Chapter 2
Finite Markov Chains

Finite Markov chains are processes with finitely many (typically only a few) states on a nominal scale (with arbitrary labels). Time runs in discrete steps, such as day 1, day 2, ..., and only the most recent state of the process affects its future development (the Markovian property). Our first objective is to compute the probability of being in a certain state after a specific number of steps. This is followed by investigating the process's *long-run* behavior.

2.1 A Few Examples

To introduce the idea of a Markov chain, we start with a few examples.

Example 2.1. Suppose that weather at a certain location can be sunny, cloudy, or rainy (for simplicity, we assume it changes only on a daily basis). These are called the states of the corresponding process.

The simplest model assumes the type of weather for the next day is chosen randomly from a distribution such as

Type	S	C	R
Pr	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

(which corresponds to rolling a biased die), *independently* of today's (and past) conditions (in Chap. 1, we called this a *trivial* stochastic process).

Weather has a tendency to resist change, for instance, sunny $\to$ sunny is more likely than sunny $\to$ rainy (incidentally, going from X_n to X_{n+1} is called a transition). Thus, we can improve the model by letting the distribution depend on the current state. We would like to organize the corresponding information in the following transition probability matrix (TPM):

J. Vrbik and P. Vrbik, *Informal Introduction to Stochastic Processes with Maple*, Universitext, DOI 10.1007/978-1-4614-4057-4_2,

$$\begin{array}{c|ccc|}
 & \text{S} & \text{C} & \text{R} \\
\hline
\text{S} & 0.6 & 0.3 & 0.1 \\
\text{C} & 0.4 & 0.5 & 0.1 \\
\text{R} & 0.3 & 0.4 & 0.3 \\
\hline
\end{array}$$

where the rows correspond to today's weather and the columns to the type of weather expected tomorrow (each row must consist of a complete distribution; thus all the numbers must be nonnegative and sum to 1). Because tomorrow's value is not *directly* related to yesterday's (or earlier) value, the process is *Markovian*.

There are several issues to investigate, for example:

1. If today is sunny, how do we compute the probability of its being rainy *two* days from now (three days from now, etc.)?
2. In the *long run*, what will be the proportion of sunny days?
3. How can we improve the model to make the probabilities depend on today's *and* yesterday's weather?

◇

To generate a possible realization of the process (starting with sunny weather) using MAPLE, we type

> *with(LinearAlgebra): with(plots): with(Statistics):*

{Henceforth we will assume these packages are loaded and will not explicitly call them (see "Library Commands" in Chap. 13).}

$$> \mathbb{P}_1 := \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.5 & 0.1 \\ 0.3 & 0.4 & .3 \end{bmatrix} :$$

$> (j, res) := (1, 1):$

$> \textbf{for } i \textbf{ from } 1 \textbf{ to } 25 \textbf{ do}$

$\quad j := Sample\left(ProbabilityTable\left(convert\left(\mathbb{P}_1[j], list\right)\right), 1\right)_1;$

$\quad j := \operatorname{trunc}(j);$

$\quad res := res, j;$

$\textbf{end do}:$

$> subs\left(1 = S, 2 = C, 3 = R, [res]\right);$

$$[S, S, C, R, S, S, S, C, C, C, R, S, S, R, R, R, C, S, S, S, S, C, S, S, S, S]$$

(The MAPLE worksheets can be downloaded from extras.springer.com.)

Example 2.2. Alice and Bob repeatedly bet \$1 on the flip of a coin. The potential states of this process are all integers, the INITIAL STATE (usually denoted X_0) may be taken as 0, and the TPM is now infinite, with each row looking like this:

$$\cdots \quad 0 \quad 0 \quad \tfrac{1}{2} \quad 0 \quad \tfrac{1}{2} \quad 0 \quad 0 \quad \cdots$$

This is an example of a so-called INFINITE MARKOV CHAIN. For the time being, we would like to investigate FINITE MARKOV CHAINS (FMCs) only, so we modify this example assuming each player has only \$2 to play with:

$$\begin{array}{c|ccccc|} & -2 & -1 & 0 & 1 & 2 \\ \hline -2 & 1 & 0 & 0 & 0 & 0 \\ -1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 1 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 2 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

The states are labeled by the amount of money Alice has won (or lost) so far. The two "end" states are called ABSORBING states. They represent the situation of one of the players running out of money; the game is over and the Markov chain is stuck in an absorbing state for good.

Now the potential *questions* are quite different:

1. What is the probability of Alice winning over Bob, especially when they start with different amounts or the coin is slightly biased?
2. How long will the game take (i.e., the distribution, expected value, and standard deviation of the number of transitions until one of the players goes broke)?

Again, we can simulate one possible outcome of playing such a game using MAPLE:

$$> \mathbb{P}_2 := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}:$$

$> (j, res) := (3, 3):$

$>$ **for** i **from** 1 **while** $(j > 1$ **and** $j < 5)$ **do**

$\quad j := Sample\,(ProbabilityTable\,(convert\,(\mathbb{P}_2[j], list)), 1)_1;$

$\quad j := \mathrm{trunc}(j);$

$\quad res := res, j:$

end do:

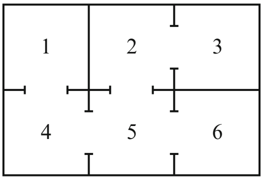
$> subs(1 = -2, 2 = -1, 3 = 0, 4 = 1, 5 = 2, [res]);$

$$[0, -1, 0, 1, 0, 1, 2]$$

(Note Alice won six rounds.)

◇

Example 2.3. Assume there is a mouse in a maze consisting of six compartments, as follows:



Here we define a transition as happening whenever the mouse changes compartments. The TPM is (assuming the mouse chooses one of the available exits perfectly randomly)

$$\begin{array}{c|cccccc|} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 & 0 \\ 4 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 5 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 6 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline \end{array}.$$

Note this example is what will be called PERIODIC (we can return to the same state only in an *even* number of transitions).

A possible REALIZATION of the process may then look like this (taking 1 as the initial state):

$$> \mathbb{P}_3 := \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}:$$

> $(j, res) := (1, 1)$:
> **for** i **from** 1 **to** 30 **do**
$j := Sample\,(ProbabilityTable\,(convert\,(\mathbb{P}_3[j], list)), 1)_1$;
$res := \mathrm{trunc}(j)$;
$res := res, j$;
end do:
> res;

$$1, 4, 1, 4, 1, 4, 5, 6, 5, 2, 3, 2, 3, 2, 5, 4, 1, 4, 1, 4, 1, 4, 1, 4, 5, 4, 1, 4, 5, 6, 5, 2$$

One of the issues here is finding the so called FIXED VECTOR (the relative frequency of each state in a long run), which we discuss in Sect. 2.5.

We modify this example by opening Compartment 6 to the outside world (letting the mouse escape, when it chooses that exit). This would then add a new "Outside" state to the TPM, a state that would be absorbing (the mouse does not return). We could then investigate the probability of the mouse's finding this exit eventually (this will turn out to be 1) and how many transitions it will take to escape (i.e., its distribution and the corresponding mean and standard deviation). ◇

Example 2.4. When repeatedly tossing a coin, we may get something like this:

$$\text{HTHHHTTHTH}\ldots.$$

Suppose we want to investigate the patterns of two consecutive outcomes. Here, the first such pattern is HT, followed by TH followed by HH, etc. The corresponding TPM is

$$\begin{array}{c|cccc|}
 & \text{HH} & \text{HT} & \text{TH} & \text{TT} \\
\hline
\text{HH} & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
\text{HT} & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\text{TH} & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
\text{TT} & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\hline
\end{array}$$

This will enable us to study questions such as the following ones:

1. What is the probability of generating TT *before* HT? (Both patterns will have to be made absorbing.)
2. How long would such a game take (i.e., what is the expected value and standard deviation of the number of flips needed)?

The novelty of this example is the initial setup: here, the very first state will itself be generated by two flips of the coin, so instead of starting in a specific initial state, we are randomly selecting it from the following INITIAL DISTRIBUTION:

State	HH	HT	TH	TT
Pr	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

In Sect. 5.1, we will extend this to cover a general situation of generating a pattern like HTTHH before THHT. ◇

2.2 Transition Probability Matrix

It should be clear from these examples that all we need to describe a Markov chain is a corresponding TPM (all of whose entries are ≥ 0 and whose row sums are equal to 1 – such square matrices are called STOCHASTIC) and the initial state (or distribution).

