Chapter 23

Markov Chains in Continuous Time

Previously we looked at Markov chains, where the transitions between states occurred at specified time-steps. That it, we made time (a continuous variable) advance in fixed, discrete steps. The transitions between each state, at each step, were defined by the transition matrix. In this chapter, we push the boundaries of theory a little further and allow time to be a continuous variable.

So let us combine these ideas of a random (discrete) variable X that may jump to other states at *random times*.

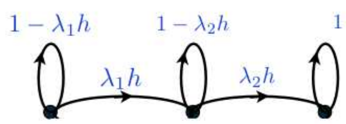
Suppose $X(t)$ is a discrete random variable. so we can label the states as follows:

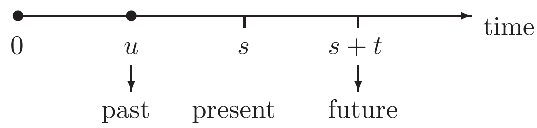
$$X \in \{0, 1, 2, \ldots, n\} \qquad .$$

And then define a *past* time (u), a *present* time (s) and a *future* time $(t + s)$.

Let us define a continuous time Markov process as follows:
For *all* $s, t \geqslant 0$ and $s + t > s > u \geqslant 0$

$$\boxed{P(X(t+s) = j | X(u) = k \text{ and } X(s) = i) = P(X(t+s) = j | X(s) = i)} \qquad . \tag{23.1}$$





That is, the process of getting from $X(u)$ to $X(s)$ (the past to the present) has no effect on what happens from $X(s)$ to $X(s + t)$ (the present to the future).

If, in addition, we make the following rule:

$$P(X(t+s) = j | X(s) = i) = P(X(t) = j | X(0) = i)$$

then the process is said to be 'stationary'. That is, the transition probability does *not* depend on the 'timing' of the process (s), just the time interval. We will assume (in this course) that the processes are 'stationary' or *homogeneous* Markov processes.

It is still possible to speak of a transition matrix, namely:

$$\boxed{P(X(t) = j|X(0) = i) = p_{ij}(t)} \quad . \tag{23.2}$$

But now the matrix depends on time.

Although the specific state of the system, at any time, is random, the system must be in one of the states. That is the total probability must add to 1.

As in the discrete case - this matrix is 'stochastic', that is for every state i:

$$\boxed{\sum_j p_{ij}(t) = 1} \quad . \tag{23.3}$$

And this must be true at all times, t. In other words, the system has to go somewhere. Summing over all j covers any possible future outcome.

23.0.1 Chapman-Kolmogorov relation

Another important property that continuous-time chains share with discrete time chains is the Chapman-Kolmogorov equation. If we consider the *meaning* of this relation for discrete time, any transition from the present to the future, has to pass through some state at some intermediate time (the 'near future').

So consider the present $t = 0$ some future time $s + t$ and some near future time, s. Conditioning on the (unkown and random) intermediate state in the near future we have:

$$p_{ij}(s+t) = P(X(s+t) = j|X(0) = i) = \sum_k P(X(s+t) = j|X(s) = k, X(0) = i)P(X(s) = k|X(0) = i) \quad . \tag{23.4}$$

But since the process is Markovian the conditional probability can be written as:

$$P(X(s+t) = j|X(s) = k, X(0) = i) = P(X(s+t) = j|X(s) = k) \tag{23.5}$$

and since the process is stationary, we can reset the time to start from zero:

$$P(X(s+t) = j|X(s) = k) = P(X(t) = j|X(0) = k) \quad . \tag{23.6}$$

So then it follows that:

$$P(X(s+t) = j|X(0) = i) = \sum_k P(X(t) = j|X(0) = k)P(X(s) = k|X(0) = i) \quad . \tag{23.7}$$

or in other words:

$$p_{ij}(s+t) = \sum_k p_{ik}(s)p_{kj}(t) \quad . \tag{23.8}$$

We recognize a matrix product on the right-hand side. So using matrix notation gives a more elegant statement of the Chapman-Kolmogorov relation:

$$\boxed{\mathsf{P}(s+t) = \mathsf{P}(s)\mathsf{P}(t)} \quad . \tag{23.9}$$

A collateral consequence of this relation is the commutivity of the transition matrices:

$$\boxed{\mathsf{P}(s)\mathsf{P}(t) = \mathsf{P}(t)\mathsf{P}(s)} \quad . \tag{23.10}$$

23.1 Kolmogorov equations

We would like to have a simple formula for the probability of being in a certain state at a certain time. That is, something akin to the formula for the Poisson process for $P(N(t) = n)$. This can be done, and we show how in this section.

