

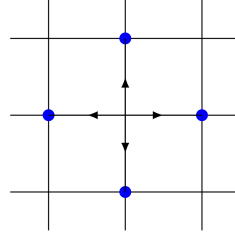
S&DS 351: Stochastic Processes - Homework 4

Bryan SebaRaj

Professor Ilias Zadik

February 14, 2025

1. (Nearest-neighbor walk in two dimensions) Let $\{X_n : n \geq 0\}$ denote the nearest-neighbor walk on \mathbb{Z}^2 : given $X_n = (a, b)$, the next position X_{n+1} is equally likely to be one of **four** possibilities: $(a + 1, b)$, $(a - 1, b)$, $(a, b + 1)$, or $(a, b - 1)$, as shown in the following picture.



- (a) (8 points) Let u_n denote the probability of returning to the origin in n steps if the chain starts from the origin. For integer $m \geq 0$, show that $u_{2m+1} = 0$ and

$$u_{2m} = \frac{1}{4^{2m}} \sum_{k=0}^m \frac{(2m)!}{\{k!(m-k)!\}^2}$$

(Hint: the walk either moves horizontally or vertically. If $2k$ steps are horizontal, then $2(m-k)$ steps are vertical.)

Observe that a return to the origin in an odd number of steps is trivially impossible, since each step changes one coordinate by ± 1 . In order to reach the origin again, the vertical and horizontal displacements must both be 0; as such, the number of steps to the left must equal the number of steps to the right, and the number of steps up must equal the number of steps down. This is impossible in a scenario with an odd number of steps. Therefore,

$$u_{2m+1} = 0 \quad \text{for all integer } m \geq 0$$

Computing u_{2m} : Since each step is equally likely to be horizontal or vertical, we may break down the $2m$ steps into $2k$ horizontal steps $2(m-k)$ vertical steps for some k with $0 \leq k \leq m$.

Within those $2k$ horizontal steps, there must be k steps to the right and left each to end at $a = 0$. The number of ways to arrange those $2k$ steps is

$$\binom{2k}{k}$$

Similarly, for the $2(m-k)$ vertical steps, $(m-k)$ must be up and $(m-k)$ must be down. The number of ways to arrange $2(m-k)$ steps is

$$\binom{2(m-k)}{m-k}$$

Note that we also need to consider k in choosing which $2k$ steps are horizontal. The number of ways to choose which $2k$ steps (out of $2m$) are horizontal is

$$\binom{2m}{2k}$$

Therefore, for a fixed k , the total number of configurations which lead back to the origin is

$$\binom{2m}{2k} \binom{2k}{k} \binom{2(m-k)}{m-k}$$

Since each of the $2m$ steps is one of four equally likely moves (left, right, up, down), each distinct path occurs with probability 4^{-2m} . Summing over all k from 0 to m yields

$$u_{2m} = 4^{-2m} \sum_{k=0}^m \binom{2m}{2k} \binom{2k}{k} \binom{2(m-k)}{m-k}$$

Simplifying the binomial terms,

$$\begin{aligned} \binom{2m}{2k} \binom{2k}{k} \binom{2(m-k)}{m-k} &= \left(\frac{(2m)!}{(2k)!(2m-2k)!} \right) \left(\frac{(2k)!}{k!k!} \right) \left(\frac{(2(m-k))!}{(m-k)!(m-k)!} \right) = \\ &= \frac{(2m)!}{k!k!(m-k)!(m-k)!} = \frac{(2m)!}{[k!(m-k)!]^2} \end{aligned}$$

Thus,

$$u_{2m} = \frac{1}{4^{2m}} \sum_{k=0}^m \frac{(2m)!}{[k!(m-k)!]^2}$$

(b) (5 points) Use the fact that $\sum_{k=0}^m \binom{m}{k}^2 = \binom{2m}{m}$ to arrive at the simple expression

$$u_{2m} = \frac{1}{4^{2m}} \binom{2m}{m}^2.$$

Rearranging the sum,

$$\begin{aligned} \sum_{k=0}^m \frac{(2m)!}{[k!(m-k)!]^2} &= \sum_{k=0}^m (2m)! \left(\frac{m!}{k!(m-k)!} \right)^2 \cdot \frac{1}{(m!)^2} = \\ \sum_{k=0}^m \frac{(2m)!}{(m!)^2} \cdot (m!)^2 \binom{m}{k}^2 \frac{1}{(m!)^2} &= \sum_{k=0}^m \binom{2m}{m} \binom{m}{k}^2 = \binom{2m}{m} \sum_{k=0}^m \binom{m}{k}^2 = \binom{2m}{m}^2. \end{aligned}$$

Thus,

$$u_{2m} = \frac{1}{4^{2m}} \binom{2m}{m}^2.$$

(c) (10 points) As we did in class, use Stirling's approximation $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$ as $n \rightarrow \infty$ to conclude that $\sum_{n \geq 0} u_n = \infty$. (Hint: recall the fact mentioned in class: $\sum_{m \geq 1} m^{-\alpha} = \infty$ if $\alpha \leq 1$ and $< \infty$ if $\alpha > 1$.)