The one-step TPM is usually denoted by $\mathbb{P}$ and is defined by

$$\mathbb{P}_{ij} \equiv \Pr(X_{n+1} = j \mid X_n = i).$$

In general, these probabilities may depend on n (e.g., the weather patterns may depend on the season, or the mouse may begin to learn its way through the maze). For the Markov chains studied here we assume this does *not* happen, and the process is thus HOMOGENEOUS IN TIME, that is,

$$\Pr(X_{n+1} = j \mid X_n = i) \equiv \Pr(X_1 = j \mid X_0 = i)$$

for all n.

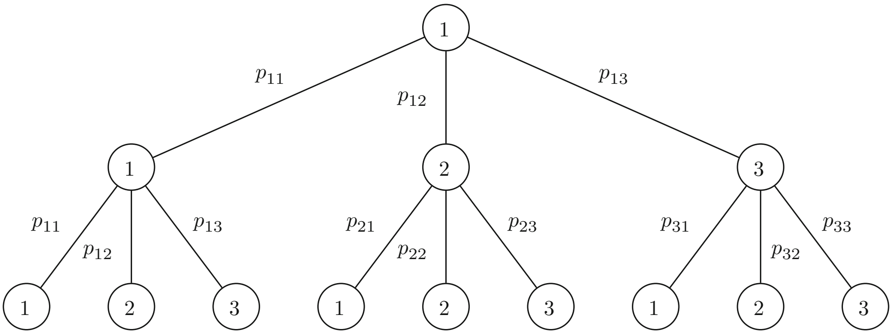
Two-Step (Three-Step, etc.) Transition Probabilities

Example 2.5. Suppose we have a three-state FMC, defined by the following (general) TPM:

$$\mathbb{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}.$$

Given we are in State 1 now, what is the probability that *two* transitions later we will be in State 1? State 2? State 3?

Solution. We draw the corresponding probability tree



and apply the formula of total probability to find the answer $p_{11}p_{11} + p_{12}p_{21} + p_{13}p_{31}$, $p_{11}p_{12} + p_{12}p_{22} + p_{13}p_{32}$, etc. These can be recognized as the $(1,1)$, $(1,2)$, etc. elements of $\mathbb{P}^2$. ◇

One can show that in general the following proposition holds.

Proposition 2.1.

$$\Pr(X_n = j \mid X_0 = i) = (\mathbb{P}^n)_{ij}.$$

Proof. Proceeding by induction, we observe this is true for $n = 1$. Assuming that it is true for $n - 1$, we show it is true for n.

We know that $\Pr(A) = \sum_k \Pr(A \mid C_k)\Pr(C_k)$ whenever $\{C_k\}$ is a partition. This can be extended to $\Pr(A \mid B) = \sum_k \Pr(A \mid B \cap C_k)\Pr(C_k \mid B)$; simply replace the original A by $A \cap B$ and divide by $\Pr(B)$. Based on this generalized formula of total probability (note $X_{n-1} = k$, with all possible values of k, is a partition), we obtain

$$\begin{aligned}&\Pr(X_n = j \mid X_0 = i)\\ &\quad = \sum_k \Pr(X_n = j \mid X_{n-1} = k \cap X_0 = i)\Pr(X_{n-1} = k \mid X_0 = i).\end{aligned}$$

The first term of the last product equals $\Pr(X_n = j \mid X_{n-1} = k)$ (by the Markovian property), which is equal to $\mathbb{P}_{kj}$ (due to time-homogeneity). By the induction assumption, the second term equals $(\mathbb{P}^{n-1})_{ik}$. Putting these together, we get

$$\sum_k \left(\mathbb{P}^{n-1}\right)_{ik} \mathbb{P}_{kj},$$

which corresponds to the matrix product of $\mathbb{P}^{n-1}$ and $\mathbb{P}$. The result thus equals $(\mathbb{P}^n)_{ij}$. □

Example 2.6. (Refer to Example 2.1). If today is cloudy, what is the probability of its being rainy three days from now?

Solution. We must compute the (2nd, 3rd) elements of $\mathbb{P}^3$, or, more efficiently

$$\begin{bmatrix} 0.4 & 0.5 & 0.1 \end{bmatrix} \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.5 & 0.1 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.1 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.5 & 0.1 \end{bmatrix} \begin{bmatrix} 0.12 \\ 0.12 \\ 0.16 \end{bmatrix} = 12.4\%.$$

Note the initial/final state corresponds to the row/column of $\mathbb{P}^3$, respectively. This can be computed more easily by

$$> \mathbb{P}_1 := \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.5 & 0.1 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} :$$

$$> (\mathbb{P}_1)^3_{2,3} ;$$

$$0.1234$$

$\square$

Similarly, if a *record* of several past states is given (such as Monday was sunny, Tuesday was sunny again, and Wednesday was cloudy), computing the probability of rainy on Saturday would yield the same answer (since we can *ignore all but the latest piece of information*).

Now we modify the question slightly: What is the probability of its being rainy on Saturday *and* Sunday? To answer this (labeling Monday as day 0), we first recall

$$\Pr(A \cap B) = \Pr(A) \Pr(B \mid A) \Rightarrow \Pr(A \cap B \mid C) = \Pr(A \mid C) \Pr(B \mid A \cap C),$$

which is the product rule, conditional upon C. Then we proceed as follows:

$$\begin{aligned} &\Pr(X_5 = R \cap X_6 = R \mid X_0 = S \cap X_1 = S \cap X_2 = C) \\ &\quad = \Pr(X_5 = R \cap X_6 = R \mid X_2 = C) \\ &\quad = \Pr(X_5 = R \mid X_2 = C) \Pr(X_6 = R \mid X_5 = R \cap X_2 = C) \\ &\quad = \Pr(X_5 = R \mid X_2 = C) \Pr(X_6 = R \mid X_5 = R) = 0.124 \times 0.3 \\ &\quad = 3.72\%. \end{aligned}$$

To summarize the basic rules of forecasting based on the past record:

1. Ignore all but the latest item of your record.
2. Given this, find the probability of reaching a specific state on the first day of your "forecast."
3. Given this state has been reached, take it to the next day of your forecast.
4. Continue until the last day of the forecast is reached.
5. Multiply all these probabilities.

If an *initial distribution* (say $\mathbf{d}$, understood to be a one-column matrix) is given (for day 0), the probabilities of being in a given State n transitions later are given by the elements of

$$\mathbf{d}^{\mathrm{T}}\mathbb{P}^n,$$

where $\mathbf{d}^{\mathrm{T}}$ is the TRANSPOSE of $\mathbf{d}$ (making it a *one-row* matrix). The result is a one-row matrix of (final-state) probabilities.

Note when $\mathbb{P}$ is stochastic, $\mathbb{P}^n$ is too for any integer n (prove by induction – this rests on the fact a product of *any* two stochastic matrices, say $\mathbb{Q}$ and $\mathbb{P}$, is also stochastic, which can be proven by summing $\sum\limits_k Q_{ik}P_{kj}$ over j).

2.3 Long-Run Properties

We now investigate the long-run development of FMCs, which is closely related to the behavior of $\mathbb{P}^n$ for large n. The simplest situation occurs when all elements of $\mathbb{P}$ are positive (a special case of the so-called REGULAR FMC, defined later).

One can show that in this case $\mathbb{P}^\infty = \lim_{n\to\infty}\mathbb{P}^n$ exists, and all of its rows are identical (this should be intuitively clear: the probability of a sunny day 10 years from now should be practically independent of the initial condition), that is,

$$\mathbb{P}^\infty = \begin{bmatrix} s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 \end{bmatrix}$$

(for a three-state chain), where $\mathbf{s}^{\mathrm{T}} = [s_1, s_2, s_3]$ is called the STATIONARY DISTRIBUTION (the individual components are called stationary probabilities). Later, we will supply a formal proof of this, but let us look at the consequences first.

