MVE550 2022 Lecture 4
Dobrow Sections 3.3 - 3.8

Petter Mostad

Chalmers University

November 10, 2022

Review

- We look at discrete time / discrete state space Markov chains $X_0, X_1, \ldots, X_n, \ldots$.
- What happens when $n \to \infty$?
- For *some* Markov chains there is a (unique) limiting distribution $\lim_{n\to\infty} P_{ij}^n = v_j$.
- Which Markov chains have a limiting distribution, and how to compute it? General results so far:
  - There is a limiting distribution when P is *regular*.
  - *If* a limiting distribution exists, there is exactly one *stationary distribution* v (fulfilling $vP = v$), and it is equal to the limiting distribution.

Example from Lecture 3: Random walks on undirected graphs

- An *undirected graph* consists of *nodes* and *undirected edges* connecting them. (An edge may connect a node with itself).
- An undirected graph defines a *random walk Markov chain* by, at every time step, following one of the edges out of a node, with equal probability. (You also need a starting distribution).
- When the graph is finite, show that the vector u is a stationary distribution, where $u_i = deg(i)/S$, where $deg(i)$ is the number of edges going into edge i and S is the sum of all weights, counting weights on edges between different nodes twice.
- Generalization: A *weighted undirected graph* is a graph with a positive weight at any edge between i and j for all i and j.
- Define the Markov chain by choosing the next node with probabilities according to the weights.
- Show that when the graph is finite, the vector u is a stationary distribution, where $u_i = w(i)/S$, where $w(i)$ is the sum of the weights of the edges going into i, and e is the total sum of all weights, counted as above.

Contents of Lecture

- Moving around: Recurrent and transient states; communication classes.
- The limit theorem for finite irreducible Markov chains.
- Periodicity.
- Classification of irreducible Markov chains.
- Time reversibility.
- Canonical decomposition and absorbing chains.

Moving between states

- State j is *accessible* from state i if $(P^n)_{ij} > 0$ for some $n \geq 0$.
- States i and j *communicate* if i is accessible from j and j is accessible from i.
- "Communication" is *transitive*, i.e., if i communicates with j and j communicates with k, then i communicates with k.
- Communication is an *equivalence relation*, subdividing all states into *communication classes*.
- Communication classes can be found for example by drawing transition graphs.
- A Markov chain is *irreducible* if it has exactly one communication class.

Recurrence and transience

- Let T_j be the *first passage time* to state j:
  $T_j = \min\{n > 0 : X_n = j\}$.
- Define f_j as the probability that a chain starting at j will return to j:

$$f_j = P(T_j < \infty \mid X_0 = j)$$

- A state j is *recurrent* if a chain starting at j will eventually revisit j, i.e., if $f_j = 1$.
- A state j is *transient* if a chain starting at j has a positive probability of never revisiting j, i.e., if $f_j < 1$.
- Note: The expected number of visits at j when the chain starts at i is given by $\sum_{n=0}^{\infty}(P^n)_{ij}$.
- j is recurrent if and only if $\sum_{n=0}^{\infty}(P^n)_{jj} = \infty$.
- j is transient if and only if $\sum_{n=0}^{\infty}(P^n)_{jj} < \infty$.

Communication classes

- The states of a communication class are either all recurrent or all transient.
- The states of a finite irreducible Markov chain are all recurrent.
- Note: There are infinite irreducible Markov chains where all states are transient.
- Example: Simple random walk with non-symmetric probabilities.
- If a state is recurrent, only states inside its communication class are accessible from it.
- If no states outside a *finite* communication class are accesible from it, then the class consists of recurrent states.

Finite irreducible Markov chains

- Recall: In a finite irreducible Markov chain, all states are recurrent.
- **Limit Theorem for Finite Irreducible Markov Chains**: Let $\mu_j = E(T_j \mid X_0 = j)$ be the expected return time to j. Then $\mu_j < \infty$ and the vector v with $v_j = 1/\mu_j$ is the **unique stationary distribution**. Furthermore,

$$v_j = \lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} (P^m)_{ij}.$$

- NOTE: All finite regular Markov chains are finite irreducible Markov chains, but not vice versa.
- NOTE: The conclusion is *weaker* than that for finite regular Markov chains: Not all finite irreducible Markov chains have limiting distributions.
- Example: The theorem holds for the chain with transition matrix $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

Extention to infinite irreducible Markov chains

- In a finite irreducible Markov chain, all states are recurrent, and all expected return times μ_j are finite.
- In a Markov chain, states may be recurrent but with infinite expected return times. Such states are called *null recurrent*, while recurrent states with finite expected return times are called *positive recurrent*.
- The previous theorem may be extended to infinite irreducible Markov chains where all states are positive recurrent.

Periodicity

- The *period* of a state i is the greatest common divisor of all $n > 0$ such that $(P^n)_{ii} > 0$.
- All states of a communication class have the same period: See proof in Dobrow.
- A Markov chain is *periodic* if it is irreducible and all states have period greater than 1.
- A Markov chain is *aperiodic* if it is irreducible and all states have period equal to 1.

