# Handout 4

### STAT5710 Applied Stochastic Processes

Here we use the canonical decomposition to compute several quantities of interest for finite chains, and then consider what changes when the state space is countably infinite.

## Computing with the canonical form

For any finite chain, the canonical decomposition gives

$$P = \begin{bmatrix} Q & R \\ 0 & B \end{bmatrix}$$

and we showed that $P^n$ converges if and only if $C_1, \ldots, C_k$ are aperiodic, in which case

$$P^n \longrightarrow \begin{bmatrix} 0 & (I-Q)^{-1}R\Lambda \\ 0 & \Lambda \end{bmatrix}, \qquad \Lambda = \operatorname{diag}(\mathbf{1}\pi_1^T, \ldots, \mathbf{1}\pi_k^T)$$

where $\pi_1, \ldots, \pi_k$ are the ________________________ of $P_1, \ldots, P_k$.

1. When the limit exists, this representation facilitates computing several quantities of interest. The matrix $F = (I-Q)^{-1}$ is called the **fundamental matrix**. We state the following without proof.
   (a) $F_{ij} =$ ______________, the expected number of visits to $j$
   (b) $(F\mathbf{1})_i =$ ______________, the expected time to absorption into some closed class
   (c) $(FR_k\mathbf{1})_i =$ ______________, the absorption probability into $C_k$
   (d) $E(T_j \mid X_0 = j) =$ ______________ for $j \in C_1 \cup \cdots \cup C_k$, the expected return times

### Examples

2. **Gambler's ruin.** Consider the gambler's ruin problem with $p = 0.6$ and $k = 5$ (Dobrow, Example 3.27), which has transition matrix

$P =$

| | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0.4 | 0 | 0.6 | 0 | 0 | 0 |
| 2 | 0 | 0.4 | 0 | 0.6 | 0 | 0 |
| 3 | 0 | 0 | 0.4 | 0 | 0.6 | 0 |
| 4 | 0 | 0 | 0 | 0.4 | 0 | 0.6 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1 |

   (a) Here $S = T \cup C_1 \cup C_2$, where $T =$ ______________, $C_1 =$ ________, and $C_2 =$ ________
   (b) Permute the states to put $P$ in canonical form, and fill in the blocks.

$Q =$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |

$R =$

| | 0 | 5 |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |

$B =$

| | 0 | 5 |
|---|---|---|
| 0 | | |
| 5 | | |

The fundamental matrix is then

$$F = (I - Q)^{-1} = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 1.54 & 1.35 & 1.07 & 0.64 \\ 2 & 0.90 & 2.25 & 1.78 & 1.07 \\ 3 & 0.47 & 1.18 & 2.25 & 1.35 \\ 4 & 0.19 & 0.47 & 0.90 & 1.54 \end{array}$$

Suppose you start with \$2, so that $X_0 = 2$ with probability 1.

(c) How many times will you have \$1? $E(N_1 \mid X_0 = 2) =$ ________ = ________

(d) How long will you play? $E(\tau_C \mid X_0 = 2) =$ ________ = ________

(e) What is the likelihood of winning? $P(X_t \text{ enters } C_2 \mid X_0 = 2) =$ ________ = ________

(f) What is the chance you get to \$1 before \$4? *Hint:* pretend that 1 and 4 are absorbing states, and rewrite $F$.

$$\tilde{Q} = \begin{array}{c|cc} & 2 & 3 \\ \hline 2 & & \\ 3 & & \end{array} \qquad \tilde{R} = \begin{array}{c|cccc} & 1 & 4 & 0 & 5 \\ \hline 2 & & & & \\ 3 & & & & \end{array} \qquad \tilde{F} = \begin{array}{c|cc} & 2 & 3 \\ \hline 2 & 1.32 & 0.79 \\ 3 & 0.53 & 1.32 \end{array}$$

So $P(\text{absorbed into } 1 \mid X_0 = 2) = (\tilde{F}\tilde{R})_{11} =$ ________________ = ________

3. **Student progression.** Each year a student drops out (d), repeats the year, or moves on to the next, until they graduate (Gr), with the probabilities given below (Dobrow, Example 3.29).

$$P = \begin{array}{c|cccccc} & \text{d} & \text{Fr} & \text{So} & \text{Jr} & \text{Se} & \text{Gr} \\ \hline \text{d} & 1 & 0 & 0 & 0 & 0 & 0 \\ \text{Fr} & 0.06 & 0.03 & 0.91 & 0 & 0 & 0 \\ \text{So} & 0.06 & 0 & 0.03 & 0.91 & 0 & 0 \\ \text{Jr} & 0.04 & 0 & 0 & 0.03 & 0.93 & 0 \\ \text{Se} & 0.04 & 0 & 0 & 0 & 0.03 & 0.93 \\ \text{Gr} & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