These probabilities have two interpretations; s_i represents

1. The probability of being in State i after many transitions (this limit is often reached in a handful of transitions);
2. The RELATIVE FREQUENCY of occurrence of State i in the long run (technically the limit of the relative frequency of occurrence when approaching an *infinite run*; again, in practice, a few hundred transitions is usually a good approximation).

By computing individual powers of the TPM for each of our four examples, one readily notices the first (weather) and the last (HT-type patterns) quickly converge to the type of matrix just described; in the latter case, this happens in one step:

$$> \mathbb{P}_1 := \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.5 & 0.1 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}:$$

> **for** i **from** 3 **by** 3 **to** 9 **do**
  $evalm\left(\mathbb{P}_1^i\right);$
end do:

$$\begin{bmatrix} 0.4900 & 0.3869 & 0.1240 \\ 0.4820 & 0.3940 & 0.1240 \\ 0.4700 & 0.3980 & 0.1320 \end{bmatrix}$$

$$\begin{bmatrix} 0.4844 & 0.3906 & 0.1250 \\ 0.4844 & 0.3906 & 0.1250 \\ 0.4842 & 0.3908 & 0.1251 \end{bmatrix}$$

$$\begin{bmatrix} 0.4844 & 0.3906 & 0.1250 \\ 0.4844 & 0.3906 & 0.1250 \\ 0.4844 & 0.3906 & 0.1250 \end{bmatrix}$$

$$> \mathbb{P}_4 := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}:$$

$> \mathbb{P}_4^2;$

$$\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

Knowing the special form of the limiting matrix, there is a shortcut to computing $\mathbf{s}$: $\mathbb{P}^\infty\mathbb{P} = \mathbb{P}^\infty \Rightarrow \mathbf{s}^\mathrm{T}\mathbb{P} = \mathbf{s}^\mathrm{T} \Rightarrow \mathbb{P}^\mathrm{T}\mathbf{s} = \mathbf{s} \Rightarrow (\mathbb{P}^\mathrm{T} - \mathbb{I})\mathbf{s} = \mathbf{0}$. Solving the last set of (homogeneous) equations yields $\mathbf{s}$. Since adding the elements of each row of $\mathbb{P} - \mathbb{I}$ results in $\mathbf{0}$, the matrix (and its transpose) is singular, so there must be at least one nonzero solution to $\mathbf{s}$. For regular FMCs, the solution is (up to a multiplicative constant) unique since the RANK of $\mathbb{P} - \mathbb{I}$ must equal $N - 1$, where N is the total number of possible states.

Example 2.7. Consider our weather example, where

$$\mathbb{P}^{\mathrm{T}} - \mathbb{I} = \begin{bmatrix} -0.4 & 0.4 & 0.3 \\ 0.3 & -0.5 & 0.4 \\ 0.1 & 0.1 & -0.7 \end{bmatrix}.$$

The matrix is of rank 2; we may thus arbitrarily discard one of the three equations. Furthermore, since the solution can be determined up to a multiplicative constant, assuming s_3 is nonzero (as it must be in the regular case), we can set it to 1, eliminating one unknown. We then solve for s_1 and s_2 and multiply $\mathbf{s}$ by a constant, which makes it into a probability vector (we call this step NORMALIZING $\mathbf{s}$). In terms of our example, we get

$$\begin{aligned} -0.4s_1 + 0.4s_2 &= -0.3, \\ 0.3s_1 - 0.5s_2 &= -0.4. \end{aligned}$$

The solution is given by

$$\begin{bmatrix} -0.5 & -0.4 \\ -0.3 & -0.4 \end{bmatrix} \begin{bmatrix} -0.3 \\ -0.4 \end{bmatrix} \frac{1}{0.08} = \begin{bmatrix} \frac{31}{8} \\ \frac{25}{8} \end{bmatrix},$$

together with $s_3 = 1$. Since, at this point, we do not care about the multiplicative factor, we may also present it as $[31, 25, 8]^{\mathrm{T}}$ (the reader should verify this solution meets all *three* equations). Finally, since the final solution must correspond to a probability distribution (the components adding up to 1), all we need to do is normalize the answer, thus:

$$\mathbf{s} = \begin{bmatrix} \frac{31}{64} \\ \frac{25}{64} \\ \frac{8}{64} \end{bmatrix} = \begin{bmatrix} 0.4844 \\ 0.3906 \\ 0.1250 \end{bmatrix}.$$

And, true enough, this agrees with what we observed by taking large powers of the corresponding TPM ◇

Example 2.8. Even though Example 2.3 is not regular (as we discover in the next section), it also has a unique solution to $\mathbf{s}^{\mathrm{T}}\mathbb{P} = \mathbf{s}^{\mathrm{T}}$. The solution is called the FIXED VECTOR, and it still corresponds to the relative frequencies of states in the long run (but no longer to the $\mathbb{P}^{\infty}$ limit). Finding this $\mathbf{s}$ is a bit more difficult now (we must solve a 5×5 set of equations), so let us see whether we can guess the answer. We conjecture that the proportion of time spent in each compartment is proportional to the number of doors to/from it. This would imply $\mathbf{s}^{\mathrm{T}}$ should be proportional to $\begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \end{bmatrix}$, implying

$\mathbf{s}^{\mathrm{T}} = \begin{bmatrix} \frac{1}{10} & \frac{2}{10} & \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{1}{10} \end{bmatrix}$. To verify the correctness of this answer, we must check that

$$\begin{bmatrix} \frac{1}{10} & \frac{2}{10} & \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{1}{10} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

equals $\begin{bmatrix} \frac{1}{10} & \frac{2}{10} & \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{1}{10} \end{bmatrix}$, which is indeed the case. But this time

$$\mathbb{P}^{100} \approx \mathbb{P}^{102} \approx \cdots \approx \begin{bmatrix} 0.2 & 0 & 0.2 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0 & 0.4 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0 & 0.4 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0 & 0.4 & 0 & 0.2 \end{bmatrix}$$

and

$$\mathbb{P}^{101} \approx \mathbb{P}^{103} \approx \cdots \approx \begin{bmatrix} 0 & 0.4 & 0 & 0.4 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0 & 0.4 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0 & 0.4 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0 & 0.6 & 0 \end{bmatrix}$$

(there appear to be two alternating limits). In the next section, we explain why. $\diamond$

Example 2.9. Recalling Example 2.4 we can easily gather that each of the four patterns (HH, HT, TH, and HH) must have the same frequency of occurrence, and the stationary probabilities should thus all equal $\frac{1}{4}$ each. This can be verified by

$$\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

And, sure enough,

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}^n = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \text{ for } n = 2, 3, \ldots ,$$

as we already discovered through MAPLE. ◇

Example 2.10. Computing individual powers of $\mathbb{P}$ from Example 2.2, we can establish the limit (reached, to a good accuracy, only at $\mathbb{P}^{30}$) is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & 0 & 0 & \frac{3}{4} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now, even though the $\mathbb{P}^\infty$ limit exists, it has a totally different structure than in the regular case. ◇

So there are several questions we would like to resolve:

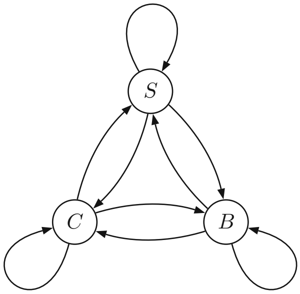
1. How do we know (without computing $\mathbb{P}^\infty$) that an FMC is regular?
2. When does a TPM have a fixed vector but not a stationary distribution, and what is the pattern of large powers of $\mathbb{P}$ in such a case?
3. What else can happen to $\mathbb{P}^\infty$ (in the nonregular cases), and how do we find this without computing high powers of $\mathbb{P}$?

To sort out all these questions and discover the full story of the long-run behavior of an FMC, a brand new approach is called for.

2.4 Classification of States

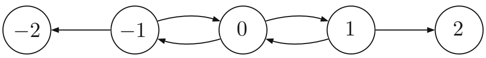
A DIRECTED GRAPH of a TPM is a diagram in which each state is represented by a small circle, and each potential (nonzero) transition by a directed arrow.

Example 2.11. A directed graph based on the TPM of Example 2.1.

