Lawler (1995)
Chapter 1
Finite Markov Chains

1.1 Definitions and Examples

- Consider a discrete time stochastic process X_n, $n = 0,1,2,\ldots$, with state space $S = \{1,\ldots,N\}$ (or $\{0,\ldots,N-1\}$)
- *Definition*:
  - A time-homogeneous **Markov chain** is a process such that

$$P\{X_n = i_n \mid X_0 = i_0,\ldots,X_{n-1} = i_{n-1}\} = P\{X_n = i_n \mid X_{n-1} = i_{n-1}\} = p(i_{n-1},i_n)$$

for some function $p : S \times S \to [0,1]$. Matrix $P = (p(i,j);\ i,j = 1,\ldots,N)$ is called the **transition matrix** of the Markov chain.

- *Notes*:
  - Given an initial distribution $\phi(i) = \mathrm{P}\{X_0 = i\}$, we have

$$P\{X_0 = i_0,\ldots,X_n = i_n\} = \phi(i_0)\,p(i_0,i_1)\cdots p(i_{n-1},i_n) \qquad (1.3)$$

  - Transition matrix P is a **stochastic matrix**, i.e.

$$0 \le p(i,j) \le 1,\ \forall i,j \qquad \sum_{j=1}^{N} p(i,j) = 1,\ \forall i$$

  - Any stochastic matrix is the transition matrix for some Markov chain

1.1 Definitions and Examples

- *Definition*:
  - Define the (conditional) ***n*-step probabilities** $p_n(i,j)$ by

$$p_n(i,j) = P\{X_n = j \mid X_0 = i\} = P\{X_{n+k} = j \mid X_k = i\}$$

- *Proposition*:
  - **Chapman-Kolmogorov equation:**

$$p_{m+n}(i,j) = \sum_{k \in S} p_m(i,k) p_n(k,j)$$

- *Idea of the proof*:
  - By conditioning at time m
- *Corollary*:
  - $p_n(i,j)$ is the (i,j)-entry in the matrix P^n,

$$p_n(i,j) = [P^n](i,j)$$

1.1 Definitions and Examples

- *Definition*:
  - Given an initial distribution ϕ, the unconditional n-step probabilities are

$$\phi_n(j) = P\{X_n = j\} = \sum_{i \in S} \phi(i) p_n(i,j)$$

- *Note*:
  - The same in the matrix form:

$$\bar{\phi}_n = \bar{\phi} P^n$$

1.2 Long-Range Behaviour and Invariant Probability

- *Definition*:
  - A probability vector π is called an **invariant distribution** for P if

$$\pi = \pi P$$

- *Notes*:
  - The system of linear equations given above for the determination of π are called **Global Balance Equations** (GBE):

$$\pi(y) = \sum_{x \in S} \pi(x) p(x, y), \quad y \in S \qquad \text{(GBE)}$$

  - Requiring that π is a probability vector (= distribution) is the so called **Normalizing Condition** (N):

$$\sum_{x \in S} \pi(x) = 1 \qquad \text{(N)}$$

1.2 Long-Range Behaviour and Invariant Probability

- *Proposition*:
  - Starting with an invariant distribution π as the initial distribution ϕ, we have, for all n,

$$\bar{\phi}_n = \pi P^n = (\pi P) P^{n-1} = \pi P^{n-1} = \cdots = \pi$$

- *Note*:
  - In fact, the chain is then stationary with **stationary distribution** π

1.2 Long-Range Behaviour and Invariant Probability

- *Proposition*:
  - Suppose π is a **limiting distribution**, i.e. for some initial distribution ϕ, we have

$$\bar{\pi} = \lim_{n\to\infty} \bar{\phi}P^n$$

  - Then it is also an invariant distribution,

$$\bar{\pi} = \lim_{n\to\infty} \bar{\phi}P^{n+1} = (\lim_{n\to\infty} \bar{\phi}P^n)P = \bar{\pi}P$$

1.3 Classification of States

- *Definition*:
  - Two states i and j **communicate** ($i \leftrightarrow j$) if there exist m and n such that $p_m(i,j) > 0$ and $p_n(j,i) > 0$
- *Notes*:
  - Relation $\leftrightarrow$ is an equivalence relation.
  - Equivalence classes are called **communication classes**
- *Definition*:
  - Markov chain is called **irreducible** if there is only one communication class

1.3 Classification of States

- *Definition*:
  - Communication class C is **recurrent** if and only if for all $i \in C$,

$$\sum_{j \in C} p(i,j) = 1$$

and **transient** if and only if for some $i \in C$,

$$\sum_{j \in C} p(i,j) < 1$$

- *Notes*:
  - A transient class is eventually left, but a recurrent class never
  - If there is only one class (i.e. the chain is irreducible), it must be recurrent