Let us consider the Chapman-Kolmogorov relation, equation (23.9), for a short time in the future:

$$\mathsf{P}(t+h) = \mathsf{P}(h)\mathsf{P}(t) = \mathsf{P}(t)\mathsf{P}(h) \quad . \tag{23.11}$$

Note that we can get to $t+h$ from $t=0$, by either a small step h then a large step t or vice versa. this explains why there are two possible expressions.

Now, subtracting $\mathsf{P}(t)$ from both sides we get:

$$\mathsf{P}(t+h) - \mathsf{P}(t) = (\mathsf{P}(h) - \mathsf{I})\mathsf{P}(t) = \mathsf{P}(t)(\mathsf{P}(h) - \mathsf{I}) \quad . \tag{23.12}$$

The symbol s I denotes the unit (identity) matrix.

Then dividing by the time increment, h, and taking the limit $h \to 0$, gives:

$$\lim_{h\to 0} \frac{\mathsf{P}(t+h) - \mathsf{P}(t)}{h} = \lim_{h\to 0} \left(\frac{\mathsf{P}(h) - \mathsf{I}}{h}\right) \mathsf{P}(t) = \mathsf{P}(t) \lim_{h\to 0} \left(\frac{\mathsf{P}(h) - \mathsf{I}}{h}\right) \quad . \tag{23.13}$$

Let us define the *jump rate* matrix (also called the *generator matrix*):

$$\boxed{\mathsf{Q} \equiv \lim_{h\to 0} \frac{\mathsf{P}(h) - \mathsf{I}}{h}} \quad . \tag{23.14}$$

This gives us (a) the Kolmogorov *forward* equation:

$$\boxed{\frac{d}{dt}\mathsf{P}(t) = \mathsf{P}(t)\mathsf{Q}} \tag{23.15}$$

and (b) the equivalent Kolmogorov *backward* equation:

$$\boxed{\frac{d}{dt}\mathsf{P}(t) = \mathsf{Q}\mathsf{P}(t)} \quad . \tag{23.16}$$

Obviously, according to (23.15) and (23.16), the matrices Q and $\mathsf{P}(t)$ commute, that is

$$\mathsf{Q}\mathsf{P}(t) = \mathsf{P}(t)\mathsf{Q} \quad . \tag{23.17}$$

Since the sum along any row of the transition matrix is one, we note that the sum of every row in a jump-rate matrix is ZERO, since:

$$\sum_j Q_{ij} = \lim_{h\to 0} \frac{1}{h}\sum_j (\mathsf{P}_{ij}(h) - \delta_{ij}) \qquad . \tag{23.18}$$

That is:

$$\boxed{\sum_j Q_{ij} = 0} \qquad , \quad \text{for all } i \qquad . \tag{23.19}$$

Theorem: The solution of (either/both) the Kolmogorov equations is:

$$\boxed{\mathsf{P}(t) = \exp(\mathsf{Q}t)} \qquad . \tag{23.20}$$

Proof:

Firstly, we *define* the exponential function for a matrix by the power series:

$$\exp(\mathsf{A}) \equiv \mathsf{I} + \frac{\mathsf{A}}{1!} + \frac{\mathsf{A}^2}{2!} + \frac{\mathsf{A}^3}{3!} + \cdots$$

Therefore: $\exp(\mathsf{0}) = \mathsf{I}$. Then according to the proposed solution (23.20):

$$\mathsf{P}(0) = \exp(\mathsf{Q}\times 0) = \exp(\mathsf{0}) = \mathsf{I} \qquad . \tag{23.21}$$

This is correct, since it is known that:

$$P(X(0)=j|X(0)=i) = 1 \quad i=j \qquad , \qquad P(X(0)=j|X(0)=i)=0 \quad i\neq j \qquad .$$

That is, given we start in a state (i) at $t=0$ we are certain to be in that state at $t=0$ and no other.