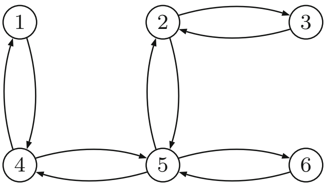


◇

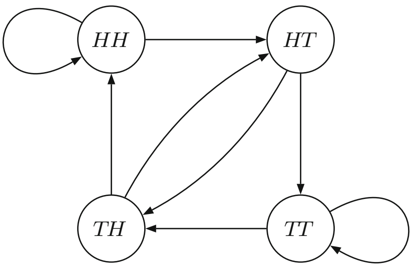
Example 2.12. A directed graph based on the TPM of Example 2.2. ◇



Example 2.13. A directed graph based on the TPM of Example 2.3. ◇



Example 2.14. A directed graph based on the TPM of Example 2.4.



◇

From such a graph one can gather much useful information about the corresponding FMC.

A natural question to ask about any two states, say a and b, is this: Is it possible to get from a to b in some (including 0) number of steps, and then, similarly, from b to a? If the answer is YES (to *both*), we say a and b COMMUNICATE (and denote this by $a \leftrightarrow b$).

Mathematically, a RELATION assigns each (ordered) pair of elements (states, in our case) a YES or NO value. A relation (denoted by $a \to b$ in general) can be symmetric ($a \to b \Rightarrow b \to a$), antisymmetric ($a \to b \Rightarrow \neg(b \to a)$), reflexive ($a \to a$ for each a), or transitive ($a \to b$ and $b \to c \Rightarrow a \to c$).

Is our "communicate" relation symmetric? (YES). Antisymmetric? (NO). Reflexive? (YES, that is why we said "including 0"). Transitive? (YES). A relation that is symmetric, reflexive, and transitive is called an EQUIVALENCE RELATION (a relation that is antisymmetric, reflexive, and transitive is called a PARTIAL ORDER).

An equivalence relation implies we can subdivide the original set (of states, in our case) into so-called EQUIVALENCE CLASSES (each state will be a member of exactly one such class; the classes are thus mutually exclusive, and their union covers the whole set – no gaps, no overlaps). To find these classes, we start with an arbitrary state (say a) and collect all states that communicate with a (together with a, these constitute Class 1); then we take, arbitrarily, any state outside Class 1 (say State b) and find all states that communicate with b (this will be Class 2), and so on till the (finite) set is exhausted.

Example 2.15. Our first, third, and fourth examples each consist of a single class of states (all states communicate with one another). In the second example, States -1, 0, and 1 communicate with one another (one class), but there is no way to reach any other state from State 2 (a class by itself) and also form State -2 (the last class). ◇

In a more complicated situation, it helps to look for closed loops (all states along a closed loop communicate with one another; if, for example, two closed loops have a common element, then they must both belong to the same class).

Once we partition our states into classes, what is the relationship among the classes themselves? It may still be possible to move from one class to another (but not back), so some classes will be connected by one-directional arrows (defining a relationship between *classes* – this relationship is, by definition, antisymmetric, reflexive, and transitive; the reflexive property means the class is connected to itself). Note this time there can be no closed loops – they would create a single class. Also note two classes being connected (say $A \to B$) implies we can get from *any* state of Class A to *any* state of Class B.

It is also possible some classes (or set of classes) are totally disconnected from the rest (no connection in either direction). In practice, this can happen only when we combine two FMCs, which have nothing to do with one another, into a single FMC – using matrix notation, something like this:

$$\begin{bmatrix} \mathbb{P}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{P}_2 \end{bmatrix},$$

where $\mathbb{O}$ represents a zero matrix. So should this happen to us, we can investigate each disconnected group of classes on its own, ignoring the rest (i.e., why this hardly ever happens – it would be mathematically trivial and practically meaningless).

There are two important definitions relating to classes (and their one-way connections): a class that *cannot be left* (found at the bottom of a connection diagram, if all arrows point down) is called RECURRENT; any other class (with an *outgoing arrow*) is called TRANSIENT (these terms are also applied to individual *states* inside these classes). We will soon discover that ultimately (in the long run), an FMC must end up in one of the recurrent classes (the probability of staying transient indefinitely is zero). Note we cannot have transient classes alone (there must always be at least one recurrent class). On the other hand, an FMC can consist of recurrent classes only (normally, only one; see the discussion of the previous paragraph).

We mention in passing that all EIGENVALUES of a TPM must be, in absolute value, less than or equal to 1. One of these eigenvalues *must* be equal to 1, and its MULTIPLICITY yields the number of *recurrent* classes of the corresponding FMC.

Example 2.16. Consider the TPM from Example 2.2.

$$> \mathbb{P}_2 := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}:$$

$> \mathit{evalf}\,(\mathit{Eigenvalues}\,(\mathbb{P}_2, \mathit{output} = \mathit{list}))\,;$

$$[0.0,\ 0.7071,\ -0.7071,\ 1.0000,\ 1.0000]$$

indicating the presence of *two* recurrent classes. ◇

After we have partitioned an FMC into classes, it is convenient to relabel the individual states (and correspondingly rearrange the TPM), so states of the same class are consecutive (the TPM is then organized into so-called BLOCKS), *starting* with recurrent classes; try to visualize what the complete TPM will then look like. Finally, $\mathbb{P}$ can be divided into four basic superblocks by separating the recurrent and transient parts only (never mind the individual classes); thus:

$$\mathbb{P} = \begin{bmatrix} \mathbb{R} & \mathbb{O} \\ \mathbb{U} & \mathbb{T} \end{bmatrix},$$

where $\mathbb{O}$ again denotes the zero matrix (there are no transitions from recurrent to transient states). It is easy to show

$$\mathbb{P}^n = \begin{bmatrix} \mathbb{R}^n & \mathbb{O} \\ ? & \mathbb{T}^n \end{bmatrix}$$

(with the lower left superblock being somehow more complicated). This already greatly simplifies our task of figuring out what happens to large powers of $\mathbb{P}$.

Proposition 2.2.

$$\mathbb{T}^\infty \to \mathbb{O},$$

meaning transient states, in the long run, disappear – the FMC must eventually enter one of its recurrent classes and stay there for good since there is no way out.

Proof. Let $P_a^{(k)}$ be the probability that, starting from a transient state a, k transitions later we will have already reached a recurrent class. These probabilities are nondecreasing in k (once recurrent, always recurrent). The fact it is *possible* to reach a recurrent class from any a (transient) effectively means this: for each a there is a number of transitions, say k_a, such that $P_a^{(k)}$ is already positive, say p_a. If we now take the largest of these k_a (say K) and the smallest of the p_a (say p), then we conclude $P_a^{(K)} \geq p$ for *each* a (transient), or, equivalently, $Q_a^{(K)} < 1-p$ (where Q_a^k is the probability that a has not yet left left the transient states after K transitions, that is, $\sum_{b\in T}(\mathbb{P}^k)_{ab}$, where T is the set of all transient states).

Now,

$$\begin{aligned}
Q_a^{(2K)} &= \sum_{b\in T}(\mathbb{P}^{2k})_{ab} \\
&= \sum_{b\in T}\sum_{c}(\mathbb{P}^k)_{ac}(\mathbb{P}^k)_{cb} \\
&\quad \text{(the } c \text{ summation is over } \textit{all} \text{ states)} \\
&\leq \sum_{b\in T}\sum_{c\in T}(\mathbb{P}^k)_{ac}(\mathbb{P}^k)_{cb} \\
&\leq (1-p)\sum_{c\in T}(\mathbb{P}^k)_{ac} \\
&\leq (1-p)^2.
\end{aligned}$$

Similarly, one can show

$$\begin{array}{rcl} Q_a^{(3K)} & \leq & (1-p)^3, \\ Q_a^{(4K)} & \leq & (1-p)^4, \\ \vdots & & \vdots \end{array}$$

implying $Q_a^{(\infty)} \leq \lim_{n\to\infty}(1-p)^n = 0$. This shows the probability that a transient state a stays transient indefinitely is zero. Thus, every transient state is eventually captured in one of the recurrent classes, with probability of 1. □

Next we tackle the upper left corner of $\mathbb{P}^n$. First of all, $\mathbb{R}$ itself breaks down into individual classes, thus:

$$\begin{bmatrix} \mathbb{R}_1 & \mathbb{O} & \cdots & \mathbb{O} \\ \mathbb{O} & \mathbb{R}_2 & \cdots & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \cdots & \mathbb{R}_k \end{bmatrix}$$

since recurrent classes do not communicate (not even one way). Clearly, then,

$$\mathbb{R}^n = \begin{bmatrix} \mathbb{R}_1^n & \mathbb{O} & \cdots & \mathbb{O} \\ \mathbb{O} & \mathbb{R}_2^n & \cdots & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \cdots & \mathbb{R}_k^n \end{bmatrix},$$

and to find out what happens to this matrix for large n, we need to understand what happens to each of the $\mathbb{R}_i^n$ individually. We can thus restrict our attention to a *single* recurrent class.