1.3 Classification of States

- *Proposition*:
  - Assume that P_R is the part of P related to a recurrent class P.
  - If there is n such that $(P_R)^n$ has all entries strictly positive, then there is a distribution π_R defined on R such that $\pi_R(i) > 0$ for all $i \in R$ and

$$\lim_{n \to \infty} (P_R)^n = \overline{1}\pi_R$$

- *Notes*:
  - As a limit, π_R is unique
  - As a limiting ditribution, π_R is invariant with respect to P_R
  - For any initial distribution ϕ_R defined on R,

$$\lim_{n \to \infty} \overline{\phi}_R (P_R)^n = \overline{\phi}_R \overline{1} \overline{\pi}_R = \overline{\pi}_R$$

  - There cannot be any other invariant distributions
- *Proposition*:
  - Assume that P_T is the part of P related to a transient class. Then

$$\lim_{n \to \infty} (P_T)^n = 0$$

1.3 Classification of States

- *Definition*:
  - Consider an irreducible Markov chain. The **period** $d(i)$ of state i is the greatest common divisor of the set

$$J_i = \{n \geq 0 \mid p_n(i,i) > 0\}$$

- *Proposition*:
  - All the states of an irreducible Markov chain have the same period $d \equiv d(i)$
- *Definition*:
  - An irreducible Markov chain is called **aperiodic** if $d = 1$
- *Notes*:
  - A self-transition (even a single one) makes an irreducible chain aperiodic
  - However, there are also aperiodic chains without any self-transitions
  - A pairwise-transition ($p(i,j) > 0$ and $p(j,i) > 0$) implies that $d \leq 2$:

$$p_2(i,i) \geq p(i,j)p(j,i) > 0$$

  - An irreducible Markov chain is aperiodic if and only if there is n such that P^n has all entries strictly positive

1.3 Classification of States

- *Theorem*:
  - An irreducible **aperiodic** Markov chain has a unique invariant distribution π
  - If ϕ is any initial distribution, then

$$\lim_{n \to \infty} \bar{\phi} P^n = \bar{\pi}$$

  - Moreover, for each i,

$$\pi(i) > 0$$

- *Notes*:
  - For any initial probability vector ϕ and any state $j \in S$:

$$\lim_{n \to \infty} \phi_n(j) = \pi(j)$$

  - For any $i,j \in S$:

$$\lim_{n \to \infty} p_n(i,j) = \pi(j)$$

1.3 Classification of States

- *Theorem*:
  - An irreducible **periodic** Markov chain with period d has a unique invariant distribution π
  - If ϕ is any initial distribution, then

$$\lim_{n\to\infty} \bar{\phi} P^n$$

does not exist, but

$$\lim_{n\to\infty} \frac{1}{d}(\bar{\phi} P^{n+1} + \cdots + \bar{\phi} P^{n+d}) = \bar{\pi}$$

  - Moreover, for each i,

$$\pi(i) > 0$$

- *Idea of the proof*:
  - Use the previous result by defining a new, aperiodic chain as follows:

$$\hat{p}(i,j) = \frac{1}{2}(\delta(i,j) + p(i,j))$$

where $\delta(i,j) = 1$ if $i = j$ and 0 otherwise. The same invariant distribution!

1.3 Classification of States

- *Theorem*:
  - Consider a Markov chain with recurrent, aperiodic classes $R_1,\ldots,R_r$ and transient classes $T_1,\ldots,T_s$. Let π^k denote the unique invariant distribution of class R_k.
  - Then any linear combination (with weights summing to 1) of π^ks is an invariant distribution for the chain
- *Idea of the proof*:
  - Use block matrices to verify the result

1.3 Classification of States

- *Theorem*:
  - Consider a Markov chain with recurrent, aperiodic classes $R_1,\ldots,R_r$ and transient classes $T_1,\ldots,T_s$. Let π^k denote the unique invariant distribution of class R_k. Let $\alpha_k(i)$ denote the probability that the chain starting in a state i eventually ends up in the recurrent class R_k.
  - Then, for any state $i \in S$ and $j \in R_k$,

$$\lim_{n\to\infty} p_n(i,j) = \alpha_k(i)\pi^k(j)$$

- *Notes*:
  - For any recurrent state $i \in R_k$, we have $\alpha_k(i) = 1$
  - If ϕ is an initial distribution, then

$$\lim_{n\to\infty} \bar{\phi}P^n$$

exists but depends on ϕ so that for any state $j \in R_k$

$$\lim_{n\to\infty} \phi_n(j) = \sum_{i\in S}\phi(i)\alpha_k(i)\pi^k(j)$$

1.4 Return Times

- Let X_n be an irreducible Markov chain with invariant distribution π
  - Assume that $X_0 = i$ and denote the first time after 0 that the Markov chain is in state i by T,

$$T = \min\{n \geq 1 \mid X_n = i\}$$

- *Proposition*:

$$E[T] = \frac{1}{\pi(i)} < \infty$$

- *Idea of the proof*:
  - By a renewal argument applying Blackwell.s Theorem. Consecutive visits to state i constitute a renewal process in discrete time.