Since $u_{2m+1} = 0$,

$$\sum_{m=0}^{\infty} u_m = \sum_{m=0}^{\infty} u_{2m}.$$

Using the closed-form expression,

$$\begin{aligned} u_{2m} &= \frac{1}{4^{2m}} \binom{2m}{m}^2, \\ m! &\sim \left(\frac{m}{e}\right)^m \sqrt{2\pi m} \end{aligned}$$

For our binomial,

$$\binom{2m}{m} = \frac{(2m)!}{(m!)^2} \sim \frac{\sqrt{4\pi m} \left(\frac{2m}{e}\right)^{2m}}{(\sqrt{2\pi m} \left(\frac{m}{e}\right)^m)^2} = \frac{\sqrt{4\pi m} (2m)^{2m} e^{-2m}}{2\pi m m^{2m} e^{-2m}} = \frac{4^m}{\sqrt{\pi m}}.$$

Therefore,

$$\binom{2m}{m}^2 \sim \frac{(4^m)^2}{\pi m}.$$

Therefore, the full equation is

$$u_{2m} = \frac{1}{4^{2m}} \binom{2m}{m}^2 \sim \frac{1}{4^{2m}} \frac{(4^m)^2}{\pi m} = \frac{1}{\pi m}.$$

So for any large m ,

$$u_{2m} \approx \frac{1}{\pi m}.$$

Recall that the following harmonic series diverges:

$$\sum_{m=1}^{\infty} \frac{1}{m}$$

$$\sum_{m=1}^{\infty} u_{2m} = \frac{1}{\pi} \sum_{m=1}^{\infty} \frac{1}{m}$$

must also diverge. Therefore,

$$\sum_{n=0}^{\infty} u_n = \sum_{m=0}^{\infty} u_{2m} = \infty.$$

- (d) (2 points) Use the classification criterion to conclude that the chain is *recurrent*.

By the standard classification criterion for Markov chains, if

$$\mathbb{P}_i\{T_i = \infty\} = 1$$

then the state is recurrent. Since $\sum_{n=0}^{\infty} u_n = \infty$,

$$\mathbb{P}_0\{T_0 = \infty\} = 1$$

As such, the 2D nearest-neighbor random walk from the origin is recurrent on the origin. Note that recurrence holds, starting from any point on the plane and ending on the same point eventually, as the original assumptions in part (a) must trivially still hold for every point on the plane.

2. (Nearest-neighbor walk in three dimensions) Now that we have warmed up, let's consider $\{X_n : n \geq 0\}$ being the nearest-neighbor walk on \mathbb{Z}^3 : given $X_n = (a, b, c)$, the next position X_{n+1} is equally likely to be one of **six** possibilities: $(a+1, b, c)$, $(a-1, b, c)$, $(a, b+1, c)$, $(a, b-1, c)$, $(a, b, c+1)$, or $(a, b, c-1)$.

- (a) (8 points) Again, let u_n denote the probability of returning to the origin in n steps if the chain starts from the origin. For integer $m \geq 0$, show that $u_{2m+1} = 0$ and

$$u_{2m} = \frac{1}{6^{2m}} \sum_{k=0}^m \sum_{j=0}^{m-k} \frac{(2m)!}{\{j!k!(m-j-k)!\}^2}$$

(Hint: Mimic the reasoning from part (a) in Problem 1, this time in three dimensions.)

Note that in each step, one of the three coordinates changes by ± 1 . After an odd number of steps, the parity of the sum of coordinates will trivially be odd. Thus, it is impossible to be at $(0, 0, 0)$, with a sum with an even parity, after an odd number of steps. Therefore,

$$u_{2m+1} = 0 \quad \text{for all } m \geq 0$$

Computing u_{2m} :

Note that in order to return to $(0, 0, 0)$ after $2m$ steps, the total net displacement in each of the three directions must to be zero.

- Define the total number of steps that change the x -coordinate be $2j$ (so there are $j+1$ and $j-1$ jumps).
- Define the total number of steps that change the y -coordinate be $2k$ (so there are $k+1$ and $k-1$ jumps).
- Define the total number of steps that change the z -coordinate be $2(m-j-k)$ (so there are $m-j-k+1$ and $m-j-k-1$ jumps).

