TRANSIENT PROPERTIES

2.1. Transition probabilities

Discrete-time Markov chains possess the Markov property. This means that we only need the conditional probabilities

$$\Pr\{X_{n+1} = j \mid X_n = i\} \,, \tag{2.1}$$

for each n, to be able to go from one time to the next. These conditional probabilities could vary from one time to another, but we are often interested in problems for which the conditional probabilities remain constant.

Definition 2.1: A discrete-time Markov chain is said to be *stationary* or *homogeneous* (in time) if the probability of going from one state to another is independent of time. That is, for all states i and j,

$$\Pr\{X_{n+1} = j \mid X_n = i\} = \Pr\{X_{n+k+1} = j \mid X_{n+k} = i\} \tag{2.2}$$

for $k = -n, -(n-1), \ldots, -1, 0, 1, 2, \ldots$. If this condition fails, the Markov chain is *nonstationary.*

For the remainder of this quarter, you should always assume that I am talking about Markov chains that are homogenous or stationary. For homogeneous Markov chains, we will write

$$p_{ij} \equiv \Pr\{X_{n+1} = j \mid X_n = i\} \tag{2.3}$$

for the *one-step transition probability* from i to j. Since the p_{ij} are condition probabilities, they satisfy the conditions

$$p_{ij} \geq 0 \quad \forall\, i, j \tag{2.4}$$

and

$$\sum_{j \in S} p_{ij} = 1 \,, \quad i = 1, 2, \ldots \,. \tag{2.5}$$

The last condition simply implies that, if you start in state i, you must end up somewhere. For fixed i, the list $\{p_{ij}\}$ is a *probability distribution* for the possible destinations.

For a finite, discrete-time, Markov chain with m states, the transition probabilities may be arrayed in an $m \times m$ *transition matrix,*

$$\mathbf{P} \equiv [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ . & . & . & . \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}. \tag{2.6}$$

This matrix contains all of the relevant information regarding transitions from one state to another.

Please note, in light of my earlier comments, that

(1) all of the entries of the transition matrix are nonnegative and

(2) the entries in each row sum to one.

Any square matrix that satisfies these two conditions is called a *stochastic matrix*. Transition matrices are examples of stochastic matrices.

A homogeneous, discrete-time, Markov chain can also be represented graphically using a transition diagram. A transition diagram is a directed graph (or *digraph*) with one node for each state in S and a directed edge from node i to node j for each $p_{ij} > 0$. There are self-loops, from node i back to itself, for each $p_{ii} > 0$.

Example:

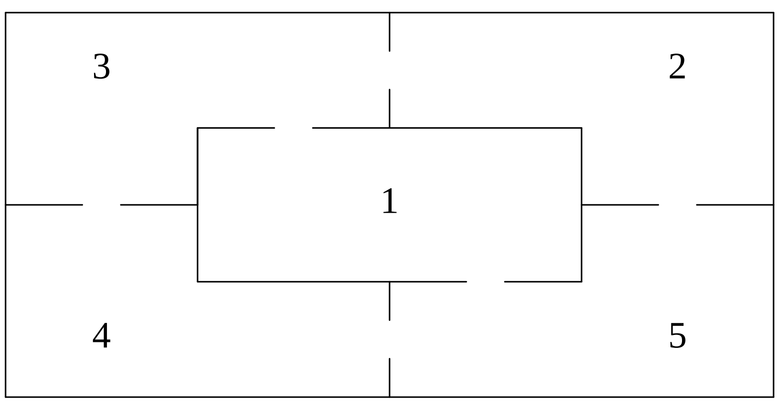

Figure 2.1 A maze

Imagine that an insect that occupies a maze of rooms (see Figure 2.1). Let us assume that time is measured in units of moves, i.e., that a transition occurs whenever the insect moves between rooms, and that each exit from a room is equally likely to be chosen.