1.4 Return Times

- Renewal theory in discrete time
  - Interarrival times T_n i.i.d. with period d

$$d = \max\{k \geq 1 \mid \sum_{n=1}^{\infty} P\{T = nk\} = 1\}$$

  - Define

$$I_n = I\{\text{arrival at time } n\}$$

- *Elementary Renewal Theorem*:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P\{I_m = 1\} = \frac{1}{E[T]}$$

- *Blackwell's Theorem*:

$$\lim_{n\to\infty} P\{I_{nd} = 1\} = \frac{d}{E[T]}$$

- *Corollary*:

$$\lim_{n\to\infty} \frac{1}{d}\sum_{m=1}^{d} P\{I_{n+m} = 1\} = \frac{1}{E[T]}$$

1.5 Transient States

- Consider a Markov chain X_n with recurrent classes $R_1,\ldots,R_r$ and transient classes $T_1,\ldots,T_s$
  - Denote the part of the transition matrix P that relates to the transient states by Q, and reorder the states so that

$$P = \begin{bmatrix} \widetilde{P} & 0 \\ S & Q \end{bmatrix}$$

  - Matrix $I - Q$ is invertible and we may define the matrix

$$M = (I - Q)^{-1} = I + Q + Q^2 + \ldots$$

  - Let i be a transient state and denote the total number of visits to i by Y_i,

$$Y_i = \sum_{n=0}^{\infty} I\{X_n = i\}$$

1.5 Transient States

- *Proposition*:
  - For any transient states i,j, we have

$$E[Y_i \mid X_0 = j] = [M](j,i)$$

- *Proof*:

$$\begin{aligned} E[Y_i \mid X_0 = j] &= E[\textstyle\sum_{n=0}^{\infty} I\{X_n = i\} \mid X_0 = j] \\ &= \textstyle\sum_{n=0}^{\infty} P\{X_n = i \mid X_0 = j\} \\ &= \textstyle\sum_{n=0}^{\infty} p_n(j,i) \\ &= \left[\textstyle\sum_{n=0}^{\infty} P^n\right](j,i) \\ &= \left[\textstyle\sum_{n=0}^{\infty} Q^n\right](j,i) \\ &= \left[(I-Q)^{-1}\right](j,i) \\ &= [M](j,i) \end{aligned}$$

1.5 Transient States

- Let X_n be an irreducible Markov chain with transition matrix P
  - Assume that $X_0 = j$ and denote the first time after 0 that the Markov chain is in state i by T_i,

$$T_i = \min\{n \geq 1 \mid X_n = i\}$$

  - Without loss of generality, we may assume that $i = 1$,

$$P = \begin{bmatrix} p(i,i) & R \\ S & Q \end{bmatrix}$$

- Consider then the modified Markov chain with transition matrix

$$\widetilde{P} = \begin{bmatrix} 1 & 0 \\ S & Q \end{bmatrix}$$

  - Now i is an absorbing state and all the other states are transient
  - Let

$$M = (I-Q)^{-1}$$

1.5 Transient States

- *Proposition*:
  - For any $j \neq i$, we have

$$E[T_i \mid X_0 = j] = \sum_{k \neq i} [M](j,k) < \infty$$

- *Proof*:
  - For any $k \neq i$, denote the total number of visits to k by $\widetilde{Y}_k$,

$$\widetilde{Y}_k = \sum_{n=0}^{\infty} I\{\widetilde{X}_n = k\}$$

  - Now

$$T_i = \sum_{k \neq i} \widetilde{Y}_k$$

  - Thus, by the previous proposition,

$$\begin{aligned} E[T_i \mid X_0 = j] &= E[\textstyle\sum_{k \neq i} \widetilde{Y}_k \mid \widetilde{X}_0 = j] \\ &= \textstyle\sum_{k \neq i} [M](j,k) \end{aligned}$$

1.5 Transient States

- Consider a Markov chain X_n with absorbing states $r_1,\ldots,r_k$ and transient states $t_1,\ldots,t_s$
  - By a suitable reordering of the states, the transition matrix P is as follows:

$$P = \begin{bmatrix} I & 0 \\ S & Q \end{bmatrix}$$

  - As before, let

$$M = (I - Q)^{-1}$$

  - Let $\alpha(t_i,r_j)$ denote the probability that the chain starting at t_i eventually ends up in recurrent state r_j,

$$\alpha(t_i, r_j) = P\{\lim_{n \to \infty} X_n = r_j \mid X_0 = t_i\}$$

  - Define an $s \times k$ matrix by $A = (\alpha(t_i,r_j);\ i = 1,\ldots,s,\ j = 1,\ldots,k)$

1.5 Transient States

- *Proposition*:
  - For any i,j, we have

$$\alpha(t_i, r_j) = [MS](i, j)$$

- *Proof*:
  - Clearly

$$\alpha(t_i, r_j) = p(t_i, r_j) + \sum_{k=1}^{s} p(t_i, t_k)\alpha(t_k, r_j)$$

  - The same in the matrix form:

$$A = S + QA \quad \Rightarrow \quad (I - Q)A = S \quad \Rightarrow \quad A = (I - Q)^{-1} S = MS$$

The End of Chapter 1