1.4.1. *larger transient classes.* Last time I explained (Theorem 1.12) that, if

$$\mathbb{P}(\text{success in one trial}) = p > 0$$

then

$$\mathbb{P}(\text{success with } \infty \text{ many trials}) = 1.$$

But you can say more:

Corollary 1.13. *Furthermore, you will almost surely succeed an infinite number of times.*

Proof. Suppose that you succeed only finitely many times, say 5 times:

$$n_1, n_2, n_3, n_4, n_5.$$

If n_5 is the last time that you succeed, it means that, after that point in time, you try over and over infinitely many times and fail each time. This has probability zero by the theorem. So,

$$\mathbb{P}(\text{only finitely many successes}) = 0.$$

But, the number of successes is either finite or infinite. So,

$$\mathbb{P}(\text{infinitely many successes}) = 1.$$

□

Apply this to Markov chains:

$$X_0, X_1, X_2, \cdots$$

These locations are random states in the finite set S of all states. This means that there is at least one state that is visited infinitely many times. Let

$$I := \{i \in S \mid X_n = i \text{ for infinitely many } n\}$$

This is the set of those states that the random path goes to infinitely many times.

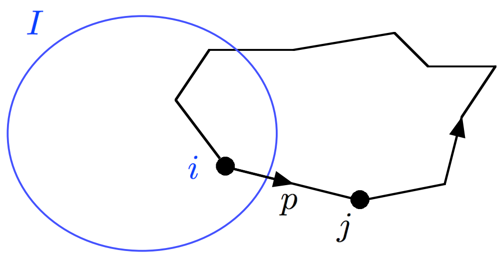
Theorem 1.14. *(A.s.) I is one recurrent class.*

At this point we had a discussion about the meaning of this. The set I is a random set. Since a general finite Markov chain has several recurrent classes, which one you end up in is a matter of chance. The probability distribution of X_n for large n will include a linear combination or "superposition" of several possible futures. So, several recurrent classes have positive probability at the beginning. However, when you actually go into the future, you pick one path and you get stuck in one recurrent class from which you cannot escape. This theorem says that you will wander around and visit every site in that recurrent class infinitely many times.

Proof. In order to prove this theorem I first proved:

Lemma (a) If $i \in I$ and $i \to j$ then $j \in I$.

This means: if it is possible to go from i to j then $j \in I$.



Proof of Lemma (a): It is given that $i \in I$. I.e., we go to i infinitely many times. Each time we go to i we have a probability $p > 0$ of going to j. Theorem 1.12 says that, with probability one, we will eventually go to j. But then (b) we have to eventually go back to i because, we are going to i infinitely many times. So, by Corollary 1.13, with probability one, you cross that bridge infinitely many times. So, $j \in I$. (The picture is a little deceptive. The path from i to j can have more than one step.)

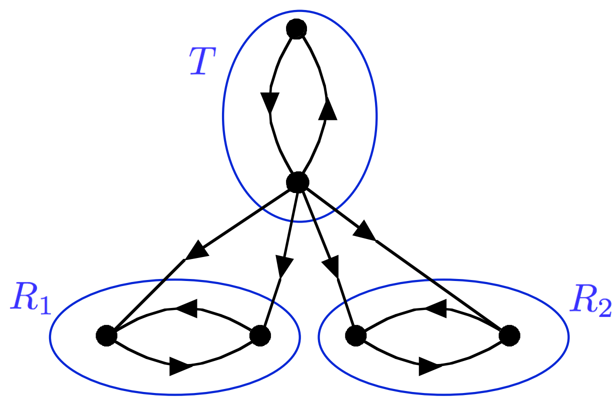
This proof also says: **(b)** $j \to i$ since you need to return to i infinitely many times. Therefore, I is one communication class. We just need to show that this class is recurrent.

But (a) implies that I is recurrent. Otherwise, there would be a j not in I so that $i \to j$ for some $i \in I$ and this would contradict (a). ☐

Corollary 1.15. *The probability is zero that you remain in a transient class indefinitely.*

1.5. **Canonical form of P.** Next, I talked about the canonical form of P which is given on page 20 of our book.