Classification of (discrete time, discrete state space) irreducible Markov chains

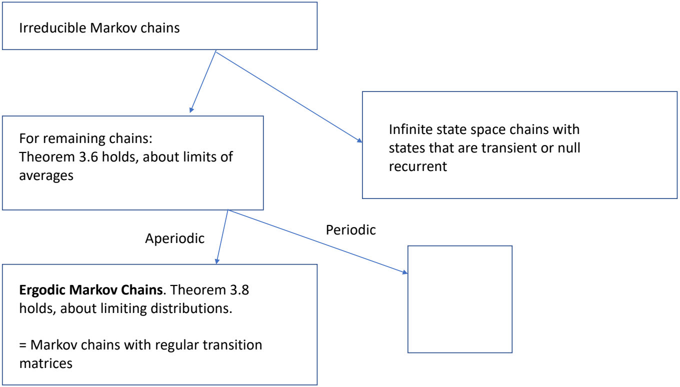

Figure: A subdivision of (discrete time, discrete state space) irreducible Markov chains

Ergodic Markov chains

- A Markov chain is *ergodic* if
  - it is irreducible
  - it is aperiodic
  - all states are positive recurrent (i.e., have finite expected return times). (Always happens if the state space is finite).
- **Fundamental Limit Theorem for Ergodic Markov Chains**: There exists a unique positive stationary distribution v which is the limiting distribution of the chain.
- We can also show that a finite Markov chain is ergodic if and only if it its transition matrix is regular.

Time reversibility

Let P be the transition matrix of an irreducible Markov chain with stationary distribution v.

- The chain is "time reversible" if, when running from its stationary distribution, it looks the same moving foreard as backwards, i.e., $\pi(X_k = i, X_{k+1} = j) = \pi(X_{k+1} = i, X_k = j)$.
- This may also be written as $v_i P_{ij} = v_j P_{ji}$ for all i, j: The *detailed balance condition*.
- Show: If x is a probability vector satisfying $x_i P_{ij} = x_j P_{ji}$ for all i, j, then necessarily x is the stationary distribution, so that $x = v$.
- Show: If a Markov chain is defined as a random walk on a weighted undirected graph, then it is time reversible.
- Show: If a finite Markov chain is time reversible, it can be represented as a random walk on a weighted undirected graph.

Canonical decomposition (assume a finite state space)

- The states of a Markov chain can be subdivided into communication classes, each consisting only of transient or recurrent states.
- Let T denote the union of all communication classes with transient states. Let remaining communication classes be $R_1, R_2, \dots, R_m$.
- Each R_i must necessarily be *closed* in the sense that no states outside R_i are accessible from R_i.
- Ordering states according to T, R_1, ..., R_m, the transition matrix can be written

$$P = \begin{bmatrix} * & * & \cdots & * \\ 0 & P_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_m \end{bmatrix}.$$

- We get

$$P^n = \begin{bmatrix} * & * & \cdots & * \\ 0 & P_1^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_m^n \end{bmatrix}$$

and can take the limits of each P_i^n, if they exist.

Absorbing chains

- State i is *absorbing* if $P_{ii} = 1$.
- A Markov chain is *absorbing* if it has at least one absorbing state.
- By reordering the states, the transition matrix for an absorbing chain can be written in block form

$$P = \begin{bmatrix} Q & R \\ \mathbf{0} & I \end{bmatrix}.$$

  where I is the identity matrix, $\mathbf{0}$ is a matrix of zeros, and Q corresponds to transient states.
- We can prove by induction that

$$P^n = \begin{bmatrix} Q^n & \left(I + Q + Q^2 + \cdots + Q^{n-1}\right) R \\ \mathbf{0} & I \end{bmatrix}.$$

- Taking the limit and using $\lim_{n\to\infty} Q^n = 0$ we get

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \mathbf{0} & (I-Q)^{-1} R \\ \mathbf{0} & I \end{bmatrix} = \begin{bmatrix} \mathbf{0} & FR \\ \mathbf{0} & I \end{bmatrix}.$$

- $F = (I - Q)^{-1} = \lim_{n\to\infty} I + Q + \cdots + Q^n$ is called the *fundamental matrix*.

Absorbing chains, cont

- The probability to be absorbed in a particular absorbing state given a start in a transient state is given by the entries of FR.
- Further, the expected number of visits in state j for a chain that starts in the transient state i is given by F_{ij}. (See proof in Dobrow).
- Thus, the expected number of steps until absorbtion is given by the vector $F\mathbf{1}^t$.
- Note: Given an irreducible Markov chain. To compute the expected number of steps needed to go from state i to the first visit to state j, one can change the chain into one where state j is absorbing, and compute the expected number of steps until absorbtion using the theory above.

Example: First detection of a particular sequence

- Assume you want to find the expected number of steps until you detect HTTH in a sequence of fair coin flips.
- Build a Markov chain where the states indicate how far into the sequence you have read so far. Make the state HTTH absorbing.
- Find the transition matrix in canonical block form.