Synopsis of Chapter 4, Ross, Markov Chains

Definition Markov chain: A sequence of discrete random variables $X_0, X_1, X_2, \ldots$ is called a Markov chain (by Ross) if

(i) $\mathrm{P}(X_n = j | X_0, X_1, X_2, \ldots, X_{n-1} = i) = \mathrm{P}(X_n = j | X_{n-1} = i)$

(ii) $\mathrm{P}(X_n = j | X_{n-1} = i) = \mathrm{P}_{ij}$ for all n

In some texts sequences satisfying (i) and (ii) would be called an homogeneous Markov chain and sequences only satisfying (i) would be called an inhomogeneous Markov chain. In the following we will use the definition of Ross.

We refer to P_{ij} as (one-step) transition probabilities. The matrix $\mathbf{P} = [\mathrm{P}_{ij}]$ of all transition probabilities is called the **transition matrix**. The row sums of $\mathbf{P}$ will always be equal to one.

Chapman-Kolmogorov equations

The n**-step transition probabilities** are defined as $\mathrm{P}^n_{ij} = \mathrm{P}(X_n = j | X_0 = i)$ and $\mathbf{P}^n = [\mathrm{P}^n_{ij}]$ is the n-step transition matrix.

The Chapman-Kolmogorov equations states

$$\mathrm{P}^{n+m}_{ij} = \sum_k \mathrm{P}^n_{ik} \mathrm{P}^m_{kj}$$

where the sum is over all possible k. Using matrix notation the Chapman-Kolomogorov equations can be written as $\mathbf{P}^{n+m} = \mathbf{P}^n \mathbf{P}^m$.

Classification of states

- A state j is said to be **accessible** from i if there exists an n such that $\mathrm{P}^n_{ij} > 0$.
- Two states i and j are said to **communicate** if j is accessible form i and i is accessible from j.
- A **class** of states is defined as the set of all communicating states.
- An **irreducible** Markov chain has only one class.
- An **absorbing** state i has $\mathrm{P}_{ii} = 1$. Absorbing states constitute one class.

- A **closed class** is a class $\mathcal{C}$ that the Markov chain can not leave, thus $\mathrm{P}_{ij} = 0$ if $i \in \mathcal{C}$ and j not in $\mathcal{C}$.
- A **transient** state/class is a state/class for which the probability of ever returning is less than one.
- A **recurrent** state/class is state/class for which the probability of ever returning equals one.
- A **positive recurrent** state/class is a state/class for which the expected return time to a state i within in the class is finite. A null-recurrent state/class is a recurrent state/class with infinite expected return time.
- The **period** of a state/class is the largest common divisor of the n such that $\mathrm{P}_{ii}^n > 0$. An **aperiodic** Markov chain has period equal to one.
- An **ergodic** Markov chain is an aperiodic and positive recurrent Markov chain.

"One-step" analysis

For several problems it is useful to (i) consider all possible outcomes of X_1 and then (ii) to condition on and sum over these possible outcomes. Such an approach is by some (not Ross) referred to as "one-step" analysis. In this section it is assumed that the state space is finite.

- Let $\mathcal{C}_0$ be the set of transient states and $\mathcal{C}_1$ the union of all closed sets. Then with T equal to the time until entering $\mathcal{C}_1$ and $\mu_i = \mathrm{E}[T|X_0 = i]$ we find μ_i solving the set of linear equations given by
$$\mu_i = \sum_{j \in \mathcal{C}_0} \mathrm{P}_{ij}(\mu_j + 1) + \sum_{j \in \mathcal{C}_1} \mathrm{P}_{ij}$$
- If there are two closed sets $\mathcal{C}_1$ and $\mathcal{C}_2$ and transient states $i \in \mathcal{C}_0$ we have that the probabilities q_i that the process eventually enters $\mathcal{C}_1$ given that $X_0 = i$ are determined solving the set of equations
$$q_i = \sum_{j \in \mathcal{C}_0} \mathrm{P}_{ij} q_j + \sum_{j \in \mathcal{C}_1} \mathrm{P}_{ij}$$
- Let s_{ij} equal the expected number of visits to j where both i and j are transients states and $\mathcal{C}_0$ is the set of all transient states. Then the s_{ij} can be found solving the linear equations
$$s_{ij} = \delta_{ij} + \sum_{k \in \mathcal{C}_0} \mathrm{P}_{ik} s_{kj}$$
where δ_{ij} equals one when $i = j$ and zero otherwise.

Stationary distribution

Assume an irreducible and ergodic Markov chain. Then we have

$$\mathrm{P}^n_{ij} \to \pi_j$$

when $n \to \infty$ where the π_j are uniquely determined by $\sum_j \pi_j = 1$ and

$$\pi_j = \sum_i \pi_i \mathrm{P}_{ij} \text{ for all } j.$$

The π_j can be interpreted as

- limiting probabilities ($\mathrm{P}^n_{ij} \to \pi_j$)
- the stationary distribution (i.e. if $\mathrm{P}(X_0 = j) = \pi_j$ then for all n we also have $\mathrm{P}(X_n = j) = \pi_j$)
- the limit of the proportion of visits to j
- the inverse of the expected return time to j (i.e. if m_{jj} is the expected number of transitions starting in j until the chain returns to j then we have $\pi_j = \frac{1}{m_{jj}}$)

Markov chain Monte Carlo methods

Assume an irreducible and ergodic Markov chain $X_0, X_1, X_2, \ldots$ with stationary distribution $\mathrm{P}^n_{ij} \to \pi_j$. Let $\theta = \sum_i h(i)\pi_i = \mathrm{E}[h(X)]$ when X has the stationary distribution. Then

$$\frac{1}{n}\sum_{i=0}^{n} h(X_i) \to \theta \text{ when } n \to \infty$$

For an irreducible and ergodic Markov chain satisfying $\pi_i \mathrm{P}_{ij} = \pi_j \mathrm{P}_{ji}$ we have

(i) the $\pi_j = \lim_{n\to\infty} \mathrm{P}(X_n = j)$ are stationary probabilities of the chain

(ii) the chain is reversible if it , i.e. $\mathrm{P}(X_n = i, X_{n+1} = j) = \mathrm{P}(X_n = j, X_{n+1} = i)$ when the chain has attained the stationary distribution

With the Hastings-Metropolis algorithm one constructs a reversible Markov chain by

- Given $X_n = i$ sample a candidate value Y_{n+1} with conditional probabilites $q_{ij} = \mathrm{P}(Y_{n+1} = j | X_n = i)$. The q_{ij} correspond to an irreducible Markov chain.
- The candidate value is accepted, i.e. $X_{n+1} = Y_{n+1} = j$, with probability

  $$\alpha_{ij} = \min(1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}})$$

  and is rejected, i.e. $X_{n+1} = X_n$, with probability $1 - \alpha_{ij}$

Branching processes

Assume that X_n is the size of the population in generation n where each individual in each generation breeds Z individuals independently with probability $P_j = \mathrm{P}(Z = j), j = 0, 1, 2, \ldots$ and expectation $\mu = \sum_{j=0}^{\infty} j P_j$. Then

- the population dies out with probability one if $\mu < 1$ and also for $\mu = 1$ (if $P_0 > 0$)
- if $X_0 = 1$ and $\mu > 1$ the probability that population dies out π_0 can be found solving the equation $\pi_0 = \sum_{j=0}^{\infty} \pi_0^j P_j$