5

Applications

Applications of Markov chains arise in many different areas. Some have already appeared to illustrate the theory, from games of chance to the evolution of populations, from calculating the fair price for a random reward to calculating the probability that an absent-minded professor is caught without an umbrella. In a real-world problem involving random processes you should always look for Markov chains. They are often easy to spot. Once a Markov chain is identified, there is a qualitative theory which limits the sorts of behaviour that can occur – we know, for example, that every state is either recurrent or transient. There are also good computational methods – for hitting probabilities and expected rewards, and for long-run behaviour via invariant distributions.

In this chapter we shall look at five areas of application in detail: biological models, queueing models, resource management models, Markov decision processes and Markov chain Monte Carlo. In each case our aim is to provide an introduction rather than a systematic account or survey of the field. References to books for further reading are given in each section.

5.1 Markov chains in biology

Randomness is often an appropriate model for systems of high complexity, such as are often found in biology. We have already illustrated some aspects of the theory by simple models with a biological interpretation. See Example 1.1.5 (virus), Exercise 1.1.6 (octopus), Example 1.3.4 (birth-and-death chain) and Exercise 2.5.1 (bacteria). We are now going to give

some more examples where Markov chains have been used to model biological processes, in the study of population growth, epidemics and genetic inheritance. It should be recognised from the start that these models are simplified and somewhat stylized in order to make them mathematically tractable. Nevertheless, by providing quantitative understanding of various phenomena they can provide a useful contribution to science.

Example 5.1.1 (Branching processes)

The original branching process was considered by Galton and Watson in the 1870s while seeking a quantitative explanation for the phenomenon of the disappearance of family names, even in a growing population. Under the assumption that each male in a given family had a probability p_k of having k sons, they wished to determine the probability that after n generations an individual had no male descendents. The solution to this problem is explained below.

The basic branching process model has many applications to problems of population growth, and also to the study of chain reactions in chemistry and nuclear fission. Suppose at time $n = 0$ there is one individual, who dies and is replaced at time $n = 1$ by a random number of offspring N. Suppose, next, that these offspring also die and are themselves replaced at time $n = 2$, each independently, by a random number of further offspring, having the same distribution as N, and so on. We can construct the process by taking for each $n \in \mathbb{N}$ a sequence of independent random variables $(N_k^n)_{k\in\mathbb{N}}$, each with the same distribution as N, by setting $X_0 = 1$ and defining inductively, for $n \geq 1$

$$X_n = N_1^n + \ldots + N_{X_{n-1}}^n.$$

Then X_n gives the size of the population in the nth generation. The process $(X_n)_{n\geq 0}$ is a Markov chain on $I = \{0, 1, 2, \ldots\}$ with absorbing state 0. The case where $\mathbb{P}(N = 1) = 1$ is trivial so we exclude it. We have

$$\mathbb{P}(X_n = 0 \mid X_{n-1} = i) = \mathbb{P}(N = 0)^i$$

so if $\mathbb{P}(N = 0) > 0$ then i leads to 0, and every state $i \geq 1$ is transient. If $\mathbb{P}(N = 0) = 0$ then $\mathbb{P}(N \geq 2) > 0$, so for $i \geq 1$, i leads to j for some $j > i$, and j does not lead to i, hence i is transient in any case. We deduce that with probability 1 either $X_n = 0$ for some n or $X_n \to \infty$ as $n \to \infty$.

Further information on $(X_n)_{n\geq 0}$ is obtained by exploiting the branching structure. Consider the probability generating function

$$\phi(t) = \mathbb{E}(t^N) = \sum_{k=0}^{\infty} t^k \mathbb{P}(N = k),$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

defined for $0 \le t \le 1$. Conditional on $X_{n-1} = k$ we have

$$X_n = N_1^n + \ldots + N_k^n$$

so

$$\mathbb{E}(t^{X_n} \mid X_{n-1} = k) = \mathbb{E}(t^{N_1^n + \ldots + N_k^n}) = \phi(t)^k$$

and so

$$\mathbb{E}(t^{X_n}) = \sum_{k=0}^{\infty} \mathbb{E}(t^{X_n} \mid X_{n-1} = k)\mathbb{P}(X_{n-1} = k) = \mathbb{E}\big(\phi(t)^{X_{n-1}}\big).$$

Hence, by induction, we find that $\mathbb{E}(t^{X_n}) = \phi^{(n)}(t)$, where $\phi^{(n)}$ is the n-fold composition $\phi \circ \ldots \circ \phi$. In principle, this gives the entire distribution of X_n, though $\phi^{(n)}$ may be a rather complicated function. Some quantities are easily deduced: we have

$$\mathbb{E}(X_n) = \lim_{t \uparrow 1} \frac{d}{dt}\, \mathbb{E}(t^{X_n}) = \lim_{t \uparrow 1} \frac{d}{dt} \phi^{(n)}(t) = \big(\lim_{t \uparrow 1} \phi'(t)\big)^n = \mu^n,$$

where $\mu = \mathbb{E}(N)$; also

$$\mathbb{P}(X_n = 0) = \phi^{(n)}(0)$$

so, since 0 is absorbing, we have

$$q = \mathbb{P}(X_n = 0 \text{ for some } n) = \lim_{n \to \infty} \phi^{(n)}(0).$$

Now $\phi(t)$ is a convex function with $\phi(1) = 1$. Let us set $r = \inf\{t \in [0,1] : \phi(t) = t\}$, then $\phi(r) = r$ by continuity. Since ϕ is increasing and $0 \le r$, we have $\phi(0) \le r$ and, by induction, $\phi^{(n)}(0) \le r$ for all n, hence $q \le r$. On the other hand

$$q = \lim_{n \to \infty} \phi^{(n+1)}(0) = \lim_{n \to \infty} \phi(\phi^{(n)}(0)) = \phi(q)$$

so also $q \ge r$. Hence $q = r$. If $\phi'(1) > 1$ then we must have $q < 1$, and if $\phi'(1) \le 1$ then since either $\phi'' = 0$ or $\phi'' > 0$ everywhere in $[0,1)$ we must have $q = 1$. We have shown that the population survives with positive probability if and only if $\mu > 1$, where μ is the mean of the offspring distribution.

There is a nice connection between branching processes and random walks. Suppose that in each generation we replace individuals by their offspring one at a time, so if $X_n = k$ then it takes k steps to obtain X_{n+1}.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

The population size then performs a random walk $(Y_m)_{m\geq 0}$ with step distribution $N-1$. Define stopping times $T_0 = 0$ and, for $n \geq 0$

$$T_{n+1} = T_n + Y_{T_n}.$$

Observe that $X_n = Y_{T_n}$ for all n, and since $(Y_m)_{m\geq 0}$ jumps down by at most 1 each time, $(X_n)_{n\geq 0}$ hits 0 if and only if $(Y_m)_{m\geq 0}$ hits 0. Moreover we can use the strong Markov property and a variation of the argument of Example 1.4.3 to see that, if

$$q_i = \mathbb{P}(Y_m = 0 \text{ for some } m \mid Y_0 = i)$$

then $q_i = q_1^i$ for all i and so

$$q_1 = \mathbb{P}(N = 0) + \sum_{k=1}^{\infty} q_1^i \mathbb{P}(N = i) = \phi(q_1).$$

Now each non-negative solution of this equation provides a non-negative solution of the hitting probability equations, so we deduce that q_1 is the smallest non-negative root of the equation $q = \phi(q)$, in agreement with the generating function approach.

The classic work in this area is *The Theory of Branching Processes* by T. E. Harris (Dover, New York, 1989).

Example 5.1.2 (Epidemics)

Many infectious diseases persist at a low intensity in a population for long periods. Occasionally a large number of cases arise together and form an epidemic. This behaviour is to some extent explained by the observation that the presence of a large number of infected individuals increases the risk to the rest of the population. The decline of an epidemic can also be explained by the eventual decline in the number of individuals susceptible to infection, as infectives either die or recover and are then resistant to further infection. However, these naive explanations leave unanswered many quantitative questions that are important in predicting the behaviour of epidemics.

In an idealized population we might suppose that all pairs of individuals make contact randomly and independently at a common rate, whether infected or not. For an idealized disease we might suppose that on contact with an infective, individuals themselves become infective and remain so for an exponential random time, after which they either die or recover. These two possibilities have identical consequences for the progress of the epidemic. This idealized model is obviously unrealistic, but it is the simplest mathematical model to incorporate the basic features of an epidemic.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

We denote the number of susceptibles by S_t and the number of infectives by I_t. In the idealized model, $X_t = (S_t, I_t)$ performs a Markov chain on $(\mathbb{Z}^+)^2$ with transition rates

$$q_{(s,i)(s-i,i+1)} = \lambda si, \quad q_{(s,i)(s,i-1)} = \mu i$$

for some $\lambda, \mu \in (0, \infty)$. Since $S_t + I_t$ does not increase, we effectively have a finite state-space. The states $(s, 0)$ for $s \in \mathbb{Z}^+$ are all absorbing and all the other states are transient; indeed all the communicating classes are singletons. The epidemic must therefore eventually die out, and the absorption probabilities give the distribution of the number of susceptibles who escape infection. We can calculate these probabilities explicitly when $S_0 + I_0$ is small.

Of greater concern is the behaviour of an epidemic in a large population, of size N, say. Let us consider the proportions $s_t^N = S_t/N$ and $i_t^N = I_t/N$ and suppose that $\lambda = \nu/N$, where ν is independent of N. Consider now a sequence of models as $N \to \infty$ and choose $s_0^N \to s_0$ and $i_0^N \to i_0$. It can be shown that as $N \to \infty$ the process (s_t^N, i_t^N) converges to the solution (s_t, i_t) of the differential equations

$$(d/dt)s_t = -\nu s_t i_t$$
$$(d/dt)i_t = \nu s_t i_t - \mu i_t$$

starting from (s_0, i_0). Here convergence means that $\mathbb{E}[|(s_t^N, i_t^N) - (s_t, i_t)|] \to 0$ for all $t \geq 0$. We will not prove this result, but will give an example of another easier asymptotic calculation.

Consider the case where $S_0 = N-1$, $I_0 = 1$, $\lambda = 1/N$ and $\mu = 0$. This has the following interpretation: a rumour is begun by a single individual who tells it to everyone she meets; they in turn pass the rumour on to everyone they meet. We assume that each individual meets another randomly at the jump times of a Poisson process of rate 1. How long does it take until everyone knows the rumour? If i people know the rumour, then $N - i$ do not, and the rate at which the rumour is passed on is

$$q_i = i(N - i)/N.$$

The expected time until everyone knows the rumour is then

$$\sum_{i=1}^{N-1} q_i^{-1} = \sum_{i=1}^{N-1} \frac{N}{i(N-i)} = \sum_{i=1}^{N-1} \left(\frac{1}{i} + \frac{1}{N-i} \right) = 2 \sum_{i=1}^{N-1} \frac{1}{i} \sim 2 \log N$$

as $N \to \infty$. This is not a limit as above but, rather, an asymptotic equivalence. The fact that the expected time grows with N is related to the fact

that we do not scale I_0 with N: when the rumour is known by very few or by almost all, the proportion of 'infectives' changes very slowly.

The final two examples come from population genetics. They represent an attempt to understand quantitatively the consequences of randomness in genetic inheritance. The randomness here might derive from the choice of reproducing individual, in sexual reproduction the choice of partner, or the choice of parents' alleles retained by their offspring. (The word *gene* refers to a particular chromosomal locus; the varieties of genetic material that can be present at such a locus are known as *alleles*.) This sort of study was motivated in the first place by a desire to find mathematical models of natural selection, and thereby to discriminate between various competing accounts of the process of evolution. More recently, as scientists have gained access to the genetic material itself, many more questions of a statistical nature have arisen. We emphasise that we present only the very simplest examples in a rich theory, for which we refer the interested reader to *Mathematical Population Genetics* by W. J. Ewens (Springer, Berlin, 1979).

Example 5.1.3 (Wright–Fisher model)

This is the discrete-time Markov chain on $\{0, 1, \dots, m\}$ with transition probabilities

$$p_{ij} = \binom{m}{j} \left(\frac{i}{m}\right)^j \left(\frac{m-i}{m}\right)^{m-j}.$$

In each generation there are m alleles, some of type A and some of type a. The types of alleles in generation $n+1$ are found by choosing randomly (with replacement) from the types in generation n. If X_n denotes the number of alleles of type A in generation n, then $(X_n)_{n\geq 0}$ is a Markov chain with the above transition probabilities.

This can be viewed as a model of inheritance for a particular gene with two alleles A and a. We suppose that each individual has two genes, so the possibilities are AA, Aa and aa. Let us take m to be even with $m = 2k$. Suppose that individuals in the next generation are obtained by mating randomly chosen individuals from the current generation and that offspring inherit one allele from each parent. We have to allow that both parents may be the same, and in particular make no requirement that parents be of opposite sexes. Then if the generation n is, for example

$$AA \quad aA \quad AA \quad AA \quad aa,$$

then each gene in generation $n+1$ is A with probability $7/10$ and a with probability $3/10$, all independent. We might, for example, get

$$aa \quad aA \quad Aa \quad AA \quad AA.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

The structure of pairs of genes is irrelevant to the Markov chain $(X_n)_{n\geq 0}$, which simply counts the number of alleles of type A.

