

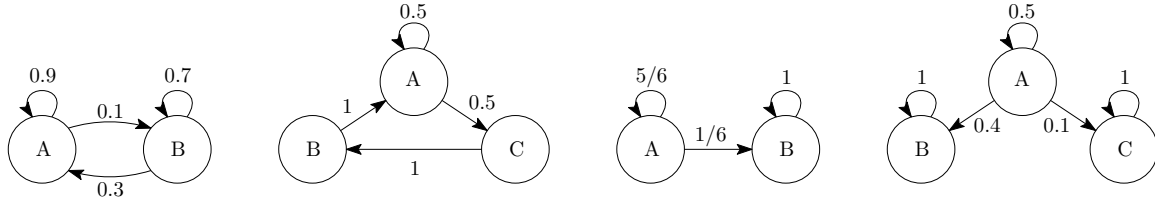
Exercise session 4 – Prob. & Stat. 2

Recall: Given a Markov chain with states $S = [n]$ and a transition matrix $P = (p_{i,j})_{i,j=1}^n$, a *stationary distribution* is any distribution $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ such that $\pi \cdot P = \pi$, that is, π satisfies the system of *balance equations*:

$$\sum_{i \in S} \pi_i \cdot p_{i,j} = \pi_j, \quad \text{for all } j \in S \quad (\dots \text{and also } \sum_{j \in S} \pi_j = 1).$$

Convergence theorem. If the Markov chain is *finite, irreducible, aperiodic* then the stationary distribution is unique and for any distribution $\rho = (\rho_1, \dots, \rho_n)$ we have $\lim_{n \rightarrow \infty} \rho \cdot P^n = \pi$.

1. (Stationary distribution) Compute the stationary distributions for each of the depicted Markov chains.



2. (Random walk) Given an undirected graph $G = (V = [n], E)$ without isolated vertices, the *random walk* $W(G)$ is a random process where we always move to a random neighboring vertex.

- Take $G = P_3$ to be a path on 3 vertices. Write the transition matrix of $W(P_3)$.
- For any G , prove that $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ defined by $\pi_i = \frac{\deg(i)}{2|E|}$ is a stationary distribution of $W(G)$.
- Prove that if G is disconnected then the walk is *not* irreducible and π is *not* unique.
- Take $G = C_n$ to be a cycle on n vertices. For what n are the states of $W(G)$ periodic?
- Characterize the graphs G , for which the walk is irreducible and all the states are periodic.

3. (Balance and detailed balance) Consider a Markov chain with states $S = [n]$ and transition matrix $(P_{i,j})_{i,j=1}^n$.

- Prove that $\pi = (\pi_1, \dots, \pi_n)$ is its stationary distribution iff $\sum_i \pi_i = 1$ and for each state $s \in S$ we have

$$\pi_1 \cdot P_{1,s} + \pi_2 \cdot P_{2,s} + \dots + \pi_n \cdot P_{n,s} = \pi_s \cdot (P_{s,1} + P_{s,2} + \dots + P_{s,n}).$$

- Subtract $\pi_s \cdot P_{s,s}$ from both sides and interpret each equation as a sort of “influx-outflux” equality at s .
- Suppose that we found a distribution ρ such that for each pair of states i, j we have $\rho_i \cdot P_{i,j} = \rho_j \cdot P_{j,i}$. Show that ρ is a stationary distribution of the chain. (This condition is called *detailed balance*.)
- Check that for every undirected G , the random walk $W(G)$ has detailed balance.
- Find a Markov chain that does *not* have detailed balance.

4. (Rook) A chess rook is sitting at a corner of an L -shape “staircase” consisting of $3 + 2 + 1 = 6$ cells. In each step, you randomly move the rook, always choosing from all the available rook moves (e.g. in the first step there are 4 possible moves).

- After 100 steps, what is roughly the chance that you are back at the corner cell?
- What if you had a chess king instead of a chess rook?

5. (Uniqueness in the convergence theorem) The first part in the convergence theorem (“ π is unique”) does not require all the assumptions – there exist finite MCs with unique π that are reducible and aperiodic. Guess a characterization of MCs with unique stationary distribution. The following notions might come in handy:

- The vertices of any directed graph can be partitioned into *strongly connected components* (SCCs), that is, equivalence classes of a relation “ u is reachable from v and vice versa”.
- An SCC X is called a *bottom SCC* (BSCC) if there is no directed edge from X to a vertex not in X .

6. (Bonus: Pagerank converges fast) Consider the PageRank algorithm on any (nice) directed graph G with “teleportation” probability ε . Let $\rho = \rho^{(0)}$ be any initial distribution and let π be the stationary distribution. Prove that for any $t \geq 0$ we have $|\rho^{(t)} - \pi|_1 \leq 2 \cdot (1 - \varepsilon)^t$, where $|a - b|_1 = \sum_i |a_i - b_i|$ is the L_1 -distance between lists a and b . (Hint: Prove (and use) that among the trajectories that “used” the teleport, the distributions are equal.)

7. (Bonus: Wolf and sheep) One wolf and $n - 1$ sheep sit at a round dinner table. At each time step, the wolf moves one seat left or one seat right (with probability 50% each) and eats a sheep on that seat (if one is there). The wolf is finished after eating $n - 2$ sheep, so one sheep survives. You are a sheep. The wolf has already taken their seat. Where do you want to sit?