The transition matrix for this discrete-time Markov chain is

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \end{bmatrix}. \qquad (2.7)$$

The corresponding digraph is

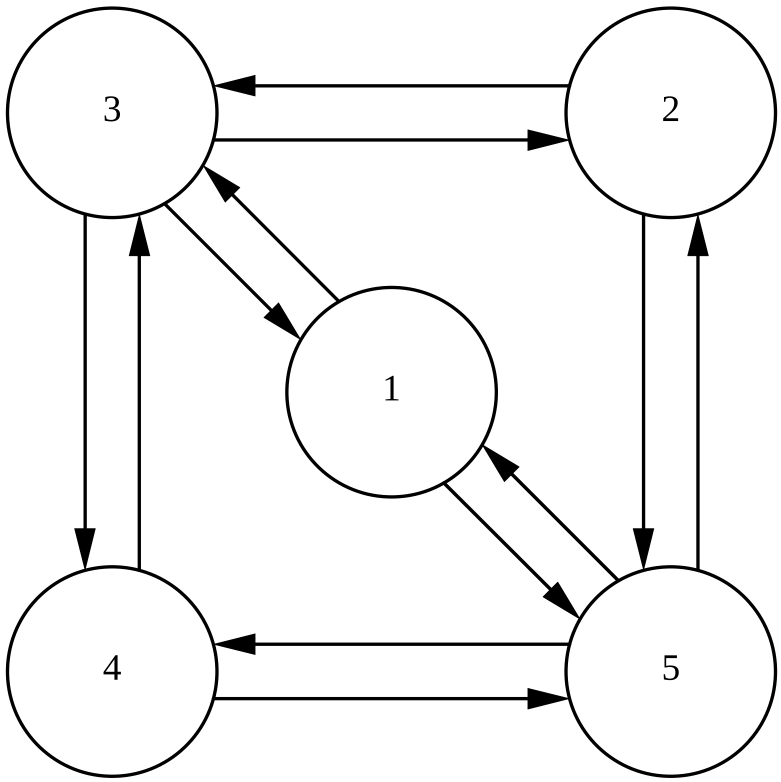

Figure 2.2 Maze digraph

What proportion of its time does the insect spend in each room? We will see that, in contrast to our desert plant communities example, the time spent in each room by our insect does not approach a stationary distribution. Perhaps you can figure out what's different by looking at the transition diagram.

Do keep in mind that the state space of a Markov chain may be countably infinite rather than finite. This can have a significant effect on the appearance of your transition matrix and transition diagram.

Example: Consider a particle that moves on a infinite, one-dimensional lattice. Let us assume that the lattice points are labeled $\ldots, -2, -1, 0, 1, 2, \ldots$

and that the particle undergoes a simple, unrestricted random walk: it moves one step to the right, with probability p, or one step to the left, with probability $q = 1 - p$, at each time step. The transition diagram for this discrete-time Markov chain is

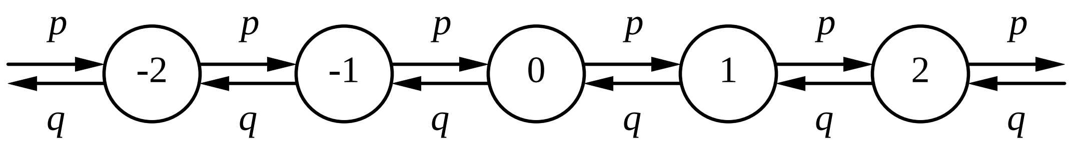

Figure 2.3 Asymmetric random walk

and the transition matrix is

$$\begin{array}{c} \cdot \\ -2 \\ -1 \\ 0 \\ +1 \\ +2 \\ \cdot \end{array} \left[\begin{array}{ccccccc} \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \\ \cdots & 0 & p & 0 & 0 & 0 & \cdots \\ \cdots & q & 0 & p & 0 & 0 & \cdots \\ \cdots & 0 & q & 0 & p & 0 & \cdots \\ \cdots & 0 & 0 & q & 0 & p & \cdots \\ \cdots & 0 & 0 & 0 & q & 0 & \cdots \\ \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \end{array} \right]. \tag{2.8}$$