Since there is $2m$ steps in total,

$$2j + 2k + 2(m-j-k) = 2m$$

Given $j, k \geq 0$ such that $j+k \leq m$, the number of ways to choose which $2j$ of the $2m$ steps affect the x -coordinate is given by the binomial terms

$$\binom{2m}{2j} \binom{2j}{j}$$

Of the remaining $2m-2j$ steps, the $2k$ that affect the y -coordinate is chosen using the binomial,

$$\binom{2m-2j}{2k} \binom{2k}{k}.$$

The remaining $2(m-j-k)$ steps must affect the z -coordinate, with a corresponding binomial $\binom{2(m-j-k)}{m-j-k}$ to consider the $+1$ and -1 on the z -coordinate.

Hence, for a given pair of j, k , the number of paths returning to the origin is given by

$$\binom{2m}{2j} \binom{2j}{j} \binom{2m-2j}{2k} \binom{2k}{k} \binom{2(m-j-k)}{m-j-k}.$$

Note that trivially, each path has probability $(1/6)^{2m}$ because at each of the $2m$ steps, there are 6 equally likely moves

Summing over all valid pairs (j, k) with $0 \leq j \leq m$ and $0 \leq k \leq m-j$,

$$u_{2m} = \frac{1}{6^{2m}} \sum_{k=0}^m \sum_{j=0}^{m-k} \binom{2m}{2j} \binom{2j}{j} \binom{2m-2j}{2k} \binom{2k}{k} \binom{2(m-j-k)}{m-j-k}.$$

Rewriting the combinatorial coefficients as factorials,

$$\begin{aligned} & \binom{2m}{2j} \binom{2j}{j} \binom{2m-2j}{2k} \binom{2k}{k} \binom{2(m-j-k)}{m-j-k} = \\ & \frac{(2m)!}{(2j)!(2m-2j)!} \times \frac{(2j)!}{j!j!} \times \frac{(2m-2j)!}{(2k)!(2m-2j-2k)!} \times \frac{(2k)!}{k!k!} \times \frac{(2(m-j-k))!}{(m-j-k)!(m-j-k)!} = \\ & \frac{(2m)!}{\{j!k!(m-j-k)!\}^2}. \end{aligned}$$

Thus,

$$u_{2m} = \frac{1}{6^{2m}} \sum_{k=0}^m \sum_{j=0}^{m-k} \frac{(2m)!}{\{j!k!(m-j-k)!\}^2}.$$

(b) (5 points) Simplify the above expression as

$$u_{2m} = \frac{1}{2^{2m}} \binom{2m}{m} \sum_{k=0}^m \sum_{j=0}^{m-k} p_{k,j}, \text{ where } p_{k,j} = \frac{1}{3^m} \binom{m}{k} \binom{m-k}{j}.$$

Note that $\binom{2m}{m} = \frac{(2m)!}{(m!)^2}$. Therefore, the inner sum can be rewritten as

$$\frac{(2m)!}{\{j!k!(m-j-k)!\}^2} = \binom{2m}{m} \frac{(m!)^2}{\{j!k!(m-j-k)!\}^2}.$$

Hence,

$$u_{2m} = \frac{1}{6^{2m}} \sum_{k=0}^m \sum_{j=0}^{m-k} \binom{2m}{m} \frac{(m!)^2}{\{j! k! (m-j-k)!\}^2} = \binom{2m}{m} \frac{1}{6^{2m}} \sum_{k=0}^m \sum_{j=0}^{m-k} \frac{(m!)^2}{\{j! k! (m-j-k)!\}^2}.$$

Notice the multinomial coefficient,

$$\frac{(m!)^2}{\{j! k! (m-j-k)!\}^2} = \left[\binom{m}{j, k, m-j-k} \right]^2$$

This suggests we factor the sum as

$$\begin{aligned} \sum_{k=0}^m \sum_{j=0}^{m-k} \binom{m}{j, k, m-j-k}^2 &= \sum_{k=0}^m \sum_{j=0}^{m-k} \left(\binom{j}{j} \binom{m-j-k+j}{j} \binom{m-j-k+j+k}{k} \right)^2 = \\ &= \sum_{k=0}^m \sum_{j=0}^{m-k} \left(\binom{m}{k} \binom{m-k}{j} \right)^2 \end{aligned}$$

Splitting up the leading term from u_{2m} ,

$$\frac{1}{6^{2m}} = \frac{1}{(2 \cdot 3)^{2m}} = \frac{1}{2^{2m}} \frac{1}{3^{2m}}$$

