Markov Chains

Markov Processes
Discrete-time Markov Chains
Continuous-time Markov Chains
A Markov Process is a stochastic process $X_t$ with the Markov property:

\[
\Pr(X_{tn} \leq x_n \mid X_{tn-1} = x_{n-1}, X_{tn-2} = x_{n-2}, \ldots, X_{t0} = x_0) = \\
\Pr(X_{tn} \leq x_n \mid X_{tn-1} = x_{n-1})
\]

\forall \text{ states } x_i \text{ in the process' state space } S \\
\forall \text{ times } t_0 \leq t_1 \ldots \leq t_n \in \mathbb{R}^+

That is, the distribution of the possible states (values) of the Markov process at a time $t_n$ (i.e. the cdf of $X_{tn}$) is dependent only on the previous state of the process $x_{n-1}$ at a time $t_{n-1}$, and is independent of the whole history of states previous to that.

Equivalently, we can say that the states of a Markov process previous to the current state of the process have no effect on determining the future path of the process.

Markov processes are very useful for analysing the performance of a wide range of computer and communications system. These processes are relatively easy to solve, given the simplified form of the joint distribution function.
In general, the distribution $\Pr(X_{t_n} \leq x_n \mid X_{t_{n-1}} = x_{n-1})$, governing the probabilities of the occurrence of different values of the process, is not only dependent on the value of the last previous state, $x_{n-1}$, but is also dependent on the current time $t_n$ at which we observe the process to be in state $x_n$ and the time $t_{n-1}$ at which it was previously observed to be in state $x_{n-1}$.

For the Markov processes of interest in this course, we may always assume that this time dependence is limited to a dependence on the difference of the times $t_n$ and $t_{n-1}$. Such Markov processes are referred to as *time homogeneous* or simply *homogeneous*.

A Markov process is said to be **time-homogeneous** if

$$\Pr(X_{t_n} \leq x_n \mid X_{t_{n-1}} = x_{n-1}) = \Pr(X_{t_{n-t_{n-1}}} \leq x_n \mid X_0 = x_{n-1}) \quad \forall n$$

Note that time homogeneity and stationarity are not the same concept. For a stationary process the (unconditional) cdf does not change with shifts in time. For a homogeneous process, the conditional cdf of $X_{t_n}$ does not change with shifts in time. That is, a homogeneous process is not necessarily a stationary process.
Markov processes may be further classified according to whether the state space and/or parameter space (time) are discrete or continuous. A Markov process which has a discrete state space (with either discrete or continuous parameter spaces) is referred to as a **Markov Chain**. For a Markov Chain, we write the Markov property in terms of a state’s conditional pmf (as opposed to conditional cdf for a continuous state space):

For a **continuous-time Markov Chain** we write:

\[
\Pr(X_{t_n} = j \mid X_{t_{n-1}} = i_{n-1}, X_{t_{n-2}} = i_{n-2}, \ldots, X_{t_0} = i_0) = \\
\Pr(X_{t_n} = j \mid X_{t_{n-1}} = i_{n-1})
\]

\[\forall j \in S, \forall i_0, i_1, \ldots \in S, \forall t_0 \leq t_1 \ldots \leq t_n \in \mathbb{R}^+\]

For a **discrete-time Markov Chain** we write:

\[
\Pr(X_n = j \mid X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \ldots, X_0 = i_0) = \\
\Pr(X_n = j \mid X_{n-1} = i_{n-1})
\]

\[\forall j \in S, \forall i_0, i_1, \ldots \in S, \forall n \in \mathbb{N}\]

In either case, the process ranges over a discrete states space \(S\). This state space may be finite or (countably) infinite.
Consider a game where a coin is tossed repeatedly and the player’s score is accumulated by adding 2 points when a head turns up and adding 1 point for a tail. Is this a Markov Chain? Is the process homogeneous?

The state space of the process is formed by all possible accumulated scores that can occur over the course of the game (i.e. \( S = \mathbb{N} \)). For any given state we see that the distribution of possible values of the state is dependent only on the previous state, that is, the state distribution is described by:

\[
\Pr(X_n = j + 1 | X_{n-1} = j) = \frac{1}{2} \\
\Pr(X_n = j + 2 | X_{n-1} = j) = \frac{1}{2}
\]

and thus the process is a Markov process. Also, the state space is discrete (countably infinite), so the process is a Markov Chain.