To be able to fully understand the behavior of any such $\mathbb{R}_i^n$ (for large n), we first need to have a closer look *inside* the recurrent class, discovering a finer structure: a division into periodic *subclasses*.

2.5 Periodicity of a Class

Let us consider a single recurrent class (an FMC in its own right). If k_1, k_2, k_3, . . . is a *complete* (and therefore infinite) list of the *number of transitions* in which one can return to the initial state (say a) – note this information can be gathered from the corresponding directed graph – then the *greatest common divisor* (say λ) of this set of integers is called the PERIOD of State a.

This can be restated as follows. If the length of every possible closed loop passing (at least once) through State a is divisible by λ, and if λ is the greatest of all integers for which this is true, then λ is the corresponding period. Note a closed loop is allowed any amount of duplication (both in terms of states and transitions) – we can go through the same loop, repeatedly, as many times as we like.

The last definition gives the impression that each state may have its own period. This is not the case.

Proposition 2.3. *The value of λ is the same regardless of which state is chosen for a. The period is thus a property of the whole class.*

Proof. Suppose State a has a period λ_a and State b has a (potentially different) period, say λ_b. Every closed loop passing through b either already passes through a or else can be easily extended (by a $b \to a \to b$ loop) to do so. Either way, the length of the loop must be divisible by λ_a (the extended loop is divisible by λ_a, and the extension itself is also divisible by λ_a; therefore, the difference between the two must be divisible by λ_a). This proves $\lambda_b \geq \lambda_a$. We can now reverse the argument and prove $\lambda_a \geq \lambda_b$, implying $\lambda_a = \lambda_b$. □

In practice, we just need to find the greatest common divisor of all closed loops found in the corresponding directed graph. Whenever there is a loop of length one (a state returning back to itself), the period must be equal to 1 (the class is then called *aperiodic* or *regular*). The same is true whenever we find one closed loop of length 2 and another of length 3 (or any other prime numbers). One should also keep in mind the period cannot be higher than the total number of states (thus, the number of possibilities is quite limited).

A trivial example of a class with a period equal to λ would be a simple CYCLE of λ states, where State 1 goes (in one transition) only to State 2, which in turn must go to State 3, etc., until State λ transits back to State 1 (visualize the directed graph). However, most periodic classes are more complicated than this!

The implication of a nontrivial (> 1) period λ is that we can further partition the set of states into λ SUBCLASSES, which are found as follows.

1. Select an arbitrary State a. It will be a member of Subclass 0 (we will label the subclasses 0, 1, 2, ..., $\lambda - 1$).
2. Find a path that starts at a and visits all states (some more than once if necessary).
3. Assign each state along this path to Subclass $k \bmod \lambda$, where k is the number of transitions to reach it.

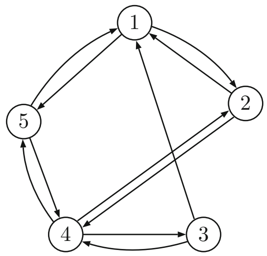
It is quite simple to realize this definition of subclasses is consistent (each state is assigned to the same subclass no matter how many times we go through it) and, up to a cyclical rearrangement, unique (we get the same answer regardless of where we start and which path we choose).

Note subclasses do not need to be of the same size!

Example 2.17. Find the subclasses of the following FMC (defined by the corresponding TMP):

$$\begin{bmatrix} 0 & 0.7 & 0 & 0 & 0.3 \\ 0.7 & 0 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0.2 & 0.2 & 0 & 0.6 \\ 0.4 & 0 & 0 & 0.6 & 0 \end{bmatrix}.$$

Solution. From the corresponding directed graph



it follows this *is* a *single* class (automatically recurrent). Since State 1 can go to State 5 and then back to State 1, there is a closed loop of length 2 (the period cannot be any higher, that is, it must be either 2 or 1). Since all closed loops we find in the directed graph are of length 2 or 4 (and higher multiples of 2), the period is equal to 2. From the path $1 \to 5 \to 4 \to 3 \to 4 \to 2$ we can conclude the two subclasses are $\{1, 4\}$ and $\{2, 3, 5\}$. Rearranging our TPM accordingly we get

$$\begin{array}{c|cc:ccc|} & 1 & 4 & 2 & 3 & 5 \\ \hline 1 & 0 & 0 & 0.7 & 0 & 0.3 \\ 4 & 0 & 0 & 0.2 & 0.2 & 0.6 \\ \hdashline 2 & 0.7 & 0.3 & 0 & 0 & 0 \\ 3 & 0.5 & 0.5 & 0 & 0 & 0 \\ 5 & 0.4 & 0.6 & 0 & 0 & 0 \\ \hline \end{array} \tag{2.1}$$

Note this partitions the matrix into corresponding subblocks. $\diamond$

One can show the last observation is true in general, that is, one can go (in one transition) only from Subclass 0 to Subclass 1, from Subclass 1 to Subclass 2, etc., until finally, one goes from Subclass $\lambda - 1$ back to Subclass 0. The rearranged TPM will then always look like this (we use a hypothetical example with four subclasses):

$$\mathbb{R} = \begin{bmatrix} \mathbb{O} & \mathbb{C}_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{C}_2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{C}_3 \\ \mathbb{C}_4 & \mathbb{O} & \mathbb{O} & \mathbb{O} \end{bmatrix},$$

where the size of each subblock corresponds to the number of states in the respective (row and column) subclasses.

Note $\mathbb{R}^\lambda$ will be (block) diagonal; for our last example, this means

$$\mathbb{R}^4 = \begin{bmatrix} \mathbb{C}_1\mathbb{C}_2\mathbb{C}_3\mathbb{C}_4 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{C}_2\mathbb{C}_3\mathbb{C}_4\mathbb{C}_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{C}_3\mathbb{C}_4\mathbb{C}_1\mathbb{C}_2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{C}_4\mathbb{C}_1\mathbb{C}_2\mathbb{C}_3 \end{bmatrix}$$

(from this one should be able to discern the general pattern). Note by taking four transitions at a time (seen as a single supertransition), the process turns into an FMC with four recurrent *classes* (no longer subclasses), which we know how to deal with. This implies $\lim_{n\to\infty} \mathbb{R}^{4n}$ will have the following form:

$$\begin{bmatrix} \mathbb{S}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{S}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{S}_3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{S}_4 \end{bmatrix},$$

where $\mathbb{S}_1$ is a matrix with identical rows, say $\mathbf{s}_1$ (the stationary probability vector of $\mathbb{C}_1\mathbb{C}_2\mathbb{C}_3\mathbb{C}_4$ – one can show each of the four new classes must be *aperiodic*); similarly, $\mathbb{S}_2$ consists of stationary probabilities $\mathbf{s}_2$ of $\mathbb{C}_2\mathbb{C}_3\mathbb{C}_4\mathbb{C}_1$ etc.