(a) Put $P$ in canonical form with $T = \{\text{Fr}, \text{So}, \text{Jr}, \text{Se}\}$, and fill in $Q$ and $R$.

$$Q = \begin{array}{c|cccc} & \text{Fr} & \text{So} & \text{Jr} & \text{Se} \\ \hline \text{Fr} & & & & \\ \text{So} & & & & \\ \text{Jr} & & & & \\ \text{Se} & & & & \end{array} \qquad R = \begin{array}{c|cc} & \text{d} & \text{Gr} \\ \hline \text{Fr} & & \\ \text{So} & & \\ \text{Jr} & & \\ \text{Se} & & \end{array}$$

(b) The absorption probabilities are

$$FR = (I-Q)^{-1}R = \begin{array}{c|cc} & \text{d} & \text{Gr} \\ \hline \text{Fr} & 0.19 & 0.81 \\ \text{So} & 0.14 & 0.86 \\ \text{Jr} & 0.08 & 0.92 \\ \text{Se} & 0.04 & 0.96 \end{array}$$

so an incoming freshman graduates with probability ________, and a junior drops out with probability ________.

(c) What is the expected time to graduation, assuming the student graduates? Write

$$h_i \;=\; h_{i,\text{Gr}} \;=\; P(\tau_{\text{Gr}} < \infty \mid X_0 = i)$$

for the probability of graduating starting from state $i$. For $i \in T$, $h_i$ is the absorption probability $(FR)_{i,\text{Gr}}$ in (b), and $h_{\text{Gr}} = 1$, $h_{\text{d}} = 0$.

Conditioning on graduation gives another Markov chain. Use the Markov property at time $t+1$ to find its transition probabilities.

$$\begin{aligned} \tilde{p}_{ij} \;&=\; P(X_{t+1} = j \mid X_t = i,\; \tau_{\text{Gr}} < \infty) \\ &=\; \frac{P(X_{t+1} = j,\; \tau_{\text{Gr}} < \infty \mid X_t = i)}{P(\tau_{\text{Gr}} < \infty \mid X_t = i)} \\ &=\; \underline{\hspace{3cm}} \big/ \underline{\hspace{1.5cm}} \end{aligned}$$

(d) So $\tilde{p}_{i,\text{d}} =$ ________ and $\tilde{p}_{ii} =$ ________, and each class year moves on to the next with probability ________. Fill in $\tilde{Q}$.

$$\tilde{Q} = \begin{array}{c|cccc} & \text{Fr} & \text{So} & \text{Jr} & \text{Se} \\ \hline \text{Fr} & & & & \\ \text{So} & & & & \\ \text{Jr} & & & & \\ \text{Se} & & & & \end{array}$$

(e) The expected time to graduation for an incoming freshman is $(\tilde{F}\mathbf{1})_{\text{Fr}} =$ ________ years.

For additional applications, see the examples in Dobrow, Section 3.8.

## Countably infinite state spaces

### Random walks on $\mathbb{Z}^d$

What about a countably infinite state space? As an example, consider a **random walk** on the integer lattice $\mathbb{Z}^d$, which starts at the origin, $X_0 = 0$, and has transition probabilities

$$p_{ij} = \begin{cases} \frac{1}{2d}, & j = i \pm e_k \text{ for some } k \in \{1, \dots, d\} \\ 0, & \text{otherwise} \end{cases} \qquad \text{where } e_k = (0, \dots, 1, \dots, 0) \text{ has a 1 in position } k.$$

We will explore this example with a simulation activity. First, consider the following questions.

4. (a) Write out the transition probabilities for $d = 1$ and $d = 2$.

$d = 1:$ $\quad p_{ij} =$ ________ for $j =$ ____________________

$d = 2:$ $\quad p_{ij} =$ ________ for $j =$ ____________________________________

(b) Draw the transition graph for $d = 1$ on the states $-3, \dots, 3$.

(c) Mark every state of $\mathbb{Z}^2$ the walk can reach from the origin in one step. Then mark, in a different way, every state it can reach in exactly two steps.

(d) $X_t$ can be represented as a cumulative sum:

$$X_t \ =$$

## Irreducibility and periodicity

5. Is the walk irreducible?
   (a) In $d = 1$, describe a path from $i$ to a state $j > i$ that has positive probability. What is its probability?
   (b) In $d = 2$, describe a path from the origin to $(3, -2)$. What is the fewest number of steps it can take?
   (c) Does this argument work in every dimension?
6. What is the period of the origin?
   (a) What is the shortest path back?
   (b) $p_{00}^{(n)} = 0$ whenever...
   (c) The period is $d(0) =$ ________
   (d) Does this argument work in every dimension?
7. How could you modify the process to make it aperiodic?