Example: Consider a gambler who bets one dollar per game and either wins or loses each game with probabilities p and q. The state of the gambler is the number of dollars in his or her possession. The gambler is ruined (and can no longer gamble) if he or she is in state zero and has no money left. I will also assume that there is no limit on how wealthy this gamber can be. This is an example of a restricted random walk on a semi-infinite domain. The restriction is the presence of an absorbing boundary state. The transition diagram for this discrete-time Markov chain is

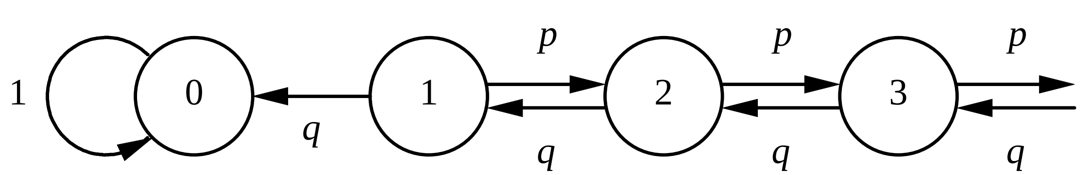

Figure 2.4 Gambler's ruin

and the transition matrix is

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ q & 0 & p & 0 & 0 & \dots \\ 0 & q & 0 & p & 0 & \dots \\ 0 & 0 & q & 0 & p & \dots \\ 0 & 0 & 0 & q & 0 & \dots \\ . & . & . & . & . & \dots \end{bmatrix} . \tag{2.9}$$

An obvious question is how long can the gambler expect to play and how does this depend on the size of his or her initial bank roll ?

For the time being, I will focus on finite chains, but I will eventually return to random walks with countably infinite state spaces. Some of these random walks are similar to some of the continuous-time processes that Hong will talk about later.

The transition matrix **P** and the transition diagram tell us about conditional probabilities. For many problems, we are also interested in *absolute* probabilities. In this case, we must also specify an initial probability distribution $\{u_i^{(0)}\}$, where

$$u_i^{(0)} = \Pr\{X_0 = i\} . \tag{2.10}$$

The notation here is far from ideal. We are letting the subscript on the left refer to the state and the superscript, in parentheses, refer to time. Clearly,

$$\sum_{i=1}^{m} u_i^{(0)} = 1$$

for a chain with m states. If we, in turn, let $u_j^{(1)}$ represent the absolute probability that state j is occupied after one step, it quickly follows that

$$u_j^{(1)} = \sum_{i=1}^{m} u_i^{(0)} p_{ij} . \tag{2.12}$$

Of course, all of this can be written, more succinctly, using vectors and matrices. If we let

$$\mathbf{u}^{(0)} = (u_1^{(0)}, u_2^{(0)}, \dots, u_m^{(0)}) \tag{2.13}$$

be a starting *probability vector* and

$$\mathbf{u}^{(1)} = (u_1^{(1)}, u_2^{(1)}, \dots, u_m^{(1)}) , \tag{2.14}$$

we may instead write

$$\mathbf{u}^{(1)} = \mathbf{u}^{(0)} \mathbf{P} , \tag{2.15}$$

where **P** is our transition matrix. In general,

$$\mathbf{u}^{(n+1)} = \mathbf{u}^{(n)}\,\mathbf{P}\,. \tag{2.16}$$

2.2. The *n*-step transition matrix

For the initial distribution $\mathbf{u}_0$, we have seen that the distribution after one year is just

$$\mathbf{u}^{(1)} = \mathbf{u}^{(0)}\,\mathbf{P}\,. \tag{2.17}$$

After two years,

$$\mathbf{u}^{(2)} = \mathbf{u}^{(1)}\,\mathbf{P} = \mathbf{u}^{(0)}\,\mathbf{P}\,\mathbf{P} = \mathbf{u}^{(0)}\,\mathbf{P}^{2}\,. \tag{2.18}$$