The communicating classes of $(X_n)_{n\geq 0}$ are $\{0\}, \{1, \ldots, m-1\}, \{m\}$. States 0 and m are absorbing and $\{1, \ldots, m-1\}$ is transient. The hitting probabilities for state m (pure AA) are given by

$$h_i = \mathbb{P}_i(X_n = m \text{ for some } n) = i/m.$$

This is obvious when one notes that $(X_n)_{n\geq 0}$ is a martingale; alternatively one can check that

$$h_i = \sum_{j=0}^{m} p_{ij} h_j.$$

According to this model, genetic diversity eventually disappears. It is known, however, that, for $p \in (0,1)$, as $m \to \infty$

$$\mathbb{E}_{pm}(T) \sim -2m\{(1-p)\log(1-p) + p\log p\}$$

where T is the hitting time of $\{0, m\}$, so in a large population diversity does not disappear quickly.

Some modifications are possible which model other aspects of genetic theory. Firstly, it may be that the three genetic types AA, Aa, aa have a relative selective advantage given by $\alpha, \beta, \gamma > 0$ respectively. This means that the probability of choosing allele A when $X_n = i$ is given by

$$\psi_i = \frac{\alpha(i/m)^2 + (1/2)\beta i(m-i)/m^2}{\alpha(i/m)^2 + \beta i(m-i)/m^2 + \gamma\big((m-i)/m\big)^2}$$

and the transition probabilities are

$$p_{ij} = \binom{m}{j} \psi_i^j (1-\psi_i)^{m-j}.$$

Secondly, we may allow genes to mutate. Suppose A mutates to a with probability u and a mutates to A with probability v. Then the probability of choosing A when $X_n = i$ is given by

$$\phi_i = \{i(1-u) + (m-i)v\}/m$$

and

$$p_{ij} = \binom{m}{j} \phi_i^j (1-\phi_i)^{m-j}.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

With $u, v > 0$, the states 0 and m are no longer absorbing, in fact the chain is irreducible, so attention shifts from hitting probabilities to the invariant distribution π. There is an exact calculation for the mean of π: we have

$$\begin{aligned}\mu &= \sum_{i=0}^{m} i\pi_i = \mathbb{E}_\pi(X_1) = \sum_{i=0}^{m} \pi_i \mathbb{E}_i(X_1) \\ &= \sum_{i=0}^{m} m\pi_i\phi_i = \sum_{i=0}^{m} \{i(1-u) + (m-i)v\}\pi_i = (1-u)\mu + mv - v\mu\end{aligned}$$

so that

$$\mu = mv/(u+v).$$

Example 5.1.4 (Moran model)

The Moran model is the birth-and-death chain on $\{0, 1, \dots, m\}$ with transition probabilities

$$p_{i,i-1} = i(m-i)/m^2, \quad p_{ii} = \big(i^2 + (m-i)^2\big)/m^2, \quad p_{i,i+1} = i(m-i)/m^2.$$

Here is the genetic interpretation: a population consists of individuals of two types, a and A; we choose randomly one individual from the population at time n, and add a new individual of the same type; then we choose, again randomly, one individual from the population at time n and remove it; so we obtain the population at time $n+1$. The same individual may be chosen each time, both to give birth and to die, in which case there is no change in the make-up of the population. Now, if X_n denotes the number of type A individuals in the population at time n, then $(X_n)_{n\geq 0}$ is a Markov chain with transition matrix P.

There are some obvious differences from the Wright–Fisher model: firstly, the Moran model cannot be interpreted in terms of a species where genes come in pairs, or where individuals have more than one parent; secondly in the Moran model we only change one individual at a time, not the whole population. However, the basic Markov chain structure is the same, with communicating classes $\{0\}, \{1, \dots, m-1\}, \{m\}$, absorbing states 0 and m and transient class $\{1, \dots, m-1\}$. The Moran model is reversible, and, like the Wright–Fisher model, is a martingale. The hitting probabilities are given by

$$\mathbb{P}_i(X_n = m \text{ for some } n) = i/m.$$

We can also calculate explicitly the mean time to absorption

$$k_i = \mathbb{E}_i(T)$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

where T is the hitting time of $\{0, m\}$. The simplest method is first to fix j and write down equations for the mean time k_i^j spent in j, starting from i, before absorption:

$$k_i^j = \delta_{ij} + (p_{i,i-1}k_{i-1}^j + p_{ii}k_i^j + p_{i,i+1}k_{i+1}^j) \quad \text{for } i = 1, \dots, m-1$$
$$k_0^j = k_m^j = 0.$$

Then, for $i = 1, \dots, m-1$

$$k_{i+1}^j - 2k_i^j + k_{i-1}^j = -\delta_{ij}m^2/j(m-j)$$

so that

$$k_i^j = \begin{cases} (i/j)k_j^j & \text{for } i \le j \\ \big((m-i)/(m-j)\big)k_j^j & \text{for } i \ge j \end{cases}$$

where k_j^j is determined by

$$\left(\frac{m-j-1}{m-j} - 2 + \frac{j-1}{j}\right) k_j^j = -\frac{m^2}{j(m-j)}$$

which gives $k_j^j = m$. Hence

$$k_i = \sum_{j=1}^{m-1} k_i^j = m\left\{\sum_{j=1}^{i}\left(\frac{m-i}{m-j}\right) + \sum_{j=i+1}^{m-1}\frac{i}{j}\right\}.$$

As in the Wright–Fisher model, one is really interested in the case where m is large, and $i = pm$ for some $p \in (0,1)$. Then

$$m^{-2}k_{pm} = (1-p)\sum_{j=1}^{mp}\frac{1}{m-j} + p\sum_{j=mp+1}^{m-1}\frac{1}{j} \to -(1-p)\log(1-p) - p\log p$$

as $m \to \infty$. So, as $m \to \infty$

$$\mathbb{E}_{pm}(T) \sim -m^2\{(1-p)\log(1-p) + p\log p\}.$$

For the Wright–Fisher model we claimed that

$$\mathbb{E}_{pm}(T) \sim -2m\{(1-p)\log(1-p) + p\log p\}$$

which has the same functional form in p and differs by a factor of $m/2$. This factor is partially explained by the fact that the Moran model deals

with one individual at a time, whereas the Wright–Fisher model changes all m at once.

Exercises

5.1.1 Consider a branching process with immigration. This is defined, in the notation of Example 5.1.1, by

$$X_n = N_1^n + \ldots + N_{X_{n-1}}^n + I_n$$

where $(I_n)_{n\geq 0}$ is a sequence of independent $\mathbb{Z}^+$-valued random variables with common generating function $\psi(t) = \mathbb{E}(t^{I_n})$. Show that, if $X_0 = 1$, then

$$\mathbb{E}(t^{X_n}) = \phi^{(n)}(t) \prod_{k=0}^{n-1} \psi\big(\phi^{(k)}(t)\big).$$

In the case where the number of immigrants in each generation is Poisson of parameter λ, and where $\mathbb{P}(N = 0) = 1 - p$ and $\mathbb{P}(N = 1) = p$, find the long-run proportion of time during which the population is zero.

5.1.2 A species of plant comes in three genotypes AA, Aa and aa. A single plant of genotype Aa is crossed with itself, so that the offspring has genotype AA, Aa or aa with probabilities 1/4, 1/2 and 1/4. How long on average does it take to achieve a pure strain, that is, AA or aa? Suppose it is desired to breed an AA plant. What should you do? How many crosses would your procedure require, on average?

5.1.3 In the Moran model we may introduce a selective bias by making it twice as likely that a type a individual is chosen to die, as compared to a type A individual. Thus in a population of size m containing i type A individuals, the probability that some type A is chosen to die is now $i/(i + 2(m - i))$. Suppose we begin with just one type A. What is the probability that eventually the whole population is of type A?

5.2 Queues and queueing networks

Queues form in many circumstances and it is important to be able to predict their behaviour. The basic mathematical model for queues runs as follows: there is a succession of customers wanting service; on arrival each customer must wait until a server is free, giving priority to earlier arrivals; it is assumed that the times between arrivals are independent random variables of the same distribution, and the times taken to serve customers are also independent random variables, of some other distribution. The main

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

quantity of interest is the random process $(X_t)_{t\geq 0}$ recording the number of customers in the queue at time t. This is always taken to include both those being served and those waiting to be served.

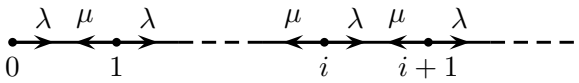
In cases where inter-arrival times and service times have exponential distributions, $(X_t)_{t\geq 0}$ turns out to be a *continuous-time Markov chain*, so we can answer many questions about the queue. This is the context of our first six examples. Some further variations on queues of this type have already appeared in Exercises 3.4.1, 3.6.3, 3.7.1 and 3.7.2.

If the inter-arrival times only are exponential, an analysis is still possible, by exploiting the memorylessness of the *Poisson process* of arrivals, and a certain *discrete-time Markov chain* embedded in the queue. This is explained in the final two examples.

In each example we shall aim to describe some salient features of the queue in terms of the given data of arrival-time and service-time distributions. We shall find conditions for the stability of the queue, and in the stable case find means to compute the equilibrium distribution of queue length. We shall also look at the random times that customers spend waiting and the length of time that servers are continuously busy.

Example 5.2.1 (M/M/1 queue)

This is the simplest queue of all. The code means: *memoryless inter-arrival times/memoryless service times/one server*. Let us suppose that the inter-arrival times are exponential of parameter λ, and the service times are exponential of parameter μ. Then the number of customers in the queue $(X_t)_{t\geq 0}$ evolves as a Markov chain with the following diagram:



To see this, suppose at time 0 there are i customers in the queue, where $i > 0$. Denote by T the time taken to serve the first customer and by A the time of the next arrival. Then the first jump time J_1 is $A \wedge T$, which is exponential of parameter $\lambda + \mu$, and $X_{J_1} = i - 1$ if $T < A$, $X_{J_1} = i + 1$ if $T > A$, which events are independent of J_1, with probabilities $\mu/(\lambda+\mu)$ and $\lambda/(\lambda+\mu)$ respectively. If we condition on $J_1 = T$, then $A - J_1$ is exponential of parameter λ and independent of J_1: the time already spent waiting for an arrival is forgotten. Similarly, conditional on $J_1 = A$, $T - J_1$ is exponential of parameter μ and independent of J_1. The case where $i = 0$ is simpler as there is no serving going on. Hence, conditional on $X_{J_1} = j$, $(X_t)_{t\geq 0}$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

begins afresh from j at time J_1. It follows that $(X_t)_{t\geq 0}$ is the claimed Markov chain. This sort of argument should by now be very familiar and we shall not spell out the details like this in later examples.

The M/M/1 queue thus evolves like a random walk, except that it does not take jumps below 0. We deduce that if $\lambda > \mu$ then $(X_t)_{t\geq 0}$ is transient, that is $X_t \to \infty$ as $t \to \infty$. Thus if $\lambda > \mu$ the *queue grows without limit* in the long term. When $\lambda < \mu$, $(X_t)_{t\geq 0}$ is positive recurrent with invariant distribution

$$\pi_i = (1 - \lambda/\mu)(\lambda/\mu)^i.$$

So when $\lambda < \mu$ the *average number of customers in the queue in equilibrium* is given by

$$\mathbb{E}_\pi(X_t) = \sum_{i=1}^{\infty} \mathbb{P}_\pi(X_t \geq i) = \sum_{i=1}^{\infty} (\lambda/\mu)^i = \lambda/(\mu - \lambda).$$

Also, the mean time to return to 0 is given by

$$m_0 = 1/(q_0\pi_0) = \mu/\lambda(\mu - \lambda)$$

so the *mean length of time that the server is continuously busy* is given by

$$m_0 - (1/q_0) = 1/(\mu - \lambda).$$

Another quantity of interest is the *mean waiting time for a typical customer*, when $\lambda < \mu$ and the queue is in equilibrium. Conditional on finding a queue of length i on arrival, this is $(i + 1)/\mu$, so the overall mean waiting time is

$$\mathbb{E}_\pi(X_t + 1)/\mu = 1/(\mu - \lambda).$$

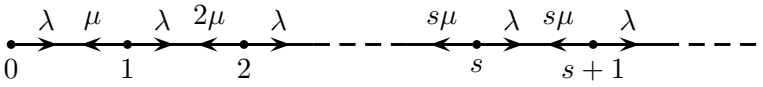
A rough check is available here as we can calculate in two ways the expected total time spent in the queue over an interval of length t: either we multiply the average queue length by t, or we multiply the mean waiting time by the expected number of customers λt. Either way we get $\lambda t/\mu - \lambda$. The first calculation is exact but we have not fully justified the second.