Further more, the distribution of possible values of a state does not depend upon the time the observation is made, so the process is a homogeneous, discrete-time, Markov Chain.
Markov Chains & Birth-Death Processes

Markov Processes
Discrete-time Markov Chains
Continuous-time Markov Chains
We have defined a **discrete-time Markov Chain** as a process having the property

\[
\Pr(X_n = j \mid X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \ldots, X_0 = i_0) = \\
\Pr(X_n = j \mid X_{n-1} = i_{n-1})
\]

\[\forall j \in S, \; \forall i_0, i_1, \ldots \in S, \; \forall n \in \mathbb{N}\]

We say that when \(X_n = j\), the process is observed to be in state \(j\) at time \(n\).

\(\Pr(X_n = j \mid X_{n-1} = i)\) is the probability of finding the process in state \(j\) at time instance \(n\) (step \(n\) of the process), given that the system was previously in state \(i\) at time instance (step) \(n - 1\).

Equivalently, we can say that \(\Pr(X_n = j \mid X_{n-1} = i)\) is the probability of the process **transitioning** from state \(i\) into state \(j\) (in a single step) at time instance \(n\).

This probability is called the **one-step transition probability**, denoted as \(p_{ij}(n)\). As we will always assume homogeneity, \(\Pr(X_n = j \mid X_{n-1} = i)\) is the same probability for any \(n\) and so \(p_{ij}(n)\) is independent of \(n\) and the one-step transition probability is:

\[p_{ij} \triangleq P[X_n = j \mid X_{n-1} = i], \quad \forall n\]
Discrete-time Markov Chains: Transition Probabilities

The set of all transition probabilities $p_{ij}$ for all $i$ and all $j$ may be expressed in matrix form as a transition probability matrix $P$ where:

$$P = \begin{bmatrix}
  p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\
  p_{1,0} & p_{1,1} & p_{1,2} & \cdots \\
  p_{2,0} & p_{2,1} & p_{2,2} & \cdots \\
  \vdots & \vdots & \vdots & \ddots
\end{bmatrix}$$

The $i$th row of $P$ are the probabilities of transitioning out of state $i$ to other states of the process (including back to state $i$). The sum of these probabilities is the probability of transitioning to some other state from $i$ and so must equal 1, i.e.

$$\sum_j p_{i,j} = 1.$$ 

The $j$th column of $P$ are the probabilities of transitioning from states of the process (including state $j$) into state $j$. The sum of these probabilities is generally $\neq 1$.

**Exercise 7**: Again, consider the game where a coin is tossed repeatedly and the player’s score is accumulated by adding 2 points when a head turns up and adding 1 point for a tail. Write down the transition probability matrix $P$ for the process.
Discrete-time Markov Chains: Transition Probabilities

We may calculate the probability of a transition from a state \( i \) to a state \( j \) in two steps, through a given intermediate state \( k \), as \( p_{ik} p_{kj} \).

This is true as, by the Markov property, the transitions \( i \rightarrow k \) and \( k \rightarrow j \) are independent events.

Additionally, by applying the Law of Total Probability, \( \mathbb{P}[A] = \sum_{\forall i} \mathbb{P}(B_i) \mathbb{P}(A \mid B_i) \), the probability of transitioning from \( i \) to \( j \) in two steps through any intermediate state \( k \) may be calculated as:

\[
p^{(2)}_{ij} = \sum_{\forall k} p_{ik} p_{kj}
\]

Furthermore, we may calculate the \textit{m-step transition probability} from state \( i \) to state \( j \) recursively:

\[
p^{(m)}_{ij} = \sum_{\forall k} p^{(m-1)}_{ik} p_{kj} \quad m = 2, 3, \ldots
\]

This equation is called the \textbf{Chapman-Kolmogorov Equation}. It is a fundamental equation in the analysis of Markov chains, allowing us to calculate of the probability that the process is in a particular state after \( m \) time steps.
Discrete-time Markov Chains: Transition Probabilities

We may express the Chapman-Kolmogorov equation in a more compact matrix form by noting that the element \( \{i, j\} \) of the matrix \( P \cdot P = P^2 \) is equal to \( \sum_k p_{ik} p_{kj} \), which is equal to the two-step transition probability \( p_{ij}^{(2)} \).

Similarly, the element \( \{i, j\} \) of \( P^m \) is equal to the \( m \)-step transition probability \( p_{ij}^{(m)} \).

Thus, we can express the Chapman-Kolmogorov Equation as:

\[
P^m = P^{m-1}P
\]

or alternatively, in a more general form, as:

\[
P^m = P^{m-n}P^n
\]

\( P^m \) is referred to as the \textit{m-step transition probability matrix} of the Markov chain.