What happens when the process undergoes one extra transition? This is quite simple: it goes from Subclass 0 to Subclass 1, (or $1 \to 2$ or $2 \to 3$ or $3 \to 0$), but the probabilities within each subclass must remain stationary. This is clear from the following limit:

$$\lim_{n\to\infty} \mathbb{R}^{4n+1} = \left(\lim_{n\to\infty} \mathbb{R}^{4n}\right)\mathbb{R}$$
$$= \begin{bmatrix} \mathbb{O} & \mathbb{S}_1\mathbb{C}_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{S}_2\mathbb{C}_2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{S}_3\mathbb{C}_3 \\ \mathbb{S}_4\mathbb{C}_4 & \mathbb{O} & \mathbb{O} & \mathbb{O} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbb{O} & \mathbb{S}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{S}_3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{S}_4 \\ \mathbb{S}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \end{bmatrix}$$

since $\mathbf{s}_1^{\mathrm{T}}\mathbb{C}_1$ is a solution to $\mathbf{s}^{\mathrm{T}}\,\mathbb{C}_2\mathbb{C}_3\mathbb{C}_4\mathbb{C}_1 = \mathbf{s}^{\mathrm{T}}$ (note $\mathbf{s}_1$ satisfies $\mathbf{s}_1^{\mathrm{T}}\,\mathbb{C}_1\,\mathbb{C}_2\mathbb{C}_3\mathbb{C}_4 = \mathbf{s}_1^{\mathrm{T}}$) and must therefore be equal to $\mathbf{s}_2$. Similarly, $\mathbf{s}_2^{\mathrm{T}}\mathbb{C}_2 = \mathbf{s}_3^{\mathrm{T}}$, $\mathbf{s}_3^{\mathrm{T}}\mathbb{C}_3 = \mathbf{s}_4^{\mathrm{T}}$, and $\mathbf{s}_4^{\mathrm{T}}\mathbb{C}_4 = \mathbf{s}_1^{\mathrm{T}}$ (back to $\mathbf{s}_1$).

This implies once we obtain *one* of the $\mathbf{s}$ vectors, we can get the rest by a simple multiplication. We would like to start from the *shortest* one since it is the easiest to find.

The fixed vector of $\mathbb{R}$ (a solution to $\mathbf{f}^{\mathrm{T}}\mathbb{R} = \mathbf{f}^{\mathrm{T}}$) is then found by

$$\mathbf{f}^{\mathrm{T}} = \frac{\left\langle \mathbf{s}_1^{\mathrm{T}}, \mathbf{s}_2^{\mathrm{T}}, \mathbf{s}_3^{\mathrm{T}}, \mathbf{s}_4^{\mathrm{T}} \right\rangle}{4},$$

and similarly for any other number of subclasses. The interpretation is clear: this yields the long-run proportion of visits to individual states of the class.

Example 2.18. Returning to our two subclasses of Example 2.17, we first compute

$$\mathbb{C}_1\mathbb{C}_2 = \begin{bmatrix} 0.7 & 0 & 0.3 \\ 0.2 & 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} 0.61 & 0.39 \\ 0.48 & 0.52 \end{bmatrix},$$

then find the corresponding $\mathbf{s}_1 = \begin{bmatrix} \frac{16}{29} \\ \frac{13}{29} \end{bmatrix}$ (a relatively simple exercise), and finally

$$\mathbf{s}_2^{\mathrm{T}} = \begin{bmatrix} \frac{16}{29} & \frac{13}{29} \end{bmatrix} \begin{bmatrix} \frac{7}{10} & 0 & \frac{3}{10} \\ \frac{2}{10} & \frac{2}{10} & \frac{6}{10} \end{bmatrix} = \begin{bmatrix} \frac{69}{145} & \frac{13}{145} & \frac{63}{145} \end{bmatrix}.$$

To verify the two answers, we can now build the (unique) fixed probability vector of the original PTM, namely, $\mathbf{f}^{\mathrm{T}} = \left[\begin{array}{ccccc} \frac{16}{58} & \frac{13}{58} & \frac{69}{290} & \frac{13}{290} & \frac{63}{290} \end{array}\right]$, and check $\mathbf{f}^{\mathrm{T}} = \mathbb{R}\,\mathbf{f}^{\mathrm{T}}$ (which is indeed the case). $\diamond$

Similarly to the previous $\lim_{n\to\infty} \mathbb{R}^{4n+1}$, we can derive

$$\lim_{n\to\infty} \mathbb{R}^{4n+2} = \begin{bmatrix} \mathbb{O} & \mathbb{O} & \mathbb{S}_3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{S}_4 \\ \mathbb{S}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{S}_2 & \mathbb{O} & \mathbb{O} \end{bmatrix} \quad \text{and} \quad \lim_{n\to\infty} \mathbb{R}^{4n+3} = \begin{bmatrix} \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{S}_4 \\ \mathbb{S}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{S}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{S}_3 & \mathbb{O} \end{bmatrix},$$

where $\mathbb{S}_i$ implies a matrix whose rows are all equal to $\mathbf{s}_i$ (but its row dimension may change from one power of $\mathbb{R}$ to the next).

So now we know how to raise $\mathbb{R}$ to any *large* power.

Example 2.19. Find

$$\begin{bmatrix} \cdot & \cdot & \cdot & 0.2 & 0.8 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.5 & 0.5 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.7 & 0.3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.4 & 0.6 \\ 0.3 & 0.2 & 0.5 & \cdot & \cdot & \cdot & \cdot \\ 0.2 & 0 & 0.8 & \cdot & \cdot & \cdot & \cdot \end{bmatrix}^{10000}$$

(dots represent zeros).

Solution. One can confirm the period of the corresponding class is 3, and the subclasses are $\{1,2,3\}$, $\{4,5\}$ and $\{6,7\}$. To get the stationary probabilities of the second subclass, we first need

$$\mathbb{C}_2\mathbb{C}_3\mathbb{C}_1 = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.2 & 0 & 0.8 \end{bmatrix} \begin{bmatrix} 0.2 & 0.8 \\ 0.5 & 0.5 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.714 & 0.286 \\ 0.768 & 0.232 \end{bmatrix}$$

whose stationary vector is $\mathbf{s}_2^{\mathrm{T}} = \left[\begin{array}{cc} 0.72865 & 0.27135 \end{array}\right]$ (verify). Then

$$\mathbf{s}_3^{\mathrm{T}} = \mathbf{s}_2^{\mathrm{T}} \mathbb{C}_2 = \left[\begin{array}{cc} 0.61860 & 0.38140 \end{array}\right]$$

and

$$\mathbf{s}_1^{\mathrm{T}} = \mathbf{s}_3^{\mathrm{T}} \mathbb{C}_3 = \left[\begin{array}{ccc} 0.26186 & 0.12372 & 0.61442 \end{array}\right].$$

Since 10000 mod $3 \equiv 1$, the answer is

$$\left[\begin{array}{ccccccc}
\cdot & \cdot & \cdot & 0.72865 & 0.27135 & \cdot & \cdot \\
\cdot & \cdot & \cdot & 0.72865 & 0.27135 & \cdot & \cdot \\
\cdot & \cdot & \cdot & 0.72865 & 0.27135 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 0.61860 & 0.38140 \\
\cdot & \cdot & \cdot & \cdot & \cdot & 0.61860 & 0.38140 \\
0.26186 & 0.12372 & 0.61442 & \cdot & \cdot & \cdot & \cdot \\
0.26186 & 0.12372 & 0.61442 & \cdot & \cdot & \cdot & \cdot
\end{array}\right]$$

Similarly, the 10001th power of the original matrix would be equal to

$$\left[\begin{array}{ccccccc}
\cdot & \cdot & \cdot & \cdot & \cdot & 0.61860 & 0.38140 \\
\cdot & \cdot & \cdot & \cdot & \cdot & 0.61860 & 0.38140 \\
\cdot & \cdot & \cdot & \cdot & \cdot & 0.61860 & 0.38140 \\
0.26186 & 0.12372 & 0.61442 & \cdot & \cdot & \cdot & \cdot \\
0.26186 & 0.12372 & 0.61442 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & 0.72865 & 0.27135 & \cdot & \cdot \\
\cdot & \cdot & \cdot & 0.72865 & 0.27135 & \cdot & \cdot
\end{array}\right]$$

At this point it should be clear what the 10002th power looks like. ◊

Remark 2.1. A recurrent class with a period of λ contributes all λ roots of 1 (each exactly once) to the eigenvalues of the corresponding TPM (the remaining eigenvalues must be, in absolute value, less than 1). Thus, eigenvalues nicely reveal the number and periodicity of all recurrent classes.

$$> \mathbb{T} := \begin{bmatrix} 0 & 0 & 0 & .2 & .8 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0.4 & 0.6 \\ 0.3 & 0.2 & 0.5 & 0 & 0 & 0 & 0 \\ 0.2 & 0 & 0.8 & 0 & 0 & 0 & 0 \end{bmatrix}:$$