Moving $\frac{1}{3^{2m}}$ into the sum,

$$u_{2m} = \binom{2m}{m} \frac{1}{2^{2m}} \sum_{k=0}^m \sum_{j=0}^{m-k} \frac{1}{3^{2m}} \left(\binom{m}{k} \binom{m-k}{j} \right)^2$$

Using the definition of $p_{k,j}$, the sum can be rewritten as

$$u_{2m} = \frac{1}{2^{2m}} \binom{2m}{m} \sum_{k=0}^m \sum_{j=0}^{m-k} (p_{k,j})^2$$

(c) (10 points) Show that

$$\sum_{k=0}^m \sum_{j=0}^{m-k} p_{k,j} = 1.$$

(Hint: how many ways to distribute m balls into 3 urns?)

Reversing the steps used in the previous part, we see that

$$\binom{m}{k} \binom{m-k}{j} = \frac{m!}{k! (m-k)!} \times \frac{(m-k)!}{j! (m-k-j)!} = \frac{m!}{k! j! (m-k-j)!}$$

Therefore,

$$\sum_{k=0}^m \sum_{j=0}^{m-k} \binom{m}{k} \binom{m-k}{j} = \sum_{k=0}^m \sum_{j=0}^{m-k} \frac{m!}{k! j! (m-k-j)!}$$

This is the standard multinomial expansion of $(1 + 1 + 1)^m = 3^m$. This can be interpreted $k, j, m-k-j$ as a partition of m balls into 3 disjoint urns, with one urn for each axis (x -urn, y -urn, z -urn), and the total number of ways is exactly 3^m .

Therefore,

$$\sum_{k=0}^m \sum_{j=0}^{m-k} \binom{m}{k} \binom{m-k}{j} = 3^m$$

Dividing both sides by 3^m ,

$$\sum_{k=0}^m \sum_{j=0}^{m-k} \frac{1}{3^m} \binom{m}{k} \binom{m-k}{j} = 1$$

By the definition of $p_{k,j}$,

$$\sum_{k=0}^m \sum_{j=0}^{m-k} p_{k,j} = 1$$

(d) (Bonus, 5 points) Using Stirling's approximation, one can show that

$$\max_{k,j} p_{k,j} \leq \frac{C}{m},$$

for some positive constant C , where the maximum is over all $k, j \geq 0$ such that $k + j \leq m$.

Expanding the definition of $p_{k,j}$,

$$\binom{m}{k} = \frac{m!}{k!(m-k)!}$$

Taking the natural logarithm,

$$\ln \binom{m}{k} = \ln m! - \ln k! - \ln(m-k)!$$

Using Stirling's approximation,

$$\ln(k!) \approx k \ln(k) - k \text{ and } \ln(m-k)! \approx (m-k) \ln(m-k) - (m-k)$$

Differentiating with respect to k and setting to zero,

$$-(\ln(m-k) + 1) + (\ln(k) + 1) = 0$$

$$\ln(k) = \ln(m-k)$$

$$k = m - k$$

$$k = \frac{m}{2}$$

Similarly for $\binom{m-k}{j}$,

$$\binom{m-k}{j} = \frac{(m-k)!}{j!(m-k-j)!}$$

Taking the natural logarithm,

$$\ln \binom{m-k}{j} = \ln(m-k)! - \ln j! - \ln(m-k-j)!$$

Using Stirling's approximation,

$$\ln(j!) \approx j \ln(j) - j \text{ and } \ln(m-k-j)! \approx (m-k-j) \ln(m-k-j) - (m-k-j)$$

Differentiating with respect to j and setting to zero,

$$-(\ln(m-k-j) + 1) + (\ln(j) + 1) = 0$$

$$\ln(j) = \ln(m-k-j)$$

$$2j = m - k$$

$$j = \frac{m-k}{2}$$

Therefore, $\binom{m}{k}$ is maximized around $k \approx m/2$, and $\binom{m-k}{j}$ is maximized around $j \approx (m-k)/2$. Approximating $\binom{m}{k}$ when $k \approx m/2$,

$$\binom{m}{m/2} \approx \frac{2^m}{\sqrt{\pi m}}$$

Similarly, for $\binom{m-k}{j}$ at $j \approx (m-k)/2$:

$$\binom{m-k}{(m-k)/2} \approx \frac{2^{m-k}}{\sqrt{\pi(m-k)}}$$

Thus, at the maximum,

$$p_{k,j} \approx \frac{1}{3^m} \cdot \frac{2^m}{\sqrt{\pi m}} \cdot \frac{2^{m-k}}{\sqrt{\pi(m-k)}}$$