After three years,

$$\mathbf{u}^{(3)} = \mathbf{u}^{(2)}\,\mathbf{P} = \mathbf{u}^{(1)}\,\mathbf{P}\,\mathbf{P} = \mathbf{u}^{(0)}\,\mathbf{P}\,\mathbf{P}\,\mathbf{P} = \mathbf{u}^{(0)}\,\mathbf{P}^{3}\,. \tag{2.19}$$

Continuing on in this way, we see that after n years,

$$\mathbf{u}^{(n)} = \mathbf{u}^{(0)}\,\mathbf{P}^{n}\,. \tag{2.20}$$

The matrix on the far right of equation (2.20) is an *n-step transition matrix*, $\mathbf{P}^{(n)}$; the elements of this matrix are the transition probabilities from states at time 0 to states at time n. Equation (2.20) makes clear that the n-step transition matrix $\mathbf{P}^{(n)}$ is just the nth power of the transition matrix **P**,

$$\mathbf{P}^{(n)} = \mathbf{P}^{n}\,. \tag{2.21}$$

In later lectures, I will often need to refer to the individual elements of the n-step transition matrix.

Definition 2.2: The *n-step transition probability,*

$$p_{ij}^{(n)} \equiv \Pr\{X_n = j \mid X_0 = i\}\,, \tag{2.22}$$

is the probability of moving from state i to state j in n time steps.

For the case $n = 0$,

$$p_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \tag{2.23}$$

where δ_{ij} is the Kronecker delta. For $n = 1$, $p_{ij}^{(1)} = p_{ij}$. In matrix form, $\mathbf{P}^{(0)} = \mathbf{I}$, where **I** is the identity matrix and $\mathbf{P}^{(1)} = \mathbf{P}$.

The n-step transition probabilities are often related to m-step and the $(n - m)$-step transition probabilities by mean of the *Chapman–Kolmogorov equations*

$$p_{ij}^{(n)} = \sum_{k \in S} p_{ik}^{(m)} \, p_{kj}^{(n-m)} , \quad 0 < m < n . \tag{2.24}$$

The left-hand side is the probability of going from i to j in n steps. The right-side side says we can think of this as the sum, over all k, of the probability of going from i to k in m steps and from k to j in $(n - m)$ step.

I hope the Chapman–Kolmogorov equations seem intuitive and reasonable to you. Even so, let me prove these equtions since the proof shows how conditional probability, the Markov property, and time homogeneity are often pulled together in the theory of Markov chains.

Now, I am sure that Hong taught you that

$$\Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}} = \frac{\Pr\{A, B\}}{\Pr\{B\}} . \tag{2.25}$$

In a moment, I will need the general equality

$$\Pr\{A \cap B \mid C\} = \Pr\{A \mid B \cap C\} \; \Pr\{B \mid C\} . \tag{2.26}$$

To see why this is true, let us start with the right-hand side of equation (2.26) and successively condition,

$$\begin{aligned} \Pr\{A \mid B \cap C\} \; \Pr\{B \mid C\} &= \frac{\Pr\{A \cap B \cap C\}}{\Pr\{B \cap C\}} \, \frac{\Pr\{B \cap C\}}{\Pr\{C\}} \qquad (2.27) \\ &= \frac{\Pr\{A \cap B \cap C\}}{\Pr\{C\}} \\ &= \Pr\{A \cap B \mid C\} \end{aligned}$$