So our proposed 'solution' satisfies the intial conditions. Does it satisfy the Kolmogorov differential equations (23.16 and 23.15).

If we substitute (plug in) the solution (23.20) for $d\mathsf{P}/dt$ we get:

$$\frac{d}{dt}\mathsf{P} = \frac{d}{dt}\left(\mathsf{I} + \frac{\mathsf{Q}t}{1!} + \frac{\mathsf{Q}^2t^2}{2!} + \frac{\mathsf{Q}^3t^3}{3!} + \cdots\right) \qquad . \tag{23.22}$$

Then differentiation of each term in turn gives:

$$\mathsf{0} + \frac{\mathsf{Q}}{1!} + \frac{2\mathsf{Q}^2t}{2!} + \frac{3\mathsf{Q}^3t^2}{3!} + \cdots \tag{23.23}$$

that is:

$$\frac{d}{dt}\mathsf{P} = \mathsf{Q}\left(\mathsf{I} + \frac{\mathsf{Q}t}{1!} + \frac{\mathsf{Q}^2t^2}{2!} + \cdots\right) = \mathsf{Q}\mathsf{P}(t) \tag{23.24}$$

which is the Kolmogorov backward equation (23.15) as required.

This all seems very straightforward. The only problem arises in calculating the exponential of a matrix. This is quite easy on a computer, but not so easy with pencil and paper. Only for very small matrices, or very simple matrices, can this be achieved. We will consider a few examples of this kind in the next chapter. Let us conclude the discussion here with a word on stationary/equilibrium states.

23.2 Equilibrium states

Suppose we have a continuous-time Markov chain, then we know that its dynamics are governed by the Kolmogorov equations:

$$\frac{d}{dt}\mathsf{P}(t) = \mathsf{Q}\mathsf{P}(t) = \mathsf{P}(t)\mathsf{Q} \qquad . \tag{23.25}$$

For an equilibrium in the system (or stationary state) the probabilities will not change in time. That is, for an equilibrium state, no matter what the initial conditions at $t = 0$:

$$\lim_{t\to\infty} \mathsf{P}_{ij}(t) = \pi_j \qquad , \qquad \text{for all } i \qquad . \tag{23.26}$$

That is, at equilibirum, in the long run:

$$\lim_{t\to\infty} \frac{d}{dt}\mathsf{P} = 0 \qquad . \tag{23.27}$$

Then the equilibrium distribution must be such that:

$$\mathsf{P}\mathsf{Q} = 0 \qquad . \tag{23.28}$$

Recalling that, at extremely long times, the transition matrix becomes a matrix with each row equal to the equilibrium distribution. This then leads to:

$$\boldsymbol{\pi}\mathsf{Q} = 0 \qquad . \tag{23.29}$$

That is, the equilbrium state is in the (left) null space of the jump rate matrix.

More explicitly, for a system with states $X \in \{0, 1, 2, 3, \ldots\}$, the equilibrium distribution is such that:

$$\sum_j \pi_j \mathsf{Q}_{jk} = 0 \qquad , \qquad k = 0, 1, 2, \ldots, n \qquad . \tag{23.30}$$

An equilbrium may, or may not, exist in the system. If an equilibirum does *not* exist, then the only solution of

$$\boldsymbol{\pi}\mathsf{Q} = 0 \tag{23.31}$$

will be the trivial solution, $\boldsymbol{\pi} = 0$.

23.3 Jump rates for the Poisson process

Consider the Poisson process as an example of a continuous-time Markov chain. We can denote the *state* of the system, X most conveniently by the number of arrivals:

$$N(t) \in \{0, 1, 2, 3, \ldots, \infty\} \qquad .$$

so we take $X = N(t)$. So for this example, what would be the *jump-rate matrix* ?

