 
    1 & \text{i} \rightarrow \text{j} \rightarrow k \\
    0 & \text{else}
    \end{cases}
\]

With some trickier manipulation, we can use the cost/reward formulation to answer the following questions:

1) Starting from state \( i \), what is the expected number of epochs until another state \( j \) is visited. No assumption on whether these states are recurrent or transient. This is a general mean first passage time calculation.

To answer this question, modify the Markov chain by making \( j \) an absorbing state (so \( P_{jj} = 1 \)). In this modified MC, \( j \) is recurrent, all
other states are transient, but the answer to the question doesn't change. But in the modified MC, this is counting the number of epochs while transient, which is a calculation we did before.

2) Starting from state \(i\), what is the expected number of epochs spent visiting state \(k\) before state \(j\) is hit?

Same trick works (make state \(j\) absorbing) and count the number of expected visits to transient state \(k\) before the recurrent class is hit.

3) Starting from state \(i\), what is the probability that state \(j\) is hit before state \(k\)?

Make both states \(j\) and \(k\) absorbing. Reduces to absorption probability calculation.

Caveat: The above tricks work perfectly well in irreducible Markov chains, but if the MC is already reducible, think through whether the modifications still work.

Example: Birth-Death Chain

From any state \(i \in \{0, ..., M\}\), allowed transitions are:

\[ i \to i + 1, i \to i - 1, i \to i \]

with probabilities \(p_i, q_i, r_i\) respectively, with the restriction that \(p_i + q_i + r_i = 1\).

Forbid running out of the state space, so insist that:

\[ p_M = 0, q_0 = 0. \]

With the Markov property that each jump only depends on the current state, not any further history.

Applications of birth-death chain:

- Population modeling
- Financial models with discrete increments of value
- Biased random walk (atomic physics, ecology)
- Molecular motors (# attached to a cargo)
- Discrete voltage model for neurons (almost, except the transition from threshold to reset \((M \to 0)\))
But we'll proceed abstractly with birth-death chains, without regard to particular application.

One can make various choices about what happens at the boundaries $0, M$. We'll study the case of absorbing boundary conditions:

- $p_0 = 0, q_M = 0$ so $r_0 = 1, r_M = 1$.

The probability transition matrix for this birth-death chain is tridiagonal.

\[
P = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}
\]

One can derive exact expressions to the following questions for birth-death chains with absorbing boundary conditions:

- With what probability does the birth-death chain become absorbed at state 0 (versus state $M$) if it starts at state $i$.
- What is the expected number of epochs required to be absorbed at an endpoint?

We'll show how to do the absorption probability calculation; the other question is answered through similar techniques; read in the texts.

One could approach the problem by writing the canonical form of the probability transition matrix, and then use the matrix inversion formulas. This is good for numerical calculation. But we'll use instead the recursion formulas because the probability transition matrix for birth-death chains has a simple enough structure, that we can get a general analytic expression.

\[
U_{ij} = P(\sum_{t=1}^{T} = j | \sum_{0}^{t} = i)
\]

\[
T = \min \{ n \geq 0 : \sum_{t=0}^{n} \in \{ 0, M \} \}
\]

\[
\forall i \in T = \{ 1, 2, \ldots, M-1 \}
\]

\[
\forall i \in T^c = \{ 0, M \}
\]
How does one solve these kinds of recursion equations? Note first of all that there is no coupling in the $j$ index; we can solve separately the equations for $j = 0$ and for $j = M$. (This will happen in general for absorption probability calculations.) The banded structure for the $i$ index makes the number of terms appearing relatively small in each equation. These often can be thought of as linear difference equations which are discrete analogs to linear differential equations. There is a parallel theory for solving linear difference equations to linear differential equations (Lawler Ch. 0, Bender & Orszag, Advanced Mathematical Methods for Scientists and Engineers). But our equations have rather general coefficients, and with 3 terms on the right hand side, it looks as hard as solving a second order linear differential equation with variable coefficients (think of a finite difference discretization) -- only in special cases does an analytical solution exist.