We now consider how the probability of being in a particular state at a particular time may be calculated from \( P^m \).
Discrete-time Markov Chains: State Probabilities

We denote the probability of the process being in state \( i \) at step \( n \) as:

\[
\pi_i^{(n)} = \Pr(X_n = i)
\]

We refer to this as a state probability. The state probabilities for all states at step \( n \) may be expressed as a state distribution vector

\[
\pi^{(n)} = (\pi_0^{(n)}, \pi_1^{(n)}, \pi_2^{(n)}, \ldots)
\]

Applying the Law of Total Probability, the probability that the process is in state \( i \) at time step 1 may be calculated as:

\[
\pi_i^{(1)} = \sum_{\forall k} \pi_k^{(0)} p_{ki}
\]

Or equivalently, in vector-matrix notation, the state distribution vector at the first step may be expressed as:

\[
\pi^{(1)} = \pi^{(0)} P
\]

\( \pi^{(0)} \) is referred to as the initial state distribution. It is an arbitrarily chosen distribution of the starting state of the process at time step 0.
The calculation of state probabilities can be continued for successive steps, as follows:

\[ \pi^{(1)} = \pi^{(0)} \mathbf{P} \]
\[ \pi^{(2)} = \pi^{(1)} \mathbf{P} = [\pi^{(0)} \mathbf{P}] \mathbf{P} = \pi^{(0)} \mathbf{P}^2 \]
\[ \vdots \]
\[ \pi^{(n)} = \pi^{(n-1)} \mathbf{P} = \pi^{(0)} \mathbf{P}^n \quad n = 1, 2, 3, \ldots \]

This gives us an important result for Markov chain analysis.

Knowing the initial distribution of states \( \pi^{(0)} \) at time step 0, and the transition probabilities between states given by \( \mathbf{P} \), we can calculate the distribution of states (the probability of being in any particular state) at any future time step \( n \).

\[ \pi^{(n)} = \pi^{(0)} \mathbf{P}^n \quad n = 1, 2, 3, \ldots \]
Exercise 8

Yet again, consider the game where a coin is tossed repeatedly and the player’s score is accumulated by adding 2 points when a head turns up and adding 1 point for a tail. We have previously found the transition probability matrix $P$ for the process.

By using the fact that $\pi^{(n)} = \pi^{(0)}P^n$, $n = 1, 2, 3, \ldots$, calculate the probability that the player’s score is at least 4 points having tossed the coin 3 times.
We have found a method to calculate $\pi^{(n)}$, which encodes the probable evolution of the Markov Chain over time starting at time 0. This vector is referred to as the transient solution of the Markov Chain.

We are also interested in how a Markov Chain behaves after a long time period (after many steps).

We can see that a process’s state distribution changes for successive steps starting at time 0. However, does a process’s state distribution ‘settle down’ over time to an equilibrium solution, that is, is there a limiting stationary distribution of the states after a long time period? Can we thus calculate the expected value of the process?

It is important to answer these questions so that we can take the analysis further and to calculate useful parameters of stochastic processes arising in communications performance analysis.

**Exercise 9**

Would you expect the process in the previous example to have a stationary distribution as $n \to \infty$? Is the process ergodic (as per the definition given in the Stochastic Processes section)?
To find the conditions under which a stationary distribution exists, we must first define some general properties of Markov Chains:

**Closure**
A subset of the states $A$ is **closed** if there is no possible transition from any state in $A$ to any state in $A^C$.

**Absorption**
A single state which is a closed set is called an **absorbing state**. That is, it is not possible to transition out of an absorbing state $i$ ($p_{ii} = 1$).

**Reducibility**
A closed set of states is called **irreducible** if it contains no closed (proper) subsets of itself. A Markov Chain is irreducible if the set of all states is irreducible.

An alternative definition is that a Markov Chain is irreducible if every state can be reached from every other state. That is, for all states $i$ and $j$, there is an integer $m$ such that $p_{ij}^{(m)} > 0$.

Conversely, a set of states is **reducible** if there exists a closed, proper subset of itself.
**Recurrence**
Further classifications are made based on whether or not, and how often, a state is revisited.

Let $f_i^{(n)}$ be the probability of the first return to state $i$ in exactly $n$ steps ($n \geq 1$).

The probability of ever returning to state $i$ is $f_i = \sum_{n=1}^{\infty} f_i^{(n)}$.