$> \mathbb{T} := \mathit{convert}\,(\mathbb{T}, \mathit{rational}) : \{$we do this to get exact eigenvalues$\}$
$> \lambda := \mathit{Eigenvalues}(\mathbb{T}, \mathit{output} = \mathit{list});$

$$\lambda := \Big[\ 0,\ \ 1,\ \ -1/2 - 1/2\,\mathbf{I}\sqrt{3},\ \ -1/2 + 1/2\,\mathbf{I}\sqrt{3},$$
$$-3/10\,\sqrt[3]{2},\ \ \tfrac{3}{20}\,\sqrt[3]{2} - \tfrac{3}{20}\,\mathbf{I}\sqrt{3}\,\sqrt[3]{2},\ \ \tfrac{3}{20}\,\sqrt[3]{2} + \tfrac{3}{20}\,\mathbf{I}\sqrt{3}\,\sqrt[3]{2}\ \Big]$$

$> \mathit{seq}\,(\mathit{evalf}\,(\mathit{abs}\,(x))\,, x \in \lambda)\,;$

$$0.000, 1.0000, 1.0000, 1.0000, 0.3780, 0.3780, 0.3780$$

$> \mathit{simplify}\,\big(\mathit{seq}\,(\lambda_i^3, i \in [2, 3, 4])\big)\,;$

$$1,\ 1,\ 1$$

This implies there is only a single recurrent class whose period is 3.

2.6 Regular Markov Chains

An FMC with a single, aperiodic class is called REGULAR. We already know that for these, $\mathbb{P}^\infty$ exists and has a stationary vector in each row. We can prove this in four steps (three propositions and a conclusion).

Proposition 2.4. *If S is a nonempty set of positive integers closed under addition and having 1 as its greatest common divisor, then starting from a certain integer, say N, all integers ($\geq N$) must be in S.*

Proof. We know (from number theory) there must be a finite set of integers from S (we call them $n_1, n_2, \ldots, n_k$) whose linear combination (with integer coefficients $a_1, a_2, \ldots, a_k$) must be equal to the corresponding greatest common divisor; thus,

$$a_1 n_1 + a_2 n_2 + \cdots + a_k n_k = 1.$$

Collecting the positive and negative terms on the left-hand side of this equation implies

$$N_1 - N_2 = 1,$$

where both N_1 and N_2 belong to S (due to its closure under addition). Let q be any integer $\geq N_2(N_2 - 1)$. Since q can be written as $a\, N_2 + b$, where $0 \leq b < N_2$ and $a \geq N_2 - 1$, and since

$$a\, N_2 + b = (a - b)N_2 + b(1 + N_2) = (a - b)N_2 + bN_1,$$

each such q must be a member of S (again, due to the closure property). □

Proposition 2.5. *The set of integers n for which $(\mathbb{P}^n)_{ii} > 0$, where $\mathbb{P}$ is regular, is closed under addition for each i. This implies, for sufficiently large n, all elements of $\mathbb{P}^N$ are strictly positive (meaning it is possible to move from State i back to State i in exactly n transitions).*

Proof. Since

$$(\mathbb{P}^{n+m})_{ij} = \sum_k (\mathbb{P}^n)_{ik} (\mathbb{P}^m)_{kj} \geq (\mathbb{P}^n)_{ii} (\mathbb{P}^m)_{ij} > 0,$$

where m is smaller than the total number of states (since State j can be reached from State i by visiting any of the other states no more than once). We can thus see, for sufficiently large n, *all* $\mathbb{P}^n_{ij}$ are strictly positive (i.e., have no zero entries). □

When a stochastic matrix $\mathbb{P}$ multiplies a column vector $\mathbf{r}$, each component of the result is a (different) weighted average of the elements of $\mathbf{r}$. The smallest value of $\mathbb{P}\mathbf{r}$ thus cannot be any smaller than that of $\mathbf{r}$ (similarly, the largest value cannot go up). We now take $\mathbb{Q} = \mathbb{P}^N$, where $\mathbb{P}$ is regular and N is large enough to eliminate zero entries from $\mathbb{Q}$. Clearly, there must be a positive ε such that all entries of $\mathbb{Q}$ are $\geq \varepsilon$. This implies the difference between the largest and smallest component of $\mathbb{Q}\mathbf{r}$ (let us call them M_1 and m_1, respectively) must be *smaller* than the difference between the largest and smallest components of $\mathbf{r}$ (let us call these M_0 and m_0) by a factor of at least $(1 - 2\varepsilon)$.

Proposition 2.6.

$$\max(\mathbb{Q}\mathbf{r}) - \min(\mathbb{Q}\mathbf{r}) \leq (1 - 2\varepsilon) \max(\mathbf{r}) - \min(\mathbf{r})\,.$$

Proof. Clearly, $m_1 \geq \varepsilon M_0 + (1-\varepsilon) m_0$ if we try to make the right-hand side as small as possible (multiplying M_0 by the smallest possible value and making all the other entries of $\mathbf{r}$ as small as possible). Similarly, $M_1 \leq \varepsilon\, m_0 + (1-\varepsilon) M_0$ (now we are multiplying m_0 by the smallest possible factor, leaving the rest for M_0). Subtracting the two inequalities yields

$$M_1 - m_1 \leq (1 - 2\varepsilon)(M_0 - m_0).$$

□

Proposition 2.7. *All rows of $\mathbb{P}^{\infty}$ are identical and equal to the stationary vector.*

Proof. Take $\mathbf{r}_1$ to be a *column* vector defined by $[1, 0, 0, \ldots, 0]^{\mathrm{T}}$ and multiply it repeatedly, say n times, by $\mathbb{Q}$, getting $\mathbb{Q}^n \mathbf{r}_1$ (the first column of $\mathbb{Q}^n$). The difference between the largest and smallest elements of the resulting vector is no bigger than $(1-2\varepsilon)^n$ – the previous proposition, applied n times – and converges to 0 when $n \to \infty$. Similarly (using the original $\mathbb{P}$) the difference between the largest and smallest elements of $\mathbb{P}^n \mathbf{r}_1$ must converge to 0 since it is a nonincreasing sequence that contains a subsequence (that, coming from $\mathbb{Q}^n \mathbf{r}_1$) converging to 0. We have thus proved the first column of $\mathbb{P}^n$ converges to a vector with constant elements. By taking $\mathbf{r}_2 = [0, 1, 0, \ldots, 0]^{\mathrm{T}}$ we can prove the same thing for each column of $\mathbb{P}^n$. □

2.A Inverting Matrices

Inverting (Small) Matrices

To invert

$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}$$

do:

1. Find the matrix of codeterminants (for each element, remove the corresponding row and column and find the determinant of what is left):

$$\begin{bmatrix} \frac{3}{4} & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix}.$$

2. Change the sign of each element of the previous matrix according to the following checkerboard scheme:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix},$$

resulting in

$$\begin{bmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

(all elements of $\mathbb{F}$ must be nonnegative).

3. Transpose the result:

$$\begin{bmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

(nothing changes in this particular case, since the matrix was symmetric).

4. Divide each element by the determinant of the original matrix (found easily as the dot product of the first row of the original matrix and the first column of the previous matrix):

$$\begin{bmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{bmatrix}$$

Remark 2.2. The number of operations required by this algorithm is proportional to $n!$ (n being the size of the matrix). This makes the algorithm practical for small matrices only (in our case, no more than 4×4) and *impossible* (even when using supercomputers) for matrices beyond even a moderate size (say 30×30).

Inverting Matrices (of Any Size)

The general procedure (easy to code) requires the following steps:

1. Append the unit matrix to the matrix to be inverted (creating a new matrix with twice as many columns as the old one), for example,

$$\left[\begin{array}{rrrr|rrrr} 2 & -3 & 5 & 1 & 1 & 0 & 0 & 0 \\ -1 & 4 & 0 & 5 & 0 & 1 & 0 & 0 \\ 2 & -6 & 2 & 7 & 0 & 0 & 1 & 0 \\ 1 & -3 & -4 & 3 & 0 & 0 & 0 & 1 \end{array}\right].$$

2. Use any number of the following ELEMENTARY OPERATIONS:

   - Multiply each element of a single row by the same nonzero constant;
   - Add/subtract a multiple of a row to/from any other row;
   - Interchange any two rows,

   to convert the original matrix to the unit matrix. Do this column by column: start with the main diagonal element (making it equal to 1), then make the remaining elements of the same column equal to 0.
3. The right side of the result (where the *original* unit matrix used to be) is the corresponding inverse.