We can now verify (prove) the Chapman-Kolmogorov equations by starting with

$$p_{ij}^{(n)} = \Pr\{X_n = j \mid X_0 = i\} . \tag{2.28}$$

Let me partition this conditional probability over mutually exclusive events at the mth stage,

$$p_{ij}^{(n)} = \sum_{k \in S} \Pr\{X_n = j, \, X_m = k \mid X_0 = i\} .$$

But, by successive-conditioning equality (2.26),

$$\Pr\{A \cap B \mid C\} = \Pr\{A \mid B \cap C\} \Pr\{B \mid C\} , \tag{2.29}$$

it now follows that

$$p_{ij}^{(n)} = \sum_{k \in S} \Pr\{X_n = j \mid X_m = k,\, X_0 = i\} \Pr\{X_m = k \mid X_0 = i\} . \tag{2.30}$$

Using the Markov property,

$$p_{ij}^{(n)} = \sum_{k \in S} \Pr\{X_n = j \mid X_m = k\} \Pr\{X_m = k \mid X_0 = i\} . \tag{2.31}$$

Finally, using time homogeneity,

$$p_{ij}^{(n)} = \sum_{k \in S} \Pr\{X_{n-m} = j \mid X_0 = k\} \Pr\{X_m = k \mid X_0 = i\} . \tag{2.32}$$

so that

$$p_{ij}^{(n)} = \sum_{k \in S} p_{kj}^{(n-m)}\, p_{ik}^{(m)} \tag{2.33}$$

$$= \sum_{k \in S} p_{ik}^{(m)}\, p_{kj}^{(n-m)}$$

Of course, the Chapman–Kolmogorov equations appear much simpler if we use matrices. Then

$$\mathbf{P}^{(n)} = \mathbf{P}^{(m)}\, \mathbf{P}^{(n-m)} \tag{2.34}$$

and all that we are really saying is that

$$\mathbf{P}^{n} = \mathbf{P}^{m}\, \mathbf{P}^{n-m} . \tag{2.35}$$

2.3. Eigenvalues and matrix powers

The transient behavior of a discrete-time Markov chain is governed by the n-step transition matrix $\mathbf{P}^{(n)} = \mathbf{P}^{n}$. It can, of course, be quite tedious to take powers of a matrix. Fortunately, there is some wonderful software out there nowadays that lets you evaluate powers of matrices. In addition, there are several, reasonably efficient, methods for evaluating $\mathbf{P}^n$ analytically. Most of these methods require knowledge of the eigenvalues and, in some cases, of the eigenvectors of the $\mathbf{P}$.

Definition 2.3: A number λ is an *eigenvalue* of the square matrix $\mathbf{P}$ if there exists a nonzero column vector $\mathbf{v}$ that satisfies

$$\mathbf{P}\,\mathbf{v} = \lambda\, \mathbf{v} . \tag{2.36}$$

The vector $\mathbf{v}$ is then called a *right eigenvector* of $\mathbf{P}$ (corresponding to the eigenvalue λ). A nonzero row vector **u** that satisfies

$$\mathbf{u}\,\mathbf{P} \;=\; \lambda\,\mathbf{u} \tag{2.37}$$

is called a *left eigenvector* of **P**. Thought problem: Convince yourself that left and right eigenvectors occur for the same values of λ.

Example: Consider the matrix

$$\mathbf{P} \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{bmatrix}. \tag{2.38}$$

The transpose of (1, 1, 1) is a right eigenvector since

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{bmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;=\; 1\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{2.39}$$

The row vector (1, 0, 0), in turn, is a left eigenvector since

$$(1,\,0,\,0)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{bmatrix} \;=\; 1\,(1,\,0,\,0)\;. \tag{2.40}$$

Both of these eigenvectors have eigenvalue 1.

There are, of course, other eigenvalues for this matrix. They satisfy

$$|\mathbf{P} - \lambda\,\mathbf{I}| \;=\; 0\;. \tag{2.41}$$

This quickly leads to the characteristic equation

$$(1-\lambda)\,(\tfrac{1}{2}-\lambda)\,(-\lambda) \;=\; 0 \tag{2.42}$$

and the eigenvalues

$$\lambda_1 = 1\,,\;\; \lambda_2 = \frac{1}{2}\,,\;\; \lambda_3 = 0\;. \tag{2.43}$$

What can we say about the eigenvalues of a Markov chain in general? Quite a bit.