### Simulation activity

Open `random-walks.R` and work through it section by section. Each item below goes with the section of the script named in bold. For a finite chain, a stationary distribution always exists, and a limit exists just in case every closed class is aperiodic. As you go, ask whether the same is true here.

8. **Is it irreducible?** Try a few targets. Does the walker always get there? Can a simulation show that the chain is irreducible?

9. **Is it periodic?** Do the steps at which the walker is at 0 agree with your answer to item 6? What about the lazy walker?

10. **Is it recurrent?**

    (a) How many walkers returned to 0? Of those that did, how long did it take on average?

    (b) Why is the running total of $p_{00}^{(n)}$ the average number of visits to 0? What would recurrence look like in that plot?

    (c) Based on your observations above, do you think that a walker is sure to return to 0 eventually?

11. **Limits.** What happens to $p_{0j}^{(n)}$ as $n$ grows, for any fixed $j$? (The simulation shows $j = 0$ and $j = 10$ as representative examples; feel free to change these to other even values to verify the pattern doesn't depend on these choices.) Is there a limiting distribution?

12. **Does the distribution of $X_t$ settle down?**

    (a) What happens to the distribution when the walkers start at 0? When they start spread out? Does it matter whether the walk is lazy?

    (b) Note that $\mu(j) = 1$ for every $j$ satisfies $\mu P = \mu$, for both the simple and the lazy walk. Why can't $\mu$ be scaled into a stationary distribution?

13. **More dimensions.** Which dimensions look recurrent, and which look transient? Can you tell for $d = 2$?

### Explaining the simulations

The simulations suggest that the walk is recurrent for $d = 1$ and transient for $d = 3$, but leave $d = 2$ unclear. Further, even when the walk is sure to return, it can take a very long time. Here we confirm both, and settle $d = 2$.

14. The random walk on $\mathbb{Z}^d$ is recurrent for $d =$ ________ and transient for $d \geq$ ________.

    **Proof sketch.** Recall that the origin is recurrent just in case $\sum_n p_{00}^{(n)} = \infty$. The partial sums of this series are the average-visits curves from the simulation.

    For the lazy walk, whose steps are uniform on $\{0, \pm e_1, \dots, \pm e_d\}$, a local central limit theorem for lattice sums shows that

    $$p_{00}^{(n)} \;\approx\; c\,n^{-d/2} \quad \text{for large } n.$$

    For the simple walk the same holds along multiples of 2 steps:

    $$p_{00}^{(2n)} \;\approx\; c\,n^{-d/2} \qquad \text{and} \qquad p_{00}^{(2n+1)} \;=\; \underline{\hspace{2cm}}$$

    Either way,

    $$\sum_n p_{00}^{(n)} \;\approx\; c \sum_n n^{-d/2}$$

    which is finite just in case $d$ ________

15. For $d = 1, 2$ the walk is recurrent, but its expected return time $\mu_0 = E(T_0 \mid X_0 = 0)$ is ________.

    **Proof.** Suppose instead that $\mu_0 < \infty$. Let $R_0 = 0$, and let $R_k$ be the time of the $k$th return to the origin. By time homogeneity, the times between returns $R_{k+1} - R_k$ are independent, each distributed as $T_0$.

    So the walk returns about once every $\mu_0$ steps, and the number of visits by step $n$ is

    $$N_0(n) \;=\; \max\{k : R_k \leq n\} \;\approx\; \underline{\hspace{2cm}}$$

    In other words, the walk spends a positive share of its time at the origin: $\frac{1}{n}N_0(n) \to 1/\mu_0 > 0$. But the expected share is

    $$E(\tfrac{1}{n}N_0(n) \mid X_0 = 0) \;=\; \frac{1}{n}\sum_{t=1}^{n} \underline{\hspace{2cm}} \;\longrightarrow\; \underline{\hspace{2cm}}$$

    by the local central limit theorem. This is a contradiction, so $\mu_0 = \infty$.

### Positive and null recurrence

16. A recurrent state $j$ is

    (a) **positive recurrent** if ________________

    (b) **null recurrent** if ________________

    So the random walk on $\mathbb{Z}^d$ is null recurrent for $d = 1, 2$. By contrast, all states in a recurrent finite closed class are positive recurrent.

17. A Markov chain is **ergodic** if it is ________________________ .

**Theorem.** A Markov chain has a strong limiting distribution, which is its unique stationary distribution, if it is ergodic.

See Dobrow, Section 3.10, for a proof.