Setting $k \approx m/2$,

$$p_{k,j} \approx \frac{1}{3^m} \cdot \frac{2^m}{\sqrt{\pi m}} \cdot \frac{2^{m/2}}{\sqrt{\pi(m/2)}} = \frac{2^{3m/2}}{3^m \pi \sqrt{m} \sqrt{m/2}} = \frac{2^{3m/2}}{3^m \pi m \sqrt{2}}$$

Given that $\frac{2^{3m/2}}{3^m \pi m \sqrt{2}} \leq 1, \forall m \geq 0$,

$$\max_{k,j} p_{k,j} \leq \frac{C}{m}$$

for some constant C .

- (e) (5 points) Combine parts (b), (c), (d) and use Stirling's approximation to show that $u_{2m} \leq \frac{C'}{m^{3/2}}$ for some constant C' and $\sum_{n \geq 0} u_n < \infty$.

Recall that

$$u_{2m} = \frac{1}{2^{2m}} \binom{2m}{m} \sum_{k=0}^m \sum_{j=0}^{m-k} p_{k,j}^2$$

Using the bound from part (d),

$$p_{k,j} \leq \frac{C}{m}$$

so

$$p_{k,j} = \frac{1}{3^m} \binom{m}{k} \binom{m-k}{j} \leq \frac{C}{m}$$

Therefore,

$$u_{2m} \leq \frac{C}{m} \frac{1}{2^{2m}} \binom{2m}{m}.$$

Using Stirling's approximation for $\binom{2m}{m}$,

$$\binom{2m}{m} = \frac{(2m)!}{(m!)^2} \sim \frac{\sqrt{4\pi m} \left(\frac{2m}{e}\right)^{2m}}{(\sqrt{2\pi m} \left(\frac{m}{e}\right)^m)^2} = \frac{4^m}{\sqrt{\pi m}}$$

Thus,

$$\frac{1}{2^{2m}} \binom{2m}{m} \sim \frac{4^m}{2^{2m} \sqrt{\pi m}} = \frac{4^m}{4^m \sqrt{\pi m}} = \frac{1}{\sqrt{\pi m}}$$

Therefore,

$$u_{2m} \leq \frac{C}{m} \frac{1}{2^{2m}} \binom{2m}{m} \sim \frac{C}{m} \frac{1}{\sqrt{\pi m}} = \frac{C}{\sqrt{\pi} m^{3/2}}$$

Define $C' = \frac{C}{\sqrt{\pi}}$ to get

$$u_{2m} \leq \frac{C'}{m^{3/2}}$$

Note that $\sum_{m=0}^{\infty} \frac{C'}{m^{3/2}} < \infty$ and converges, as $d = 3$ in $\sum_{m=0}^{\infty} \frac{1}{m^{d/2}}$ (from lecture). Since $u_{2m} \leq \frac{C'}{m^{3/2}}$ for all $m \geq 0$,

$$\sum_{m=0}^{\infty} u_{2m} \text{ converges}$$

As such,

$$\sum_{n=0}^{\infty} u_n = \sum_{m=0}^{\infty} u_{2m} < \infty.$$

- (f) (2 points) Use the classification criterion to conclude that the chain is *transient*.

By the standard classification criterion for Markov chains, a state is transient if

$$\sum_{n=0}^{\infty} u_n < \infty$$

Since we have shown

$$\sum_{n=0}^{\infty} u_n < \infty$$

it follows that the origin is transient. Since the underlying assumptions from part (a) are required for the return to the initial state for all states in the 3D state space, all points in S are also transient. Thus, the 3D nearest-neighbor random walk is transient.

3. Let i be a recurrent state of a Markov Chain. Assume another state j is accessible from i .

- (a) (3 points) Prove that

$$\mathbb{P}(N_i = \infty | X_0 = i) = 1.$$

Define the first return time to i (after time 0) as

$$T_i^{(1)} = \inf\{n \geq 1 : X_n = i\}.$$

Define recursively the k -th return time to i as

$$T_i^{(k+1)} = \inf\{n > T_i^{(k)} : X_n = i\}, \quad k \geq 1$$

Note that $T_i^{(k)} = \infty$ if no further visit occurs. Since i is a recurrent state, $\mathbb{P}_i(T_i^{(k)} < \infty) = 1$ for all $k \geq 1$. As such, with probability 1, we have infinitely many finite return times to i when starting from state i . Therefore,

$$N_i = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = i\}} = \infty \quad \text{almost surely, when } X_0 = i$$

Thus,

$$\mathbb{P}_i(N_i = \infty | X_0 = i) = 1$$

- (b) (10 points) Prove that

$$\mathbb{P}(T_j < \infty | X_0 = i) = 1.$$

Note that state i is recurrent. Let us define the random sequence of return times to i as

$$T_i^{(k)}, \quad k = 1, 2, 3, \dots$$

Trivially, with probability 1, there are infinitely many such return times (by definition of recurrence) when starting at state i .