Theorem 2.1: One of the eigenvalues of a stochastic matrix **P** is always $\lambda = 1$. All of the eigenvalues satisfy $|\lambda| \leq 1$.

Proof: Since every row of **P** sums to one,

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ \cdot \\ 1 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 1 \\ \cdot \\ 1 \end{pmatrix} \tag{2.44}$$

so that 1 is an eigenvalue of **P** with a right eigenvector that is a column of 1's.

To prove that $|\lambda| \leq 1$, let λ be a root of

$$|\mathbf{P} - \lambda \mathbf{I}| = 0 \tag{2.45}$$

with a right eigenvector that satisfies

$$(\mathbf{P} - \lambda \mathbf{I}) \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ v_m \end{pmatrix} = \mathbf{0} \, . \tag{2.46}$$

If we look at each row of this last equation, we see that

$$\lambda \, v_i = p_{i1} \, v_1 + p_{i2} \, v_2 + \cdots + p_{im} \, v_m \, , \quad i = 1, \ldots, m \tag{2.47}$$

Take the modulus of each side. Since the transition probabilities p_{ij} are non-negative,

$$|\lambda \, v_i| \leq \sum_{j=1}^{m} p_{ij} \, |v_j| , \quad i = 1, \ldots, m \, . \tag{2.48}$$

Now, if v_k is the largest component, in modulus, of our right eigenvector,

$$|\lambda \, v_i| \leq |v_k| , \quad i = 1, \ldots, m \, . \tag{2.49}$$

This last equation is true for every i and so it is certainly true for $i = k$. For $i = k$,

$$|\lambda \, v_k| \leq |v_k| , \tag{2.50}$$

but this now implies that

$$|\lambda| \leq 1 \, . \tag{2.51}$$

This proves our theorem.

How does knowing the eigenvalues and/or eigenvectors of **P** help us evaluate $\mathbf{P}^n$? A number of clever methods allow you to efficiently evaluate

$\mathbf{P}^n$ using the eigenvalues or the eigenvalues and eigenvectors of $\mathbf{P}$. My own favorite method uses the Cayley–Hamilton theorem from linear algebra. I will instead show you a method that is quite common in the Markov chain literature. This method involves diagonalizing the transition matrix.

An $m \times m$ transition matrix is diagonalizable if and only if it possesses m linearly independent (right) eigenvectors. This will always occur if all of the eigenvalues of your transition matrix are distinct. It may even occur if some your eigenvalues are equal but you are lucky. If you are unlucky, you may need to use generalized eigenvectors and Jordan canonical forms. Jordan canonical forms are more, however, than I need for this course.

Let us assume, therefore, that $\mathbf{P}$ has m distinct eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_m$. I will write the corresponding right eigenvectors as $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. These eigenvectors satisfy

$$(\mathbf{P} - \lambda_i \mathbf{I})\, \mathbf{v}_i = \mathbf{0}\,, \quad i = 1, 2, \ldots, m\,. \tag{2.52}$$

Now construct the modal matrix $\mathbf{M}$,

$$\mathbf{M} = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \ldots \;\; \mathbf{v}_m]\,, \tag{2.53}$$

that has the right eigenvectors as its columns. By matrix multiplication,

$$\begin{aligned} \mathbf{P}\mathbf{M} &= \mathbf{P}\,[\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \ldots \;\; \mathbf{v}_m] = [\mathbf{P}\mathbf{v}_1 \;\; \mathbf{P}\mathbf{v}_2 \;\; \ldots \;\; \mathbf{P}\mathbf{v}_m] \\ &= [\lambda_1 \mathbf{v}_1 \;\; \lambda_2 \mathbf{v}_2 \;\; \ldots \;\; \lambda_m \mathbf{v}_m] \\ &= [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \ldots \;\; \mathbf{v}_m] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ . & . & \ldots & . \\ 0 & 0 & \ldots & \lambda_m \end{bmatrix} \\ &= \mathbf{M}\mathbf{D}\,, \end{aligned} \tag{2.54}$$