If you fail to complete these steps (which can happen only when getting a *zero* on the *main diagonal and* every other element of the same column *below* the main diagonal), the original matrix is SINGULAR.

The number of operations required by this procedure is proportional to n^3 (n being the size of the matrix). In practical terms, this means even a standard laptop can invert matrices of huge size (say, $1{,}000^2$ elements) in a fraction of a second.

EXERCISES

Exercise 2.1. Consider a simple Markov chain with the following TPM:

$$\mathbb{P} = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \end{bmatrix}.$$

Assuming X_0 is generated from the distribution

X_0	1	2	3
Pr	0.6	0.0	0.4

find:

(a) $\Pr(X_2 = 3 \mid X_4 = 1)$;
(b) The stationary vector;
(c) The expected number of transitions it will take to enter, for the first time, State 2 and the corresponding standard deviation.

Exercise 2.2. Find (in terms of exact fractions) the fixed vector of the following TPM:

$$\mathbb{P} = \begin{bmatrix} 0 & 0.3 & 0.4 & 0 & 0.3 \\ 0.4 & 0 & 0 & 0.6 & 0 \\ 0.7 & 0 & 0 & 0.3 & 0 \\ 0 & 0.5 & 0.1 & 0 & 0.4 \\ 0.5 & 0 & 0 & 0.5 & 0 \end{bmatrix}$$

and the limit $\lim\limits_{n\to\infty} \mathbb{P}^{2n}$.

(a) What is the long-run percentage of time spent in State 4?
(b) Is this Markov chain reversible (usually one can get the answer by constructing only a single element of $\overset{\circ}{\mathbb{P}}$)?

Exercise 2.3. Find the exact (in terms of fractions) answer to

$$\lim_{n\to\infty} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.15 & 0.85 & 0 & 0 \\ 0 & 0.55 & 0.45 & 0 & 0 \\ 0.12 & 0.18 & 0.21 & 0.26 & 0.23 \\ 0.19 & 0.16 & 0.14 & 0.27 & 0.24 \end{bmatrix}^n .$$

Exercise 2.4. Do the complete classification of the following TPM ($\times$ indicates a nonzero entry, $\cdot$ denotes zero):

$$\begin{bmatrix} \cdot & \cdot & \times & \times & \cdot & \cdot & \cdot \\ \cdot & \cdot & \times & \times & \cdot & \cdot & \cdot \\ \cdot & \cdot & \times & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \times & \cdot & \cdot & \times & \times \\ \times & \cdot & \times & \cdot & \cdot & \cdot & \cdot \\ \cdot & \times & \times & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \times & \cdot & \times & \cdot & \cdot \end{bmatrix} .$$

Are any of the TPM's classes periodic?

Exercise 2.5. Using the TPM

$$\begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.5 & 0.1 & 0.4 \end{bmatrix}$$

find $\Pr(X_3 = 2 \cap X_1 = 3)$ given that the initial state is drawn from the distribution

X_0	1	2	3
Pr	0.25	0.40	0.35

.

Also, find the probability of visiting State 2 before State 3.

Exercise 2.6. Find the fixed probability vector of the following TPM:

$$\begin{bmatrix} 0 & 0.4 & 0 & 0.2 & 0 & 0.4 \\ 0 & 0 & 0.7 & 0 & 0.3 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.4 & 0 & 0.6 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0.8 & 0 \end{bmatrix}.$$

Also, find (in exact fractions) $\lim_{n\to\infty} \mathbb{P}^{3n+1}$.

Exercise 2.7. Find the fixed probability vector of

$$\begin{bmatrix} 0 & 0.4 & 0 & 0.6 \\ 0.2 & 0 & 0.8 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.7 & 0 & 0.3 & 0 \end{bmatrix}.$$

Starting in State 1, what is the probability of being in State 4 after 1,001 transitions?

Exercise 2.8. Calculate exactly using fractions:

$$\lim_{n\to\infty}\begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0.2 & 0.8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0.4 & 0.6 & 0 & 0 \\ 0.1 & 0 & 0 & 0.2 & 0.3 & 0.4 \end{bmatrix}^{2n+1}.$$

Exercise 2.9. Do a complete classification of

$$\begin{bmatrix} \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \times \\ \cdot & \cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot & \times \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \times & \cdot \\ \cdot & \cdot & \times & \cdot & \times & \cdot & \cdot & \cdot \\ \cdot & \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \times & \cdot & \cdot & \times & \cdot \\ \times & \cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot \end{bmatrix}.$$

Exercise 2.10. Do a complete classification of

$$\begin{bmatrix} \times & \cdot & \cdot & \cdot & \times & \cdot & \times & \cdot & \cdot & \cdot \\ \cdot & \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \times \\ \cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot & \times & \cdot & \cdot \\ \cdot & \cdot & \cdot & \times & \cdot & \cdot & \cdot & \cdot & \times & \cdot \\ \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \times & \cdot & \times & \cdot & \cdot \\ \times & \cdot & \cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot & \cdot \\ \cdot & \cdot & \times & \cdot & \cdot & \times & \cdot & \times & \cdot & \cdot \\ \cdot & \times & \cdot & \times & \cdot & \cdot & \cdot & \times & \cdot & \cdot \\ \cdot & \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}.$$

Exercise 2.11. Do a complete classification of

$$\begin{bmatrix}
\times & \cdot & \times & \cdot & \cdot & \times & \times & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \times & \cdot & \cdot & \cdot & \times & \cdot & \cdot \\
\times & \times & \cdot & \times & \times & \times & \cdot & \times & \times \\
\cdot & \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \times \\
\times & \cdot & \times & \cdot & \cdot & \times & \times & \cdot & \cdot \\
\cdot & \cdot & \times & \cdot & \cdot & \cdot & \times & \cdot & \cdot \\
\times & \times & \cdot & \times & \times & \times & \cdot & \times & \times \\
\cdot & \cdot & \cdot & \cdot & \times & \cdot & \cdot & \cdot & \cdot
\end{bmatrix}.$$

Exercise 2.12. Calculate exactly using fractions:

$$\lim_{n\to\infty} \begin{bmatrix}
0 & 0.5 & 0 & 0.2 & 0 & 0 & 0.3 & 0 \\
0 & 0 & 0.1 & 0 & 0.5 & 0.4 & 0 & 0 \\
0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8 \\
0 & 0 & 0.3 & 0 & 0.4 & 0.3 & 0 & 0 \\
0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 \\
0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\
0 & 0 & 0.2 & 0 & 0.6 & 0.2 & 0 & 0 \\
0 & 0.7 & 0 & 0.1 & 0 & 0 & 0.2 & 0
\end{bmatrix}^{3n+2}.$$

Exercise 2.13. Calculate exactly using fractions:

$$\lim_{n\to\infty} \begin{bmatrix}
0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 \\
0 & 0 & 0.2 & 0.8 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.4 & 0.6 & 0 \\
0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \\
0.6 & 0.4 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.4
\end{bmatrix}^{3n+1}.$$

Exercise 2.14. For

$$\mathbb{P} = \begin{bmatrix} 0.2 & 0 & 0.3 & 0.1 & 0 & 0.3 & 0.1 \\ 0 & 0 & 0 & 0.7 & 0 & 0 & 0.3 \\ 0.1 & 0.2 & 0.2 & 0 & 0 & 0.5 & 0 \\ 0 & 0.7 & 0 & 0 & 0 & 0.3 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.2 & 0.2 & 0 & 0.6 \\ 0 & 0.4 & 0 & 0 & 0 & 0.6 & 0 \end{bmatrix}$$

find the exact (i.e., use fractions) value of $\lim_{n\to\infty} \mathbb{P}^{2n}$ and $\lim_{n\to\infty} \mathbb{P}^{2n+1}$.

http://www.springer.com/978-1-4614-4056-7

Informal Introduction to Stochastic Processes with Maple
Vrbik, J.; Vrbik, P.
2013, X, 287 p. 54 illus. With online files/update., Softcover
ISBN: 978-1-4614-4056-7