Thus, once the queue size is identified as a Markov chain, its behaviour is largely understood. Even in more complicated examples where exact calculation is limited, once the Markovian character of the queue is noted we know what sort of features to look for – transience and recurrence, convergence to equilibrium, long-run averages, and so on.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Example 5.2.2 (M/M/s queue)

This is a variation on the last example where there is one queue but there are s servers. Let us assume that the arrival rate is λ and the service rate by each server is μ. Then if i servers are occupied, the first service is completed at the minimum of i independent exponential times of parameter μ. The first service time is therefore exponential of parameter $i\mu$. The total service rate increases to a maximum $s\mu$ when all servers are working. We emphasise that the queue size includes those customers who are currently being served. By an argument similar to the preceding example, the queue size $(X_t)_{t\geq 0}$ performs a Markov chain with the following diagram:



So this time we obtain a birth-and-death chain. It is transient in the case $\lambda > s\mu$ and otherwise recurrent. To find an invariant measure we look at the detailed balance equations

$$\pi_i q_{i,i+1} = \pi_{i+1} q_{i+1,i}.$$

Hence

$$\pi_i/\pi_0 = \begin{cases} (\lambda/\mu)^i/i! & \text{for } i = 0, 1, \dots, s \\ (\lambda/\mu)^i/(s^{i-s}s!) & \text{for } i = s+1, s+2, \dots. \end{cases}$$

The queue is therefore positive recurrent when $\lambda < s\mu$. There are two cases when the invariant distribution has a particularly nice form: when $s = 1$ we are back to Example 5.2.1 and the invariant distribution is geometric of parameter λ/μ:

$$\pi_i = (1 - \lambda/\mu)(\lambda/\mu)^i.$$

When $s = \infty$ we normalize π by taking $\pi_0 = e^{-\lambda/\mu}$ so that

$$\pi_i = e^{-\lambda/\mu}(\lambda/\mu)^i/i!$$

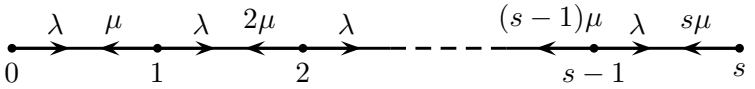
and the invariant distribution is Poisson of parameter λ/μ.

The number of arrivals by time t is a Poisson process of rate λ. Each arrival corresponds to an increase in X_t, and each departure to a decrease. Let us suppose that $\lambda < s\mu$, so there is an invariant distribution, and consider the queue in equilibrium. The detailed balance equations hold and $(X_t)_{t\geq 0}$ is non-explosive, so by Theorem 3.7.3 for any $T > 0$, $(X_t)_{0\leq t\leq T}$

and $(X_{T-t})_{0\le t\le T}$ have the same law. It follows that, in equilibrium, the number of departures by time t is also a Poisson process of rate λ. This is slightly counter-intuitive, as one might imagine that the departure process runs in fits and starts depending on the number of servers working. Instead, it turns out that the process of departures, in equilibrium, is just as regular as the process of arrivals.

Example 5.2.3 (Telephone exchange)

A variation on the M/M/s queue is to turn away customers who cannot be served immediately. This might serve as a simple model for a telephone exchange, where the maximum number of calls that can be connected at once is s: when the exchange is full, additional calls are lost. The maximum queue size or *buffer size* is s and we get the following modified Markov chain diagram:



We can find the invariant distribution of this finite Markov chain by solving the detailed balance equations, as in the last example. This time we get a *truncated Poisson distribution*

$$\pi_i = \frac{(\lambda/\mu)^i}{i!} \Big/ \sum_{j=0}^{s} \frac{(\lambda/\mu)^j}{j!}\,.$$

By the ergodic theorem, the long-run proportion of time that the exchange is full, and hence the long-run proportion of calls that are lost, is given by

$$\pi_s = \frac{(\lambda/\mu)^s}{s!} \Big/ \sum_{j=0}^{s} \frac{(\lambda/\mu)^j}{j!}\,.$$

This is known as *Erlang's formula*. Compare this example with the bus maintenance problem in Exercise 3.7.1.

Example 5.2.4 (Queues in series)

Suppose that customers have two service requirements: they arrive as a Poisson process of rate λ to be seen first by server A, and then by server

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

B. For simplicity we shall assume that the service times are independent exponentials of parameters α and β respectively. What is the average queue length at B?

Let us denote the queue length at A by $(X_t)_{t\geq 0}$ and that by B by $(Y_t)_{t\geq 0}$. Then $(X_t)_{t\geq 0}$ is simply an M/M/1 queue. If $\lambda > \alpha$, then $(X_t)_{t\geq 0}$ is transient so there is eventually always a queue at A and departures form a Poisson process of rate α. If $\lambda < \alpha$, then, by the reversibility argument of Example 5.2.2, the process of departures from A is Poisson of rate λ, *provided queue A is in equilibrium*. The question about queue length at B is not precisely formulated: it does not specify that the queues should be in equilibrium; indeed if $\lambda \geq \alpha$ there is no equilibrium. Nevertheless, we hope you will agree to treat arrivals at B as a Poisson process of rate $\alpha \wedge \lambda$. Then, by Example 5.2.1, the average queue length at B when $\alpha \wedge \lambda < \beta$, in equilibrium, is given by $(\alpha \wedge \lambda)/\big(\beta - (\alpha \wedge \lambda)\big)$. If, on the other hand, $\alpha \wedge \lambda > \beta$, then $(Y_t)_{t\geq 0}$ is transient so the queue at B grows without limit.

There is an equilibrium for both queues if $\lambda < \alpha$ and $\lambda < \beta$. The fact that in equilibrium the output from A is Poisson greatly simplifies the analysis of the two queues in series. For example, the average time taken by one customer to obtain both services is given by

$$1/(\alpha - \lambda) + 1/(\beta - \lambda).$$

Example 5.2.5 (Closed migration process)

Consider, first, a single particle in a finite state-space I which performs a Markov chain with irreducible Q-matrix Q. We know there is a unique invariant distribution π. We may think of the holding times of this chain as service times, by a single server at each node $i \in I$.

Let us suppose now that there are N particles in the state-space, which move as before except that they must queue for service at every node. If we do not care to distinguish between the particles, we can regard this as a new process $(X_t)_{t\geq 0}$ with state-space $\widetilde{I} = \mathbb{N}^I$, where $X_t = (n_i : i \in I)$ if at time t there are n_i particles at state i. In fact, this new process is also a Markov chain. To describe its Q-matrix $\widetilde{Q}$ we define a function $\delta_i : \widetilde{I} \to \widetilde{I}$ by

$$(\delta_i n)_j = n_j + \delta_{ij}.$$

Thus δ_i adds a particle at i. Then for $i \neq j$ the non-zero transition rates are given by

$$\widetilde{q}(\delta_i n, \delta_j n) = q_{ij}, \qquad n \in \widetilde{I}, \quad i, j \in I.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Observe that we can write the invariant measure equation $\pi Q = 0$ in the form

$$\pi_i \sum_{j \neq i} q_{ij} = \sum_{j \neq i} \pi_j q_{ji}. \tag{5.1}$$

For $n = (n_i : i \in I)$ we set

$$\widetilde{\pi}(n) = \prod_{i \in I} \pi_i^{n_i}.$$

Then

$$\begin{aligned} \widetilde{\pi}(\delta_i n) \sum_{j \neq i} \widetilde{q}(\delta_i n, \delta_j n) &= \prod_{k \in I} \pi_k^{n_k} \left(\pi_i \sum_{j \neq i} q_{ji} \right) \\ &= \prod_{k \in I} \pi_k^{n_k} \left(\sum_{j \neq i} \pi_j q_{ji} \right) \\ &= \sum_{j \neq i} \widetilde{\pi}(\delta_j n) \widetilde{q}(\delta_j n, \delta_i n). \end{aligned}$$

Given $m \in \widetilde{I}$ we can put $m = \delta_i n$ in the last identity whenever $m_i \geq 1$. On summing the resulting equations we obtain

$$\widetilde{\pi}(m) \sum_{n \neq m} \widetilde{q}(m, n) = \sum_{n \neq m} \widetilde{\pi}(n) \widetilde{q}(n, m)$$

so $\widetilde{\pi}$ is an invariant measure for $\widetilde{Q}$. The total number of particles is conserved so $\widetilde{Q}$ has communicating classes

$$C_N = \left\{ n \in \widetilde{I} : \sum_{i \in I} n_i = N \right\}$$

and the unique invariant distribution for the N-particle system is given by normalizing $\widetilde{\pi}$ restricted to C_N.

Example 5.2.6 (Open migration process)

We consider a modification of the last example where new customers, or particles, arrive at each node $i \in I$ at rate λ_i. We suppose also that customers receiving service at node i leave the network at rate μ_i. Thus customers enter the network, move from queue to queue according to a Markov chain and eventually leave, rather like a shopping centre. This model includes the closed system of the last example and also the queues

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

in series of Example 5.2.4. Let $X_t = (X_t^i : i \in I)$, where X_t^i denotes the number of customers at node i at time t. Then $(X_t)_{t\geq 0}$ is a Markov chain in $\widetilde{I} = \mathbb{N}^I$ and the non-zero transition rates are given by

$$\widetilde{q}(n, \delta_i n) = \lambda_i, \quad \widetilde{q}(\delta_i n, \delta_j n) = q_{ij}, \quad \widetilde{q}(\delta_j n, n) = \mu_j$$

for $n \in \widetilde{I}$ and distinct states $i, j \in I$. We shall assume that $\lambda_i > 0$ for some i and $\mu_j > 0$ for some j; then $\widetilde{Q}$ is irreducible on $\widetilde{I}$.

The system of equations (5.1) for an invariant measure is replaced here by

$$\pi_i \left(\mu_i + \sum_{j \neq i} q_{ij} \right) = \lambda_i + \sum_{j \neq i} \pi_j q_{ji}.$$

This system has a unique solution, with $\pi_i > 0$ for all i. This may be seen by considering the invariant distribution for the extended Q-matrix $\overline{Q}$ on $I \cup \{\partial\}$ with off-diagonal entries

$$\overline{q}_{\partial j} = \lambda_j, \quad \overline{q}_{ij} = q_{ij}, \quad \overline{q}_{i\partial} = \mu_i.$$

On summing the system over $i \in I$ we find

$$\sum_{i \in I} \pi_i \mu_i = \sum_{i \in I} \lambda_i.$$

As in the last example, for $n = (n_i : i \in I)$ we set

$$\widetilde{\pi}(n) = \prod_{i \in I} \pi_i^{n_i}.$$

Transitions from $m \in \widetilde{I}$ may be divided into those where a new particle is added and, for each $i \in I$ with $m_i \geq 1$, those where a particle is moved from i to somewhere else. We have, for the first sort of transition

$$\begin{aligned} \widetilde{\pi}(m) \sum_{j \in I} \widetilde{q}(m, \delta_j m) &= \widetilde{\pi}(m) \sum_{j \in I} \lambda_j \\ &= \widetilde{\pi}(m) \sum_{j \in I} \pi_j \mu_j = \sum_{j \in I} \widetilde{\pi}(\delta_j m) \widetilde{q}(\delta_j m, m) \end{aligned}$$

and for the second sort

$$\begin{aligned} &\widetilde{\pi}(\delta_i n) \Big(\widetilde{q}(\delta_i n, n) + \sum_{j \neq i} \widetilde{q}(\delta_i n, \delta_j n) \Big) \\ &= \prod_{k \in I} \pi_k^{n_k} \left(\pi_i (\mu_i + \sum_{j \neq i} q_{ij}) \right) \\ &= \prod_{k \in I} \pi_k^{n_k} \left(\lambda_i + \sum_{j \neq i} \pi_j q_{ji} \right) \\ &= \widetilde{\pi}(n) \widetilde{q}(n, \delta_i n) + \sum_{j \neq i} \widetilde{\pi}(\delta_j n) \widetilde{q}(\delta_j n, \delta_i n). \end{aligned}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

On summing these equations we obtain

$$\widetilde{\pi}(m)\sum_{n\neq m}\widetilde{q}(m,n)=\sum_{n\neq m}\widetilde{\pi}(n)\widetilde{q}(n,m)$$

so $\widetilde{\pi}$ is an invariant measure for Q. If $\pi_i<1$ for all i then $\widetilde{\pi}$ has finite total mass $\prod_{i\in I}(1-\pi_i)$, otherwise the total mass if infinite. Hence, $\widetilde{Q}$ is positive recurrent if and only if $\pi_i<1$ for all i, and in that case, in equilibrium, the individual queue lengths $(X_t^i : i\in I)$ are *independent* geometric random variables with

$$\mathbb{P}(X_t^i=n_i)=(1-\pi_i)\pi_i^{n_i}.$$

Example 5.2.7 (M/G/1 queue)

As we argued in Section 2.4, the Poisson process is the natural probabilistic model for any uncoordinated stream of discrete events. So we are often justified in assuming that arrivals to a queue form a Poisson process. In the preceding examples we also assumed an exponential service-time distribution. This is desirable because it makes the queue size into a continuous-time Markov chain, but it is obviously inappropriate in many real-world examples. The service requirements of customers and the duration of telephone calls have observable distributions which are generally not exponential. A better model in this case is the M/G/1 queue, where G indicates that the service-time distribution is general.

We can characterize the distribution of a service time T by its distribution function

$$F(t)=\mathbb{P}(T\leq t),$$

or by its Laplace transform

$$L(w)=\mathbb{E}(e^{-wT})=\int_0^\infty e^{-wt}dF(t).$$

(The integral written here is the Lebesgue–Stieltjes integral: when T has a density function $f(t)$ we can replace $dF(t)$ by $f(t)dt$.) Then the mean service time μ is given by

$$\mu=\mathbb{E}(T)=-L'(0+).$$

To analyse the M/G/1 queue, we consider the queue size X_n immediately following the nth departure. Then

$$X_{n+1}=X_n+Y_{n+1}-1_{X_n>0} \tag{5.2}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

where Y_n denotes the number of arrivals during the nth service time. The case where $X_n = 0$ is different because then we get an extra arrival before the $(n+1)$th service time begins. By the Markov property of the Poisson process, $Y_1, Y_2, \ldots$ are independent and identically distributed, *so* $(X_n)_{n\geq 0}$ *is a discrete-time Markov chain*. Indeed, except for visits to 0, $(X_n)_{n\geq 0}$ behaves as a random walk with jumps $Y_n - 1$.