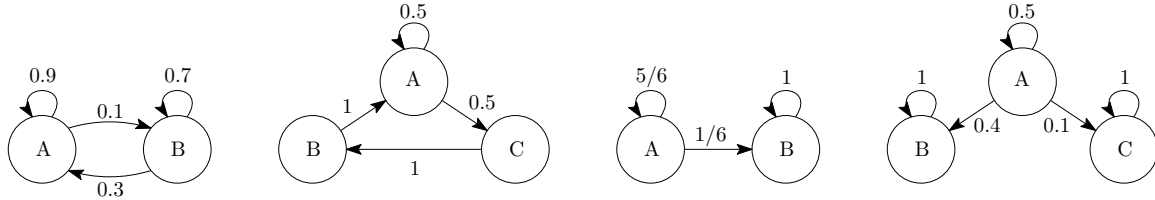
Exercise session 4 – Prob. & Stat. 2

Recall: Given a Markov chain with states $S = [n]$ and a transition matrix $P = (p_{i,j})_{i,j=1}^n$, a *stationary distribution* is any distribution $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ such that $\pi \cdot P = \pi$, that is, π satisfies the system of *balance equations*:

$$\sum_{i \in S} \pi_i \cdot p_{i,j} = \pi_j, \quad \text{for all } j \in S \quad (\dots \text{and also } \sum_{j \in S} \pi_j = 1).$$

Convergence theorem. If the Markov chain is *finite, irreducible, aperiodic* then the stationary distribution is unique and for any distribution $\rho = (\rho_1, \dots, \rho_n)$ we have $\lim_{n \rightarrow \infty} \rho \cdot P^n = \pi$.

1. (Stationary distribution) Compute the stationary distributions for each of the depicted Markov chains.



2. (Random walk) Given an undirected graph $G = (V = [n], E)$ without isolated vertices, the *random walk* $W(G)$ is a random process where we always move to a random neighboring vertex.

- Take $G = P_3$ to be a path on 3 vertices. Write the transition matrix of $W(P_3)$.
- For any G , prove that $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ defined by $\pi_i = \frac{\deg(i)}{2|E|}$ is a stationary distribution of $W(G)$.
- Prove that if G is disconnected then the walk is *not* irreducible and π is *not* unique.
- Take $G = C_n$ to be a cycle on n vertices. For what n are the states of $W(G)$ periodic?
- Characterize the graphs G , for which the walk is irreducible and all the states are periodic.

3. (Balance and detailed balance) Consider a Markov chain with states $S = [n]$ and transition matrix $(P_{i,j})_{i,j=1}^n$.

- Prove that $\pi = (\pi_1, \dots, \pi_n)$ is its stationary distribution iff $\sum_i \pi_i = 1$ and for each state $s \in S$ we have

$$\pi_1 \cdot P_{1,s} + \pi_2 \cdot P_{2,s} + \dots + \pi_n \cdot P_{n,s} = \pi_s \cdot (P_{s,1} + P_{s,2} + \dots + P_{s,n}).$$

- Subtract $\pi_s \cdot P_{s,s}$ from both sides and interpret each equation as a sort of “influx-outflux” equality at s .
- Suppose that we found a distribution ρ such that for each pair of states i, j we have $\rho_i \cdot P_{i,j} = \rho_j \cdot P_{j,i}$. Show that ρ is a stationary distribution of the chain. (This condition is called *detailed balance*.)
- Check that for every undirected G , the random walk $W(G)$ has detailed balance.
- Find a Markov chain that does *not* have detailed balance.

4. (Rook) A chess rook is sitting at a corner of an L -shape “staircase” consisting of $3 + 2 + 1 = 6$ cells. In each step, you randomly move the rook, always choosing from all the available rook moves (e.g. in the first step there are 4 possible moves).

- After 100 steps, what is roughly the chance that you are back at the corner cell?
- What if you had a chess king instead of a chess rook?

5. (Uniqueness in the convergence theorem) The first part in the convergence theorem (“ π is unique”) does not require all the assumptions – there exist finite MCs with unique π that are reducible and aperiodic. Guess a characterization of MCs with unique stationary distribution. The following notions might come in handy:

- The vertices of any directed graph can be partitioned into *strongly connected components* (SCCs), that is, equivalence classes of a relation “ u is reachable from v and vice versa”.
- An SCC X is called a *bottom SCC* (BSCC) if there is no directed edge from X to a vertex not in X .

6. (Bonus: Pagerank converges fast) Consider the PageRank algorithm on any (nice) directed graph G with “teleportation” probability ε . Let $\rho = \rho^{(0)}$ be any initial distribution and let π be the stationary distribution. Prove that for any $t \geq 0$ we have $|\rho^{(t)} - \pi|_1 \leq 2 \cdot (1 - \varepsilon)^t$, where $|a - b|_1 = \sum_i |a_i - b_i|$ is the L_1 -distance between lists a and b . (Hint: Prove (and use) that among the trajectories that “used” the teleport, the distributions are equal.)

7. (Bonus: Wolf and sheep) One wolf and $n - 1$ sheep sit at a round dinner table. At each time step, the wolf moves one seat left or one seat right (with probability 50% each) and eats a sheep on that seat (if one is there). The wolf is finished after eating $n - 2$ sheep, so one sheep survives. You are a sheep. The wolf has already taken their seat. Where do you want to sit?