No class next week. No office hours either.

Next class will be 11/01.

The cost/reward formula has two specific widely used applications:

- Take the cost/reward function to be $f = 1$. Then $w_i$ will be the expected number of epochs spent in a transient state before reaching a recurrent state, given that the MC starts in a transient state $i$.
- Take the cost reward function to be:

  $$\delta_{i,k} = \begin{cases} 
  1 & \text{for } i = k \\
  0 & \text{for } i \neq k
  \end{cases}$$

  Where $k$ is a transient state. Then $w_i$ will be the expected number of visits to state $k$, starting from state $i$.

A simple trick will make the absorption probability and cost/reward calculations apply to many more questions than one might have naively thought:

- Starting from state $i$, what is the expected number of epochs until another state $j$ is visited? (first passage time problem)
- Starting from state $i$, how many epochs is the Markov chain expected to be in state $k$ before state $j$ is visited.
- Starting from state $i$, what is the probability that state $k$ is visited before state $j$?

First let's consider the case of an irreducible Markov chain.

For the first question, modify the Markov chain by making state $j$ absorbing (probability one to remain in
that state). Then j will be recurrent, and all other states of the Markov chain will become transient. Now one can apply to this modified Markov chain the formula for the expected time to reach a recurrent state.

For the second question, make the same modification, and now use the cost/reward function

\[ f(s) = \delta_{1/k} \]

For the third question, make both states j and k absorbing, then use absorption probability formula.

If the Markov chain is reducible, then one uses similar ideas, but have to be careful about whether some answers may be infinite.

Example

**Birth-death chain** (biased random walk)

\[
\begin{array}{cccccccc}
& & & & & & \rho_1 & & \rho_2 \\
\rho_0 & \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 & \rho_6 & \rho_7 & \rho_8 \\
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{array}
\]

The Markov chain dynamics are prescribed by the above diagram, giving the probabilities for transitions to the same or neighboring state during an epoch.

\[
\begin{align*}
\rho_i + q_i + r_i &= 1 \\
q_0 &= 0 \\
\rho_m &= 0 \\
p_i, q_i, r_i &\geq 0
\end{align*}
\]

Birth-death chains also arise in:
### Biased random walk model (with inhomogeneities)
- Atomic physics models, including defects in lattices
- Chemical reactions
- Financial price models
- Molecular motor models with multiple motors bound to a cargo

But we'll just approach the birth-death chains with a generic analytical approach without consideration of a specific application.

We'll in particular take the special case of **absorbing boundary conditions**.

\[
\begin{align*}
\rho_0 &= q_0 = 0 & r_0 &= 1 \\
\rho_m &= q_m = 0 & r_m &= 1
\end{align*}
\]

We'll take this as the original Markov chain, but it could also have arisen from a modification to a Markov chain without absorbing boundary conditions, but where we are trying to compute statistics about what happens before I reach one of the endpoints.

The stochastic update rule is easy to write down, but we have developed the formulas to more naturally refer to the probability transition matrix, so that's how we'll encode the Markov chain.

\[
P = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
1 & 0 & 2 & \cdots & 1 \\
0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & 0 \\
1 & 0 & \cdots & 0 & 0
\end{pmatrix}
\]

Let's consider how to compute analytically the answers...
to two basic questions involving this Markov chain:

• If the MC starts in state $I$, what is the probability that it will become absorbed in state $0$ (or $M$).
• What is the expected number of epochs until the Markov chain becomes absorbed at an endpoint?

Optionally, we can prepare for these questions by forming the canonical decomposition of the probability transition matrix:

Using our notation from last time:

$$V_{ij} = P(X_T = j \mid X_0 = i)$$

$$i \in \{1, 2, \ldots, M-1\}$$

$$j \in \{0, M\}$$

$$T = \min\{n > 0 : X_n \in \{0, M\}\}$$

$$V = (I - W)^+ R$$

This would be a good way to compute the answers.
numerically using numerical linear algebra.