where

$$\mathbf{D} \equiv \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ . & . & \ldots & . \\ 0 & 0 & \ldots & \lambda_m \end{bmatrix} \tag{2.55}$$

is the diagonal matrix of eigenvalues. We can now multiply each side of

$$\mathbf{P}\mathbf{M} = \mathbf{M}\mathbf{D} \tag{2.56}$$

on the right by the inverse matrix $\mathbf{M}^{-1}$ to obtain

$$\mathbf{P} \;=\; \mathbf{M}\,\mathbf{D}\,\mathbf{M}^{-1}\;. \tag{2.57}$$

The last equation can now be used to determine powers of **P**. For example,

$$\begin{aligned} \mathbf{P}^2 \;&=\; (\mathbf{M}\,\mathbf{D}\,\mathbf{M}^{-1})\,(\mathbf{M}\,\mathbf{D}\,\mathbf{M}^{-1}) \\ &=\; (\mathbf{M}\,\mathbf{D})\,(\mathbf{M}^{-1}\,\mathbf{M})\,(\mathbf{D}\,\mathbf{M}^{-1}) \\ &=\; (\mathbf{M}\,\mathbf{D})\,\mathbf{I}\,(\mathbf{D}\,\mathbf{M}^{-1}) \\ &=\; \mathbf{M}\,\mathbf{D}^2\,\mathbf{M}^{-1}\;. \end{aligned} \tag{2.58}$$

Higher powers can be generated in a similar manner,

$$\mathbf{P}^n \;=\; \mathbf{M}\,\mathbf{D}^n\,\mathbf{M}^{-1}\;. \tag{2.59}$$

Please note that

$$\mathbf{D}^n \;\equiv\; \begin{bmatrix} \lambda_1^n & 0 & \ldots & 0 \\ 0 & \lambda_2^n & \ldots & 0 \\ . & . & \ldots & . \\ 0 & 0 & \ldots & \lambda_m^n \end{bmatrix} \tag{2.60}$$

has a particularly simple form.

As part of the above process, you must compute the inverse matrix $\mathbf{M}^{-1}$. A simple way of doing this is to apply elementary row operations to the augmented matrix $[\mathbf{M}|\mathbf{I}]$ to get $[\mathbf{I}|\mathbf{M}^{-1}]$. For small matrices, you can also use the cofactor matrix and adjugate (or classical adjoint) matrix of the original matrix or whatever other means of matrix inversion you have learned.

Example: (To be read by class)

Consider the transition matrix

$$\mathbf{P} \;=\; \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}. \tag{2.61}$$

The eigenvalues of **P** satisfy the characteristic equation

$$|\mathbf{P} - \lambda\,\mathbf{I}| \;=\; \begin{vmatrix} 1/4-\lambda & 1/2 & 1/4 \\ 1/2 & 1/4-\lambda & 1/4 \\ 1/4 & 1/4 & 1/2-\lambda \end{vmatrix} \;=\; 0\,. \tag{2.62}$$

Before expanding this determinant, it is good to remember that we can always a multiple of one row to another row and not change the value of the determinant. Adding the second and third rows to the first row gives

$$\begin{vmatrix} 1-\lambda & 1-\lambda & 1-\lambda \\ 1/2 & 1/4-\lambda & 1/4 \\ 1/4 & 1/4 & 1/2-\lambda \end{vmatrix} \;=\; 0 \tag{2.63}$$

or

$$(1-\lambda)\begin{vmatrix} 1 & 1 & 1 \\ 1/2 & 1/4-\lambda & 1/4 \\ 1/4 & 1/4 & 1/2-\lambda \end{vmatrix} \;=\; 0\,. \tag{2.64}$$