Let T_n denote the nth service time. Then, conditional on $T_n = t$, Y_n is Poisson of parameter λt. So

$$\mathbb{E}(Y_n) = \int_0^\infty \lambda t dF(t) = \lambda\mu$$

and, indeed, we can compute the probability generating function

$$\begin{aligned} A(z) = \mathbb{E}(z^{Y_n}) &= \int_0^\infty \mathbb{E}(z^{Y_n} \mid T_n = t) dF(t) \\ &= \int_0^\infty e^{-\lambda t(1-z)} dF(t) = L\big(\lambda(1-z)\big). \end{aligned}$$

Set $\rho = \mathbb{E}(Y_n) = \lambda\mu$. We call ρ the *service intensity*. Let us suppose that $\rho < 1$. We have

$$X_n = X_0 + (Y_1 + \ldots + Y_n) - n + Z_n$$

where Z_n denotes the number of visits of X_n to 0 before time n. So

$$\mathbb{E}(X_n) = \mathbb{E}(X_0) - n(1-\rho) + \mathbb{E}(Z_n).$$

Take $X_0 = 0$, then, since $X_n \geq 0$, we have for all n

$$0 < 1 - \rho \leq \mathbb{E}(Z_n/n).$$

By the ergodic theorem we know that, as $n \to \infty$

$$\mathbb{E}(Z_n/n) \to 1/m_0$$

where m_0 is the mean return time to 0. Hence

$$m_0 \leq 1/(1-\rho) < \infty$$

showing that $(X_n)_{n\geq 0}$ is positive recurrent.

Suppose now that we start $(X_n)_{n\geq 0}$ with its equilibrium distribution π. Set

$$G(z) = \mathbb{E}(z^{X_n}) = \sum_{i=0}^{\infty} \pi_i z^i$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

then

$$\begin{aligned} zG(z) &= \mathbb{E}(z^{X_{n+1}+1}) = \mathbb{E}(z^{X_n+Y_{n+1}+1_{X_n=0}}) \\ &= \mathbb{E}(z^{Y_{n+1}})\Big(\pi_0 z + \sum_{i=1}^{\infty} \pi_i z^i\Big) \\ &= A(z)\big(\pi_0 z + G(z) - \pi_0\big) \end{aligned}$$

so

$$\big(A(z) - z\big)G(z) = \pi_0 A(z)(1 - z). \tag{5.3}$$

By l'Hôpital's rule, as $z \uparrow 1$

$$\big(A(z) - z\big)/(1 - z) \to 1 - A'(1-) = 1 - \rho.$$

Since $G(1) = 1 = A(1)$, we must therefore have $\pi_0 = 1 - \rho, m_0 = 1/(1 - \rho)$ and

$$G(z) = (1 - \rho)(1 - z)A(z)/\big(A(z) - z\big).$$

Since A is given explicitly in terms of the service-time distribution, we can now obtain, in principle, the full equilibrium distribution. The fact that generating functions work well here is due to the additive structure of (5.2).

To obtain the *mean queue length* we differentiate (5.3)

$$\big(A(z) - z\big)G'(z) = \big(A'(z) - 1\big)G(z) = (1 - \rho)\big\{A'(z)(1 - z) - A(z)\big\},$$

then substitute for $G(z)$ to obtain

$$G'(z) = (1-\rho)A'(z)\frac{(1-z)}{\big(A(z)-z\big)} - (1-\rho)A(z)\frac{\big\{\big(A'(z)-1\big)(1-z) + A(z) - z\big\}}{\big(A(z)-z\big)^2}.$$

By l'Hôpital's rule:

$$\lim_{z\uparrow 1} \frac{\big(A'(z)-1\big)(1-z) + A(z) - z}{\big(A(z)-z\big)^2} = \lim_{z\uparrow 1} \frac{A''(z)(1-z)}{2\big(A'(z)-1\big)\big(A(z)-z\big)} = \frac{-A''(1-)}{2(1-\rho)^2}.$$

Hence

$$\begin{aligned} \mathbb{E}(X_n) &= G'(1-) = \rho + A''(1-)/2(1 - \rho) \\ &= \rho + \lambda^2 L''(0+)/2(1 - \rho) = \rho + \lambda^2 \mathbb{E}(T^2)/2(1 - \rho). \end{aligned}$$

In the case of the M/M/1 queue $\rho = \lambda/\mu$, $\mathbb{E}(T^2) = 2/\mu^2$ and $\mathbb{E}(X_n) = \rho/(1 - \rho) = \lambda/(\mu - \lambda)$, as we found in Example 5.2.1.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

We shall use generating functions to study two more quantities of interest – the queueing time of a typical customer and the busy periods of the server.

Consider the queue $(X_n)_{n\in\mathbb{Z}}$ in equilibrium. Suppose that the customer who leaves at time 0 has spent time Q queueing to be served, and time T being served. Then, conditional on $Q+T=t$, X_0 is Poisson of parameter λt, since the customers in the queue at time 0 are precisely those who arrived during the queueing and service times of the departing customer. Hence

$$G(z) = \mathbb{E}\big(e^{-\lambda(Q+T)(1-z)}\big) = M\big(\lambda(1-z)\big)L\big(\lambda(1-z)\big)$$

where M is the Laplace transform

$$M(w) = \mathbb{E}(e^{-wQ}).$$

On substituting for $G(z)$ we obtain the formula

$$M(w) = (1-\rho)w/\Big(w - \lambda\big(1-L(w)\big)\Big).$$

Differentiation and l'Hôpital's rule, as above, lead to a formula for the *mean queueing time*

$$\mathbb{E}(Q) = -M'(0+) = \frac{\lambda L''(0+)}{2\big(1+\lambda L'(0+)\big)^2} = \frac{\lambda\mathbb{E}(T^2)}{2(1-\rho)}\,.$$

We now turn to the busy period S. Consider the Laplace transform

$$B(w) = \mathbb{E}(e^{-wS}).$$

Let T denote the service time of the first customer in the busy period. Then conditional on $T=t$, we have

$$S = t + S_1 + \ldots + S_N,$$

where N is the number of customers arriving while the first customer is served, which is Poisson of parameter λt, and where $S_1, S_2, \ldots$ are independent, with the same distribution as S. Hence

$$\begin{aligned} B(w) &= \int_0^\infty \mathbb{E}(e^{-wS} \mid T=t)dF(t) \\ &= \int_0^\infty e^{-wt}e^{-\lambda t(1-B(w))}dF(t) = L\Big(w+\lambda\big(1-B(w)\big)\Big). \end{aligned}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Although this is an implicit relation for $B(w)$, we can obtain moments by differentiation:

$$\mathbb{E}(S) = -B'(0+) = -L'(0+)\big(1 - \lambda B'(0+)\big) = \mu\big(1 + \lambda\mathbb{E}(S)\big)$$

so the *mean length of the busy period* is given by

$$\mathbb{E}(S) = \mu/(1-\rho).$$

Example 5.2.8 (M/G/∞ queue)

Arrivals at this queue form a Poisson process, of rate λ, say. Service times are independent, with a common distribution function $F(t) = \mathbb{P}(T \le t)$. There are infinitely many servers, so all customers in fact receive service at once. The analysis here is simpler than in the last example because customers do not interact. Suppose there are no customers at time 0. What, then, is the distribution of the number X_t being served at time t?

The number N_t of arrivals by time t is a Poisson random variable of parameter λt. We condition on $N_t = n$ and label the times of the n arrivals randomly by $A_1, \dots, A_n$. Then, by Theorem 2.4.6, $A_1, \dots, A_n$ are independent and uniformly distributed on the interval $[0, t]$. For each of these customers, service is incomplete at time t with probability

$$p = \frac{1}{t}\int_0^t \mathbb{P}(T > s)ds = \frac{1}{t}\int_0^t \big(1 - F(s)\big)ds.$$

Hence, conditional on $N_t = n$, X_t is binomial of parameters n and p. Then

$$\begin{aligned}
\mathbb{P}(X_t = k) &= \sum_{n=0}^{\infty} \mathbb{P}(X_t = k \mid N_t = n)\mathbb{P}(N_t = n) \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda t}(\lambda t)^n/n! \\
&= e^{-\lambda t}(\lambda p t)^k/k! \sum_{n=k}^{\infty} \big(\lambda(1-p)t\big)^{n-k}/(n-k)! \\
&= e^{-\lambda p t}(\lambda p t)^k/k!
\end{aligned}$$

So we have shown that X_t is Poisson of parameter

$$\lambda \int_0^t \big(1 - F(s)\big)ds.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Recall that

$$\int_0^\infty \big(1 - F(s)\big)ds = \int_0^\infty \mathbb{E}(1_{T>t})dt = \mathbb{E}\int_0^\infty 1_{T>t}dt = \mathbb{E}(T).$$

Hence if $\mathbb{E}(T) < \infty$, the queue size has a limiting distribution, which is Poisson of parameter $\lambda\mathbb{E}(T)$.

For further reading see *Reversibility and Stochastic Networks* by F. P. Kelly (Wiley, Chichester, 1978).

5.3 Markov chains in resource management

Management decisions are always subject to risk because of the uncertainty of future events. If one can quantify that risk, perhaps on the basis of past experience, then the determination of the best action will rest on the calculation of probabilities, often involving a Markov chain. Here we present some examples involving the management of a resource: either the stock in a warehouse, or the water in a reservoir, or the reserves of an insurance company. See also Exercise 3.7.1 on the maintenance of unreliable equipment. The statistical problem of estimating transition rates for Markov chains has already been discussed in Section 1.10.

Example 5.3.1 (Restocking a warehouse)

A warehouse has a capacity of c units of stock. In each time period n, there is a demand for D_n units of stock, which is met if possible. We denote the residual stock at the end of period n by X_n. The warehouse manager restocks to capacity for the beginning of period $n+1$ whenever $X_n \le m$, for some threshold m. Thus $(X_n)_{n\ge 0}$ satisfies

$$X_{n+1} = \begin{cases} (c - D_{n+1})^+ & \text{if } X_n \le m \\ (X_n - D_{n+1})^+ & \text{if } m < X_n \le c. \end{cases}$$

Let us assume that $D_1, D_2, \dots$ are independent and identically distributed; then $(X_n)_{n\ge 0}$ is a Markov chain, and, excepting some peculiar demand structures, is irreducible on $\{0, 1, \dots, c\}$. Hence $(X_n)_{n\ge 0}$ has a unique invariant distribution π which determines the long-run proportion of time in each state. Given that $X_n = i$, the expected unmet demand in period $n+1$ is given by

$$u_i = \begin{cases} \mathbb{E}\big((D-c)^+\big) & \text{if } i \le m \\ \mathbb{E}\big((D-i)^+\big) & \text{if } m < i \le c. \end{cases}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Hence the long-run proportion of demand that is unmet is

$$u(m) = \sum_{i=0}^{c} \pi_i u_i.$$

The long-run frequency of restocking is given by

$$r(m) = \sum_{i=0}^{m} \pi_i.$$

Now as m increases, $u(m)$ decreases and $r(m)$ increases. The warehouse manager may want to compute these quantities in order to optimize the long-run cost

$$ar(m) + bu(m)$$

where a is the cost of restocking and b is the profit per unit.

There is no general formula for π, but once the distribution of the demand is known, it is a relatively simple matter to write down the $(c+1) \times (c+1)$ transition matrix P for $(X_n)_{n \geq 0}$ and solve $\pi P = \pi$ subject to $\sum_{i=0}^{c} \pi_i = 1$. We shall discuss in detail a special case where the calculations work out nicely.