From the Markov property and the fact that j is accessible from i , each time the chain enters i , the event $\{X_{T_i^{(k)}+m} = j\}$ has probability $p > 0$, regardless of the states prior to $X_{T_i^{(k)}} = i$.

Suppose, for the sake of contradiction, that with some positive probability the chain never hits j . On that event, there would be infinitely many returns to i due to i 's recurrence, but every time the chain is at i , the chain fails to proceed to j at time $T_i^{(k)} + m$.

The probability of failing to hit j on all of the infinitely many attempts approaches 0, as at each attempt there is at least probability $p > 0$ of reaching state j since j is accessible from i , i.e.

$$\mathbb{P}_i\left(X_{T_i^{(k)}+m} \neq j, \forall k\right) \leq \lim_{K \rightarrow \infty} (1-p)^K = 0$$

Hence, the chain must eventually hit j with probability 1 when starting at state i . Thus

$$\mathbb{P}(T_j < \infty | X_0 = i) = 1$$

(c) (5 points) Prove that

$$\mathbb{P}(T_i < \infty | X_0 = j) = 1.$$

Proof by contradiction: Let

$$A = \{T_i = \infty\} \quad \text{and} \quad A^c = \{T_i < \infty\}$$

Suppose, for the sake of contradiction, that $\mathbb{P}_j(A) > 0$. Then, on event A , starting from j , the chain never reaches i . Now consider a chain which starts from i . From part (b),

$$\mathbb{P}(T_j < \infty | X_0 = i) = 1.$$

As such, the chain starting from i reaches j at some finite time with positive probability. Consider the paths, starting from i , that reach j and then remain in event A , i.e. never returning to i after j . By the strong Markov property,

$$\mathbb{P}_i(\text{reach } j \text{ and stay in } A) = \mathbb{P}_i(T_j < \infty) \mathbb{P}_j(A) = 1 \cdot \mathbb{P}_j(A) = \mathbb{P}_j(A)$$

Recall that we assume $\mathbb{P}_j(A) > 0$.

As such, from i , there is a strictly positive probability $\mathbb{P}_j(A)$ of eventually ending up in a path that never returns to i .

However, i is a recurrent state, which by definition, must be revisited infinitely often with probability 1. This yields a contradiction with the existence of a path starting from i through j that may never return back to i . Hence the assumption $\mathbb{P}_j(A) > 0$ was false.

Therefore,

$$\mathbb{P}(T_i = \infty | X_m = j, \dots, X_0 = i) = 0$$

Note that by the Markov property, the chain starting at i and reaching j is independent of the chain's progress after j . Therefore,

$$\mathbb{P}(T_i = \infty | X_m = j) = 0$$

At $m = 0$,

$$\mathbb{P}(T_i = \infty | X_0 = j) = 0$$

Thus,

$$\mathbb{P}(T_i < \infty | X_0 = j) = 1.$$

(d) (2 points) Is it true that j needs also to be recurrent?

Recall from part (c)

$$\mathbb{P}(T_i < \infty | X_0 = j) = 1.$$

Hence, starting at j , the chain will reach i in finite time.

Recall from part (b)

$$\mathbb{P}(T_j < \infty | X_0 = i) = 1.$$

As such, there is a probability of 1 of returning to j from j (via j to i , then i to j), or

$$\mathbb{P}(T_j < \infty | X_0 = j) = 1$$

$$\mathbb{P}(N_j = \infty | X_0 = j) = 1$$

Thus, j must trivially be recurrent.

Using the previous statements and general theorems:

Note that this fits the standard theorem that all states in the same communicating class share the same classification as either all recurrent or all transient. In other words, if x and y are two states that communicate, then x is recurrent if and only if y is recurrent.

Since part (b) shows j is accessible from i and (c) shows i is accessible from j , the states i and j communicate. Because i is assumed to be recurrent, it follows that j must be recurrent as well.

4. Let j be a transient state of a Markov chain and i an arbitrary state of the Markov chain.

(a) (10 points) Prove that for some constant $C > 0$ and some $p \in (0, 1)$ it holds for all $t \geq 1$

$$\mathbb{P}(N_j \geq t | X_0 = i) = 1 \leq Cp^t.$$

Define the *first hitting time* of j as

$$T_j := \inf\{n \geq 0 : X_n = j\}.$$

(We may set $T_j = \infty$ if the chain never visits j .)