The *mean recurrence time* of state $i$ is $M_i = \sum_{n=1}^{\infty} n f_i^{(n)}$.

A state $i$ is said to be **transient** if the probability of returning to the state is less $< 1$ ($f_i < 1$). A state $i$ is said to be **recurrent** if $f_i = 1$.

A state $i$ is **positive recurrent** if its mean recurrence time is finite ($M_i < \infty$).

A state $i$ is **null recurrent** if its mean recurrence time is infinite ($M_i = \infty$).

**Periodicity**
A state is **periodic** if its first return time $f_i^{(n)}$ can only be a multiple of an integer greater than 1. Otherwise, the state is called **aperiodic**.
Theorem 1

The states of an irreducible Markov Chain are either

- all transient or
- all null-recurrent or
- all positive-recurrent

Additionally, all states are either aperiodic or periodic.

We are most interested in 'well-behaved' Markov Chains which have a unique stationary distribution. These types of chains are described as **ergodic**.

**Definition:** An irreducible Markov Chain which is positive-recurrent and aperiodic is said to be **ergodic**.

The next theorem states that a unique stationary distribution always exists for a homogeneous Markov Chain that is ergodic.
**Theorem 2 - Kolmogorov’s Theorem**

In an irreducible, aperiodic, time-homogeneous Markov Chain the limits

\[
\pi_j = \lim_{n \to \infty} \pi_j^{(n)} \quad \forall j
\]  

always exist and are independent of the initial state probability distribution \( \pi^{(0)} \). Moreover, either:

(a) all states are transient or all states are null-recurrent in which case \( \pi_j = 0 \) for all \( j \) and there exists no stationary distribution, or

(b) all states are positive-recurrent and then \( \pi_j > 0 \) for all \( j \), in which case the set \( \{\pi_j\} \) is a stationary probability distribution and

\[
\pi_j = \frac{1}{M_j} \quad \text{where } M_j \text{ is the mean recurrence time for state } j
\]

In this case the quantities \( \pi_j \) are uniquely determined by:

\[
\pi_j = \sum_{\forall i} \pi_i p_{ij} \quad (2) \quad \text{and} \quad 1 = \sum_{\forall i} \pi_i \quad (3)
\]
Discrete-time Markov Chains: Stationary Distribution

We may re-write (1) above, using our previous vector-matrix notation, as

\[ \pi = \lim_{n \to \infty} \pi^{(n)} \]

\( \pi \) is the **stationary distribution** of the Markov Chain, also know as the equilibrium solution or the equilibrium (or steady-state) state probability vector.

We may write (2) and (3) using more compact vector/matrix notation as

\[ \pi = \pi \mathbf{P} \] (2) \quad \text{Called the 'global balance equation'}

\[ \pi.1^T = 1 \] (3) \quad \text{Called the 'normalisation condition'}

where \( 1 \) is the unit row vector. We note that (2) can be obtained from the state probabilities equation, by way of

\[ \pi^{(n)} = \pi^{(n-1)} \mathbf{P} \]

\[ \lim_{n \to \infty} \pi^{(n)} = \lim_{n \to \infty} \pi^{(n-1)} \mathbf{P} \]

\[ \pi = \pi \mathbf{P} \]
Consider the **State Transition Diagram** below showing the probabilities of a Markov chain transitioning between its three states. (i) Explain why this chain must have a stationary distribution and (ii) calculate the equilibrium probabilities $\pi$. (iii) Calculate the transient solution $\pi^{(n)}$ for the first four steps of the process ($n = 1, 2, 3, 4$) for each of the initial state distribution vectors $\pi^{(0)} = (1, 0, 0)$, $\pi^{(0)} = (0, 1, 0)$ and $\pi^{(0)} = (0, 0, 1)$.
Discrete-time Markov Chains: Example

Solution

(i) We can see that the process is a homogeneous Markov Chain. It is also irreducible as all states are reachable from any state. It is recurrent as we will surely return to a state over time and is positive-recurrent as the expected first return time is less than infinity. Furthermore it is aperiodic. Thus, by Theorem 2, the process must have a stationary distribution.