Suppose that the capacity $c = 3$, so possible threshold values are $m = 0, 1, 2$. Suppose that the profit per unit $b = 1$, and that the demand satisfies

$$\mathbb{P}(D \geq i) = 2^{-i} \quad \text{for } i = 0, 1, 2, \dots .$$

Then

$$\mathbb{E}\big((D-i)^+\big) = \sum_{k=1}^{\infty} \mathbb{P}\big((D-i)^+ \geq k\big) = \sum_{k=1}^{\infty} \mathbb{P}(D \geq i+k) = 2^{-i}.$$

The transition matrices for $m = 0, 1, 2$ are given, respectively, by

$$\begin{pmatrix} 1/8 & 1/8 & 1/4 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/2 & 0 \\ 1/8 & 1/8 & 1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 1/8 & 1/8 & 1/4 & 1/2 \\ 1/8 & 1/8 & 1/4 & 1/2 \\ 1/4 & 1/4 & 1/2 & 0 \\ 1/8 & 1/8 & 1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 1/8 & 1/8 & 1/4 & 1/2 \\ 1/8 & 1/8 & 1/4 & 1/2 \\ 1/8 & 1/8 & 1/4 & 1/2 \\ 1/8 & 1/8 & 1/4 & 1/2 \end{pmatrix}$$

with invariant distributions

$$(1/4, 1/4, 1/4, 1/4), \quad (1/6, 1/6, 1/3, 1/3), \quad (1/8, 1/8, 1/4, 1/2).$$

Hence

$$u(0) = 1/4, \quad u(1) = 1/6, \quad u(2) = 1/8$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

and

$$r(0) = 1/4, \quad r(1) = 1/3, \quad r(2) = 1/2.$$

Therefore, to minimize the long-run cost $ar(m) + u(m)$ we should take

$$m = \begin{cases} 2 & \text{if } a \leq 1/4 \\ 1 & \text{if } 1/4 < a \leq 1 \\ 0 & \text{if } 1 < a. \end{cases}$$

Example 5.3.2 (Reservoir model – discrete time)

We are concerned here with a storage facility, for example a reservoir, of finite capacity c. In each time period n, A_n units of resource are available to enter the facility and B_n units are drawn off. When the reservoir is full, surplus water is lost. When the reservoir is empty, no water can be supplied. We assume that newly available resources cannot be used in the current time period. Then the quantity of water X_n in the reservoir at the end of period n satisfies

$$X_{n+1} = \big((X_n - B_{n+1})^+ + A_{n+1}\big) \wedge c.$$

If we assume that A_n, B_n and c are integer-valued and that $A_1, A_2, \ldots$ are independent and identically distributed, likewise $B_1, B_2, \ldots$, then $(X_n)_{n\geq 0}$ is a Markov chain on $\{0, 1, \ldots, c\}$, whose transition probabilities may be deduced from the distributions of A_n and B_n. Hence we know that the long-run behaviour of $(X_n)_{n\geq 0}$ is controlled by its unique invariant distribution π, assuming irreducibility. For example, the long-run proportion of time that the reservoir is empty is simply π_0. So we would like to calculate π.

A simplifying assumption which makes some calculations possible is to assume that consumption in each period is constant, and that our units are chosen to make this constant 1. Then the infinite capacity model satisfies a recursion similar to the M/G/1 queue:

$$X_{n+1} = (X_n - 1)^+ + A_{n+1}.$$

Hence, by the argument used in Example 5.2.7, if $\mathbb{E}(A_n) < 1$, then $(X_n)_{n\geq 0}$ is positive recurrent and the invariant distribution π satisfies

$$\sum_{i=0}^{\infty} \pi_i z^i = (1 - \mathbb{E}A_n)(1 - z)A(z)/\big(A(z) - z\big)$$

where $A(z) = \mathbb{E}(z^{A_n})$. In fact, whether or not $\mathbb{E}(A_n) < 1$, the equation

$$\sum_{i=0}^{\infty} \nu_i z^i = (1 - z)A(z)/\big(A(z) - z\big)$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

serves to define a solution to the infinite system of linear equations

$$\nu_0 = \nu_0(a_0 + a_1) + \nu_1 a_0$$
$$\nu_i = \nu_{i+1}a_0 + \sum_{j=0}^{i} \nu_j a_{i-j+1}, \quad \text{for } i \geq 1$$

where $a_i = \mathbb{P}(A_n = i)$.

Note that $(X_n)_{n\geq 0}$ can only enter $\{0, 1, \ldots, c\}$ through c. Hence, by the strong Markov property, $(X_n)_{n\geq 0}$ observed whilst in $\{0, 1, \ldots, c\}$ is simply the finite-capacity model. In the case where $\mathbb{E}(A_n) < 1$, we can deduce for the finite-capacity model that the long-run proportion of time in state i is given by $\nu_i/(\nu_0 + \ldots + \nu_c)$. In fact, this is true in general as the equilibrium equations for the finite-capacity model coincide with those for ν up to level $c-1$, and the level c equation is redundant.

In reality, it is to be hoped that, in the long run, supply will exceed demand, which is true if $\mathbb{E}(A_n) > 1$. Then $(X_n)_{n\geq 0}$ is transient, so ν must have infinite total mass. The problem faced by the water company is to keep the long-run proportion of time $\pi_0(c)$ that the reservoir is empty below a certain acceptable fraction, $\varepsilon > 0$ say. Hence c should be chosen large enough to make

$$\nu_o/(\nu_0 + \ldots + \nu_c) < \varepsilon$$

which is always possible in the transient case.

Example 5.3.3 (Reservoir model – continuous time)

Consider a reservoir model where fresh water arrives at the times of a Poisson process of rate λ. The quantities of water $S_1, S_2, \ldots$ arriving each time are assumed independent and identically distributed. We assume that there is a continuous demand for water of rate 1. For a reservoir of infinite capacity, the quantity of water held $(W_t)_{t\geq 0}$ is just the stored work in an M/G/1 queue with the same arrival times and service times $S_1, S_2, \ldots$. The periods when the reservoir is empty correspond to idle periods of the queue. Hence in the positive recurrent case where $\lambda\mathbb{E}(S_n) < 1$, the long-run proportion of time that the reservoir is empty is given by $\mathbb{E}(S_n)/(1 - \lambda\mathbb{E}(S_n))$. Note that $(W_t)_{t\geq 0}$ can enter $[0, c]$ only through c. As in the preceding example we can obtain the finite capacity model by observing $(W_t)_{t\geq 0}$ whilst in $[0, c]$, but we shall not pursue this here.

The next example is included, in part, because it illustrates a surprising and powerful connection between *reflected random walks* and the *maxima*

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

of random walks, which we now explain. Let $X_1, X_2, \ldots$ denote a sequence of independent, identically distributed random variables. Set $S_n = X_1 + \ldots + X_n$ and define $(Z_n)_{n\geq 0}$ by $Z_0 = 0$ and

$$Z_{n+1} = (Z_n + X_{n+1})^+.$$

Then, by induction, we have

$$Z_n = \max\{0, X_n, X_{n-1} + X_n, \ldots, X_1 + \ldots + X_n\}$$

so Z_n *has the same distribution as* M_n where

$$M_n = \max\{0, X_1, X_1 + X_2, \ldots, X_1 + \ldots + X_n\} = \max_{m\leq n} S_m.$$

Example 5.3.4 (Ruin of an insurance company)

An insurance company receives premiums continuously at a constant rate. We choose units making this rate 1. The company pays claims at the times of a Poisson process of rate λ, the claims $Y_1, Y_2, \ldots$ being independent and identically distributed. Set $\rho = \lambda\mathbb{E}(Y_1)$ and assume that $\rho < 1$. Then in the long run the company can expect to make a profit of $1 - \rho$ per unit time. However, there is a danger that large claims early on will ruin the company even though the long-term trend is good.

Denote by S_n the cumulative net loss following the nth claim. Thus $S_n = X_1 + \ldots + X_n$, where $X_n = Y_n - T_n$ and T_n is the nth inter-arrival time. By the strong law of large numbers

$$S_n/n \to \mathbb{E}(Y_n) - 1/\lambda < 0$$

as $n \to \infty$. The maximum loss that the company will have to sustain is

$$M = \lim_{n\to\infty} M_n$$

where

$$M_n = \max_{m\leq n} S_m.$$

By the argument given above, M_n has the same distribution as Z_n, where $Z_0 = 0$ and

$$Z_{n+1} = (Z_n + Y_n - T_n)^+.$$

But Z_n is the queueing time of the nth customer in the M/G/1 queue with inter-arrival times T_n and service times Y_n. We know by Example 5.2.7 that the queue-length distribution converges to equilibrium. Hence, so does the

queueing-time distribution. Also by Example 5.2.7, we know the Laplace transform of the equilibrium queueing-time distribution. Hence

$$\mathbb{E}(e^{-wM}) = (1-\rho)w/\Big(w - \lambda\big(1 - \mathbb{E}(e^{-wY_1})\big)\Big).$$

The probability of eventual bankruptcy is $\mathbb{P}(M > a)$, where a denotes the initial value of the company's assets. In principle, this may now be obtained by inverting the Laplace transform.

5.4 Markov decision processes

In many contexts costs are incurred at a rate determined by some process which may best be modelled as a Markov chain. We have seen in Section 1.10 and Section 4.2 how to calculate in these circumstances the long-run average cost or the expected total cost. Suppose now that we are able to choose the transition probabilities for each state from a given class and that our choice determines the cost incurred. The question arises as to how best to do this to minimize our expected costs.

Example 5.4.1

A random walker on $\{0, 1, 2, \dots\}$ jumps one step to the right with probability p and one step to the left with probability $q = 1 - p$. Any value of $p \in (0, 1]$ may be chosen, but incurs a cost

$$c(p) = 1/p.$$

The walker on reaching 0 stays there, incurring no further costs.

If we are only concerned with minimizing costs over the first few time steps, then the choice $p = 1$ may be best. However, in the long run the only way to avoid an infinite total cost is to get to 0. Starting from i we must first hit $i-1$, then $i-2$, and so on. Given the lack of memory in the model, this makes it reasonable to pick the same value of p throughout, and seek to minimize $\phi(p)$, the expected total cost starting from 1. The expected total cost starting from 2 is $2\phi(p)$ since we must first hit 1. Hence

$$\phi(p) = c(p) + 2p\phi(p)$$

so that

$$\phi(p) = \begin{cases} c(p)/(1-2p) & \text{for } p < 1/2 \\ \infty & \text{for } p \geq 1/2. \end{cases}$$

Thus for $c(p) = 1/p$ the choice $p = 1/4$ is optimal, with expected cost 8. The general discussion which follows will make rigorous what we claimed was reasonable.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Generally, let us suppose given some distribution $\lambda = (\lambda_i : i \in I)$ and, for each *action* $a \in A$, a transition matrix $P(a) = \big(p_{ij}(a) : i, j \in I\big)$ and a cost function $c(a) = \big(c_i(a) : i \in I\big)$. These are the data for a *Markov decision process*, though so far we have no process and when we do it will not in general be Markov. To get a process we must choose a policy, that is, a way of determining actions by our current knowledge of the process. Formally, a *policy* u is a sequence of functions

$$u_n : I^{n+1} \to A, \qquad n = 0, 1, 2, \dots \ .$$

Each policy u determines a probability law $\mathbb{P}^u$ for a process $(X_n)_{n\geq 0}$ with values in I by

(i) $\mathbb{P}^u(X_0 = i_0) = \lambda_{i_0}$;

(ii) $\mathbb{P}^u(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) = p_{i_n i_{n+1}}\big(u_n(i_0, \dots, i_n)\big)$.

A *stationary policy* u is a function $u : I \to A$. We abuse notation and write u also for the associated policy given by

$$u_n(i_0, \dots, i_n) = u(i_n).$$

Under a stationary policy u, the probability law $\mathbb{P}^u$ makes $(X_n)_{n\geq 0}$ Markov, with transition probabilities $p_{ij}^u = p_{ij}\big(u(i)\big)$.

We suppose that a cost $c(i, a) = c_i(a)$ is incurred when action a is chosen in state i. Then we associate to a policy u an *expected total cost* starting from i, given by

$$V^u(i) = \mathbb{E}^u \sum_{n=0}^{\infty} c\big(X_n, u_n(X_0, \dots, X_n)\big).$$

So that this sum is well defined, we assume that $c(i, a) \geq 0$ for all i and a. Define also the *value function*

$$V^*(i) = \inf_u V^u(i)$$

which is the minimal expected total cost starting from i.

The basic problem of Markov decision theory is how to minimize expected costs by our choice of policy. The minimum expected cost incurred before time $n = 1$ is given by

$$V_1(i) = \inf_a c(i, a).$$

Then the minimum expected cost incurred before time $n = 2$ is

$$V_2(i) = \inf_a \big\{ c(i, a) + \sum_{j \in I} p_{ij}(a) V_1(j) \big\}.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Define inductively

$$V_{n+1}(i) = \inf_a \big\{ c(i,a) + \sum_{j \in I} p_{ij}(a) V_n(j) \big\}. \tag{5.4}$$

It is easy to see by induction that $V_n(i) \le V_{n+1}(i)$ for all i, so $V_n(i)$ increases to a limit $V_\infty(i)$, possibly infinite. We have

$$V_{n+1}(i) \le c(i,a) + \sum_{j \in I} p_{ij}(a) V_n(j) \qquad \text{for all } a$$

so, letting $n \to \infty$ and then minimizing over a,

$$V_\infty(i) \le \inf_a \big\{ c(i,a) + \sum_{j \in I} p_{ij}(a) V_\infty(j) \big\}. \tag{5.5}$$

It is a reasonable guess that $V_\infty(i)$, being the limit of minimal expected costs over finite time intervals, is in fact the value function $V^*(i)$. This is not always true, unless we can show that the inequality (5.5) is actually an equality. We make three technical assumptions to ensure this. *We assume that*

(i) *for all i, j the functions $c_i : A \to [0, \infty)$ and $p_{ij} : A \to [0, \infty)$ are continuous;*

(ii) *for all i and all $B < \infty$ the set $\{a : c_i(a) \le B\}$ is compact;*

(iii) *for each i, for all but finitely many j, for all $a \in A$ we have $p_{ij}(a) = 0$.*

A simple case where (i) and (ii) hold is when A is a finite set. It is easy to check that the assumptions are valid in Example 5.4.1, with $A = (0, 1]$, $c_i(a) = 1/a$ and

$$p_{ij}(a) = \begin{cases} a & \text{if } j = i+1 \\ 1-a & \text{if } j = i-1 \\ 0 & \text{otherwise,} \end{cases}$$

with obvious exceptions at $i = 0$.