But birth-death chains are simple enough that we can actually obtain the answers analytically, and to do so, it's more convenient to work with the recursion equations:

\[
U_{ij}^* = P_{ij} + \sum_{k=1}^{M-1} P_{ik} U_{kj}^*
\]

for \( i \in \{1, \ldots, M-1\} \), \( j \in \{0, \ldots, M\} \).

We now substitute in the structure of the probability transition matrix to write this system of equations for \( U_{ij} \) explicitly.

For \( 2 \leq j \leq M-2 \):

\[
U_{ij} = 0 + q_{ij} U_{i-j,j} + r_{ij} U_{ij} + P_{ij} U_{i+1,j}
\]

For \( i=1 \),

\[
U_{1j} = q_{1j} \left[ \sum_{k=1}^{M} U_{0k} \right] + r_{1j} U_{1j} + P_{1j} U_{2j}
\]

For \( i=M-1 \),

\[
U_{(M-1)j} = \sum_{k=1}^{M} U_{M-k,j}
\]
For \( i = M - 1 \),
\[
\begin{align*}
U_{M, j} & = \rho_{M, j} \delta_{M, j} + \gamma_{M, j} U_{M-1, j} + \eta_{M, j} U_{M-2, j} \\
\end{align*}
\]

Now the problem boils down to solving a system of linear equations with a tridiagonal structure. The calculation is a bit technical though not difficult, once the right ideas are introduced.

First we'll strengthen the recursive structure by noting that we can make the equations for \( i = 1, M - 1 \) look the same as for all other states, if we simply define:

\[
\begin{align*}
U_{M, 0} &= 0 & U_{M, M} &= 1 \\
U_{0, 0} &= 1 & U_{0, M} &= 0
\end{align*}
\]

By introducing this extended definition of \( U \), we have to solve the following equations:

\[
\begin{align*}
U_{i, j} &= q_i U_{i-1, j} + \gamma_i U_{i, j} + \rho_i U_{i, j+1} \quad j \in \{ 0, \ldots, M-1 \} \\
U_{0, 0} &= 1 & U_{0, M} &= 0 & U_{M, 0} &= 0 & U_{M, M} &= 1
\end{align*}
\]

These are just linear difference equations, whose solution technique can be found, for example, in Lawler Ch. 0. Because our coefficients are fairly arbitrary (though they have a tridiagonal structure), simple approaches don’t necessarily work so well.

Note the equations for \( j = 0 \) and \( j = M \) decouple so we can just solve them separately. But we can solve the \( j = 0 \) equations because

\[
U_{1, M} = 1 - U_{1, 0}
\]
separately. But we can solve the $j=0$ equations because

$$V_{j=M} = -V_{j=0}$$

We can show that these equations have a special structure that can be summed "by quadrature." if we note by using

$$r_i = 1 - p_i - q_i$$

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

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

We'll now apply a trick that one would use differential equations $y'' = a(x) y'(x)$. Solve by introducing the function $v = y'$, and then solve first order equation for $v$.

$$V_i = U_{i} - U_{i-1} \quad \text{for} \quad 1 \leq i \leq M$$

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

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

By induction:

$$V_i = \prod_{1 \leq k \leq i} \frac{q_k}{p_k} V_1$$

How do we get a boundary condition so we can solve for $V_1$? Check that by using telescoping sum ideas:

$$\sum_{i=1}^{M} V_i = \sum_{i=1}^{M} (U_i - U_{i-1})$$
\[ V_i = \sum_{j=1}^{n} (V_{ij0} - V_{i,j-1}) \]

\[ = V_{M_{ij0}} - V_{ij,0} = 0 \]

\[ \sum_{j=1}^{M} f_j V_i = -1 \]

\[ V_i = \frac{-1}{\sum_{j=1}^{M} f_j} \]

\[ V_{ij} = \sum_{k=1}^{M} f_k \]

Sum over these to get \( U_i \).

\[ U_{i+1} = \sum_{k=1}^{M} V_k \]

\[ = \sum_{l=1}^{i} \sum_{k=1}^{M} f_k \]

\[ V_{00} = 1 \]

\[ 1 - \sum_{j=1}^{i} f_j \]
Similar techniques yield the mean time until absorption via analytical formulas; see the texts.