Since j is transient, there is a probability strictly less than 1 of ever making a *return* to j starting from j . More concretely, let

$$r := \mathbb{P}_j(T_j < \infty \text{ for a second time}),$$

i.e. r is the probability that the chain, starting in j , returns to j at least once more in the future. If j is transient, we must have

$$r < 1.$$

(A recurrent state j would have $r = 1$, but we know j is transient.)

Step 2. The event $\{N_j \geq t\}$ implies $t - 1$ returns to j .

Let us define the successive hitting times of j :

$$T_j^{(1)} := T_j, \quad T_j^{(k+1)} := \inf\{n > T_j^{(k)} : X_n = j\}.$$

We say $T_j^{(k)} = \infty$ if the chain does not visit j for the k -th time. Then

$$\{N_j \geq t\} \Leftrightarrow \{T_j^{(t)} < \infty\}.$$

Hence, to have at least t visits to j , the chain must:

- First, hit j once (for time $T_j^{(1)}$).
- Then, starting from that first visit j , *return* to j at least $(t - 1)$ additional times.

Step 3. Bounding $\mathbb{P}_i(N_j \geq t)$.

We now use a decomposition over the time of the *first* visit to j . From an arbitrary state i , let $a := \mathbb{P}_i(T_j < \infty)$. (Note a could be zero if it's impossible or very unlikely to reach j , or it might be positive. We do *not* need $a > 0$ for the final statements to hold, but usually $a > 0$ is typical if j communicates with i .)

Once the chain has arrived in j , the probability of returning j an additional $(t - 1)$ times is at most r^{t-1} by the Markov property (each new return from j to j has probability at most r).

Hence, we can bound

$$\mathbb{P}_i(N_j \geq t) = \mathbb{P}_i(T_j^{(t)} < \infty) \leq \mathbb{P}_i(T_j^{(1)} < \infty) \times (\text{max prob of } (t - 1) \text{ further returns to } j).$$

But

$$\mathbb{P}_i(T_j^{(1)} < \infty) = a, \quad \text{and} \quad (\text{max prob of } (t - 1) \text{ returns}) \leq r^{t-1}.$$

Therefore

$$\mathbb{P}_i(N_j \geq t) \leq a r^{t-1}.$$

Setting $p = r < 1$ and $C = ar^{-1}$, we get

$$\mathbb{P}_i(N_j \geq t) \leq ar^{t-1} = ar^{-1}r^t = Cp^t, \quad \text{where } C = ar^{-1}, p = r \in (0, 1).$$

That completes the exponential bound.

(If $a = 0$, this inequality trivially holds for all t , picking any $C > 0$ and $p \in (0, 1)$, since the LHS is always zero.)

(b) (5 points) Conclude that

$$\lim_{n \rightarrow \infty} P_{i,j}^n = 0.$$

Given the exponential bound

$$\mathbb{P}_i(N_j \geq t) \leq Cp^t \quad (\text{for some } C > 0, p < 1),$$

observe that

$$\{X_n = j\} \subseteq \{N_j \geq 1\}.$$

But more precisely, if $X_n = j$, then necessarily $N_j \geq 1$. Even if $n = 0$, we interpret it similarly ($X_0 = i$ assumed in $P^n(i, j)$, so $n > 0$). So

$$P^n(i, j) = \mathbb{P}_i(X_n = j) \leq \mathbb{P}_i(N_j \geq 1).$$

In fact, we can be more refined: to be in j at time n , the chain must visit j at least $\lceil n/\tau \rceil$ times for some threshold τ , etc. But the simpler inclusion above suffices to deduce a limit of zero. Specifically, from part (a) with $t = 1$, we know

$$\mathbb{P}_i(N_j \geq 1) \leq Cp^1 = Cp < 1$$

and indeed for any $t \geq 1$.

Actually, to conclude

$$\lim_{n \rightarrow \infty} P^n(i, j) = 0,$$

we can use a common argument: if $\{P^n(i, j)\}_{n \geq 1}$ did *not* tend to 0, there would exist some $\epsilon > 0$ and an infinite subsequence n_k such that $P^{n_k}(i, j) > \epsilon$ for all k . That would imply that with probability at least ϵ (from state i), we hit j at times n_k . But that would mean N_j is arbitrarily large with probability at least ϵ , contradicting the exponential bound on $\mathbb{P}_i(N_j \geq t)$ for large t .