Here's the differential equation analogy of the structure and strategy we will employ for the recursion analogy:

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

$$\begin{align*}
V_{ij} &= q_i V_{i-1,j} + r_i V_{ij} + p_i \sum_{k \neq i} \rho_{ik} V_{kj} \\
\end{align*}$$

For $i=1$

$$\begin{align*}
V_{1j} &= q_1 \delta_{1j} + r_1 V_{1j} + p_1 \sum_{k \neq 1} \rho_{1k} V_{kj} \\
\end{align*}$$

For $i=M-1$

$$\begin{align*}
V_{M-1,j} &= p_{M-1} \delta_{M-1,j} + r_{M-1} V_{M-1,j} + q_{M-1} V_{M-2,j} \\
\end{align*}$$
employ for the recursion analogy:

\[ y'' + a(y) y' = 0 \]

\[ \text{Change variable } \eta = y' \]

\[ \eta' + a(y) \eta = 0 \]

Note now that \( r_i = 1 - p_i - q_i \) so that we can write

\[ U_{i0} = q_i U_{i-1,0} + (1 - p_i - q_i) U_{i,0} + p_i U_{i+1,0} \]

\[ 0 = r_i (U_{i+1,0} - U_{i,0}) + q_i (U_{i-1,0} - U_{i,0}) \]

Define

\[ V_i = U_{i,0} - U_{i-1,0} \quad \text{for} \quad 1 \leq i \leq M \]

Then

\[ \rho_i V_{i+1} + q_i (-V_i) = 0 \quad \text{for} \quad 1 \leq i \leq M-1 \]

\[ V_{i+1} = \frac{q_i}{\rho_i} V_i \]

To implement this idea, it's helpful first to define:

\[ U_{0j} = \delta_{0j} \text{ and } U_{Mj} = \delta_{Mj} \]

which is intuitive way to extend definition of absorption probabilities to the recurrent endpoints. Then we can write the equations for the three cases in a unified fashion:

\[ U_{ij} = q_i U_{i-1,j} + r_i U_{i,j} + p_i U_{i+1,j} \quad \text{for } i \in \{1, 2, ..., M - 1\} \text{ and } j = \{0, M\}. \]

We just need to solve the \( j = 0 \) equations, since once we have \( U_{0j} \), we obtain \( U_{Mj} = 1 - U_{M0} \).
\[ V_i = \prod_{k=1}^{i-1} \left( \frac{q_k}{P_k} \right) V_i \]

Need to apply boundary conditions to determine \( V_i \).

Note the telescoping sum:

\[ \sum_{i=1}^{M} V_i = \sum_{i=1}^{M} (V_{i,0} - V_{i-1,0}) = V_{M,0} - V_{0,0} = 0 - 1 = -1 \]

\[ \sum_{i=1}^{M} V_i \sigma_i = -1 \]

\[ \therefore V_j = -\frac{1}{\sum_{i=1}^{M} \sigma_i} \]

\[ \therefore V_j = -\frac{1}{\sum_{k=1}^{M} \sigma_k} \]

Sum up to get \( V_{i,0} \):

\[ V_{i,0} = V_{0,0} - \sum_{j=1}^{i} (V_{j,0} - V_{j-1,0}) \]

\[ = \sum_{k=1}^{i} V_k = \frac{\sum_{k=1}^{i} \sigma_k}{\sum_{k=1}^{M} \sigma_k} \]

\[ V_{0,0} \]

\[ \therefore V_{i,0} = 1 - \frac{i}{\sum_{k=1}^{M} \sigma_k} \]
\[ \sum_{k=1}^{M} \sum_{i=1}^{j} \delta_{k, i} = \sum_{k=1}^{M} \delta_{k} - \sum_{k=1}^{i} \delta_{k, i} \]

\[ \sum_{k=1}^{M} \sum_{i=1}^{j} \delta_{k} \]

\[ V_{i, 0} = \sum_{k=1}^{M} \sum_{j=1}^{i} \delta_{k} \quad V_{i, M} = \sum_{k=1}^{M} \delta_{k} \]

where \( \gamma_j = \prod_{k=1}^{j-1} \left( \frac{\ell_k}{p_k} \right) \)