Lemma 5.4.2. *There is a stationary policy u such that*

$$V_\infty(i) = c\big(i, u(i)\big) + \sum_{j \in I} p_{ij}\big(u(i)\big) V_\infty(j). \tag{5.6}$$

Proof. If $V_\infty(i) = \infty$ there is nothing to prove, so let us assume that $V_\infty(i) \le B < \infty$. Then

$$V_{n+1}(i) = \inf_{a \in K} \left\{ c(i,a) + \sum_{j \in J} p_{ij}(a) V_n(j) \right\}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

where K is the compact set $\{a : c(i,a) \le B\}$ and where J is the finite set $\{j : p_{ij} \not\equiv 0\}$. Hence, by continuity, the infimum is attained and

$$V_{n+1}(i) = c\big(i, u_n(i)\big) + \sum_{j\in J} p_{ij}\big(u_n(i)\big)V_n(j) \tag{5.7}$$

for some $u_n(i) \in K$. By compactness there is a convergent subsequence $u_{n_k}(i) \to u(i)$, say, and, on passing to the limit $n_k \to \infty$ in (5.7), we obtain (5.6). □

Theorem 5.4.3. *We have*

(i) $V_n(i) \uparrow V^*(i)$ *as* $n \to \infty$ *for all* i;

(ii) *if* u^* *is any stationary policy such that* $a = u^*(i)$ *minimizes*

$$c(i,a) + \sum_{j\in I} p_{ij}(a)V^*(j)$$

for all i, *then* u^* *is optimal, in the sense that*

$$V^{u^*}(i) = V^*(i) \qquad \text{for all } i.$$

Proof. For any policy u we have

$$\begin{aligned} V^u(i) &= \mathbb{E}_i^u \sum_{n=0}^{\infty} c\big(X_n, u_n(X_0, \dots, X_n)\big) \\ &= c\big(i, u_0(i)\big) + \sum_{j\in I} p_{ij}\big(u_0(i)\big)V^{u[i]}(j) \end{aligned}$$

where $u[i]$ is the policy given by

$$u[i]_n(i_0, \dots, i_n) = u_{n+1}(i, i_0, \dots, i_n).$$

Hence we obtain

$$V^u(i) \ge \inf_a \big\{c(i,a) + \sum_{j\in I} p_{ij}(a)V^*(j)\big\}$$

and, on taking the infimum over u

$$V^*(i) \ge \inf_a \big\{c(i,a) + \sum_{j\in I} p_{ij}(a)V^*(j)\big\}. \tag{5.8}$$

Certainly, $V_0(i) = 0 \le V^*(i)$. Let us suppose inductively that $V_n(i) \le V^*(i)$ for all i. Then by substitution in the right sides of (5.4) and (5.8) we find $V_{n+1}(i) \le V^*(i)$ and the induction proceeds. Hence $V_\infty(i) \le V^*(i)$ for all i.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Let u^* be any stationary policy for which

$$V_\infty(i) \geq c\big(i, u^*(i)\big) + \sum_{j \in I} p_{ij}\big(u^*(i)\big) V_\infty(j).$$

We know such a policy exists by Lemma 5.4.2. Then by Theorem 4.2.3 we have $V^{u^*}(i) \leq V_\infty(i)$ for all i. But $V^*(i) \leq V^{u^*}(i)$ for all i, so

$$V_\infty(i) = V^*(i) = V^{u^*}(i) \qquad \text{for all } i$$

and we are done. □

The theorem just proved shows that the problem of finding a good policy is much simpler than we might have supposed. For it was not clear at the outset that there would be a single policy which was optimal for all i, even less that this policy would be stationary. Moreover, part (i) gives an explicit way of obtaining the value function V^* and, once this is known, part (ii) identifies an optimal stationary policy.

In practice we may know only an approximation to V^*, for example V_n for n large. We may then hope that, by choosing $a = u(i)$ to minimize

$$c(i, a) + \sum_{j \in I} p_{ij}(a) V_n(j)$$

we get a nearly optimal policy. An alternative means of constructing nearly optimal policies is sometimes provided by the method of *policy improvement*. Given one stationary policy u we may define another θu by the requirement that $a = (\theta u)(i)$ minimizes

$$c(i, a) + \sum_{j \in I} p_{ij}(a) V^u(j).$$

Theorem 5.4.4 (Policy improvement). *We have*

(i) $V^{\theta u}(i) \leq V^u(i)$ *for all* i;

(ii) $V^{\theta^n u}(i) \downarrow V^*(i)$ *as* $n \to \infty$ *for all* i, *provided that*

$$\mathbb{E}_i^{u^*}\big(V^u(X_n)\big) \to 0 \quad \text{as} \quad n \to \infty \qquad \text{for all } i. \tag{5.9}$$

Proof. (i) We have, by Theorem 4.2.3

$$\begin{aligned} V^u(i) &= c\big(i, u(i)\big) + \sum_{j \in I} p_{ij}\big(u(i)\big) V^u(j) \\ &\geq c\big(i, \theta u(i)\big) + \sum_{j \in I} p_{ij}\big(\theta u(i)\big) V^u(j) \end{aligned}$$

so $V^u(i) \geq V^{\theta u}(i)$ for all i, by Theorem 4.2.3.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

(ii) We note from part (i) that

$$V^{\theta u}(i) \le c(i,a) + \sum_{j\in I} p_{ij}(a)V^u(j) \qquad \text{for all } i \text{ and } a. \tag{5.10}$$

Fix $N \ge 0$ and consider for $n = 0, 1, \dots, N$ the process

$$M_n = V^{\theta^{N-n}u}(X_n) + \sum_{k=0}^{n-1} c\big(X_k, u^*(X_k)\big).$$

Recall the notation for conditional expectation introduced in Section 4.1. We have

$$\begin{aligned}\mathbb{E}^{u^*}(M_{n+1} \mid \mathcal{F}_n) &= \sum_{j\in I} p_{X_n j}\big(u^*(X_n)\big)V^{\theta^{N-n-1}u}(j) + c\big(X_n, u^*(X_n)\big) \\ &\quad + \sum_{k=0}^{n-1} c\big(X_k, u^*(X_k)\big) \\ &\ge M_n\end{aligned}$$

where we used (5.10) with u replaced by $\theta^{N-n-1}u$, $i = X_n$ and $a = u^*(X_n)$. It follows that $\mathbb{E}^{u^*}(M_{n+1}) \ge \mathbb{E}^{u^*}(M_n)$ for all n. Hence if we assume (5.9), then

$$\begin{aligned}V^{\theta^N u}(i) &= \mathbb{E}_i^{u^*}(M_0) \le \mathbb{E}_i^{u^*}(M_N) \\ &= \mathbb{E}_i^{u^*}\big(V^u(X_N)\big) + \mathbb{E}^{u^*}\left(\sum_{n=0}^{N-1} c\big(X_n, u^*(X_n)\big)\right) \\ &\to V^*(i) \qquad \text{as } N \to \infty.\end{aligned}$$

□

We have been discussing the minimization of expected total cost, which is only relevant to the transient case. This is because we will have $V^*(i) = \infty$ unless for some stationary policy u, the only states j with positive cost $c(j, u(j)) > 0$, accessible from i, are transient. The recurrent case is also of practical importance and one way to deal with this is to discount costs at future times by a fixed factor $\alpha \in (0,1)$. We now seek to minimize the *expected total discounted cost*

$$V_\alpha^u(i) = \mathbb{E}_i^u \sum_{n=0}^{\infty} \alpha^n c\big(X_n, u_n(X_0, \dots, X_n)\big).$$

Define the *discounted value function*

$$V_\alpha^*(i) = \inf_u V_\alpha^u(i).$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

In fact, the discounted case reduces to the undiscounted case by introducing a new absorbing state ∂ and defining a new Markov decision process by

$$\widetilde{p}_{ij}(a) = \alpha p_{ij}(a), \quad \widetilde{p}_{i\partial}(a) = 1 - \alpha,$$
$$\widetilde{c}_i(a) = c_i(a), \qquad \widetilde{c}_\partial(a) = 0.$$

Thus the new process follows the old until, at some geometric time of parameter α, it jumps to ∂ and stays there, incurring no further costs.

Introduce $V_{0,\alpha}(i) = 0$ and, inductively

$$V_{n+1,\alpha}(i) = \inf_a \big\{ c(i,a) + \alpha \sum_{j \in J} p_{ij}(a) V_{n,\alpha}(i) \big\}$$

and, given a stationary policy u, define another $\theta_\alpha u$ by the requirement that $a = (\theta_\alpha u)(i)$ minimizes

$$c(i,a) + \alpha \sum_{j \in J} p_{ij}(a) V^u(j).$$

Theorem 5.4.5. *Suppose that the cost function $c(i,a)$ is uniformly bounded.*

(i) *We have $V_{n,\alpha}(i) \uparrow V^*_\alpha(i)$ as $n \to \infty$ for all i.*

(ii) *The value function V^*_α is the unique bounded solution to*

$$V^*_\alpha(i) = \inf_a \big\{ c(i,a) + \alpha \sum_{j \in I} p_{ij}(a) V^*_\alpha(j) \big\}. \tag{5.11}$$

(iii) *Let u^* be a stationary policy such that $a = u^*(i)$ minimizes*

$$c(i,a) + \alpha \sum_{j \in I} p_{ij}(a) V^*_\alpha(j)$$

for all i. Then u^* is optimal in the sense that

$$V^{u^*}_\alpha(i) = V^*_\alpha(i) \qquad \text{for all } i.$$

(iv) *For all stationary policies u we have*

$$V^{\theta^n_\alpha u}_\alpha(i) \downarrow V^*_\alpha(i) \qquad \text{as } n \to \infty \text{ for all } i.$$

Proof. With obvious notation we have

$$V^u_\alpha = \widetilde{V}^u, \quad V^*_\alpha = \widetilde{V}^*, \quad V_{n,\alpha} = \widetilde{V}_n, \quad \theta_\alpha u = \widetilde{\theta}_u$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

so parts (i), (ii) and (iii) follow directly from Theorems 5.4.3 and 5.4.4, except for the uniqueness claim in (ii). But given any bounded solution V to (5.11), there is a stationary policy u such that

$$V(i) = c\big(i, u(i)\big) + \alpha \sum_{j \in I} p_{ij}\big(u(i)\big) V(j).$$

Then $V = V_\alpha^u$, by Theorem 4.2.5. Then $\theta_\alpha u = u$ so (iv) will show that u is optimal and $V = V_\alpha^*$.

We have $c(i,a) \leq B$ for some $B < \infty$. So for any stationary policy u we have

$$V_\alpha^u(i) = \mathbb{E}_i^u \sum_{n=0}^{\infty} \alpha^n c\big(X_n, u(X_n)\big) \leq B/(1-\alpha)$$

and so

$$\widetilde{\mathbb{E}}_i^{u^*}\big(\widetilde{V}^u(X_n)\big) = \alpha^n \mathbb{E}_i^{u^*}\big(V_\alpha^u(X_n)\big) \leq B\alpha^n/(1-\alpha) \to 0$$

as $n \to \infty$. Hence (iv) also follows from Theorem 5.4.4. □

We finish with a discussion of long-run average costs. Here we are concerned with the limiting behaviour, as $n \to \infty$, of

$$\overline{V}_n^u(i) = \mathbb{E}_i^u \left(\frac{1}{n} \sum_{k=0}^{n-1} c\big(X_k, u_k(X_0, \dots, X_k)\big) \right).$$

We assume that

$$|c(i,a)| \leq B < \infty \qquad \text{for all } i \text{ and } a.$$

This forces $|\overline{V}_n^u(i)| \leq B$ for all n, but in general the sequence $\overline{V}_n^u(i)$ may fail to converge as $n \to \infty$. In the case of a stationary strategy u for which $(X_n)_{n \geq 0}$ has a unique invariant distribution π^u, we know by the ergodic theorem that

$$\frac{1}{n} \sum_{k=0}^{n-1} c\big(X_k, u(X_k)\big) \to \sum_{j \in I} \pi_j^u c\big(j, u(j)\big)$$

as $n \to \infty$, $\mathbb{P}_i^u$-almost surely, for all i. So $\overline{V}_n^u(i)$ does converge in this case by bounded convergence, with the same limit. This suggests that one approach to minimizing long-run costs might be to minimize

$$\sum_{j \in I} \pi_j^u c\big(j, u(j)\big).$$

But, although this is sometimes valid, we do not know in general that the optimal policy is positive recurrent, or even stationary. Instead, we use a martingale approach, which is more general.

Theorem 5.4.6. *Suppose we can find a constant* $\overline{V}^*$ *and a bounded function* $W(i)$ *such that*

$$\overline{V}^* + W(i) = \inf_a \big\{ c(i,a) + \sum_{j \in I} p_{ij}(a) W(j) \big\} \qquad \text{for all } i. \tag{5.12}$$

Let u^* *be any stationary strategy such that* $a = u^*(i)$ *achieves the infimum in* (5.12) *for each* i. *Then*

(i) $\overline{V}_n^{u^*}(i) \to \overline{V}^*$ *as* $n \to \infty$ *for all* i;

(ii) $\liminf_{n\to\infty} \overline{V}_n^u(i) \geq \overline{V}^*$ *for all* i, *for all* u.