1.5.1. *definition.* Suppose that $R_1, R_2, \cdots, R_r$ are the recurrent classes of a Markov chain and $T_1, T_2, \cdots, T_s$ are the transient classes. I drew a picture similar to the following to illustrate this.



Then the *canonical form* of the transition matrix P is given by the following "block" form of the matrix: (In the book, all transient classes are combined. So, I will do the same here.)

$$P = \begin{array}{c} \\ R_1 \\ R_2 \\ T \end{array} \begin{array}{c} \begin{array}{ccc} R_1 & R_2 & T \end{array} \\ \begin{pmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ S_1 & S_2 & Q \end{pmatrix} \end{array}$$

If you start in the recurrent class R_1 then you can't go anywhere else. So, there is only P_1 in the first row. In the example, it is a 2×2 matrix. Similarly, the second row has only P_2 since, if you start in R_2 you can't get out.

The matrices P_1 and P_2 are *stochastic matrices.* Their rows add up to one since, in the entire matrix P, there are no other numbers in those rows. This also reflects the fact that the recurrent classes R_1 and R_2 are, in themselves, (irreducible) Markov chains.

The transient class T is not a Markov chain. Why not? There are several reasons. If you look at the picture, you see that you can leave the transient class out of the bottom. So, it is not a "closed system." Another reason is that the matrix Q is not stochastic. Its rows do not add up to one. So, Q does not define a Markov chain.

The bottom row in the canonical form describes what happens if you start in any transient class. You either go to another transient state or you go to a recurrent state. The matrix Q is the *transient-to-transient* matrix. The matrix

$$S = (S_1, S_2)$$

is the *transient-to-recurrent* matrix. It has one block S_i for every recurrent state R_i.

Since each recurrent state R_i is an irreducible Markov chain, it has a unique invariant distribution π_i.

Theorem 1.16. *If π_i is the invariant distribution for P_i then the invariant distributions for P are the positive linear combinations of the π_i (with coefficients adding to 1). In other words,*

$$\pi = \sum t_i \pi_i$$

where $t_i \geq 0$ and $\sum t_i = 1$. In the case of two recurrent states, this is:

$$\pi = t\pi_1 + (1-t)\pi_2$$

where $0 \leq t \leq 1$.

Proof. Suppose that π_1, π_2 are invariant distributions for P_1, P_2. Then they are row vectors of the same size as P_1, P_2, respectively, and

$$\pi_1 P_1 = \pi_1, \quad \pi_2 P_2 = \pi_2.$$

When $t = 1/3$ we get the invariant distribution:

$$\pi = \left(\tfrac{1}{3}\pi_1, \tfrac{2}{3}\pi_2, 0\right).$$

You need to multiply by $1/3$ and $2/3$ (or some other numbers ≥ 0 which add to 1) so that the entries of π add up to 1. Block matrix multiplication show that this is an invariant distribution:

$$\pi P = \left(\tfrac{1}{3}\pi_1, \tfrac{2}{3}\pi_2, 0\right) \left(\begin{array}{cc|c} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ \hline S_1 & S_2 & Q \end{array}\right)$$

$$= \left(\tfrac{1}{3}\pi_1 P_1, \tfrac{2}{3}\pi_2 P_2, 0\right) = \left(\tfrac{1}{3}\pi_1, \tfrac{2}{3}\pi_2, 0\right) = \pi$$

This shows that the positive linear combinations of the invariant distributions π_i are invariant distributions for P.