Well, one can analyse the *transition matrix* over a short time h. This can be represented by the *transition graph* (23.2)

For the Poisson process $P(N(h) = j|N(0) = i)$ is well known. This is just the extra events that occur

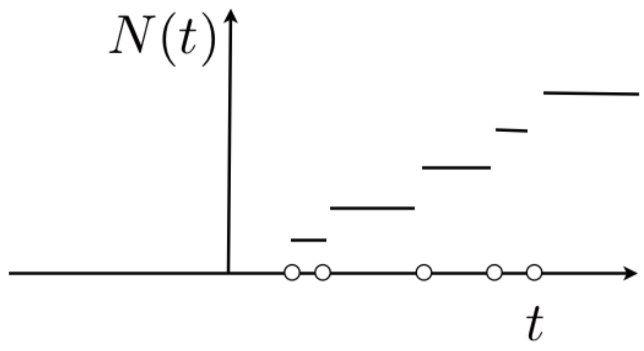


Figure 23.1: Counting process for Poisson arrivals.

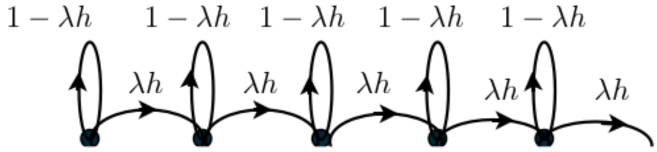


Figure 23.2: Transition graph for a (homogeneous) Poisson process over a short time h. Note that, when the time is very short we never observe more than one arrival. So the jumps are only between nearest-neighbour states $n \to n+1$, with probability proportional to the rate λ and the observation time h. By very short we mean $\lambda h \ll 1$, so the transition probabilities are very small to the neighbouring states.

in the short time h, by very short we mean $\lambda h \ll 1$. We know that, for $j < i$, there is zero probability. That is the number of events cannot decrease (there are no 'negative' events to reduce the number), $N(t)$ always stays the same or increases.

We know that there is a negligible chance of more than one event happening in such a short time. In fact, as discussed, we always arrange the time h such that this is the case! Really, we only need be concerned with one event of no event in this time. In mathematical terms we have:

$$P(N(h) = j | N(0) = i) = \begin{cases} \lambda h + o(h) & j = i+1 \quad \textsf{one event} \\ 1 - \lambda h + o(h) & j = i \quad \textsf{no event} \\ o(h) & j > i+1 \quad \textsf{multiple events} \\ 0 & j < i \quad \textsf{'negative' events} \end{cases}$$

By definition:

$$\mathsf{Q}_{ij} = \lim_{h \to 0} \frac{P(X(h) = j | X(0) = i) - \delta_{ij}}{h}$$

where we use the Kronecker delta (δ_{ij}) short-hand for the unit matrix:

$$\mathsf{I}_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Therefore the jump rate matrix is:

$$\mathsf{Q}_{ij} = \lim_{h\to 0} \frac{P(N(h)=j|N(0)=i)-\delta_{ij}}{h} = \begin{cases} \lambda & j=i+1 & \text{one event} \\ -\lambda & j=i & \text{no event} \\ 0 & j>i+1 & \text{multiple events} \\ 0 & j<i & \text{'negative' events} \end{cases}$$

That is, in explicit form we have a matrix with the following structure:

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & & \ddots & & \ddots \end{pmatrix} \tag{23.32}$$

The sum along every row of the jump-rate matrix is zero (as it should always be).

There is no equilibrium state to this system. That is, as time goes on, arrivals will continue to add up indefinitely so that $N(t) \to \infty$, as $t \to \infty$. The system never settles into a *stationary state* (a time-average equilibrium). This is confirmed by the solution to the equilibrium equation:

$$\boldsymbol{\pi}\mathsf{Q} = 0 \tag{23.33}$$

that is:

$$\begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \cdots \end{pmatrix} \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & & \ddots & & \ddots \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \cdots \end{pmatrix}$$

If we multiply out the vector ($\boldsymbol{\pi}$) by first column of Q, we get:

$$-\lambda\pi_0 = 0$$

which gives us: $\pi_0 = 0$. Then $\boldsymbol{\pi}$ times the second column of Q gives:

$$\lambda\pi_0 - \lambda\pi_1 = 0$$

giving: $\pi_1 = 0$. Continuing in this manner we see that the only solution is the trivial solution: $\pi_0 = 0, \pi_1 = 0, \pi_2 = 0 \ldots$. That is, there is no equilibrium distribution for such a system.