Proof. Fix a strategy u and set $U_n = u_n(X_0, \dots, X_n)$. Consider

$$M_n = W(X_n) - n\overline{V}^* + \sum_{k=0}^{n-1} c(X_k, U_k).$$

Then

$$\begin{aligned} &\mathbb{E}^u(M_{n+1} \mid \mathcal{F}_n) \\ &= M_n + \big\{ c(X_n, U_n) + \sum_{j \in I} p_{X_n j}(U_n) W(j) \big\} - \big(\overline{V}^* + W(X_n) \big) \\ &\geq M_n \end{aligned}$$

with equality if $u = u^*$. Therefore

$$W(i) = \mathbb{E}_i^u(M_0) \leq \mathbb{E}_i^u(M_n) = \mathbb{E}_i^u W(X_n) - n\overline{V}^* + n\overline{V}_n^u(i).$$

So we obtain

$$\overline{V}^* \leq \overline{V}_n^u(i) + 2 \sup_i |W(i)|/n.$$

This implies (ii) on letting $n \to \infty$. When $u = u^*$ we also have

$$\overline{V}_n^{u^*}(i) \leq \overline{V}^* + 2 \sup_i |W(i)|/n$$

and hence (i). □

The most obvious point of this theorem is that it identifies an optimal stationary policy when the hypothesis is met. Two further aspects also deserve comment. Firstly, if u is a stationary policy for which $(X_n)_{n \geq 0}$ has an invariant distribution π^u, then

$$\begin{aligned} \sum_{i \in I} \pi_i^u \big(\overline{V}^* + W(i) \big) &\leq \sum_{i \in I} \pi_i^u \left(c\big(i, u(i)\big) + \sum_{j \in I} p_{ij}\big(u(i)\big) W(j) \right) \\ &= \sum_{i \in I} \pi_i^u c\big(i, u(i)\big) + \sum_{j \in I} \pi_j^u W(j) \end{aligned}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

so

$$\overline{V}^* \leq \sum_{i\in I} \pi_i^u c\big(i, u(i)\big)$$

with equality if we can take $u = u^*$.

Secondly, there is a connection with the case of discounted costs. Assume that I is finite and that $P(a)$ is irreducible for all a. Then we can show that as $\alpha \uparrow 1$ we have

$$V_\alpha^*(i) = \overline{V}^*/(1-\alpha) + W(i) + o(1-\alpha).$$

On substituting this into (5.11) we find

$$\begin{aligned}&\overline{V}^*/(1-\alpha) + W(i) + o(1-\alpha)\\ &= \inf_a \left\{ c(i,a) + \alpha \sum_{j\in I} p_{ij}(a)\big(\overline{V}^*/(1-\alpha) + W(j) + o(1-\alpha)\big) \right\}\end{aligned}$$

so

$$\overline{V}^* + W(i) = \inf_a \left\{ c(i,a) + \alpha \sum_{j\in I} p_{ij}(a) W(j) \right\} + o(1-\alpha)$$

which brings us back to (5.12) on letting $\alpha \uparrow 1$.

The interested reader is referred to S. M. Ross, *Applied Probability Models with Optimization Applications* (Holden-Day, San Francisco, 1970) and to H. C. Tijms, *Stochastic Models – an algorithmic approach* (Wiley, Chichester, 1994) for more examples, results and references.

5.5 Markov chain Monte Carlo

Most computers may be instructed to provide a sequence of numbers

$$\begin{aligned}u_1 &= 0.u_{11}u_{12}u_{13}\ldots u_{1m}\\ u_2 &= 0.u_{21}u_{22}u_{23}\ldots u_{2m}\\ u_3 &= 0.u_{31}u_{32}u_{33}\ldots u_{3m}\end{aligned}$$

written as decimal expansions of a certain length, which for many purposes may be regarded as sample values of a sequence of independent random variables, uniformly distributed on $[0,1]$:

$$U_1(\omega), U_2(\omega), U_3(\omega), \ldots .$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

We are cautious in our language because, of course, $u_1, u_2, u_3, \ldots$ are actually all integer multiples of 10^{-m} and, more seriously, they are usually derived sequentially by some entirely deterministic algorithm in the computer. Nevertheless, the generators of such *pseudo-random numbers* are in general as reliable an imitation as one could wish of $U_1(\omega), U_2(\omega), U_3(\omega), \ldots$. This makes it worth while considering how one might construct Markov chains from a given sequence of independent uniform random variables, and then might exploit the observed properties of such processes.

We shall now describe one procedure to simulate a Markov chain $(X_n)_{n\geq 0}$ with initial distribution λ and transition matrix P. Since $\sum_{i\in I} \lambda_i = 1$ we can partition $[0,1]$ into disjoint subintervals $(A_i : i \in I)$ with lengths

$$|A_i| = \lambda_i.$$

Similarly for each $i \in I$, we can partition $[0,1]$ into disjoint subintervals $(A_{ij} : j \in I)$ such that

$$|A_{ij}| = p_{ij}.$$

Now define functions

$$G_0 : [0,1] \to I,$$
$$G : I \times [0,1] \to I$$

by

$$G_0(u) = i \quad \text{if } u \in A_i,$$
$$G(i,u) = j \quad \text{if } u \in A_{ij}.$$

Suppose that $U_0, U_1, U_2, \ldots$ is a sequence of independent random variables, uniformly distributed on $[0,1]$, and set

$$X_0 = G_0(U_0),$$
$$X_{n+1} = G(X_n, U_{n+1}) \quad \text{for } n \geq 0.$$

Then

$$\mathbb{P}(X_0 = i) = \mathbb{P}(U_0 \in A_i) = \lambda_i,$$
$$\mathbb{P}(X_{n+1} = i_{n+1} \mid X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(U_{n+1} \in A_{i_n i_{n+1}}) = p_{i_n i_{n+1}}$$

so $(X_n)_{n\geq 0}$ is Markov(λ, P).

This simple procedure may be used to investigate empirically those aspects of the behaviour of a Markov chain where theoretical calculations become infeasible.

The remainder of this section is devoted to one application of the simulation of Markov chains. It is the application which finds greatest practical use, especially in statistics, statistical physics and computer science, known as *Markov chain Monte Carlo*. Monte Carlo is another name for computer simulation so this sounds no different from the procedure just discussed. But what is really meant is simulation *by means of* Markov chains, the object of primary interest being the invariant distribution of the Markov chain and not the chain itself. After a general discussion we shall give two examples.

The context for Markov chain Monte Carlo is a state-space in product form

$$I = \prod_{m \in \Lambda} S_m$$

where Λ is a finite set. For the purposes of this discussion we shall also assume that each component S_m is a finite set. A random variable X with values in I is then a family of component random variables $\big(X(m) : m \in \Lambda\big)$, where, for each site $m \in \Lambda$, $X(m)$ takes values in S_m.

We are given a distribution $\pi = (\pi_i : i \in I)$, perhaps up to an unknown constant multiple, and it is desired to compute the number

$$\sum_{i \in I} \pi_i f_i \tag{5.13}$$

for some given function $f = (f_i : i \in I)$. The essential point to understand is that Λ is typically a large set, making the state-space I very large indeed. Then certain operations are computationally infeasible – performing the sum (5.13) state by state for a start.

An alternative approach would be to simulate a large number of independent random variables $X_1, \ldots, X_n$ in I, each with distribution π, and to approximate (5.13) by

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k).$$

The strong law of large numbers guarantees that this is a good approximation as $n \to \infty$ and, moreover, one can obtain error estimates which indicate how large to make n in practice. However, simulation from the distribution π is also difficult, unless π has product form

$$\pi(x) = \prod_{m \in \Lambda} \pi_m\big(x(m)\big).$$

For recall that a computer just simulates sequences of independent $U[0,1]$ random variables. When π does not have product form, Markov chain Monte Carlo is sometimes the only way to simulate samples from π.

The basic idea is to simulate a Markov chain $(X_n)_{n\geq 0}$, which is constructed to have invariant distribution π. Then, assuming aperiodicity and irreducibility, we know, by Theorem 1.8.3, that as $n \to \infty$ the distribution of X_n converges to π. Indeed, assuming only irreducibility, Theorem 1.10.2 shows that

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \to \sum_{i\in I} \pi_i f_i$$

with probability 1. But why should simulating an entire Markov chain be easier than simulating a simple distribution π? The answer lies in the fact that the state-space is a product.

Each component $X_0(m)$ of the initial state X_0 is a random variable in S_m. It does not matter crucially what distribution X_0 is given, but we might, for example, make all components independent. The process $(X_n)_{n\geq 0}$ is made to evolve by changing components one site at a time. When the chosen site is m, we simulate a new random variable $X_{n+1}(m)$ with values in S_m according to a distribution determined by X_n, and for $k \neq m$ we set $X_{n+1}(k) = X_n(k)$. Thus at each step we have only to simulate a random variable in S_m, not one in the much larger space I.

Let us write $i \overset{m}{\sim} j$ if i and j agree, except possibly at site m. The law for simulating a new value at site m is described by a transition matrix $P(m)$, where

$$p_{ij}(m) = 0 \quad \text{unless } i \overset{m}{\sim} j.$$

We would like π to be invariant for $P(m)$. A sufficient condition is that the detailed balance equations hold: thus for all i, j we want

$$\pi_i p_{ij}(m) = \pi_j p_{ji}(m).$$

There are many possible choices for $P(m)$ satisfying these equations. Indeed, given any stochastic matrix $R(m)$ with

$$r_{ij}(m) = 0 \quad \text{unless } i \overset{m}{\sim} j$$

we can determine such a $P(m)$ by

$$\pi_i p_{ij}(m) = \big(\pi_i r_{ij}(m)\big) \wedge \big(\pi_j r_{ji}(m)\big)$$

for $i \neq j$, and then

$$p_{ii}(m) = 1 - \sum_{j\neq i} p_{ij}(m) \geq 0.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

This has the following interpretation: if $X_n = i$ we simulate a new random variable Y_n so that $Y_n = j$ with probability $r_{ij}(m)$, then if $Y_n = j$ we set

$$X_{n+1} = \begin{cases} Y_n & \text{with probability } \big(\pi_j r_{ji}(m)/\pi_i r_{ij}(m)\big) \wedge 1 \\ X_n & \text{otherwise.} \end{cases}$$

This is called a *Hastings algorithm*.

There are two commonly used special cases. On taking

$$r_{ij}(m) = \Big(\sum_{k \overset{m}{\sim} i} \pi_k\Big)^{-1} \pi_j \qquad \text{for } i \overset{m}{\sim} j$$

we also find

$$p_{ij}(m) = \Big(\sum_{k \overset{m}{\sim} i} \pi_k\Big)^{-1} \pi_j \qquad \text{for } i \overset{m}{\sim} j.$$

So we simply resample $X_n(m)$ according to the conditional distribution under π, given the other components. This is called the *Gibbs sampler*. It is particularly useful in Bayesian statistics.

On taking $r_{ij}(m) = r_{ji}(m)$ for all i and j we find

$$p_{ij}(m) = \big((\pi_j/\pi_i) \wedge 1\big) r_{ij}(m) \qquad \text{for } i \overset{m}{\sim} j, i \neq j.$$

This is called a *Metropolis algorithm*. A particularly simple case would be to take

$$r_{ij}(m) = 1/(N_m - 1) \qquad \text{for } i \overset{m}{\sim} j, i \neq j$$

where $N_m = |S_m|$. This amounts to choosing another value j_m at site m uniformly at random; if $\pi_j > \pi_i$, then we adopt the new value, whereas if $\pi_j \leq \pi_i$ we adopt the new value with probability π_j/π_i.

We have not yet specified a rule for deciding which site to visit when. In practice this may not matter much, provided we keep returning to every site. For definiteness we mention two possibilities. We might choose to visit every site once and then repeat, generating a sequence of sites $(m_n)_{n\geq 0}$. Then $(m_n, X_n)_{n\geq 0}$ is a Markov chain in $\Lambda \times I$. Alternatively, we might choose a site randomly at each step. Then $(X_n)_{n\geq 0}$ is itself a Markov chain with transition matrix

$$P = |\Lambda|^{-1} \sum_{m \in \Lambda} P(m).$$

We shall stick with this second choice, where the analysis is simpler to present. Let us assume that P is irreducible, which is easy to ensure in the examples. We know that

$$\pi_i p_{ij}(m) = \pi_j p_{ji}(m)$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

for all m and all i, j, so also

$$\pi_i p_{ij} = \pi_j p_{ji}$$

and so π is the unique invariant measure for P. Hence, by Theorem 1.10.2, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \to \sum_{i\in I} \pi_i f_i$$

as $n \to \infty$ with probability 1. Thus the algorithm works eventually. In practice one is concerned with how fast it works, but useful information of this type cannot be gained in the present general context. Given more information on the structure of S_m and the distribution π to be simulated, much more can be said. We shall not pursue the matter here. It should also be emphasised that there is an empirical side to simulation: with due caution informed by the theory, the computer output gives a good idea of how well we are doing. For further reading we recommend *Stochastic Simulation* by B. D. Ripley (Wiley, Chichester, 1987), and *Markov Chain Monte Carlo in practice* by W. R. Gilks, S. Richardson and D. J. Spiegelhalter (Chapman and Hall, London, 1996). The recent survey article *Bayesian computation and stochastic systems* by J. Besag, P. Green, D. Higdon and K. Mengersen (*Statistical Science*, 10 (1), pp. 3–40, 1995) contains many interesting references. We finish with two examples.