Hence the only possibility is $P^n(i, j) \rightarrow 0$ as $n \rightarrow \infty$. This completes the proof of part (b).

Chang 1.25. Prove the following proposition:

The total variation distance $\|\lambda - \mu\|$ may also be expressed in the alternative forms:

$$\|\lambda - \mu\| = \sup_{A \subseteq S} |\lambda(A) - \mu(A)| = \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)| = 1 - \sum_{i \in S} \min\{\lambda(i), \mu(i)\}.$$

textbf1. Showing $\|\lambda - \mu\| = \frac{1}{2} \sum_i |\lambda(i) - \mu(i)|$. (a) **Upper bound:** For any set $A \subseteq S$, we have

$$\lambda(A) - \mu(A) = \sum_{i \in A} \lambda(i) - \sum_{i \in A} \mu(i) = \sum_{i \in A} (\lambda(i) - \mu(i)).$$

Hence

$$|\lambda(A) - \mu(A)| \leq \sum_{i \in A} |\lambda(i) - \mu(i)| \leq \sum_{i \in S} |\lambda(i) - \mu(i)|.$$

It follows that

$$\|\lambda - \mu\| = \sup_{A \subseteq S} |\lambda(A) - \mu(A)| \leq \sum_{i \in S} |\lambda(i) - \mu(i)|.$$

However, we will refine this to show it is exactly $\frac{1}{2} \sum_i |\lambda(i) - \mu(i)|$. (b) **Constructing a maximizing**

set: Observe that

$$\sum_{i \in S} |\lambda(i) - \mu(i)| = \sum_{i \in S} (\lambda(i) - \mu(i))_+ + \sum_{i \in S} (\mu(i) - \lambda(i))_+,$$

where $(x)_+ := \max\{x, 0\}$. Note that exactly one of $(\lambda(i) - \mu(i))_+$ or $(\mu(i) - \lambda(i))_+$ is nonzero for each i . This means

$$\sum_{i \in S} |\lambda(i) - \mu(i)| = 2 \sum_{i \in S} (\lambda(i) - \mu(i))_+ = 2 \sum_{i \in S} (\mu(i) - \lambda(i))_+.$$

Hence,

$$\sum_{i \in S} (\lambda(i) - \mu(i))_+ = \sum_{i \in S} (\mu(i) - \lambda(i))_+ = \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)|.$$

Define the set

$$A^* = \{i \in S : \lambda(i) \geq \mu(i)\}.$$

Then

$$\lambda(A^*) - \mu(A^*) = \sum_{i \in A^*} (\lambda(i) - \mu(i)) = \sum_{i \in S} (\lambda(i) - \mu(i))_+,$$

because $\lambda(i) - \mu(i)$ is nonnegative for $i \in A^*$ and nonpositive for $i \notin A^*$. Therefore,

$$|\lambda(A^*) - \mu(A^*)| = \sum_{i \in S} (\lambda(i) - \mu(i))_+ = \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)|.$$

Since $\|\lambda - \mu\| \geq |\lambda(A^*) - \mu(A^*)|$, we conclude

$$\|\lambda - \mu\| \geq \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)|.$$

Combined with the earlier upper bound, this proves

$$\|\lambda - \mu\| = \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)|.$$

2. Showing $\|\lambda - \mu\| = 1 - \sum_i \min\{\lambda(i), \mu(i)\}$. From the expression

$$\sum_{i \in S} |\lambda(i) - \mu(i)| = \sum_{i \in S} (\lambda(i) + \mu(i)) - 2 \sum_{i \in S} \min\{\lambda(i), \mu(i)\},$$

we see that

$$\sum_{i \in S} |\lambda(i) - \mu(i)| = \underbrace{(1+1)}_{=2} - 2 \sum_{i \in S} \min\{\lambda(i), \mu(i)\} = 2 \left[1 - \sum_{i \in S} \min\{\lambda(i), \mu(i)\} \right],$$

because λ and μ are probability measures, i.e. $\sum_{i \in S} \lambda(i) = \sum_{i \in S} \mu(i) = 1$. Dividing both sides by 2, and recalling from step (1) that $\|\lambda - \mu\| = \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)|$, we obtain:

$$\|\lambda - \mu\| = 1 - \sum_{i \in S} \min\{\lambda(i), \mu(i)\}.$$

Conclusion. The three expressions

$$\sup_{A \subseteq S} |\lambda(A) - \mu(A)|, \quad \frac{1}{2} \sum_{i \in S} |\lambda(i) - \mu(i)|, \quad 1 - \sum_{i \in S} \min\{\lambda(i), \mu(i)\}$$

are all equal. This completes the proof of the proposition.