(ii) From the state transition diagram we have

\[ P = \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \]

Applying \( \pi = \pi P \) we have

\[
\pi_0 = 0\pi_0 + \frac{1}{4}\pi_1 + \frac{1}{4}\pi_2 \\
\pi_1 = \frac{3}{4}\pi_0 + 0\pi_1 + \frac{1}{4}\pi_2 \\
\pi_2 = \frac{1}{4}\pi_0 + \frac{3}{4}\pi_1 + \frac{1}{2}\pi_2
\]
In addition, the normalisation condition \( \pi_0 \cdot \mathbf{1}^T = 1 \) gives

\[
\pi_0 + \pi_1 + \pi_2 = 1
\]

We now have a system of four simultaneous equations which must have a unique solution (by Theorem 2). Solving the equations gives the solution:

\[
\pi_0 = \frac{1}{5}, \quad \pi_1 = \frac{7}{25}, \quad \pi_2 = \frac{13}{25}
\]

or in vector notation:

\[
\pi = \left[ \frac{1}{5}, \frac{7}{25}, \frac{13}{25} \right]
\]

These are the equilibrium state probabilities (the stationary distribution of the Markov Chain).

(iii) We may calculate the \textbf{transient solution} for the first few steps of the process using \( \pi^{(n)} = \pi^{(n-1)} \mathbf{P} \) iteratively, that is:

\[
\begin{align*}
\pi^{(1)} &= \pi^{(0)} \mathbf{P} \\
\pi^{(2)} &= \pi^{(1)} \mathbf{P} \\
\pi^{(3)} &= \pi^{(2)} \mathbf{P} \\
&\vdots
\end{align*}
\]
We do this for the different starting points (the different initial state probabilities) $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. The results are:

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\pi_0^{(n)}$</th>
<th>$\pi_1^{(n)}$</th>
<th>$\pi_2^{(n)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.25</td>
<td>0.75</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>0.19</td>
<td>0.36</td>
<td>0.69</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
<td>0.25</td>
<td>0.45</td>
</tr>
<tr>
<td>4</td>
<td>0.20</td>
<td>0.28</td>
<td>0.55</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.20</td>
<td>0.28</td>
<td>0.52</td>
</tr>
</tbody>
</table>

We note that after four steps the state distribution vectors are very close in value to the stationary distribution vector, regardless of the starting point of the process.
Markov Chains & Birth-Death Processes

Markov Processes
Discrete-time Markov Chains
Continuous-time Markov Chains
Continuous-time Markov Chains

In the case of a continuous-time Markov Chain, state transitions may occur at any arbitrary time instance and not merely at fixed, discrete time points, as was the case for discrete-time chains.

Thus, for the continuous-time Markov Chain, the Markov property is expressed in terms of the continuous time parameter $t \in \mathbb{R}^+$ as

\[
\Pr(X_{t_n} = j \mid X_{t_{n-1}} = i_{n-1}, X_{t_{n-2}} = i_{n-2}, \ldots, X_0 = i_0) = \\
\Pr(X_{t_n} = j \mid X_{t_{n-1}} = i_{n-1}) \\
\forall j \in \mathcal{S}, \quad \forall i_0, i_1, \ldots \in \mathcal{S}, \quad \forall t_0 \leq t_1 \ldots \leq t_n \in \mathbb{R}^+
\]

That is, the process is entirely described by the state transition probability $\Pr(X_{t_n} = j \mid X_{t_{n-1}} = i)$, the probability of finding the process in state $j$ at a time $t_n$, given that it was in state $i$ at time $t_{n-1}$, for all $i, j, t_{n-1}$ and $t_n$.

As with the discrete time case, we only concern ourselves with chains that are time-homogeneous. In this case, for any times $t_1 < t_2$, and $t \triangleq t_2 - t_1$

\[
\Pr(X_{t_2} = j \mid X_{t_1} = i) = \Pr(X_t = j \mid X_0 = i)
\]

and we denote this **transition probability** as

\[
p_{ij}(t) \equiv p_{ij}(0, t) = \Pr(X_t = j \mid X_0 = i) = \Pr(X_{t_2} = j \mid X_{t_1} = i)
\]
Continuous-time Markov Chains

$p_{ij}(0, t)$ is analogous to the $m$-step transition probability in the discrete-time case, where the discrete-time variable $m$ is replaced by the continuous-time variable $t$.