Example 5.5.1 (Bayesian statistics)

In a statistical problem one may be presented with a set of independent observations $Y_1, \ldots, Y_n$, which it is reasonable to assume are normally distributed, but with unknown mean μ and variance τ^{-1}. One then seeks to draw conclusions about μ and τ on the basis of the observations. The Bayesian approach to this problem is to assume that μ and τ are themselves random variables, with a given prior distribution. For example, we might assume that

$$\mu \sim N(\theta_0, \phi_0^{-1}), \quad \tau \sim \Gamma(\alpha_0, \beta_0),$$

that is to say, μ is normal of mean θ_0 and variance ϕ_0^{-1}, and τ has gamma distribution of parameters α_0 and β_0. The parameters θ_0, ϕ_0, α_0 and β_0 are known. Then the prior density for (μ, τ) is given by

$$\pi(\mu, \tau) \propto \exp\{-\phi_0(\mu - \theta_0)^2/2\}\tau^{\alpha_0 - 1}\exp\{-\beta_0\tau\}.$$

The posterior density for (μ, τ), which is the conditional density given the observations, is then given by Bayes' formula

$$\pi(\mu, \tau \mid y) \propto \pi(\mu, \tau) f(y \mid \mu, \tau)$$

$$\propto \exp\{-\phi_0(\mu - \theta_0)^2/2\} \exp\left\{-\tau \sum_{i=1}^{n} (y_i - \mu)^2/2\right\} \tau^{\alpha_0 - 1 + n/2} \exp\{-\beta_0\tau\}.$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Note that the posterior density is no longer in product form: the conditioning has introduced a dependence between μ and τ. Nevertheless, the *full conditional distributions* still have a simple form

$$\pi(\mu \mid y, \tau) \propto \exp\{-\phi_0(\mu-\theta_0)^2/2\} \exp\left\{-\tau \sum_{i=1}^{n}(y_i-\mu)^2/2\right\} \sim N(\theta_n, \phi_n^{-1}),$$

$$\pi(\tau \mid y, \mu) \propto \tau^{\alpha_0-1+n/2} \exp\left\{-\tau\left(\beta_0 + \sum_{i=1}^{n}(y_i-\mu)^2/2\right)\right\} \sim \Gamma(\alpha_n, \beta_n)$$

where

$$\theta_n = \left(\phi_0\theta_0 + \tau\sum_{i=1}^{n} y_i\right)/(\phi_0 + n\tau), \quad \phi_n = \phi_0 + n\tau,$$

$$\alpha_n = \alpha_0 + n/2, \quad \beta_n = \beta_0 + \sum_{i=1}^{n}(y_i-\mu)^2/2.$$

Our final belief about μ and τ is regarded as measured by the posterior density. We may wish to compute probabilities and expectations. Here the *Gibbs sampler* provides a particularly simple approach. Of course, numerical integration would also be feasible as the dimension is only two. To make the connection with our general discussion we set

$$I = S_1 \times S_2 = \mathbb{R} \times [0, \infty).$$

We wish to simulate $X = (\mu, \tau)$ with density $\pi(\mu, \tau \mid y)$. The fact that $\mathbb{R}$ and $[0, \infty)$ are not finite sets does not affect the basic idea. In any case the computer will work with finite approximations to $\mathbb{R}$ and $[0, \infty)$. First we simulate X_0, say from the product form density $\pi(\mu, \tau)$. At the kth stage, given $X_k = (\mu_k, \tau_k)$, we first simulate μ_{k+1} from $\pi(\mu \mid y, \tau_k)$ and then τ_{k+1} from $\pi(\tau \mid y, \mu_{k+1})$, then set $X_{k+1} = (\mu_{k+1}, \tau_{k+1})$. Then $(X_k)_{k\geq 0}$ is a Markov chain in I with invariant measure $\pi(\mu, \tau \mid y)$, and one can show that

$$\frac{1}{k}\sum_{j=0}^{k-1} f(X_j) \to \int_I f(x)\pi(x \mid y)dx \quad \text{as } k \to \infty$$

with probability 1, for all bounded continuous functions $f : I \to \mathbb{R}$. This is not an immediate consequence of the ergodic theorem for discrete state-space, but you may find it reasonable at an intuitive level, with a rate of convergence depending on the smoothness of π and f.

We now turn to an elaboration of this example where the Gibbs sampler is indispensible. The model consists of m copies of the preceding one, with

different means but a common variance. Thus there are mn independent observations Y_{ij}, where $i = 1, \dots n$, and $j = 1, \dots, m$, normally distributed, with means μ_j and common variance τ^{-1}. We take these parameters to be independent random variables as before, with

$$\mu_j \sim N(\theta_0, \phi_0^{-1}), \quad \tau \sim \Gamma(\alpha_0, \beta_0).$$

Let us write $\mu = (\mu_1, \dots, \mu_n)$. The prior density is given by

$$\pi(\mu, \tau) \propto \exp\left\{-\phi_0 \sum_{j=1}^{m} (\mu_j - \theta_0)^2/2\right\} \tau^{\alpha_0 - 1} \exp\{-\beta_0 \tau\}$$

and the posterior density is given by

$$\pi(\mu, \tau \mid y) \propto \exp\left\{-\phi_0 \sum_{j=1}^{m} (\mu_j - \theta_0)^2/2\right\}$$
$$\times \exp\left\{-\tau \sum_{i=1}^{n} \sum_{j=1}^{m} (y_{ij} - \mu_j)^2/2\right\} \tau^{\alpha_0 - 1 + mn/2} \exp\{-\beta_0 \tau\}.$$

Hence the full conditional distributions are

$$\pi(\mu_j \mid y, \tau) \sim N(\theta_{jn}, \phi_n^{-1}), \quad \pi(\tau \mid y, \mu) \sim \Gamma(\alpha_n, \beta_n)$$

where

$$\theta_{jn} = \left(\phi_0 \theta_0 + \tau \sum_{i=1}^{n} y_{ij}\right) / (\phi_0 + n\tau), \quad \phi_n = \phi_0 + n\tau,$$
$$\alpha_n = \alpha_0 + mn/2, \quad \beta_n = \beta_0 + \sum_{i=1}^{n} \sum_{j=1}^{m} (y_{ij} - \mu_j)^2/2.$$

We can construct approximate samples from $\pi(\mu, \tau \mid y)$, just as in the case $m = 1$ discussed above, by a Gibbs sampler method. Note that, conditional on τ, the means μ_j, for $j = 1, \dots, m$, remain independent. Thus one can update all the means simultaneously in the Gibbs sampler. This has the effect of speeding convergence to the equilibrium distribution. In cases where m is large, numerical integration of $\pi(\mu, \tau \mid y)$ is infeasible, as is direct simulation from the distribution, so the Markov chain approach is the only one available.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

Example 5.5.2 (Ising model and image analysis)

Consider a large box $\Lambda = \Lambda_N$ in $\mathbb{Z}^2$

$$\Lambda = \{-N, \dots, -1, 0, 1, \dots, N\}^2$$

with boundary $\partial\Lambda = \Lambda_N \backslash \Lambda_{N-1}$, and the *configuration space*

$$I = \{-1, 1\}^\Lambda.$$

For $x \in \Lambda$ define

$$H(x) = \tfrac{1}{2} \sum \big(x(m) - x(m')\big)^2$$

where the sum is taken over all pairs $\{m, m'\} \subseteq \Lambda$ with $|m - m'| = 1$. Note that $H(x)$ is small when the values taken by x at neighbouring sites are predominantly the same. We write

$$I^+ = \{x \in I : x(m) = 1 \quad \text{for all } m \in \partial\Lambda\}$$

and for each $\beta > 0$ define a probability distribution $\big(\pi(x) : x \in I^+\big)$ by

$$\pi(x) \propto e^{-\beta H(x)}.$$

As $\beta \downarrow 0$ the weighting becomes uniform, whereas, as $\beta \uparrow \infty$ the mass concentrates on configurations x where $H(x)$ is small. This is one of the fundamental models of statistical physics, called the *Ising model*. A famous and deep result of Onsager says that if X has distribution π, then

$$\lim_{N\to\infty} \mathbb{E}\big(X(0)\big) = \big[\big(1 - (\sinh 2\beta)^{-4}\big)^+\big]^{1/8}.$$

In particular, if $\sinh 2\beta \le 1$, the fact that X is forced to take boundary values 1 does not significantly affect the distribution of $X(0)$ when N is large, whereas if $\sinh 2\beta > 1$ there is a residual effect of the boundary values on $X(0)$, uniformly in N.

Here we consider the problem of simulating the Ising model. Simulations may sometimes be used to guide further developments in the theory, or even to detect phenomena quite out of reach of the current theory. In fact, the Ising model is rather well understood theoretically; but there are many related models which are not, where simulation is still possible by simple modifications of the methods presented here.

First we describe a Gibbs sampler. Consider the sets of even and odd sites

$$\Lambda^+ = \{(m_1, m_2) \in \Lambda : m_1 + m_2 \text{ is even}\},$$
$$\Lambda^- = \{(m_1, m_2) \in \Lambda : m_1 + m_2 \text{ is odd}\}$$

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

and for $x \in I$ set

$$x^{\pm} = \big(x(m) : m \in \Lambda^{\pm}\big).$$

We can exploit the fact that the conditional distribution $\pi(x^+ \mid x^-)$ has product form

$$\pi(x^+ \mid x^-) \propto \prod_{m \in \Lambda^+ \setminus \partial\Lambda} e^{\beta x(m) s(m)}$$

where, for $m \in \Lambda^+ \setminus \partial\Lambda$

$$s(m) = \sum_{|m'-m|=1} x^-(m').$$

Therefore, it is easy to simulate from $\pi(x^+ \mid x^-)$ and likewise from $\pi(x^- \mid x^+)$. Choose now some simple initial configuration X_0 in I^+. Then inductively, given $X_n^- = x^-$, simulate firstly X_{n+1}^+ with distribution $\pi(\cdot \mid x^-)$ and then given $X_{n+1}^+ = x^+$, simulate X_{n+1}^- with distribution $\pi(\cdot \mid x^+)$. Then according to our general discussion, for large n, the distribution of X_n is approximately π. Note that we did not use the value of the normalizing constant

$$Z = \sum_{x \in I^+} e^{-\beta H(x)}$$

which is hard to compute by elementary means when N is large.

An alternative approach is to use a Metropolis algorithm. We can again exploit the even/odd partition. Given that $X_n = x$, independently for each $m \in \Lambda^+ \setminus \partial\Lambda$, we change the sign of $X_n^+(m)$ with probability

$$p(m, x) = \big(\pi(\hat{x})/\pi(x)\big) \wedge 1 = e^{2\beta x(m) s(m)} \wedge 1$$

where $\hat{x} \overset{m}{\sim} x$ with $\hat{x}(m) = -x(m)$. Let us call the resulting configuration Y_n. Next we apply the corresponding transformation to $Y_n^-(m)$ for the odd sites $m \in \Lambda^- \setminus \partial\Lambda$, to obtain X_{n+1}. The process $(X_n)_{n \geq 0}$ is then a Markov chain in I^+ with invariant distribution π.

Both methods we have described serve to simulate samples from π; there is little to choose between them. Convergence is fast in the subcritical case $\sinh 2\beta < 1$, where π has an approximate product structure on large scales.

In a Bayesian analysis of two-dimensional images, the Ising model is sometimes used as a prior. We may encode a digitized image on a two-dimensional grid as a particular configuration $\big(x(m) : m \in \Lambda\big) \in I$, where $x(m) = 1$ for a white pixel and $x(m) = -1$ for a black pixel. By varying the parameter β in the Ising model, we vary the tendency of black pixels

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press

to clump together; the same for white pixels. Thus β is a sort of texture parameter, which we choose according to the sort of image we expect, thus obtaining a prior $\pi(x)$. Observations are now made at each site which record the true pixel, black or white, with probability $p \in (0,1)$. The posterior distribution for X given observations Y is then given by

$$\pi(x \mid y) \propto \pi(x) f(y \mid x) \propto e^{-\beta H(x)} p^{a(x,y)} (1-p)^{d(x,y)}$$

where $a(x,y)$ and $d(x,y)$ are the numbers of sites at which x and y agree and disagree respectively. 'Cleaned-up' versions of the observed image Y may now be obtained by simulating from the posterior distribution. Although this is not exactly the Ising model, the same methods work. We describe the appropriate Metropolis algorithm: given that $X_n = x$, independently for each $m \in \Lambda^+ \backslash \partial\Lambda$, change the sign of $X_n^+(m)$ with probability

$$\begin{aligned} p(m,x,y) &= \big(\pi(\hat{x} \mid y)/\pi(x \mid y)\big) \wedge 1 \\ &= e^{-2\beta x(m) s(m)} \big((1-p)/p\big)^{x(m)y(m)} \end{aligned}$$

where $\hat{x} \overset{m}{\sim} x$ with $\hat{x}(m) = -x(m)$. Call the resulting configuration $X_{n+1/2}$. Next apply the corresponding transformation to $X_{n+1/2}^-$ for the odd sites to obtain X_{n+1}. Then $(X_n)_{n \geq 0}$ is a Markov chain in I^+ with invariant distribution $\pi(\cdot \mid y)$.

https://doi.org/10.1017/CBO9780511810633.007 Published online by Cambridge University Press