The converse, which I did not prove in class is easy: Suppose that $\pi = (\alpha, \beta, \gamma)$ is an invariant distribution. Then we must have $\gamma = 0$, since otherwise

$$(\alpha, \beta, \gamma)P^n = (\alpha, \beta, \gamma)$$

indicating that we have a positive probability of remaining in a transient state indefinitely, a contradiction to what we just proved. So, $\pi = (\alpha, \beta, 0)$ and

$$(\alpha, \beta, 0) \left(\begin{array}{cc|c} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ \hline S_1 & S_2 & Q \end{array} \right) = (\alpha P_1, \beta P_2, 0) = (\alpha, \beta, 0)$$

which means that $\alpha P_1 = \alpha$ and $\beta P_2 = \beta$. So, α, β are scalar multiples of invariant distributions for P_1, P_2. □

The next two pages are what I handed out in class, although the page numbers have shifted.

1.5.2. *example.* The problem is to find all invariant distributions of the following transition matrix.

$$P = \begin{pmatrix} 1/2 & 0 & 0 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1 & 0 \\ 1/4 & 0 & 0 & 3/4 \end{pmatrix}$$

An invariant distribution π is the solution of:

$$\pi P = \pi.$$

This equation can be rewritten as:

$$\pi(P - I) = 0$$

where $I = I_4$ is the identity matrix. In other words π is a left null vector of

$$P - I = \begin{pmatrix} -1/2 & 0 & 0 & 1/2 \\ 1/4 & -3/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 \\ 1/4 & 0 & 0 & -1/4 \end{pmatrix}$$

Corollary 1.17. *The dimension of the null space of $P - I$ is equal to the number of recurrent classes. A basis is given by the invariant distributions of each recurrent class.*

Note that the numbers in each row of $P - I$ adds up to zero. This is the same as saying that the column vectors of $P - I$ add up to the zero vector.

In order to find the left null space of $P - I$ we have to do *column operations* on $P - I$ to reduce it to *column echelon form*! This is not such a terrible thing. For example, you can always eliminate the last column using column operations, namely, add the first three columns to the last column. It becomes all zero! So we have:

$$\begin{pmatrix} -1/2 & 0 & 0 & 0 \\ 1/4 & -3/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 \\ 1/4 & 0 & 0 & 0 \end{pmatrix}$$

Now, multiply the fourth column by 4 then, using column operations, clear the 2nd row:

$$\begin{pmatrix} -1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1/4 & 0 & 0 & 0 \end{pmatrix}$$

This is not quite in column echelon form. But it is good enough to answer all the questions because:

Every row has at most one nonzero entry.

(1) The rank of $P - I$ is 2, the number of nonzero columns.
(2) The dimension of the null space of $P - I$ is 2 since

$$\dim \text{Null space} = \text{size} - \text{rank} = 4 - 2 = 2.$$

Therefore, there are 2 recurrent classes.
(3) A basis for the null space is given by
   (a) $(0, 0, 1, 0)$
   (b) $(1, 0, 0, 2)$
(4) If we normalize these two vectors (divide by the sum of the coordinates), we get the *basic invariant distributions*:
   (a) $\beta = (0, 0, 1, 0)$
   (b) $\gamma = (1/3, 0, 0, 2/3)$
(5) These are the unique invariant distributions for the two recurrent classes. So, their *supports* $\{3\}$ and $\{1, 4\}$ are the recurrent classes.
(6) Now we can find all invariant distributions. They are given by

$$\pi = t\gamma + (1-t)\beta = \left(\frac{t}{3}, 0, 1-t, \frac{2t}{3}\right)$$

for $0 \le t \le 1$.
(7) This represents the long term distribution where t is the probability of ending up in the recurrent class $\{1, 4\}$ and $1 - t$ is the probability of ending up in the other recurrent class $\{3\}$. For example, if the initial distribution is

$$\alpha = (1/4, 1/4, 1/4, 1/4)$$

then

$$t = \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{4} + 0 + \frac{1}{4} = \frac{5}{8}.$$

So, in the long run (as $n \to \infty$) we get:

$$\lim_{n\to\infty} \alpha P^n = \left(\frac{t}{3}, 0, 1-t, \frac{2t}{3}\right) = \left(\frac{5}{24}, 0, \frac{3}{8}, \frac{5}{12}\right)$$