Subtracting one quarter of the first row from the third row and one half of the first row from the second row now gives us

$$(1-\lambda)\begin{vmatrix} 1 & 1 & 1 \\ 0 & -1/4-\lambda & -1/4 \\ 0 & 0 & 1/4-\lambda \end{vmatrix} \;=\; 0\,. \tag{2.65}$$

It is now easy to expand the determinant:

$$(1-\lambda)\,(1/4-\lambda)\,(-1/4-\lambda) \;=\; 0\,. \tag{2.66}$$

The eigenvalues are clearly

$$\lambda_1 = 1,\;\; \lambda_2 = \frac{1}{4}\,,\;\; \lambda_3 = -\,\frac{1}{4}\;. \tag{2.67}$$

As we expected, one of the eigenvalues is $\lambda = 1$ and all of the eigenvalues satisfy $|\lambda| \le 1$. Fortunately, the eigenvalues are distinct.

The right eigenvectors $\mathbf{v}_i$ (i = 1, 2, 3) satisfy

$$(\mathbf{P} - \lambda_i\,\mathbf{I})\,\mathbf{v}_i \;=\; \mathbf{0}\,. \tag{2.68}$$

The first eigenvalue, $\lambda_1 = 1$, has, we know, the right eigenvector

$$\mathbf{v}_1 \;=\; \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{2.69}$$

For the second eigenvalue, $\lambda_2 = 1/4$, we have the linear system

$$\begin{bmatrix} 0 & 1/2 & 1/4 \\ 1/2 & 0 & 1/4 \\ 1/4 & 1/4 & 1/4 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.70}$$

and the eigenvector

$$\mathbf{v}_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}. \tag{2.71}$$

For the third eigenvalue, $\lambda_3 = -1/4$, we have the linear system

$$\begin{bmatrix} 1/2 & 1/2 & 1/4 \\ 1/2 & 1/2 & 1/4 \\ 1/4 & 1/4 & 3/4 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.72}$$

and the eigenvector

$$\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \tag{2.73}$$

The modal matrix is

$$\mathbf{M} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{bmatrix}. \tag{2.74}$$

We also need the inverse, $\mathbf{M}^{-1}$ of the modal matrix. $\mathbf{M}$ is a small matrix and so I will use a fairly inefficient method to determine this inverse. Let me first take the cofactor matrix,

$$\mathbf{M}^c = [(-1)^{i+j} \det \mathbf{M}(i|j)] = \begin{bmatrix} 2 & -1 & 3 \\ 2 & -1 & -3 \\ 2 & 2 & 0 \end{bmatrix}, \tag{2.75}$$

where $\mathbf{M}(i|j)$ stands for the submatrix of bold M obtained by deleting row i and column j. Now, I take the transpose of the cofactor matrix to form the adjugate (or adjoint),

$$\mathbf{M}^a = (\mathbf{M}^c)^T = \begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & 2 \\ 3 & -3 & 0 \end{bmatrix}. \tag{2.76}$$

Finally,

$$\mathbf{M}^{-1} = \frac{1}{\det \mathbf{M}} \mathbf{M}^a = \frac{1}{6}\begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & 2 \\ 3 & -3 & 0 \end{bmatrix}. \qquad (2.77)$$

With all these ingredients in hand, I can now write

$$\mathbf{P}^n = \mathbf{M}\,\mathbf{D}^n\,\mathbf{M}^{-1} \qquad (2.78)$$

or

$$\mathbf{P}^n = \frac{1}{6}\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & (1/4)^n & 0 \\ 0 & 0 & (-1/4)^n \end{bmatrix}\begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & 2 \\ 3 & -3 & 0 \end{bmatrix}. \qquad (2.79)$$

I will let you do that last bit of matrix multiplication. For $n = \infty$,

$$\mathbf{P}^\infty = \frac{1}{6}\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & 2 \\ 3 & -3 & 0 \end{bmatrix} \qquad (2.80)$$

$$= \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$