Similarly to the discrete-time case, the **Chapman-Kolmogorov** equation for the continuous-time Markov chain may be written as:

\[
p_{ij}(t) = p_{ij}(0, t) = \sum_{\forall k \in S} p_{ik}(0, t')p_{kj}(t', t), \quad 0 < t' < t
\]  

(4)

Or in matrix notation we may write equation (4) as \( \mathbf{P}(t) = \mathbf{P}(t')\mathbf{P}(t - t') \)

If we denote the **state probabilities**, the probability of being in state $i$ at a time $t$, as $\pi_i(t) = \Pr(X_t = i)$, we may write the relationship between state probabilities and transition probabilities, for the continuous-time Markov chain, as:

\[
\pi_j(t) = \sum_{\forall i \in S} p_{ij}(t)\pi_i(0)
\]

(5)

Or in vector-matrix notation we may write equation (5) as $\mathbf{\pi}(t) = \mathbf{\pi}(0)\mathbf{P}(t)$. 
In summary, similar to the discrete-time case, we have two equations in terms of the state probability distribution vector \( \pi(t) \) and the transition probability matrix \( P(t) \), from which we wish to resolve the transient and steady-state (stationary) state probabilities.

As with the discrete-time case, the existence of the stationary distribution of the state probabilities depends on the properties of the chain. We first resolve the transient solution and then examine the steady-state solution.

In the discrete-time case \( p_{ij} \) did not have a dependence on the length of time between state changes and the solution was obtained by solving a set of linear equations. In the continuous-time case, the transition probabilities \( p_{ij}(t) \) are dependent on \( t \) and the solution is more involved.

Instead of resolving the probabilities \( p_{ij}(t) \) directly, we express the probabilities in terms of a quantity \( q_{ij} \) for which, for time-homogeneous chains, we can remove the dependence on the transition time.

\( q_{ij} \) is termed the **state transition rate**, the rate at which the chain leaves state \( i \) in order to transition in state \( j \).
We define this transition rate as

\[ q_{ij} \triangleq \lim_{\Delta t \to 0} \frac{p_{ij}(t, t + \Delta t)}{\Delta t}, \quad i \neq j \]

\[ q_{ii} \triangleq \lim_{\Delta t \to 0} \frac{p_{ii}(t, t + \Delta t) - 1}{\Delta t} \]

Now, considering the Chapman-Kolmogorov equation (4) at time points 0, t and t + Δt

\[ p_{ij}(t + \Delta t) = \sum_{\forall k \in S} p_{ik}(t)p_{kj}(\Delta t) \]

Dividing by Δt and taking the limit as Δt → 0

\[ \lim_{\Delta t \to 0} \frac{p_{ij}(t + \Delta t)}{\Delta t} = \sum_{\forall k \in S} p_{ik}(t) \lim_{\Delta t \to 0} \frac{p_{kj}(\Delta t)}{\Delta t} \]

\[ \frac{d}{dt} p_{ij}(t) = \sum_{\forall k \in S} p_{ik}(t)q_{kj} \quad (6) \]
Continuous-time Markov Chains

Now, differentiating (5) w.r.t. $t$ and then substituting from (6):

$$\frac{d}{dt} \pi_j(t) = \sum_{\forall i \in S} \frac{d}{dt} p_{ij}(t) \pi_i(0)$$

$$= \sum_{\forall i \in S} \sum_{\forall k \in S} p_{ik}(t) q_{kj} \pi_i(0)$$

$$= \sum_{\forall k \in S} q_{kj} \sum_{\forall i \in S} p_{ik}(t) \pi_i(0)$$

$$\frac{d}{dt} \pi_j(t) = \sum_{\forall k \in S} q_{kj} \pi_k(t)$$

The solution of this system of differential equations gives the **transient solution** $\pi = \{\pi_i(t)\}$ of the continuous-time (homogeneous) Markov Chain. We can express these equations in vector-matrix form as

$$\dot{\pi}(t) = \pi(t)Q$$

The structure of $Q$ determines how easy/difficult it is to find a solution for $\pi$. We are most interested in processes called *Birth-Death processes*, which have a 'nice' $Q$ matrix that makes the system relatively easy to solve.
Similarly to the discrete-time case, it can be shown that a homogeneous continuous-time chain that is *ergodic* has a unique stationary (steady-state) distribution $\pi$.

In this case, $\pi(t)$ tends to the constant $\pi$ as $t \to \infty$, so

$$\lim_{t \to \infty} \frac{d\pi_i(t)}{dt} = 0, \quad \forall i \in S$$

which yields a simple set of linear equations for resolving $\pi$

$$\sum_{\forall i \in S} q_{ij} \pi_i = 0, \quad \forall j \in S$$

or, in vector-matrix notation, the steady-state solution is given by resolution of

$$0 = \pi Q$$

To find a unique solution, we must also impose the *normalisation condition*

$$\pi 1^T = \sum_{\forall i \in S} \pi_i = 1$$